Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 16, Issue. 2 : 2025
ISSN : **1906-9685**

Deep learning based segmentation of the coronary artery region in the Cardiac MRI images using U-Net model

Mr. P Vijay Kumar

Assistant Professor ,Department of Computer Science,Sir C R Reddy College(A),ELURU

Email Id:gani.1239@gmail.com

Mrs K Haritha

Assistant Professor ,Department of Computer Science,Sir C R Reddy College(A),ELURU

Email Id:kharitha8492@gmail.com

Mrs Mehaboob Karishma

Assistant Professor ,MCA Department,Sir C R Reddy College(A),ELURU

Email Id:mehaboob.karishma@gmail.com

Mr T Pardha Saradhi

Pursuing II MCA ,MCA Department ,Sir C R Reddy College(A),ELURU

Email Id:pardhutirumani9@gmail.com

Abstract: Since the past few years, in the province of computer vision the availability of medical imaging technology has helped the doctors to visualize and analyze the interior parts of the body with ease. Previously, classical methods have been employed to deal with medical image analysis. Later with the emergence of deep learning concepts and techniques, essentially semantic segmentation, its use has rapidly increased and became one of the most preferred choices of the researchers. That too, in cardiac pathologies, automated segmentation of known object of interest in cardiac images from MRI datasets is a very crucial step for timely diagnosis. So, after the thorough research it is perceived that, U-Net architecture which is based upon deep fully convolutional encoder-decoder model has been accredited for its triumph in the biomedical image processing community. In this work, to meet the requirements of the system objective, the implementation of detection and segmentation of coronary artery on the cardiac images of MRI modality has turned out to be rewarding using U-Net architecture. Evaluation metrics presented in this paper, showcases that this model achieved 96.62% mean IoU, 77.95% precision, 81.72% recall and 89.22% F1-score on the test dataset.

Index terms: deep learning, U-Net, medical imaging technology, object of interest, fully convolutional encoder-decoder model, cardiac image dataset, semantic segmentation.

1. Introduction

Over the last few decades there have been efforts to train machines to replicate the human brain which gave rise to the concept of machine learning (ML) [1] where one can train machines to make human like decisions by recognizing certain features. In other words, ML quantifies the contents in the images employing hand-crafted features. Feature extraction descriptors such as texture, shape and color are applied for all the images using some ML algorithm that returns a feature vector (a list of numbers). This resultant feature vectors are then directed as inputs to the machine learning architectures.

Since 2006 a complex yet effective technique termed deep learning (DL) [2] which is the leading-edge in the manifestation of neural networks (NN), has emerged as a new area of ML research. It has been impacting a wide range of real time information processing work. The advantage of DL is that the program automatically builds the feature set to classify data, as opposed to most traditional ML algorithms, which require intense time and effort on the part of data scientists. The DL techniques depend on learning the object features instead of extracting them which makes it very useful in cases where the set of features cannot accurately describe the objects by the user. By the implementation of DL techniques it is possible to train the images and search through millions of available images so as to accurately identify the object of interest in the image within a short span. In order to achieve an acceptable level of accuracy, DL programs require access to immense amount of training data and processing power. But the emergence of hardware specifications enabled the possibilities of creating and testing big NN [3] with larger training data and hence it is much faster, efficient and accurate. The captivating factor of DL from the earlier incarnations of NN is that feature extraction step is bypassed and the emphasis is laid on learning the filters in the process of training the network.

The remaining work in this paper is organized as follows: section 2 elucidates the related works, section 3 discusses about neural network architecture for segmentation the cardiac images, section 4 presents the evaluation metrics, section 5 describes the results and discussions and finally section 6 with conclusions followed by acknowledgements.

2. Related Work

In the domain of medical imaging [4], paramount importance is conferred to quantitative analysis of clinical parameters associated with shape and volume and DL techniques are the best choice for performing any tasks related to classification, object detection or segmentation. The number of medical imaging strategies being carried out around the globe is enhancing much more times faster than the number of doctors available who can really construe them. Every scan is rich in information and much more complex than some time ago. At present, magnetic resonance imaging (MRI) is one of the most appropriate and precise non-invasive diagnostic tools to assess the structure and function of an organ in medical imaging studies. The quantity of information keep accumulating as does the load of the physician that they wish to process, prior to arriving at a diagnosis. In this scenario, the risk of human error intensifies and thereby the doctors may miss a diagnosis that is tangible in retrospection. DL algorithms meticulously discover and highlight anomalies, therein decreasing the likelihood of over passing a diagnosis. DL transforms how machines learn from having to manually delineated features to automatically learning features from the set of labeled images data. In recent times, DL has been able to accomplish well on tasks which call for human expertise gained over years of training. The advantage machines have in diagnosing is that they can scrutinize thousands of images in a matter of few hours, a number which may deplete many years or even a whole life time of a human. Furthermore, healthcare is rushing through a process of revolution of its own. The quantity of medical image data produced is thriving exponentially. The tendency of diagnostic imaging has outstretched significantly and endures to evolve.

Although there are many works related to medical image analysis pertaining to individual human parts like kidney, liver, brain, heart etc., the work in this paper is confined only to segmentation of cardiac images as it is beyond the scope of this paper to discuss all the work done on the medical images of various organs. Especially, detecting and segmenting the substructures of cardiac images is essential in determining their structural and behavioral patterns. Prior to automation, cardiologists manually segmented the images which is in fact very slow and has got poor accuracy. With the advent of fully convolutional neural networks (FCNs) [5] automatic segmentation of medical images has grown importance to find a fast and accurate method to analyze them [6].

The list of papers [7-15] emphasize the research works carried out on segmentation of left ventricle (LV) as it is the biggest chamber of the heart and right ventricle (RV) as several cardiac diseases such as cardiomyopathy, pulmonary hypertension, dysplasia are related to RV. When compared to LV, the segmentation of RV poses more difficulties due to their complex crescent shape, relatively thinner wall of ventricle etc. So far, much of the research has been done on cardiac images for automatic segmentation of left and right ventricles using classical methods such as atlas-based models, statistical models, deformable models, traditional ML techniques and DL techniques as there are a number of open technical challenges in automated LV and RV segmentation. But for automatic segmentation of the coronary artery(s) there have been works done using these classical methods whereas still it is a high demanding task to improve the segmentation accuracy further by using deep learning architectures. The job of coronary artery segmentation does not need such complex models to train with and can be achieved with good accuracy through the implementation of convolutional encoder-decoder models that are found to be appropriate for this task. U-Net architecture [16, 17 and 7] is chosen as a pertinent one to ensue with this problem.

3. Network Architecture

CNNs [18] combine features with the aim of classifying their input whereas; convolutional autoencoders (CAEs) are trained only to learn filters to be able to extract features that can be used to reconstruct the input. Since, CAEs are convolutional in nature regardless of image size they scale well to the realistic-sized high-dimensional images as the number of parameters necessary to yield an activation map is all the time alike. Hence CAEs are general purpose feature extractors [19].

Employing convolutional networks in encoder-decoder architecture as a competent method to accomplish segmentation can be recommended for this task. An encoder-decoder convolutional network which is composed of a stack of encoders followed by a corresponding decoder stack that feeds into a soft-max classification layer, functions greatly like a CAEs except that it is trained to yield segmentation instead of a reconstruction of the input. Particularly the encoder maps the input to a representation which captures information about the structure in the image. That representation which is a low resolution feature maps received as the output of the encoder is then mapped back to full input image size feature maps of the original resolution by the decoder, creating a pixel-by-pixel segmentation of the image. This booms a key shortcoming of present-day DL methodologies which have adopted networks intended for object categorization for pixel wise labeling [20]. U-Net architecture uses convolutional encoder-decoder model which is discussed in detail as follows.

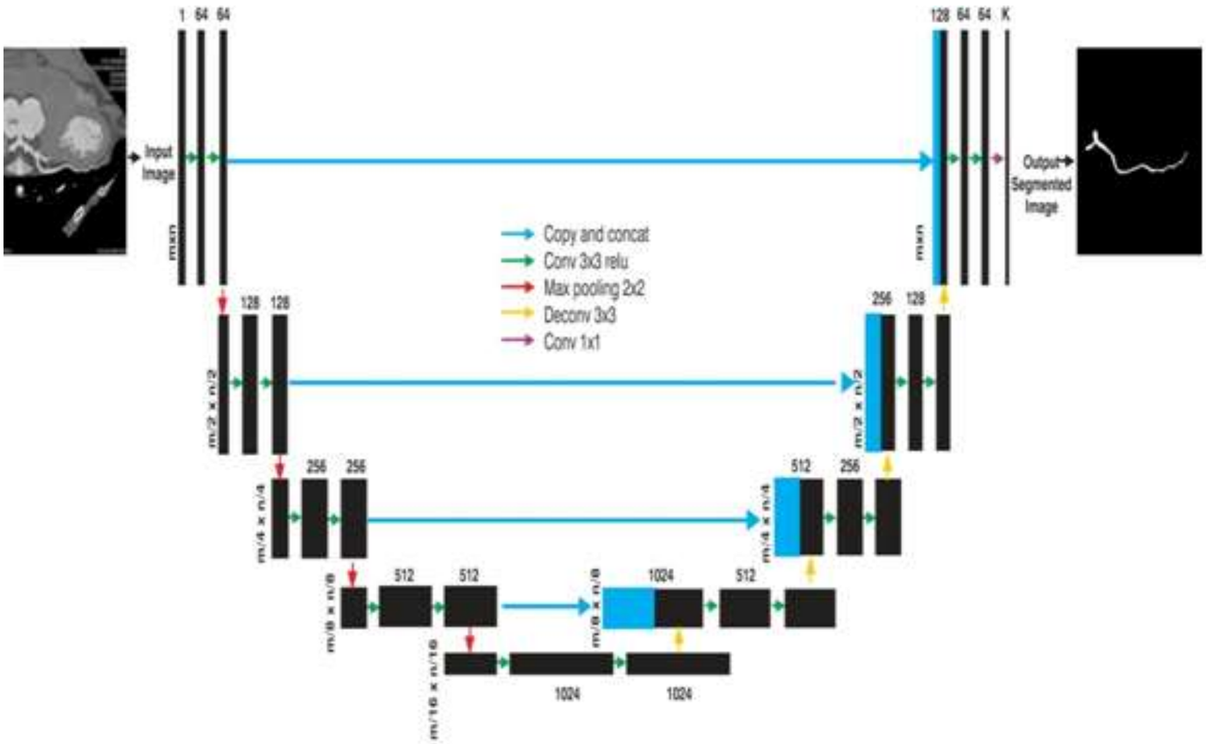

Fig. 1 U-Net Architecture. Image reproduced from [22]

3.1 U-Net

Ronneberger et al. [16] in his network architecture named U-Net as illustrated in figure. 1 has skip connections [21] from encoder layers (usually acts as an analysis arm) to decoder layers (usually a synthesis arm) that are on the same level. In total, the U-Net network has 23 convolutional layers. It consists of a contracting path on left side and an expansive path on right side. As like in the architecture of a typical convolutional network, U-Net also has contracting path which consists of the repeated application of two 3x3 convolutions, each followed by a rectified linear unit (ReLU) [22] and a 2x2 max pooling operation with stride 2 for downsampling and at each downsampling step the number of feature channels are doubled. Every step in the expansive path consists of an upsampling of the feature map followed by a 2x2 convolution. This halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3x3 convolutions, each followed by a ReLU. The cropping is necessary due to the loss of border pixels in every convolution. A convolution of 1x1 is applied on the final layer that can be used to map each 64 component feature vector to the desired number of classes.

4. Evaluation Metrics

Once the model is built, the most important element that crops up is assessing that specific model. To acquire competent results in the zone of semantic segmentation, it is essential to evaluate the performance of a segmentation method [23] with rigor as the users expect accurate results. This is a vital job in a particular DL system which will formulate how smart the

Table 1: Depicts the confusion matrix

		Predicted class	
Actual Class	Class:Yes	True Positive(TP)	False Negative(FN)
	Class:No	False Positive(FP)	True Negative(TN)

predictions are. The above table 1 depicts the confusion matrix followed by the evaluation measures such as IoU, precision, recall and F1-score.

$$IoU = \frac{TP}{(TP + FP + FN)} \qquad \text{Eq. (1)}$$

The acquired IoU is calculated on class-wise basis and is then averaged.

$$precision = \frac{TP}{TP + FP} \qquad \text{Eq. (2)}$$

$$recall = \frac{TP}{TP + FN} \qquad \text{Eq. (3)}$$

$$F_1 = \frac{2TP}{2TP + FN + FP} \qquad \text{Eq. (4)}$$

5. Results and Discussions

Here, the dataset on cardiac MRI images is provided by ExaWizards Inc., [24]. The whole data set was comprised of 1964 images in DICOM format upon which standardization and normalization are applied. The cardiac dataset in gray scale is treated as input to the U-Net model towards detection and segmentation of coronary artery which is the known object of interest. The model is implemented in Keras with TensorFlow backend with python as coding language.

Table 2: Depicts the parameters with all the 5 Trials

PARAMETERS	TRIAL 1	TRIAL 2	TRIAL 3	TRIAL 4	TRIAL 5
Optimizer	Adam	Adam	Stochastic Gradient Descent	Adam	Adam
Loss function	Binary cross entropy	Binary cross entropy	Binary cross entropy	Binary cross entropy	Binary cross entropy
Learning rate	1e-5	1e-5	0.0001	1e-5	1e-5
Epoch size	50	50	70	70	100
Batch size	4	4	4	4	4
Training accuracy	97.46%	98.10%	97.61%	98.13%	98.17%
Validation accuracy	97.88%	98.28%	97.89%	98.24%	96.2%
Testing accuracy	97.13%	97.20%	97.06%	98.26%	97.20%
Train loss	0.04040%	0.01033%	NA	0.0086%	0.0072%
Validation loss	0.03186%	0.01735%	NA	0.01627%	0.0192%
Early stopping	04-(No. of epochs trained)	NA	NA	NA	NA
Total time taken for training	31m 59s	8h 43m 50s	12h 18m 9s	12h 19m 15s	17h 44m 33s

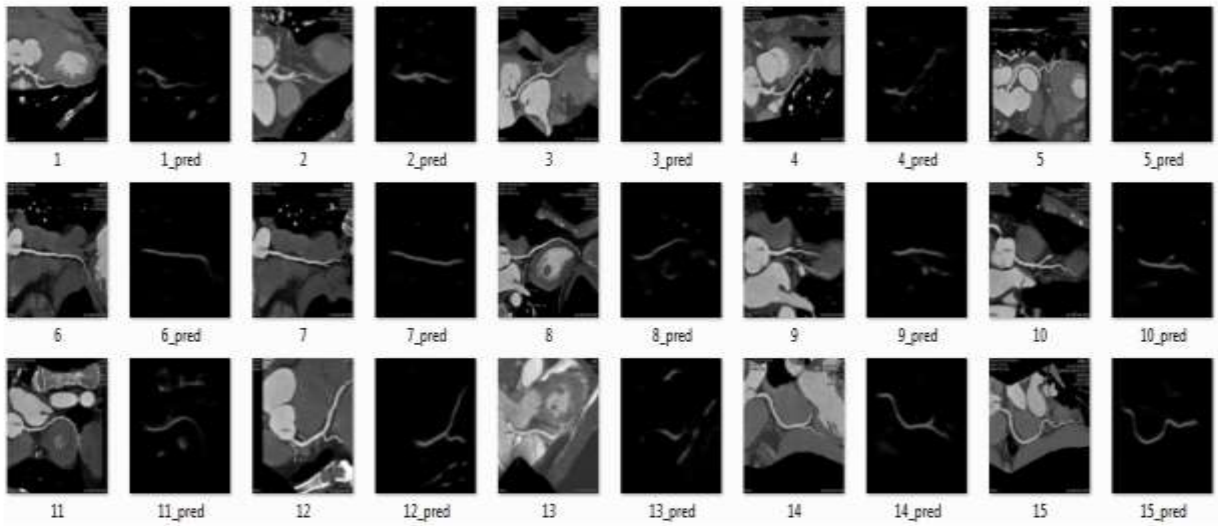

Fig. 2 Qualitative results after using early stopping callback at the end of the first trail with both original and segmented images

The implemented model is trained with 488 training images, validated on 40 images and tested on 17 images along with their corresponding ground truth images which are in fact manually delineated for the purpose of evaluation. All the training, validation and testing datasets are gray scale cardiac images of dimensions 512*512. A softmax activation function is arranged at the last layer so as to obtain the desired predictions to reflect the probability of a particular pixel being pertinent to a coronary artery or not. There are all together 5 different trials executed with this U-Net model by varying the parameters to perform the fine-tuning and the results are tabulated in table2.

From figure.2 which is the initial trail of the system, it has been perceived that the model is predicting parts of the coronary artery but not accurately due to the early stopping callback used in training process. Figure. 3 is the final trail of the system which depicts the original images and their corresponding segmented images.

The below table 3 illustrates the performance of the various evaluation metrics after imposing the U-Net model over the cardiac MRI scans.

Because of its wide application for both research and production on deep learning, TensorFlow [25] is used throughout this project and all the results are thus obtained from it. All the experiments pertaining to cardiac MRI dataset run a cloud machine armed with K80 GPU [26] with 12GB RAM on Microsoft Windows Azure. The U-Net architecture has consumed ~640 seconds on Tesla K80 for each epoch on training.

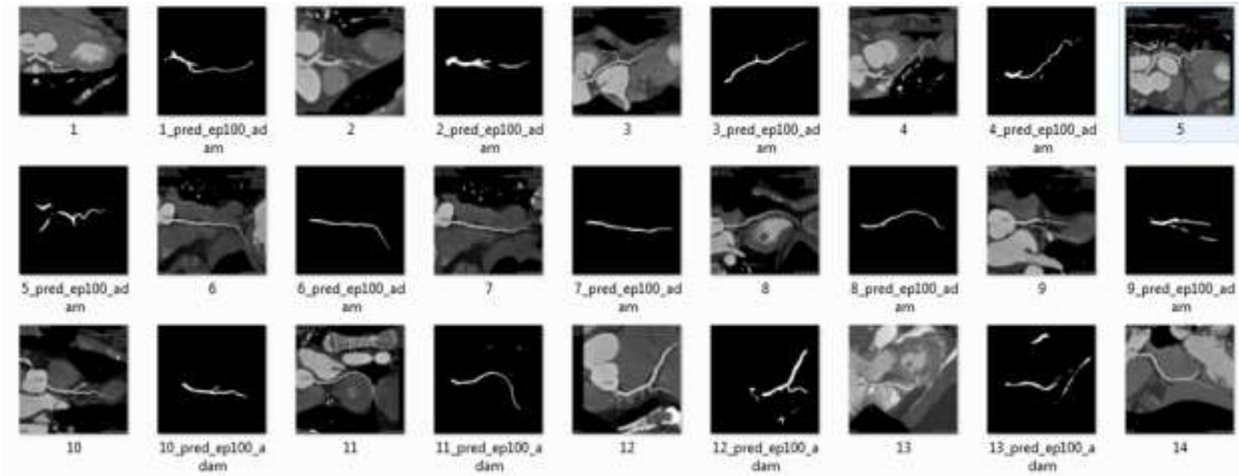

Fig. 3 Qualitative results at the end of the trail 5 with both original and segmented images

Table 3: Depicts the performance of various evaluation metrics over the U-Net model

Evaluation metric / Architecture	Mean IoU (%)		Precision(%)		Recall (%)		F1-score (%)	
	Train	Test	Train	Test	Train	Test	Train	Test
U-Net	97.16	96.62	80.75	77.95	84.12	81.72	92.77	89.22

6. Conclusions

In the arena of biomedical imaging community, U-Net is an ideal model that has turned out with very impressive results. It is a FCN based CAEs which is regarded as one of the most popular and successful methodology. Particularly, in the breadth of clinical applications this versatile model contributes greater accuracy provided decent training, sufficient dataset and adequate training time. Finally, the system is trained with U-Net architecture which has delivered venerable results of 96.62% mean IoU.

Currently there is a drastic increase in the number of complexities in the modalities of medical images such as CT, PET, MRI, fusion imaging and ultrasound etc. So, it is quite difficult for the radiologists to perform imaging tests to fetch adequate time in reading them and end up with precise reports. Yet, with the development of deep learning technology, it is becoming easy for the radiologists to analyze lesions which generate reports automatically that are suspicious within a short period of time.

Acknowledgements

This research work is supported by ExaWizards Inc., Tokyo, Japan under MoU with Andhra University.

References

[1] Chen, Z., & Liu, B, "Lifelong Machine Learning", *Synthesis Lectures On Artificial Intelligence And Machine Learning*, 10(3), 1-145. doi: 10.2200/s00737ed1v01y201610aim033, 2016.

[2] LeCun, Y., Bengio, Y., & Hinton, G, "Deep learning", *Nature*, *521*(7553), pp.436-444, 2015.

[3] Schmidhuber, J.,"Deep learning in neural networks", *An overview. Neural Networks*, 61, pp.85-117, doi: 10.1016/j.neunet.2014.09.003, 2015.
[4] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ...& Sánchez, C. I, "A survey on deep learning in medical image analysis", *arXiv preprint arXiv:1702.05747*, 2017.

[5] Long, J., Shelhamer, E., & Darrell, T, "Fully convolutional networks for semantic segmentation", In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431-3440, 2015.

[6] Wang, J., Mackenzie, J. D., Ramachandran, R., & Chen, D. Z. "A Deep Learning Approach for Semantic Segmentation in Histology Tissue Images. Medical Image Computing and Computer-Assisted Intervention – MICCAI", Lecture Notes in Computer Science,176-184. doi:10.1007/978-3-319-46723-8_21, 2016.

[7] Lieman-Sifry, J., Le, M., Lau, F., Sall, S., & Golden, D. "FastVentricle: Cardiac Segmentation with ENet", In: *International Conference on Functional Imaging and Modeling of the Heart* (pp. 127-138), Springer, Cham, 2017.

[8] Mortazi, A., Karim, R., Kawal, R., Burt, J., &Bagci, U, "CardiacNET: Segmentation of Left Atrium and Proximal Pulmonary Veins from MRI Using Multi-View CNN", *arXiv preprint arXiv:1705.06333*. 2017.

[9] Luo, G., An, R., Wang, K., Dong, S., & Zhang, H, "A deep learning network for right ventricle segmentation in short-axis MRI", In: *Computing in Cardiology Conference (CinC), 2016* (pp. 485-488). IEEE, 2016.

[10] Avendi, M. R., Kheradvar, A., &Jafarkhani, H. "Combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI", *Medical image analysis*, *30*, pp.108-119, 2016.

[11] Emad, O., Yassine, I. A., &Fahmy, A. S. "Automatic localization of the left ventricle in cardiac MRI images using deep learning", In: *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* ,pp. 683-686), IEEE, 2015.

[12] Tran, P. V. "A fully convolutional neural network for cardiac segmentation in short-axis MRI", *arXiv preprint arXiv:1604.00494*, 2015.

[13] Zotti, C., Luo, Z., Lalande, A., Humbert, O., &Jodoin, P. M, " Novel Deep Convolution Neural Network Applied to MRI Cardiac Segmentation", *arXiv preprint arXiv:1705.08943*, 2017.

[14] Poudel, R. P., Lamata, P., & Montana, G. "Recurrent fully convolutional neural networks for multi-slice mri cardiac segmentation", In: *International Workshop on Reconstruction and Analysis of Moving Body Organs*, pp. 83-94, Springer, Cham, 2016.

[15] Dong, S., Luo, G., Sun, G., Wang, K., & Zhang, H. "A left ventricular segmentation method on 3D echocardiography using deep learning and snake", In: *Computing in Cardiology Conference (CinC), 2016*, pp. 473-476, IEEE, 2016.

[16] Ronneberger, O., Fischer, P., &Brox, T, "U-net: Convolutional networks for biomedical image segmentation", In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234-241, Springer, Cham, 2016.

[17] Dong, H., Yang, G., Liu, F., Mo, Y., &Guo, Y. "Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks", *Communications in Computer and Information Science Medical Image Understanding and Analysis*, 506-517. doi:10.1007/978-3-319-60964-5_44, 2017.

[18] Ciresan, D. C., Meier, U., Masci, J., Maria Gambardella, L., & Schmidhuber, J. Flexible, "High performance convolutional neural networks for image classification", In: *IJCAI Proceedings-International Joint Conference on Artificial Intelligence* ,Vol. 22, No. 1, pp. 1237, 2011.

[19] Makhzani, A., & Frey, B. "A winner-take-all method for training sparse convolutional autoencoders", *arXiv preprint arXiv:1409.2752*, *6*, 2014.

[20] Sammons, D., Winfree, W. P., Burke, E., &Ji, S. "Segmenting delaminations in carbon fiber reinforced polymer composite CT using convolutional neural networks", doi:10.1063/1.4940585, 2016.

[21] Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., & Pal, C. "The importance of skip connections in biomedical image segmentation", In: *International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*, pp. 179-187, Springer International Publishing, 2016.

[22] Nair, V., & Hinton, G. E. "Rectified linear units improve restricted boltzmann machines", In: *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 807-814, 2010.

[23] Taha, A. A., & Hanbury, A. "Evaluation Metrics for Medical Organ Segmentation and Lesion Detection." *Cloud-Based Benchmarking of Medical Image Analysis*, pp.87-105. doi:10.1007/978-3-319-49644-3_6, 2017.

[24] (n.d.). Retrieved from https://exawizards.com/en/

[25] TensorFlow – Medium. (n.d.). Retrieved from https://medium.com/tensorflow

[26] Shi, L., Liu, W., Zhang, H., Xie, Y., & Wang, D. "A survey of GPU-based medical image computing techniques." *Quantitative imaging in medicine and surgery*, 2(3), 188, 2012.