

# AN OVERVIEW OF DEEP LEARNING TECHNIQUES IN ECHOCARDIOGRAPHY IMAGE SEGMENTATION

MEHDI SAMIEIYEGANEH<sup>1</sup>, PROFESSOR RAHMITA WIRZA BT O. K. RAHMAT<sup>2</sup>,  
Dr. FATIMAH BINTI KHALID<sup>3</sup>, Dr. KHAIRUL AZHAR BIN KASMIRAN<sup>4</sup>

<sup>1</sup>) Faculty of Computer Science and Information Technology, University Putra Malaysia (UPM), Serdang, MALAYSIA,

<sup>2</sup>) Faculty of Computer Science and Information Technology, University Putra Malaysia (UPM), Serdang, MALAYSIA,

<sup>3</sup>) Faculty of Computer Science and Information Technology, University Putra Malaysia (UPM), Serdang, MALAYSIA,

<sup>4</sup>) Faculty of Computer Science and Information Technology, University Putra Malaysia (UPM), Serdang, MALAYSIA,

E-mail: <sup>1</sup>gs51922@student.upm.edu.my, <sup>2</sup>rahmita@upm.edu.my, <sup>3</sup>fatimahk@upm.edu.my ,

<sup>4</sup>k\_azhar@upm.edu.my

## ABSTRACT

Machine Learning (ML) has been a remarkable success in the last few years, Sequential -Decision Making tasks are a main topic in ML, these are tasks based on deciding, the sequence of actions from experience carry out in an environment that is uncertain to achieve goals. these tasks cover so many ranges of applications such as healthcare, robotic, finance and many more. The design and extracting of features in ML were done based on defining (hand-crafting features), which is a weak point for this model. Due to ML problem as well as advances in computer hardware Machine Deep Learning (DL) has entered the field of image processing. In fact, DL is a type of Function approximator. To solve ML tasks Function approximator is required and this idea is core of ML. There are many type of Function approximator such as, linear models, gaussian process, support vector machine and decision tree. In this paper, considering the importance of segmenting in medical images, we will review works that have utilized DL methods, as well as our focus is based on Echocardiography Image Segmentation.

**Keywords:** *Machine Learning, Deep Learning, Medical Image Segmentation, Echocardiography.*

## 1. INTRODUCTION

Basically, machine learning methods can be grouped into three categories: Supervised Learning, Unsupervised Learning and Reinforcement Learning. Until in 1960s, there was confusion about the two modes of reinforcement learning and supervised learning, at this time, Sutton and Barto [1] came up with solutions. To concisely define the Supervised Method, the data need to be trained to becomes feature vectors, each data need to be labelled by an external supervisor and based on ML algorithms, generate the predictive model, For example in the neural structure, this model is generated in the Neural Network (NN) weights, the model, based on new data samples that do not have labels, should find the appropriate label for them, In Supervise Learning, model discovery is based on the natural structure of training data so that new data should fit into appropriate clusters and the

training data in this model is not labelled. Reinforcement learning is the process in which the agent is used.

This process operates based on trial and error with the environment and the agent receives signals from the environment that are rewarded or punished that are commensurate with the agent's performance. (Figure 1).

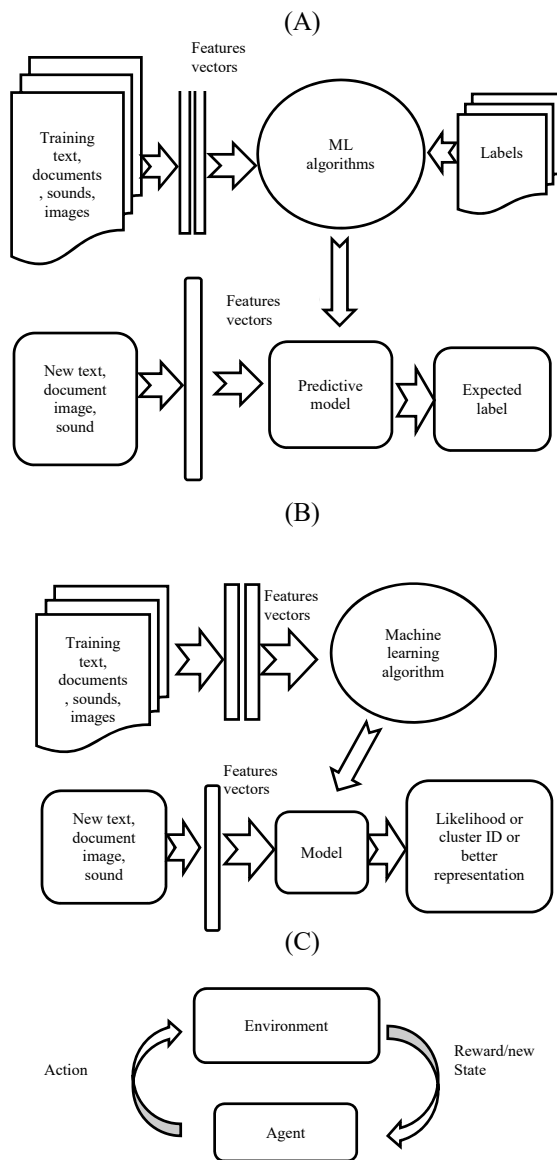


Figure 1- Shows here the flow of (A) Supervised Learning. (B) Unsupervised Learning and (C) Reinforcement Learning.

From the year 2000 onwards, with the advancement of computer hardware, deep learning has entered the field of image processing and has achieved excellent results [2]. One of the important aspects of DL is the use of Neural Networks (NNs) which can find compact representation in high dimensional data. In this case, the problem of manual features in ML will disappear [3- 4]. In fact, Deep Neural Networks (DNNs) consists of multiple layers of neurons or perceptron that are connected in an inter-layer fashion [5]. Learning of Machine in highly complex mathematical models can be demonstrated by deep

learning framework that can be used to perform accurate analysis of data.

In this paper, we do not separate the use of DNNs in unsupervised learning, supervised learning, or reinforcement learning so they are concerned with the overall context. It is important to note that the key concepts of DNNs are better understood in supervised learning settings. The same ideas are easily applicable in unsupervised learning. But the important point is that in many literature reviews, such as [2 -19, 24], there are no any issues about Deep Reinforcement Learning (DRL) in segmenting of medical images, and this is an important challenge for future work. It should be noted that by combining deep learning and reinforcement learning, deep reinforcement learning has emerged [3].

According to opinion of World Health Organization (WHO), cardiovascular disease is the leading cause of death in the world, with approximately eighteen million people dying from cardiovascular disease in the year 2016[6]. That number is increasing every year. This could be a reason for increased research in this area, and the number of articles on the use of deep learning in this area is increasing. Techniques such as Magnetic Resonance Imaging (MRI), Echocardiography (Cardiac Ultrasound) and Computed Tomography (CT) are used to diagnose and predict cardiovascular disease, each of which offers different ways and structures to evaluate cardiovascular disease. Because of echo features, such as its low cost of the machine, which costs about one hundred and two hundred and fifty thousand dollars compared to the MRI, which is between one and a half to two and a half million dollars, as well as easy portability as the first method used in the diagnosis and imaging of heart and heart disease.

Image segmentation is a fundamental task as well as a challenge on the computer vision, in this work two types of information is need, including, global object context and local boundary position. Pixel detection of lesions or organs in medical images such as MRI or US involves medical image segmentation. This is a challenging task that provides us with important information about the shapes and sizes of organs. For these reasons, in this article we review papers based on segmentation using DL in Cardiac Ultrasound (US).

It is necessary to mention that in 1960s, ML was used in medical imaging. [7- 8]. However, DL techniques emerged in 1990s in the medical imaging, as mentioned in [9- 10- 11- 12, 13]. In this review paper, general review of recent DL techniques in medical imaging segmentation focuses on US has been prepared. There have been many

articles on deep learning contributions in medical image, such as segmentation, Detection/localization, and Registration and Landmark detection, in the various organs of the body such as the eye, brain, chest, Abdomen, etc. Please refer readers to the [19] for reading.

In search engines like PubMed and Scopus and also conference proceedings for ISBI, EMBC and MICCAI ,we were looking for articles that included keywords like “Deep Learning”, “Machine Learning” “Convolutional Neural Networks(CNN)”, “Echocardiography or Cardiac Ultrasound” and “image segmentation or medical image segmentation”. Based on our studies, many articles have been published on the application of DL methods for segmenting and analyzing medical images. [14-15- 16- 17, 18], However, none of these studies have systematically focused on segmenting US medical images. This paper consists of three important sections, in the second section provides a brief overview of the structures of DNN most commonly used in medical image processing and analysis. In the third section, mentioned review network training techniques; fourth section is based on review articles that have done some work on Cardiac Ultrasound image segmentation. In the final section, both the challenges and future work will be examined. In the reviewed literature, most articles focused on the Left Ventricle (LV) and the Left Atrium (LA), and this may be because the function of the left heart is more important in the diagnosis of cardiovascular problems.

## 2. STRUCTURES OF DEEP NEURAL NETWORKS

### 2.1 Convolutional Neural Networks (CNNs)

In deep learning techniques, the use of CNNs is important in image analysis [20-21]. Such networks have been successfully applied to advance the state of art on image processing and computer vision, such as image segmentation, classification, object detection and landmark detection. These networks consist of several layers like convolution, pooling and fully connected layers and each layer performing a specific operation. The CNNs have an initial layer as the input layer and this layer is directly connected to the input image and the number of pixels in the image is equal to the number of neurons in the input layer. In convolutional neural networks, there are a stack of functional layers as well as an output layer. These functional layers are responsible for converting inputs into outputs in forms such as vectors. These layers are also called convolution layers and their aim is to learn weights.

Each convolution uses filters or kernels that perform the convolution operation on the image these filters or kernels are of arbitrary size that are based on designer , for example, if the input image is two-dimensional  $n \times n$  kernel or if it is three-dimensional  $n \times n \times n$  kernel followed by batch normalization can be used.

After which the output is passed through a nonlinear activation function such as ReLU which is used to feature extraction maps from an image.[22] The pooling layers reduce redundant properties, which improves statistical efficiency and reduced the height and width of the activation maps [5]. Fully connected layers were used for reducing the dimension of properties, which also for finding properties of task-relevant for inference [23]. The structure of a CNN and the sample used for the cardiac MRI image showed in Figure2.

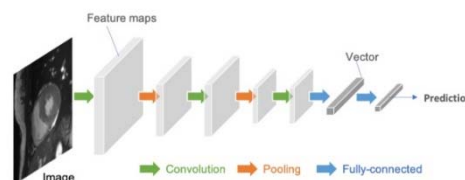
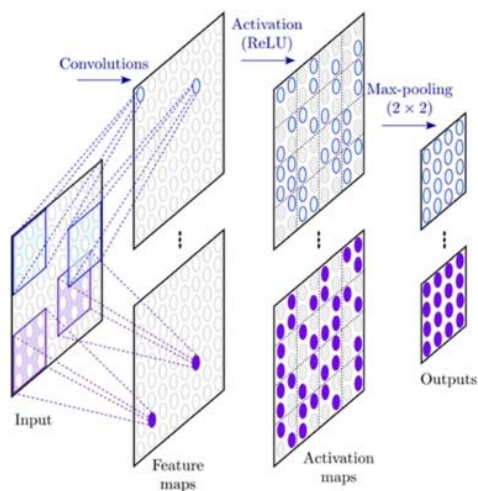


Figure 2- The structure of a CNN [23-24].

Without major adaptations to the network architecture, CNNs for classification of image can also be used for segmentation of image [25]. In this case, patches can be obtained by dividing the image. Then a CNN is trained to predict the central pixel

class label for every patch. The problem of inference time in this method is obvious and is a weak point for it. For each patch individually the network needs to be deployed. There is also a great deal of redundancy due to overlapping patches in the image. Because of this problem aims of CNNs with fully connected layers is to estimate the bounding box of interested object in an image that called object localization. Crop the image is based on this bounding box, creating an image preprocessing step to reduce segmentation computational cost [26]. In the following, the application of CNNs structures in segmentation of medical image will be mentioned.

## 2.2 CNN in 2D and 2.5D

In this case, segmentation will be performed with applying filters that are 2D upon 2D input image. In [27] images in the form of 2D based on multiple sources of information are moved to a CNN input layer, images are in various channels such as R, G, B. The purpose of this survey is to investigate whether segmentation results would be improved if use multi-modality images i.e., breast or Brain MRI, and Cardiac CT images as inputs.

2.5 D methods have lower computational costs than 3D and the wealthy spatial information of neighboring pixels [28- 29, 30]. With the two-dimensional kernel, 3 orthogonal two-dimensional patches are extracted, respectively, in the XY, YZ, and XZ. For multitask segmentation in [28] and knee cartilage segmentation in [31] this idea applied.

Learning the 2.5D approaches has the advantage that the system works with 2D labeled data, which are much more accessible than labeled 3D data and hardware fits in with such systems. This approach has its disadvantages as well, for example, in [32] authors believe that using volumetric medical data would not be optimal to use three orthogonal views rather than more views in 3D images. The 2.5D methods use two-dimensional kernels and are limited to such kernels and cannot be used in 3D filters.

## 2.3 CNN in 3D

3D CNNs is using to derive a stronger volumetric representation in all three X and Z axes. Depend on the content of surrounding 3D patches the central voxel label based on 3D network training will be predicted. The structural difference that 3D CNNs have with 2D CNNs is that in 3D networks applying 3D modules in each section. In [33] first 3d models was applied to brain tumor segmentation with arbitrary size.in [34] authors using 3D CNN and a cut-off threshold function performed brain boundary detection. Deep model to extract highly

informative features is required to segment complicated volumetric images of an organ but training of these types of networks is considered as a significant challenge in 3D models [5].

## 2.4 Fully Convolutional Neural Networks (FCNs)

FCNs [35] which is a special type of CNNs with fully connected layer, later was replaced with a fully convolutional layer. Structure of encoder and decoder designed in FCNs, able to take the input of the desired size and produce the output of the same size. FCNs have a dense pixel- wise prediction, decoder based on given an input image converts the input into high level feature representation and decoder responsible for interprets maps of feature and recovers spatial details back to the image space for pixel-wise prediction through a series of transposed convolution and convolution operations [24]. Figure3. shows the structure of the FCNs.

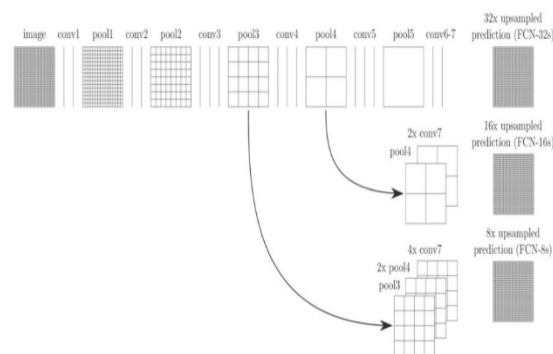


Figure 3- The structure of the FCNs [35].

With Similar modalities as well as identical datasets but with application of [36] in [37] have shown that the FCN has performed better than the CNN.

## 2.5 U-Net in 2D and 3D

One of the most well-known and popular type of FCNs that is actually built on the beautiful FCN architecture is the U-Net that proposed by [38].U-Net engages skip connections between different stages of network (encoder and decoder) with this structure it can recover spatial context loss in the down-sampling path, making more precise segmentation and also In this model the network depth is increased to nineteen layers [39]. Figure 4. shows the structure of the U-Net.

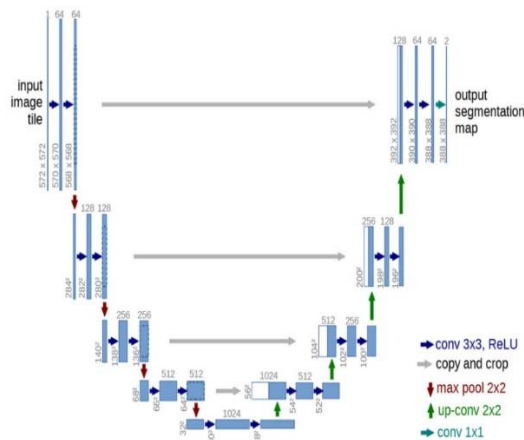


Figure 4- The structure of the U-Net [38].

Many changes were made to the segmentation of medical images according to the structure of the U-Net, and this structure was highly regarded [40- 41, 42]. Many of State-of-the-art methods have used the U-Net structure to medical images segmentation, for example in [41] lung segmentation explored in X-ray scans with use of U-Net and their results showed that the U-Net is capable of segmenting accurately and at high speed.

In an attempt to reinforce the structure of U-Net networks, a 3D model of such networks was introduced by [40]. Generate dense volumetric segmentation with this model from some 2D annotated slices is one of the capabilities of this model. Much work has been done on this model, for example vascular boundary detection with 3D U-Net was proposed by [43]. One of the extractions of the U-Net Networks is V-Net that proposed by [44]. And given the success of such networks, there were hopes for a good and accurate segmentation of the cardiac images.

### 2.6 Recurrent Neural Networks (RNNs)

Another type of Artificial Neural Networks (ANNs) is RNNs used for sequential data such as MRI and US. RNNs can enable the network to remember and memorize the patterns from the past so it can use past knowledge to make decisions. Other capability of such networks includes that such network from the input slices can extract inter slice contexts. By combining RNN and 2D FCN to segmentation of cardiac, a network is created that can receive information from adjacent slices to improve the inter-slice coherence of segmentation results [45]. To model long-term memory, two architectures of the RNNs family are used, Long Short-Term Memory (LSTM) proposed by [46] and Gated Recurrent Unit (GRU) proposed by [47]. In fact, GRU is an alteration of LSTM. Because of

standard LSTM network inputs need to be vectorized inputs, the spatial information for medical image segmentation will be lost. To solve this problem, [48, 49] proposed a convolutional LSTM network to use the convolutional operation instead of the multiplication vector. In GRU the memory cells are deleted and has been simplified without restructuring, and this simplification does not change the performance of such networks [50].

### 3. TECHNIQUES FOR TRAINING NEURAL NETWORKS

Neural network training prior to reviewing its results requires a dataset for training and testing, an optimizer, and a loss function for updating model parameters. In fact, the network prediction error and sending signals in the form of backpropagation that are used to optimize and update the network parameters, are tasks of this function. Minimizing loss function based on finding proper values of the networks parameters is goal of network training.

#### 3.1 Deeply Supervised

In this method, direct observation of the hidden layers is performed and then propagated to the lower layers. This technique has been used by [51], albeit non-medical cases. For medical purposes, this technique was used by [52] to 3D-segment liver in CT.

#### 3.2 Weakly Supervised

One of the problems that the supervised technique for automatic segmentation of medical images face is the need for pixel-level annotation that is not available in many cases. And even if this annotation exists, it's a very boring and expensive [53]. This problem can be solved by the outsource labeling services, the same as Amazon MTurk did, though not related to medical images. Another solution implemented by [54], is the use of labeled image data, which indicates the presence or absence of a pattern.

#### 3.3 Transfer Learning

If a system, for a novel task, recognizes and uses the knowledge learned from the previous source, it will be called transfer learning [55]. Transfer learning can be used in general as well as medical images in form of fine tuning a network pre-trained [56]. This technique has been shown to work better if the target network tasks and the source network tasks are the same and similar [57]. It is important to note that due to the different appearance and size of different organs in medical images, the

transfer learning in models that are highly variable may not change the segmentation result [2].

### 3.4 Common Loss Functions

Many Loss Functions are used for neural network training, which is mentioned in this section, which are used for tasks such as regression, classification and segmentation.

A simple loss function in regression tasks such as Landmark Detection, Heart Localization, and Image Reconstruction is the use of Mean Squared Error (MSE) as follows (Formulae 1):

$$L_{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

$y_i$  And  $\hat{y}_i$  represent vector of target values and vector of the predicted values, respectively, and also the samples of data is displayed with  $n$ .

For functions such as classification and segmentation of images, Cross-Entropy (CE) Loss Function can be used as follows (Formulae 2):

$$L_{CE} = -\frac{1}{n} \sum_{i=1}^n \sum_{c=1}^C y_i^c \log(p_i^c) \quad (2)$$

For segmentation this loss function summarizes the pixel-wise probability errors between the predicted probabilistic output  $p$  and its corresponding target segmentation map  $y$  for each class  $c$ , number of all classes demonstrated with  $c$  [24].[58] Offers a Loss Function that is used only for object segmentation that called Soft-Dice loss function as follows (Formulae 3):

$$L_{Dice} = 1 - \frac{2 \sum_{i=1}^n \sum_{c=1}^C y_i^c p_i^c}{\sum_{i=1}^n \sum_{c=1}^C (y_i^c + p_i^c)} \quad (3)$$

In this loss function, penalizes the mismatch between a predicted segmentation map and its target map at pixel-level [24].

## 4. SEGMENTATION OF ECHOCARDIOGRAPHY IMAGES

This section deals with the segmentation of Echocardiography images, most work has been done on the LA and LV.

Voxel-wise tissue classification because of the nature and quality of the US imaging is very challenging. To solve this problem in [59-60- 61- 62- 63- 64, 65], on 2D images of the LV, the DL technique is combined with deformable model. As mentioned before, the features are extracted using trained DNNs which are much more accurate than handcrafted features. Applied DL in a two-stage pipeline increases robustness of segmentation framework and also reduces the search region of the segmentation, this was done by [60, 61] on the left ventricle in apical long axis echocardiograms. Given

that the data recorded in the Cardiac ultrasound are timed and sequential, several approaches aim to leverage the connection between temporally close frames to improve the accuracy and robustness of the LV segmentation. In [61, 62] based on a Sequential Monte Carlo (SMC) proposed a dynamic method, that with using pervious cardiac phases, current cardiac phase will be done. The results are better than the work done in [59]. In By combining LSTM, U-Net and inter-frame optical flow to employ multiple frames for segmenting one target frame showed that this method was more accurate [66]. One of the challenges that will address in the next section is the lack of training data in deep learning. One of the methods used to increase the number of training data is proposed by [67, 68]. In fact, they use algorithms that do not work as deep learning to generate labels of unlabeled data. Another work that has been done in this field is by [69], with using a Kalman filtering based method and the images annotated with this filter they have trained a U-Net. Tasks of segmenting US images are not limited to the use of labelled data. For example, a CNN network trained in [70] using a data set that is partially labelled. Also enables training on both the labelled and unlabeled images with a semi-supervised framework proposed by [71].

Segmenting 3D US images is much more difficult and challenging than 2D images. The reason is that 3D networks have more parameters than 2D networks, so 3D networks for optimization have complex computations, and due to the lack of 3D training data set they may encounter with Over-Fitting. The Over-Fitting problem is explained in the next section. One solution to reduce computational costs in 3D data is coarse 2D segmentation maps [72], with this solution, avoided direct processing on 3D data. For example, in [73] a two-stage method based on the coarse 2D segmentation maps solution is proposed.in this work first applying a 2D CNN to produce coarse segmentation maps on 2D slices from a 3D volume. Other work done to segment the left ventricle and the left atrium include [74- 75, 76]. In [74] view classification, LA and LV segmentation and detection of cardiovascular diseases demonstrated based on CNNs applicability. By combining the three methods of marginal space learning (MSL), deep neural networks (DNNs) and active shape model (ASM) to segment the aortic valve in 3D cardiac ultrasound volumes in [77], The authors show that they have achieved better results than [78].

## 5. CHALLENGES IN DEEP MODELS

### 5.1 Overfitting

If a model obtains the regularities and patterns in the training data with greater accuracy than unprocessed instances of the problem, overfitting has occurred [79]. In fact, the small number training data causes this problem. Any solution used to increase the amount of training data is helpful to solve this problem. Examples of solutions to reduce this problem are, weight initialization [80], Dropout [81], Transfer Learning and Ensemble learning [82]. A public dataset for 2D US were made available in the year 2019 with name of CAMUS and reference of [76].

### 5.2 Organ Appearance

Another major challenge in segmenting medical images is the heterogeneous appearance of the organs being examined, in fact, segmentation of such organs is an important challenge as body parts appear in different shapes and sizes in different patients. It is also important to note that this is a very difficult task in US because of moving the Heart. Increasing the depth of network is reported as an effective solution [83].

## 6. FUTURE WORK AND CONCLUSION

Segmenting medical images with deep learning methods is expanding in this field many DL methods use Supervised or Unsupervised learning methods. As a result, very little work has been done on the processing of medical images based on RL. For example, by [84], in the heart and brain, based on DRL method, Landmark Detection is done. However, DRL has not been done in medical images segmentation, especially in US. A challenging task can be posed by the question: How can we simulate a physician's performance in US image segmentation using one or more agents that are based on DRL? In fact, what is the use of agents to simulate when a doctor picks up a pen to segment different parts of the heart? This is exactly what we are looking for. As mentioned above, by combining deep learning and reinforcement learning, deep reinforcement learning has emerged, reinforcement learning algorithms work on a single agent, and the environment of such algorithms is formulated as a Markov Decision Process [3]. To increase one agent to several, it needs a combination of DRL and Multi-agent methods. The combination of these creates a new context called Multi-agent Deep Reinforcement Learning (MADRL). For further study in this field, two review papers are presented in [3, 85].

In this review paper, we provided an overview of DL techniques used in US images segmentation. And also challenges of using DL techniques and exploring the types of DNNs used in medical Images segmentation was examined.

## ACKNOWLEDGEMENT

Interpolation Mesh Fitting to Reconstruct Quantitative 3D Geometrical Model of the Tricuspid Aortic Valve in 3D Echocardiographic Images for Cardiac Disease Prediction, FRGS/2/2014/ICT05/UPM/02/2

## REFERENCES

- [1] Sutton, R. S., and Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT press.
- [2] M. H. Hesamian, W. Jia, X. He, and P. Kennedy. "Deep learning techniques for medical image segmentation: Achievements and challenges". *Journal of Digital Imaging*, pages 1–15, 2019.
- [3] Hernandez-Leal, P., Kartal, B., and Taylor, M. E. (2018). "Is multiagent deep reinforcement learning the answer or the question? a brief survey". *arXiv preprint arXiv:1810.05587*
- [4] K. Arulkumaran, M. P. Deisenroth, M. Brundage, A. A. Bharath, *A Brief Survey of Deep Reinforcement Learning*. [online] available at: <http://arXiv.org/abs/1708.05866v2>
- [5] F. Altaf, S. Islam, N. Akhtar, N. K. Janjua, *Going deep in medical image analysis: Concepts methods challenges and future directions* in arXiv:1902.05655, 2019, [online] Available at : <https://arxiv.org/abs/1902.05655>.
- [6] [online] available at : [https://www.who.int/cardiovascular\\_diseases/about\\_cvd/en/](https://www.who.int/cardiovascular_diseases/about_cvd/en/)
- [7] H. Becker, W. Nettleton, P. Meyers, J. Sweeney, and C. Nice, "Digital computer determination of a medical diagnostic index directly from chest x-ray images," *IEEE Transactions on Biomedical Engineering*, no. 3, pp. 67–72, 1964.
- [8] G. S. Lodwick, T. E. Keats, and J. P. Dorst, "The coding of roentgen images for computer analysis as applied to lung cancer," *Radiology*, vol. 81, no. 2, pp. 185–200, 1963.
- [9] Y. Wu, M. L. Giger, K. Doi, C. J. Vyborny, R. A. Schmidt, and C. E. Metz, "Artificial neural networks in mammography: application to decision making in the diagnosis of breast

- cancer.” *Radiology*, vol. 187, no. 1, pp. 81–87, 1993.
- [10] S.-C. Lo, S.-L. Lou, J.-S. Lin, M. T. Freedman, M. V. Chien, and S. K. Mun, “Artificial convolution neural network techniques and applications for lung nodule detection,” *IEEE Transactions on Medical Imaging*, vol. 14, no. 4, pp. 711–718, 1995.
- [11] B. Sahiner, H.-P. Chan, N. Petrick, D. Wei, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, “Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images,” *IEEE transactions on Medical Imaging*, vol. 15, no. 5, pp. 598–610, 1996.
- [12] H.-P. Chan, S.-C. B. Lo, B. Sahiner, K. L. Lam, and M. A. Helvie, “Computer-aided detection of mammographic microcalcifications: Pattern recognition with an artificial neural network,” *Medical Physics*, vol. 22, no. 10, pp. 1555–1567, 1995.
- [13] W. Zhang, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, “An improved shift-invariant artificial neural network for computerized detection of clustered microcalcifications in digital mammograms,” *Medical Physics*, vol. 23, no. 4, pp. 595–601, 1996.
- [14] Litjens G, Kooi T, Bejnordi BE, Setio A, Ciampi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI: “A survey on deep learning in medical image analysis”. *Med Image Anal* 42:60–88, 2017
- [15] Shen D, Wu G, Suk HI: “Deep learning in medical image analysis”. *Annu Rev Biomed Eng* 19:221–248, 2017
- [16] Greenspan, H., Van Ginneken, B., and Summers, R. M. (2016). Guest editorial deep learning in medical imaging: “Overview and future promise of an exciting new technique”. *IEEE Transactions on Medical Imaging*, 35, 1153–1159
- [17] Gandhi, S., Mosleh, W., Shen, J., and Chow, C.-M. (2018). “Automation, machine learning, and artificial intelligence in echocardiography: A brave new world”. *Echocardiography* 35, 1402–1418. doi:10.1111/echo.14086
- [18] Mazurowski, M. A., A., and Bashir, M. R. (2019). “Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI”. *Journal of magnetic resonance imaging* 49, 939–954. doi:10.1002/jmri.26534
- [19] F. Altaf, S. Islam, N. Akhtar, N. K. Janjua, “Going deep in medical image analysis: Concepts methods challenges and future directions” in arXiv:1902.05655, 2019, [online] Available at: <https://arxiv.org/abs/1902.05655>.
- [20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [21] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [22] Ioffe, S. and Szegedy, C. (2015). “Batch normalization: accelerating deep network training by reducing internal covariate shift”. In *ICML (JMLR.org)*, 448–456
- [23] Commandeur F, Goeller M, Betancur J, Cadet S, Doris M, Chen X, Berman DS, Slomka PJ, Tamarappoo BK, Dey D: “Deep learning for quantification of epicardial and thoracic adipose tissue from non-contrast CT”. *IEEE Trans Med Imaging* 37(8):1835–1846, 2018.
- [24] Chen Chen , Chen Qin , Huaqi Qiu , Giacomo Tarroni<sup>1,2</sup>, Jinming Duan, Wenjia Bai, and Daniel Rueckert , “Deep learning for cardiac image segmentation: A review”. [online] available at: <https://www.frontiersin.org/articles/10.3389/fcvm.2020.00025/full>
- [25] Ciresan, D. C. and Giusti, A. (2012). “Deep neural networks segment neuronal membranes in electron microscopy images”. In *Conference on Neural Information Processing Systems*. 2852–2860
- [26] Avendi, M. R., Kheradvar, A., and Jafarkhani, H. (2016). “A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac mri”. *Medical Image Analysis* 30, 108–119.
- [27] Zhang W, Li R, Deng H, Wang L, Lin W, Ji S, Shen D “Deep convolutional neural networks for multi-modality isointense infant brain image segmentation”. *NeuroImage* 108:214–224, 2015
- [28] Moeskops P, Wolterink JM, van der Velden BH, Gilhuijs KG, Leiner T, Viergever MA, Išgum I: “Deep learning for multi-task medical image segmentation in multiple modalities”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2016, pp 478–486
- [29] Roth HR, Lu L, Seff A, Cherry KM, Hoffman J, Wang S, Liu J, Turkbey E, Summers RM: “A new 2.5 D representation for lymph node detection using random sets of deep



- convolutional neural network observations”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2014, pp 520–527
- [30] Prasoon A, Petersen K, Igel C, Lauze F, Dam E, Nielsen M:” Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2013, pp 246–253
- [31] Simonyan K, Zisserman A (2014) “Very deep convolutional networks for large-scale image recognition”[online] available at: *arXiv:1409.1556*.
- [32] Kamnitsas K, Ledig C, Newcombe VF, Simpson JP, Kane AD, Menon DK, Rueckert D, Glocker B: “Efficient multi-scale 3DCNN with fully connected CRF for accurate brain lesion segmentation”. *Med Image Anal* 36:61–78, 2017.
- [33] Urban G, Bendszus M, Hamprecht F, Kleesiek J (2014) “Multi-modal brain tumor segmentation using deep convolutional neural networks. MICCAI braTS (Brain Tumor Segmentation) Challenge”. *Proceedings, winning contribution*
- [34] Kleesiek J, Urban G, Hubert A, Schwarz D, Maier-Hein K, Bend- szus M, Biller A: “DeepMRI brain extraction: a 3Dconvolutional neu- ral network for skull stripping”. *NeuroImage* 129:460–469, 2016.
- [35] Long, J., Shelhamer, E., and Darrell, T. (2015). “Fully convolutional networks for semantic segmentation”. In *Conference on Computer Vision and Pattern Recognition*. 3431–3440.
- [36] Nie D, Wang L, Gao Y, Sken D: “Fully convolutional networks for multi-modality isointense infant brain image segmentation”. In: *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2016, pp 1342–1345.
- [37] Zhang W, Li R, Deng H, Wang L, Lin W, Ji S, Shen D: “Deep convolutional neural networks for multi-modality isointense infant brain image segmentation”. *NeuroImage* 108:214–224, 2015.
- [38] Ronneberger, F. P., Olaf and Brox, T. (2015). U-Net: “Convolutional networks for biomedical image segmentation”. In *Medical Image Computing and Computer Assisted Intervention*, Springer, 234–241.
- [39] Christ PF, Elshaer MEA, Ettliger F, Tatavarty S, Bickel M, Bilic P, Rempfler M, Armbruster M, Hofmann F, D’Anastasi M, et al: “Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2016, pp 415–423.
- [40] C,ic,ek O,“ Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O: “3F U-Net: learning dense volumetric segmentation from sparse annotation”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2016, pp 424–432.
- [41] Gordienko Y, Gang P, Hui J, Zeng W, Kochura Y, Alienin O, Rokovyi O, Stirenko S: “Deep learning with lung segmentation and bone shadow exclusion techniques for chest X-ray analysis of lung cancer”. In: *International Conference on Theory and Applications of Fuzzy Systems and Soft Computing*. Springer, 2018, pp 638– 647
- [42] Zeng G, Yang X, Li J, Yu L, Heng PA, Zheng G: “3D U-net with multi-level deep supervision: fully automatic segmentation of proximal femur in 3D MR images”. In: *International Workshop on Machine Learning in Medical Imaging*. Springer, 2017, pp 274– 282.
- [43] Kleesiek J, Urban G, Hubert A, Schwarz D, Maier-Hein K, Bend- szus M, Biller A: “DeepMRI brain extraction: a 3Dconvolutional neu- ral network for skull stripping.” *NeuroImage* 129:460–469, 2016
- [44] Milletari F, Navab N, Ahmadi SA: “V-net: fully convolutional neural networks for volumetric medical image segmentation”. In: *2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, 2016, pp 565–571.
- [45] Poudel, R. P. K., Lamata, P., and Montana, G. (2016). “Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation”. In *1st International Workshops on Reconstruction and Analysis of Moving Body Organs, RAMBO 2016 and 1st International Workshops on Whole-Heart and Great Vessel Segmentation from 3D Cardiovascular MRI in Congenital Heart Disease, HVSMR 2016*. 83–94.
- [46] Hochreiter, S. and Schmidhuber, J. (1997). “Long short-term memory”. *Neural computation* 9, 1735–1780. doi:10.1162/neco.1997.9.8.1735.
- [47] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., et al. (2014).” Learning phrase representations using RNN Encoder-Decoder for statistical machine

- translation.” In *Conference on Empirical Methods in Natural Language Processing (ACL)*, 1724–1734.
- [48] Srivastava N, Mansimov E, Salakhudinov R: “Unsupervised learning of video representations using lstms”. In: *International Conference on Machine Learning*, 2015, pp 843–852.
- [49] Xingjian S, Chen Z, Wang H, Yeung DY, Wong WK, Woo WC: “Convolutional LSTM network: a machine learning approach for precipitation nowcasting”. In: *Advances in Neural Information Processing Systems*, 2015, pp 802–810.
- [50] Cheng D, Liu M: “Combining convolutional and recurrent neural networks for Alzheimer’s disease diagnosis using pet images”. In: *2017 IEEE International Conference on Imaging Systems And Techniques (IST)*. IEEE, 2017, pp 1–5.
- [51] Lee CY, Xie S, Gallagher P, Zhang Z, Tu Z: “Deeply-supervised nets”. In: *Artificial Intelligence and Statistics*, 2015, pp 562–570.
- [52] Dou Q, Chen H, Jin Y, Yu L, Qin J, Heng PA: “3D deeply supervised network for automatic liver segmentation from CT volumes.” In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2016, pp 149–157.
- [53] Hwang S, Kim HE: “Self-transfer learning for weakly supervised lesion localization.” In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2016, pp 239–246
- [54] Anirudh R, Thiagarajan JJ, Bremer T, Kim H: “Lung nodule detection using 3D convolutional neural networks trained on weakly labeled data.” In: *Medical Imaging 2016: Computer-Aided Diagnosis, vol 9785, 2016, p 978532. International Society for Optics and Photonics*.
- [55] Shie CK, Chuang CH, Chou CN, Wu MH, Chang EY: “Transfer representation learning for medical image analysis.” In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp 711–714.
- [56] Hoo-Chang S, Roth HR, Gao M, Lu L, Xu Z, Nogue I, Yao J, Mollura D, Summers RM: “Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning.” *IEEE Trans Med Imaging* 35(5):1285, 2016.
- [57] Yosinski J, Clune J, Bengio Y, Lipson H: “How transferable are features in deep neural networks?.” In: *Advances in Neural Information Processing Systems*, 2014, pp 3320–3328.
- [58] Milletari, F., Navab, N., and Ahmadi, S. (2016). “V-Net: Fully convolutional neural networks for volumetric medical image segmentation.” In *2016 Fourth International Conference on 3D Vision (3DV)*. 565–571. doi:10.1109/3DV.2016.79.
- [59] Carneiro, G., Nascimento, J., and Freitas, A. (2010). “Robust left ventricle segmentation from ultrasound data using deep neural networks and efficient search methods.” In *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. 1085–1088 Carneiro,
- [60] Carneiro, G., Nascimento, J. C., and Freitas, A. (2012). “The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods.” *IEEE Transactions on Image Processing* 21, 968–982. doi:10.1109/TIP.2011.2169273
- [61] Carneiro, G. and Nascimento, J. C. (2010). Multiple dynamic models for tracking the left ventricle of the heart from ultrasound data using particle filters and deep learning architectures. In *Conference on Computer Vision and Pattern Recognition (IEEE)*, 2815–2822.
- [62] Carneiro, G. and Nascimento, J. C. (2013). “Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data.” *IEEE transactions on pattern analysis and machine intelligence* 35, 2592–2607.
- [63] Nascimento, J. C. and Carneiro, G. (2014). “Non-rigid segmentation using sparse low dimensional manifolds and deep belief networks”. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (cv-foundation.org)*, 288–295.
- [64] Nascimento, J. C. and Carneiro, G. (2019). One shot segmentation: “unifying rigid detection and non-rigid segmentation using elastic regularization.” *IEEE Trans. Pattern Anal. Mach. Intell.* doi:10.1109/TPAMI.2019.2922959.
- [65] Veni, G., Moradi, M., Bulu, H., Narayan, G., and Syeda-Mahmood, T. (2018). “Echocardiography segmentation based on a shape-guided deformable model driven by a fully convolutional network prior.” In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (IEEE)*, 898–902.
- [66] Jafari, M. H., Girgis, H., Liao, Z., Behnami, D., Abdi, A., Vaseli, H., et al. (2018). “A unified

- framework integrating recurrent Fully-Convolutional networks and optical flow for segmentation of the left ventricle in echocardiography data.” In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (Springer International Publishing)*, 29–37.
- [67] Carneiro, G. and Nascimento, J. C. (2011). “Incremental on-line semi-supervised learning for segmenting the left ventricle of the heart from ultrasound data.” In *2011 International Conference on Computer Vision (IEEE)*, 1700–1707.
- [68] Carneiro, G. and Nascimento, J. C. (2012). “The use of on-line co-training to reduce the training set size in pattern recognition methods: Application to left ventricle segmentation in ultrasound.” In *Conference on Computer Vision and Pattern Recognition (IEEE)*, 948–955 Carneiro.
- [69] Smistad, E., Ostvik, A., Haugen, B. O., and Lovstakken, L. (2017). “2D left ventricle segmentation using deep learning.” In *2017 IEEE International Ultrasonics Symposium (IUS) (IEEE)*, 1–4.
- [70] Yu, L., Guo, Y., Wang, Y., Yu, J., and Chen, P. (2017b). “Segmentation of fetal left ventricle in echocardiographic sequences based on dynamic convolutional neural networks.” *IEEE Trans. Biomed. Eng.* 64, 1886–1895.
- [71] Jafari, M. H., Girgis, H., Abdi, A. H., Liao, Z., Pesteie, M., Rohling, R., et al. (2019). “Semi-Supervised learning for cardiac left ventricle segmentation using conditional deep generative models as prior.” In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) (IEEE)*, 649–652.
- [72] Kass, M., Witkin, A., and Terzopoulos, D. (1988). “Snakes: Active contour models.” *Int. J. Comput. Vis.* 1, 321–331.
- [73] Dong, S., Luo, G., Wang, K., Cao, S., Li, Q., and Zhang, H. (2018a). “A combined fully convolutional networks and deformable model for automatic left ventricle segmentation based on 3D echocardiography.” *Biomed Res. Int.* 2018, 5682365. doi:10.1155/2018/5682365
- [74] Zhang, J., Gajjala, S., Agrawal, P., Tison, G. H., Hallock, L. A., Beussink-Nelson, L., et al. (2018a). “Fully automated echocardiogram interpretation in clinical practice: feasibility and diagnostic accuracy.” *Circulation* 138, 1623–1635.
- [75] Smistad, E., Ostvik, A., Haugen, B. O., and Lovstakken, L. (2017). “2D left ventricle segmentation using deep learning.” In *2017 IEEE International Ultrasonics Symposium (IUS) (IEEE)*, 1–4 Smistad,.
- [76] Leclerc, S., Smistad, E., Pedrosa, J., Ostvik, A., Cervenansky, F., Espinosa, F., et al. (2019). “Deep learning for segmentation using an open Large-Scale dataset in 2D echocardiography.” *IEEE Trans. Med. Imaging* <https://www.creatis.insalyon.fr/Challenge/camus>.
- [77] Ghesu, F. C., Krubasik, E., Georgescu, B., Singh, V., Yefeng Zheng, Hornegger, J., et al. (2016). “Marginal space deep learning: Efficient architecture for volumetric image parsing.” *IEEE Transactions on Medical Imaging* 35, 1217–1228
- [78] Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., and Comaniciu, D. (2008). “Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features.” *IEEE Transactions on Medical Imaging* 27, 1668–1681.
- [79] Golan R, Jacob C, Denzinger J:” Lung nodule detection in CT images using deep convolutional neural networks.” In: *2016 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2016, pp 243–250.
- [80] He, K., Zhang, X., Ren, S., and Sun, J. (2015). “Delving deep into rectifiers: Surpassing Human-Level performance on ImageNet classification.” In *International Conference on Computer Vision (IEEE Computer Society)*, 1026–1034.
- [81] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). “Dropout: A simple way to prevent neural networks from overfitting. Journal of machine learning research” *JMLR* 15, 1929–1958.
- [82] Kamnitsas, K., Bai, W., Ferrante, E., McDonagh, S., and Sinclair, M. (2017a). “Ensembles of multiple models and architectures for robust brain tumour segmentation.” In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers*. 450–462.
- [83] Yu L, Chen H, Dou Q, Qin J, Heng PA: “Automated melanoma recognition in dermoscopy images via very deep residual networks.” *IEEE Trans Med Imaging* 36(4):994–1004, 2017.

- [84]. A. Alansary, O. Oktay, Y. Li, L. Le Folgoc, B. Hou, G. Vaillant, B. Glocker, B. Kainz, and D. Rueckert, “Evaluating reinforcement learning agents for anatomical landmark detection,” *Med. Image Ana.*, vol. 53, pp. 156-164, Apr. 2019.
- [85] T. T. Nguyen, N. D. Nguyen, S. Nahavandi, “Deep reinforcement learning for multi-agent systems: A review of challenges, solutions and applications,” [online] available at: *arXiv preprint arXiv:1812.11794*.