

# Application of Deep Learning in Cardiovascular Medicine using Multiple Biomedical Data

Giovanah Gogi  
PhD student, School of Computing  
University of Portsmouth, UK  
giovanah.gogi@port.ac.uk

Dr Mohamed Bader-El-Den  
Senior Lecturer  
University of Portsmouth, UK  
Mohamed.Bader@port.ac.uk

Dr Alexander Gegov  
Reader in Computational Intelligence  
University of Portsmouth, UK  
alexander.gegov@port.ac.uk

Dr Boriana Vatchova  
Bulgarian Academy of Sciences  
Institute of Information and Communication Technologies  
boriana.vatchova@gmail.com

**Abstract** – Today with the help of emerging technologies scientists can gather data in many sectors which includes the Healthcare sector. The amount of data generated by machines and humans is overwhelming and is growing faster than it has ever had before. Taking advantage of all the power that exist in that data we have come up with a novel idea where we have used the development of machine learning techniques in diagnosing cardiovascular disease. Here, we fit a function to examples and using the function to generalize and make predictions about new diagnosis. In other words, the machine learning models learn from past data to make predictions about a patient diagnosis. This research provides a detailed work with the application of planning techniques is used in order to diagnose CVDs. A combination of biological data and medical imaging has been used. Ultimately, as this is a challenging research area open issues and its possible future works have also been discussed.

**Keywords:** *Cardiovascular Diseases; Deep Learning; Convolutional Neural Network; U-Net*

## I. INTRODUCTION

Cardiovascular Diseases (CVDs) are the number one Cause of death worldwide. It is estimated that 17.9 million people died from CVDs in 2016 [1] and it has been reported that in the US, every 37 seconds, one person dies from CVDs [2].

People are more at risk if they have atrial fibrillation, high blood pressure, high cholesterol, diabetes or have a family history of heart disease. However, analysing medical images can be challenging. It takes a long time to be interpreted by several medical professionals to maintain accuracy in diagnosing. This is therefore very time consuming and can as well as lead to misdiagnosis at times if not enough Cardiologists are present. This project has taken these issues into consideration and has introduced a solution that can be implemented to help the above-mentioned concerns. This is

possible today thanks to the emerging technology that are in place. The aim of this research is to assist cardiologists interpreting the medical images and as well as by using the chosen biological data, make the most effective for a cardiac patient. During the validation stage, the results will be compared to the different algorithms in order to ensure the best accuracy rate of diagnosis.

## II. CHOICE OF ANALYSIS

ML brings the promise of deriving meaning from all the data. Many techniques of AI, particularly ML, have been recommended and is being used to help recognise, classify and predict patterns in data in different sectors.

Some of the current tests to diagnose CVDs are ECG, echocardiogram, stress test, coronary angiography, computerised tomography scan (CT) and cardiac magnetic resonance imaging (MRI) [3,4].

Considering the pros and cons of each of the equipment and the beneficial result that it provides the medical professional with, this research has considered CT and MRI as the chosen attributes.

A key challenge that is facing the Healthcare industry is the quality of service which depends on correctly diagnosing diseases effectively and efficiently. In the cardiac domain, a diagnosis is made based on the Medical professional's which is mainly a Cardiologist's knowledge and point of view through their experiences.

No matter what type of heart disease a patient may be suffering from, the Medical Professional performs a physical exam and ask about the patient's medical and family history followed by a blood test to access the patient's cholesterol level.

Analysing the results from the Biological Data and from the decision of the Cardiologist, they can decide whether a cardiac MRI and/or CT Scan should be done. The system will take in the scanned images as input and give the results as a probability score where the highest score will correspond to the nearest match of the diagnosed heart condition.

Figure 1 shows a prototype of this research project where there are two sections namely: Biological Data and Medical Imaging.

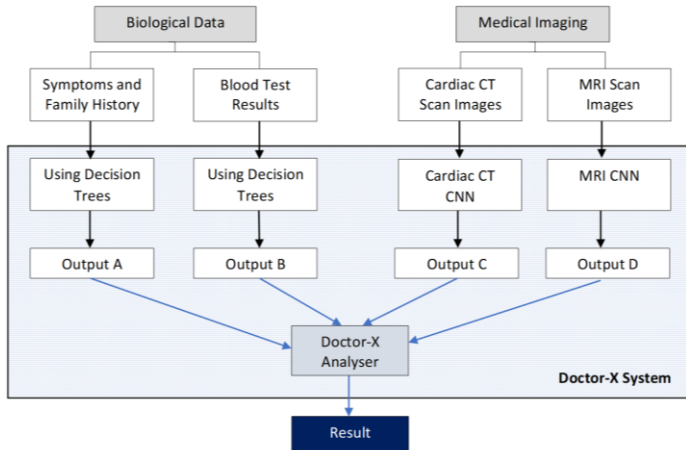


Figure 1: Prototype of the proposed research project

### III. BIOLOGICAL DATA

#### A. Symptoms and Family History

Firstly, the Cardiologist will carry out a risk assessment on the patient asking about their symptoms that they are displaying followed by their family history. For developing this section of the prototype Decision Trees and Naïve Bayes have been taken into consideration.

While some patients may not develop any symptoms from Coronary Artery Disease, therefore, in this case, a Blood Test and Medical scans are vital.

#### B. Blood Test Results

Blood Test (BT) is a parameter that has been considered for this research in view of the advantages that it brings in diagnosing a Cardiac patient. BT provides the measurement amount of certain fats (cholesterol and triglycerides), minerals and sugars in the blood. For risk of a cardiac condition, the cardiologist will check the cholesterol levels as well as the level of Triglycerides which is another type of blood fat [5, 6, 7]

- Decision Tree for classification of the Biological Data

Decision Tree (DT) algorithm is useful and powerful in the healthcare domain for decision making and analysis of prediction. There are several papers that reviewed and discussed about the decision tree algorithm and its benefits for classification problems [8, 9].

DT is one of the popular algorithms in ML. It starts from the root node and splits into different nodes, where the splitting of the root node is done based on probabilities. Each node extract information about feature. The tree is mapped based on two rules, namely: The Gini index and Entropy. Some of the famously known DT algorithms are ID3, C4.5 (extension of ID3), CART and J48 [10].

- Naive Bayes Classifier for classification of the Biological Data.

The Naïve Bayes (NB) classifier is based on Bayes Theorem. The classifier proved to work well where it has outperformed other sophisticated classification algorithms. NB classifier uses the Bayes theorem with independence assumptions between features.

- Experimental Analysis between the two chosen algorithms

To evaluate the performance of the model, we will be using the confusion matrix. True Positive (TP): the patient has CVDs and the test is positive. False Positive (FP): the patient does not have CVDs, but the test is positive. True Negative (TN): the patient does not have CVDs and the test is negative. False Negative (FN): the patient has CVDs, but the test is negative.

*Accuracy* is used to measure the ratio of the correctly classified predictions. *Sensitivity* is used to correctly identify the patients with the disease. *Specificity* is used to correctly identify the patients that do not have the disease that the model got right as well [11].

So, we want something with high sensitivity and high specificity. F-score is used as a combined measure of both sensitivity and specificity.

TABLE I. CONFUSION MATRIX FOR THIS RESEARCH.

	Has CVDs	No CVDs
Has CVDs	TP	FN
No CVDs	FP	TN

From the literature review, it can be concluded that DT and NB classifiers provide high accuracy rate in comparison with other algorithms [12,13,14]

Thus, DT will be used for the development of the Biological Data and as well as for the data aggregation. Nevertheless, DT and NB will be considered for comparison during the validation phase.

#### IV. MEDICAL IMAGING

In the Healthcare area, Deep Learning has been beneficial in analysing medical images. It is excellently being implemented for tasks such as: segmentation, classification and detection of abnormalities, hence making diagnosis successfully correct. Another research in this domain has conducted a Fully CNN for segmenting a short axis MRI for the left ventricle [15].

CNN has proved to be efficient in image analysis such as for classification and segmentation with great accuracy rate[16]. CNN has been considered in developing part B and C of this research.

##### A. Cardiac CT scan Images

A cardiac CT scan is a test that is conducted to help diagnose patients with narrowing or blockage of the arteries around the heart. This scan produces detailed images of the heart and the surrounding blood vessels (coronary arteries).

##### B. MRI Scan Images

Cardiac MRI scan is an advanced and effective scan that complement the CT scan. It is advantageous in detecting several cardiac conditions, namely: Coronary artery disease, congenital heart defects, heart failure, and defects in the heart valves. It also evaluates the anatomy of the heart, its chambers, valves, the blood flow through the vessels and the structure of the pericardium.

- Deployment of CNN

Inside the DL black box, the CNN is not just one process. It is a network of interconnected processes that are organised in layers, where each layer helps the CNN in detecting higher-level and more abstract features. The CNN can identify these features with the use of filters. The Cardiac MRI and CT scanned images are characterised by local structures such as: edges and textures. These filters constitute little versions of each of the little building blocks where during the convolution process, the CNN looks for those building blocks.

The Cardiac MRI and CT scan images are made up of pixels where each pixel is a representation of a tiny portion of the original scanned image. The pixel has a value and an intensity.

The filter that the CNN uses in the first layer are small squares of pixels that correspond to textures and contrast

between the edges. Next, the CNN looks for a combination of the filters. As the complexity rises, eventually the CNN gets closer to identifying whether the patient’s MRI and/or CT image detects a heart condition.

If we combine different types of edges and surroundings, we might get a contouring line that resembles coronary heart disease. As more of the specific features are combined, it becomes more high level at detecting whether there is a blockage, a calcium plaque or a heart valve defect in the scanned image. In the end, after several layers, the network then summarises all the information into a final prediction of whether the image detects a coronary heart disease, an atrial septal defect and so forth.

- CNN architecture for image classification

There are various CNN architectures such as U-Net, GoogleNet, AlexNet, VGGNet, LeNet and ResNet [17]. Figure 2 illustrate an example of a CNN architecture for a scanned medical image.

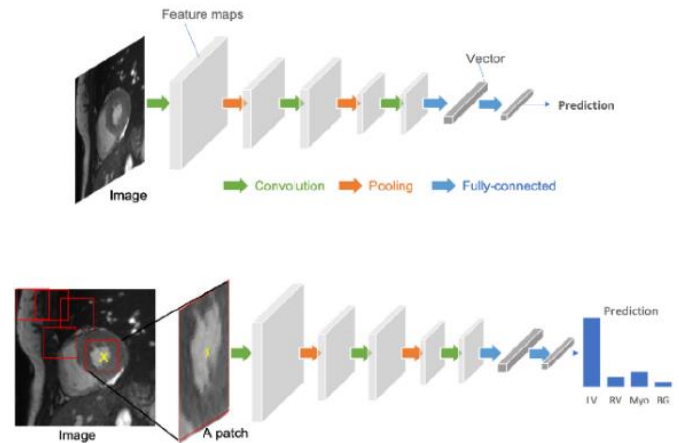


Figure 2: CNN architecture for a cardiac MRI scan [18]

#### V. U-NET ARCHITECTURE FOR THE DEPLOYMENT OF THE CHOSEN MEDICAL IMAGING

A general CNN focuses its task on image classification where an input is an image and the output is a label. However, in this research, it does not only differentiate from the image whether there is a disease but also localise the area of abnormality in the Cardiac MRI and CT scan. Considering the advantages of using U-Net in related works [19], U-Net is being deployed here in both cases of the Medical Imaging to localise and distinguish borders by doing classification on every pixel and hence the input and the output shares the same size. It works with few training images and gives precise segmentation [20].

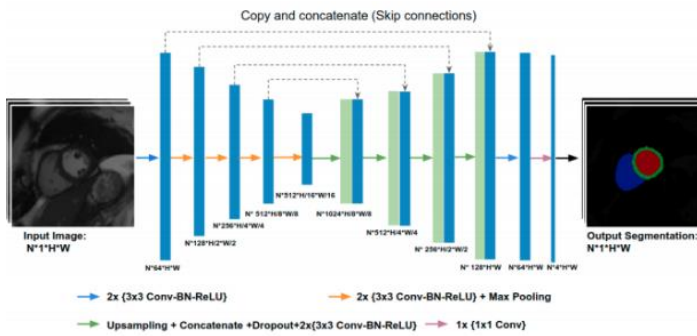


Figure 3: Development of a U-Net architecture [21]

We have an input image of an MRI scan with size 512x512 with 64 feature maps. There is a sequence of two convolutions both with 64 feature maps with padded convolution so that the size is preserved. In this network, there is a convolution of 3x3. The red arrow indicates the max pooling process (2x2) where the 512x512 becomes 256x256 while the number of channels remain the same. Again, a 3x3 convolution is done to produce a size of 128 feature maps. Again, a 3x3 convolution is done to provide 256 and 512 feature maps followed by pooling respectively.

On the decoder side, the operation is duplicated. However, instead of max pooling, up sampling is done. The convolutions are again 3x3 convolutions and the number of feature maps are reduced as we go towards the output. Also, as we go up the decoder side, the feature maps from the encoder side is taken at the same resolution and added to the decoder side.

To finish, an output with a requisite number of classes is produced, where the output represents the reconstructed input. Hence, it will have the same dimension.

## VI. EVALUATION METRICS

The performance of the techniques that has been proposed on the MRI and CT scan images will be evaluated using the Dice Similarity Coefficient which is similarly used in ACDC study [22]. The dice score evaluates the overlap between the manual segmentation from the Radiologists and the automated segmentation. The dice score formula used is:

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

The value of the Dice score should range between 1 and 0. 1 will represent True Negative where there are no cardiac abnormalities and the processing algorithm has correctly detected the defects. 0 will represent False Positive where there are no abnormalities in the heart, but the processing algorithm has incorrectly segmented certain parts as defects.

## VII. CONCLUSION

In this research Biological Data and Medical Images have been combined to achieve effectiveness and better diagnosis in the field of Cardiovascular Diseases. DT and CNN have

been chosen. DT is also considered for data aggregation. Furthermore, U-Net architecture has been designed and being deployed in the Dr X system as shown in the prototype. This research project will use pre-trained CNN models such as GoogleNet, AlexNet, VGGNet and LeNet upon completion using the Sciam Supercomputer from the Institute of Cosmology and Gravitation. We also planned to implement cross-validation for possible results.

## REFERENCES

- [1] World Health Organisation. 2017. Cardiovascular diseases. [ONLINE] Available at: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)). [Accessed 12 January 2020].
- [2] Centers for Disease Control and Prevention. 2019. Heart Disease Facts. [ONLINE] Available at: <https://www.cdc.gov/heartdisease/facts.htm>. [Accessed 12 January 2020].
- [3] MayoClinic. 2018. Heart disease. [ONLINE] Available at: <https://www.mayoclinic.org/diseases-conditions/heart-disease/diagnosis-treatment/drc-20353124>. [Accessed 12 January 2020].
- [4] British Heart Foundation. Tests for heart and circulatory conditions. [ONLINE] Available at: <https://www.bhf.org.uk/informationsupport/tests>. [Accessed 12 January 2020].
- [5] CardioSmart American College of Cardiology. 2012. Cholesterol and Triglycerides Tests. [ONLINE] Available at: <https://www.cardiosmart.org/Healthwise/hw20/7814/hw207814>. [Accessed 13 January 2020].
- [6] Clevelandclinic. 2019. Blood Tests to Determine Risk of Coronary Artery Disease. [ONLINE] Available at: <https://my.clevelandclinic.org/health/diagnostics/16792-blood-tests-to-determine-risk-of-coronary-artery-disease>. [Accessed 13 January 2020].
- [7] mayoclinic. 2019. Blood tests for heart disease. [ONLINE] Available at: <https://www.mayoclinic.org/diseases-conditions/heart-disease/in-depth/heart-disease/art-20049357>. [Accessed 13 January 2020].
- [8] A. Pandey et al., 2013. A Heart Disease Prediction Model using Decision Tree. IOSR Journal of Computer Engineering (IOSR-JCE), [Online]. Volume 12, 83-86. Available at: <https://pdfs.semanticscholar.org/22b5/100b7a51cfbb30559457d385070bfd4b1083.pdf> [Accessed 14 January 2020].
- [9] S. Mirza et al., 2018. Applying Decision Tree for Prognosis of Diabetes Mellitus. International Journal of Applied Research on Information Technology and Computing, [Online]. Volume 9, 15-20. Available at: DOI: 10.5958/0975-8089.2018.00002.7 [Accessed 14 January 2020].

- [10] A. Karthiga, 2017. Early Prediction of Heart Disease Using Decision Tree Algorithm. International Journal of Advanced Research in Basic Engineering Sciences and Technology, [Online]. Volume 3, Issue 3, 1-16. Available at: <https://www.ijarbest.com/Archive>. [Accessed 15 January 2020].
- [11] A. Lalkhen, 2008. Clinical tests: sensitivity and specificity. Continuing Education in Anaesthesia, Critical Care & Pain, [Online]. Volume 8, Issue 6, 221–223. Available at: <https://academic.oup.com/bjaed/article-abstract/8/6/221/406440>. [Accessed 15 January 2020].
- [12] C. Latha, 2019. Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. Informatics in Medicine Unlocked, [Online]. Volume 16, pp 1-9. Available at: <https://doi.org/10.1016/j.imu.2019.100203> [Accessed 15 January 2020].
- [13] R. Patra et al., "Predictive Analysis of Rapid Spread of Heart Disease with Data Mining," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 2019, pp. 1-4, doi: 10.1109/ICECCT.2019.8869194. [Accessed 15 January 2020].
- [14] P.N. and PR. Kumar, "Usage of data mining techniques in predicting the heart diseases — Naïve Bayes & decision tree," 2017 International Conference on Circuit , Power and Computing Technologies (ICCPCT), Kollam, 2017, pp. 1-7, doi: 10.1109/ICCPCT.2017.8074215. [Accessed 15 January 2020].
- [15] P. V. Tran, "A Fully Convolutional Neural Network for cardiac segmentation in short-axis MRI", Available at: <https://arxiv.org/pdf/1604.00494.pdf>. [Accessed 19 January 2020].
- [16] N. Sharma, 2018. An Analysis Of Convolutional Neural Networks For Image Classification. International Conference on Computational Intelligence and Data Science, [Online]. Volume 132, 377-384. Available at: <https://doi.org/10.1016/j.procs.2018.05.198> [Accessed 16 January 2020].
- [17] A. Khan, A. Sohail, U. Zahoora et al., A survey of the recent architectures of deep convolutional neural networks. Artificial Intelligence Review (2020). Available at: <https://doi.org/10.1007/s10462-020-09825-6> [Accessed 11 May 2020].
- [18] C. Chen et al., 2020. Deep learning for cardiac image segmentation: A review. Frontiers in Cardiovascular Medicine, [Online]. Volume 7 Article 25, 1-33. Available at: <https://doi.org/10.3389/fcvm.2020.00025> [Accessed 20 April 2020].
- [19] M. T. Rostami et al., "Deep learning network for fully automatic left ventricle segmentation," 2018 4th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), Tehran, Iran, 2018, pp. 141-144, Available at: doi: 10.1109/ICSPIS.2018.8700528 [Accessed 20 April 2020].
- [20] K. Hu, C. Liu, X. Yu et al., "A 2.5D Cancer Segmentation for MRI Images Based on U-Net," 2018 5th International Conference on Information Science and Control Engineering (ICISCE), Zhengzhou, 2018, pp. 6-10, doi: 10.1109/ICISCE.2018.00011 [Accessed 20 April 2020].
- [21] arXiv. 2019. Improving the generalizability of convolutional neural network-based segmentation on CMR images. [ONLINE] Available at: <https://arxiv.org/pdf/1907.01268.pdf>. [Accessed 28 April 2020].
- [22] O. Bernard et al., "Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved?," in IEEE Transactions on Medical Imaging, vol. 37, no. 11, pp. 2514-2525, Nov. 2018, doi: 10.1109/TMI.2018.2837502. [Accessed 29 April 2020].