

CNN-KCL: Automatic Myocarditis Diagnosis using Convolutional Neural Network Combined with K-means Clustering

Danial Sharifrazi¹, Roohallah Alizadehsani², Javad Hassannataj Joloudari³, S Shamshirband⁴, Sadiq Hussain⁵, Zahra Alizadeh Sani^{6,7}, Fereshteh Hasanzadeh⁸, Afshin Shoaibi⁹, Abdollah Dehzangi¹⁰, Hamid Alinejad-Rokny^{11, 12,13}

¹ Department of Computer Engineering, School of Technical and Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran

² Institute for Intelligent Systems Research and Innovation, Deakin University, Geelong, VIC 3216, Australia

³ Department of Computer Engineering, Faculty of Engineering, University of Birjand, Birjand, Iran

⁴ Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam

⁵ System Administrator, Dibrugarh University, Assam 786004, India

⁶Rajaie Cardiovascular Medical and Research Center, Iran University of Medical Sciences, Tehran, Iran

⁸Omid hospital, Iran University of Medical Sciences, Tehran, Iran

⁹Computer Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran

¹⁰ Department of Computer Science, Morgan State University, Baltimore, MD, 21251, USA.

¹¹ Systems Biology and Health Data Analytics Lab, The Graduate School of Biomedical Engineering, UNSW Sydney, Sydney, NSW, 2052, AU.

¹²School of Computer Science and Engineering, UNSW Sydney, Sydney, NSW, 2052, AU.

¹³Health Data Analytics Program Leader, AI-enabled Processes (AIP) Research Centre, Macquarie University, Sydney, 2109, AU.

Abstract

Myocarditis is the form of an inflammation of the middle layer of the heart wall which is caused by a viral infection and can affect the heart muscle and its electrical system. It has remained as one of the most challenging diagnoses in cardiology. Myocardial is the prime cause of unexpected death in approximately 20% of adults less than 40 years of age. Cardiac MRI (CMR) has been considered as a noninvasive and golden standard diagnostic tool for suspected myocarditis and plays an indispensable role in diagnosing various cardiac diseases. However, the performance of CMR is heavily dependent on the clinical presentation and non-specific features such as chest pain, arrhythmia, and heart failure. Besides, other imaging factors like artifacts, technical errors, pulse sequence, acquisition parameters, contrast agent dose, and more importantly qualitatively visual interpretation can affect the result of the diagnosis. This paper introduces a new deep learning-based model called Convolutional Neural Network-Clustering (CNN-KCL) to diagnose the Myocarditis. The hybrid CNN-KCL method performs the early and accurate diagnosis of Myocarditis. To the best-of-our-knowledge, a Convolutional neural network has never been used before for the diagnosis of Myocarditis. In this study, we used 47 subjects to diagnose myocarditis patients from Tehran's Omid Hospital. The total number of data examined is 10425. Our results demonstrate that CNN-KCL achieves 92.3% in terms of diagnosis myocarditis prediction accuracy which is significantly better than those reported in previous studies.

Keywords: Myocarditis; Diagnosis; Convolutional Neural Network; Cardiac MRI; prediction

1. Introduction

Cardiovascular Diseases (CVD) are among the most important causes of mortality around the world [1, 2]. Atherosclerosis as the leading cause of CVD referring to build-up of different substances including cholesterol and fat in and on the walls of the arteries. The flow of blood is constrained in such a situation which in turn affect the whole body. CVDs due to atherosclerosis include cerebrovascular disease (e.g. stroke), ischemic heart disease (e.g. heart attack), and hypertensive heart disease. Other CVDs embrace rheumatic heart disease, inflammatory heart disease, congenital heart disease, cardiac arrhythmias, and heart failure [2].

Inflammation of the heart muscles is clinically termed as myocarditis [3]. In **Figure 1**, discrimination of normal and myocarditis affected heart can be seen. The inflammation of heart muscles is evident in myocarditis affected heart as shown in **Figure 1**. The symptoms of myocarditis include chest pain or mild dyspnea. Common viral infections such as hepatitis B and C, parvovirus and the recent one, COVID-19 may also cause myocarditis. Other specific forms of myocarditis include sarcoidosis, giant-cell myocarditis, hypersensitivity drug reactions, toxic or pathogens which may occur less commonly [4]. Patients that are diagnosed with myocarditis should be referred to the heart specialist for endomyocardial biopsy. The endomyocardial biopsy enables the clinicians to check the presence of premorbid in the patients of myocardial inflammation. Myocardial is the prime cause of unexpected death approximately 20% in adults less than 40 years old [5]. Although it had been centuries since the recognition of this enigmatic disease, effective treatment strategies are yet to develop because of several issues such as: insensitivity to diagnostic tests, complex relations between maladaptive and adaptive immune responses [4]. A recent development in the genetic basis of immune-mediated heart disease and studies in animals provided key information in treating the disease.

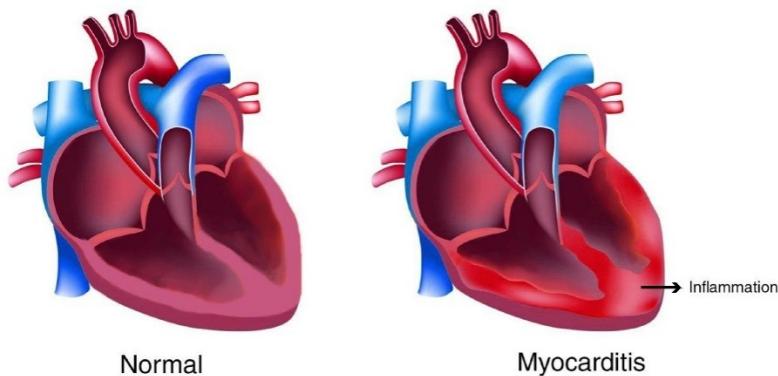


Figure 1: Demonstration of Myocarditis (on the right) and normal heart (on the left). As shown here, the inflammation of the heart muscles is evident in myocarditis affected heart.

Early detection of heart diseases and hazards is crucial for decreasing the mortality rate [6]. It is thorny for the specialist as well to interpret accurately different images such as electrocardiograms, echocardiograms or magnetocardiogram for diagnosis of heart related diseases as it is time consuming, costly, and error-prone. Hence, automatic computer-based diagnosis systems are in high demand to assist the clinicians to detect and diagnose the diseases more accurately, faster, and less costly. Such systems are used in various diseases including heart-related diseases. These systems utilize various machine learning and data mining techniques to effectively resolve issues [6]. The aim is to develop prognostic models that effectively predict (with reasonable specificity and sensitivity) the occurrence and outcome of disease in the near future.

The screening methods provided above help physicians diagnose the disease at an early stage and they can manually analyze the data to know if one has the disease. However, using imaging techniques (for diagnosing the disease) often causes a misdiagnosis. Factors such as fatigue and excessive workload of the physician as well as other factors such as the presence of various noises in the images or the presence of a mass or lesion that is not visible to the eye can cause misdiagnosis [7]. A wrong diagnosis in the early stages, can be seriously challenging for a person's health leading to possible death. To address this problem,

computerized aided diagnosis (CADs) can help clinicians in diagnosing the diseases in their early stages [8].

Machine learning methodologies are deployed to predict and analysis different disease-related issues [9]. Analyzing large numbers of images collected from patients and using non-deep learning algorithms is difficult because when the data becomes too large the algorithm becomes complicated due to the number of segmentation algorithms, required feature extraction, and classification process [10]. In addition, non-deep learning methods can partially receive the input data, and if the data exceeds a certain limit, they become unable to adequately process it. However, the deep learning-based methods can solve these complex problems. In fact, for deep learning-based models, the more the input data increases, the more they become accurate and efficient. Deep learning plays a crucial role in successfully analyzing heart-related images [11]. The performance of the classification methods is enhanced by utilizing deep learning techniques.

To the best-of-our-knowledge, there is no scientific literature related to deep learning-based studies on the detection and diagnosis of myocarditis disease. This is the first step towards such a study. We, therefore, systematically reviewed the literature based on Cardiomyopathy oriented diseases. The rest of the paper is organized as follows: Section 2 describes literature review; Section 3 presents Z-Alizadeh Sani myocarditis dataset; Section 4 represents method while section 5 describes experimental results; Section 6 presents the discussion which is follows by section 7 that presents conclusion and future work.

2. Literature Review

Baeßler *et al.* [12] examined whether machine learning-based techniques might be utilized for the recognition of myocardial tissue alterations in hypertrophic cardiomyopathy (HCM) on T1- weighted non-contrast cardiac magnetic resonance (CMR) images using texture analysis (TA). Texture feature selection and step-wise dimension reduction were exploited for feature selection and identification of myocardial tissue alterations on non-contrast T1-weighted CMR images in HCM patients.

In another study, Ovreiu and Simon [13] investigated the diagnosis of cardiomyopathy in its two common forms: hypertrophic and dilated via *P* wave features. They applied a novel evolutionary technique dubbed as biogeography-based optimization (BBO) and developed a neuro-fuzzy network. They demonstrated that cardiomyopathy could be successfully diagnosed by applying a neuro-fuzzy model. Later Ali *et al.* presented a computerized framework for detection of cardiomyopathy diseases utilizing a multilayered perceptron (MLP) neural network [14]. In this study, the high-frequency noise removal method was employed using moving and median average filters at the preprocessing stage.

More recently Alis *et al.* [15] exploited a machine learning approach for texture feature analysis of cardiac magnetic resonance imaging (MRI) for examining the incidence of ventricular tachyarrhythmia (VT) in hypertrophic cardiomyopathy patients. Similarly, Borkar *et al.* [16] proposed a machine learning approach for the automatic detection of Atrial Septal Defect (ASD) and dilated cardiomyopathy (DCM) diseases. Their dataset comprised of the ultrasound videos of DCM, ASD, and normal cases. In another study, Sengupta *et al.* [17] designed a machine learning algorithm based on an associative memory classifier using echocardiographic and clinical records of 44 patients with restrictive cardiomyopathy and 50 with constrictive pericarditis. To discriminate constrictive pericarditis from restrictive cardiomyopathy, they

normalized the speckle tracking echocardiography images with respect to 47 controls with no structural heart disease and evaluated the diagnostic area under the ROC curve of the associative memory classifier.

At the same time, Begum *et al.* [18] developed an automated diagnostic system for cardiomyopathy disease using feed-forward backpropagation neural network and SVM classifiers. They utilized online PTB diagnostic ECG database and preprocessed it for baseline correction and noise cancellation. They then proposed four time-based features and classified them using artificial neural networks and SVM [19]. More recently, Green *et al.* [20] studied the echocardiograms and photo plethysmography data from a control cohort of 64 healthy volunteers and 19 HCM patients with left ventricular outflow tract obstruction (oHCM). In another study, Tsai and Kojima[21] utilized four texture features of ultrasonic images for heart disease classification. Their proposed method took as input the heart images that measured texture features by generating a gray-level co-occurrence matrix.

More recently, Narula *et al.* [22] examined the potential for a machine learning system that integrated speckle-tracking echocardiographic recordings for discrimination of HCM from physiological hypertrophy seen in athletes (ATH). Similarly, Rahman *et al.* [23] proposed an HCM patient classifier utilizing standard 12-lead and 10-seconds ECG signals. They derived 504 temporal and morphological features including both newly-developed and commonly used ones from ECG signals for heartbeat classification.

Recently, Shao *et al.* [24] examined whether TA parameters on magnetic resonance T1 mapping could be applied for the diagnosis of DCM. In this study, Modified Look-Locker Inversion Recovery (MOLLI) sequence at a 3.0 T MR scanner was used to acquire T1 maps. The epicardium and endocardium were strained on the short-axis slices of the T1 maps by a skilled radiologist. Most recently, Capture *et al.* [25] performed plasma proteomics and exploratory myocardial screens and consequently devised a multiplexed targeted liquid chromatography-tandem/mass spectrometry-based assay to examine 26 peptide biomarkers to recognize novel plasma biomarkers for patients with HCM.

Based on the related works, there is no deep learning based method for diagnosing myocarditis. In the proposed CNN-KCL method, a combination of k-means clustering and CNN classifier was used. Because of the large number of image sections, we first used clustering method. In k-means algorithm, due to the reduction of low slope K, the elbow method cannot be used. For this reason, we increased the value of K from 2 clusters to 24 clusters to determine the effect of increasing the number of clusters on the accuracy of CNN classification.

3. Z-Alizadeh Sani myocarditis dataset

Cardiac MR Imaging is considered as the noninvasive and diagnostic golden standard of myocarditis in the absence of biopsy. CMR provides the possibility of anatomical and functional imaging and accurate assessment of heart. However, in this respect, its ability of tissue characterization is even more important[26]. Three diagnostic targets for the three recommended Cardiac MRI criteria are myocardial edema, hyperemia/capillary leak, and scar which is known as Lake Louise Criteria (LLC) [27]. Existence of contrast enhancement (CE-GD) affirms myocardial injury (i.e. scar, fibrosis) while T2-weighted images show interstitial edema, known as an integral part of the inflammatory response. The pre and post-contrast T1-Weighted image indicates presents of hyperemia/capillary leak in myocardial tissue. The LLC has been

accepted beyond clinical criteria and endomyocardial biopsy. Two out of three “Lake Louise Criteria” have 80.0% accuracy for a correct diagnosis of acute myocarditis [28]. In this study, we developed a model to diagnose myocarditis based on three indexes of Lake Louise Criteria (LLC).

CMR imaging protocols

This study was prospectively performed from September 2018 to September 2019 at the CMR department of OMID hospital at Tehran, IRAN. The study was approved by the local ethical committee of the OMID hospital. CMR examination was performed using a 1.5-T system (MAGNETOM Aera Siemens, Erlangen Germany) [29]. All patients were scanned with dedicated body coils in the standard supine position. The CMR protocols are as follows:

We performed CINE-segmented images and pre-contrast T2-Weighted (trim) images in short and long-axis views. Pre contrast T1-Weighted relative images were acquired in axial views of the myocardium. Immediately after Gadolinium injection (DOTAREM 0/1 mmol/kg) the T1-Weighted relative sequence was repeated and after 10-15 minutes, Late Gadolinium Enhancements (LGE- high-resolution PSIR) sequences in short and long-axis views were performed. **Table 1** shows the parameters and details of CMR sequences:

Table 1.Z-Alizadeh Sani myocarditis dataset description*

Protocols & Parameters	TE (ms)	TR (ms)	Segment	slice thickness (mm)	Concatenation and Slice number	NEX	Breath-hold time (s)
CINE_segmented (true FISP) Long Axis (LAX):	1.15	33.60	15	7	3	1	8
CINE_segmented (true FISP) Short Axis (SAX):	1.11	31.92	15	7	15	1	8
T2-Weighted (TIRM) LAX, pre-contrast	52	800	Non-cine	10	3	1	9
T2-Weighted (TIRM) SAX, pre-contrast	52	800	Non-cine	10	5	1	10
T1 Relative-Weighted TSE (Trigger)- AXIAL-dark blood pre and post-contrast	24	525	Non-cine	8	5	1	7
Late-GD Enhancement LGE(high-resolution PSIR) SAX and LAX	3.16	666	Non-cine	8	1	1	7

*Sequences'/protocols' name and parameters description Used by CMR in Patients of this study. **TE**: Time Echo, **TR**: Time Repetition, **Segment**: numbers of frames (Segmented acquisition to produce a series of images that can be displayed as a movie of cardiac function (cine)), **Slice Thickness**: how thick the slices are, **Concatenation**: distribution of the slices to be measured over multiple TR, **NEX**: Number of Excitations (How many times each line of k-space data is acquired during the scan), **Breath-hold time**: Duration of time that patient should hold his/her breath in order to avoiding chest motion artifacts.

4. Method

4.1 Convolutional Neural Network (CNN)

A CNN is a deep learning technique that takes an image as input, allocates importance (biases and learnable weights) to different objects/aspects in the image, and then classify them [30]. CNN requires much lower pre-processing compared to the other deep learning architectures. Generally, filters are hand-engineered in primitive methods while CNN has the ability to learn these characteristics/filters. The connectivity pattern of neurons in the Human Brain and the architecture of CNN are analogous. In fact, CNN is developed based on the organization of the Visual Cortex. A CNN is comprised of an input layer, multiple hidden layers, and an output layer. The hidden layers of CNN are typically created with convolutional layers that convolve with others [31]. The activation function is usually a Rectified Linear Units (ReLU) layer and additional convolutions such as normalization layers, fully connected layers, and pooling layers are consequently followed [32]. There are various architectures of CNNs that are extensively used in the deep learning domain. Some of such architectures are LeNet [33], AlexNet [34], VGGNet [35], GoogleNet [36], and ResNet [37]. CNN is widely used in many different areas of research including computer vision, medical image analysis, recommendation system, financial time series, natural language processing, image, and video recognition and classification [38]. CNN has been applied successfully in detection and diagnosis of various diseases using various image modalities [31, 39-43].

4.2 K-means clustering

K-means clustering is one of the popular and simplest unsupervised technique used in machine learning domain [44]. It is based on vector quantization that aims to partition n observations into k clusters. Each observation belongs to a cluster with nearest mean value (cluster centroid or cluster centers). The Euclidean distance (or any other distance metrics) is usually utilized to compute k groups of similarity [45]. Each item is categorized to its closest mean and the mean's coordinates are updated as the averages of the items categorized in that cluster so far.

4.3 Elbow method

A basic step in any clustering method is to estimate the optimal number of clusters in which data needs to be clustered to. To estimate the optimal value of k (number of clusters), the elbow method [45] is considered as one of the popular techniques. In this method, the mean of the squared distances from the cluster centers of the relevant clusters is termed as distortion. The sum of squared distances of samples to their closest cluster center is dubbed as inertia. To find the optimal number of the clusters, the values of k is iterated (from 1 to n) and the values of distortions for each value of k is computed to measure the inertia and distortion. The optimal number of clusters is then computed by selecting the value of k at the “elbow” i.e. the point after which the inertia/distortion begins to decrease linearly [46].

4.4 Proposed CNN-KCL scheme

The block diagram of the proposed clustering method namely Convolutional Neural Network-Clustering (CNN-KCL) is shown in **Figure 2**, which includes the steps of data entry, clustering, classification, and final prediction.

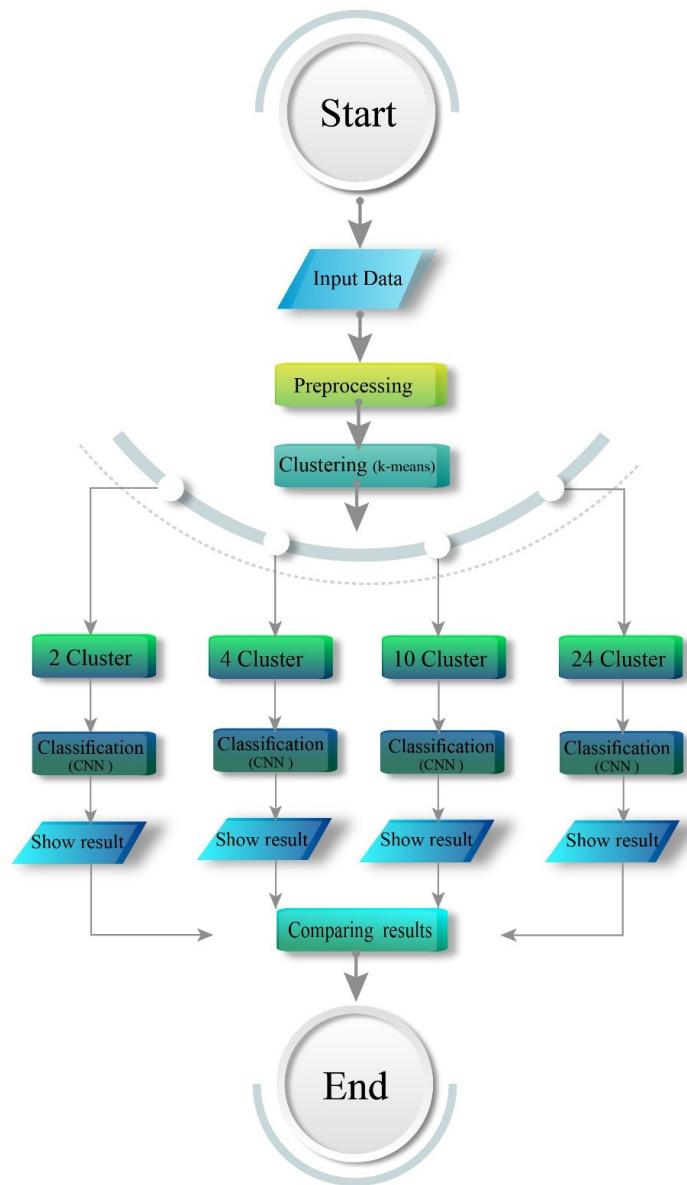


Figure 2. The block diagram of the proposed CNN-KCL method

According to Figure 2, the implementation steps are shown briefly:

Step 1: In the first step, the proposed dataset is entered into the system.

Step 2: The second step is related to the initial pre-processing. All dataset images are on a Gray level. So there is no need to delete the image channels. Dataset images are available in a variety of sizes. For this purpose, all images are resized to a size of 100 x 100. After resizing the images in the data pre-processing step, the image normalization operation is performed. Normalization of images transforms the light intensity of all image pixels to the range of zero and one (0 and 1).

Step 3: Due to the great variety of image sections, clustering is used in this step. There are different methods for clustering data that in this research, the k-means method is used to cluster images. It is important to note

that the k-means algorithm is not capable of clustering two-dimensional images. For this reason, all images are reshaped into the vector of pixels, and then these are used in the k-means algorithm.

Step 4: In this step, the data is divided into 2, 4, 10, and 24 clusters by the k-means algorithm.

Step 5: In this step, the data of each cluster are classified using CNN separately. Convolution layers are for feature extraction and fully connected layers act as classifiers in this network.

Step 6: In step 6, the results of the classification of 2, 4, 10, and 24 clusters are displayed separately.

Step 7: Finally, the results of the classification of clustered images performed by k-means and CNN algorithms are compared.

We explain the above steps in more detail. When we used different images as input to CNN-KCL, it did not have an accurate detection of incoming images. It was mainly due to the fact that our employed dataset contains images with a very large difference, meaning that images were taken from different angles such as side, front, bottom, and top. Hence, it was decided to use several clusters to incorporate the accuracy in different modes with better results. In order to cluster the data, the K-means method has been used where K is the number of clusters. We also used the Elbow method to find the best K .

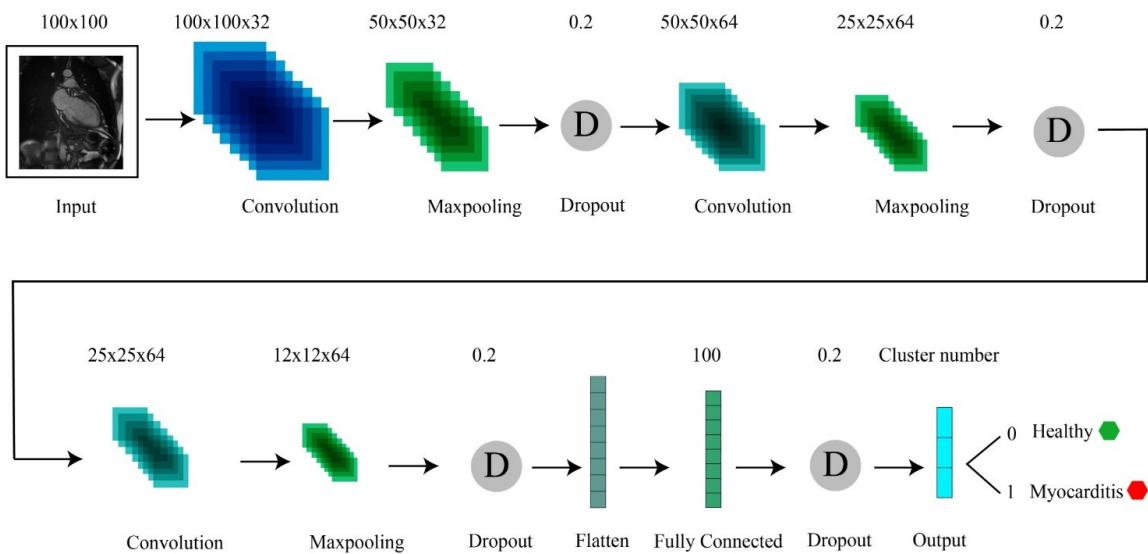
In this study, we implemented the Elbow method from one cluster to 30 clusters ($K = 1$ to 30). The resulting graph was declining very slightly, so-called Smooth, and we could not identify the Elbow inside (that is, the point with a very steep slope, and once the chart breaks and becomes horizontal). Hence, with the Elbow method, we could not find the best number of clusters. For this reason, the K-means method was used with $K=4$, $k=10$, and $k=24$ to determine the trend of accuracy changes with increasing and decreasing K . In fact, we increase the number of clusters from 1 to 30 clusters and examine the results. First, we have 2 clusters ($K = 2$) that have 1 class for healthy people and 1 class for sick people. To divide the number of clusters even further, we divide healthy people into 2 clusters and divide sick people into 2 clusters (a total of 4 clusters ($K = 4$)). In this way, input images should be classified by CNN-KCL to 4 classes. We also increased the number of clusters to 10, so that 5 clusters were created for sick people and 5 clusters for healthy people. As a result, a total of 10 clusters or 10 classes ($K = 10$) were created to identify healthy and sick people at CNN-KCL output. Finally, the number of clusters increased to 24 ($K = 24$), which we divided into 12 clusters for healthy people and 12 clusters for sick people. An input image is classified as healthy if it is categorized in one of the healthy clusters. All modes of 2 clusters, 4 clusters, 10 clusters, and 24 clusters in conditions 30-70 (70% data for training and 30% data for a test) and 20-80 (80% of the data for training and 20% of the data for the test), have been examined.

4.6 Recommended CNN-KCL architecture

The proposed CNN-KCL architecture is shown in **Figure 3**. CNN-KCL includes input layers, convolution (32 kernels), Maxpooling (2×2), Dropout, re-convolution layers (64 kernels), Maxpooling (2×2), and Dropout. It also has the Flatten layer, fully connected, and the dropout layer, and finally the desired output. CNN-KCL is described in detail in **Table 2**.

Table 2. Details explanation of CNN-KCL

Number of Kernel from first to third connection	Size of the convolution kernels	Size of the max pooling kernels	Number of neurons in the Fully Connected layer	Number of neurons in the output layer	Size of the Dropout	Number of batch size	Number of epochs	Value of validation data	Optimizer function	Activator function	Output layer classifier in multi-class mode	Output layer classifier for the pre-cluster mode
32,64, and 64	3*3	2*2	100	4, 10 and 24	0.2	32	70	0.3 and 0.2	Adam	ReLU	SoftMax	sigmoid

**Figure 3.** The geneal archutecture of CNN-KCL

The proposed CNN architecture is presented in Figure 3. Images with a size of $100 * 100$ enter the CNN network. As the first part of the network, a convolution layer with a $3 * 3$ kernel size, a Maxpooling layer with a $2 * 2$ kernel size, and a dropout layer with a size of 0.2 have been selected. In this network, the convolution layer is used to extract the feature from the dataset. The dropout layer is used to randomly remove a number of neurons to avoid overfitting in the network. The combination of these three layers with the activation function on CNN is usually referred to as one layer. The proposed CNN network consists of three layers with the mentioned features. The numbers of filters in the first, second, and third convolution layers are 32, 64, and 64, respectively. After passing the data through these three layers, the Flatten layer is used to change the form of the data from two-dimensional to one-dimensional mode. Then, the Fully Connected layer is used to classify the network. After the Fully Connected layer, the Dropout layer is used again to avoid the network overfit. As the output layer, there is a vector with the number of neurons equal to the number of clusters, i.e., 2, 4, 10, and 24.

5. Experimental results

In this section, the proposed algorithm is applied on Z-Alizadeh Sani myocarditis dataset described in before. We also used Python programming language and the Keras library used with the Tensorflow to implement CNN-KCL. The results of clustering and classification are presented in this section.

In the first step, CNN-KCL was applied for the classification purpose. Initially, the Z-Alizadeh Sani myocarditis dataset was given to the CNN-KCL model without clustering. In fact, just a cluster for healthy persons and a cluster for sick persons considered. Due to a great variety in the employed dataset and existence of images in different parts of the heart, the initial accuracy of CNN-KCL was very low. For this reason, the data was divided into more clusters to represents the data with more distinctive pattern.

In order to cluster the data, the K-means method was used where K is the number of clusters. The Elbow method was used to find the best K. The value of K was examined for the Elbow method from 1 to 30. **Figures 4** shows the Elbow method's result to identify the optimal K. As shown in this figure, since we did not observe a sharp slope, the exact amount of the best K was not determined. For this reason, the K-means method was used with different values of K (4, 10, and 24) to determine the process of accuracy changes considering different values of K. In the first steps, K = 4 was placed (that means 2 clusters for healthy persons and 2 clusters for sick persons). Previously, 1 cluster for healthy persons and 1 cluster for sick persons were examined.

For the case where K = 4, two clusters were obtained for healthy individuals and 2 clusters for individuals. As the value of K increased to 10 and 24, the accuracy of CNN-KCL decreased. In each of the K states (2 (before clustering), 4, 10, and 24) CNN-KCL networks were executed, and the results were obtained. The results achieved using different number of clusters are shown in **Figures 5 to 10**, respectively. The figures represent Accuracy [47], LOSS [48], and ROC [49]. Also, the figures are created for 2 clusters, 4 clusters, and 24 cluster modes.

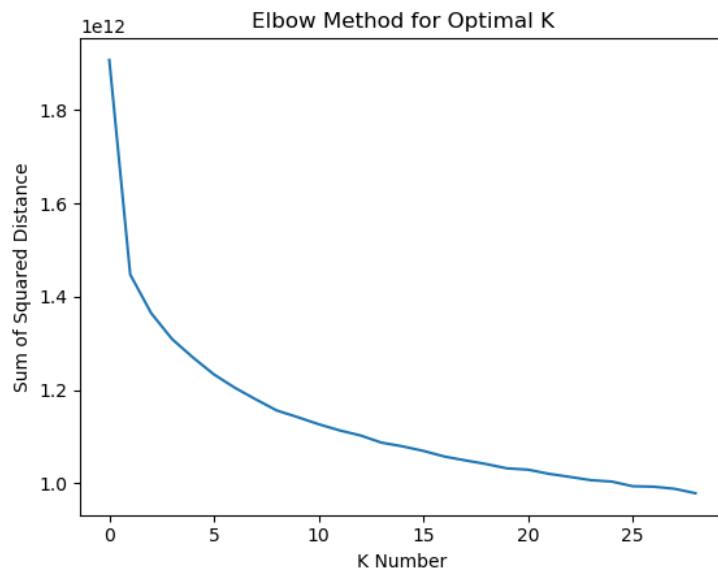


Figure 4. Elbow method for optional K

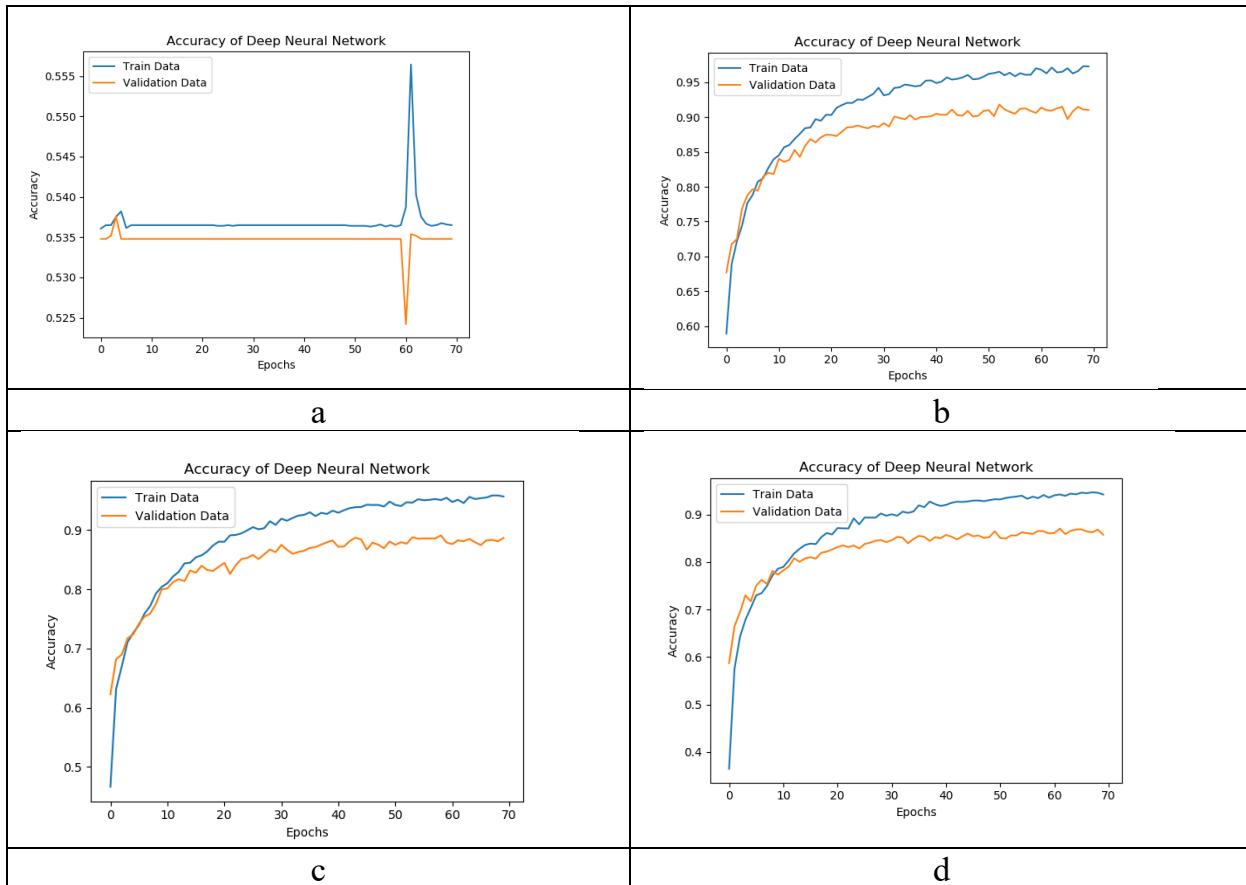


Figure 5. Accuracy of CNN-KCL for different number of clusters for 70/30 training/ testing. **a)** before clustering. **b)** 4 clusters. **c)** 10 clusters **d)** 24 clusters.

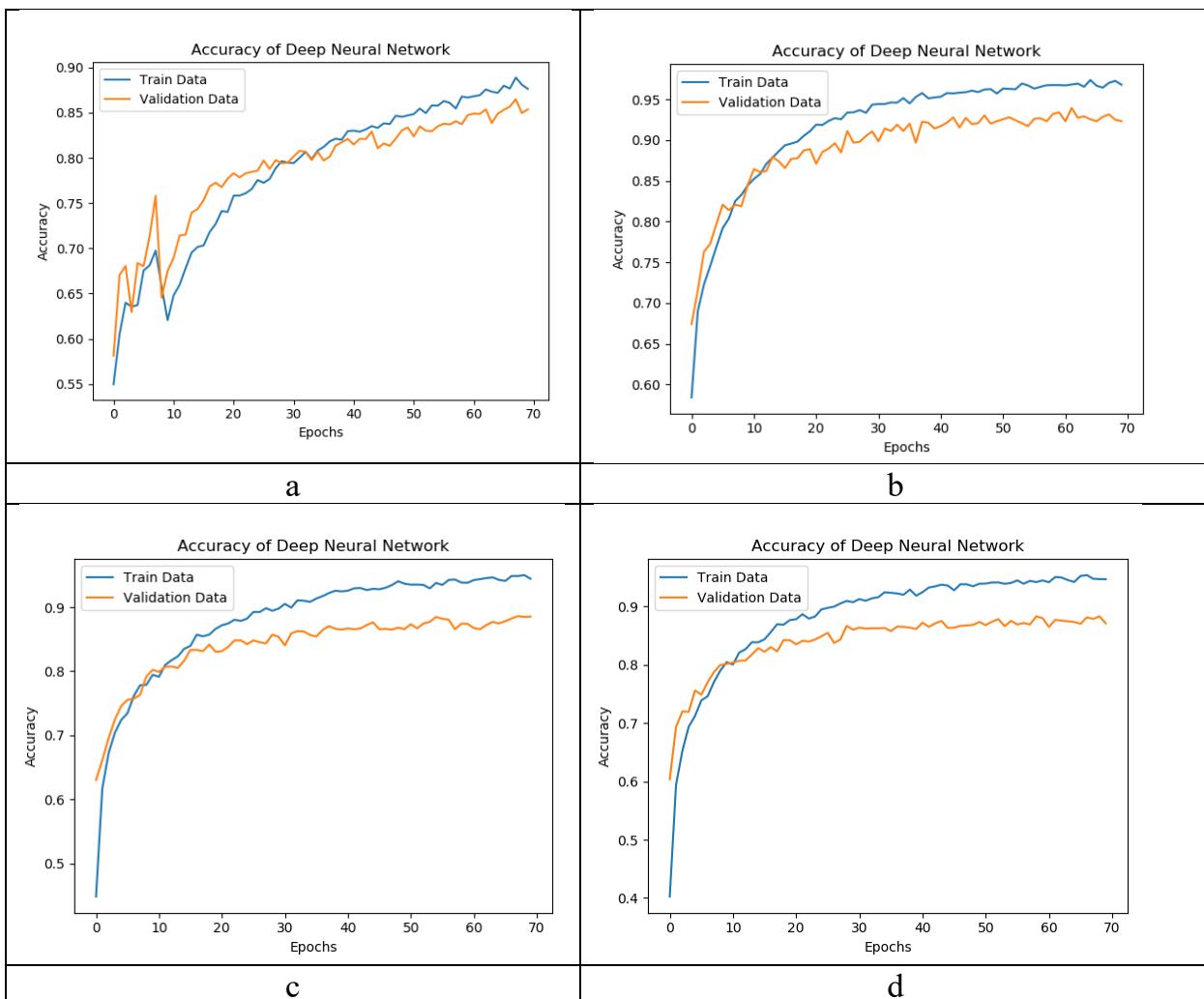


Figure 6. Accuracy of CNN-KCL for different numbers of clusters for 80/20 training/ testing. **a)** before clustering. **b)** 4 clusters. **c)** 10 clusters. **d)** 24 clusters.

Figures 5 and 6 show the accuracy in 30-70 modes (70% of the data for training and 30% of the data for testing) and 20-80 (80% of the data for training and 20% of the data for testing), respectively. In each of these two figures, the results with 4 clusters, 10 clusters, and 24 clusters are shown, and the blue and orange lines represent the training and validation process over 70 epochs, respectively. In **Figures 5 and 6**, in the 30-70 mode without clustering, the CNN-KCL network is not well trained and has low accuracy. While in 4-cluster, 10-cluster, and 24-cluster modes, the training and validation process are almost identical as shown in these figures, the best results achieved using CNN-KCL and $K = 4$ clusters.

It was observed in **Figures 7 and 8** that as loss of training data approached to zero, less error in detecting the input images was recorded. In general, the validation data should move in accordance with the training chart. If the distance between these two graphs increases, the difference between the two training and validation graphs increases, and as a result, the error increases and the network tends to overfit.

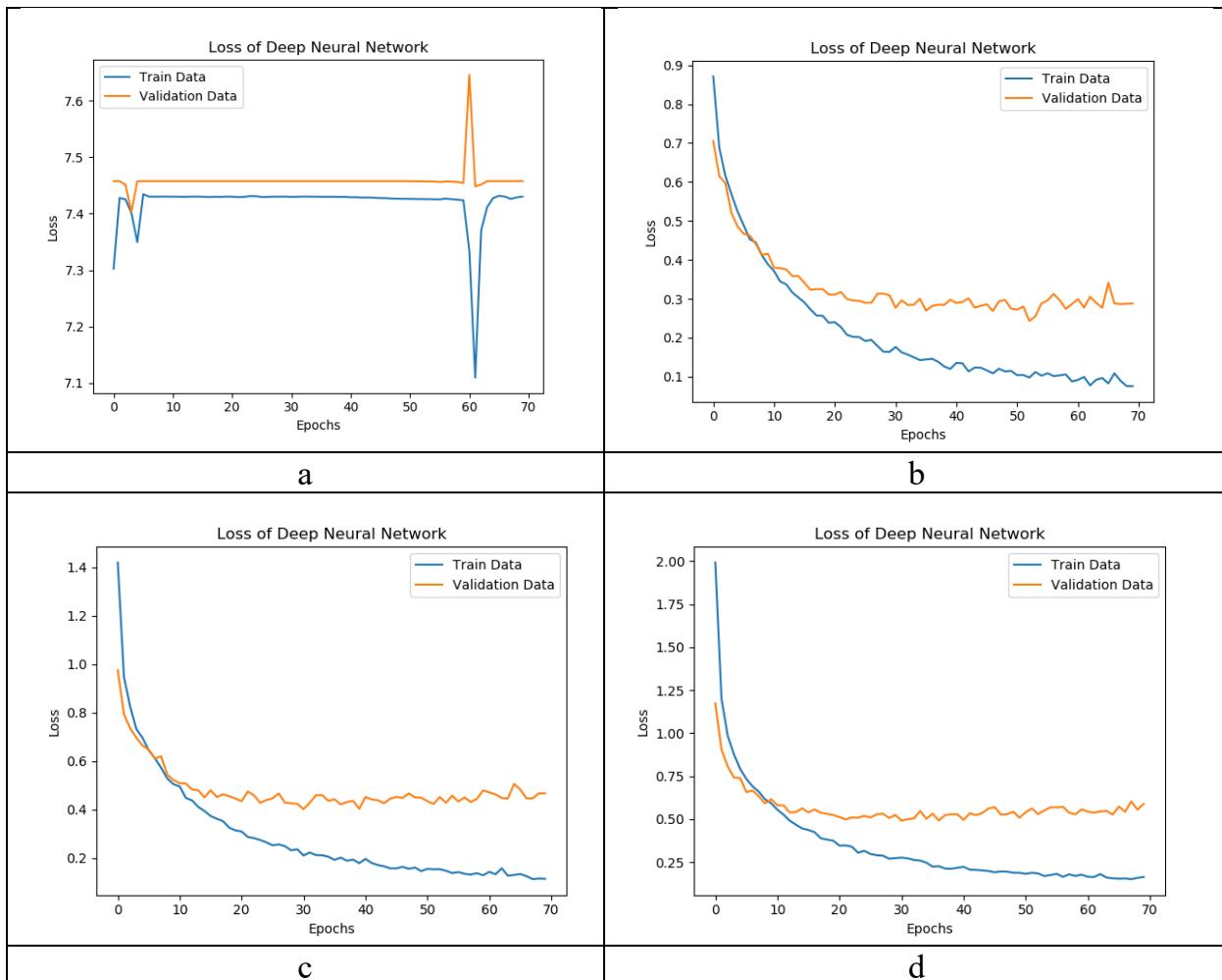


Figure 7. Loss of CNN-KCL for different number of clusters for 70/30 training/ testing. **a)** before clustering. **b)** 4 clusters. **c)** 10 clusters. **d)** 24 clusters.

As shown in **Figure 7**, in the case of 2 clusters, by dividing the data by 30-70 ratio, the error increases dramatically, and in the cases of 4, 10, and 24 clusters, the errors decrease. In the case of 4 clusters, we observe the least error. We also observe similar results in the 20-80 ratio data division in 4, 10, and 24 cluster modes in **Figure 8**.

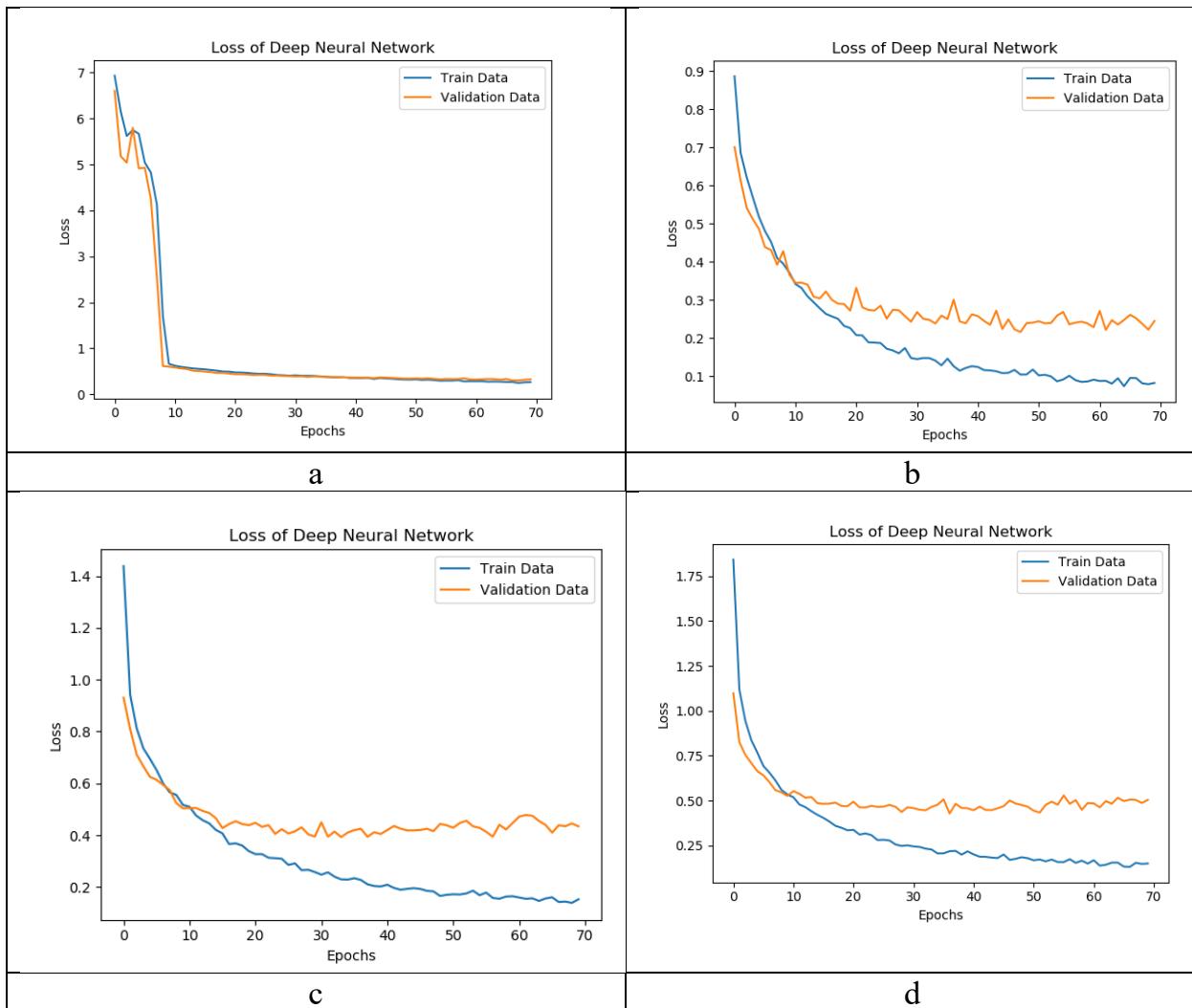


Figure 8. Loss of CNN-KCL for a different number of clusters for 80/20 training/ testing. **a)** before clustering. **b)** 4 clusters. **c)** 10 clusters. **d)** 24 clusters.

Finally, as demonstrated in **Figures 9** and **10**, we obtained the ROC diagram for the 4, 10, and 24 cluster modes to divide the data into 30-70 and 20-80 ratios. Here, the ROC used to evaluate different models. As shown in this figure, the proposed method can detect the disease very well in 4, 10 and 24 cluster modes. In **Figure 9**, in the case of 2 clusters, we have the lowest AUC value, while in cases of 4, 10, and 24 clusters, the performance of ROC chart are increasing. As shown in this figure, using 4 clustering, ROC chart demonstrates the best performance compared to other modes.

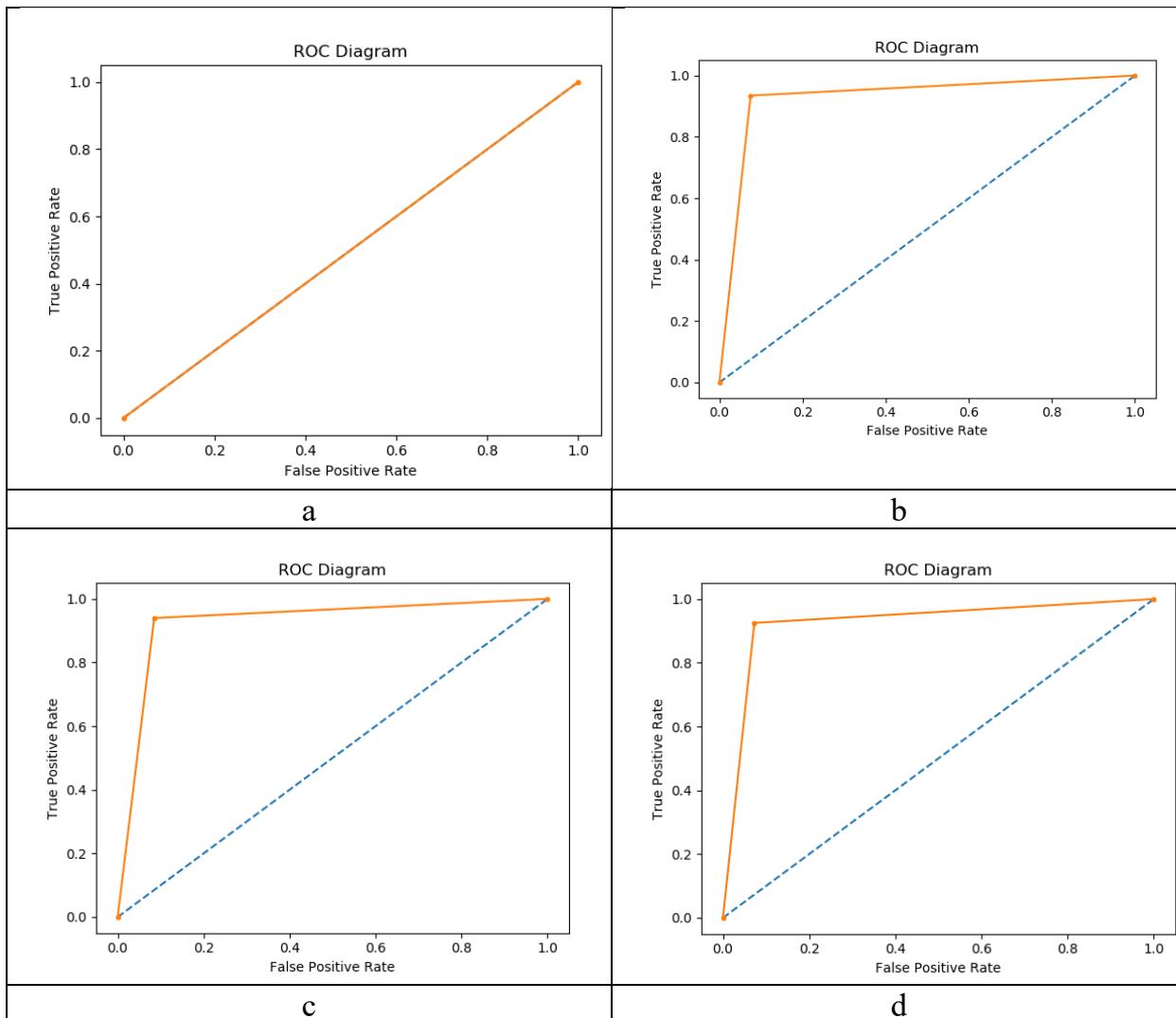


Figure 9. ROC of CNN-KCL for different number of clusters for 70/30 training/ testing. **a)** before clustering **b)** 4 clusters. **c)** 10 clusters. **d)** 24 clusters.

In addition, in **Figure 10**, we observe similar results in terms of ROC for the above modes (70/30 and 80/20 ratios). Finally, the results of the comparison for the diagnosis of myocarditis using CNN-KCL with 4, 10, and 24 clusters for data division in 30-70 and 20-80 ratios are presented in **Table 3**.

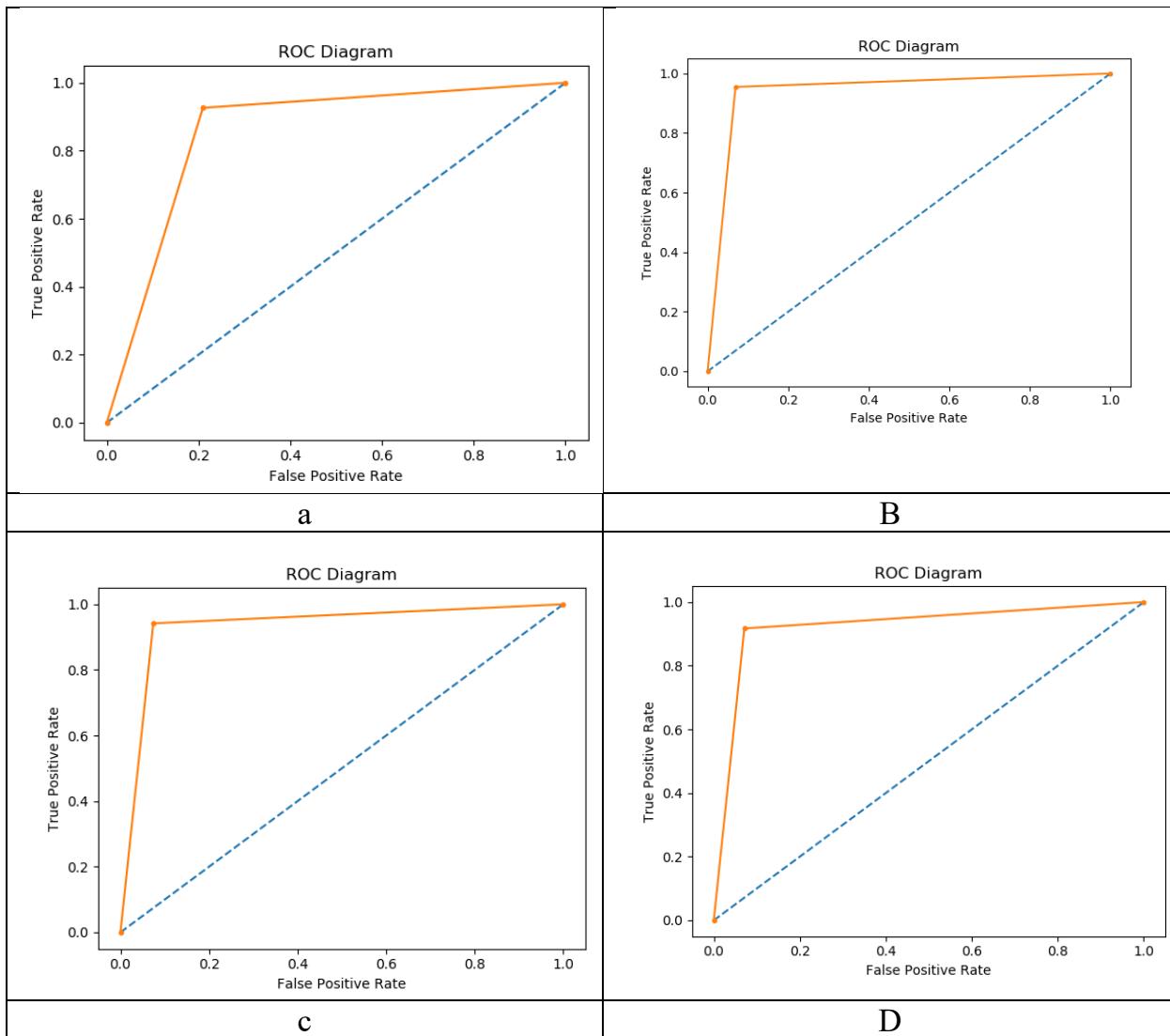


Figure 10. ROC of CNN-KCL for a different number of clusters for 80/20 training/ testing. **a)** before clustering. **b)** 4 clusters. **c)** 10 clusters. **d)** 24 clusters.

Table 3. The classification results achieved using CNN-KCL when using different number of clusters.

Number of clusters	Train/test	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1-score (%)	Loss
Without clustering	70/30	52.8	52.8	100.0	0	69.1	7.565
	80/20	85.0	92.8	79.1	92.6	85.4	0.317
4 clusters	70/30	91.9	94.3	92.5	93.5	93.4	0.272
	80/20	92.3	95.9	93.1	95.5	94.5	0.301
10 clusters	70/30	87.8	94.1	91.5	94.0	92.8	0.563
	80/20	88.1	94.7	92.6	94.2	93.6	0.394
24 clusters	70/30	84.9	93.8	92.7	92.5	93.3	0.691
	80/20	86.0	92.4	92.9	91.7	92.6	0.585

According to **Table 3**, we investigated various evaluation criteria including: precision, accuracy, recall, specificity, F1-Score, and Loss. We have examined 30-70 and 20-80 ratios to divide the data to build independent test set. In the non-clustering mode, by dividing the data using 30-70 ratio, the values of accuracy, precision, recall, specificity, F1-score, and Loss are 52.8%, 52.8%, 100%, 0.0, 69.1% and 7.6, respectively. By dividing the data using 20-80 ratio, the values of accuracy, precision, recall, specificity, F1-score, and Loss are 85.0, 92.8, 79.0, 92.6, 85.4 and 0.32, respectively. According to the obtained results, using the CNN-KCL method, when $K = 4$, by dividing the data into 30-70 and 20-80 ratios, the best performance is achieved compared to other modes.

6. Conclusion and future works

Myocarditis which remains one of the most challenging diagnoses in cardiology can affect the heart muscle and its electrical system. The most common cause of myocarditis is viral infection but other potential causes include Bacteria, Parasites, Fungi, Medications, or illegal drugs that might cause an allergic or toxic reaction, chemicals, or radiation [50]. Patients might present with a wide variety of symptoms like palpitations, dizziness, or syncope, serious ventricular arrhythmia, and angina like chest pain [4]. Sudden cardiac death and acute coronary syndrome in healthy young adults can be caused by heart failure or severe arrhythmia [51].

Nowadays, thanks to the advent of technology, diagnostic possibilities have expanded and improved but there is still room for additional modifications and improvements [52]. Endomyocardial biopsy (EMB) as the invasive gold standard for diagnosis of myocarditis is subject to sampling error, false negative, false-positive results, procedural risks, or lack of local expertise therefore due to these limitations is infrequently performed [53]. Among non-invasive methods, Electrocardiogram (ECG) and echocardiography are the primary diagnostic tools for the heart diseases which are applicable for excluding other causes of cardiomyopathy. They can help to document disease development because temporal changes in systolic function, chamber size, and thickness can be evaluated regularly. However, they do not provide determined signs for myocarditis. Their utility base more on excluding other causes in assessing heart function. Even normal findings in a patient do not exclude myocarditis [54]. When it is compared, cardiac MRI (CMR) has been considered as a noninvasive and golden standard diagnostic tool for suspected myocarditis and plays an indispensable role in diagnosing various cardiac diseases. CMR provides the possibility of anatomical and functional imaging and accurate assessment of heart. However, in this respect what is more important is its ability of tissue characterization [26, 52].

Based on our experimental results and studies on diseases, the convolutional neural network has an acceptable function for detecting and classifying images. However, no studies have used CNN-KCL to diagnose myocarditis. The images were very diverse because they were taken from different angles. Therefore, in dealing with such images, we used the K-means clustering method. We used the Elbow method to find the optimal K , which in the end, this method was not successful. For this reason, in this study, in this study, a novel hybrid CNN-KCL method was first used to diagnose and classify myocarditis on Z-Alizadeh Sani myocarditis dataset. Our proposed method consists of two parts: clustering and convolutional neural network. Due to the diversity and heterogeneity of the input data in that heart images,

we adopted clustering method to separate healthy samples from unhealthy samples. As a result, the non-tagged method of clustering with K-means was used. We considered different values for K (from 1 to 30). On the other hand, for testing, the data set was divided into 30-70 and 20-80 ratios. We demonstrated our results using various evaluation criteria including accuracy, F1-score, sensitivity, transparency, recall, and Loss. As shown in this paper, CNN-KCL when using 4-cluster mode demonstrates the best performance compared to other modes. As shown in Table 3, CNN-KCL achieves to 92.3%, 95.9%, 93.1%, 95.5%, 94.5%, and 0.3 in terms of accuracy, precision, recall, specificity, F1-score, and Loss, respectively. As of future works, we aim at using semi-supervised methods, reinforcement learning e.g., deep reinforcement learning, and supervised methods such as Recurrent Neural Networks (RNN) to enhance the prediction performance. We also aim at using explainable AI and Case-Based Reasoning (CBR).

Author contributions

SS, RA, and ZAS designed the study. DS, RA, JHJ, SH, AS, and FH wrote the manuscript; FH, ZAS, and AS collected data. SS, AD, and HAR edited the manuscript; DS, RA, and JHJ carried out the analyses. DS, RA, and JHJ generated all figures and tables. All authors have read and approved the final version of the paper.

Ethics approval and consent to participate

NA

Conflict of interest

The authors declare no competing financial and non-financial interests.

Funding

HAR is supported by UNSW Scientia Program Fellowship and is a member of the UNSW Graduate School of Biomedical Engineering.

Acknowledgments

HAR is supported by a UNSW Scientia Program Fellowship.

References

- [1] "Different heart diseases - World Heart Federation <https://www.world-heart-federation.org/resources/different-heart-diseases/>, accessed on 05.04.2020." (accessed.
- [2] J. H. Joloudari *et al.*, "Coronary Artery Disease Diagnosis; Ranking the Significant Features Using a Random Trees Model," *International Journal of Environmental Research and Public Health*, vol. 17, no. 3, p. 731, 2020. [Online]. Available: <https://www.mdpi.com/1660-4601/17/3/731>.
- [3] L. T. Cooper, "Myocarditis," *New England Journal of Medicine*, vol. 360, no. 15, pp. 1526-1538, 2009, doi: 10.1056/NEJMra0800028.
- [4] L. A. Blauwet and L. T. Cooper, "Myocarditis," *Progress in Cardiovascular Diseases*, vol. 52, no. 4, pp. 274-288, 2010/01/01/ 2010, doi: <https://doi.org/10.1016/j.pcad.2009.11.006>.

[5] A. M. Feldman and D. McNamara, "Myocarditis," *New England Journal of Medicine*, vol. 343, no. 19, pp. 1388-1398, 2000, doi: 10.1056/nejm200011093431908.

[6] M. Abdar, E. Nasarian, X. Zhou, G. Bargshady, V. N. Wijayaningrum, and S. Hussain, "Performance Improvement of Decision Trees for Diagnosis of Coronary Artery Disease Using Multi Filtering Approach," in *IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*, 2019, pp. 26-30.

[7] E. Nasarian *et al.*, "Association between work-related features and coronary artery disease: A heterogeneous hybrid feature selection integrated with balancing approach," *Pattern Recognition Letters*, vol. 133, pp. 33-40, 2020/05/01/2020, doi: <https://doi.org/10.1016/j.patrec.2020.02.010>.

[8] R. Alizadehsani *et al.*, "Hybrid genetic-discretized algorithm to handle data uncertainty in diagnosing stenosis of coronary arteries," *Expert Systems*, vol. n/a, no. n/a, doi: 10.1111/exsy.12573.

[9] R. Alizadehsani *et al.*, "Model uncertainty quantification for diagnosis of each main coronary artery stenosis," *Soft Computing*, vol. 24, no. 13, pp. 10149-10160, 2020/07/01 2020, doi: 10.1007/s00500-019-04531-0.

[10] H. Greenspan, B. v. Ginneken, and R. M. Summers, "Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1153-1159, 2016, doi: 10.1109/TMI.2016.2553401.

[11] B. C. S. Loh and P. H. H. Then, "Deep learning for cardiac computer-aided diagnosis: benefits, issues & solutions," *mHealth*, vol. 3, pp. 45-45, 2017.

[12] B. Baeßler, M. Mannil, D. Maintz, H. Alkadhi, and R. Manka, "Texture analysis and machine learning of non-contrast T1-weighted MR images in patients with hypertrophic cardiomyopathy—Preliminary results," *European Journal of Radiology*, vol. 102, pp. 61-67, 2018.

[13] M. Ovreiu and D. Simon, "Biogeography-based optimization of neuro-fuzzy system parameters for diagnosis of cardiac disease," presented at the Proceedings of the 12th annual conference on Genetic and evolutionary computation, Portland, Oregon, USA, 2010.

[14] M. Ali, M. F. Rani, A. H. Jahidin, M. F. Saaid, and M. Z. H. Noor, "Identification of cardiomyopathy disease using hybrid multilayered perceptron network," in *IEEE International Conference on Control System, Computing and Engineering*, 2012, pp. 23-27.

[15] D. Alis, A. Guler, M. Yergin, and O. Asmakutlu, "Assessment of ventricular tachyarrhythmia in patients with hypertrophic cardiomyopathy with machine learning-based texture analysis of late gadolinium enhancement cardiac MRI," *Diagnostic and Interventional Imaging*, vol. 101, no. 3, pp. 137-146, 2020.

[16] S. Borkar and M. N. Annadate, "Supervised Machine Learning Algorithm for Detection of Cardiac Disorders," in *Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, 2018, pp. 1-4.

[17] P. P. Sengupta *et al.*, "Cognitive Machine-Learning Algorithm for Cardiac Imaging," *Circulation: Cardiovascular Imaging*, vol. 9, no. 6, p. e004330, 2016, doi: doi:10.1161/CIRCIMAGING.115.004330.

[18] R. Begum and M. Ramesh, "Detection of cardiomyopathy using support vector machine and artificial neural network," *International Journal of Computer Applications*, vol. 133, no. 14, pp. 29-34, 2016.

[19] J. H. Joloudari, H. Saadatfar, A. Dehzangi, and S. Shamshirband, "Computer-aided decision-making for predicting liver disease using PSO-based optimized SVM with feature selection," *Informatics in Medicine Unlocked*, vol. 17, p. 100255, 2019/01/01/ 2019, doi: <https://doi.org/10.1016/j.imu.2019.100255>.

[20] E. M. Green, R. v. Mourik, C. Wolfus, S. B. Heitner, O. Dur, and M. J. Semigran, "Machine Learning Detection of Obstructive Hypertrophic Cardiomyopathy Using a Wearable Biosensor," *Circulation*, vol. 136, pp. A24031-A24031, 2017.

[21] D.-Y. Tsai and K. Kojima, "Measurements of texture features of medical images and its application to computer-aided diagnosis in cardiomyopathy," *Measurement*, vol. 37, no. 3, pp. 284-292, 2005.

[22] S. Narula, K. Shameer, A. M. Salem Omar, J. T. Dudley, and P. P. Sengupta, "Machine-Learning Algorithms to Automate Morphological and Functional Assessments in 2D Echocardiography," *Journal of the American College of Cardiology*, vol. 68, no. 21, p. 2287, 2016, doi: 10.1016/j.jacc.2016.08.062.

[23] Q. A. Rahman, L. G. Tereshchenko, M. Kongkatong, T. Abraham, M. R. Abraham, and H. Shatkay, "Utilizing ECG-Based Heartbeat Classification for Hypertrophic Cardiomyopathy Identification," *IEEE Transactions on NanoBioscience*, vol. 14, no. 5, pp. 505-512, 2015, doi: 10.1109/TNB.2015.2426213.

[24] X.-N. Shao *et al.*, "Texture analysis of magnetic resonance T1 mapping with dilated cardiomyopathy: A machine learning approach," *Medicine*, vol. 97, no. 37, pp. e12246-e12246, 2018.

[25] G. Captur *et al.*, "Identification of a Multiplex Biomarker Panel for Hypertrophic Cardiomyopathy Using Quantitative Proteomics and Machine Learning," *Molecular & Cellular Proteomics*, vol. 19, no. 1, p. 114, 2020, doi: 10.1074/mcp.RA119.001586.

[26] M. Chetrit and M. G. Friedrich, "The unique role of cardiovascular magnetic resonance imaging in acute myocarditis," *F1000Research*, vol. 7, pp. F1000 Faculty Rev-1153, 2018.

[27] M. D. Cornicelli, C. K. Rigsby, K. Rychlik, E. Pahl, and J. D. Robinson, "Diagnostic performance of cardiovascular magnetic resonance native T1 and T2 mapping in pediatric patients with acute myocarditis," *Journal of Cardiovascular Magnetic Resonance*, vol. 21, no. 1, pp. 40-48, 2019.

[28] M. A. G. M. Olimulder, J. van Es, and M. A. Galjee, "The importance of cardiac MRI as a diagnostic tool in viral myocarditis-induced cardiomyopathy," *Netherlands Heart Journal*, vol. 17, no. 12, pp. 481-486, 2009.

[29] C. Moenninghoff *et al.*, "Workflow Efficiency of Two 1.5 T MR Scanners with and without an Automated User Interface for Head Examinations," *Academic Radiology*, vol. 20, no. 6, pp. 721-730, 2013/06/01/ 2013, doi: <https://doi.org/10.1016/j.acra.2013.01.004>.

[30] N. Ghassemi *et al.*, "Material Recognition for Automated Progress Monitoring using Deep Learning Methods," *arXiv preprint arXiv:2006.16344*, 2020.

[31] A. Shoeibi *et al.*, "Epileptic seizure detection using deep learning techniques: A Review," *arXiv preprint arXiv:2007.01276*, 2020.

[32] A. Shoeibi *et al.*, "Automated Detection and Forecasting of COVID-19 using Deep Learning Techniques: A Review," *arXiv preprint*, pp. 1-20, 2020.

[33] Y. LeCun *et al.*, "Handwritten digit recognition with a back-propagation network," in *Advances in neural information processing systems*, 1990, pp. 396-404.

[34] 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.

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

[36] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1-9.

[37] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.

[38] S. Lawrence, C. L. Giles, T. Ah Chung, and A. D. Back, "Face recognition: a convolutional neural-network approach," *IEEE Transactions on Neural Networks*, vol. 8, no. 1, pp. 98-113, 1997, doi: 10.1109/72.554195.

[39] H. Shin *et al.*, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285-1298, 2016, doi: 10.1109/TMI.2016.2528162.

[40] U. R. Acharya *et al.*, "Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals," *Applied Intelligence*, vol. 49, no. 1, pp. 16-27, 2019/01/01 2019, doi: 10.1007/s10489-018-1179-1.

[41] U. R. Acharya, H. Fujita, O. S. Lih, M. Adam, J. H. Tan, and C. K. Chua, "Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network," *Knowledge-Based Systems*, vol. 132, pp. 62-71, 2017/09/15/ 2017, doi: <https://doi.org/10.1016/j.knosys.2017.06.003>.

[42] J. H. Tan *et al.*, "Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals," *Computers in Biology and Medicine*, vol. 94, pp. 19-26, 2018/03/01/ 2018, doi: <https://doi.org/10.1016/j.combiomed.2017.12.023>.

[43] M. Khodatars *et al.*, "Deep Learning for Neuroimaging-based Diagnosis and Rehabilitation of Autism Spectrum Disorder: A Review," *arXiv preprint arXiv:2007.01285*, 2020.

[44] K. Wagstaff, C. Cardie, S. Rogers, and S. Schrödl, "Constrained k-means clustering with background knowledge," 2001.

[45] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651-666, 2010/06/01/ 2010, doi: <https://doi.org/10.1016/j.patrec.2009.09.011>.

[46] P. Bholowalia and A. Kumar, "EBK-means: A clustering technique based on elbow method and k-means in WSN," *International Journal of Computer Applications*, vol. 105, no. 9, 2014.

[47] R. Alizadehsani *et al.*, "A database for using machine learning and data mining techniques for coronary artery disease diagnosis," *Scientific Data*, vol. 6, no. 1, p. 227, 2019/10/23 2019, doi: 10.1038/s41597-019-0206-3.

[48] E. M. Song *et al.*, "Endoscopic diagnosis and treatment planning for colorectal polyps using a deep-learning model," *Scientific Reports*, vol. 10, no. 1, p. 30, 2020/01/08 2020, doi: 10.1038/s41598-019-56697-0.

[49] S. Hussain and G. Hazarika, "Educational data mining model using rattle," *Editorial Preface*, vol. 5, no. 6, 2014.

[50] I. Kindermann *et al.*, "Update on Myocarditis," *Journal of the American College of Cardiology*, vol. 59, no. 9, p. 779, 2012, doi: 10.1016/j.jacc.2011.09.074.

[51] T. S. Kafil and N. Tzemos, "Myocarditis in 2020: Advancements in Imaging and Clinical Management," *JACC: Case Reports*, vol. 2, no. 2, pp. 178-179, 2020.

[52] A. d. Roos, "Diagnosis of Myocarditis at Cardiac MRI: The Continuing Quest for Improved Tissue Characterization," *Radiology*, vol. 292, no. 3, pp. 618-619, 2019, doi: 10.1148/radiol.2019191476.

[53] F. Dominguez, U. Kühl, B. Pieske, P. Garcia-Pavia, and C. Tschöpe, "Update on Myocarditis and Inflammatory Cardiomyopathy: Reemergence of Endomyocardial Biopsy," *Revista Española de Cardiología (English Edition)*, vol. 69, no. 2, pp. 178-187, 2016.

[54] C. Buttà, L. Zappia, G. Laterra, and M. Roberto, "Diagnostic and prognostic role of electrocardiogram in acute myocarditis: A comprehensive review," *Annals of Noninvasive Electrocardiology*, vol. 1, pp. 1-10.