

Computer Vision-based Efficient Segmentation Method for Left Ventricular Epicardium and Endocardium using Deep Learning

A F M Saifuddin Saif, Trung Duong, Zachary Holden

Computer Science and Information Systems, West Virginia University Institute of Technology,
Beckley, West Virginia 25801, USA¹

School of Engineering, Colorado State University Pueblo, Pueblo, CO 81001, USA^{2,3}

Abstract—Segmentation of the Left Ventricular Epicardium and Endocardium remains challenging and significant for valuable investigation of cardiac image classification. Previous research methods did not consider the flexibility of the heart area, so measurements needed to be more consistent and accurate. In addition, previous methods ignored the presence of affectability and additional parts, such as the lung organ inside the frame, during segmentation. Deep learning architectures, specifically convolutional neural networks, have become the primary choice for assessing cardiac medical images. In this context, a Convolutional Neural Network (CNN) can be an effective way to segment the left ventricular epicardium and endocardium as CNN can take data pictures, move enormity to various centers or objects in the image and have the choice to separate one from the other. This research proposes an efficient method for segmenting the left ventricular epicardium and endocardium using the InceptionV3 convolutional neural network. Rather than including fully connected layers on the head of the component maps, the proposed method considers the average of each element map, and the subsequent vector was taken care of legitimately into the SoftMax layer. Data augmentation technique was used to validate the proposed method on large number of dataset images. Besides, the proposed method was validated in publicly available MRI cardiac image datasets. Comprehensive experimental analysis was done by analyzing a large number of performance metrics, i.e., cosine similarity, log cos error, mean absolute error, mean absolute percentage error, mean squared error, mean squared logarithmic error, and root mean squared error. The proposed method depicted superior performance for localization of the left ventricular epicardium and endocardium in terms of all these performance metrics. In addition, the proposed method performed efficiently to get smooth curve for covering the region due to usage of interpolation technique to draw the curve, which made it smoother compared with previous research.

Keywords—Convolutional neural network; segmentation; computer vision; deep learning

I. INTRODUCTION

Coronary artery disease (CAD) has the highest morbidity and mortality rates globally [1]. For this, localization of the left ventricular epicardium and endocardium using deep learning approach can be automated to provide a robust tool for imaging the structure of the human heart. Generally, LV

segmentation methods are formed based on area and time, where the area locates the heart within the midpoint to the indicated frame [2]. Previous research methods did not consider the flexibility of the heart area, so measurements needed to be more consistent and accurate. Time-based strategies acknowledge the heart to be the central working fact within the frame [3]. These methods endured the absence of affectability and additional parts, such as the lung, which is an active organ inside the frame in expansion for movement production [2]. More estimation has been put forward to handle this issue for LV segmentation in MRI [4] [5]. Besides, due to excellent efficacy or cost ratio in evaluating left ventricular function, gated myocardial perfusion SPECT (MPS) is widely investigated for non-destructive diagnosis of CAD [6]. Endocardium and epicardium must be accurately delineated on perfusion images for quantitative analysis of the left ventricle (LV) in MPS followed by measurement of LV functional parameters. In this context, manual segmentation is time-consuming and needs more reproducibility [6]. Therefore, to improve the accuracy of quantitative analysis, it is necessary to develop a precise, reproducible, and fully automated localization method.

At present, industrial software extracts the left ventricular epicardium and endocardium surface features by estimating the maximum myocardial counts. Then, Gaussian fit is implicated with empirical standard deviation or threshold method to evaluate the details. However, this method needs to be investigated again in assessing myocardial functions. In particular, left ventricular ejection fraction (LVEF) is often overestimated in patients with tiny hearts, and the error is more pronounced in females than males [7].

Traditional computer vision-based image processing methods have demonstrated significant improvement in cardiac image segmentation, such as atlas and model based methods [8, 9]. In recent years, deep learning models, which automatically learn high-level features of the potential distribution of data, outperformed traditional images segmentation methods in accuracy and time efficiency [10]. In this context, multi-class three-dimensional (3D) V-Net was proposed to automatically segment the endocardium and epicardium in gated MPS, which exhibits improved performance [11]. The average Dice similarity coefficient (DSC) values of the model in the endocardium and epicardium

of regular patients were 0.907 and 0.965, respectively. However, previous deep learning methods still need to be improved for accurate LV segmentation in MPS towards localization. The object shapes extracted from previous image segmentation methods for localization have shown great success as prior knowledge in refining the deep learning models for medical image classification [12,13]. In this context, combination of prior knowledge reduces the potential output space of model partitioning and speeds up the convergence during the model training. However, prior knowledge is generally used as the input which was hard to extract.

This research uses Inception-v3 CNN architecture for segmenting the left epicardium and endocardium. The innovations and our contributions are listed as follows:

1) This research proposes an efficient method for segmenting the left ventricular epicardium and endocardium using Inception-v3 CNN architecture where the average of each element map and the subsequent vector were taken care of legitimately into the SoftMax layer rather than including fully connected layers on the head of the component maps.

2) Data augmentation technique was used to overcome the shortage of dataset images by the proposed method, which allows the proposed method to be validated robustly.

3) The proposed method was validated by analyzing a large number of performance metrics, i.e., cosine similarity, log cosh error, mean absolute error, mean absolute percentage error, mean squared error, mean squared logarithmic error, and root mean squared error, where proposed method depicted superior performance.

The remainder of this paper is organized as follows: Section I gives the introduction. Critical previous research is illustrated Section II, comprehensive details of the proposed methodology are elaborated in Section III, details of experimental results with analysis for validation are presented in Section IV, and finally, Section V concludes the paper.

II. PREVIOUS RESEARCH METHODS

Manual left ventricular epicardium and endocardium segmentation are crucial for risk stratification, diagnosis, and treatment evaluation. However, manual segmentation has been suffering from various issues, i.e., time-consuming, tedious, and lack of generalization, which can impact the reproducibility of the results [14]. Automatic segmentation for segmentation can overcome some of these limitations [15], where deep learning methods are under investigation to develop accurate, robust, and fast computer vision techniques. As the clinical application of MRI is rapidly growing, robust computer vision techniques are required, which will not need any supervision for acceptable accuracy. This research developed and implemented a robust method for segmenting the Left Ventricular (LV) Epicardium and Endocardium using efficient deep learning method.

Previous Research reported the application of Deep Learning (DL) to segment the left ventricular epicardium and endocardium. Research in [16] evaluated multiple DL

methods for left ventricular endocardium segmentation and found the superiority of encoder–decoder-based architectures over non deep learning methods. Research in [17] implicated U-Net to segment the left ventricle by changing UNet architecture in MFP-U-Net. Their proposed CNN added additional convolution layers for producing fixed size feature maps and efficient left ventricular segmentation performance. Research in [18] combined a modified U-Net architecture with an FCN encoder to influence feature extraction and allow the system to learn from execution. Research in [19] implemented bilateral segmentation network to extract deep features and a pyramid local attention algorithm to extract significant features within compact and sparse neighboring contexts. Research in [20] used multiple parallel pipelines for ES and ED frame segmentation using DeepResU-Net. Distinct from the other Research, Research in [22] used self-supervised algorithms [21] to separate the left ventricle to reduce the issue for the lack of labeled data. Research in [23] addressed object detection method and YOLOv3 algorithm, to detect three points of the ventricular chamber and segment the ventricles. Despite these method's innovativeness and high performance, previous methods focused on segmenting the ventricle for segmentation but not all its anatomical structures.

Convolutional Neural Network (CNN) is a deep learning strategy that can extract data features, move enormity to various centers or objects in the image, and have the choice to separate one from the other. While in harsh strategies for CNN, channels are hand-worked with enough preparation, ConvNets can get capacity with these channels. Most LV limitation techniques are primarily founded on spot-based, time-based, and shape-based speculations, which refer to the areas of strategies except the heart in the picture [24]. A combination of the dynamic figure model and dynamic appearance models were used to confine the left and right ventricles of customary and Tetralogy of Fallot (TOF) hearts on 4-D (3-D+time) MR pictures [25] [26]. For each ventricle, a 4-D model was first used to accomplish incredible essential confinement on all heart stages, and a 3-D model was applied to each stage to increase the exactness while keeping up the complete heartiness of the 4-D division [27]. Another procedure was introduced in Deep CNN to restrict the Left Ventricular in cardiovascular MRI. A six-layered Convolutional Neural Network with different part estimations was used to separate highlights trailed by SoftMax, a connected layer for portrayal.

The pyramids of scales assessment were familiar with the record of the different dimensions of the heart [28]. A range-based device was produced to draw closer to experiencing the ill effects of affectability [29]. Automatic-Image-Driven technique's suppositions depend on the heart, roughly in the middle of the genuine picture. In this context, the LV blood pool is more roundabout than the Right Ventricular blood pool, which has an upper sign force [30]. Artificial Intelligence (AI), Computer Vision (CV), and Image Processing (IP) calculations have been likewise introduced to handle the segmentation of issues by isolating the frontal regional object from the foundation. While hardly any specific methods have been proposed to deal with the issue of LV restriction in X-beam, a couple of computer vision and image

processing methods have been familiar with limited unmistakable body parts in modalities, i.e., Ultrasound and Computed Tomography (CT). Kellman assessed and limited the LV posture using probabilistic boosting trees and minimal space learning [31]. Research in [32] utilized nonlinear planning through relapse to limit in echodiographic dataset. Modified LV (Left Ventricle) limitation in cardiovascular MRI pictures is a significant development for programmed division and practical perception. In this context, a comparative investigation should be combined into the severe degree of chance for backslide to improve the restriction task [33]. Recently, substantial convolutional systems have accomplished magnificent execution in many pictorial division assignments [34] and are excitedly applied in the field of clinical picture appraisal [35]. For instance, research in [36] suggested a 3D essentially oversight system for the robotized division of the liver and the entire heart, which needs further investigation due to a lack of datasets. Different study attempts were made to approach the problem of segmentation of left ventricular epicardium and endocardium activity. This research proposes an efficient method for segmenting the left ventricular epicardium and endocardium using the InceptionV3 CNN model. The proposed method considers the average of each element map, and the subsequent vector was taken care of legitimately into the SoftMax layer. The proposed method depicted superior performance for segmentation of left ventricular epicardium and endocardium.

III. PROPOSED RESEARCH METHODOLOGY

This research used InceptionV3 as the CNN architecture. The overall proposed methodology is shown in Fig. 1.

Top layer with a custom network is trained rather than including fully connected layers on the head of the component maps. The average of each element map was used, and the subsequent vector was taken care of legitimately into the SoftMax layer. Global average pooling was aggregated with spatial data for spatial solid interpretations of the features. The input layer was normalized by adjusting and scaling the activations. To build the solidness of the network, batch normalization normalized the yield of a previous actuation layer by removing the batch mean and separating it by batch standard deviation.

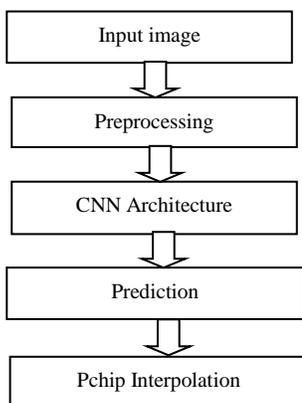


Fig. 1. Proposed methodology.

A. Input Images

Cardiac MR Image sequences with short-axis were used from 33 subjects for 7980 2D images. All the subjects were under the age of 18 where each patient's image sequence consisted of exactly 20 frames, and the number of slices collected along the long axis of the subjects ranges between 8 and 15. Spacing between slices ranged between 6 and 13 mm. Each image slice consisted of $256 * 256$ pixels with a pixel spacing of 0.93–1.64 mm.

B. Preprocessing

Each subject's arrangement was comprised of 20 frames and 8 to 15 slices along the long axis for an aggregate of 7980 pictures. However, images were raw and unprocessed, kept as 16-bit DICOM images. So, this research converted 16-bit DICOM input images into $256*256*3$ by reshaping them, as shown in Fig. 2.

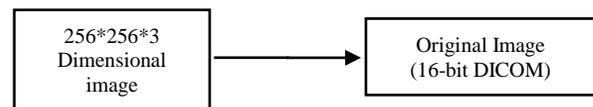


Fig. 2. Conversion of 16-bit DICOM input images into $256*256*3$.

C. Data Augmentation

The paucity of data is another main problem in establishing deep learning models like CNN. Data augmentation is a helpful strategy in building a convolutional neural network that can expand the size of the training set without procuring new pictures. In this context, frames are copied with some variety. This research expanded the image to safeguard the highlights key to make predictions yet revamp the pixels enough that it includes some noise. In addition, this research rescaled images by dividing 255 with every pixel. In this context, insufficient data for model training was a typical scenario, whereas 5011 segmentation images were available. The data augmentation technique provided strong support with reduced loss in that context.

D. CNN Architecture

Each input image was passed through convolution layers with kernels, max pooling, and fully connected layers. SoftMax function was applied to segment objects with probabilistic values between 0 and 1. Convolution preserved the correlation between pixels by understanding features using squares of input data. Max pooling is considered the most significant element from the dense feature map. Fully connected layers are used where all the inputs from one layer are added to every activation unit of the next layer. In this context, the last few layers were fully connected layers, which compiled the data extracted by previous layers to form the final output. CNN architecture deployed by this research is shown in Fig. 3.

E. Prediction

The cardiac MRI dataset provides short-axis cardiac MR images and ground truth of their left ventricles endocardial and epicardial segmentations. Each image was manually segmented for a total of 7980 images where both the

endocardium and epicardium of the left ventricle were visible, for a total of 5011 segmented MR images and 10022 contours. Each contour was described by 32 points given in pixel coordinates. This research trained the proposed method based on these 32 points (target variable). After training the model, a prediction of 64 points (32 for epicardium and 32 for endocardium) was acquired.

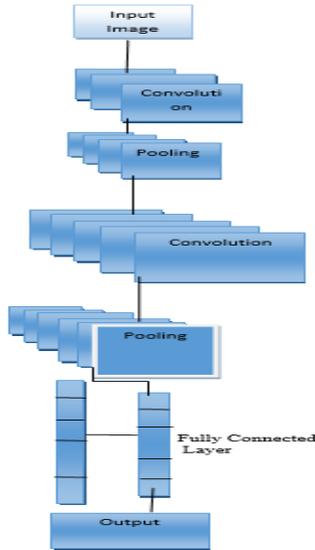


Fig. 3. CNN architecture used in the proposed research.

F. Interpolation

The proposed method used interpolation to draw the smooth curve. In this context, for each image, the epicardium and endocardium contour were depicted by 32 points in pixel coordinates (so, in total, 64 points). In this context, Pchip stands for Piecewise Cubic Hermite Interpolating polynomial used by the proposed method that interpolates data and specified derivatives at the interpolation points. As two points determine a linear function, two points and two given slopes determine a cubic. The data points are known as "knots." Y-values remain at the knots, so to get a particular PCHIP, the proposed method specified the values of the derivative y' at the knots. In addition, these two cubic polynomials were considered in x on the interval $1 \leq x \leq 2$. These functions were formed by adding cubic terms that vanish at the endpoints to the linear interpolant. After getting the predicted values of epicardium and endocardium, those points were passed for interpolation.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Dataset

Cardiac MRI datasets [30] were developed from 33 subjects. This dataset contains 20 frames and 8-15 slices along the axis for individual subject sequences. Most slices in this dataset contain heart anomalies because of some heart diseases. So, this dataset consists of a total of 7980 images. In Cardiac MRI datasets, every patient contains 20 timeframes for 8-15 short-axis slices with matrix size 256x256, 6-13 mm slice thickness, and pixel resolution of 0.93-1.64mm [31]. The

images were originally stored as 16-bit DICOM images. Form 7980 images, there were 5011 segmented images with 10022 contours. Each shape was pointed by 32 given in pixel coordinates.

In Cardiac MRI datasets, 4008 images were used by this research for training purposes. This research used 501 images for validation, and for testing, this research used 502 images. Besides, "Adam" as an optimization algorithm was used by the proposed method [37][38]. The preliminary learning rate was 0.02, the learning rate decreased with factor = 0.4, and patience was three by monitoring the 'validation loss.' Relu was used as an activation function and in the outcome layer, proposed used liner as an activation function.

B. Hardware and Software Set Up

For experimental purposes, a Windows platform was used with an 8th generation Intel Quad-Core i7-7300HQ processor (14MB Cache, 4.0GHz), 16GB DDR4 DRAM, and NVIDIA GeForce GTX 1050 with 16GB VRAM. This research used TensorFlow v2.13.0 and Keras 2.13.1 RC1 [39] [40]. This research also used some modules, i.e., Numpy, Matplotlib, Itk, Seaborn, Sklearn, and Scipy, for experimentation. The whole dataset was divided into two sections, i.e., one for training and the other for testing. The training section contains 80 % of the dataset images, and the remaining 20% was used for testing.

C. Performance Metrics

To validate the performance of the proposed method, various performance metrics were used, i.e., cosine similarity, log cosh error, mean absolute error, mean absolute percentage error, mean squared error, mean squared logarithmic error, and root mean squared error for graphical representation. Details of these metrics are mentioned in subsequent sections.

1) *Cosine similarity*: Estimating the comparability between at least two vectors is called cosine similarity [41]. The vectors are ordinarily nonzero and are inside an inward item space. In practice, cosine similarity is used to decide how comparable the documents are regardless of their size, which is a value equal to the division between the dot product of vectors and the product of the Euclidean norms or magnitude of each vector mentioned in Eq. (1).

$$similarity = \cos \theta = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (1)$$

Here, A and B are two vectors. A.B is the dot product of those two vectors.

2) *Log cosh error*: Log Cosh Error is used in a regression task, a logarithm of hyperbolic cosine of the prediction error. Log Cosh Error works like mean squared error, which is not firmly influenced by the occasional wildly incorrect prediction. Log Cost Error was estimated using Eq. (2).

$$L(y, y^p) = \sum_{i=1}^n \log(\cosh(y_i^p - y_i)) \quad (2)$$

In this equation, L denotes Log Cosh Error, $\log()$ denotes logarithm function, $\cosh()$ denotes hyperbolic cosine function, $(y_i^p - y_i)$ denotes difference between two points in y-axis.

3) Mean Absolute Error (MAE): Mean Absolute Error (MAE) is a widely recognized metric that quantifies precision for persistent factors [42]. MAE estimates the average magnitude of the errors in a set of predictions without thinking about their path. Besides, MAE is the difference between true or actual values and predicted values measured using Eq. (3).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_j - \hat{y}_j| \quad (3)$$

Here, MAE denotes Mean Absolute Error, $y_j - \hat{y}_j$ denotes difference between prediction and true value, n denotes total number of data points.

4) Mean absolute percentage error: Mean Absolute Percentage Error (MAPE) is the mean or normal of the total rate errors of forecasts [43]. Error is characterized as the result of the observed value by subtracting the forecasted value. Percentage errors are added regardless of the sign to register MAPE. This measurement is straightforward because it gives the error as far as percentage. Likewise, absolute percentage errors are utilized, and the issue of positive and negative errors offsetting each other is dodged. Thus, MAPE has administrative allure and is regularly utilized in estimation. MAPE can be well defined by middling the Absolute Percentage Errors of forecasts. MAPE is estimated using Eq. (4).

$$MAPE = \{(Estimated\ Value - Actual\ Value) \times 100\} \quad (4)$$

5) Mean Squared Error (MSE): The average of the squared error is called the Mean Squared Error (MSE) [44]. MSE reveals how close a regression line is to a set of points, which is done by taking the distances from the points to the regression line. The squaring is essential to eliminate any negative signs, which likewise gives more weight to bigger contrasts using Eq. (5).

$$MSE = \frac{1}{y} \sum_{i=1}^n |y_j - \hat{y}_i| \quad (5)$$

Here, MSE denotes Mean squared error, n denotes number of data points, y_j denotes observed values and \hat{y}_i denotes predicted values.

6) Root Mean Squared Logarithmic Error (RMSLE): Root-mean-squared-logarithmic error (RMSLE) is a function mentioned in Eq. (6) used for finding the difference between predicted values and the actual values [45]. To comprehend the manipulation of RMSLE and corresponding disparities, it is imperative to calculate means squared Error (MSE) mean where MSE joins both fluctuation and inclination of the indicator. MSLE cares about the relative difference between predicted value and actual value.

$$RSMLE = \{\log(prediction + 1) - \log(actual + 1)\}^2 \quad (6)$$

7) Root Mean Squared Error (RMSE): Root Mean Squared Error (RMSE) is a recognized way to find the error of a regression model, which defines how close a fitted line is to data points [44]. RMSE was estimated using Eq. (7).

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x}_i)^2}{N}} \quad (7)$$

Here, i denotes variable, N denotes number of non-missing data points, x_i denotes actual observed values and \hat{x}_i denotes estimated time series.

D. Experimental Results

The proposed method received the best cosine similarity of 0.9977 for the Cardiac MRI dataset after 76 epochs, as shown in Fig. 4. For each of the performance metrics, the graphical representation is shown in Fig. 4 to 9, where the orange line was used for training, and the blue line was used for validation. This research considered epoch along with the x-axis and cosine similarity along with the y-axis. Estimation of cosine similarity turned into an essential factor for understanding likenesses between objects and provided the strong assumption between the train set and validation set's similarity. In addition, loss differences can be calculated from this estimation, which helped to ensure that the proposed method never faces an overfitting problem.

Cosine similarity calculates the comparability between two vectors of an inner product space. The output produces a value ranging from -1 to 1, indicating similarity where -1 is non-similar, 0 is orthogonal (perpendicular), and 1 represents total similarity. From Fig. 4, it can be observed that in 76 epochs, the value of cosine similarity was 0.9977, and then the model stopped learning. If the model continued learning, the proposed method would face the overfitting problem.

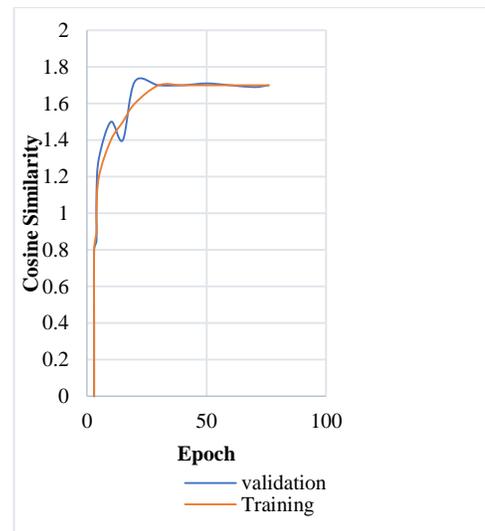


Fig. 4. Cosine similarity of cardiac MRI dataset.

This research received Log Cosh Error of 1.4585. Log Cosh Error was strongly affected by the occasional wide incorrection prediction and thus considered an improved version of MSE. Log Cosh Error is preferable to use when dataset contains significant errors due to more sensitivity to errors than the MSE. For this reason, this research used Log cos error for validating the proposed method. From Fig. 5, in 76 epochs, the model stopped learning. From 1 to 20 epoch, it had a large loss difference between train and validation data. But after the 20th epoch, loss difference was less.

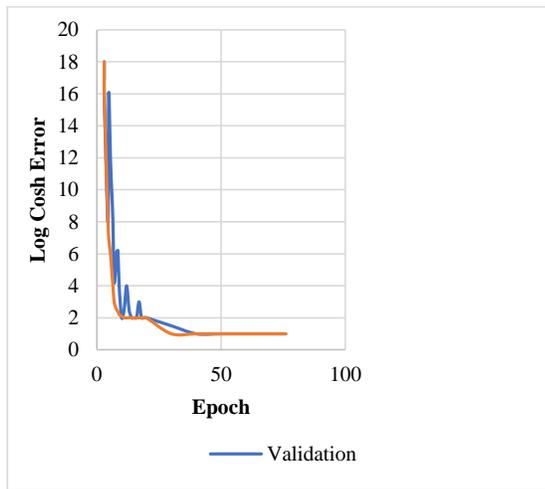


Fig. 5. Log cosh error of cardiac MRI dataset.

This research received Mean Absolute Error (MAE) of 2.0178. In many regression scenarios, Mean Absolute Error is preferable when the average error becomes very large, which is the main reason to use MAE in this research for validating the proposed method. In Fig. 6, in 76 epochs, MAE was 2.0178, then the model stopped learning. From 1 to 25 epochs, a significant loss difference was observed between train and validation data. However, after the 25th epoch, the loss difference was less.

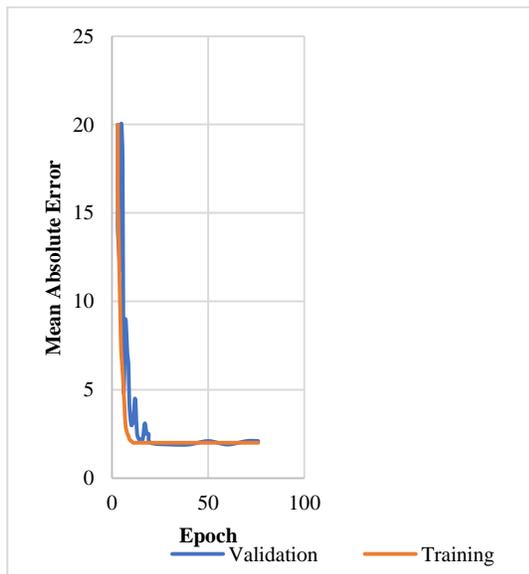


Fig. 6. Mean absolute error of cardiac MRI dataset.

This research received Mean-Squared-Logarithmic Error (MSLE) of $4.1755e-04$. Usage of MSLE during regression prevented significant errors from being significantly more penalized than small ones. For cases where the target value range was large, this was the main reason to use MSLE for validating the proposed method. In Fig. 7, in 76th epochs learning was stopped where from 1 to 20th epoch, the proposed method had a significant loss difference between train and validation data. However, after 20 epochs, the loss difference was less.

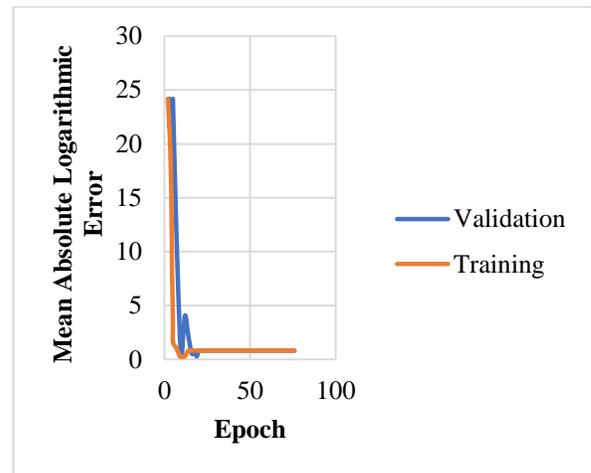


Fig. 7. Mean absolute percentage error on cardiac MRI dataset.

This research received Mean Squared Error (MSE) of 7.0577. MSE is preferable to use when the average error is very small. One minor difference with MAE was that the result is squared, which introduced benefits during optimization, which was the main reason for this research to use MSE for validation. In Fig. 8, in 76 epochs, the result of the Mean Squared Error was 7.0577, then the model stopped learning. From 1 to 25 epochs, it had a significant loss difference between train and validation data. However, after 25 epochs, the loss difference was less.

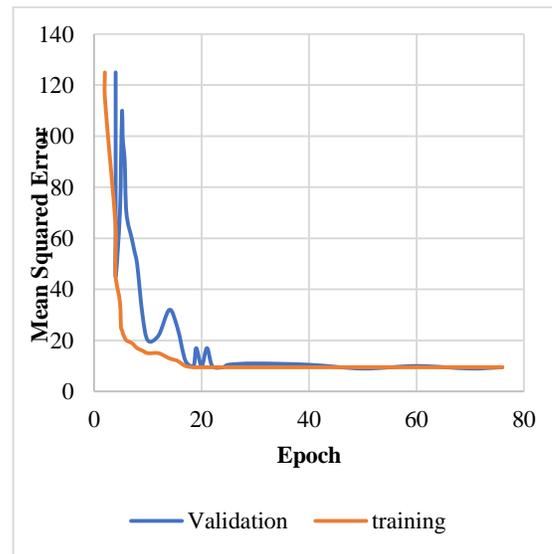


Fig. 8. Mean squared error of cardiac MRI dataset.

The proposed method received Root Mean Squared Error (RMSE) of 2.6561. RMSE was mostly applicable when significant errors were undesirable, which is the main reason for using RMSE in this research for validation. In Fig. 9, in 76 epochs, the Root Mean Squared Error result was 2.6561, then the model stopped learning. From 1 to 20 epochs, it had a significant loss difference between train and validation data. However, after 20 epochs, the loss difference was less. Overall experimental results for the proposed method are in Table I. Sample resultant outputs as image our show in Fig. 10.

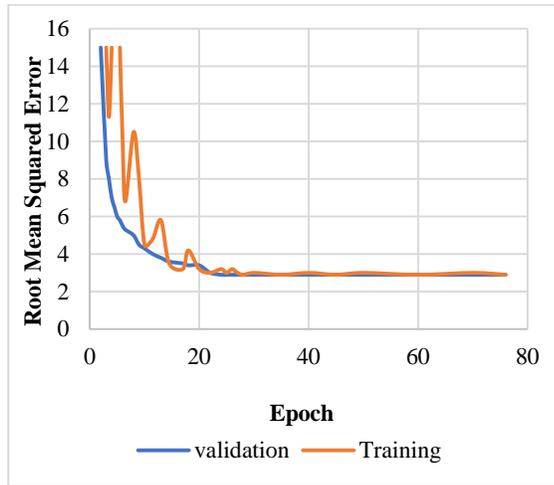


Fig. 9. Root mean square error of cardiac MRI dataset.

TABLE I. OVERALL EXPERIMENTAL RESULTS BASED ON CARDIAC MRI DATASET

Performance Metrics	Result
Cosine similarity	0.9977
Log Cosh Error	1.4585
Mean Absolute Error	2.0178
Mean Squared Logarithmic Error	4.1755e-04
Mean Squared Error	7.0577
Root Mean Squared Error	2.6561
Training Processing Time in Seconds	2500 s

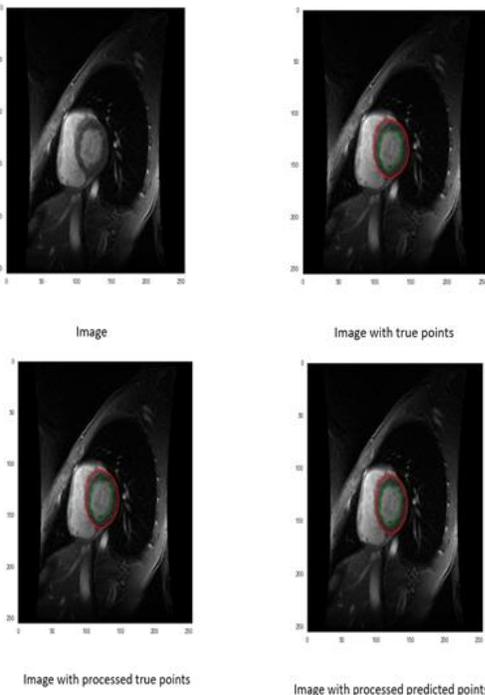


Fig. 10. Same output from Cardiac MRI dataset.

The proposed method was compared with existing research based on several metrics, i.e., Mean Absolute Error (MAE), Mean Absolute Percentage Error and Root Mean Squared Error (RMSE). Mean Absolute Error (MAE) was used for summarizing and measuring the quality of a deep learning model. The proposed method received MAE 2.0178, shown in Table II, after the compilation of training. Previously, research in [46] estimated the mean absolute error of 2.34 using a lightweight left ventricle localizer approach. The performance difference between the research in [46] and the proposed method by this research is that the number of data presented in the dataset used by research in [46] was significantly low to train the model perfectly. This research overcame the obstacle by using data augmentation approach to balance the sample data. Data augmentation is a significant strategy in building a convolutional neural network that can expand the size of the training set without procuring new frames. The proposed method rescaled the image by dividing 255 with every pixel. The proposed research could feed a decent amount of data through the network through this approach.

TABLE II. COMPARISON BASED ON MEAN ABSOLUTE ERROR

Method	Mean Absolute Error
Proposed Inception-V3 Convolutional Neural Network	2.0178
LVLNET and Fully Convolutional Neural Network [46]	2.34

The proposed method received Mean Absolute Percentage Error value of 1.5661, shown in Table III. Research in [47] received Mean Absolute Percentage Error of 1.43 using 3D active appearance models (AAM) [47]. Compared with the proposed method by this research with the 3D active appearance models on short axis cardiac, the proposed method provided better performance in terms of Mean Absolute Percentage Error. Research in [47] used the Gauss-Newton optimization technique. In contrast, this research did not use the Gauss-Newton optimization technique because even though Gauss-Newton optimization is accurate and reliable, the Gauss-Newton optimization technique is slow.

TABLE III. COMPARISON BASED ON MEAN ABSOLUTE PERCENTAGE ERROR

	Mean Absolute Percentage Error
Inception-V3 Convolutional Neural Network	1.5661
3D Active Appearance Model (AAM) and 2D + time active shape model (ASM) [47]	1.43

Root Mean Squared Error (RMSE) is primarily useful when significant errors are particularly undesirable. For RMSE, the lower value is better. In the proposed method, this research received RMSE of 2.6561. By using the multiple linear regression (MLR) model, research in [48] received an RMSE of 6.24. Using the random-forest regression (RFR) model, research in [46] received RMSE of 5.72. So, in both cases, this research's proposed method received better results compared with research in [46] and [48].

V. CONCLUSION

This research proposed an efficient method to segment left ventricular epicardium and endocardium using convolutional neural network. Data augmentation technique implicated by the proposed method expanded the size of the training set without procuring new images thus assisted to validate the proposed method on large number of dataset images. This research used Cardiac MRI dataset to validate the proposed methodology and estimated various performance metrics to justify the effectiveness robustly. After reshaping and rescaling the datasets images, proposed method used Inception V3 without top layer. Pchip interpolation technique was applied to smoothen the curve in the output images. This research observed the loss 2.011 because of the paucity of data which indicates effective localization of epicardium and endocardium with less error. Besides, due to usage of data augmentation, proposed research was able to feed a decent amount of data through network comparing with existing research. In addition, in comparison with other research, for 3D active appearance models on short axis cardiac, the proposed method provided better performance. In future, this research aims to improve CNN framework using various recent deep learning architecture such as Faster R-CNN and YOLOv7 for vast varieties of dataset images using data augmentation technique. Proposed method is expected to contribute significantly for more precise discovery and treatment of cardiovascular disease.

ACKNOWLEDGMENT

Authors would like to thank West Virginia University Institute of Technology and School of Engineering, Colorado State University Pueblo for various support for this research.

REFERENCES

- [1] A. Garavand, A. Behmanesh, N. Aslani, H. Sadeghsalehi, and M. Ghaderzadeh, "Towards diagnostic aided systems in coronary artery disease detection: a comprehensive multiview survey of the state of the art," *International Journal of Intelligent Systems*, vol. 2023, 2023.
- [2] M. A. Shoaib, J. H. Chuah, R. Ali, S. Dhanalakshmi, Y. C. Hum, A. Khalil, and K. W. Lai, "Fully Automatic Left Ventricle Segmentation Using Bilateral Lightweight Deep Neural Network," *Life*, vol. 13, p. 124, 2023.
- [3] W. H. Marshall, "Development of a competency-based curriculum in advanced heart failure training for the adult congenital heart disease cardiology fellow," The Ohio State University, 2023.
- [4] N. Das and S. Das, "A review on right ventricle cardiac MRI segmentation," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 11, pp. 1348-1358, 2023.
- [5] S. H. Chuah, L. K. Tan, N. A. Md Sari, B. T. Chan, K. Hasikin, E. Lim, N. M. Ung, Y. F. Abdul Aziz, J. Jayabalan, and Y. M. Liew, "Remodeling in Aortic Stenosis With Reduced and Preserved Ejection Fraction: Insight on Motion Abnormality Via 3D+ Time Personalized LV Modeling in Cardiac MRI," *Journal of Magnetic Resonance Imaging*, 2023.
- [6] F. Zhu, L. Li, J. Zhao, C. Zhao, S. Tang, J. Nan, Y. Li, Z. Zhao, J. Shi, and Z. Chen, "A new method incorporating deep learning with shape priors for left ventricular segmentation in myocardial perfusion SPECT images," *Computers in biology and medicine*, vol. 160, p. 106954, 2023.
- [7] K. Nakajima, T. Shibutani, F. Massanes, T. Shimizu, S. Yoshida, M. Onoguchi, S. Kinuya, and A. H. Vija, "Myocardial perfusion imaging with retrospective gating and integrated correction of attenuation, scatter, respiration, motion, and arrhythmia," *Journal of Nuclear Cardiology*, pp. 1-17, 2023.
- [8] X. Shu, Y. Yang, J. Liu, X. Chang, and B. Wu, "ALVLS: Adaptive local variances-Based levelset framework for medical images segmentation," *Pattern Recognition*, vol. 136, p. 109257, 2023.
- [9] F. Yuan, Z. Zhang, and Z. Fang, "An effective CNN and Transformer complementary network for medical image segmentation," *Pattern Recognition*, vol. 136, p. 109228, 2023.
- [10] Z. R. Mahayuddin and A. Saif, "A comprehensive review towards segmentation and detection of cancer cell and tumor for dynamic 3D reconstruction," *Asia-Pacific Journal of information technology and multimedia*, vol. 9, pp. 28-39, 2020.
- [11] Y. Zhang, F. Wang, H. Wu, Y. Yang, W. Xu, S. Wang, W. Chen, and L. Lu, "An automatic segmentation method with self-attention mechanism on left ventricle in gated PET/CT myocardial perfusion imaging," *Computer Methods and Programs in Biomedicine*, vol. 229, p. 107267, 2023.
- [12] S. Barraza-Aguirre, J. Diaz-Roman, C. Ochoa-Zezzatti, B. Mederos-Madrado, J. Cota-Ruiz, and F. Enriquez-Aguilera, "Segmentation of Lung Lesions Caused by COVID-19 in Computed Tomography Images Using Deep Learning," in *Internet of Everything for Smart City and Smart Healthcare Applications*, ed: Springer, 2023, pp. 237-259.
- [13] C. Zhao, Y. Xu, Z. He, J. Tang, Y. Zhang, J. Han, Y. Shi, and W. Zhou, "Lung segmentation and automatic detection of COVID-19 using radiomic features from chest CT images," *Pattern Recognition*, vol. 119, p. 108071, 2021/11/01/ 2021.
- [14] N. Vladimirov, E. Brui, A. Levchuk, W. Al-Haidri, V. Fokin, A. Efimtcev, and D. Bendahan, "CNN-based fully automatic wrist cartilage volume quantification in MR images: A comparative analysis between different CNN architectures," *Magnetic Resonance in Medicine*, 2023.
- [15] H. Dang, M. Li, X. Tao, G. Zhang, and X. Qi, "LVSegNet: A novel deep learning-based framework for left ventricle automatic segmentation using magnetic resonance imaging," *Computer Communications*, vol. 208, pp. 124-135, 2023.
- [16] C. D. Reddy, L. Lopez, D. Ouyang, J. Y. Zou, and B. He, "Video-Based Deep Learning for Automated Assessment of Left Ventricular Ejection Fraction in Pediatric Patients," *Journal of the American Society of Echocardiography*, vol. 36, pp. 482-489, 2023.
- [17] S. Moradi, M. G. Oghli, A. Alizadehasl, I. Shiri, N. Oveisi, M. Oveisi, M. Maleki, and J. Dhooge, "MFP-Unet: A novel deep learning based approach for left ventricle segmentation in echocardiography," *Physica Medica*, vol. 67, pp. 58-69, 2019.
- [18] K. B. Girum, G. Créhange, and A. Lalande, "Learning with context feedback loop for robust medical image segmentation," *IEEE transactions on medical imaging*, vol. 40, pp. 1542-1554, 2021.
- [19] F. Liu, K. Wang, D. Liu, X. Yang, and J. Tian, "Deep pyramid local attention neural network for cardiac structure segmentation in two-dimensional echocardiography," *Medical image analysis*, vol. 67, p. 101873, 2021.
- [20] M. G. R. Alam, A. M. Khan, M. F. Shejuty, S. I. Zubayear, M. N. Shariar, M. Altaf, M. M. Hassan, S. A. AlQahtani, and A. Alsanad, "Ejection Fraction estimation using deep semantic segmentation neural network," *The Journal of Supercomputing*, vol. 79, pp. 27-50, 2023.
- [21] M. Saeed, R. Muhtaseb, and M. Yaqub, "Is Contrastive Learning Suitable for Left Ventricular Segmentation in Echocardiographic Images?," *arXiv*, 2022.
- [22] M. Fischer, T. Hepp, S. Gatidis, and B. Yang, "Self-supervised contrastive learning with random walks for medical image segmentation with limited annotations," *Computerized Medical Imaging and Graphics*, vol. 104, p. 102174, 2023.
- [23] Z. Zhuang, P. Jin, A. N. Joseph Raj, Y. Yuan, and S. Zhuang, "Automatic segmentation of left ventricle in echocardiography based on YOLOv3 model to achieve constraint and positioning," *Computational and Mathematical Methods in Medicine*, vol. 2021, pp. 1-11, 2021.
- [24] H. Zhang, A. Wahle, R. K. Johnson, T. D. Scholz, and M. Sonka, "4-D cardiac MR image analysis: left and right ventricular morphology and function," *IEEE transactions on medical imaging*, vol. 29, pp. 350-364, 2009.
- [25] C. Payer, D. Štern, H. Bischof, and M. Urschler, "Multi-label whole heart segmentation using CNNs and anatomical label configurations," in

- International Workshop on Statistical Atlases and Computational Models of the Heart*, 2017, pp. 190-198.
- [26] D. F. Pace, A. V. Dalca, T. Geva, A. J. Powell, M. H. Moghari, and P. Golland, "Interactive whole-heart segmentation in congenital heart disease," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* 18, 2015, pp. 80-88.
- [27] S. P. O'Brien, O. Ghita, and P. F. Whelan, "A novel model-based 3D $\{+\}$ Time left ventricular segmentation technique," *IEEE transactions on medical imaging*, vol. 30, pp. 461-474, 2010.
- [28] S. Zambal, A. Schöllhuber, K. Bühler, and J. Hladuvka, "Fast and robust localization of the heart in cardiac MRI series," *Proc VISAPP*, vol. 1, pp. 341-6, 2008.
- [29] L. Zhong, J.-M. Zhang, X. Zhao, R. S. Tan, and M. Wan, "Automatic localization of the left ventricle from cardiac cine magnetic resonance imaging: a new spectrum-based computer-aided tool," *PloS one*, vol. 9, p. e92382, 2014.
- [30] Y. Lu, P. Radau, K. Connelly, A. Dick, and G. A. Wright, "Segmentation of left ventricle in cardiac cine MRI: An automatic image-driven method," in *Functional Imaging and Modeling of the Heart: 5th International Conference, FIMH 2009, Nice, France, June 3-5, 2009. Proceedings* 5, 2009, pp. 339-347.
- [31] P. Kellman, X. Lu, M.-P. Jolly, X. Bi, R. Kroeker, M. Schmidt, P. Speier, C. Hayes, J. Guehring, and E. Mueller, "Automatic LV localization and view planning for cardiac MRI acquisition," *Journal of Cardiovascular Magnetic Resonance*, vol. 13, pp. 1-2, 2011.
- [32] S. K. Zhou, B. Georgescu, X. S. Zhou, and D. Comaniciu, "Image based regression using boosting method," in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, 2005, pp. 541-548.
- [33] S. K. Zhou, J. Zhou, and D. Comaniciu, "A boosting regression approach to medical anatomy detection," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1-8.
- [34] 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.
- [35] H. R. Roth, L. Lu, N. Lay, A. P. Harrison, A. Farag, A. Sohn, and R. M. Summers, "Spatial aggregation of holistically-nested convolutional neural networks for automated pancreas localization and segmentation," *Medical image analysis*, vol. 45, pp. 94-107, 2018.
- [36] Y. Lu, P. Radau, K. Connelly, A. Dick, and G. A. Wright, "Segmentation of left ventricle in cardiac cine MRI: An automatic image-driven method," in *Functional Imaging and Modeling of the Heart: 5th International Conference, FIMH 2009, Nice, France, June 3-5, 2009. Proceedings* 5, 2009, pp. 339-347.
- [37] A. S. Saif, E. D. Wollega, and S. A. Kalevela, "Spatio-Temporal Features based Human Action Recognition using Convolutional Long Short-Term Deep Neural Network," *International Journal of Advanced Computer Science and Applications*, vol. 14, 2023.
- [38] R. Chowdhury and A. S. Saif, "Efficient mathematical procedural model for brain signal improvement from human brain sensor activities," *International Journal of Image, Graphics and Signal Processing*, vol. 10, p. 46, 2018.
- [39] M. H. Al Walid, D. Anisuzzaman, and A. S. Saif, "Data analysis and visualization of continental cancer situation by Twitter scraping," *International Journal of Modern Education and Computer Science*, vol. 11, p. 23, 2019.
- [40] A. M. Choudhury, A. S. Saif, and M. Rahman, "Toddler Sensory-Motor Development for Object Manipulation by Analyzing Hand-Pose," in *Proceedings of the International Conference on Computing Advancements*, 2020, pp. 1-7.
- [41] P. Xia, L. Zhang, and F. Li, "Learning similarity with cosine similarity ensemble," *Information sciences*, vol. 307, pp. 39-52, 2015.
- [42] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature," *Geoscientific model development*, vol. 7, pp. 1247-1250, 2014.
- [43] A. De Myttenaere, B. Golden, B. Le Grand, and F. Rossi, "Mean absolute percentage error for regression models," *Neurocomputing*, vol. 192, pp. 38-48, 2016.
- [44] A. Rabbani, H. Gao, A. Lazarus, D. Dalton, Y. Ge, K. Mangion, C. Berry, and D. Husmeier, "Image-based estimation of the left ventricular cavity volume using deep learning and Gaussian process with cardio-mechanical applications," *Computerized Medical Imaging and Graphics*, vol. 106, p. 102203, 2023.
- [45] K. Škardová, T. Hussain, M. Genet, and R. Chabiniok, "Effect of Spatial and Temporal Resolution on the Accuracy of Motion Tracking Using 2D and 3D Cine Cardiac Magnetic Resonance Imaging Data," in *International Conference on Functional Imaging and Modeling of the Heart*, 2023, pp. 235-244.
- [46] D. Abdelraouf, M. Essam, and M. Elattar, "LVLNET: lightweight left ventricle localizer using encoder-decoder neural network," in *2019 Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, 2019, pp. 235-238.
- [47] J. Park, D. Metaxas, and L. Axel, "Analysis of left ventricular wall motion based on volumetric deformable models and MRI-SPAMM," *Center for Human Modeling and Simulation*, p. 99, 1996.
- [48] S. Zhou, A. AbdelWahab, J. L. Sapp, J. W. Warren, and B. M. Horáček, "Localization of ventricular activation origin from the 12-lead ECG: a comparison of linear regression with non-linear methods of machine learning," *Annals of biomedical engineering*, vol. 47, pp. 403-412, 2019.