ISSN 2278-3091
Volume 9, No.5, September - October 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse314952020.pdf
https://doi.org/10.30534/ijatcse/2020/314952020

A Hybrid Spatio-Temporal Network for Cardiac MRI Segmentation

Surayya Ado Bala [1], Shri Kant [2]
[1]Department of Computer Science and Engineering, Sharda University, India, surayyaadob@gmail.com
[2]Research and Technology Development Centre, Sharda University, India, shrikant.ojha@gmail.com

ABSTRACT

In cardiac magnetic resonance imaging (CMR) analysis, Spatio-temporal consistency in segmentation is desirable for estimating cardiac motion, which often characterizes cardiac pathologies and reliably calculates the ejection fraction of the left and right ventricle. In this paper, we combine the convolutional neural network (CNN) and a recurrent neural network (RNN) to propose a hybrid Spatio-temporal (HST-Net) network. Our network modifies U-Net by adding an improved GRU called the Bidirectional Gated Recurrent Unit (BGRU). On the ACDC dataset, the network was trained and assessed. The network performance can be evaluated using the Dice score and Hausdorff distance to accurately delineate the myocardium (MYO), right ventricle (RV) and left ventricle (LV), and myocardium (MYO) from end-systole (ES) and end-diastole (ED) cycles. The proposed network hits high efficiency from ED and ES on LV, RV, and MYO by capturing image data flow from past, present, and subsequent 2D MRI sequence frames. It achieves a 0.970, 0.951, and 0.933 Dice score for LV, RV, and MYO at ED and 0.939, 0.930, 0.920 for LV, RV, and MYO at ES. The Hausdorff distance is 4.400 mm, 4.173 mm, and 4.310 mm at ED for LV, RV, and MYO and 6.500 mm, 6.976 mm, and 5.814 mm at ES for LV, RV, and MYO. Compared with previous related work, the proposed network shows a promising result.

Key words: Cardiac MRI Segmentation, Convolutional Neural Network. Recurrent Neural Network, U-Net, Spatio-temporal consistency.

1. INTRODUCTION

Segmentation of the heart's essential structures from cardiac MRI is one of the critical tasks, the most crucial prerequisite for cardiac MRI analysis. Accurate LV, RV, MYO delineation is needed to assess parameters of cardiac such as fraction of ejection, the volume of stroke, volume of right and left ventricular, left ventricular mass, and thickness of the myocardium [1]. Initially, expert radiologists manually delineate cardiac MRI, which is rather monotonous, laborious, and vulnerable to intra and inter-observer variability [2]. Hence, an accurate, reliable, automated segmentation model is required to facilitate cardiovascular disease diagnosis.

Cardiac MRI segmentation has various difficulties, as identified by [3][4],[5]. The problems identified include first, cardiac MRI variability from a different institution, scanners, and patient. Second, extreme imbalance in the number of pixels belonging to the object class and background. Third, Spatio-temporal consistency in segmentation. To tackle these difficulties, many segmentation models have been proposed, including the graph-based model (graph cut) [6], deformable models[7],atlas-based model [8], active shape and appearance model[9]. However, still, there is a need for improvement to be usable in the clinical setting. These models relied totally on handcrafted engineering features that required expertise and domain knowledge. Thus, handcrafted features have restricted emblematic abilities to deal with the significant disparity in the heart structure's appearance and shape. Deep learning (DL) based models, including convolutional neural network (CNN) [10] and it is variant such as U-Net [11], have been utilized to search for more robust features to overcome these limitations. U-Net is a 2 dimensional CNN model develop based on a fully convolutional network (FCN) [12] by introducing a skip connection between two paths, which are the path of the encoder (analysis) and the path of the decoder (synthesis) to merge the features. It is widely adopted in medical image segmentation[13] [14] due to its adaptability to small data set and efficiency.

Despite the successes achieved by U-Net architecture in segmenting cardiac MRI, it fails to capture the temporal information (features) between the sequences of the cardiac MRI. It is also essential to evaluate the cardiac motion, cardiac pathologies description [15] and accurately determine the left and right ventricle's ejection fraction. A model needs to capture both spatial and temporal features for a highly accurate multi-slice segmentation of the cardiac MRI.

Recently, studies show that the recurrent neural network (RNN) and its variant, such as GRU, can capture temporal information; they help the network with recurrent information feedback loops[16]. RNN is benevolent to feed-forward neural networks with a persistent hidden state, triggered at a given time by previous states. Inspired by the fact that RNN can dynamically model the contextual information, it handles the variable-length sequences and learns dependencies over longer intervals efficiently [17].

This paper combines CNN and RNN to propose a hybrid Spatio-temporal network (HST-Net) for fully automatic segmentation of cardiac MRI with Spatio-temporal consistency by extracting the spatial and temporal features from the cardiac MRI sequence. Our network modifies U-Net[11] by adding an improved GRU called the BGRU. To capture temporal information flow from past, current, and next 2D MRI sequence frames and use time distribution to exploit inter and intra-slice feature representation.

To sum up, our main contribution is summarized as follows:

- We propose an end-to-end and lightweight cardiac MRI segmentation network called HST-Net by combining CNN and RNN to leverage the slices' spatial information and the temporal information that was not exploited by other CNN-based models.
- We proposed a ROINet for extracting the region of interest to reduce the HST-Net training computational cost.
- We conduct ablation studies to evaluate BGRU's contribution to our proposed HST-Net's overall results.

The remaining part of the paper is as follows: Section 2 highlights recent cardiac MRI segmentation methods. The proposed HST-Net is presented in section 3. We explained the experimental setup in section 4. The experimental results and discussion are described in section 5.Finally; we conclude in section 6.

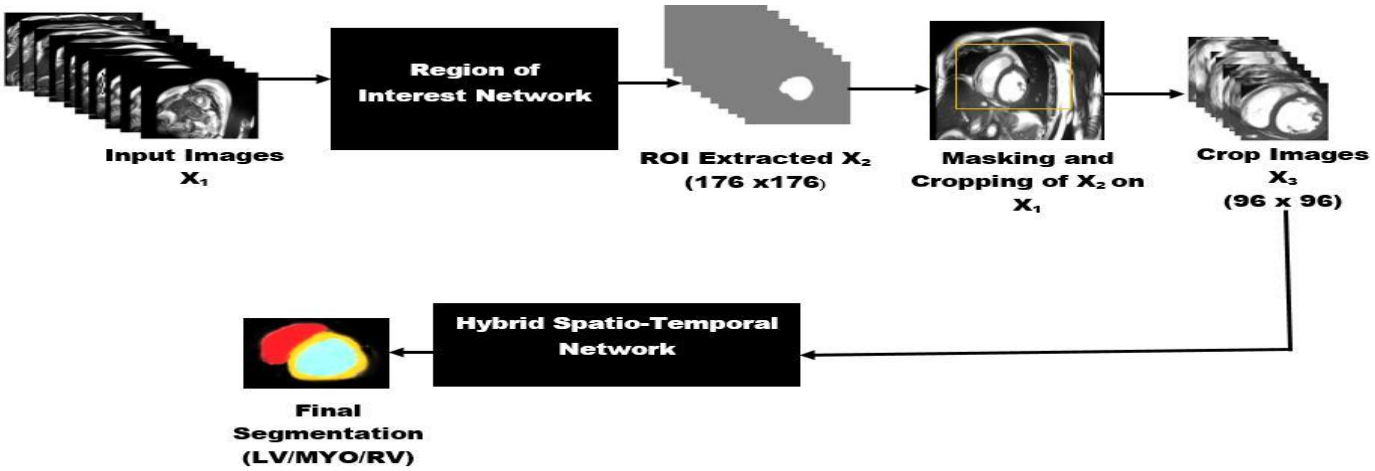

Figure 1: A framework of our proposed method

2. Related Work

Nowadays, segmentation of the state-of-the-art cardiac MRIs is built on deep learning models that ultimately supersede existing non-deep learning models. Several researchers have recently used deep convolutional architecture with hierarchical learning capability [18] in cardiac MRI segmentation [19], most U-Net-like networks. In their review, Bernard et al. [19] reported that nine of the ten participants involved in the ACDC cardiac MRI assessment challenge implemented deep convolutional architecture, most of whom are U-Net and its 3D extensions [14]. Isensee et al. conducted an ensemble of 2D and 3D U-Net architectures with layers of upsampling and residual connectivity[20]. 2D and 3D U-Nets evaluated by Baumgartner et al. [21], and their 2-dimensional U-Net performs better with cross-entropy loss. Patravali et al. [22] present 2D U-Net and 3D U-Net with different dice and cross-entropy loss; 2D U-Net architecture performed better. Yang et al. [23], however, replace their 3D U-Net's skip connection with the residual connection.

Similarly, Jang et al. [24] modified the U-Net and introduced M-Net; they connected it to previous layers. Khened et al. [25] had built a dense U-Net. Instead of the basic convolutional block and inception layer, they use dense blocks in their network's first layer. Rohe et al.[26] created an SVF-Net using a multi-atlas algorithm to capture a target image with all training dataset images. Zotti et al.[27] introduced the Grid-Net architecture that, along with the skip connections, resembles U-Net with convolutional layers.

Zhang Q-Zhang Q et al.[28] using 2D U-Net as well as 3D U-Net architectures introduced 3D consistent and robust spatial precision segmentation accuracy.

3. Proposed Method

Figure 1 shows the structure of our proposed method. The method was conducted in two stages, including

I) The extraction of a region of interest (ROI) using a region of interest (RIONet). II) The HST-Net segmentation. The

ROINet used a stack of 2 dimensional MRI size X_1 images (176 x 176) as input and output the extracted 176 x 176 ROI image X_2. We masked the ROI (X_2) on X_1 after ROI extraction and then cropped X_3 with size (96 x 96), as shown in Figure 1. The cropped X_3 is further piped to the HST-Net as input for the final segmentation of LV, RV, and MYO.

3.1 Region of Interest Extraction

In this stage, we extract the heart, which is the region of interest (ROI) similar to the work of [10] using ROINet from the CMR images. ROI extraction has the following advantages; 1) It removes the unwanted part considering CMR images contain the heart, lungs, and diaphragm; this reduces memory consumption. 2) It reduces the time taken to train the segmentation network and increases the segmentation accuracy. The ROINet used stack 2D size X_1 images (176 x 176) as input and output of extracted ROI images X2 with the size 176 x 176 as binary segmentation, 0 as background 1 for the heart (LV, RV, and MYO union). ROINet, as shown in Figure 2, is the U-Net version. It consists of 3 x 3 convolutions, batch normalizations (BN)[29], and exponential linear units (ELUs)[30] as building blocks of the encoder and decoder paths with 2 x 2 max-pooling and 2 x 2 upsampling, a bottleneck, and skip connection amid the encoder and the decoder route. A sigmoid function is used to output pixel-wise probabilities.

convolution layer in the bottleneck with BGRU. As illustrated in Figure 3, our network has three main blocks: encoding path, bottleneck, and decoding path. The encoding path is the spatial extraction component consisting of three convolution layers with two (2x2) convolutions, followed by an exponential logic unit (ELU) and batch normalization (BN), then a (2x2) max pooling with a downsampling stride of two. The bottleneck is the temporal feature extraction component that consists of BGRU. The decoding path is the segmentation component that consists of three convolution layers with two (3x3) convolution followed by ELU and BN with upsampling, then 1x1 convolution, and a softmax classifier to map the feature maps to classes. The features extracted from the encoder are compressed features that contain the global context, which is pass to the bottleneck to extract the temporal information. Meanwhile, the decoder performs segmentation with the concatenated features from the skip connection and the bottleneck, as illustrated in Figure
3. We further discussed BGRU in section 3.2.1.

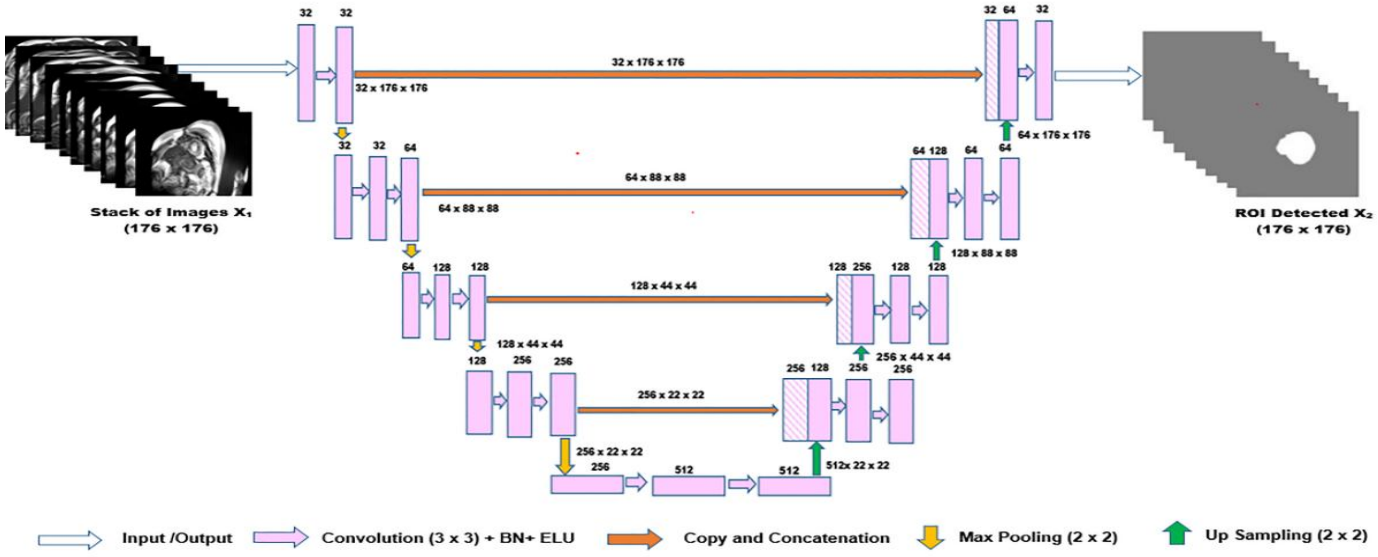

Figure 2: The Structure of ROINet

3.2 Hybrid Spatio-temporal Network

Hybrid Spatio-temporal network (HST-Net) is an end to end cardiac MRI segmentation model design based upon U-Net architecture. We modify U-Net architecture by replacing the

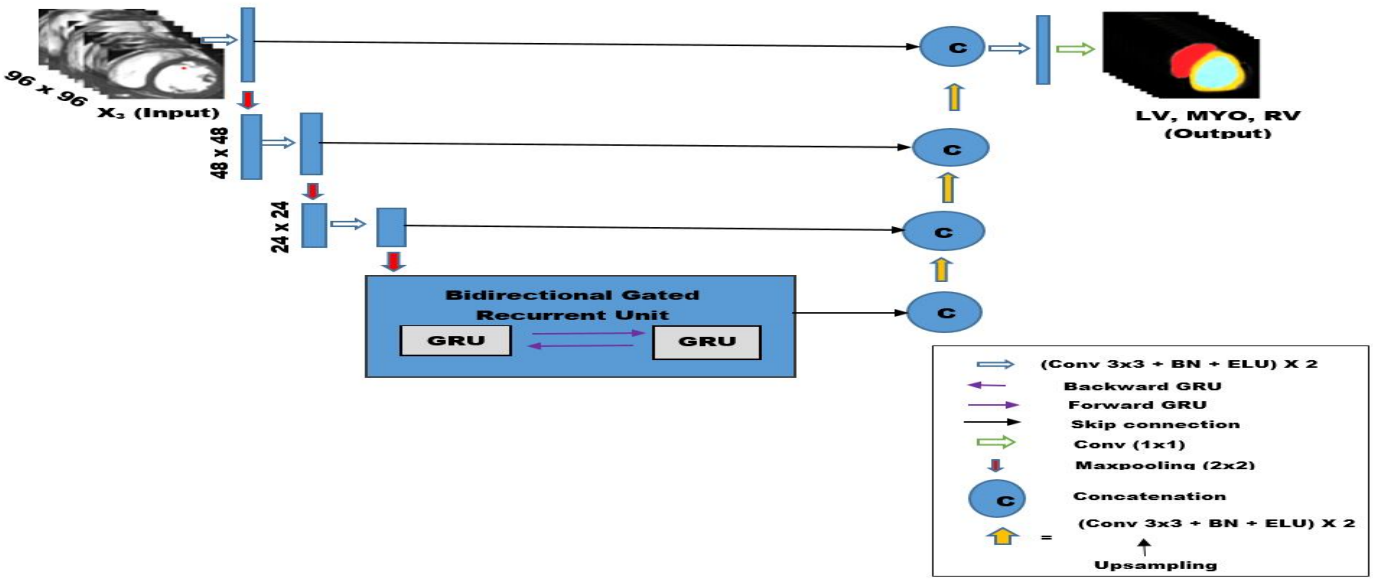

Figure 3: Our proposed Hybrid Spatio-temporal Network

3.2.1 Bidirectional Gated Recurrent Unit

Bidirectional gated recurrent unit (BGRU) is a significant factor in processing sequential data and enhancing the Spatio-temporal consistency in segmentation. BGRU, as shown in Figure 3a, consists of two GRUs in the opposite direction, allowing the network model to learn more features both forward and backward hidden layers. GRU[31] is an RNN variant that can process sequential data by storing previous inputs in an internal network state. They are proposed to support RNN with a skillful model complexity that needs less parameter to generalize and requires less computational power compared to long-term memory (LSTM)[30][32]. Similar to LSTM, GRU uses the hidden state to transfer information. This hidden state (h_t) is the memory unit, which holds the information of the previous information the network has seen before. GRU has two gates, the reset gate r_t, which regulates the integration of new input with the previous memory (h_{t-1}), and the update gates (z_t), which perform two functions as forgot gate and input forget. Forget gate determines what data to keep or throw away. The information from the previous h_t and the current input (x_t) is passed via the sigmoid activation function (σ)[33], outputting values amid 0 and 1; those closer to 0 are discarded whereas those closer to 1 are detained.

Meanwhile, in the input gate, we update the current state (c_t), move the previous h_t, and x_t to σ, which determines which values to change by changing values between 0 and 1.0 means not significant while 1 means were significant. The h_t and the c_t are passed through the tanh activation function. (tanh) [34] to squeeze the values between -1 and 1 to adjust the network. Mathematically represented as follows;

$$z_t = \sigma(W^z x_t + V^z h_{t-1} + b^z) \quad (1)$$

$$r_t = \sigma(W^r x_t + V^r h_{t-1} + b^r) \quad (2)$$

$$c_t = tanh(W^c x_t + V^c(r_t \odot h_{t-1})) \quad (3)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot c_t \quad (4)$$

The model parameters are W, V, and b, and Θ denote the element-wise product. The BGRU is integrated into the bottleneck of our network to access the data in both forward and backward directions with two different hidden layers to properly scrutinize the previous and future data in each state to capture the representation of the inter and intra slice function. The results of both directions are concatenated to the decoder As follows;

$$h_T = h_{tf} \odot h_{tb} \quad (5)$$

Where h_{tf} and h_{tb} denote forward and backward process, respectively. The hidden vector is updated as follows;

$$h_{tf} = \mathbb{H}(x_{tf}, h_{tf-1}) \quad (6)$$

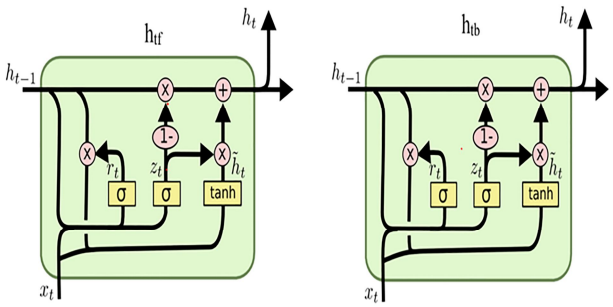

Figure 3a: The structure of the Forward (h_{tf}) and the backward (h_{tb}) Gated Recurrent Unit

$$h_{tf} = \mathbb{H}\,(x_{tf}, h_{tf+1}) \qquad (7)$$

 represents a non-linear differentiable transformation function.

4. EXPERIMENTAL SETUP

The experimental setup is presented in this section, including the preparation of the dataset, evaluation metrics, and details for the implementation.

4.1 Datasets

We use the "Automated Challenge for Cardiac Diagnosis (ACDC) dataset. The ACDC dataset is a 3D cine MRI acqu ired at the University of Dijon (France) using 1.5 T and3Tsc anners over six years in a clinical routine.The training datas et consisted of 100 patients with 20 patients with their corre sponding manual LV, RV, and MYO labels, segmented by medical experts from ED as well as ES cycle from 5 classes, including 1) Normal (NOR), 2) Patients with past Myocardial Infraction (MINF), 3) Dilated cardiomyopat hy patients (DCM), 4) Hypertrophic cardiomyopathy (HCM) patients, 5) Abnormal Right Ventricle (ARV) patients". In this work, the training, validation, and testing data are in the ratio of 80:10:10.80 patients for training, randomly subdivided.

4.1.1 Preprocessing of Dataset

The dataset was acquired in clinical routine [19] that prone to artifacts, including intensity inhomogeneity, patient movement, MRI low-frequency intensity fluctuation, the variability of the intensity ranges, and contrast these leads to discrepancies in the image analysis. Additionally, multi-slice cine MR with different protocols and hardware is used for image acquisition. Therefore, preprocessing is required to convey all data to the same distributions. Our training data was prepared by extracting 2D MR slices from all patient 3D cine MR sequences for both ED and ES at each time frame of the cine sequences.

Since the cardiac MRI has a different size, we first resized the 2D CMR image slices of both ED and ES images X of size (731 x 447) to X_1 of size (176 x 176). We applied a bias field correction on the datasets to correct the strength non-uniformity using N4ITK [35]. Then, using the z-score [36], each 2D cardiac slice is normalized. Z-score is defined as follows:

$$X_{norm} = \frac{X - \mu[X]}{\sigma[X]} \qquad (9)$$

Where, respectively, μ and σ denote the mean and the standard deviation of image X.

4.2 Evaluation Metrics

Using the Dice Score (Dice) and Hausdorff distance (HD) [2], we evaluated our work. The Dice score is a measure of the overlap between a method's extracted segmented volumes P and the corresponding annotated volume Q (ground truth). Mathematically represented as;

$$Dice\,(P,Q) = \frac{2|P \wedge Q|}{(|P|+|Q|)} \qquad (10)$$

Hausdorff distance (HD) calculates the distance between Prediction (P) and Ground Truth (G) segmentation. Described mathematically as;

$$Haus(P,G) = \max\{Dist(P,G), Dist(G,P)\} \qquad (11)$$

4.3 Implementation Details

Our model was executed with a tensorflow library [38] in keras [37]. We performed cross-validation with 5-fold and selected 80 patients randomly for model training, 10 patients for validation, then 10 patients for testing. For the duration of the training, random rotation, translation, rescaling, and flipping (horizontal and vertical) are used to augment our data. The model used Adam optimizer [39] to minimize loss (cross-entropy) features. He normal initializer [36][40] was used to initialize the network weights and learning rate of 0.001. The training batch includes 16 ROI-cropped 2D images 96 x 96. It took us 12 hours in training our network for a total of 200 epochs on NVidia GeForce GTX1080Ti GPU, developed on 64-bit windows Intel Core(TM) i7-7700HQCPY@2.80GHZ-2.81GHZ. In all the layers, exponential linear units (ELUs) are used as the activation function apart from the output layer that applies softmax. To overcome the problem of difficulty in training and vanishing gradient, we use batch normalization to normalize the feature maps[45] and stabilize the network while training.

Table 1: Parameters and hyperparameters used in our proposed model

Stage	Parameters/ Hyper parameters	Value
Initialization of network	Bias Weight	0 He normal
Adam Optimizer	β_1 β_2 α epsilon	0.9 0.999 0.001 10^{-8}
Training	Learning Rate Batch Size Epoch Momentum Decay	0.001 16 200 0.9 0.9

The details of the parameters and the hyperparameters used are in Table 1.

5. EXPERIMENTAL RESULTS AND DISCUSSION

We validated our HST-Net performance on ACDC datasets and conducted various further experiments to test our proposed model's efficacy.

5.1 Ablation Study

To understand the importance and contribution of GRU and BGRU on our HST-Net, we conduct an ablation analysis. We investigate whether our proposed HST-Net is essential for the extraction of Spatio-temporal features and in improving segmentation accuracy.

Table 3: Segmentation results of the myocardium (MYO), right ventricle (RV) and, left ventricle (LV) in end-systole (ES) and end-diastole (ED)performance based on the Hausdorff distance of our proposed methods and U-Net baseline

Method	Hausdorff Distance					
	LV		RV		MYO	
	ED	ES	ED	ES	ED	ES
U-Net (Baseline)	10.951	11.290	12.210	12.573	9.601	12.134
U-Net + GRU	4.824	7.570	6.205	8.100	5.482	7.852
U-Net + BGRU	4.400	6.500	4.173	6.976	4.310	5.814

Table 2: The segmentation output of myocardium (MYO), right ventricle (RV) and, left ventricle (LV) in end-systole (ES) and end-diastole (ED) based on the dice score of our proposed methods and U-Net baseline

Method	Dice Score					
	LV		RV		MYO	
	ED	ES	ED	ES	ED	ES
U-Net (Baseline)	0.940	0.897	0.880	0.793	0.841	0.804
U-Net + GRU	0.956	0.918	0.905	0.827	0.896	0.852
U-Net + BGRU	0.970	0.939	0.951	0.890	0.933	0.920

To test the robustness of HST-Net, we performed an extensive set of experiments to evaluate its effectiveness.
Different experiments based on ACDC datasets were carried out to study how our proposed HST-Net influences the cardiac MRI segmentation. We performed three experiments and their comparisons on Dice and Hausdorff's distance, as shown in Table 2, Table 3, Figure 4, and Figure 5. We use standard U-Net architecture as our baseline in experiment 1. We use U-Net + GRU in experiment 2. Finally, U-Net + BGRU (our HST-Net) achieves the best results in the cardiac segmentation of MRIs in experiment 3. The ablation studies show that our proposed HST-Net helps boost the efficiency and accuracy of cardiac MRI segmentation with Spatio-temporal consistency.

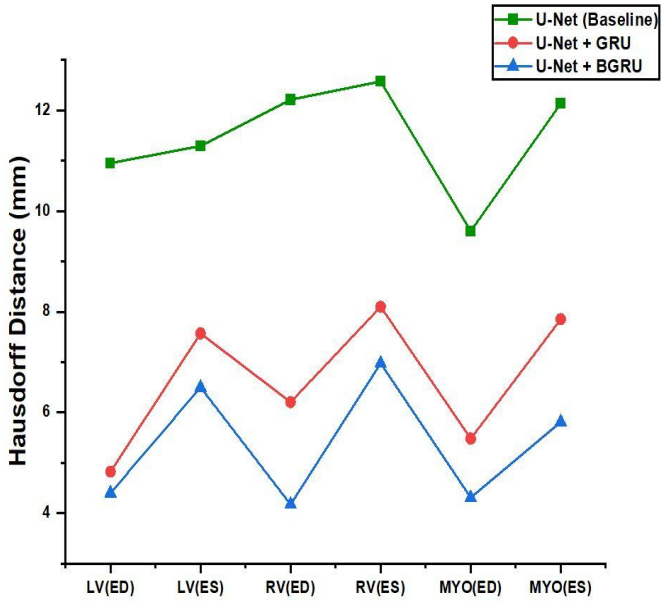

Figure 4: Details of U-Net (Baseline), U-Net + GRU, and U-Net + BGRU (HST-Net) segmentation comparison on the Dice metrics

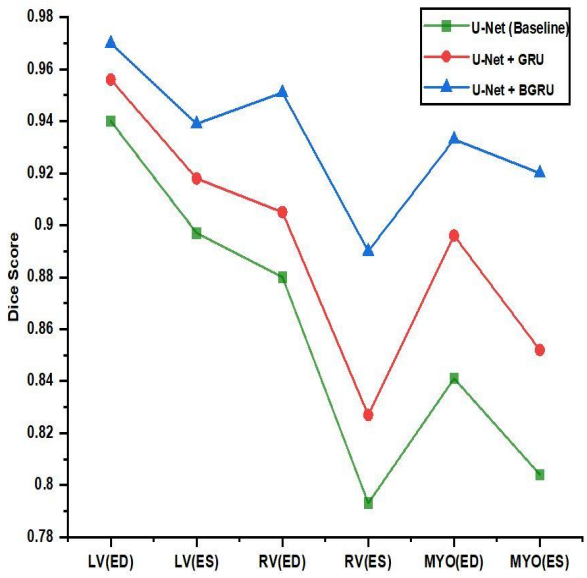

Figure 5: Details of U-Net (Baseline), U-Net + GRU and, U-Net + BGRU (HST-Net) segmentation on the Hausdorff Distance (mm) metrics

5.2 Discussion

As presented in Table 2, Table 3 and depicted in Figure 4, Figure 5, GRU implementation in U-Net increases the Dice score with 1.6%, 2.5%, and 5.5% in LV, RV, and MYO of ED and 2.1%, 3.4%, and 4.8% of LV, RV, and MYO of ES.

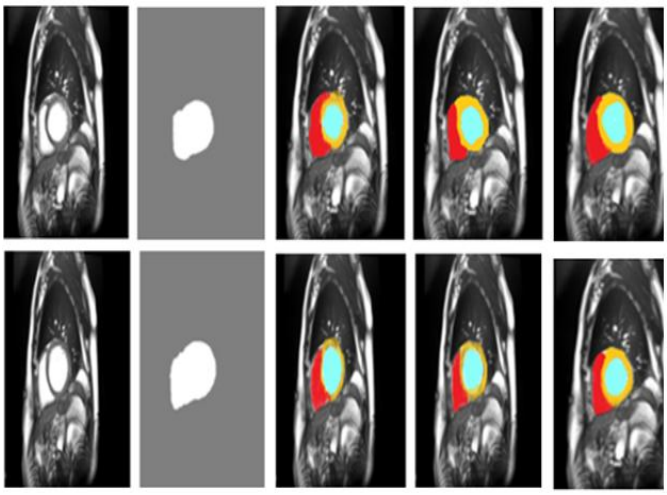

Figure 6: Details of comparison between U-Net (Baseline), U-Net+GRU, and U-Net + BGRU. The first row presents the patient's 033 end-diastole cardiac MRI, and the second row shows the end-systole cardiac MRI of patient 033 from the ACDC dataset. Columns 1-5 in both rows 1 and 2 present original cardiac MRI, ground truth image, U-Net segmented image, U-Net + GRU segmented image, U-Net + BGRU segmented image, respectively. The cyan, red, and yellow color represents the segmented LV, RV, and MYO

Meanwhile, implementing BGRU in the U-Net increases the Dice score with 3%, 7.1%, and 9.2% in LV, RV, and MYO of ED; and 4.2%, 9.7%, and 11.6% of LV, RV, MYO of ES. However, implementing GRU in the U-Net decreases the Hausdorff distance with 6.127mm, 6.005mm, and 4.119mm of LV, RV, and MYO of ED and 3.720mm, 4.473mm, and 4.282mm in LV, RV, and MYO of ES. Likewise, implementing BGRU in the U-Net decreases the Hausdorff distance with 6.551mm, 8.037mm, and 5.291mm in LV, RV, and MYO of ED 4.790mm, 5.597mm, and 6.320mm in LV, RV, and MYO of ES.

Comparing the above results, our U-Net + BGRU (HST-Net) achieves better accuracy than the other methods. However, the decrease in the Hausdorff distance and an increase in the Dice score shows strong evidence that our proposed HST-Net will perform better in segmenting cardiac MRI with spatial-temporal consistency from apex to the heart base. Furthermore, this indicates that the enhanced BGRU can access the temporal information in both forward and backward directions; to adequately scrutinize the previous and future information at each state to capture the inter and intra slice feature representation. Figure 6 visually contrasts the result obtained for cardiac MRI segmentation with the three studies, i.e., U-Net, U-Net + GRU, and U-Net + BGRU.

Compared to the expert annotated ED and ES cardiac cycle, as shown in Table 2 and Figure 4, We achieve the Dice score performance output of 0.933, 0.951, and 0.970 in MYO, RV, and LV at ED and 0.920, 0.890, and 0.939 in MYO, RV, and LV, and MYO at ES. Similarly, we obtain the Hausdorff distance performance of 4.310mm, 4.173mm, and 4.400 mm, in MYO, RV, and LV at ED and 5.814mm,6.976mm and 6.500 mm in MYO, RV, and LV at ES, as shown in Table 3 and Figure 5.As we observed from the result it shows that our method performs better at the ED.

6. CONCLUSION

A new end to end deep learning model called HST-Net is proposed in this paper for automatic segmentation of cardiac MRI with Spatio-temporal consistency. In cooperating BGRU to U-Net architecture, our model not only extracts the Spatio-temporal features but also decreases the Hausdorff distance and raises the Dice score in contrast to the U-Net model. The results from the experiments prove that our model boosts the efficiency of the segmentation of cardiac MRIs. We plan to examine our HST-Net's performance using Semi-supervised Learning on more clinical data such as the UK Biobank.

REFERENCES

1. P. Peng, K. Lekadir, A. Gooya, L. Shao, S. E. Petersen, and A. F. Frangi. **A review of heart chamber segmentation for structural and functional analysis**

using cardiac magnetic resonance imaging, Magn. Reson. Mater. Phys. Biol. Med., vol. 29, no. 2, pp. 155-195, Apr. 2016.

2. C.A Miller, P.Jordan, A.Borg, R.Argyle, D.Clark, K.Pearce, and M.Schmitt. **Quantification of left ventricular indices from SSFP cine imaging: Impact of real-world variability in analysis methodology and utility of geometric modeling**, J. Magn. Reson. Image. vol. 37, no. 5, pp. 1213-1222, 2013.
3. C. Petitjean and J. N. Dacher. **A review of segmentation methods in short axis cardiac MR images**, Med. Image Anal., vol. 15, pp. 169-184, 2011.
4. W. Bai, W. Shi, C. Ledig, and D. Rueckert. **Multi-atlas segmentation with augmented features for cardiac MR images**, Med. Image Anal., vol. 19, no. 1, pp. 98-109, 2015.
5. S.C Mitchell, J.G Bosch, B.P.F Lelieveldt, R.J van der Geest, J.H.C Reiber, and M. Sonka, **3-D Active Appearance Models: Segmentation of Cardiac MR and Ultrasound Images**, IEEE Trans. Medical Imaging, vol. 21, pp. 1167-1178, Sept. 2002.
6. T. Chen, J. Babb, P. Kellman, L. Axel, and D. Kim, **Semi-automated segmentation of myocardial contours for fast strain analysis in cine displacement-encoded MRI**, IEEE Trans. Med. Imag., vol. 27, no. 8, pp. 1084-1094, Aug. 2008.
7. S. Queirós et al., **Fast, automatic myocardial segmentation in 4D cine CMR datasets**, Med. Image Anal., vol. 18, no. 7, pp. 1115-1131, 2014.
8. H. Liu, H. Hu, X. Xu, E. Song, **Automatic left ventricle segmentation in cardiac MRI using topological stable-state thresholding and region-restricted dynamic programming**, Acad. Radiol., vol. 19, no. 6, pp. 723-731, 2012.
9. I. B. Ayed, H.M. Chen, K. Punithakumar, I. Ross and S. Li, **Max-flow segmentation of the left ventricle by recovering subject-specific distributions via a bound of the Bhattacharyya measure**, Med. Image Anal., vol. 16, no. 1, pp. 87-100, 2012.
10. M. Avendi, A. Kheradvar, and H. Jafarkhani, **A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI**, Med. Image Anal., vol. 30, pp. 108-119, 2016.
11. O. Ronneberger, P. Fischer and T. Brox, **U-Net: Convolutional Networks for Biomedical Image Segmentation**, MICCAI, pp. 234-241, 2015.
12. J. Long, E. Shelhamer, and T. Darrell, **Fully convolutional networks for semantic segmentation**, Proc. IEEE Conf. Comput. Vision Pattern Recog., pp. 3431-3440. 2015.
13. F. Milletari, N. Navab, and S. A. Ahmadi, **V-net: Fully convolutional neural networks for volumetric medical image segmentation**, Proc. IEEE 3D Vis., pp. 565-571, Oct. 2016.
14. Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox and O. Ronneberger**, 3D U-net: Learning dense volumetric segmentation from sparse annotation**, Proc. MICCAI, pp. 424-432, 2016.
15. D. P. O' Regan, **Putting machine learning into motion: applications in cardiovascular imaging**, Clinical Radiology, *75*(1), pp.33-37, 2019.
16. Q.Tong et al., **RIANet: Recurrent interleaved attention network for cardiac MRI segmentation**, Computers in biology and medicine, Elsevier, vol.109, pp. 290-302, 2019.
17. A. Jagannatha and H. Yu, **Bidirectional recurrent neural networks for medical event detection in electronic health records**, Proc. conf. Assoc. Computer. Linguistics North American Chapter Meeting, pp. 473-482, 2016.
18. G. Litjens et al**., A survey on deep learning in medical image analysis**, Medical image analysis, vol.42, pp.60-88, 2017.
19. O. Bernard et al., **Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved**?, IEEE Trans. Med. Imag., vol. 37, no. 11, pp. 2514-2525, Nov. 2018.
20. F. Isensee, P. Jaeger, P. M. Full, I. Wolf, S. Engelhardt, and K. H. Maier-Hein, **Automatic cardiac disease assessment on cine-MRI via time-series segmentation and domain specific features**, Proc. STACOM-MICCAI, pp. 120-129, 2017.
21. C. F. Baumgartner, L. M. Koch, M. Pollefeys, and E. Konukoglu**, An exploration of 2D and 3D deep learning techniques for cardiac MR image segmentation**, Proc. STACOM-MICCAI, pp. 111-119, 2017.
22. J. Patravali, S. Jain, and S. Chilamkurthy, **2D–3D fully convolutional neural networks for cardiac MR segmentation**, Proc. STACOM MICCAI, pp. 130-139, 2017.
23. X. Yang, C. Bian, L. Yu, D. Ni, and P.A. Heng, **Class-balanced deep neural network for automatic ventricular structure segmentation**, Proc. STACOM MICCAI, pp. 152-160, 2017.
24. Y. Jang, Y. Hong, S. Ha, S. Kim, and H.-J. Chang, **Automatic segmentation of LV and RV in cardiac MRI**, Proc. STACOM MICCAI, pp. 161-169, 2017.
25. M. Khened, V. Alex, and G. Krishnamurthi, **Densely connected fully convolutional network for short-axis cardiac cine MR image segmentation and heart diagnosis using random forest**, Proc. STACOM MICCAI, pp. 140-151, 2017.
26. M.-M. Rohe, M. Sermesant and X. Pennec, **Automatic multi-atlas segmentation of myocardium with SVF-net**, Proc. STACOM MICCAI, pp. 170-177, 2017.
27. C. Zotti, Z. Luo, O. Humbert, A. Lalande, and P.-M. Jodoin, **GridNet with automatic shape prior registration for automatic MRI cardiac segmentation**, Proc. STACOM MICCAI, pp. 73-81, 2017.

28. Q. Zheng, N. Delingette, H. Duchateau, and N. Ayache, **3-D consistent and robust segmentation of cardiac images by deep learning with spatial propagation**, IEEE Trans. Med. Imag., vol. 37, no. 9, pp. 2137-2148, Sep. 2018.
29. S. Ioffe and C. Szegedy, **Batch normalization: Accelerating deep network training by reducing internal covariate shift**, in ICML, 2015.
30. D.-A. Clevert, T. Unterthiner, and S. Hochreiter, **Fast and accurate deep network learning by exponential linear units (ELUs)**, arXiv preprint arXiv:1511.07289, 2015.
31. J. Chung, C. Gulcehre, K. Cho, Y. Bengio, **Empirical evaluation of gated recurrent neural networks on sequence modeling**, Proc. NIPS Workshop on Deep Learning, 2014.
32. R. Zhao, R. Yan, J. Wang, and K. Mao, **Learning to monitor machine health with convolutional bi-directional LSTM networks**, Sensors, vol. 17, no. 2, pp. 273-290, 2017.
33. J. Han and C. Moraga, **The influence of the sigmoid function parameters on the speed of backpropagation learning** International Workshop on Artificial Neural Networks. Springer, Berlin, Heidelberg, 1995.
34. A. Namin, K. Leboeuf, R. Muscedere, H. Wu and M. Ahmadi, **Efficient hardware implementation of the hyperbolic tangent sigmoid function**, Proc. IEEE Int. Symp. Circuits Syst., pp. 2117-2120, 2009.
35. NJ. Tustison et al., **N4ITK: improved N3 bias correction**, Medical Imaging IEEE Transactions, vol. 29, no. 6, pp. 1310-1320, 2010.
36. S. G. K. Patro and K. Kumar, **Normalization: A Preprocessing Stage**, 2015.
37. François Chollet, **Keras: The python deep learning library**, [online] Available: https://keras.io/.
38. M. Abadi et al., **Tensorflow: Large-scale machine learning on heterogeneous systems**, arXiv: 1603.04467, Nov. 2015.
39. D. P. Kingma and J. L. Ba, **Adam: A method for stochastic optimization**, Int. Conf. Learn. Represent. pp. 1-13, 2015.
40. K. He, X. Zhang, S. Ren, and J. Sun, **Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification,** Proc. IEEE Int. Conf. Computer. Vis., pp. 1026-1034, 2015.