Fully automated deep learning based segmentation of normal, infarcted and edema regions from multiple cardiac MRI sequences

Xiaoran Zhang[1], Michelle Noga[2,3], Kumaradevan Punithakumar[2,3]

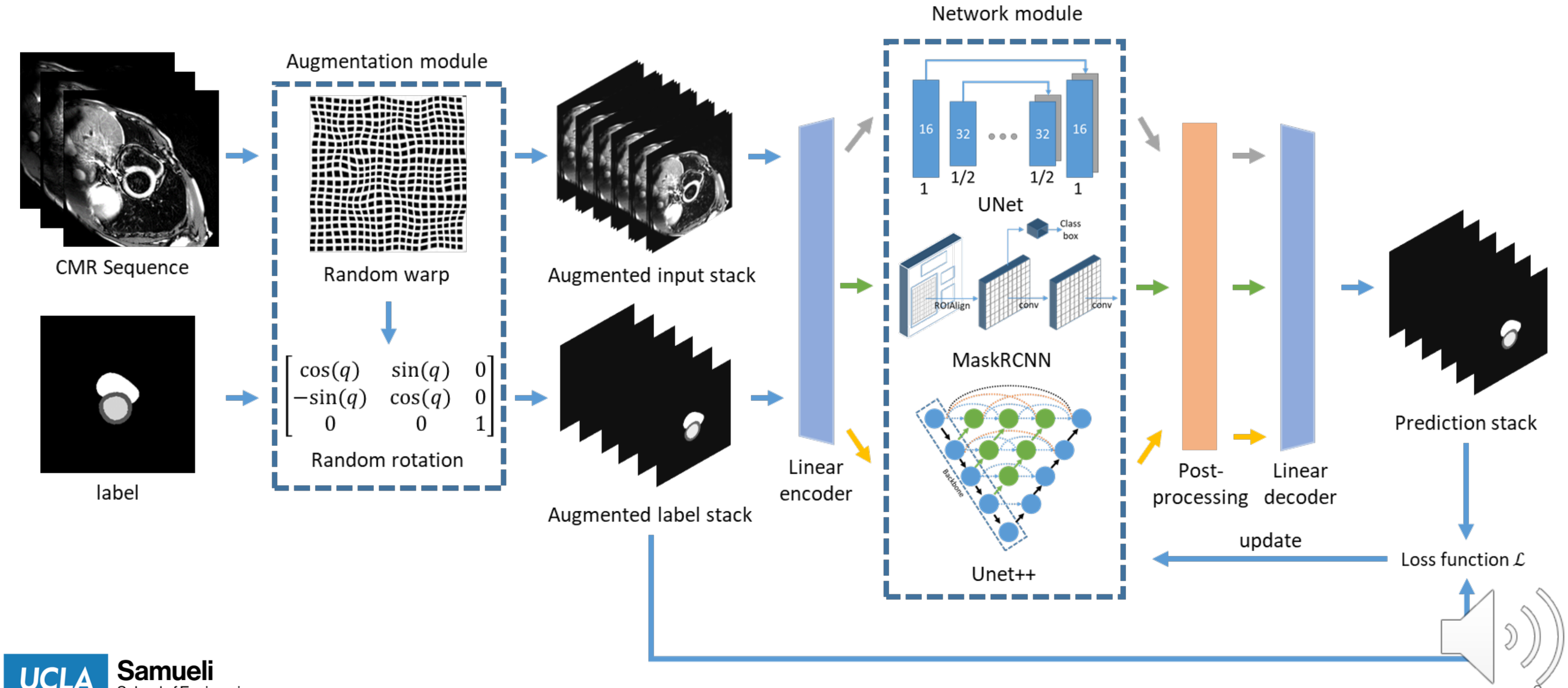
[1] Department of Electrical and Computer Engineering, UCLA
[2] Department of Radiology & Diagnostic Imaging, University of Alberta
[3] Servier Virtual Cardiac Centre, Mazankowski Alberta Heart Institute

Introduction

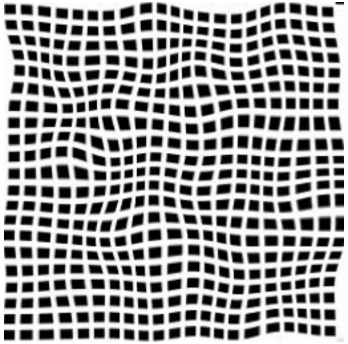
- Late gadolinium-enhanced (LGE) imaging is a commonly used cardiac magnetic resonance (CMR) sequence to diagnose myocardial infarction and acute injury can be detected using T2-weighted CMR.
- Accurately detecting ventricular boundaries using the LGE or T2-weighted images is challenging.
- Utilizing the combination of CMR sequences will allow a more robust and accurate diagnostic information for myocardial infarction [1].
- This study proposes a fully automated approach by utilizing deep convolutional neural networks to delineate left ventricle (LV) blood pool (BP), right ventricle (RV) BP, LV normal myocardium (NM), LV myocardial edema (ME) and LV myocardial scars (MS).

[1] Zhuang, Xiahai. "Multivariate mixture model for myocardial segmentation combining multi-source images." IEEE Transactions on Pattern Analysis and Machine Intelligence 41.12 (2018): 2933-2946.

Overall Architecture



Augmentation Module



Random warp

$$\begin{bmatrix} \cos(q) & \sin(q) & 0 \\ -\sin(q) & \cos(q) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Random rotation

- Random warping is performed by initially generating a 8x8x2 uniformly distributed random matrix, with each entry in range [-5,5]. We then resize the matrix to the image size and apply bilinear interpolation.

- We utilize random rotation in 90, 180, 270 degree with equal probability.

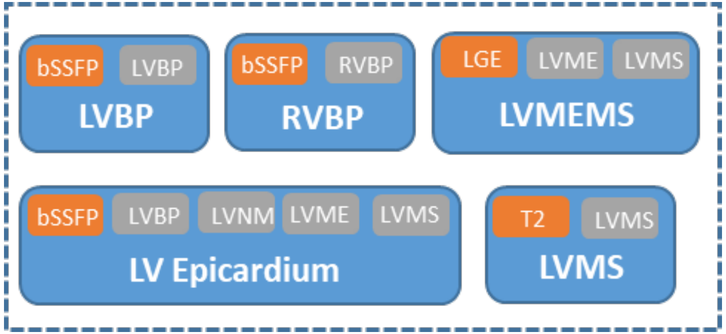
UCLA Samueli School of Engineering

Preprocessing and Linear Encoder

Preprocessing:

- All training and validation images are normalized using 5th and 95th percentile values of the intensity distribution of the extracted 2D slices.

$$I_n = \frac{I - I_{05}}{I_{95} - I_{05}}$$

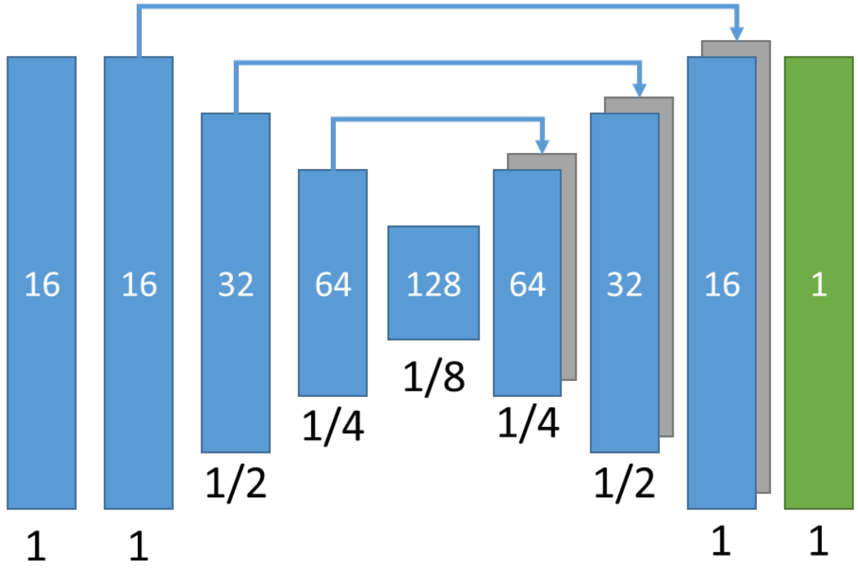


Linear encoder:

- We encode the augmented input stack after preprocessing and produce five input blocks instead of blindly concatenating the CMR sequences.

UCLA Samueli School of Engineering

Network Module



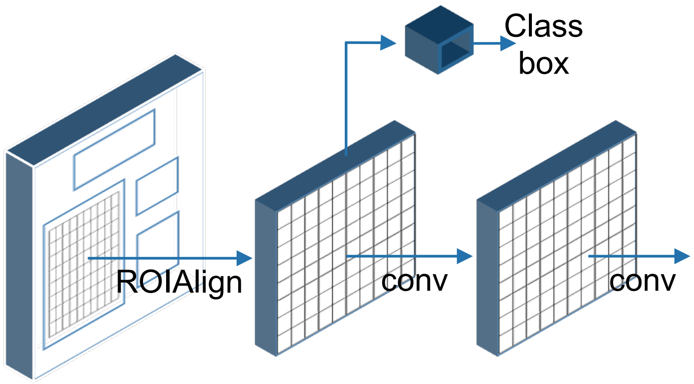
U-net:

- We utilize a shallow version of the standard U-net [2].
- The U-net is trained on all five input blocks produced by the linear encoder.
- The loss function of the U-net is selected as the negative of dice.

UCLA Samueli School of Engineering

[2] Ronneberger, Olaf, et al. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2015.

Network Module



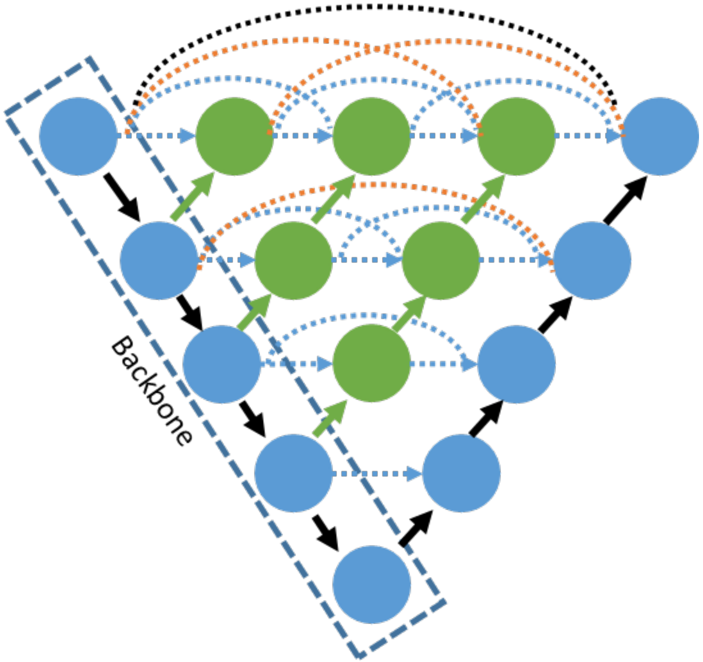
Mask-RCNN:

- We utilize ResNet 50 as the backbone for the Mask-RCNN [3] in Matterport library [4].
- The Mask-RCNN is trained on the LVMEMS and LVMS blocks using Adam optimizer.

UCLA Samueli School of Engineering

[3] He, Kaiming, et al. "Mask R-CNN." Proceedings of the IEEE International Conference on Computer Vision. 2017.
[4] Abdulla, Waleed. Mask R-CNN for object detection and instance segmentation on keras and tensorflow. https://github.com/matterport/Mask_RCNN (2017)

Network Module



U-net++:

- We utilize VGG 16 as the backbone for the U-net++ [5].
- The U-net++ is trained on the LVMEMS and LVMS blocks using Adam optimizer.
- The loss function of the U-net++ is selected as the negative of dice.

[5] Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." IEEE Transactions on Medical Imaging 39.6 (2019): 1856-1867.

Post-processing and Linear Decoder

Post-processing:

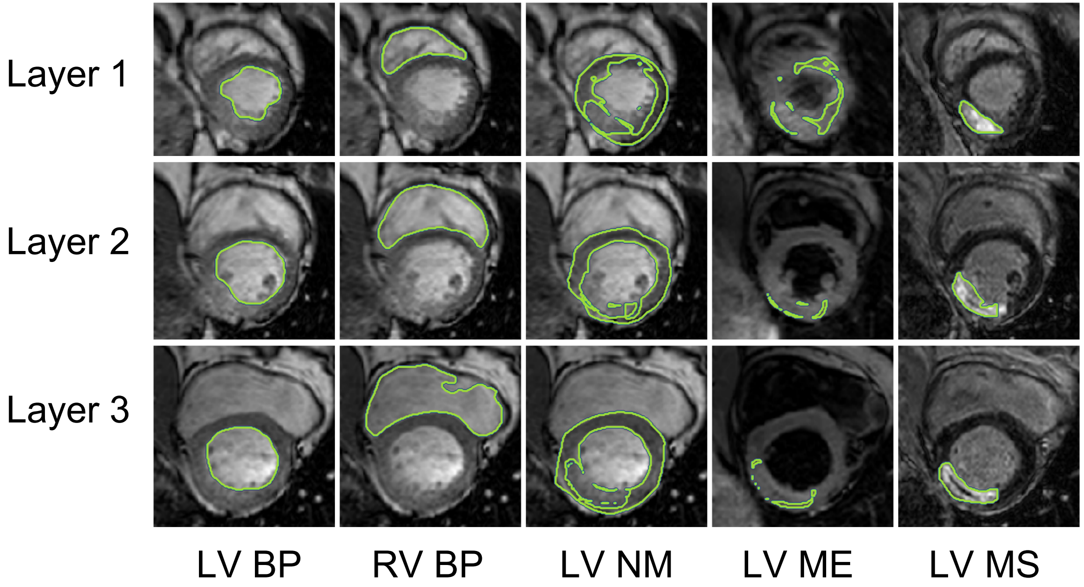
- We applied post-processing to retain only the largest connected component for the predictions of LV BP and LV epicardium by U-net.
- We applied an operation to remove holes that appear inside the foreground masks.

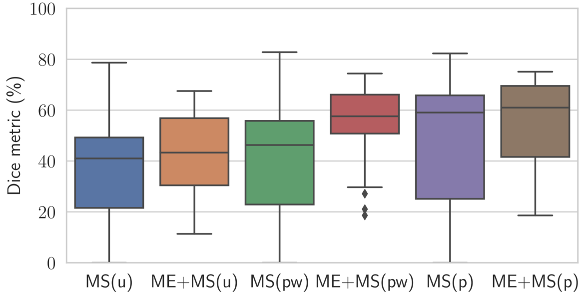
Linear decoder:

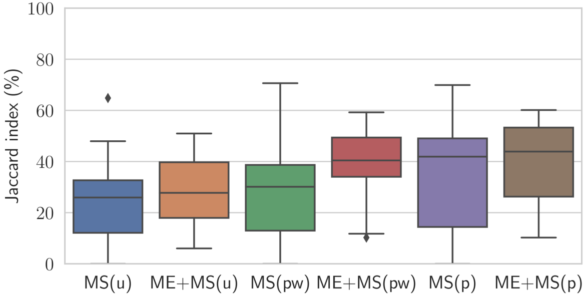
- We perform linear subtraction on the predicted masks to extract each corresponding class.
- We perform a binary constraint for LV ME and LV MS by applying myocardium mask.

$$P_i = \begin{cases} \sigma\left(\widetilde{P_0}\right) & i = 0 \\ \sigma\left(\widetilde{P_1}\right) & i = 1 \\ \sigma\left(\widetilde{P_2} - \widetilde{P_0} - \widetilde{P_3}\right) & i = 2 \\ \sigma\left(\widetilde{P_2} - \widetilde{P_0}\right) \circ \sigma\left(\widetilde{P_3} - \widetilde{P_4}\right) & i = 3 \\ \sigma\left(\widetilde{P_2} - \widetilde{P_0}\right) \circ \sigma\left(\widetilde{P_4}\right) & i = 4 \end{cases}$$

Results







Methods	Dice metric (%)		Jaccard index(%)	
	MS	ME+MS	MS	ME+MS
UNet	36.2 ± 23.2	43.2 ± 16.0	24.5 ± 18.1	28.8 ± 13.1
Proposed†	38.5 ± 24.3	54.2 ± 17.1	26.5 ± 18.9	38.9 ± 15.0
Proposed	$\mathbf{46.8 \pm 26.8}$	$\mathbf{55.7 \pm 18.3}$	$\mathbf{34.2 \pm 22.2}$	$\mathbf{40.5 \pm 16.3}$

† indicates without the linear encoder and decoder module.

UCLA Samueli School of Engineering

Conclusion

- We propose a fully automated approach to segment LV ME and LV MS from multi-sequence CMR data.
- We introduce an augmentation module to enhance the training set and a linear encoder and decoder along with a network module to improve the segmentation performance.
- The algorithm is trained using the 25 cases provided by the challenge organizer on evaluated another 20 cases which are not included in the training set.
- Our proposed method achieves a overall mean dice metric of 46.8%, 55.7% for LV ME and LV ME+MS delineations respectively.

References

[1] Zhuang, Xiahai. "Multivariate mixture model for myocardial segmentation combining multi-source images." IEEE Transactions on Pattern Analysis and Machine Intelligence 41.12 (2018): 2933-2946.
[2] Ronneberger, Olaf, et al. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2015.
[3] He, Kaiming, et al. "Mask R-CNN." Proceedings of the IEEE International Conference on Computer Vision. 2017.
[4] Abdulla, Waleed. Mask R-CNN for object detection and instance segmentation on keras and tensorflow. https://github.com/matterport/Mask_RCNN (2017)
[5] Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." IEEE Transactions on Medical Imaging 39.6 (2019): 1856-1867.