STACKED BCDU-NET WITH SEMANTIC CMR SYNTHESIS: APPLICATION TO THE MYOCARDIAL PATHOLOGY SEGMENTATION CHALLENGE

Carlos Martín-Isla[1], Maryam Asadi-Aghbolaghi[2], Polyxeni Gkontra[1], Victor M. Campello[1], Sergio Escalera[1,3], Karim Lekadir[1]

1. Departament de Matemàtiques & Informàtica, Universitat de Barcelona, Spain, 2. Institute for Research in Fundamental Sciences (IPM), Iran, 3. Computer Vision Center, Univeritat Autònoma de Barcelona, Spain

MYOPS Challenge

Automatically segment scar tissue and edemic regions from multi-sequence cardiac magnetic resonance images (bSSFP, LGE and T2).

- Training data: 25 patients from Shanghai Renji Hospital with registered and interpolated multi-sequence MRI [1].
- Testing data: 20 unseen patients

Background

- Accurate segmentation of scar tissue and edema from CMR is fundamental to the assessment of the severity of myocardial infarction and viability.
- Automatic segmentation is particularly challenging due to:
  - Variability in texture of infarcted and edemic areas
  - Limited input data to train models
  - Imaging acquisition protocol & inter-observer variability

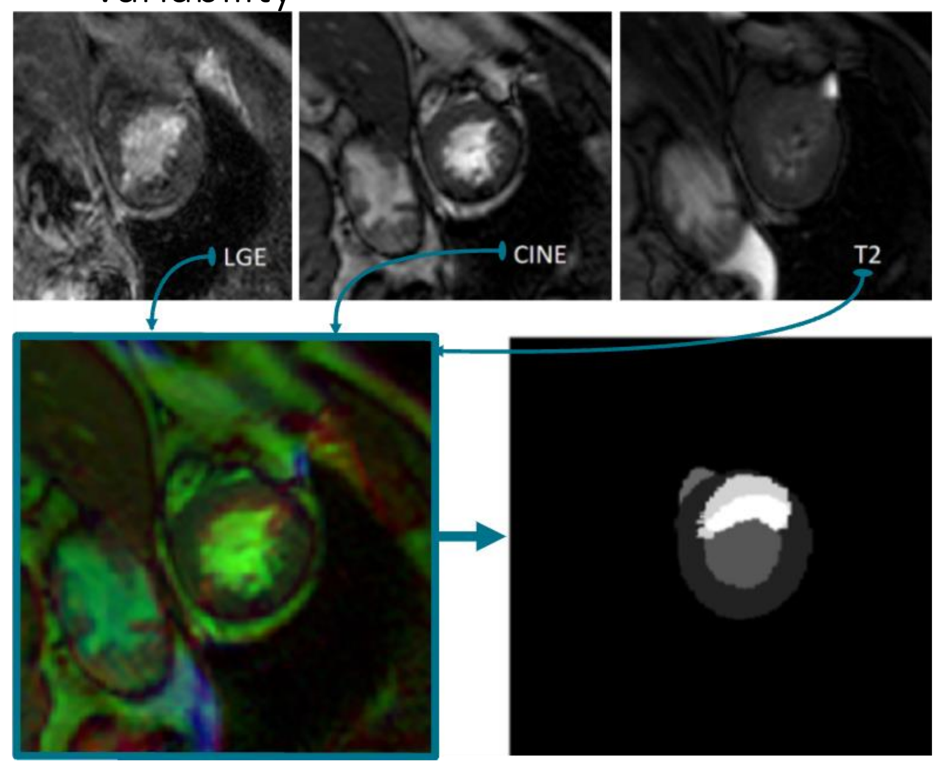


Fig. 1. LGE, bSSFP, T2, their combination, their manual segmentation

The solution

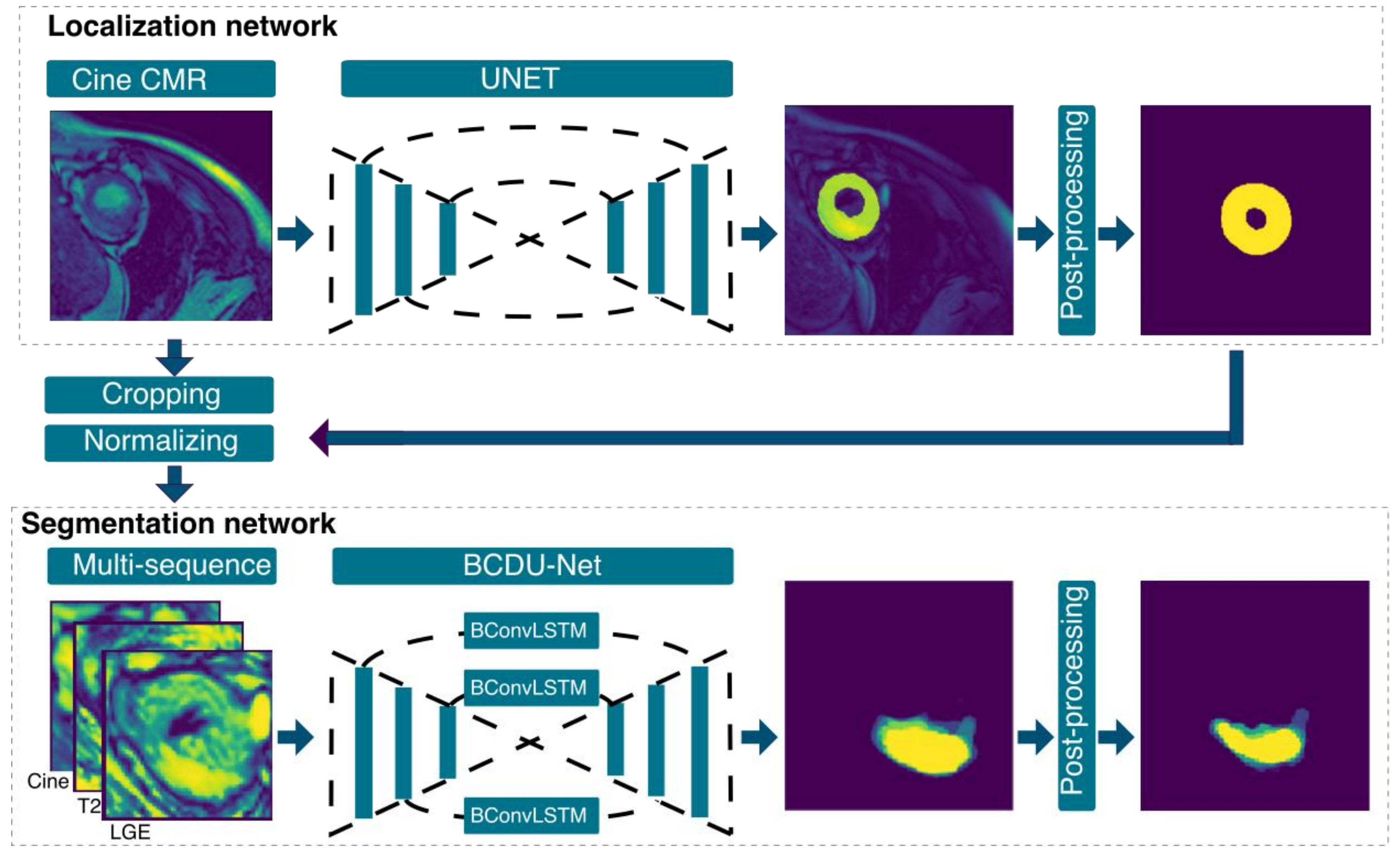
- To address these challenges, we propose:
  - Deep learning stacked BCDU-NET architecture
  - Localisation and segmentation stages
  - Multi-modal Semantic Image Synthesis with Spatially-Adaptive Normalization (SPADE) [2]

Proposed architecture & augmentation strategy

Proposed 2-stage architecture based on BCDU-Net

- STEP 1: LOCALIZATION
  - Binary segmentation network to localize the myocardium
  - Cine-MRI as input modality
  - U-net architecture
- STEP 2: SEGMENTATION
  - BCDU-Net [3] to segment the scar and edema
  - Normalized myocardium of the three input modalities
  - Averaging the confidence maps of an ensemble of 15 models.

Fig. 2. Overview of the proposed stacked localisation – segmentation network.



Data augmentation strategy

- New multi-modality images with multi-styles from altered versions of real annotations
- Morphological operations: 1. contour warpings between pairs of annotations, 2. scar tissue and edema rotations, and 3. dilations/erosions over the original

Results – Ablation study

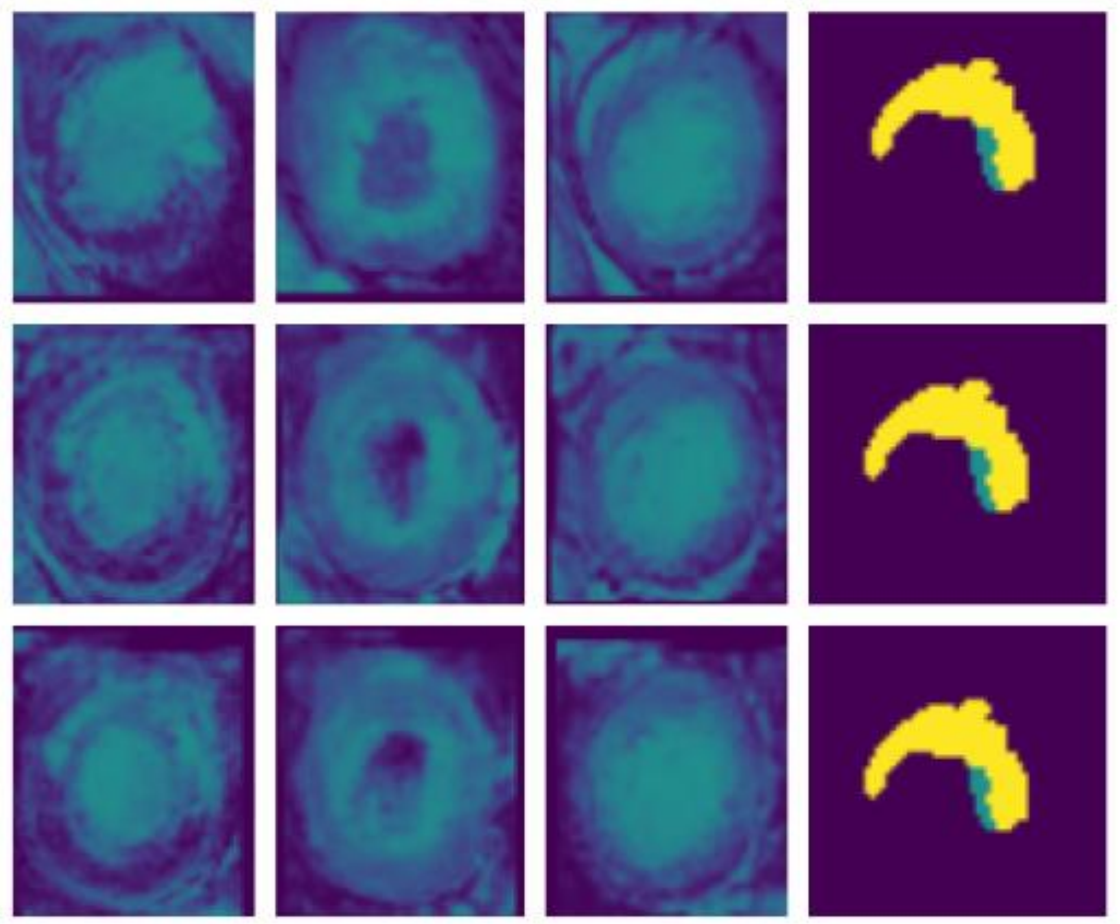


Fig. 3. Example style modifications. The proposed data augmentation strategy allows us to generate realistic images with different styles.

Table 1. 2D Dice score (mean ± standard deviation) of the proposed method for scar and scar+edema using different training data.

Training data	Scar	Scar + Edema
Original data	0.202 ± 0.286	0.170 ± 0.253
Original data + cropping + normalizing	0.449 ± 0.261	0.508 ± 0.243
Style transfer	0.548 ± 0.250	0.640 ± 0.192
Epicardium warping	0.490 ± 0.260	0.586 ± 0.222
Scar + edema rotation	0:466 ± 0:241	0:554 ± 0.224
Scar + edema dilation + erosion	0:458 ± 0.299	0:600 ± 0.224
All spade	0.518 ± 0:286	0:617 ± 0.253

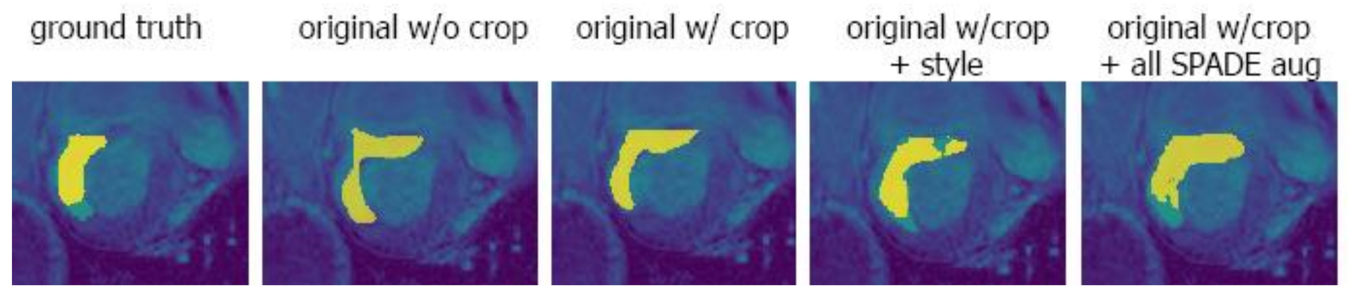


Fig. 4. Example of improvement offered by the proposed data augmentation technique

Results – Test set

- Post-processing based on unconnected components deletion and convex hull merging was applied to ensure segmentations satisfy anatomical constrains
- Two ensembles with and without such process are evaluated

Table 2. 3D Dice score for the final testing set of 20 subjects.

Approach	Scar	Scar + Edema
5 models ensemble	0.625 ± 0.255	0.677 ± 0.146
5 models ensemble + post-processing	0.635 ± 0.281	0.692 ± 0.143
15 models ensemble	0.636 ± 0.243	0.687 ± 0.131
15 models ensemble + post-processing	0.665 ± 0.241	0.698 ± 0.128

Conclusions

- Larger ensemble results in improved performance (larger training sizes).
- The effect of the low validation size was noticeable as a noisier validation curve, and attenuated by means of a greater regularization power, with an overall improved accuracy.
- Consistent results across the different semantic manipulations and their respective synthesis, indicate the potential of this set of transformations for improving generalization of multi-modality cardiac pathology segmentation algorithms.

References

1. Zhuang, X.: Multivariate mixture model for myocardial segmentation combiningmulti-source images IEEE Trans. Pattern Anal. Mach. Intell, 41(12), 2933–2946, 2019.
2. Park, T., et al.: Semantic image synthesis with spatially-adaptive normalization. CVPR, 2019.
3. Azad, R., et al.: Bi-directional ConvLSTM u-net with densley connected convolutions. ICCVW, 2019.

Acknowledgements

This work was partly funded by the European Union'sHorizon 2020 research and innovation programme under grant agreement no.825903 (euCanSHare project). This work has been partially supported by the Spanish project PID2019-105093GB-I00 (MINECO/FEDER, UE) and CERCA Programme/Generalitat de Catalunya.). This work is partially supported by ICREA under the ICREA Academia programme. KL is supported by the Ramon y Cajal Program of the the Spanish Ministry of Economy and Competitiveness under grant no. RYC-2015-17183.