

# Direct Learning Left Ventricular Meshes from CMR Images

Lukasz Romaszko<sup>1</sup>, Agnieszka Borowska<sup>1</sup>, Alan Lazarus<sup>1</sup>, Hao Gao<sup>1</sup>, Xiaoyu Luo<sup>1</sup>, Dirk Husmeier<sup>1</sup>

<sup>1</sup>School of Mathematics and Statistics, University of Glasgow,  
Glasgow G12 8SQ, UK

lukasz.romaszko@glasgow.ac.uk; agnieszka.borowska@glasgow.ac.uk; a.lazarus.1@research.gla.ac.uk;  
hao.gao@glasgow.ac.uk; xiaoyu.luo@glasgow.ac.uk; dirk.husmeier@glasgow.ac.uk

**Abstract** - Biomechanical studies of the left ventricle (LV) typically rely on a mesh of finite element nodes for a discrete representation of the LV geometry, which is used in an approximate numerical solution of the cardio-mechanical equations based on finite-element discretisation. This is typically done by first manually annotating cardiovascular magnetic resonance (CMR) scans, second creating a preliminary mesh, third manually correcting the mesh to account for motion. The whole process requires specialist knowledge, is time consuming and prone to human error, which prohibits its common adoption in the clinics. We propose to overcome these shortcomings by applying statistical pattern recognition techniques to CMR images. In particular, we train a convolutional neural network (CNN) to predict the LVM via learning its principal component representation directly from CMR scans. As a useful side-product we obtain a low-dimensional representation of the LVM, which is of interest for surrogate models (emulators) of the myocardium constitutive models.

**Keywords:** Statistical pattern recognition, convolutional neural networks, CMR images, left ventricular mesh, dimensionality reduction, principle component analysis.

## 1. Introduction

Biomechanical studies of the left ventricle (LV) typically rely on a mesh of finite element nodes for a discrete representation of the LV geometry, see e.g. [1]. This left ventricular mesh (LVM) is used in an approximate numerical solution of the cardio-mechanical equations based on finite-element discretisation. Hence, obtaining the LVM is an important task in numerous studies, especially those focused on personalised health care, in which a mesh needs to closely represent a given subject's physiology. This is typically done by first manually annotating a subject's cardiovascular magnetic resonance (CMR) scans, second creating a preliminary mesh, third manually correcting the mesh to account for motion. The whole process requires specialist knowledge, is time consuming and prone to human error, which prohibits its common adoption in the clinics. We propose to overcome these shortcomings by applying statistical pattern recognition techniques to CMR images. In particular, we train a convolutional neural network (CNN) to predict the LVM via learning its principal component representation directly from CMR scans. Our aim is to achieve this without the need to perform manual tasks, such as pixel labelling (classification) or shape annotations. As a useful side-product we obtain a low-dimensional representation of the LVM, which is of interest for surrogate models (statistical emulators) of the myocardium constitutive models.

CNNs have been successfully applied in the context of cardiac image analysis to different tasks. Automated segmentation of CMR scans is considered in [2]. This work, however, analyses only short-axis images, while long axis scans provide important information, e.g. regarding the length of the ventricle, for an accurate reconstruction of the LVM. The issue of survival prediction based on sequences of CMR scans is addressed in [3]. The geometries there, however, are not learned directly from the images, but rely on a non-rigid registration approach to map each patient's data into a template space.

The outline of this paper is as follows. Section 2 describes our mesh dataset constructed from CMR scans. In Section 3 we present the developed CNN, which uses as its only input CMR scans and outputs the reconstructed LVMs, together with the corresponding PCA representations. Section 4 reports the results suggesting that CNNs can provide a useful tool for automatising the process of LVM generation. In Section 5 we discuss our findings and plans for further research.

## 2. Data

**CMR Scans** The meshes of the left ventricle are constructed from CMR scans, which in our dataset come from 70 healthy volunteers (HVs) and 128 patients with acute myocardial infarction (MI), which amounts to 198 subjects. MI patients were selected from a population of 324 patients with acute ST-elevation MI analysed by the British Heart Foundation (ClinicalTrials.gov identifier: NCT02072850) between 17 July 2011 and 22 November 2012.

For the patients the CMR scans were carried out at 1.5 Tesla (Siemens Avanto, Siemens Healthcare, Erlangen, Germany) within 2 days after their hospital admission. Steady-state free precession cine scans were taken, in which a short-axis (SA) stack ranged from the base to the apex, with the slice thickness of 7 mm with a 3 mm gap. The remaining cine imaging parameters were set in a typical manner, with matrix =  $180 \times 256$ , flip angle =  $80^\circ$ , TR: 3.3 ms, TE: 1.2 ms, bandwidth: 930 Hz/pixel. The voxel size was  $1.32 \times 1.32 \times 7 \text{ mm}^3$ . Cine images were acquired in the three-chamber, horizontal long-axis (LA), and vertical LA planes. There was no difference in the scanning procedure for the HVs compared to the MI patients other than that for HVs younger than 45 years no intravenous gadolinium contrast was used.

**Left Ventricular Meshes** We fitted 3-dimensional LV geometries using 4–8 SA scans (number depends on the subject’s heart size) and 3 LA cine images at early diastole. To this end, we first manually annotated the boundaries of the LV wall in each image using an in-house MATLAB code. Second, we aligned the SA LV wall boundaries to the boundaries in LA images. In the third step segmented LV boundaries were fitted to a prolate spheroidal B-spline described template LV mesh with one layer of hexahedron elements across the wall, shown in Figure 1, Panel A. The mesh consists of  $2 \times 2896 = 5792$  vertices (each with 3D coordinates), which means that it is represented by  $3 \times 5792 = 17368$  values. We refer to the original input as “17k”.

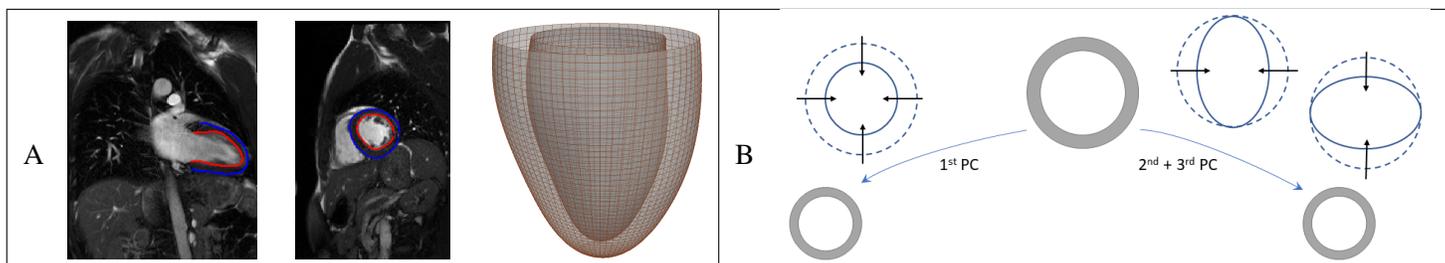


Fig. 1: A: examples of annotated CMR scans and a constructed left ventricular mesh (LVM). Left: long-axis view (one chamber plane); middle: short-axis view; right: LVM composed of two identical meshes, one for the epicardium (inside layer, annotated red in scans) and one for the endocardium (outside layer, annotated blue in scans). B: applying the deformation defined by an idealised 1st principal component (PC), controlling the overall size of the short-axis view of the LVM, can be equivalently achieved by applying the deformations defined by idealised 2nd and 3rd PCs, which control the width and length.

**Data Preprocessing and Cleaning** To prepare the CMR images to be inputs to the CNN we needed to appropriately unify the images, vastly varying among subjects. Especially for the LA views care needed to be taken to guarantee that they are oriented in the same direction. Pixel spacing was applied to resize the images to the same units. Subjects with some images missing and/or less than 6 SA slices were excluded (16 subjects in total). Image cropping was performed with the image patches being centred at the top of the LV. We allowed for the reflect option if outside of the image frame.

### 3. Method: Convolutional Neural Network

**Preparation: PCA and Baseline** Given the limited size of our dataset, predicting high dimensional LVM directly from images is difficult. Therefore we decided to train the CNN to predict a lower dimensional representation of the LVM. Such a low dimensional representation was obtained by applying principle component analysis (PCA) to the mesh dataset. This approach has an additional advantage of delivering as a side product an encoded representation of the LVM, important in several other statistical problems related to the LV (e.g. for constructing statistical emulators). We used the first 4 principal components as first, they can explain almost 86% of the total variation of the data, and second, as encoding into 4 values is useful from the perspective of designing efficient statistical emulators. Finally, we set up a baseline method referred to as the mean mesh method (MMM) which predicts the mean mesh from our mesh dataset for any image input.

**CNN Architecture and Training** To predict LVM from CMR images, we use a CNN which we train to predict the ground-truth (GT) meshes by minimising the reconstruction error set to the mean squared error (MSE) between the GT mesh and its reconstruction. We use the same metric for evaluation. Alternative approaches of predicting either 4 PCA components instead of meshes or PCA-reconstructed mesh are disadvantageous, as the former can be ambiguous without the associated PCA mapping *within* the CNN, see Figure 1, Panel B, while the latter introduces unnecessary noise into training. We use 14 fold cross-validation (CV) for training, i.e. we split out dataset consisting of 182 subjects into 14 splits of 13 subjects (with HV and MI patients randomly shuffled) and use one split for validation and one for testing.

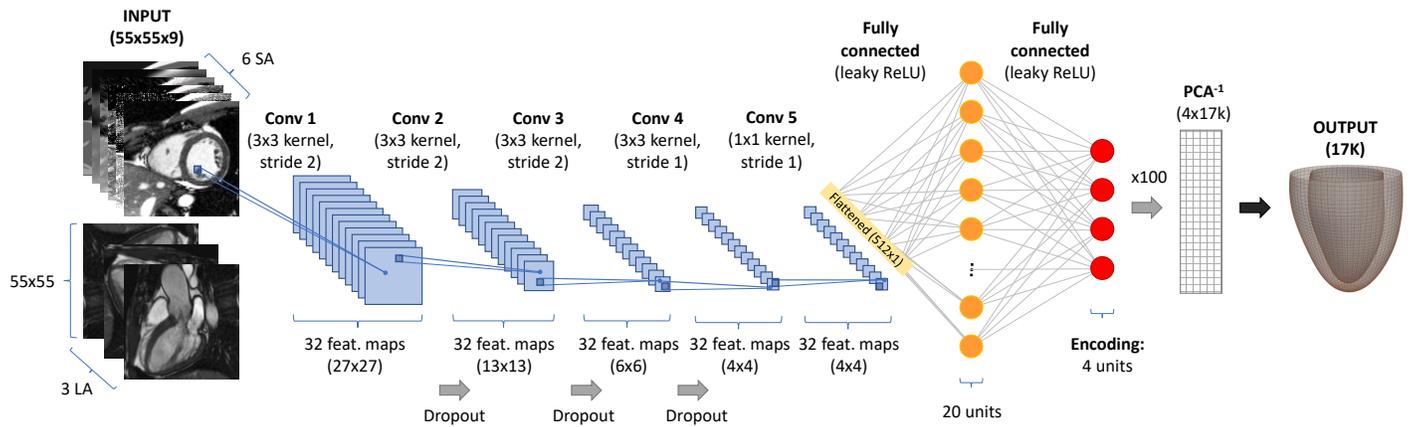


Fig. 2: Scheme of the convolutional neural network for prediction of the left ventricular mesh directly from CMR images. Input: 9 CMR scans, 6 for short-axis (SA) and 3 for long-axis (LA). Output: a mesh represented by approx. 17 thousand values (17k). There are 5 convolutional layers (Conv. 1–5) and next 2 fully connected layers with leaky ReLU activations.

Figure 2 presents the CNN architecture. We note that all activations except the linear output are leaky ReLU ( $x = \max(0.2x, x)$ ). To prevent overfitting we take random crops of the image during the training, and the central crop at the test time. Images are resized into  $64 \times 64$  pixels and  $55 \times 55$  crops are used. Random cropping is an instance of data augmentation in deep learning, which is a standard practice in computer vision, see [4]. Given that the pixel values of the original images are large, i.e.  $\in [0, 255]$ , while the CNN initialisation assumes the inputs have unit standard deviation, we standardised the image inputs by dividing by 100 and demeaning. Similarly, since PCA coefficients values can be approximately  $\pm 100$ , we also let the network predict lower values by multiplication of its output by 100 on top of it.

The CNN hyperparameters, the L2 regularisation and the learning rate, are set for each CV split to 0.0001 and 0.0002, respectively. These values were chosen manually by first investigating the validation split performance of the CNN for 400 epochs and then choosing to stop early at the 300th epoch.

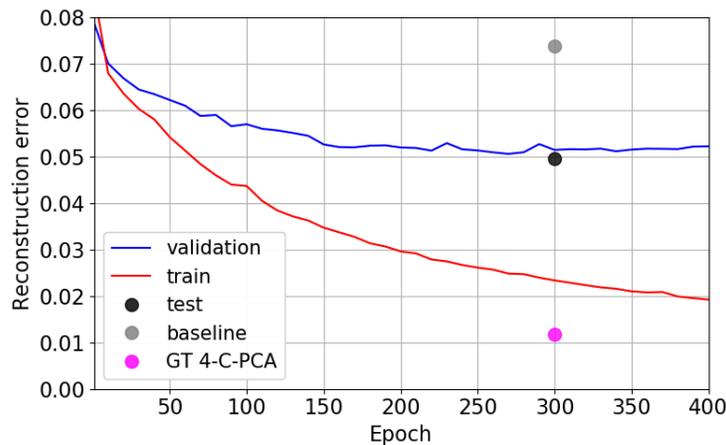


Fig. 3: Reconstruction error (mean squared errors) in predicting left ventricular ground-truth meshes: convolutional neural network (CNN) on the training set, validation set and the test set against the baseline (mean mesh method). The error of the ground truth 4 component PCA (GT 4-C-PCA) constitutes a lower bound on the attainable error for our CNN specification with a 4-dimensional bottleneck. The baseline error, the PCA lower bound and the CNN error on the test set do not depend on the number of epochs. Early stopping after 300 epochs was used to prevent overfitting.

Figure 3 illustrates the dynamics (as a function of the number of epochs) of the reconstruction error on the training and validation datasets, and compares these with the test set performance as well as with the baseline error of 0.074 recorded for the MMM prediction. Approximately in the 300th epoch the CNN error on the validation set is in the minimum, so we perform early stopping at this point to prevent overfitting. The validation set error is roughly constant after 150 epochs and

equal to approximately 0.050. The final error on the test set is equal to 0.049. This means that our CNN provides a noticeable improvement over the baseline (by more than one third).

#### 4. Results

Since we trained the CNN to predict the 4-dimensional representation of the LVM, the associated PCA error constitutes a lower bound on the reconstruction error for our network. The reconstruction error is equal to 0.011 and is equal to the mean error in a leave-one-out CV in which an independent PCA is carried out on each training split and the first 4 principal components are used to predict the remaining validation mesh.

Another finding from our analysis is that using all 9 slices (6 SA and 3 LA images) is useful for reducing the reconstruction error. This was not obvious *a priori*, since the LA images tend to be much more heterogeneous than the SA images, with the associated boundaries being much harder to detect (even in manual segmentation) than for circular SA boundaries. Finally, we found out that the bottom 3 SA slices are the most informative for the correct LVM prediction. This is a consequence of the LVMs in our dataset being aligned at the top, hence wider or narrower bottom SA slices indicate the length of the LV.

#### 5. Conclusion

We have developed a convolutional neural network able to learn the high-dimensional mesh of the left ventricle directly from CMR scans. Our results demonstrate that this approach can provide a useful alternative compared with the traditional approach to generating the LVMs, which is based on tedious manual segmentation of cine images and a following motion correction. The developed CNN is 33% more accurate than the benchmark of the mean mesh.

A natural strand of further research is performing a sensitivity analysis with respect to the network architecture, noise set-ups or adopting data augmentation. Of particular interest is investigating the impact on the prediction accuracy of the size of the encoded output, which is now 4 dimensional. This choice was dictated by our interest in obtaining a low-dimensional representation of the LVM, required by e.g. statistical emulators of cardiovascular biomechanical models. However, we note that it may be easier for a network to learn a richer representation, from which one can then select the number of components as required. Another direction we aim to pursue in the future is to combine the existing reconstruction CNN with a segmentation CNN detecting the boundaries of the LV in CMR scans. We note that the pipeline will still be fully automatic, without a need of human intervention, but separating the task which the current network completes simultaneously (detection and reconstruction) into two steps is likely to enhance the overall performance.

#### Acknowledgements

This work was funded by the UK Engineering and Physical Sciences Research Council (EPSRC), grant number EP/N014642/1. Alan Lazarus is partially funded by a grant from GlaxoSmithKline plc. Dirk Husmeier is supported by a grant from the Royal Society of Edinburgh, award number 62335.

#### References

- [1] H. Gao, H. Wang, C. Berry, X. Luo, and B. E. Griffith, "Quasi-static image-based immersed boundary-finite element model of left ventricle under diastolic loading," *International journal for numerical methods in biomedical engineering*, vol. 30, no. 11, pp. 1199–1222, 2014.
- [2] W. Bai, M. Sinclair, G. Tarroni, O. Oktay, M. Rajchl, A. M. L. Ghislain Vaillant, N. Aung, E. Lukaschuk, M. M. Sanghvi, F. Zemrak, J. M. P. Kenneth Fung, V. Carapella, Y. J. Kim, H. Suzuki, B. Kainz, P. M. Matthews, S. E. Petersen, S. K. Piechnik, S. Neubauer, B. Glocker, and D. Rueckert, "Automated cardiovascular magnetic resonance image analysis with fully convolutional networks," *Journal of Cardiovascular Magnetic Resonance*, vol. 20, no. 1, p. 65, 2018.
- [3] G. A. Bello, T. J. Dawes, J. Duan, C. Biffi, A. de Marvao, L. S. Howard, J. S. R. Gibbs, M. R. Wilkins, *et al.*, "Deep-learning cardiac motion analysis for human survival prediction," *Nature machine intelligence*, vol. 1, no. 2, p. 95, 2019.
- [4] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *arXiv preprint arXiv:1712.04621*, 2017.