Medical Image Analysis 80 (2022) 102498

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Automatic 3D+t four-chamber CMR quantification of the UK biobank: integrating imaging and non-imaging data priors at scale

Yan Xia[a,b,1], Xiang Chen[a,1], Nishant Ravikumar[a], Christopher Kelly[a], Rahman Attar[a,b], Nay Aung[d,e], Stefan Neubauer[c], Steffen E. Petersen[d,e], Alejandro F. Frangi[a,b,f,g,h,*]

[a] Centre for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), School of Computing, University of Leeds, Leeds, UK
[b] Biomedical Imaging Department, Leeds Institute for Cardiovascular and Metabolic Medicine (LICAMM), School of Medicine, University of Leeds, Leeds, UK
[c] Oxford Centre for Clinical Magnetic Resonance Research (OCMR), Division of Cardiovascular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK
[d] William Harvey Research Institute, NIHR Barts Biomedical Research Centre, Queen Mary University of London, London, UK
[e] Barts Heart Centre, St Bartholomew's Hospital, Barts Health NHS Trust, London, UK
[f] Department of Cardiovascular Sciences, KU Leuven, Leuven, Belgium
[g] Department of Electrical Engineering, KU Leuven, Leuven, Belgium
[h] Alan Turing Institute, London, UK

ARTICLE INFO

Article history:
Received 2 November 2020
Revised 14 May 2022
Accepted 24 May 2022
Available online 27 May 2022

Keywords:
UK Biobank
Cardiac MR
Deep learning
Statistical shape models
Population imaging
Fully automatic analysis
Cardiac functional indexes
Cardiac morphological analysis

ABSTRACT

Accurate 3D modelling of cardiac chambers is essential for clinical assessment of cardiac volume and function, including structural, and motion analysis. Furthermore, to study the correlation between cardiac morphology and other patient information within a large population, it is necessary to automatically generate cardiac mesh models of each subject within the population. In this study, we introduce MCSI-Net (Multi-Cue Shape Inference Network), where we embed a statistical shape model inside a convolutional neural network and leverage both phenotypic and demographic information from the cohort to infer subject-specific reconstructions of all four cardiac chambers in 3D. In this way, we leverage the ability of the network to learn the appearance of cardiac chambers in cine cardiac magnetic resonance (CMR) images, and generate plausible 3D cardiac shapes, by constraining the prediction using a shape prior, in the form of the statistical modes of shape variation learned a priori from a subset of the population. This, in turn, enables the network to generalise to samples across the entire population. To the best of our knowledge, this is the first work that uses such an approach for patient-specific cardiac shape generation. MCSI-Net is capable of producing accurate 3D shapes using just a fraction (about 23% to 46%) of the available image data, which is of significant importance to the community as it supports the acceleration of CMR scan acquisitions. Cardiac MR images from the UK Biobank were used to train and validate the proposed method. We also present the results from analysing 40,000 subjects of the UK Biobank at 50 time-frames, totalling two million image volumes. Our model can generate more globally consistent heart shape than that of manual annotations in the presence of inter-slice motion and shows strong agreement with the reference ranges for cardiac structure and function across cardiac ventricles and atria.

1. Introduction

According to the World Health Organisation (WHO, 2020), cardiovascular disease (CVD) is the most prevalent cause of death worldwide, accounting for nearly 18 million deaths each year. Identifying individuals at risk of CVDs and ensuring they receive appropriate and timely treatment can help prevent premature deaths.

Early quantitative assessment of cardiac structure, motion, and function support preventive care and early cardiovascular treatment. Therefore, fully automated analysis and interpretation of large-scale population-based cardiovascular magnetic resonance

Abbreviations: CMR, Cardiac Magnetic Resonance; CVD, Cardiovascular Disease; UKB, UK Biobank; MLP, Multi-level perceptron; FCN, Fully Convolutional Networks; CNN, Convolutional Neural Networks; LAX, Long axis; SAX, Short axis; LV, Left ventricle; RV, Right ventricle; LA, Left atrium; RA, Right atrium; ED, End diastole; ES, End systole; HdMM, Hybrid Mixture Model; CPD, Coherent point drift; PCA, Principal component analysis; TSP, Thin-plate spline; gCPD, Generalised coherent point drift.

* Corresponding author.
E-mail address: a.frangi@leeds.ac.uk (A.F. Frangi).
[1] Contributed equally.

https://doi.org/10.1016/j.media.2022.102498

(CMR) imaging studies is of high importance. This analysis helps to identify patterns and trends across population groups, and accordingly, provides insights into key risk factors before CVDs fully develop.

UK Biobank (UKB) is currently the world's most extensive prospective population study (Petersen et al., 2015), which contains questionnaire data, biological samples, physical measurements, CMR images, and so forth (Petersen et al., 2013). CMR is an essential element of multi-organ, multi-modality imaging visits for patients in multiple dedicated UKB imaging centres that will acquire and store imaging data from 100,000 participants by 2022. At this time, CMR scans of 40,000 subjects have been released and are available for health-related research.

We believe that 3D analysis is critical for the accurate clinical assessment of cardiac function. In this work, we introduce a new approach that ensures the global coherence of cardiac anatomy and naturally lends itself to any further analysis requiring the full 3D anatomy; for example, in interventional treatment planning requiring precise volumetric quantification, mechanical and flow simulations, motion analysis, and modelling the associations between cardiac structure and patient clincial data (such as socio-demographic, lifestyle and environmental factors, or family history, genetic, and omics data). Though fully automatic 3D segmentation is required to facilitate such analyses, the complexity of anatomical structures, intensity and morphology variation across a population cohort, and the sparse information available from CMR images (typically on average around 12 image slices covering the full heart) make this task challenging.

In our previous studies, (Attar et al., 2018) and (Attar et al., 2019b), we showed that 3D statistical shape model-based approaches have the power and potential to automatically segment cardiac structures, and generate associated cardiac function indexes. This success is attributed to the inclusion of prior knowledge of cardiac shape, within the segmentation method. These segmentation approaches typically use simple sets of features to fit a shape model through an iterative process and the goal is to minimise the Mahalanobis distance between an intensity profile sampled at a candidate position and its corresponding intensity appearance model, by deforming the shape within its range of normal variation to match the image data.

On the other hand, in the last decade, fully convolutional networks (FCN) have shown great potential in image-based pattern recognition in a variety of tasks, including cardiac segmentation. However, their output results are, by nature, 2D segmentation masks for every short axis (SAX) and long-axis (LAX) CMR slice. Although these 2D masks are sometimes extended via a further step of non-rigid registration to a 3D atlas to produce a 3D cardiac shape (Duan et al., 2019), this is not efficient for learning topological shape information. Furthermore, this is based on the strong assumption that all the 2D segmentation masks are always correct and meaningful, however, in practice, there are often errors in pixel-wise segmentation approaches due to spurious false positives. For instance, Painchaud et al. (2020) recently proposed a generative model based on a variational autoencoder to identify anatomically implausible results following 2D segmentation, and corrected these to fit the closest anatomically correct contours, based on the learned latent space. This provides further evidence that, conducting CMR segmentation in 2D requires a subsequent quality control step, in order to ensure that the downstream quantitative analyses are accurate. Consequently, large-scale studies would benefit from an efficient approach for reconstructing cardiac shapes in 3D, as it would remove the need for multiple sequential steps involving pixel-wise 2D segmentation, followed by quality control, and iterative registration-based 3D shape reconstruction, which can be cumbersome and time-consuming.

In order to address these limitations, in a recent study (Attar et al., 2019a) published at MICCAI 2019, we proposed an approach to exploit image features obtained using deep FCNs trained on both SAX and LAX views, along with a rich shape prior learned using a statistical shape model, to directly predict the shape-space parameters required to reconstruct 3D cardiac shapes. Another significant aspect of this study was the integration of patient information into the process of shape prediction using a Multilayer Perceptron (MLP). This information, which is currently ignored by most cardiac segmentation or shape generation techniques, has been shown in different clinical studies to have an impact on cardiac morphology and structure (Gilbert et al., 2019). To evaluate our method, in addition to comparing against manual measurements, we also compared our performance against two state-of-the-art methods, namely, the work by Bai et al. (2018) in which the authors propose a 2D convolutional neural network (CNN)-based segmentation method and our previous work (Attar et al., 2019b) where, we analysed and reported cardiac functional indexes of 20,000 subjects of the UKB through a fully automatic quality-controlled image parsing framework.

In this paper, we have substantially extended our previous work presented at MICCAI 2019 (Attar et al., 2019a). The main contributions of this paper that were not addressed in our previous work are as follows:

- We extended the shape model from biventricular to a four-chamber cardiac mesh model. We segmented all four cardiac chambers; namely, left ventricle (LV), right ventricle (RV), left atrium (LA), and right atrium (RA). For the LV, we segmented both the endocardium and epicardium surfaces; while for the other chambers, we only segment the endocardium. This is because the myocardium is too thin to reliably distinguish epicardium from the endocardium. Consequently, there is no label available to use for the training. This was achieved by first generating a reference cohort of four-chambered cardiac meshes through the non-rigid registration of a four-chambered cardiac atlas, to a set of 3D points obtained from manual delineations. To ensure high accuracy when fitting the 3D shape to the stack of manually delineated contours, we adopted an alternating global-local affine and non-rigid registration approach, based on the generalised coherent point drift (gCPD) algorithm proposed previously by our group Ravikumar et al. (2017), and thin plate spline (TPS) based mesh warping. TPS-based mesh warping is used to prevent mesh intersections between adjacent cardiac chambers, and helps ensure shared nodes/node-connectivity between adjacent structures as defined in the original atlas mesh is preserved following registration.
- We standardised our network in the feature extraction phase so that any FCNs can be used for this task. It is an innovative end-to-end deep neural network that directly predicts 3D shape parameters derived from a Principal Component Analysis (PCA) space. The network was optimised using a loss function defined in the domain of shape space parameters which weights each PCA mode of variation independently, prioritising the more significant modes and leading to more accurate shape prediction. In addition to the shape parameters, we extended our network to also learn the similarity transformation parameters required to transform the generated shape back from the normalised PCA space, to the image space. We achieved this through a multitask learning approach, where the extracted features are used to jointly optimise a second loss function defined over the desired transformation parameters.
- We adopted a novel approach of exploiting all the CMR image views (short axis and long axis two-, three- and four-chamber CMR images) and a wider range of patient data (almost tripled the quantity to cover more attributes) simultaneously, to pre-

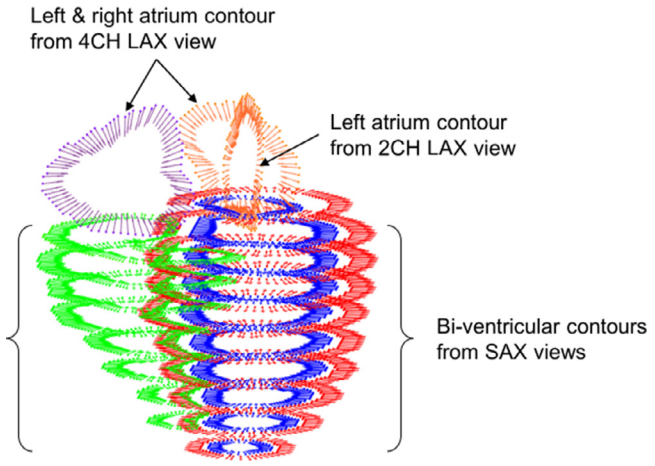

Fig. 1. An example stack of manual contours represented as a hybrid point set comprising spatial positions and associated surface normal vectors, extracted from SAX, two- (2CH) and four-chamber (4CH) LAX view slices. Contour colours denote the following - Blue: LV-endocardium; Red: LV-epicardium; Green: RV-endocardium; Orange: LA-endocardium; Purple: RA-endocardium.

dict 3D four-chambered cardiac shapes. The introduction of the patient information yielded a positive impact on shape prediction. We hope this work inspires other researchers to exploit such informative priors in their applications, to improve the performance of their models.

- We investigated the importance of the available CMR slices towards the accuracy of the final predicted shape, and evaluated the performance of our approach in scenarios where only a few SAX and/or LAX slices are available. These experiments were conducted to validate the hypothesis that 3D cardiac shapes can be predicted accurately, given sparse CMR acquisitions. This is particularly relevant to applications requiring acceleration of CMR scan acquisitions, at minimal cost to subsequent cardiac quantification accuracy.

2. Method

In the sections that follow, we describe the data and methods used within each step of our framework, and our corresponding design choices. These are ordered as follows — first, we describe the generation of reliable 3D reference shapes from the stack of 2D manual contours; next, the construction of the point distribution model; and finally, we provide details of the input data used to train and validate MCSI-Net, and describe its architecture.

2.1. Reference 3D cardiac shapes

To generate the reference cohort of 3D cardiac shapes, we first construct the 3D stack of 2D manual delineations by combining all the available contours from the three views (i.e. SAX, two- and four-chamber LAX slices) while exploiting the recorded orientation and position information available in the DICOM header to be aligned with their actual corresponding CMR images, as shown in Fig. 1. The mean shape of a high-resolution atlas of the human heart available from a recent study by Rodero et al. (2021) was then registered to the stack of manual contours using an initial affine and subsequent non-rigid registration steps, to produce the patient-specific, four-chamber cardiac shape model. The resulting 3D mesh model comprises five structures; the Left Ventricle (LV) endo- and epicardium, Right Ventricle (RV), Left Atrium (LA) and Right Atrium (RA).

To generate the reference cardiac meshes, we used the gCPD algorithm (Ravikumar et al., 2017) for affine and non-rigid registration of the atlas to each subjects' stack of manual contour points. The gCPD algorithm is based on a hybrid mixture model that uses shapes represented as 6D hybrid point sets. Here, each point in each shape is a 6D vector created by concatenating the spatial position vector with the associated surface normal vector. An example of the 6D-hybrid point set representation of a stack of manual contours is shown in Fig. 1. Such a hybrid representation of shape provides a richer description of local variations in geometry than spatial positions alone, providing more discriminative information than the latter to guide the registration of the cardiac atlas mesh to the stack of manual contours of each subject. For each subject comprising manual delineations, we start with affine registration to align the two hybrid point sets and then perform non-rigid registration to deform the cardiac atlas mesh to fit the set of 6D points obtained from manual delineations. Finally, using the point-wise correspondences established between the atlas and the stack of manual contours, we warp the cardiac atlas mesh to the latter using TPS interpolation. This helps ensure shared nodes/node-connectivity across adjacent structures in the atlas surface mesh are preserved following registration.

Here, we briefly review the gCPD algorithm and outline the registration steps followed to generate the reference cohort of subject-specific cardiac meshes. gCPD treats the problem of registering a source hybrid point set $\mathbf{X} \in \mathbb{R}^{n\times 6}$ to a target hybrid point set $\mathbf{Y} \in \mathbb{R}^{n\times 6}$ as one of probability density estimation. $\mathbf{X} = [\mathbf{X}_p, \mathbf{X}_n]$ is considered to represent the centroids of a Hybrid Mixture Model (HdMM), from which a transformed set of observations $\mathbf{Y} = [\mathbf{Y}_p, \mathbf{Y}_n]$ are sampled. Here, $\mathbf{X}_p$ & $\mathbf{X}_n$ represent the spatial positions and associated surface normal vectors of the cardiac atlas mesh, while $\mathbf{Y}_p$ & $\mathbf{Y}_n$ denote the spatial positions and associated surface normal vectors of each subjects' stack of manual contours from the UK Biobank cohort. Each mixture component in the HdMM is formulated as a product of two distinct types of distributions, namely, a Student's t-distribution and a Von-Mises-Fisher distribution, designed to model the joint distribution of point spatial positions and their associated unit normal vectors, respectively. Consequently, by fitting the HdMM to $\mathbf{Y}$, in a manner analogous to data clustering, the underlying spatial transformations that map $\mathbf{X}$ to $\mathbf{Y}$ and maximise the likelihood of the latter being sampled from the former, are estimated. This is achieved by maximising the log-likelihood function using the expectation-maximisation (EM) algorithm, given as follows:

$$p(\mathbf{Y}|\mathbf{X}, \mathcal{T}) = \sum_{j=1}^{M} \log \sum_{i=1}^{N} \pi_i \mathcal{S}(\mathbf{y}_p^j | \mathcal{T}\mathbf{x}_p^i, \sigma^2, \nu_i) \times \mathcal{F}(\mathbf{y}_n^j | \mathcal{T}\mathbf{x}_n^i, \kappa_i), \quad (1)$$

where, $[\mathbf{x}_p^i, \mathbf{x}_n^i]_{i=1..N} \in \mathbf{X}$ denote N mixture components, $[\mathbf{y}_p^j, \mathbf{y}_n^j]_{i=1..N} \in \mathbf{Y}$ are the observed data points, and subscripts p and n denote the spatial position and normal vector components of the hybrid point sets. Additionally, $\mathcal{T}$ are the spatial transformation parameters to be estimated, $\pi_{i=1..N}$ are mixture coefficients, $\mathcal{S}$ is the Student's t-distribution function with variance σ^2 and shape parameter ν_i and $\mathcal{F}$ is a Von-Mises-Fisher distribution function with dispersion κ_i, associated with each mixture component i.

EM is necessary as no tractable solution exists for directly maximising Eq. (1). The EM algorithm iteratively alternates between two steps. In the expectation (E)-step, the posterior probabilities that describe the responsibility of each mixture component $\mathbf{x}^i = [\mathbf{x}_p^i, \mathbf{x}_n^i]$, in describing the observed data points $\mathbf{y}^j = [\mathbf{y}_p^j, \mathbf{y}_n^j]$, are estimated. In the maximisation (M)-step, the posterior probabilities estimated in the preceding E-step are used to maximise Eq. (1) with respect to the unknown spatial transformation parameters $\mathcal{T}$, mixture coefficients π_i, the covariance $\sigma^2\mathbf{I}$ and the shape parameters ν_i associated with each Student's t-distribution in the mixture, and the dispersion parameters κ_i associated with each Von-Mises-Fisher distribution (analogous to the variance of a Gaus-

sian distribution) in the mixture. gCPD considers an isotropic and shared covariance across all components' Student's t-distributions in the HdMM, however, a unique shape parameter ν_i is estimated for each component in the mixture, which controls the heaviness of the tails of each t-distribution and hence provides automatic, locally adaptive robustness to noise/outliers that may be present in the data. Both affine and non-rigid registration are achieved by alternating between these two steps of EM, until a suitable convergence criterion is reached.

The gCPD algorithm was chosen over the more widely used coherent point drift (CPD) point set registration algorithm proposed by Myronenko and Song (2010), as the former offers the following advantages over the latter: (1) by utilising Student's t-distributions in the HdMM, gCPD enables data-driven estimation of the constituent shape parameters of each t-distribution, thereby providing automatic and locally adaptive robustness to noise and outliers. Conversely, CPD handles noise and outliers through a linear combination of a uniform distribution component within a Gaussian mixture model formulation, requiring manual tuning of a hyperparameter that controls the degree of global robustness to noise and outliers enforced on the registration process. Consequently, it is difficult to scale up generation of a reference cohort of cardiac meshes to large numbers using CPD (as required in this study), whilst ensuring robustness across all pair-wise registration executions. gCPD provides an elegant solution to this problem by guaranteeing automatic robustness to noise and outliers. (2) gCPD utilises hybrid shape representations that comprise 6D points to guide the registration of the cardiac atlas to each UK Biobank subjects' stack of manual contours. By utilising surface normals vectors as additional features alongside point spatial positions to drive the registration, the topology and localised variations in geometry of each cardiac chamber is better preserved than afforded by purely spatial position-based point set registration approaches such as CPD. Using spatial positions alone to guide cardiac atlas to manual contour registration may result in the estimation of anatomically incorrect/implausible correspondences, leading to errors in the warped atlas mesh such as intersections between adjacent surfaces of cardiac chambers, for example. However, by incorporating surface normals as additional features, more discriminative information is provided to help improve estimation of correspondences, and thereby preserve the interfaces between adjacent cardiac structures. Further details on the HdMM formulation of the gCPD algorithm and evidence of its superior registration performance to CPD are presented in Ravikumar et al. (2017); Ravikumar (2017).

The following registration steps are used to warp the cardiac atlas mesh obtained from Rodero et al. (2021) to each subjects' stack of manual contours available in the UK Biobank - (1) Consider the target hybrid point set (observed data, see Fig. 1) **Y** represents a stack of manual contours, and the source hybrid point set **X** represents centroids of a HdMM defined by vertices and associated surface normal vectors of the cardiac atlas mesh. Initially **X** is registered to **Y** by estimating a global 3D affine transformation that recovers the rotation, translation and scaling necessary to align the former to the latter. (2) Next, we conduct region-wise non-rigid registration, wherein, each chamber of the heart in **X**, is registered independently to the corresponding region-specific contours in **Y** (i.e. for example, hybrid points of the LV mesh in **X**, are only registered non-rigidly to the LV hybrid contour points in **Y**). This step is necessary as non-rigid registration using gCPD does not explicitly account for multiple objects/regions and enforces a global smoothing constraint on the estimated deformation field through Tikhonov regularisation. The non-rigid transformation in gCPD (as in the original CPD algorithm Myronenko and Song (2010)) is parameterised as a linear combination of Gaussian radial basis functions. Two parameters control the degree of smoothness of the deformation field, namely, λ a weight which controls the trade-off between smoothness and registration accuracy, and β, which represents the width of the Gaussian kernel, used to parameterise the deformation. Region-wise non-rigid registration using gCPD, leverages the additional discriminative information provided by incorporating surface normal vectors as point-wise features to guide the estimation of valid region-wise correspondences and helps prevent gross topological changes during registration. This also limits intersections between adjacent regions in the registered atlas mesh. (3) Finally, TPS-based mesh warping/interpolation is used to refine the non-rigid transformations estimated in step (2), by using the established point-wise correspondences between the atlas and contours (from step (2)), and information regarding the shared nodes/node connectivity between adjacent chambers (available from the original cardiac atlas mesh) as targets to drive the transformation of the all cardiac chamber meshes jointly/globally. This final global non-rigid registration step helps correct for any existing mesh intersections between adjacent cardiac chambers by preserving the shared nodes/node connectivity between the same (as defined by the original atlas mesh). This in turn helps ensure that the topology of the whole heart meshes and the spatial relationships between its constituent chambers are maintained.

We empirically determined that registering **X** to **Y** in this manner, and adopting distinct, sequential steps for region-wise and global non-rigid registration, provided better registration quality than using either approach alone. The values for the hyperparameters that control the smoothness of the region-wise non-rigid transformations estimated in step (2), namely, λ and β, were 0.5 and 1.0, respectively, for all regions. These were determined empirically. Increasing the value for β increases interaction between the points in the point cloud, and results in coherent motion of larger neighbourhoods of points (i.e. similar displacements are estimated for larger proportions of points). Alternatively, decreasing this value reduces interaction between the points in the point cloud and conduces localised displacement of points. Similarly, increasing the value of λ produces more coherent motion across all points globally. More details of the non-rigid registration algorithm we used to achieve this effect can be found in Ravikumar et al. (2017); Ravikumar (2017) or in the original paper on the CPD algorithm Myronenko and Song (2010).

Finally, all the reference shapes were quality controlled to maintain high accuracy in the generated shapes. As a first check, we computed the point-to-point distance of the generated shape to the stack of manual contours, and if the average error was less than half of the in-plane pixel spacing, we used the shape for the PDM construction. We then visually checked all the shapes overlaid on the stack of contours to discard any sub-optimal shapes from the dataset. Ultimately, 4,525 subjects were available after quality control and were randomly split into two sets of 3,925/600 for training/test, i.e. 3925 subjects for training the neural network, and 600 subjects for evaluating performance.

In the next section, we describe how we use the generated 3D reference shapes to construct the point distribution model, as illustrated in Fig. 2. We would like to highlight that the cardiac atlas mesh used in the study (available from Rodero et al. (2021)) includes the base of the aorta and pulmonary artery, and all cardiac valve planes in addition to all four cardiac chambers. However, as the manual delineations available for the UK Biobank cohort do not include these structures, the deformation of these additional structures is driven purely by the TPS-based mesh warping step, when generating the subject-specific meshes. Thus the aortic vessel surface visible in Fig. 2 for example, is inferred based on the predicted deformations for the adjacent cardiac chambers alone.

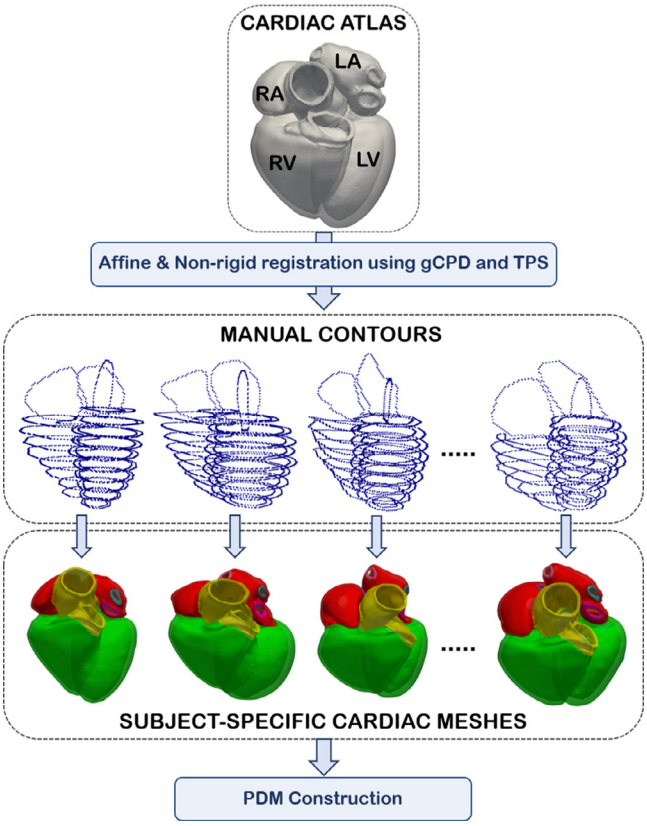

Fig. 2. The cardiac atlas mesh is registered to each stack of manual contours to produce subject-specific, high-resolution, smooth triangular meshes. Then all the new generated reference shapes are used to create the point distribution model (PDM).

2.2. Point distribution model (PDM)

To encode the mean and variance of the 3D cardiac shapes, we use a PCA-based PDM. We construct the PDM by applying PCA on a set of generalised Procrustes aligned shapes.

Following Procrustes analysis, for a training set of M shapes $\mathbf{z}_i$, we get M new shapes $\mathbf{s}_i$ (representing the i-th shape of the dataset) where all the nuisance pose parameters, i.e., translation $\mathbf{t} = (T_x, T_y, T_z)$, rotation $\mathbf{r} = (R_{q1}, R_{q2}, R_{q3}, R_{q4})$ and scaling (C) were removed. Here the desired 3D rotations are parameterised as quarternions due to the improved numeric stability they provide and their compact nature, relative to Euler angles and 3D rotation matrices, respectively. Hence, using these 8 transformation parameters, we can transform back the Procrustes-aligned shapes to their original (image) coordinates as follows:

$$\mathbf{z}_i = C_i \times \mathbf{s}_i \times \mathbf{r}_i + \mathbf{t}_i, \tag{2}$$

where $\mathbf{s}_i \in \mathbb{R}^{3N}$ represents the i-th shape as $(\mathbf{x}_1^i, \mathbf{y}_1^i, \mathbf{z}_1^i, \ldots, \mathbf{x}_N^i, \mathbf{y}_N^i, \mathbf{z}_N^i)^T$ vector.

The shape class mean $\bar{\mathbf{s}}$ and covariance $\mathbf{cov}$ of the training set of M shapes are calculated as follows:

$$\bar{\mathbf{s}} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{s}_i \tag{3}$$

$$\mathbf{cov} = \frac{1}{M-1}\sum_{i=1}^{M}(\mathbf{s}_i - \bar{\mathbf{s}})(\mathbf{s}_i - \bar{\mathbf{s}})^T \tag{4}$$

The shape covariance is represented in a low-dimensional PCA space providing $l < \min(M, N)$ eigenvectors $\mathbf{\Phi} = [\boldsymbol{\varphi}_1 \boldsymbol{\varphi}_2 \ldots \boldsymbol{\varphi}_l]$, and corresponding eigenvalues $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_l)$ computed through the Singular Value Decomposition of the covariance matrix.

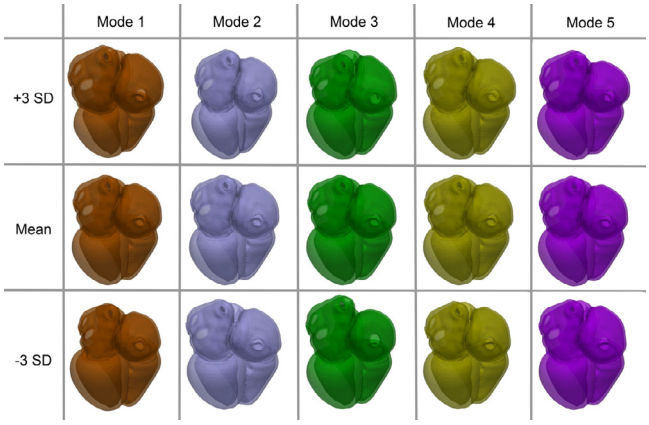

Fig. 3. Representation of the mean ± 3 standard deviation (SD) of the first five modes of variation in the 3D shape models of four-chambered cardiac.

Thus, assuming the group of shapes follows a multi-dimensional Gaussian distribution, we can approximate any shape in the group using the following linear generative model:

$$\mathbf{s} \approx \bar{\mathbf{s}} + \mathbf{\Phi}\mathbf{b} \tag{5}$$

where, $\mathbf{b} \in \mathbb{R}^l$ are shape parameters restricted to $|\mathbf{b}_i| \leq \beta\sqrt{\lambda_i}$; to capture 99.7% of shape variability, we set $\beta = 3$. The shape parameters of $\mathbf{s}$ can then be estimated as follows:

$$\mathbf{b} = \mathbf{\Phi}_l^T(\mathbf{s} - \bar{\mathbf{s}}). \tag{6}$$

where the entries of $\mathbf{b}$ are the projection coefficients of mean-centred shapes $(\mathbf{s} - \bar{\mathbf{s}})$ along the first l columns of $\mathbf{\Phi}_l$.

Fig. 3 shows the mean ± 3 standard deviation (SD) of the first five PCA modes variation and illustrate the variations present in the training dataset.

2.3. Network architecture and loss function

The overall architecture of MCSI-Net is shown in Fig. 4. The network consists of two sub-networks: MMF-Net and Loc-Net. The MMF-Net has five inputs: SAX view images, two-, three- and four-chamber LAX view images, and patient clinical data and outputs shape parameters $\mathbf{b^P} = \{b_j^P | j = 1, \ldots, k\}$. The Loc-Net takes SAX view images as inputs and predicts the transformation parameters $\mathbf{t^P} = \{T_x, T_y, T_z, R_{q1}, R_{q2}, R_{q3}, R_{q4}, C\}$ which are shown in the bottom branch of the network. The two tasks are learned through these two sub-networks independently. The proposed network is trained using the following loss functions:

$$L_b = \sum_{j=1}^{k} f(b_j^P(\theta),\ b_j^R) \tag{7}$$

and,

$$L_t = \sum_{l=1}^{8} f(t_l^P(\theta),\ t_l^R) \tag{8}$$

where k is the number of shape parameters, θ denotes the network parameters, $f(.)$ denotes the loss functions used to minimise the difference between the reference values (b_j^R and t_l^R) and the values predicted by the network ($b_j^P(\theta)$ and $t_l^P(\theta)$). For the term L_b, we used a L2 loss function to minimise the difference between b_j^R and $b_j^P(\theta)$. For the term L_t, we employed the L2 loss for minimising the translation and scaling errors, and the geodesic distance loss for the rotation formulated as quaternion parameterisations.

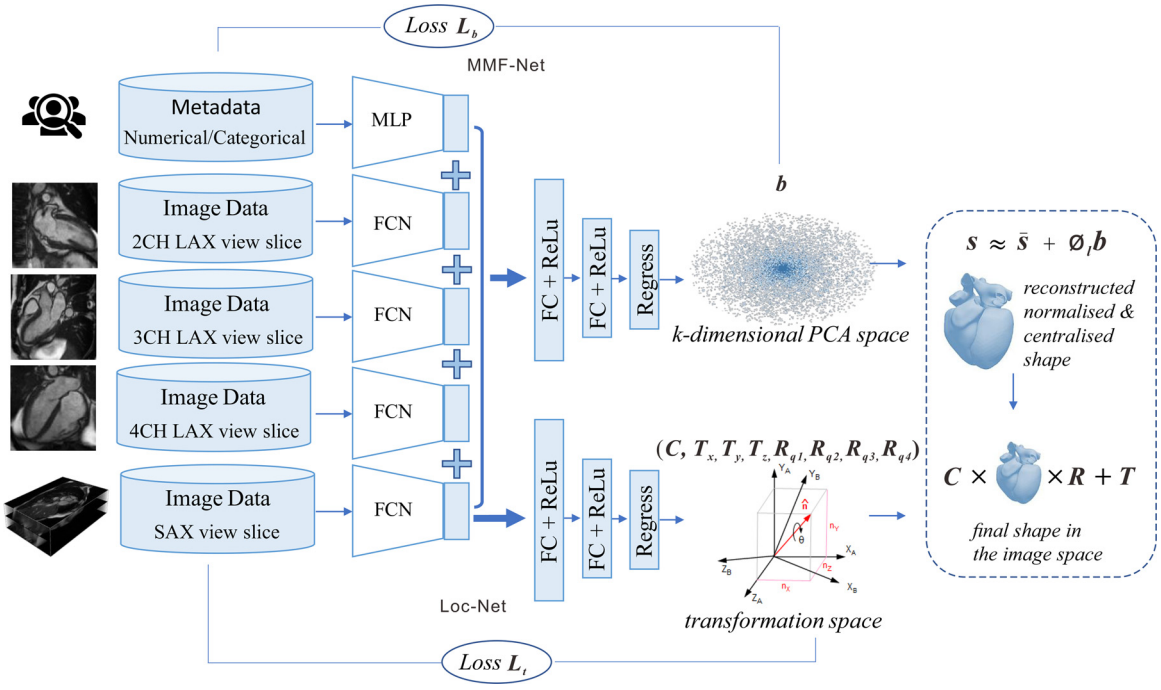

Fig. 4. MCSI-Net consists of two subnets: MMF-Net and Loc-Net. The MMF-Net extracts a high-level representation of the image from SAX and LAX views using four FCNs and concatenates the image features together with the output of an MLP network applied to the patient data. The MMF-Net finally produces the k-dimentional parameters in PCA space, whereas the Loc-Net predicts the corresponding transformation parameters, with which the 3D cardiac shape can be predicted..

We set the weights [1, 100, 100] where the network optimises the sub-branches of translation, rotation and scaling.

Ultimately, having the mean shape, eigenvectors and predicted shape parameters, the final shape can be reconstructed using Eq. (5) and Eq. (2). Once the network is trained, it is capable of segmenting all five sub-parts of the heart in different views simultaneously. While other state-of-the-art approaches have been proposed to segment the heart in multi-view CMR images, to the best of our knowledge, none of them exploit the contextual shape information provided by each view, to enhance cardiac shape reconstruction in 3D, as done by MCSI-Net.

As shown in Fig. 4, we used FCNs to extract image features. In this study, we looked into recent deep learning networks that have demonstrated the best performance in regression problems (Lathuilière et al., 2019), namely Alex-Net (Krizhevsky et al., 2012), VGG-16 (Simonyan and Zisserman, 2014), DenseNet (Huang et al., 2017), and ResNet (He et al., 2016).

All these FCNs are available in their standard architecture, in a Pytorch package called *torchvision*, and are used here for the purpose of feature extraction. We added an extra convolutional layer with the kernel size equal to the feature map dimensions to produce a vector of features. We examined the performance of the above mentioned FCNs in our application, and found that the ResNet outperforms the other architectures. We computed the segmentation accuracy using three key metrics: Dice Similarity Coefficient ($\mathcal{D}$), Mean Contour Distance ($\mathcal{M}$) and Hausdorff Distance ($\mathcal{H}$). $\mathcal{D}$ is between 0 and 1, with a higher $\mathcal{D}$ indicating a better match between the two shapes. $\mathcal{M}$ and $\mathcal{H}$ measure the mean and maximum distance, respectively, between the manual and automatic results, with a lower value indicating a better the agreement. These metrics are defined in Eqs. (10), (11) and (12), respectively, in Section 3.4. Table 1 summarises the performance of each architecture investigated within our framework. This indicates that the best performing FCN is ResNet, which appears at the right-most column of the table. It consistently achieves the largest overlap and lowest point-to-point distance, across all sub-parts of the heart.

Table 1
Comparison of shape prediction accuracy using different FCNs in terms of $\mathcal{D}$, $\mathcal{M}$(mm) and $\mathcal{H}$(mm) for LV endo-/epicardium, RV, LA and RA. The results were assessed at both ED and ES phases. **Bold** indicates the best performing method.

		Alex-Net	VGG-16	DenseNet	ResNet
LV_{endo}	$\mathcal{D}$	0.84 ± 0.07	0.85 ± 0.07	0.85 ± 0.07	**0.88 ± 0.05**
	$\mathcal{M}$	2.51 ± 1.00	2.28 ± 0.94	2.02 ± 0.98	**1.86 ± 0.79**
	$\mathcal{H}$	6.09 ± 2.26	5.63± 2.10	5.13 ± 2.18	**4.74 ± 1.75**
LV_{myo}	$\mathcal{D}$	0.69 ± 0.11	0.72 ± 0.09	0.76 ± 0.09	**0.78 ± 0.08**
	$\mathcal{M}$	2.55 ± 1.03	2.27 ± 0.95	2.04 ± 0.91	**1.86 ± 0.82**
	$\mathcal{H}$	6.17 ± 2.22	5.69 ± 2.05	5.36 ± 2.21	**4.75 ± 1.76**
RV_{endo}	$\mathcal{D}$	0.81 ± 0.07	0.82 ± 0.07	0.85 ± 0.06	**0.85 ± 0.06**
	$\mathcal{M}$	3.05 ± 1.09	2.98 ± 1.12	2.46 ± 1.24	**2.27 ± 0.95**
	$\mathcal{H}$	8.78 ± 2.94	8.77 ± 3.20	7.47 ± 3.55	**7.06 ± 2.64**
LA_{2ch}	$\mathcal{D}$	0.82 ± 0.09	0.85 ± 0.07	**0.87 ± 0.07**	0.86 ± 0.07
	$\mathcal{M}$	3.67 ± 1.46	3.12 ± 1.10	2.96 ± 1.68	**2.87 ± 1.21**
	$\mathcal{H}$	8.95 ± 3.67	7.65 ± 3.48	**7.11 ± 3.62**	7.19 ± 2.83
LA_{4ch}	$\mathcal{D}$	0.83 ± 0.08	0.85 ± 0.07	0.87 ± 0.08	**0.88 ± 0.07**
	$\mathcal{M}$	3.54 ± 1.39	3.11 ± 1.35	2.89 ± 1.59	**2.65 ± 1.11**
	$\mathcal{H}$	8.61 ± 3.05	7.89 ± 3.64	7.11 ± 3.10	**6.82 ± 2.47**
RA_{4ch}	$\mathcal{D}$	0.83 ± 0.11	0.84 ± 0.10	0.86 ± 0.06	**0.88 ± 0.06**
	$\mathcal{M}$	3.57 ± 1.58	3.42 ± 1.32	2.95 ± 1.33	**2.79 ± 1.23**
	$\mathcal{H}$	8.57 ± 3.46	8.47 ± 3.20	7.73 ± 3.34	**7.34 ± 2.99**

Thus, we used ResNet as the deep feature extractor within our approach, for all subsequent experiments conducted throughout the study.

As shown in Fig. 4, we also used an MLP to learn features from patient data and integrate it with the extracted image-based features. The MLP is a two-hidden-layer function denoted $g(\boldsymbol{\eta})$, where, $\boldsymbol{\eta}$ is the input (patient data feature vector). It comprises two hidden layers (with 64 and 128 neurons), and an output layer (with 256 neurons). ReLU is used in the hidden and output layers as an activation function, such that:

$$g(\boldsymbol{\eta}) = ReLu\left(\boldsymbol{c}^{(2)} + \boldsymbol{w}^{(2)}\left(ReLu\left(\boldsymbol{c}^{(1)} + \ldots \boldsymbol{w}^{(1)}\left(ReLu\left(\boldsymbol{c}^{(0)} + \boldsymbol{w}^{(0)}\boldsymbol{\eta}\right)\right)\right)\right)\right) \tag{9}$$

where $\{\boldsymbol{c}^{(0)}, \boldsymbol{c}^{(1)}, \boldsymbol{c}^{(2)}\}$ and $\{\boldsymbol{w}^{(0)}, \boldsymbol{w}^{(1)}, \boldsymbol{w}^{(2)}\}$ denote biases and weights for the input and two hidden layers, respectively.

In the MMF-Net, the outputs of the five sub-networks are concatenated to construct one feature vector that contains the behavioural, phenotypic, and demographic information derived from the patient data in addition to appearance information from the imaging data. This information is fed into a fully connected layer, with ReLU activation functions, so that, by minimising Eq. (7), they produce the first k parameters in PCA space, which describe the 3D shape of the cardiac chambers. To capture 99.7% of shape variability in the training dataset, we set $k = 70$ and regress only those parameters from randomly initialised weights. Similarly, in the Loc-Net, three different fully connected layers are applied to predict the scale (C), rotation ($R_{q1}, R_{q2}, R_{q3}, R_{q4}$) and translation ($T_x, T_y, T_z$) parameters, respectively.

3. Experiments and results

3.1. Data and annotations

We collected the UKB data under access applications number 2964 and 11350. This study complies with the Declaration of Helsinki. The study was covered by the ethical approval for UKB studies from the National Health Service National Research Ethics Service on 17th June 2011 (Ref 11/NW/0382) and extended on 10th May 2016 (Ref 16/NW/0274) with informed consent obtained from all participants. The full CMR protocol in the UKB has been described in detail elsewhere (Petersen et al., 2015). Researchers can apply to use the UKB resource for health-related research that is in the public interest[2].

We performed train/test experiments on CMR images of 4.525 subjects from the UKB using both end-diastolic (ED) and end-systolic (ES) time points. In terms of population sample size, experimental setup, and quality control, the most reliable reference annotations of cardiovascular structure and function found in the literature are those reported by Petersen et al. (2017a), in which CMR scans were manually delineated and analysed by a team of eight expert observers. These delineations were used to generate the reference 3D shapes, as explained in Section 2.1.

Regarding the image data, each slice (SAX and LAX views) was intensity- and spatially-normalised similar to our previous work (Attar et al., 2019a). After the pre-processing, every slice has a size of 128 x 128 px and intensity values ranging between 0 and 1. As the number of SAX slices in CMR images varies typically from 7 to 15 slices, the SAX image volumes were resampled to a fixed volume size of 15 slices, using cubic B-spline interpolation. Data augmentation with the random rotation and translation on the SAX and LAX images was applied when training the MMF-Net as it aims to produce the normalised and centralised shape models. In the training stage, we constructed the training set using the subject-specific reference whole heart shapes at two cardiac phases, i.e., from both ED and ES phases. Once the model was trained at ED and ES phases, we can infer not only the heart shape at the ED and ES phases but also any time point in the cardiac cycle.

With respect to the patient data, based on available literature and advice from our clinical collaborators, we selected a list of attributes that might directly/indirectly contribute to variations in cardiac morphology. Table 2 shows the summary of the patient data available for every subject in the UKB. All variables were scaled to the range [0, 1], including categorical variables, which were first indexed by grouping variables and then scaled (viz. *sex* $\in (0, 1)$, or *smoking status* $\in (0, 0.5, 1)$).

[2] https://www.ukbiobank.ac.uk/register-apply.

3.2. Competing methods

In addition to manual delineations and our previous work (Attar et al., 2019b), we compare the performance of MCSI-Net against the fully automated CNN-based method by Bai et al. (2018). In Table 3, we present the data we used for training, testing and evaluating different methods. Here, the proposed method MCSI-Net is labelled as C, our previous work (Attar et al., 2019b) is labelled as A, and the method by Bai et al. (2018) is labelled as B. As an additional assessment, we conducted a quantitative evaluation of human performance by measuring the inter-observer variability among the segmentations performed manually by three different clinical experts. Here, we randomly selected 50 subjects, where, each subject was analysed independently by three expert observers labelled O1, O2 and O3.

3.3. Implementation and training

The method was implemented using Python and Pytorch. The MMF-Net was trained using Adam for optimising the loss function (Eq. (7)) through 100 epochs with a learning rate of 0.00007 and batch size of 8 subjects. The Loc-Net was trained using Adam for optimising the loss function (Eq. (8)) through 300 epochs with a learning rate of 0.0001 and batch size of 32 subjects. Training took ~24 hours on Nvidia Tesla V100 GPUs hosted by Amazon Web Service and accessed using the MULTI-X platform (de Vila et al., 2018). At test time, it took about 5 seconds to predict the shape parameters of the full cardiac cycle.

3.4. Accuracy of predicted shapes

Fig. 5 shows several samples of our 3D cardiac shape results, at ED and ES phases, representing as corresponding 2D contours overlaid with the SAX and LAX images. It confirms that the system is capable of producing accurate shape parameters to generate shapes very similar to the reference contours.

We evaluated the performance of our approach using standard metrics for assessing segmentation accuracy. These include, the Dice index ($\mathcal{D}$), and the mean ($\mathcal{M}$) and Hausdorff distance ($\mathcal{H}$) between reference and predicted contours. Note that the manual annotation was performed for each slice independently so the contours are not globally consistent across slices in the presence of inter-slice motion (see Fig. 6). Thus, to better quantify the segmentation results, we used the target meshes, i.e., the reference 3D cardiac shapes used for the training of MCSI-Net, as the reference. $\mathcal{D}$ evaluates the overlap between automated segmentation $\mathbf{A}$ and manual segmentation $\mathbf{M}$; we define $\mathcal{D}$ as follows:

$$\mathcal{D} = \frac{2|\mathbf{A} \cap \mathbf{M}|}{|\mathbf{A}| + |\mathbf{M}|} \tag{10}$$

The $\mathcal{M}$ and $\mathcal{H}$ metrics evaluate the mean and maximum distance, respectively, between segmentation contours $\partial\mathbf{A}$ and $\partial\mathbf{M}$. These measures are defined as follows:

$$\mathcal{M} = \frac{1}{2|\partial\mathbf{A}|} \sum_{p \in \partial\mathbf{A}} d(p, \partial\mathbf{M}) + \frac{1}{2|\partial\mathbf{M}|} \sum_{q \in \partial\mathbf{M}} d(q, \partial\mathbf{A}) \tag{11}$$

$$\mathcal{H} = max(\max_{p \in \partial\mathbf{A}} d(p, \partial\mathbf{M}), \max_{q \in \partial\mathbf{M}} d(q, \partial\mathbf{A})) \tag{12}$$

where $d(p, \partial)$ denotes the minimal distance from point p to contour ∂. The lower the distance metric, the stronger the agreement.

Table 4 presents $\mathcal{D}$, $\mathcal{M}$ and $\mathcal{H}$ measures that compare automated segmentation results and reference. The mean and standard deviations of $\mathcal{D}$ for the LV_{endo}, LV_{myo}, RV_{endo}, LA_{2ch}, LA_{4ch} and RA_{4ch} with $n = 600$ are 0.88 ± 0.05, 0.78 ± 0.09, 0.85 ± 0.06, 0.86 ± 0.07, 0.88 ± 0.08 and 0.88 ± 0.06 respectively, indicating good agreement between the target mesh and automated segmentation with

Table 2
Summary of the patient data of 40,000 subjects used in this study. According to the nature of the data, we have grouped them into four categories, i.e. (1) demographics, (2) blood pressure, (3) biological samples, and (4) lifestyle. Continuous values indicate mean ± standard deviation.

Type	Patient data	Range
Demographics	Sex	Male / Female
	Age (years)	57 ± 8
	Height (cm)	170 ± 9
	Body mass index (kg/m2)	27 ± 4
	Weight (kg)	78 ± 16
	Body surface area (m2)	1.8 ± 0.2
	Ethnic background	White / Mixed / Other
Blood pressure	Systolic blood pressure (mmHg)	139 ± 19
	Diastolic blood pressure (mmHg)	82 ± 11
	Diastolic brachial blood pressure (mmHg)	69 ± 12
	Systolic brachial blood pressure (mmHg)	137 ± 22
	Pulse rate (bpm)	69 ± 12
	Central systolic blood pressure during PWA (mmHg)	134 ± 21
	End systolic pressure during PWA (mmHg)	113 ± 19
	Mean arterial pressure during PWA (mmHg)	96 ± 14
Biological samples	HbA1c (mmol/mol)	36.1 ± 6.7
	Cholesterol (mmol/L)	5.7 ± 1.1
	C-reactive protein (mg/L)	2.5 ± 4.3
	Glucose (mmol/L)	5.1 ± 1.2
	HDL Cholesterol (mmol/L)	1.4 ± 0.3
	IGF-1 (nmol/L)	21.3 ± 5.6
	LDL direct Cholesterol (mmol/L)	3.5 ± 0.8
	Triglycerides (mmol/L)	1.7 ± 1.0
Lifestyle	Smoking status	Never / Previous / Current
	Number of cigarettes smoked daily	15 ± 8
	Sleep duration (hours/day)	7 ± 1
	Duration of moderate activity (minutes/day)	66 ± 75
	Duration of vigorous activity (minutes/day)	44 ± 47
	Duration of walks (minutes/day)	61 ± 77
	Alcohol drinker status	Never / Previous / Current
	Alcohol intake frequency	- Never - Daily or almost daily - Three or four times a week - Once or twice a week - One to three times a month - Special occasions only

Table 3
Specific UKB datasets used for training and testing the methods proposed and presented in this paper.

Label	Method	Training/Tuning Data	Test Data
A	Attar et al. (2019b)	4275 subjects from SAX view	600 subjects
B	Bai et al. (2018)	Trained **separately** on: 4,275 subjects from SAX view 4,123 subjects from LAX 2CH view 4,082 subjects from LAX 4CH view	600 subjects
C	MCSI-Net	Trained **jointly** on: 3,925 subjects from SAX and LAX views	50 subjects 600 subjects
O1-O3	Three expert readers	Manual contours	50 subjects

Table 4
Segmentation results based on the different test sets (n=50 and 600), where we compare the target mesh with automated segmentation by our proposed MCSI-Net. LV_{endo} represents LV endocardium, LV_{myo} represents LV myocardium, RV_{endo} represents RV endocardium, LA_{2ch} represents LA in two chamber LAX view, LA_{4ch} represents RA in four chamber LAX view, and RA_{4ch} represents represents RA in four chamber LAX view. Table values are shown as mean ± standard deviation. The results were assessed at both ED and ES phases..

	(a) $\mathcal{D}$		(b) $\mathcal{M}$ (mm)		(c) $\mathcal{H}$ (mm)	
	Pred vs Target	Pred vs Target	Pred vs Target	Pred vs Target	Pred vs Target	Pred vs Target
Test-set	(n=50)	(n=600)	(n=50)	(n=600)	(n=50)	(n=600)
LV_{endo}	0.88 ± 0.05	0.88 ± 0.05	1.90 ± 0.86	1.86 ± 0.87	4.84 ± 1.89	4.74 ± 1.75
LV_{myo}	0.76 ± 0.08	0.78 ± 0.09	1.92 ± 0.81	1.86 ± 0.82	4.79 ± 1.69	4.75 ± 1.76
RV_{endo}	0.85 ± 0.06	0.85 ± 0.06	2.29 ± 0.95	2.27 ± 0.95	7.05 ± 2.57	7.06 ± 2.64
LA_{2ch}	0.86 ± 0.07	0.86 ± 0.07	2.77 ± 1.18	2.87 ± 1.21	7.32 ± 2.93	7.19 ± 2.83
LA_{4ch}	0.87 ± 0.05	0.88 ± 0.08	2.70 ± 1.21	2.65 ± 1.11	6.87 ± 2.51	6.82 ± 2.47
RA_{4ch}	0.88 ± 0.06	0.88 ± 0.06	2.76 ± 1.22	2.79 ± 1.23	7.12 ± 2.61	7.34 ± 2.99

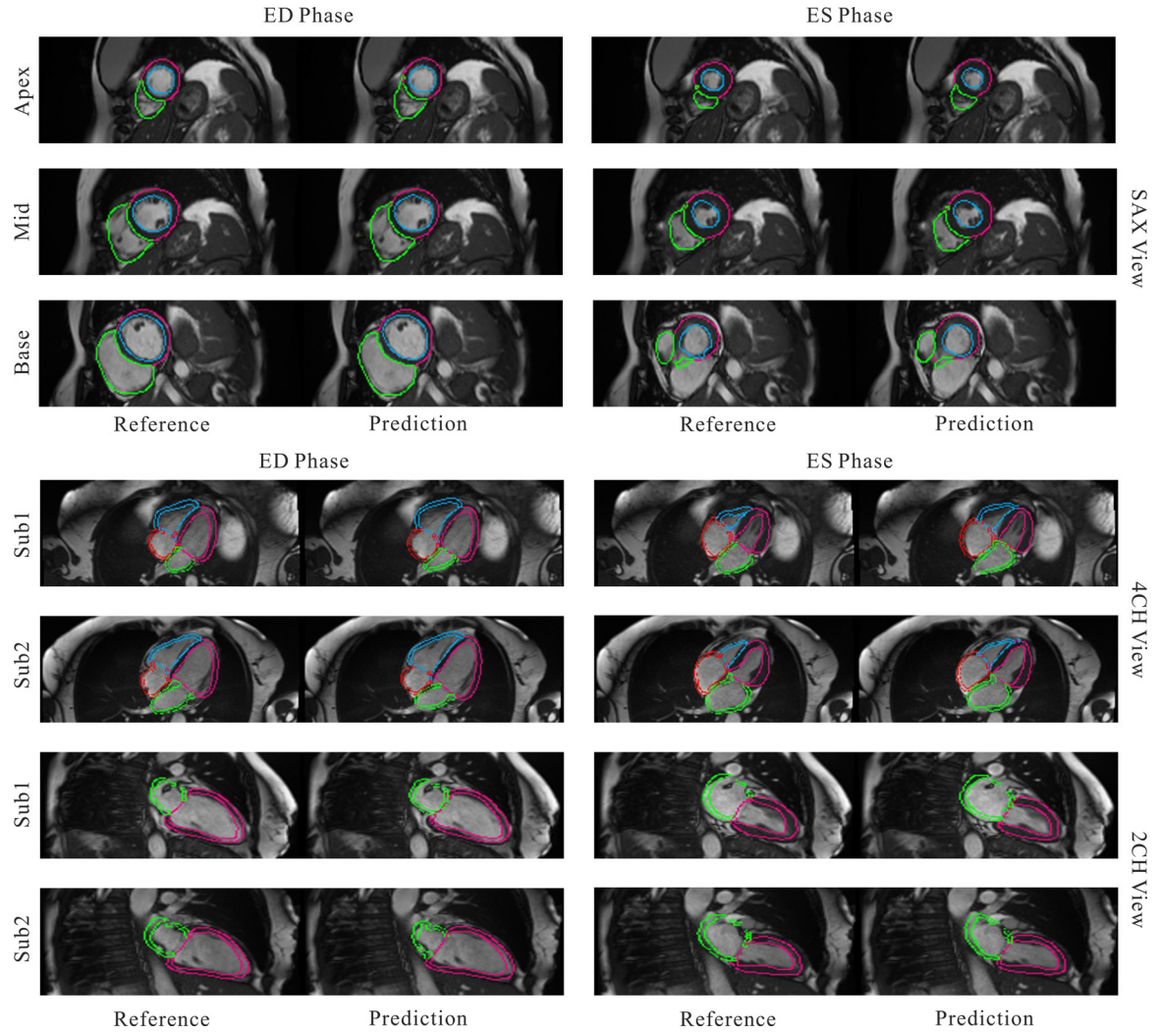

Fig. 5. Example segmentation results for short-axis and long-axis slices at the ED and ES phases illustrating the quality for automated segmentation contours versus the reference. The automated method segments all the time frames. However, only ED and ES frames are shown, as GTs are only available at the ED and ES frames. The CMR images were reproduced with the permission of UK Biobank..

the average $\mathcal{D}$ of 0.85. Due to the annular shape of LV_{myo} which has a larger perimeter (i.e. endo and epicardial boundary) and causes some overlap shifts, its $\mathcal{D}$ value is less than that of the other cases.

In addition, we observe that the $\mathcal{M}$ is 1.86 ± 0.87 mm for the LV_{endo}, 1.86 ± 0.82 mm for the LV_{myo}, 2.27 ± 0.95 mm for the RV_{endo}, 2.87 ± 1.21 mm for the LA_{2ch}, 2.65 ± 1.11 mm for the LA_{4ch}, and 2.79 ± 1.23 mm for the RA_{4ch}, all of which are close to the in-plane pixel spacing range of 1.8 to 2.3 mm. The $\mathcal{H}$ measures are 4.74 ± 1.75 mm, 4.75 ± 1.76 mm, 7.06 ± 2.64 mm, 7.19 ± 2.83 mm, 6.82 ± 2.47 mm and 7.34 ± 2.99 mm for the LV_{endo}, LV_{myo}, RV_{endo}, LA_{2ch}, LA_{4ch}, RA_{4ch}, respectively.

3.5. Accuracy of cardiac function indexes

Here, we report clinical cardiac functional indexes derived from manual and automated segmentation such as atrial and ventricular EDV and ESV. To reproduce the reference ranges reported by Petersen et al. (2017a), we first extract contours corresponding to the intersection between our 3D triangular meshes and CMR image slices. For the ventricular indexes calculated on SAX slices, we use the method of disks, whereby a cardiac 3D volume can be approximated by summing the areas within 2D segmentation contours and multiplying by the inter-slice spacing. For the atrial indexes, we calculated the volume according to the area-length method on LAX slices. Specifically, the LA volume was calculated using the biplane area–length method as $V = \frac{8}{3\pi}\frac{A_{\text{2ch}} \cdot A_{\text{4ch}}}{L}$, where A_{2ch} and A_{4ch} represent the atrial area on the 2 chamber and 4 chamber views, respectively, and L is the longitudinal diameter averaged across two views. Similarly, the RA volume was calculated using $V = \frac{8}{3\pi}\frac{A^2_{\text{4ch}}}{L}$.

We processed and quantified all the available 40,000 subjects of the UKB; each imaged at 50 time points, i.e. in total two million image volumes. Fig. 7 illustrates the mean and standard deviation of LV mass and the ventricular and atrial volumes of all subjects at each time point. The volumes trend matches well with the physiology of the heart where ventricles have the highest blood volume at ED, and lowest at ES phase. At the same time, the atria have the opposite behaviour where, it contains the lowest blood volume at ED and highest at ES phase. LV mass remains consistent over the cardiac cycle, which is apparent in the diagram.

Having quantified cardiac chamber volumes across all 50 time points, we report the following sets of indexes: the LV end-diastolic volume (LVEDV) and end-systolic volume (LVESV), LV stroke volume (LVSV), LV ejection fraction (LVEF), LV myocardial

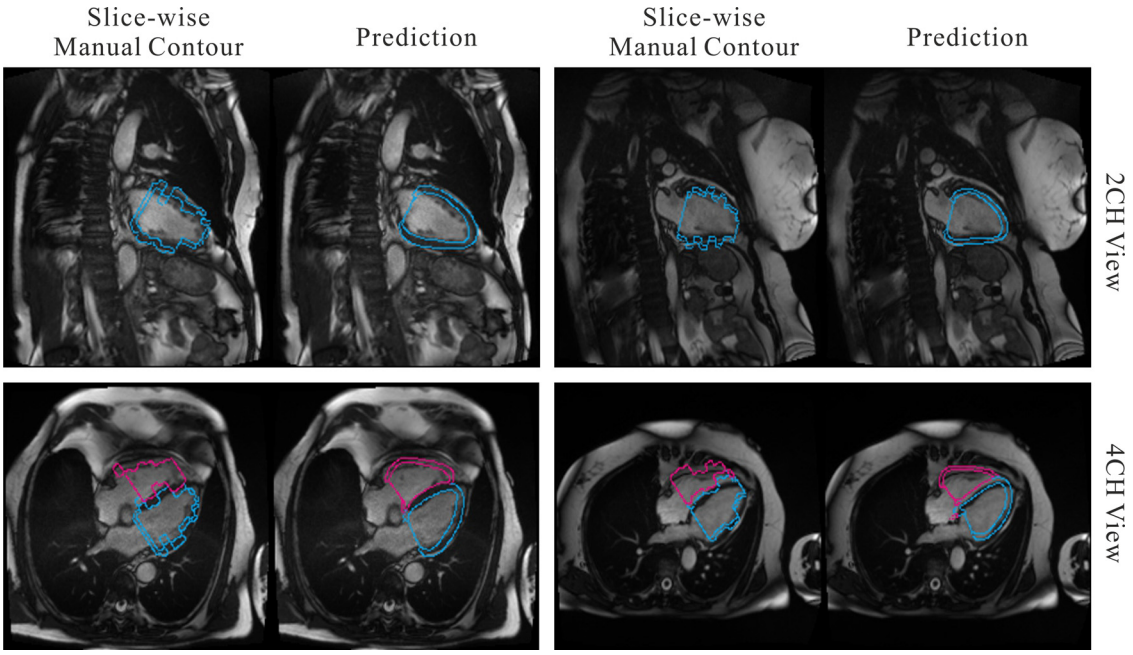

Fig. 6. Example segmentation results for short-axis slices overlayed on the long-axis views illustrating the quality of our automated segmentation contours versus manual contours. Note that the manual annotation was performed for each slice independently so the contours are not globally consistent in the presence of inter-slice motion, whereas our approach can generate globally consistent and plausible shapes and contours of heart. The CMR images were reproduced with the permission of UK Biobank. .

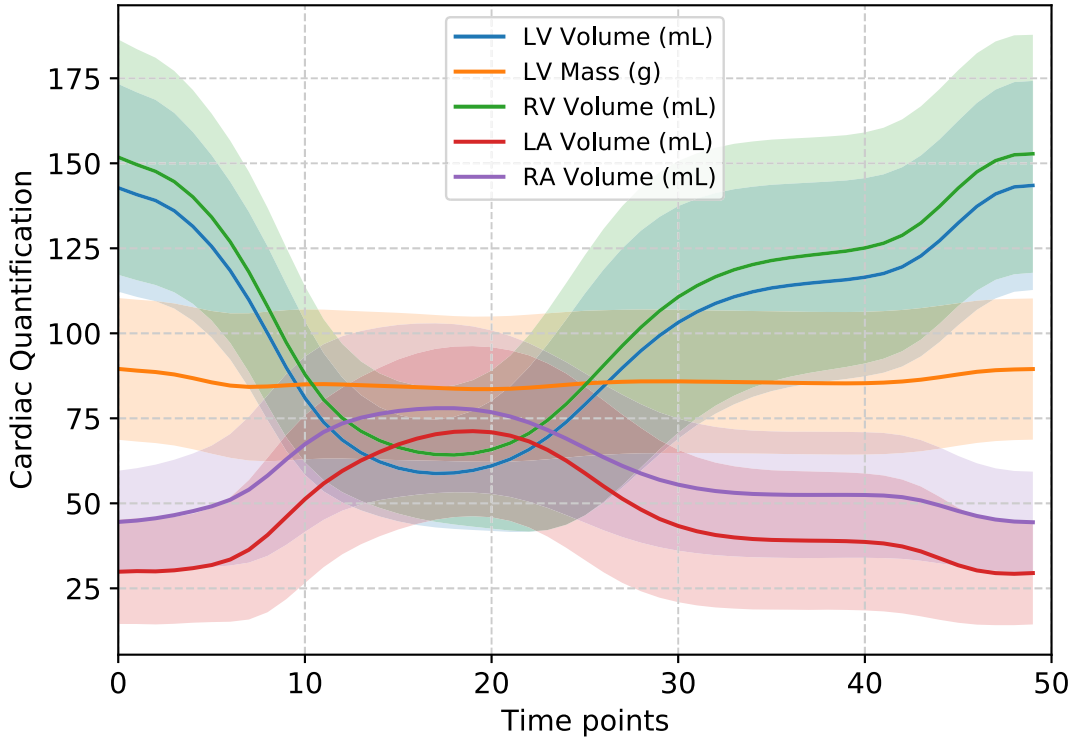

Fig. 7. Quantification results of **2 million** CMR image volumes. Illustrating the mean (solid lines) and standard deviation (shaded area) of the ventricular and atrial volume (ml), LV mass (g) of 40k subjects of the UKB over the 50 time points.

mass (LVM), RV end-diastolic volume (RVEDV) and end-systolic volume (RVESV), RV stroke volume (RVSV), RV ejection fraction (RVEF), LA end-diastolic volume (LAEDV) and end-systolic volume (LAESV), LA stroke volume (LASV), LA ejection fraction (LAEF), RA end-diastolic volume (RAEDV) and end-systolic volume (RAESV), RA stroke volume (RASV), and RA ejection fraction (RAEF). Table 5 shows the main aforementioned cardiac functional indexes, with the first column representing the indexes derived from all the available manual segmentation. We observe here that there is excellent agreement between the clinical indexes of our proposed method and the ground truth values for all the indexes listed in the table.

Table 6 presents the relative differences between the automated and manual measurements, as well as between the automated and

Table 5
Summarising the differences in clinical measures derived from our proposed method and manual segmentation. Here, GT represents the ground-truth values provided by manual segmentation from (Petersen et al., 2017a). A, B, and C represents the quantification results derived from automated segmentation by Attar et al. (2019b), Bai et al. (2018), and our proposed MCSI-Net, respectively. Values indicate mean ± standard deviation.

Indexes	GT (n=4,875)	GT (n=600)	A (n=600)	B (n=600)	C (n=600)
LVEDV (ml)	144 ± 34	143 ± 34	144 ± 32	148 ± 35	145 ± 32
LVESV (ml)	59 ± 20	58 ± 19	57 ± 18	62 ± 20	59 ± 17
LVSV (ml)	85 ± 19	85 ± 20	87 ± 20	86 ± 20	86 ± 19
LVEF (%)	59 ± 6	60 ± 6	61 ± 7	58 ± 6	59 ± 5
LVM (g)	90 ± 25	88 ± 25	89 ± 25	91 ± 24	89 ± 20
RVEDV (ml)	153 ± 37	150 ± 39	153 ± 36	156 ± 39	152 ± 36
RVESV (ml)	67 ± 22	65 ± 23	63 ± 20	68 ± 21	62 ± 20
RVSV (ml)	85 ± 19	85 ± 20	90 ± 23	88 ± 22	90 ± 20
RVEF (%)	56 ± 6	57 ± 7	59 ± 8	57 ± 6	59 ± 6
LAEDV (ml)	29 ± 14	29 ± 14	NA	29 ± 16	29 ± 11
LAESV (ml)	67 ± 21	67 ± 21	NA	67 ± 23	67 ± 19
LASV (ml)	39 ± 11	40 ± 11	NA	39 ± 11	38 ± 11
LAEF (%)	59 ± 9	59 ± 8	NA	58 ± 10	58 ± 10
RAEDV (ml)	46 ± 19	46 ± 18	NA	49 ± 21	45 ± 15
RAESV (ml)	79 ± 26	79 ± 26	NA	82 ± 28	74 ± 19
RASV (ml)	33 ± 11	33 ± 13	NA	34 ± 13	31 ± 11
RAEF (%)	42 ± 11	42 ± 10	NA	41 ± 11	41 ± 10

manual measurements computed by different expert human observers. We observe here that the relative differences for the two subsets of 50 and 600 subjects matched well, and the difference range was either very close or slightly larger, to the difference ranges obtained by the different expert observers. Due to large variations in the right ventricle appearance near the basal slices at the ES phase, the ventricular contours are more complex and thus lead to more ambiguous, inconsistent manual annotations amongst the human observers and consequently more difficult to segment with automated segmentation methods. We also looked into the automated values from a built-in automated segmentation software of the scanner device (i.e. *inlineVF D13A*), which was only available for the LV. Overall, A, B, and C performed substantially better than *inlineVF*, achieving an improvement of nearly 45% for all metrics. Note that these data were retrieved for every subject from the central UKB database.

Next, in Fig. 8 and 9, we present Bland-Altman plots and correlation plots of the main ventricular and atrial parameters computed using our approach and the manual reference covering 600 test subjects. The Bland-Altman plot is commonly used for analysing agreement and bias between two measurements. In Fig. 8, we have reported the mean difference (i.e. bias) and limits of agreement (LoA), i.e. ± 1.96 standard deviations from the mean.

Table 7
Example comparison of cardiac function on large-scale datasets using automatically derived clinical indexes of 1600 subjects. Illustration of cardiac remodelling on healthy subjects vs subjects with myocardial infarction (MI).

	Healthy	MI	*p*-value
LVEDV (ml)	143 ± 31	157 ± 32	< 0.001
LVESV (ml)	59 ± 17	67 ± 22	< 0.001
LVM (g)	89 ± 21	102 ± 21	< 0.001
RVEDV (ml)	150 ± 35	163 ± 32	< 0.001
RVESV (ml)	62 ± 20	67 ± 18	< 0.001
LAEDV (ml)	29 ± 12	38 ± 16	< 0.001
LAESV (ml)	67 ± 20	82 ± 24	< 0.001
RAEDV (ml)	46 ± 13	49 ± 17	< 0.001
RAESV (ml)	76 ± 20	79 ± 25	=0.108

The Bland-Altman plots show strong agreement and a mean difference line at nearly zero, suggesting that the clinical indexes estimated using our approach have little bias.

Fig. 9 presents correlation plots comparing the manual and automated methods, for different cardiac function indexes. The correlation coefficient (corr) measures the strength of the relationship between two sets of observations. The strength and direction of the relationship indicate the predictive power of our framework. Coefficients for ventricle measurements, such as LVEDV, LVESV, LVM, RVEDV, RVESV ranged between 0.87 and 0.93, indicating a strong relationship between our approach and the manual delineations. The atrium measurements were slightly worse, ranged between 0.76 and 0.86.

Furthermore, Fig. 10, illustrates the Q-Q plots of the same indexes on the test of 600 subjects, confirming the agreement on different quantiles over all the indexes, where we see the quantiles points roughly forming a straight line.

3.6. Sub-Cohort analysis

The proposed approach enables us to perform large-scale population-wide analysis of CMR images, demonstrated thus far on data from the UKB, with statistics on the most commonly used clinical indexes derived from CMR exams. We believe, however, that the power of population studies lies in the opportunity to define and characterise human sub-populations, and investigate the patterns and trends across different sub-populations.

Table 7 compares the ventricular and atrial volume and LV mass, derived from our approach, between two groups of subjects, namely, healthy subjects (n=800), and subjects with myocardial infarction (MI) (n=800). The table shows that MI is associated with increased ventricular and atrial volume and LV mass with

Table 6
The difference in clinical measures between the automatic and manual segmentations, as well between measurements by different human observers. M: ground truth provided by manual segmentation (Petersen et al., 2017a). A, B, and C represents the clinical indexes derived from the automated segmentation by Attar et al. (2019b), Bai et al. (2018), and our proposed MCSI-Net, respectively. Values indicate mean ± standard deviation. .

Relative difference (%)									
	O1 vs O2 (n=50)	O2 vs O3 (n=50)	O3 vs O1 (n=50)	A vs M (n=50)	B vs M (n=50)	C vs M (n=50)	A vs M (n=600)	B vs M (n=600)	C vs M (n=600)
LVEDV	4.2 ± 3.1	6.3 ± 3.3	3.4 ± 2.2	4.2 ± 3.0	2.9 ± 3.6	5.7 ± 4.5	4.7 ± 3.3	4.1 ± 3.5	7.4 ± 5.9
LVESV	6.8 ± 7.5	12.5 ± 8.5	11.7 ± 5.1	10.2 ± 8.1	12.5 ± 11.2	12.1 ± 11.1	9.3 ± 9.4	9.5 ± 9.5	11.8 ± 10.1
LVM	4.4 ± 3.3	6.0 ± 3.7	6.7 ± 4.6	6.5 ± 4.1	8.0 ± 4.8	9.3 ± 7.1	8.3 ± 7.7	8.3 ± 7.6	11.3 ± 9.3
RVEDV	8.0 ± 5.0	4.2 ± 3.1	5.7 ± 3.6	7.3 ± 4.2	5.7 ± 4.3	11.3 ± 10.6	5.4 ± 4.7	5.6 ± 4.6	8.1 ± 6.4
RVESV	30.6 ± 15.5	10.9 ± 8.3	16.9 ± 9.2	22.0 ± 8.4	29.8 ± 22.1	17.0 ± 6.9	12.4 ± 9.0	11.8 ± 12.2	12.5 ± 10.3
LAEDV	13.3 ± 9.8	11.0 ± 7.8	7.8 ± 5.3	NA	13.1 ± 10.0	6.1 ± 4.2	NA	13.1 ± 10.3	6.8 ± 6.9
LAESV	6.6 ± 5.4	7.3 ± 5.6	6.0 ± 5.0	NA	7.1 ± 6.9	10.7 ± 8.6	NA	7.5 ± 6.8	11.6 ± 10.5
RAEDV	6.5 ± 7.6	8.0 ± 4.9	8.4 ± 5.7	NA	12.1 ± 14.3	8.4 ± 5.3	NA	10.0 ± 18.7	8.6 ± 7.3
RAESV	4.7 ± 6.4	5.2 ± 6.4	4.9 ± 4.3	NA	7.7 ± 6.6	9.8 ± 10.1	NA	7.4 ± 6.9	11.8 ± 11.3

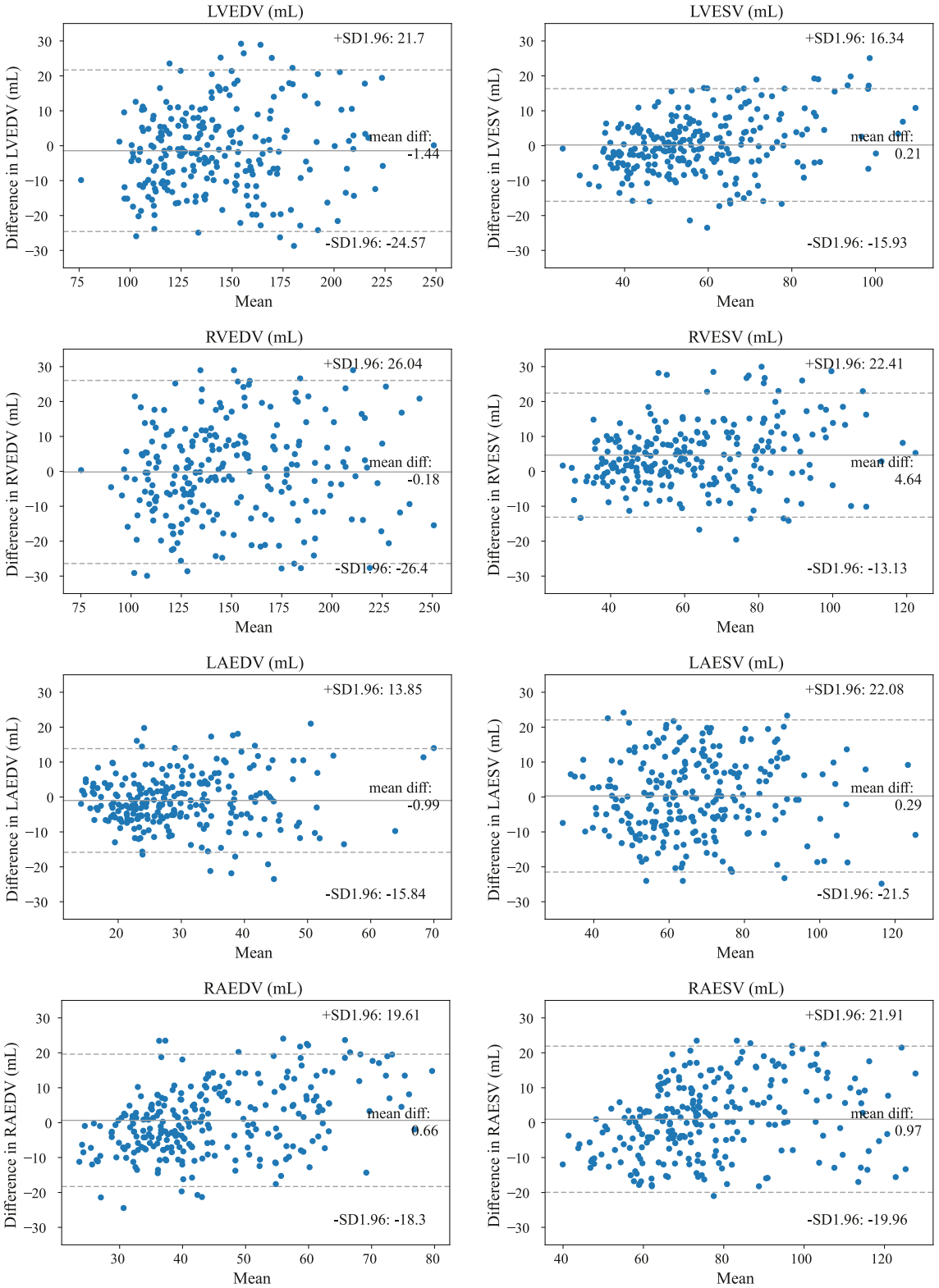

Fig. 8. Illustrating the repeatability of various cardiac functional indexes comparing the manual and automated analysis of 600 subjects from the UKB cohort; Bland-Altman plots for various cardiac functional indexes computed both manually and automatically in which manual segmentation was available. The solid horizontal lines denote the bias or the mean difference (Automatic - Manual), whereas the two dashed lines denote limits of agreement (LoA) i.e. ± 1.96 standard deviations from the mean..

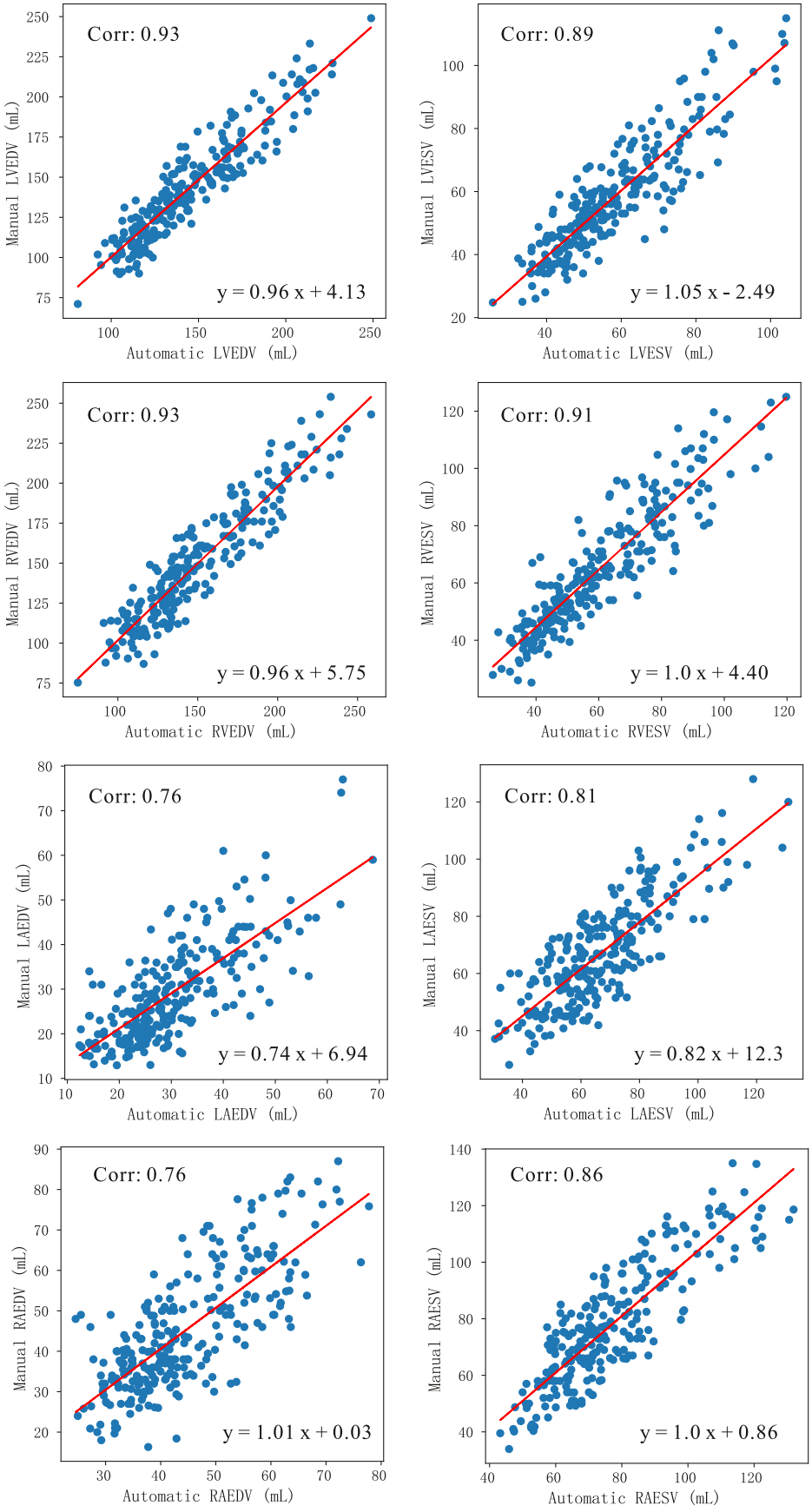

Fig. 9. Illustrating the repeatability of various cardiac functional indexes comparing the manual and automated analysis of 600 subjects from the UKB cohort; Correlation plots for various cardiac functional indexes computed both manually and automatically in which manual segmentation was available. .

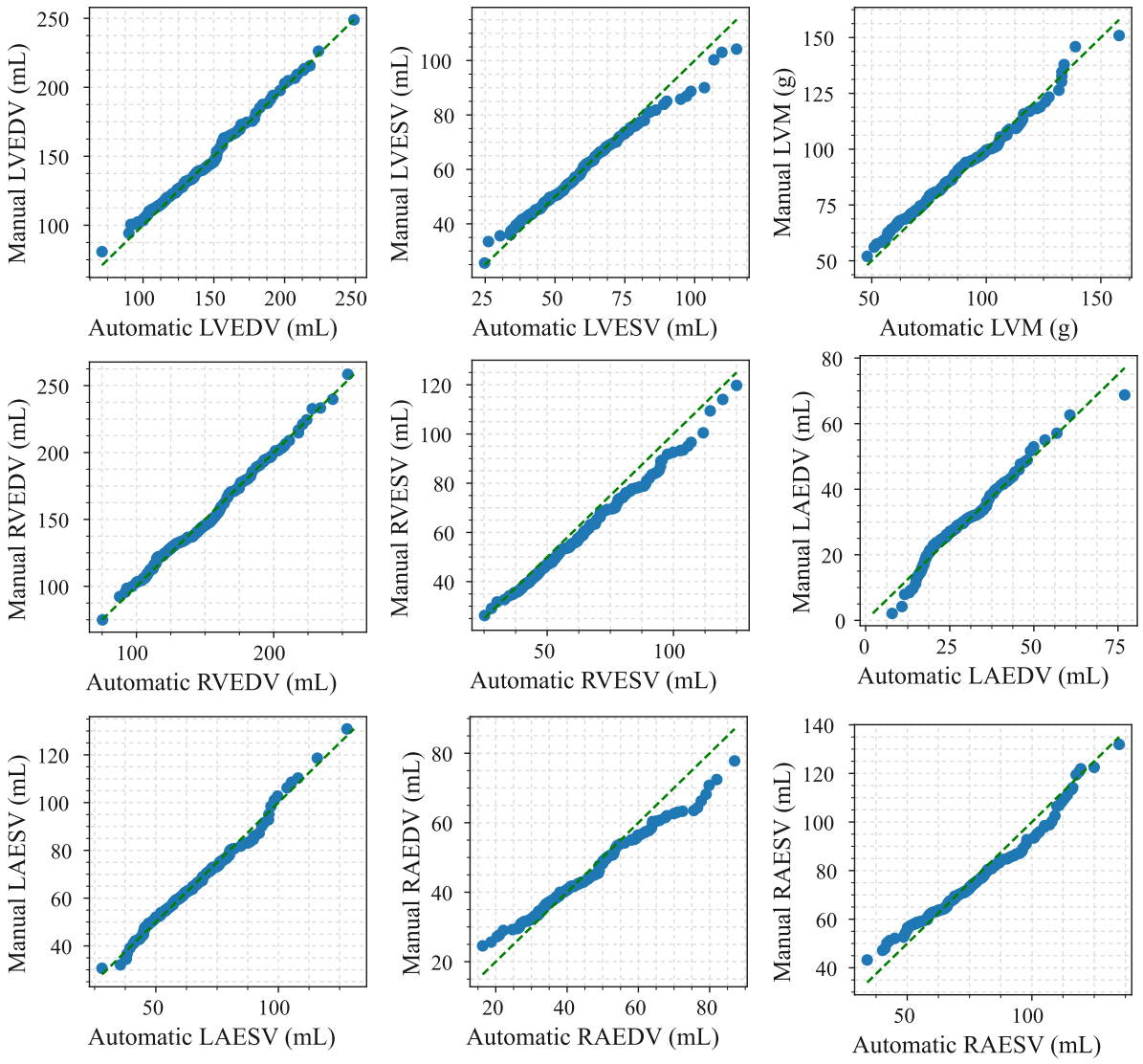

Fig. 10. Q-Q plots for various cardiac functional indexes computed both manually and automatically of 600 subjects..

statistical significance (p-value< 0.001). This is consistent with previous findings, from manual analysis of a few dozens of subjects, where cardiac remodelling and enlargement happens in patients with MI. Now we can confirm similar findings and perform similar studies, with the advantage of reliable automated analysis on really large-scale datasets such as UKB, where, there is tons of data available to discover different associations and patterns.

3.7. Impact of patient data on shape accuracy

Several studies (Petersen et al. (2017b), Gilbert et al. (2019) and Jensen et al. (2019), just to name a few) have shown the correlation between baseline features (such as lifestyle and demographics) and cardiac morphology and structure.

In this subsection, we report the effect of including patient data (as shown in Table 2) as additional information in our network. As expected, the use of patient data alongside image information improves the network, leading to more accurate prediction in all cardiac substructures.

Table 8 summarises the effect of including patient data on our method. The first column shows the accuracy metrics of $\mathcal{D}$, $\mathcal{M}$ and $\mathcal{H}$ when we only used the image data (labelled as IMG). The sec-

Table 8
Comparison of shape prediction accuracy using only images (IMG) or images with different type of patient data (IMG+MD) in terms of $\mathcal{D}$, $\mathcal{M}$(mm) and $\mathcal{H}$(mm) for LV endo-/epicardium and RV, LA, and RA on 600 UKB subjects. **Bold** indicates the best performing method..

		IMG	IMG+MD
LV_{endo}	$\mathcal{D}$	0.87 ± 0.05	**0.88 ±0.05**
	$\mathcal{M}$	1.90 ± 0.81	**1.86 ±0.79**
	$\mathcal{H}$	4.77 ± 1.77	**4.74 ±1.75**
LV_{myo}	$\mathcal{D}$	0.77 ± 0.09	**0.78 ±0.08**
	$\mathcal{M}$	1.89 ± 0.85	**1.86 ±0.82**
	$\mathcal{H}$	4.79 ± 1.78	**4.75 ±1.76**
RV_{endo}	$\mathcal{D}$	0.85 ± 0.06	**0.85 ±0.06**
	$\mathcal{M}$	2.29 ± 0.95	**2.27 ±0.95**
	$\mathcal{H}$	**7.03 ±2.54**	7.06 ± 2.64
LA_{2ch}	$\mathcal{D}$	0.86 ± 0.08	**0.86 ±0.07**
	$\mathcal{M}$	2.86 ± 1.21	**2.87 ±1.21**
	$\mathcal{H}$	7.24 ± 2.95	**7.19 ±2.83**
LA_{4ch}	$\mathcal{D}$	0.87 ± 0.07	**0.88 ±0.07**
	$\mathcal{M}$	2.65 ± 1.12	**2.65 ±1.11**
	$\mathcal{H}$	6.87 ± 2.53	**6.82 ±2.47**
RA_{4ch}	$\mathcal{D}$	0.87 ± 0.06	**0.88 ±0.06**
	$\mathcal{M}$	**2.78 ±1.21**	2.79 ± 1.23
	$\mathcal{H}$	7.43 ± 3.06	**7.34 ±2.99**

Table 9
Comparison of shape prediction accuracy using different set of images i.e. only SAX images, only LAX images, only three main SAX images (BMA: base, mid, and apex) or combination of SAX and LAX images in terms of $\mathcal{D}$, $\mathcal{M}$(mm) and $\mathcal{H}$(mm) for LV endo-/epicardium and RV, LA, and RA on 600 UKB subjects.

		BMA	SAX	LAX	BMA+LAX	SAX+LAX
LV_{endo}	$\mathcal{D}$	0.85 ± 0.07	0.87 ± 0.05	0.78 ± 0.10	0.85 ± 0.07	0.88 ± 0.05
	$\mathcal{M}$	2.62 ± 1.21	1.93 ± 0.83	3.45 ± 1.59	2.34 ± 1.11	1.86 ± 0.79
	$\mathcal{H}$	5.83 ± 2.55	5.13 ± 1.97	7.98 ± 3.34	5.77 ± 2.42	4.74 ± 1.75
LV_{myo}	$\mathcal{D}$	0.72 ± 0.12	0.76 ± 0.09	0.59 ± 0.14	0.72 ± 0.11	0.78 ± 0.08
	$\mathcal{M}$	2.45 ± 1.28	1.97 ± 0.95	3.56 ± 1.76	2.38 ± 1.21	1.86 ± 0.82
	$\mathcal{H}$	6.01 ± 2.59	5.31 ± 1.98	8.25 ± 3.54	5.91 ± 2.57	4.75 ± 1.76
RV_{endo}	$\mathcal{D}$	0.83 ± 0.06	0.85 ± 0.06	0.72 ± 0.12	0.83 ± 0.07	0.85 ± 0.06
	$\mathcal{M}$	2.77 ± 1.25	2.34 ± 0.98	4.53 ± 1.76	2.63 ± 1.06	2.27 ± 0.95
	$\mathcal{H}$	8.01 ± 3.03	7.11 ± 2.78	12.5 ± 4.65	7.94 ± 2.98	7.06 ± 2.64
LA_{2ch}	$\mathcal{D}$	0.82 ± 0.08	0.84 ± 0.08	0.78 ± 0.11	0.83 ± 0.08	0.86 ± 0.07
	$\mathcal{M}$	3.62 ± 1.54	3.26 ± 1.47	4.42 ± 1.96	3.46 ± 1.53	2.87 ± 1.21
	$\mathcal{H}$	8.96 ± 3.79	8.21 ± 3.81	10.3 ± 3.98	8.09 ± 3.39	7.19 ± 2.83
LA_{4ch}	$\mathcal{D}$	0.83 ± 0.10	0.86 ± 0.08	0.80 ± 0.10	0.85 ± 0.08	0.88 ± 0.07
	$\mathcal{M}$	3.43 ± 1.55	3.26 ± 1.47	4.30 ± 1.93	3.12 ± 1.29	2.65 ± 1.11
	$\mathcal{H}$	8.48 ± 3.78	7.20 ± 2.51	10.4 ± 2.14	7.49 ± 2.60	6.82 ± 2.47
RA_{4ch}	$\mathcal{D}$	0.84 ± 0.08	0.86 ± 0.08	0.80 ± 0.10	0.85 ± 0.09	0.88 ± 0.06
	$\mathcal{M}$	3.57 ± 1.73	2.92 ± 1.44	4.78 ± 2.16	3.19 ± 1.61	2.79 ± 1.23
	$\mathcal{H}$	9.09 ± 4.08	7.81 ± 3.58	11.2 ± 4.55	8.03 ± 3.96	7.34 ± 2.99

ond column report the accuracy of the predicted shapes using patient data. We observe that using the patient data alongside the image data had a positive impact on the shape accuracy, achieving improvement compared with the first column (p-value< 0.001), where we only use the image data. This can be attributed to the combined use of image and patient data within a single network to directly predict shape parameters. It is intuitive that patient data provides subject-specific prior information on the heart shape and including this information is helpful particularly for shape parameter estimation (B-value) that approximates the patient' s normalised and centralised cardiac shape.

3.8. Contribution of slices on shape accuracy

We further investigated generating accurate 3D cardiac shapes when only a fraction of the image data is available. This was considered highly relevant to this study for the following reasons: (1) Despite all the advantages of CMR scans, there is an explicit limitation of acquisition time. A routine CMR scan takes 20–60 minutes, which is time-consuming and expensive, especially in environments where resources are strained and availability of scan time is limited. Additionally, such long scan times may be infeasible entirely, for certain niche groups of patients. Besides, CMR often requires breath-holds, which could be not easy for many patients. Accordingly, it is essential to decrease the acquisition time, while maintaining the quality of subsequent quantitative analyses. Consequently, accurate cardiac quantification using fewer CMR slices, would help facilitate the acceleration of CMR scan acquisitions, improving patient comfort and reducing expenses. (2) Typical artefacts found during CMR image acquisitions are missing slices, which result in missing contours from 2D segmentation methods (such as the previously mentioned by Bai et al. (2018)) across the heart and, as a result, sub-optimal estimation of volumetric indexes. Consequently, a 3D cardiac shape generation framework that is robust to the presence of such variations would be of significant clinical value as it would enable accurate quantification of cardiac functional indexes, despite the presence of such artefacts.

To address these issues, in addition to the default image data set where we use all the available SAX and LAX image slices, we used incomplete samples to train our network, and subsequently, predicted complete cardiac shapes. We extended our training dataset by adding new cases where the image data in SAX and LAX are not all present, i.e. retaining only

- the three Basal, Mid and Apical slices (labelled as BMA) without any LAX slices,
- SAX slices (on average ten SAX slices without any LAX slices), and
- the three LAX view slices (two-, three and four-chamber view) without any SAX slices.

Using this process, we generated three new samples, from every sample in the original dataset. To keep the network architecture, input format, and structure consistent across all experiments, slices were excluded during training by replacing them with an empty slice, i.e. with zero values for all pixels.

Table 9 summarises the shape prediction accuracy of our approach for all cardiac structures, compared with the manual delineations, where, we do not use all the available image data. As expected, we observe that the most accurate results are obtained when we use all the available image data, reported in the rightmost column of the table.

The first column shows the accuracy metrics of different cardiac structures when we only use Basal, Mid and Apical slices (BMA). We see the accuracy of our approach is satisfactory (average $\mathcal{D}$ of 0.85, 0.72, 0.83 for LV endo, LV myo and RV endo, and 0.83, 0.84 for LA 4ch and RA 4ch) using three SAX slices that do not cover the atria. It is encouraging to see that the network can still predict the atrial shapes to a reasonable degree of accuracy based on the minimal cues available in the three SAX slices.

The second column presents the results associated with the SAX slices only. We see that the results become better in terms of both ventricles and atria segmentation accuracy when the network has access to more image data. Note, on average the number of slices is at least three times higher than the BMA case, but the improvement is about 2% on average. The third column shows the results of segmentation when three LAX images are used, and we get much worse results for the ventricles and atria.

We see the addition of the three LAX slices to the three BMA slices (labelled as BMA+LAX), results in higher accuracy compared with BMA and LAX, as the visual cues exist to describe both ventricles and atria regions, covering the sagittal and transverse representation of the heart.

Overall, we see the potential strength of this method to produce accurate and high-resolution 3D meshes of cardiac shape using only a fraction (about 23% to 46%) of image data.

4. Discussion and conclusions

In this study, we presented a fully automatic method capable of performing high-throughput end-to-end 3D cardiac MR image analysis via the simultaneous use of images and patient data. We validated our workflow on a reference cohort of 600 subjects for which both manual delineations and reference functional indexes exist, and the full dataset with 40,000 subjects. We also presented the positive impact of the inclusion of patient data on the accuracy of the predicted shapes. Furthermore, we showed that the proposed method achieves comparable results using just a fraction of image data, which supports the idea of accelerating the CMR scan acquisition by capturing significantly fewer image slices. Note that, although we also investigated the results of discarding the LAX images as input and observed that the network can still predict the atrial shapes to a reasonable degree of accuracy, presumably by scaling them with the ventricles, it may yield less satisfactory results for the patients who might have relative enlargement of their atria. As future work, we would like to increase the robustness of our method to handle severe morphological variations due to pathology and variable image quality, and extend its generalisability to cope with other CMR sequences and image modalities. Specifically, to tackle inaccurate segmentation results on exceptional cases, the following directions can be considered. First, an adversarial discriminator can be introduced into the framework to classify whether the predicted shape masks are generated or real and encourages the network to capture object structure implicitly, which could be particularly beneficial for realistic shape reconstruction. In addition, adversarial discriminator provides a powerful 3D shape descriptor, which learned without supervision, can be used to criticize the reconstruction quality (e.g., 3D heart shape being plausible or not). Second, slice-wise quantification of cardiac morphological indices such as myocardial thickness could identify whether there is a smooth transition in the reconstructed shape masks from the adjacent slices. Large changes in these values in the segmentation results might indicate incorrect prediction. We foresee including 3D+t quality control methods to assess the 3D shape accuracy over the full cardiac cycle. In addition to increasing the robustness of our system, we would like to extend our analyses to establish reference ranges for different sub-populations. Furthermore, future work could also look at modelling the relationship between cardiac morphology and function, and other available demographics, genetic, and omics data, which would help improve our understanding of disease progression, potentially leading to improved treatment planning and delivery.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Yan Xia: Conceptualization, Methodology, Visualization, Validation, Formal analysis, Writing – original draft. **Xiang Chen:** Conceptualization, Methodology, Validation, Formal analysis, Writing – review & editing. **Nishant Ravikumar:** Conceptualization, Methodology, Visualization, Writing – review & editing. **Christopher Kelly:** Formal analysis, Writing – review & editing. **Rahman Attar:** Conceptualization, Methodology, Validation, Writing – original draft. **Nay Aung:** Data curation. **Stefan Neubauer:** Data curation. **Steffen E. Petersen:** Data curation. **Alejandro F. Frangi:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition.

Acknowledgements

This research has been conducted using the UK Biobank Resource under Application 11350. The authors are grateful to all UK Biobank participants and staff. AFF acknowledges support from the Royal Academy of Engineering Chair in Emerging Technologies Scheme (CiET1819/19), EPSRC-funded Grow MedTech CardioX (POC041), and the MedIAN Network (EP/N026993/1) funded by the Engineering and Physical Sciences Research Council (EPSRC).

References

Attar, R., Pereanez, M., Bowles, C., Piechnik, S.K., Neubauer, S., Petersen, S.E., Frangi, A.F., 2019. 3D Cardiac Shape Prediction with Deep Neural Networks: Simultaneous Use of Images and Patient Metadata. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 586–594.

Attar, R., Pereañez, M., Gooya, A., Albà, X., Zhang, L., Piechnik, S.K., Neubauer, S., Petersen, S.E., Frangi, A.F., 2018. High Throughput Computation of Reference Ranges of Biventricular Cardiac Function on the UK Biobank Population Cohort. In: International Workshop on Statistical Atlases and Computational Models of the Heart. Springer, pp. 114–121.

Attar, R., Pereañez, M., Gooya, A., Albà, X., Zhang, L., de Vila, M.H., Lee, A.M., Aung, N., Lukaschuk, E., Sanghvi, M.M., et al., 2019. Quantitative CMR population imaging on 20,000 subjects of the UK biobank imaging study: LV/RV quantification pipeline and its evaluation. Med. Image Anal. 56, 26–42.

Bai, W., Sinclair, M., Tarroni, G., Oktay, O., Rajchl, M., Vaillant, G., Lee, A.M., Aung, N., Lukaschuk, E., Sanghvi, M.M., et al., 2018. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. J. Cardiovascul. Magnetic Resonance 20 (1), 65.

Duan, J., Bello, G., Schlemper, J., Bai, W., Dawes, T.J., Biffi, C., de Marvao, A., Doumou, G., O'Regan, D.P., Rueckert, D., 2019. Automatic 3D bi-ventricular segmentation of cardiac images by a shape-refined multi-task deep learning approach. IEEE Trans. Med. Imag.

Gilbert, K., Bai, W., Mauger, C., Medrano-Gracia, P., Suinesiaputra, A., Lee, A.M., Sanghvi, M.M., Aung, N., Piechnik, S.K., et al., 2019. Independent left ventricular morphometric atlases show consistent relationships with cardiovascular risk factors: a UK biobank study. Sci. Rep. 9 (1), 1130.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700–4708.

Jensen, M.T., Fung, K., Aung, N., Sanghvi, M.M., Chadalavada, S., Paiva, J.M., Khanji, M.Y., de Knegt, M.C., Lukaschuk, E., Lee, A.M., et al., 2019. Changes in cardiac morphology and function in individuals with diabetes mellitus: the UK biobank cardiovascular magnetic resonance substudy. Circulation: Cardiovascul. Imag. 12 (9), e009476.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 25.

Lathuilière, S., Mesejo, P., Alameda-Pineda, X., Horaud, R., 2019. A comprehensive analysis of deep regression. IEEE Trans. Pattern Anal. Mach. Intell..

Myronenko, A., Song, X., 2010. Point set registration: coherent point drift. IEEE Trans. Pattern. Anal. Mach. Intell. 32 (12), 2262–2275.

Painchaud, N., Skandarani, Y., Judge, T., Bernard, O., Lalande, A., Jodoin, P.-M., 2020. Cardiac segmentation with strong anatomical guarantees. IEEE Trans. Med. Imag..

Petersen, S.E., Aung, N., Sanghvi, M.M., Zemrak, F., Fung, K., Paiva, J.M., Francis, J.M., Khanji, M.Y., Lukaschuk, E., Lee, A.M., et al., 2017. Reference ranges for cardiac structure and function using cardiovascular magnetic resonance (CMR) in caucasians from the UK biobank population cohort. J. Cardiovascul. Magnet. Resonance 19 (1), 18.

Petersen, S.E., Matthews, P.M., Bamberg, F., Bluemke, D.A., Francis, J.M., Friedrich, M.G., Leeson, P., Nagel, E., Plein, S., Rademakers, F.E., et al., 2013. Imaging in population science: cardiovascular magnetic resonance in 100,000 participants of UK biobank - rationale, challenges and approaches. J. Cardiovascul. Magnetic Resonance 15 (1), 46.

Petersen, S.E., Matthews, P.M., Francis, J.M., Robson, M.D., Zemrak, F., Boubertakh, R., Young, A.A., Hudson, S., Weale, P., Garratt, S., et al., 2015. UK Biobanks cardiovascular magnetic resonance protocol. J. Cardiovascul. Magnetic Resonance 18 (1), 8.

Petersen, S.E., Sanghvi, M.M., Aung, N., Cooper, J.A., Paiva, J.M., Zemrak, F., Fung, K., Lukaschuk, E., Lee, A.M., Carapella, V., et al., 2017. The impact of cardiovascular risk factors on cardiac structure and function: insights from the UK biobank imaging enhancement study. PLoS ONE 12 (10), e0185114.

Ravikumar, N., 2017. A probabilistic framework for statistical shape models and atlas construction: Application to neuroimaging. University of Sheffield.

Ravikumar, N., Gooya, A., Frangi, A.F., Taylor, Z.A., 2017. Generalised coherent point drift for group-wise registration of multi-dimensional point sets. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 309–316.

Rodero, C., Strocchi, M., Marciniak, M., Longobardi, S., Whitaker, J., O Neill, M.D., Gillette, K., Augustin, C., Plank, G., Vigmond, E.J., et al., 2021. Linking statistical shape models and simulated function in the healthy adult human heart. PLoS Comput. Biol. 17 (4), e1008851.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition arXiv preprint arXiv:1409.1556.

de Vila, M.H., Attar, R., Pereanez, M., Frangi, A.F., 2018. MULTI-X, a state-of-the-art cloud-based ecosystem for biomedical research. In: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, pp. 1726–1733.

WHO, 2020. Cardiovascular Diseases. https://www.who.int/health-topics/cardiovascular-diseases/tab=tab_1/. Last checked on Jul 01, 2020. https://www.who.int/health-topics/cardiovascular-diseases/tab=tab_1/.