

Multitask Learning for Estimating Multitype Cardiac Indices in MRI and CT Based on Adversarial Reverse Mapping

Chengjin Yu, Zhifan Gao^{ID}, Member, IEEE, Weiwei Zhang^{ID}, Guang Yang^{ID}, Member, IEEE,
Shu Zhao^{ID}, Heye Zhang^{ID}, Yanping Zhang, and Shuo Li^{ID}

Abstract—The estimation of multitype cardiac indices from cardiac magnetic resonance imaging (MRI) and computed tomography (CT) images attracts great attention because of its clinical potential for comprehensive function assessment. However, the most exiting model can only work in one imaging modality (MRI or CT) without transferable capability. In this article, we propose the multitask learning method with the reverse inferring for estimating multitype cardiac indices in MRI and CT. Different from the existing forward inferring methods, our method builds a reverse mapping network that maps the multitype cardiac indices to cardiac images. The task dependencies are then learned and shared to multitask learning networks using an adversarial training approach. Finally, we transfer the parameters learned from MRI to CT. A series of experiments were conducted in which we first optimized the performance of our framework via ten-fold cross-validation of over 2900 cardiac MRI images. Then, the fine-tuned network was run on an independent data set with 2360 cardiac CT images. The results of all the experiments conducted on the proposed adversarial reverse mapping show excellent performance in estimating multitype cardiac indices.

Index Terms—Adversarial training, multitask learning, multitype cardiac indices, reverse mapping.

Manuscript received July 8, 2019; revised December 20, 2019 and February 25, 2020; accepted March 28, 2020. Date of publication April 16, 2020; date of current version February 4, 2021. This work was supported in part by the Guangdong Natural Science Funds for Distinguished Young Scholar under Grant 2019B151502031, in part by the Key Research and Development Program of Guangdong Province under Grant 2019B010110001, in part by the Key Program for International Cooperation Projects of Guangdong Province under Grant 2018A050506031, and in part by the National Natural Science Foundation of China under Grant 61771464, Grant U1801265, Grant U1908211, Grant 61876001, and Grant 61673020. (Corresponding authors: Heye Zhang; Yanping Zhang.)

Chengjin Yu, Shu Zhao, and Yanping Zhang are with the Key Laboratory of Intelligent Computing and Signal Processing, Ministry of Education, Anhui University, Hefei 230601, China, and also with the School of Computer Science and Technology, Anhui University, Hefei 230601, China (e-mail: chengjin.yuahu@gmail.com; zhaoshuzs2002@hotmail.com; zhangyp2@gmail.com).

Zhifan Gao and Heye Zhang are with the School of Biomedical Engineering, Sun Yat-sen University, Guangzhou 510275, China (e-mail: gaozhifan@gmail.com; zhangheye@mail.sysu.edu.cn).

Weiwei Zhang is with the School of Biomedical Engineering, Sun Yat-sen University, Guangzhou 510275, China (e-mail: mr.ww.zhang@gmail.com).

Guang Yang is with the National Heart and Lung Institute, Imperial College London, London SW3 6NP, U.K. (e-mail: g.yang@imperial.ac.uk).

Shuo Li is with the Department of Medical Biophysics, Western University, London, ON N6A 3K7, Canada (e-mail: slishuo@gmail.com).

This article has supplementary downloadable material available at <https://ieeexplore.ieee.org>, provided by the authors.

Color versions of one or more of the figures in this article are available online at <https://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TNNLS.2020.2984955

I. INTRODUCTION

MULTITYPE cardiac indices (see Fig. 1) estimation faces two challenges. Primarily, the complicated relationships between multitype cardiac indices are the main difficult (i.e., the mining and the representation of commonalities and differences among indices) for learning proper task dependencies. These indices have considerably different dimensions (i.e., 2-D cardiac segmentation, 1-D linear indices, and planar indices). Concerning specific indices, regional wall thicknesses (WT) also vary according to the orientation of the myocardium segments in different regions [1]–[4]. Furthermore, most of existing multitask methods can only work on one image modality because of the significantly different in appearance between magnetic resonance imaging (MRI) and computed tomography (CT). MRI can accurately define cardiovascular structures and characterize tissue composition [5], while short- and long-axis CT can be used for cardiac function analysis [6]. This diagnostic capability has attracted considerable attention because these MRI and CT techniques allow angiography to be performed noninvasively [1], [6]. However, MRI may not be used for patients with cardiac rhythm management devices for safety reasons [7]. In such circumstances, knowledge transfer from CT to MRI is necessary to obtain a better clinical evaluation.

Existing multitask learning methods may be unsuitable for estimating multitype cardiac indices in different modalities for two reasons. First, they are unable to learn the complex task dependencies among multitype cardiac indices.

- 1) For the assumption-based regularization methods, task dependencies are assumed to be known as the prior information, which includes a shared set of features [8], and shared cluster structures among tasks [9]. However, these assumptions are not always accurate or suited to all tasks.
- 2) Data-driven methods can learn task dependencies automatically from training data without any assumptions [10], [11]. However, most of these methods express task dependencies via a covariance matrix, which might be insufficient for 2-D and 1-D indices because the relationship between the indices cannot be represented directly owing to their different dimensions [12].

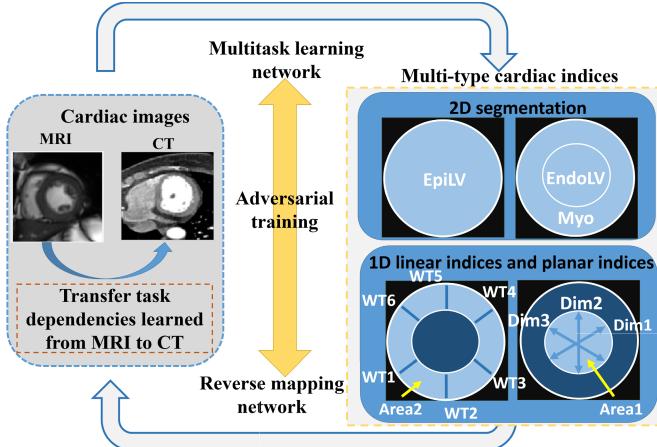


Fig. 1. Multitype cardiac indices are presented on the right [1]–[3], including 2-D cardiac segmentation: left ventricle epicardium (EpiLV), left ventricular endocardium (EndoLV), and left ventricular myocardium (Myo); 1-D linear indices: six regional WTs (WT₁, ..., WT₆) and 3-D cavity directions (Dim₁, Dim₂, and Dim₃); and 1-D planar indices: area of the cavity (Area₁) and the myocardium (Area₂). The complex relationships between these indices pose significant difficulties in learning the proper task dependencies. Two image modalities (MRI and CT) are shown on the left, and the significant difference in appearance poses great difficulties to knowledge transfer from MRI to CT. To address these difficulties, we investigated the problem from the perspective of reverse generation.

Second, existing multitask learning methods cannot efficiently transfer the task dependencies learned to other modalities. This is because most existing multitask learning methods only mine the relationships among tasks [12]–[14] but lack a mechanism for sharing the learned knowledge between different image modalities, especially for complex task dependencies.

To address the abovementioned challenges, a multitask learning method based on adversarial reverse mapping is proposed for estimating multitype cardiac indices in different imaging modalities (in this case, MRI and CT). The proposed method investigates multitask learning and the knowledge transfer problem from the perspective of a reverse generation that further learns the mapping from multitype cardiac indices to cardiac images via an adversarial training approach. This is loosely inspired by the recent progress of reverse mapping and adversarial training [15], [16], which shows the promising ability for modeling the data distribution, especially for assessing joint distributions of complex semantic variables [17]. As shown in Fig. 2, in addition to the multitask learning network that learns the mapping from cardiac image to multitype cardiac indices, we propose a reverse mapping network that further learns the reverse mapping to facilitate mining and the representation of task dependencies. Then, an adversarial training approach is applied to integrate the multitask learning network and the reverse mapping network. Finally, the parameters from the two networks learned from the source modality (MRI) are shared with the target modality (CT).

Specifically, our contributions also include the following.

- 1) We propose an efficient multitask learning framework based on adversarial reverse mapping that can obtain task dependencies and can also be applied to data in a different modality.

- 2) We propose a reverse mapping network that can facilitate the mining of task dependencies. This network reveals the role of each cardiac index and the dependencies among different cardiac indices from a generational perspective. Thus, the reverse mapping network can act as a regularization enabling the multitask network to learn task dependencies.
- 3) We propose a symmetric adversarial training for realizing the regularization of the reverse network to the multitask learning network. Learning occurs in this adversarial training, and the joint distribution from the reverse mapping network and the multitask network are matched. The convexity of the matching problem is then proven, which indicates that the reverse mapping network can effectively constrain the learning of the multitask network.
- 4) Furthermore, we propose a bidirectional parameter sharing scheme that shares the parameters learned from both the multitask learning network and the reverse mapping network to facilitate training on different modalities. Compared with parameters from one-way deep neural networks, our strategy is more powerful for representing complex task dependencies and is, therefore, more effective at sharing them.
- 5) Finally, our comprehensive experiments show promising results on estimating multitype cardiac indices from MRI and CT, which validates the feasibility of our adversarial reverse mapping-based multitask learning framework.

A. Estimation of Multitype Cardiac Indices

Existing cardiac indices estimation methods can be categorized into two classes: 1) segmentation methods and 2) regression methods.

1) Segmentation Methods: Existing segmentation-based method, Max Flow [18], initially performs cardiac segmentation and then computes other cardiac indices from the cardiac segmentation results. However, cardiac segmentation is still a challenge. Previously, cardiac segmentation methods were based on traditional techniques, such as region growing [19] and clustering [20], and energy minimization techniques, such as graph cuts [21]. Other model-based methods include the multivariate mixture model for cardiac segmentation from multisource MRI [22] and the statistical shape model-based methods [23]. Recently, deep neural networks have demonstrated powerful capability for cardiac image segmentation. FCN [24] is a basic method for biomedical image segmentation. GAN [25] improved the FCN based on adversarial training, and V-net [26] improved the FCN based on 3-D convolutional neural network (CNN) that is often used in 3-D image. Patravali *et al.* [27] used a 2-D CNN to implement slice-by-slice segmentation and a 3-D CNN for 2D + T MRI. In contrast to most existing 3-D methods, Qiao *et al.* [28] proposed a more advanced approach to obtaining 3-D context information using spatial propagation and have achieved promising results. Poudel *et al.* [29] and Alayba *et al.* [30] combined a 2-D CNN and recurrent neural

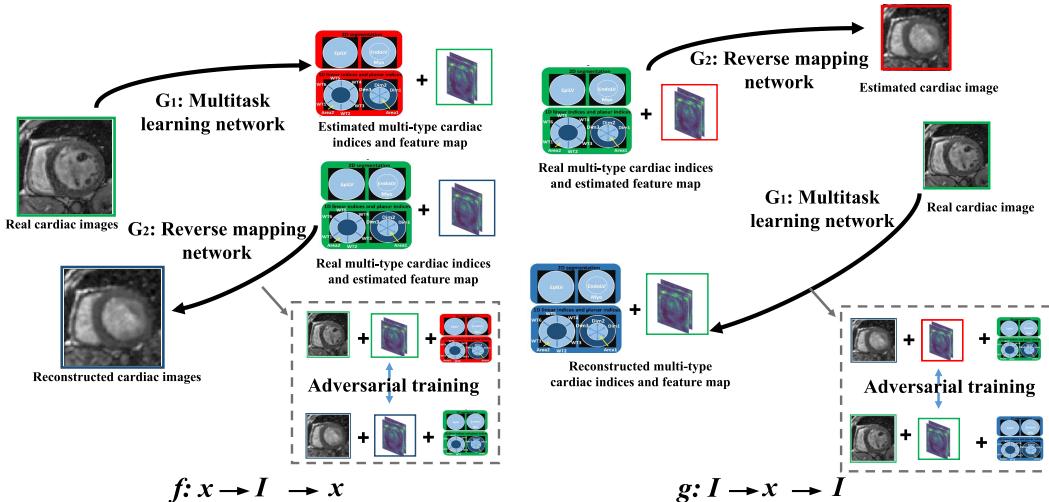


Fig. 2. Workflow of our framework. We integrate the multitask learning network and reverse mapping network in two ways. f : the multitask learning network first estimates the multitype cardiac indices I and generates a feature map z from a cardiac image x , and then, the reverse mapping network regenerates the cardiac image from the estimated feature map z and the real multitype cardiac indices. g : the reverse mapping first estimates z from the real I and then reconstructs x from z . The multitask learning network then estimates I and z from x . Finally, the discriminator networks try to distinguish the joint distributions sampled from the multitask learning network and the reverse mapping network.

network (RNN) to segment cardiac structures from 3-D MR images. Shi *et al.* proposed a 3-D CNN-based dilated residual network for whole heart segmentation [31]. Ahmed *et al.* [32] also used a CNN to segment cardiac structures from CT images.

However, for segmentation methods, additional computational cost and intermediate operations are required to calculate 1-D cardiac indices from segmentation results. In contrast, our method estimates 1-D cardiac indices and 2-D segmentation jointly. This is because the joint estimation directly models the relationship between image appearance and the cardiac indices, which allows the model to take further advantage of the dependencies between 1-D cardiac indices and 2-D segmentation. Therefore, our method provides more information more efficiently.

2) Regression Methods: Regression methods that calculate the cardiac indices directly from cardiac images instead of using cardiac segmentation [2], [33] have recently attracted a lot of attention. The benefit of regression methods is that they can estimate these indices directly from cardiac images more conveniently, efficiently, and accurately by directly modeling the relationship between the cardiac image appearance and cardiac indices. Therefore, regression methods reduce computational cost and avoid errors induced by intermediate operations that may be applicable in clinical applications [34]. Primarily, regression methods can be divided into two categories: two-phase methods and multitask learning methods. Most two-phase methods usually extract a feature representation and then employ a regression model to achieve the final estimation. The feature representation can include statistical features [35], meshfree representation [36], CNN-based feature representation [37], multifeatures composed of pyramidal Gabor features (Multifeature + RF) [38], steerable features [39] and a histogram of oriented gradients [40], the features extracted by an unsupervised multiscale convolutional deep belief network (MCDBN + RF) [41], CNN-based feature representation [42],

and a supervised descriptor (SDL + AKSF) [43]. Regression models include artificial neural networks [44], [45], random forests [46], K-cluster regression forests [47], and deep regression networks [48]. Regarding multitask learning methods, recent efforts are primarily based on deep multitask learning. Xue *et al.* [2] proposed a framework of joint representation and regression learning for extracting task-aware features (Indices-Net). Several other studies (FullLVNet [49] and DMTRL [3]) have also modeled the dependencies among different tasks and obtained promising results.

However, most existing methods focus only on MRI but pay no attention to CT, which is also important for clinical diagnosis. In contrast, the proposed method can estimate multitype cardiac indices from MRI and can also be transferred to CT.

B. Reverse Mapping and Generative Adversarial Networks

Reverse mapping aims to improve the learned model via learning the reverse process of the main task. An important research area is dual learning used for translation tasks [15]. It has two models: one model that translates English into French and a reverse model that translates French into English. During the training process, the two models would be complementary via the performance of dual tasks. Recently, reverse mapping has also been employed for multitask learning. A representative study is [50], which first predicts segmentation and tags via a multitask network and then regenerates original images from the predicted results via a reverse mapping network.

Generative adversarial networks (GANs) [16] show impressive results for data synthesis. In a standard GAN, the generator would try to generate fake samples to fool the discriminator, and the discriminator would try to distinguish between fake and real samples. In this game, if the discriminator cannot distinguish between fake and real samples correctly, we suppose that the generator would model the

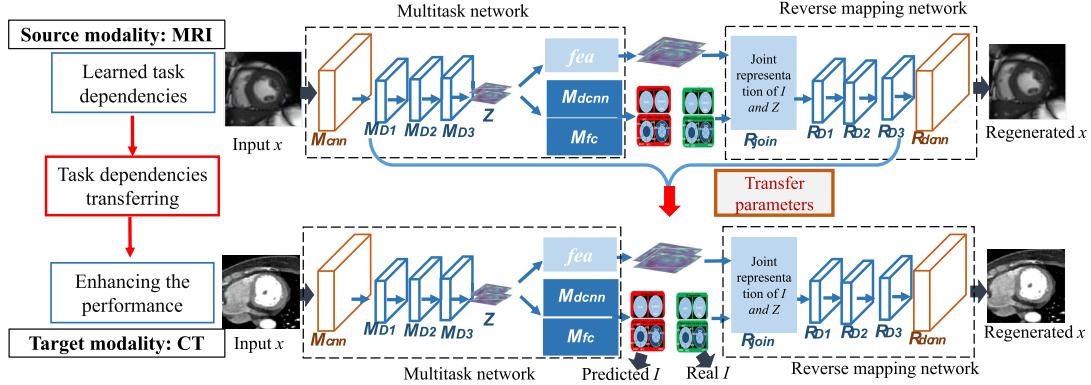


Fig. 3. Entire framework is first trained on the source modality (MRI). Then, the learned parameters (the layers colored by blue, i.e., M_{D1} , M_{D2} , M_{D3} , M_{dcnn} , M_{fc} , R_{join} , R_{D1} , R_{D2} , and R_{D3}) of both the multitask learning network and the reverse mapping network are transferred to the target modality (CT). Finally, we fine-tune the trained multitask learning network and reverse mapping network with labeled CT images.

data distribution accurately. Recently, some studies also take advantage of both GAN and reverse mapping (e.g., in realizing image translation). Dual GAN [50] employs two GANs to learn dual tasks. Cycle GAN [51] and Disco GAN [52] concatenate the two generators to ensure cyclic consistency.

II. METHOD

A. Overview of Our Method

Our proposed method uses a multitask learning framework with a reverse mapping network that reconstructs the image from learned features and multitype cardiac indices (see Section II-B), an adversarial training approach to solve multitask learning (see Section II-C), and a bidirectional parameter sharing mechanism (see Section II-D). Notably, we first inference a shared feature representation to predict multitype cardiac indices. Then, we also reconstruct the original images from the learned feature representation and predicted multitype cardiac indices, as illustrated in Fig. 2. Furthermore, we propose a discriminator network to distinguish between the two joint distributions from the multitask learning network and the reverse mapping network. In this case, we treat the discriminator network as a binary classification subtask in the multitask learning framework. We then explicitly learn the multitask relation. Finally, we transfer the learned parameters and multitask relation to different imaging modalities, as illustrated in Fig. 3.

B. Multitask Learning With Reverse Mapping for Estimating Multitype Cardiac Indices

Consider multitype cardiac indices $I = \{i_{seg}, i_{linear}, i_{planar}\}$, and i_{seg} , i_{linear} , and i_{planar} denote segmentation, linear indices, and planar indices, respectively. The prior distribution on the multitype cardiac indices is denoted by $p(I)$. Considering an observed cardiac image x , the prior distribution on the cardiac image is denoted by $q(x)$. Most existing methods aim to establish multitask learning mapping $G_1 : x \rightarrow I$ to accurately predict I . In this study, we further learn a reverse mapping $G_2 : I \rightarrow x$ that generates x from I . This reverse mapping establishes the mapping relationship from a reverse

perspective capable of revealing the complex relationships between x and I .

A common practice to learn G_1 and G_2 is maximum likelihood estimation based on parameterized conditional distributions $q_\varphi(I|x)$ and $p_\theta(x|I)$

$$G_1(x; \varphi) = \arg \max_I q_\varphi(I|x) \quad (1)$$

$$G_2(I; \theta) = \arg \max_x p_\theta(x|I) \quad (2)$$

where φ and θ denote the learned parameters of the two distributions. Apparently, G_1 and G_2 are reverse processes of each other. We aim to learn G_1 and G_2 jointly and enable them to complement each other. As shown in Fig. 2, we consider two ways to learn G_1 and G_2 .

- 1) We assume that X are the observable variables, thus concatenating G_1 and G_2 . Then, we have a function $f : \hat{X} \rightarrow \hat{X}$, defined by $f(x) = G_2(G_1(x))$ for all x in \hat{X} . In this case, G_1 first predicts I from the observed cardiac image x , and then, G_2 regenerates x from the predicted I .
- 2) We assume that \hat{I} are the observable variables, thus concatenating G_2 and G_1 . Then, we have another function $g : \hat{I} \rightarrow \hat{I}$, defined by $g(I) = G_1(G_2(I))$, for all I in \hat{I} . In this case, G_2 first predicts x from the multitype cardiac index I , and then, G_1 regenerates I from the predicted x . A mathematical description of the learning of f and g is as follows:

$$f(x; \theta, \varphi) = \arg \max_x p_\theta(x | \arg \max_I q_\varphi(I|x)) \quad (3)$$

$$g(I; \theta, \varphi) = \arg \max_I q_\varphi(I | \arg \max_x p_\theta(x|I)). \quad (4)$$

In previous studies, f and g are solved using an L_1 -based reconstruction loss that focuses on image translation but pays no attention to multitask problems [53]. The reverse mapping network reveals the role of each index (task) and dependence among different indices (tasks) from a generational perspective. It can act as a regularization enabling the multitask network to learn task dependencies via checking the cardiac image reconstructed from predicted cardiac indices. Notably, the reverse mapping network is trained to generate

corresponding cardiac images from given multitype cardiac indices. When inaccurate indices are predicted from a cardiac image, the reverse network would also generate an inaccurate cardiac image. Therefore, the reverse mapping network enables the multitask network to learn the correct task dependence by comparing the regenerated image with the real image. However, if f and g are well learned, the joint distribution $q_\phi(x, I) = q_\phi(I|x)q(x)$ learned by the multitask learning network and the joint distribution $p_\theta(x, I) = p_\theta(x|I)p(I)$ learned by the reverse mapping network would be matched. Then, we consider the Kullback–Leibler (KL) divergence [16] between $q_\phi(x, I)$ and $p_\theta(x, I)$

$$\mathcal{L}_f = \mathbb{E}_{q(x)} \log p_\theta(x) - \text{KL}(q_\phi(x, I) || p_\theta(x, I)) \quad (5)$$

$$\mathcal{L}_g = \mathbb{E}_{p(I)} \log q_\phi(I) - \text{KL}(p_\theta(x, I) || q_\phi(x, I)) \quad (6)$$

where $p_\theta(x) = \int p_\theta(x|I)p(I)dI$ such that it synthesizes samples that are well matched to those drawn from $q(x)$, and $q_\phi(I) = \int q_\phi(I|x)q(x)dx$ such that it synthesizes samples that are well matched to those drawn from $p(I)$. Therefore, (5) and (6) can be regarded as automatic encoders, which would be well learned if $q_\phi(x, I)$ and $p_\theta(x, I)$ were matched. However, it is difficult to directly learn (5) and (6) because the information entropy of I is lossy compared with that of x . To address this problem, we introduce a feature map z to increase the information entropy. On implementation, our reverse mapping reconstructs the cardiac image from z and I . Then, we can rewrite (5) and (6) as

$$\mathcal{L}_f = \mathbb{E}_{q(x)} \log p_\theta(x) - \text{KL}(q_\phi(x, z, I) || p_\theta(x, z, I)) \quad (7)$$

$$\mathcal{L}_g = \mathbb{E}_{p(I)} \log q_\phi(I) - \text{KL}(p_\theta(x, z, I) || q_\phi(x, z, I)). \quad (8)$$

Based on (7) and (8), our goal becomes learning and matching $q_\phi(x, z, I)$ and $p_\theta(x, z, I)$. In Section II-C, we employ an adversarial training approach to learn and match the two distributions simultaneously.

C. Adversarial Training Approach for Joint Distribution Matching

To effectively learn and match $q_\phi(x, z, I)$ and $p_\theta(x, z, I)$, we consider the sampling method shown in Fig. 2: For the function f , we can first sample (x, z, I) , where x is the real cardiac image, z is the feature map predicted by G_1 , and I is estimated by G_1 . Then, we can sample (x, z, I) , where x is the image reconstructed from z and the real I by G_2 . For the function g , we can first sample (x, z, I) , where x is a cardiac image reconstructed by G_2 , z is a feature map predicted by G_2 , and I is a real multitype cardiac index. Then, we can sample (x, z, I) , where x is the real image, and both z and I are predicted by G_1 . Because z is a feature representation of x , we do not consider the local relationship between x and z ; and then, let $X = (x, z)$ for convenience. We use $q(X)$ to denote the prior distribution of X . Then, we have $q_\phi(X, I) = q_\phi(I|X)q(X)$ and $p_\theta(X, I) = p_\theta(X|I)p(I)$. Furthermore, we have $p_\theta(X) = \int p_\theta(X|I)p(I)dI$ and $q_\phi(I) = \int q_\phi(I|X)q(X)dX$.

Then, we use two real-valued discriminator networks, $T_{\psi_1}(X, I)$ and $T_{\psi_2}(X, I)$, parameterized using ψ_1 and ψ_2 ,

respectively [54], to distinguish between joint distributions (x, I) . Precisely, we consider the following objectives:

$$\begin{aligned} & \max_{T_{\psi_1}} \mathbb{E}_{X \sim q(X)} \mathbb{E}_{I \sim q_\phi(I|X)} \log \sigma(T_{\psi_1}(X, I)) \\ & + \mathbb{E}_{X \sim p_\theta(X|I), I' \sim p(I)} \mathbb{E}_{I \sim p(I)} \log(1 - \sigma(T_{\psi_1}(X, I))) \end{aligned} \quad (9)$$

$$\begin{aligned} & \max_{T_{\psi_2}} \mathbb{E}_{I \sim p(I)} \mathbb{E}_{X \sim p_\theta(X|I)} \log \sigma(T_{\psi_2}(X, I)) \\ & + \mathbb{E}_{I \sim q_\phi(I|X'), X' \sim q(X)} \mathbb{E}_{X \sim q(X)} \log(1 - \sigma(T_{\psi_2}(X, I))). \end{aligned} \quad (10)$$

In (9) and (10), $\sigma(t) = (1 + e^{-t})^{-1}$ denotes the sigmoid function. Intuitively, $T_{\psi_1}(X, I)$ tries to distinguish between joint distributions (X, I) in f , while $T_{\psi_2}(X, I)$ tries to distinguish between joint distributions (X, I) in g .

If parameters θ and φ are given, we can rewrite (9) and (10) as

$$T_{\psi_1^*}(X, I) = \log q_\phi(X, I) - \log p_\theta(X)p(I) \quad (11)$$

$$T_{\psi_2^*}(X, I) = \log p_\theta(X, I) - \log q_\phi(I)q(X) \quad (12)$$

where $T_{\psi_1^*}(X, I)$ and $T_{\psi_2^*}(X, I)$ are defined as the functions that maximize (9) and (10), respectively. In this case, the parameters θ and φ can be optimized with ψ_1^* and ψ_2^* . Therefore, the complete object can be written as

$$\begin{aligned} & \min_{\theta, \varphi} \max_{\psi_1, \psi_2} \mathbb{E}_{q_\phi(X, I)} (-T_{\psi_1}(X, I) + \log p_\theta(X|I)) \\ & + \mathbb{E}_{p_\theta(X, I)} (-T_{\psi_2}(X, I) + \log q_\phi(I|X)). \end{aligned} \quad (13)$$

Algorithm 1 Multitask Learning With Adversarial Reverse Mapping

```

1:  $\varphi, \theta, \psi_1, \psi_2 \leftarrow$  Initialize parameters
2: repeat
3:    $\{x^1 \dots x^M\}, \{I^1 \dots I^M\} \leftarrow$  Random minibatch of  $M$  images and corresponding  $M$  multitype cardiac indices.
4:   Compute  $\varphi$  gradient:  

 $gr_\varphi \leftarrow \frac{1}{M} \sum_{k=1}^M \Delta \{-\log \sigma(T_{\psi_1}(x, I)) + \log p_\theta(x|I)\}$ 
5:   Compute  $\theta$  gradient:  

 $gr_\theta \leftarrow \frac{1}{M} \sum_{k=1}^M \Delta \{-\log \sigma(T_{\psi_2}(x, I)) + \log q_\phi(I|x)\}$ 
6:   Compute  $\psi_1$  gradient:  

 $gr_{\psi_1} \leftarrow \frac{1}{M} \sum_{k=1}^M \Delta \{\log \sigma(T_{\psi_1}(x, I)) + \log(1 - \sigma(T_{\psi_1}(x, I)))\}$ 
7:   Compute  $\psi_2$  gradient:  

 $gr_{\psi_2} \leftarrow \frac{1}{M} \sum_{k=1}^M \Delta \{\log \sigma(T_{\psi_2}(x, I)) + \log(1 - \sigma(T_{\psi_2}(x, I)))\}$ 
8:    $\varphi, \theta, \psi_1, \psi_2 \leftarrow$  Update parameters using gradients  $gr_\varphi, gr_\theta, gr_{\psi_1}, gr_{\psi_2}$  based on the SGD.
9: until convergence with parameters  $(\varphi, \theta, \psi_1, \psi_2)$ 
10: return  $\theta, \varphi, \psi_1, \psi_2$ 

```

Based on (12), we propose using Algorithm 1 to optimize the objects in (9) and (10) via the stochastic gradient descent (SGD). We regard (13) as a min–max game. As shown in Algorithm 1, we first update the parameters θ and φ while keeping ψ_1 and ψ_2 fixed. Then, we update ψ_1 and ψ_2 using the adversarial objects in (9) and (10).

Finally, if the Nash-equilibrium of the min–max game in (13) is achieved via parameters $\{\theta^*, \varphi^*, \psi_1^*, \psi_2^*\}$, we have $p_{\theta^*}(X, I) \approx q_{\varphi^*}(X, I)$, which indicates that the knowledge

learned by the reverse mapping network is shared to the multitask learning network. Furthermore, due to $X = (x, z)$, we can get $p_{\theta^*}(x, z, I) \approx q_{\phi^*}(x, z, I)$. This indicates that our method can learn the joint distribution $p_{\theta^*}(x, z, I)$ and $q_{\phi^*}(x, z, I)$ via adversarial training. Furthermore, this also shows that the proposed method can establish a one-to-one relationship between I and x as well as a one-to-one relationship between I and z . Furthermore, the information on x can be maintained in z to avoid steganography or the overfitting effect [55].

D. Bidirectional Parameter Sharing Scheme

Although the knowledge leveraged from MRI can reflect the common characteristics, such as cardiac structures and dependencies among multitype cardiac indices, shared in MRI and CT, additional effort is still required to represent and transfer such common characteristics.

In this section, we propose a new transfer learning scheme (as shown in Fig. 3), i.e., a bidirectional parameter sharing scheme that transfers the learned parameters from deep layers (colored blue) in both the multitask learning network and the reverse mapping network learned from MRI.

1) Structure of the Multitask Learning Network: We improved a dense convolutional network (DenseNet) [56] to fit our multitask learning network, which has one convolution layer (M_{cnn}) and three dense blocks with four, eight, and eight layers (M_{D1} , M_{D2} , and M_{D3}) (as shown in Fig. 3). Our model takes 2-D cardiac image slices across one cardiac cycle as input. Thus, a 3-D CNN is employed for all convolution layers because of its excellent capability in action modeling. For the first convolution layer, we use three different sizes with convolution kernel (3^3 , 5^3 , and 7^3) to generate hierarchical information for further learning. For the rest of the 3-D convolution layers in dense blocks, we use a (3^3) kernel. Then, we can obtain a feature map (z) output using the third dense block. Finally, one 3-D deconvolution-based pixel-level classifier (M_{dcnn}) is employed to generate 2-D segmentation results from z , and one fully connected network-based regression network M_{fc} is employed to estimate 1-D cardiac indices from z . Furthermore, the feature map z is also output by a branch fea. Notably, the outputs of M_{dcnn} and M_{fc} correspond to the 2-D segmentation and 1-D cardiac indices, respectively, in 2-D slices across one cardiac cycle.

2) Structure of the Reverse Mapping Network: Our reverse mapping network has a similar but reverse structure of the multitask learning network. It regenerates corresponding cardiac images from the given multitype cardiac indices. In particular, for f , we first employ both convolution layers and fully connected layers as a joint learning network (R_{join}) to learn a joint representation from z and the real I . For g , the joint representation is directly generated from the real I . Then, three dense blocks with eight, eight, and four layers (R_{D1} , R_{D2} , and R_{D3}) are employed for further learning. Finally, we reconstruct final cardiac images using deconvolution layers (R_{dcnn}).

In contrast to existing methods that only share the parameters from one-way network, our bidirectional parameter sharing scheme shares the parameters of the multitask learning

network and the reverse mapping network. The entire model is first trained on the MRI data using the adversarial training approach mentioned in Section II-B. Then, the parameters of layers M_{D1} , M_{D2} , M_{D3} , M_{dcnn} , and M_{fc} in the multitask learning network and layers R_{join} , R_{D1} , R_{D2} , and R_{D3} in the reverse mapping network are shared to CT images (as shown in Fig. 3, the shared layers are the blue parts of the networks). In this step, we concatenate new untrained layers M_{cnn} and R_{dcnn} with the shared layers. Then, we fine-tune the new network using labeled CT images. For the multitask learning network, we employ the Dice-based [26] loss L_{Dice} for 2-D cardiac segmentation and L_2 (ridge) [57] loss for 1-D cardiac indices. For the reverse mapping network, we employ the reconstruction loss $L_{\text{recon}} = \|x - G_2(G_1(x))\|_1$ [50]. Then, the total loss can be written as

$$L_{M\&R} = \lambda_1 L_{\text{Dice}} + \lambda_2 L_{\text{recon}} + \lambda_3 L_2. \quad (14)$$

It is important to note that only $L_{M\&R}$ is used for model fine-tuning on the CT data. For the model training on the MRI data, we use a joint loss $L_{\text{join}} = \lambda_4 L_{\text{Adv}} + \lambda_5 L_{M\&R}$ that combines the adversarial training loss mentioned in Section II-C and $L_{M\&R}$ loss mentioned in Section II-D.

III. EXPERIMENTS

A. Data Acquisition

Our data include one MRI data set and two CT data sets. The MRI data are collected from a public data set (Left Ventricle Full Quantification Challenge of MICCAI 2018). There are 2900 2-D short-axis cine MR images of 145 subjects (20 frames for each subject) collected from three hospitals affiliated with two healthcare centers (London Healthcare Center and St. Josephs HealthCare). The subjects' age is from 16 to 97 years, with an average of 58.9 years. The pixel spacings of the MR images range from 0.6836 to 2.0833 mm/pixel, with a mode of 1.5625 mm/pixel. The CT data sets are private, which includes one testing set with 360 images and one training set with 2000 images. All the images are 2-D short-axis CT images of 118 subjects (20 frames for each subject) collected from the Beijing Anzhen Hospital. The subjects' age is from 38 to 85 years, with an average of 60.4 years. The pixel spacing of these CT images range from 0.5627 to 2.1232 mm/pixel, with a mode of 1.6435 mm/pixel.

All cardiac images undergo several preprocessing steps, including landmark labeling, rotation, ROI cropping, and resizing. The resulted images are approximately aligned with dimensions of 80×80 . Then, these cardiac images are manually contoured to obtain the epicardial and endocardial borders, which are double-checked by four experienced cardiac radiologists (two for MRI and two for CT). The interobserver error ranged between 0.33% and 4.11% (average: 1.21%) for MRI and between 0.42% and 5.57% (average: 1.63% for CT). The ground-truth values of the 1-D cardiac indices can be obtained from the two borders. The values of WTs and cavity dimensions are normalized by the image dimension, while the areas are normalized by the pixel number (6400). During the evaluation, the obtained results are converted to

physical thickness (in mm) and area (in mm²) by reversing the resizing procedure and multiplying by the pixel spacing for each subject.

B. Experimental Settings and Evaluation Criteria

In our experiments, in order to test the performance of our method on multitype cardiac indices estimation, ten-fold cross-validation is employed for performance evaluation for all comparison and ablation studies in MRI data set, and the reported results are the average of ten-fold cross-validation. To test the performance on transfer learning from MRI to CT, the model is trained with all MRI data, and the CT data set with 360 images are used for fine-tuning and testing. From CT to MRI, the other CT data set with 2000 images are first used to train the model, and then, the randomly selected 360 MRI images are used for fine-tuning and testing. The training set and test set are separated to ensure that the data from one patient will not appear in both training and testing simultaneously. In particular, we implemented all of the codes using Python on a Linux (Ubuntu 16.04) desktop computer with an Intel Xeon CPU E5-2650 and 64-GB DDR2 memory. The graphics card used is an Nvidia GTX1080 GPU. The deep learning libraries are implemented using Tensorflow. The hyperparameters are determined by model performance, and we adopt an SGD with 0.9 momenta. The learning rate is given by $\eta_p = (\eta_0 / (1 + \alpha p)^\beta)$, where $\eta_0 = 0.01$, $\alpha = 10$, $\beta = 0.75$, and p is the training progress changing from 0 to 1 in proportion to the training step. Finally, we also balance the different losses used in the model: $\lambda_1 = \lambda_2 = 0.01$, and $\lambda_3 = \lambda_4 = \lambda_1 = \lambda_5 = 1.0$.

To measure the performance of the proposed method, we use two metrics for different cardiac indices. First, for the 2-D cardiac segmentation, the average Dice metric (Dice) is computed. Then, for the 1-D linear indices and planar indices, mean absolute error (MAE) is computed. The formal description $\text{Dice} = (2P_{TE}/P_T + P_E)$, where P_T denotes all pixels of manually segmented contour area and P_E denotes all pixels of automatically segmented contour area by different methods, while P_{TE} denotes the pixels of overlapping between P_T and P_E . Dice always ranges between [0, 1]. The larger Dice value is, the higher consistency between manual and automated segmentation. Then, $\text{MAE}(y, \hat{y}) = (1/N) \sum_{i=1}^N |y_i - \hat{y}_i|$, where $y \in R^N$ and $\hat{y} \in R^N$ are two vectors that denote true indices and estimated indices, respectively. To verify the effectiveness of multitask learning for the proposed method, we use two metrics to measure the multitask relationship. First, the multitask correlation matrix C for the cardiac indices, which is calculated from the cardiac indices covariate matrix [59] (the correlation matrix C only can be calculated from regression tasks and logical regression tasks according to [59]; therefore, we only calculate C from cardiac indices in our work.). In C , the positive and negative values indicate positive and negative correlation, respectively. The higher absolute value indicates a strong correlation. Then, to further consider the segmentation and reverse generation tasks, we measure the performance gain of task correlation between two tasks T_i and T_j , and we calculate the pairwise

performance gain [60] as $\text{PPG}_{ij} = (P'_i P'_j / P_i P_j)^{1/2}$, where P_i , P_j , P'_i , and P'_j are the performances of tasks T_i and T_j when learned individually and jointly. (MAE is used to measure the PPG for all tasks. Thus, the smaller the PPG value, the greater the gain between two tasks.)

C. Experiments

To verify the effectiveness and performance of our method, we have done rich experiments as follows.

First, to evaluate the performance of our method for multitype cardiac indices estimation, we apply our method to cardiac MR images. Then, we fine-tune the trained model and apply it to a different imaging modality, i.e., CT images.

Second, to validate the ability to explore task dependencies of our adversarial reverse mapping network, we performed the following comparison and ablation studies. For comparison, we test our fully framework (Ours) on MRI, and then, we compared it with segmentation-based methods, e.g., Max Flow [18], two-phase methods, e.g., Multifeature + RF [38], SDL + AKSF [43], MCDBN + RF [41], and deep multitask learning methods, e.g., Indices-Net [2], FullILVNet [49], and DMTRL [3] due to their good results on MRI data. Then, we calculated the Dice value of 2-D segmentation results of our method and compared it with GAN [25], FCN [24], U-net [58], and V-net [26]. For ablation studies, we compare the performance from two aspects that include the different network architectures and different cardiac indices that the model predicted. For the different network architectures, both the multitask learning network and the reverse mapping network are employed. In this case, we first use $L_{M\&R}$ loss to train the model ($M + R + L_{M\&R}$) without the adversarial loss. Then, we use the $L_{M\&R}$ loss and the adversarial loss that only match the predicted I and real I to train the model ($M + R + L_{M\&R+I-GAN}$), and we train the framework without a reverse mapping network. In this case, we first use the $L_{Dice} + L_2$ loss ($M + L_{DL}$) to train the model. Then, we use the $L_{Dice} + L_2$ loss that matches only the predicted I and real I to train the model ($M + L_{DL+I-GAN}$) to train the model. For different cardiac indices, we employ the full model to predict different cardiac indices: 1) each type of cardiac indices independently (segmentation only, linear indices only, and planar indices only); 2) segmentation + linear indices; 3) segmentation + planar indices; 4) linear indices + planar indices; and 5) finally, when the model predicts only the segmentation, we also calculate the 1-D cardiac indices from the segmentation results.

Third, to show the effectiveness of our bidirectional parameter sharing scheme, we performed the following studies. First, we establish several multitask learning frameworks over different baseline deep networks. These multitask learning frameworks share the same structure of pixel-level classifier, regression network, and joint representation network but have different feature extraction layers. In particular, we employ 2-D/3-D residual nets and dense nets for feature extraction. Then, for different multitask learning frameworks that we established, we test their performance under different conditions: 1) the framework is trained directly on target data

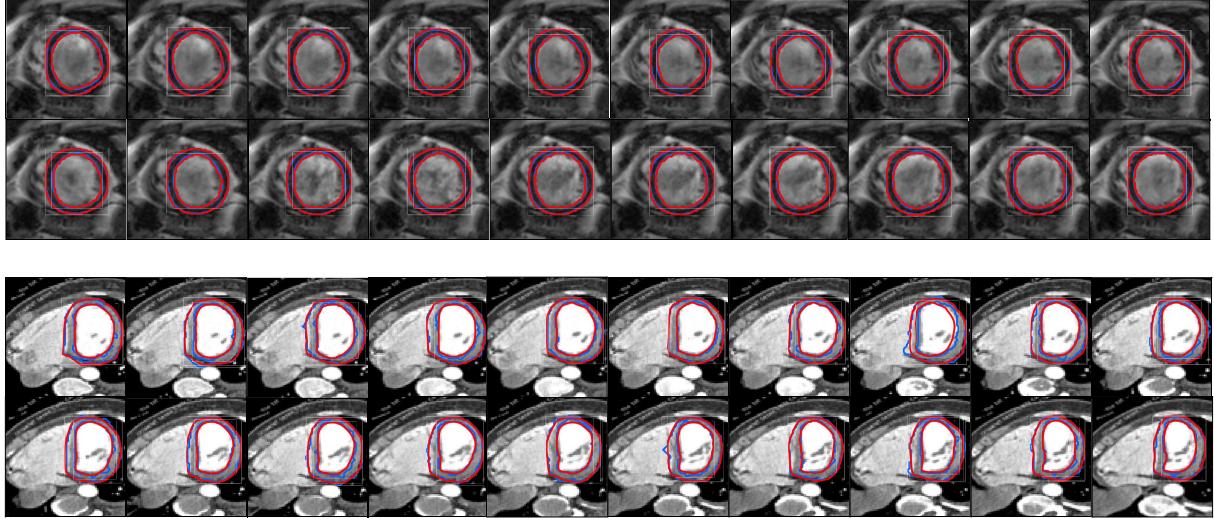


Fig. 4. Visualization of the segmentation results of 20 frames across the entire cardiac cycle for two representative subjects using MRI (top two rows) and CT (bottom two rows). Red lines in each frame represent the ground truth. Blue lines in each frame represent the automated segmentation results using our method.

without parameter sharing (No sharing); 2) the traditional one-way parameter sharing mechanism is employed (One-way-para) [61]; and 3) the bidirectional parameter sharing mechanism is employed (Bi-para). Finally, we transfer the parameters learned from the MRI to the CT. To further validate the importance of our method, we also transfer the parameters learned from the CT to the MRI.

D. Results and Discussion

1) Performance: Fig. 4 shows the segmentation results estimated by our method across the entire cardiac cycle for two representative subjects from the MRI and CT data, respectively. Red lines in each frame represent ground-truth contours. Blue lines in each frame represent the automated segmentation results of our method. The complex task dependencies and different imaging modalities pose great challenges for cross-modal multitype cardiac indices estimation. Even so, our method accurately locate the cardiac structure, as shown in Fig. 4, and also accurately estimate the WT, Dim, and Area indices, as comparison in Fig. 5 for MRI and shown in Table IV for CT, respectively. It means that our method is efficient for the multitype cardiac indices estimation in the MRI and the CT.

A comprehensive quantified validation is given in the last column of Tables I–IV. Our method estimates WTs, Dims, and Areas with average MAE of 1.33 ± 0.65 mm, 2.36 ± 1.45 mm, and 168 ± 130 mm² for the MRI and 1.56 ± 0.89 mm, 2.67 ± 1.97 mm, and 188 ± 144 mm² for the CT, respectively. In these indices, the areas of the WT are difficult to estimate due to the changeable shape of the myocardium across the cardiac cycle (see Fig. 4) and almost invisible border near the lateral free wall [2]. However, our method still achieves average MAE 1.33 ± 0.82 mm in MRI and 1.56 ± 0.89 mm in the CT for WTs estimation, which are less than the mode of pixel spacing (1.5625 mm/pixel in MRI and 1.6435 mm/pixel in CT). Then, the effectiveness for cardiac

segmentation of our approach is validated; for the segmentation of EpiLV, EndoLV, and Myo, our approach obtains the Dice values of 0.976, 0.952, and 0.916 in the MRI and 0.966, 0.923, and 0.860 in the CT, respectively.

2) Adversarial Reverse Mapping Network for Task Dependencies Learning:

a) Task dependencies: Fig. 6(a) shows the multitask correlation matrix of the cardiac indices, which is calculated from the predicted results of our method. Specifically, the myocardium area is correlated with WTs, while the cavity area is correlated with dimensions. Strong correlations exist within indices of WTs and dimensions. Furthermore, Fig. 6(b) uncovers pairwise gains (PPG) for every two tasks, where the darkness of colors indicates the strength of the gains. It is intuitive that the linear indices (WTs + Dims) with higher pairwise gains with planar indices (Areas). Meanwhile, both the linear and planar indices can achieve positive pairwise gains with the Seg and the Reverse tasks, which indicates that the Seg and reverse tasks can facilitate model learning.

b) Comparison studies: Our method achieves better performance than existing one-way methods on MRI data. For linear and planar indices estimation, our method outperforms segmentation-based method Max Flow [18], two-phase methods multifeature + RF [38], SDL + AKSF [43], MCDBN + RF [41], and deep multitask learning-based methods, e.g., Indices-Net [2], FullILVNet [49], and DMTRL [3], as shown in Table I. For the cardiac segmentation, our method outperforms than adversarial training method GAN [25], FCN [24], and V-Net [26], as shown in Table II.

Compared with Max Flow, our method achieves average MAE reductions of 56.7%, 5.28%, and 27.1% for WTs, Dims, and Areas. MaxFlow method first estimates segmentation, and then, it predicts the cardiac indices from segmentation. Although the MaxFlow method is based on the optimization of global geometry and intensity constraints, it fails to distinguish between the papillary muscles and the myocardium due to extreme deformation and almost the

TABLE I

OUR RESULTS OF LINEAR INDICES AND PLANNER INDICES AND COMPARED WITH THE PERFORMANCE USING THE MAX FLOW [18], MULTIFEATURE + RF [38], SDL + AKSF [43], MCDBN + RF [41], INDICES-NET [2], FULLLVNET [49], AND DEEP MULTITASK REPRESENTATION LEARNING (DMTRL) [3]

	Max Flow	Multi-feature+RF	SDL+AKRF	MCDBN+RF	Indices-Net	FullLVNet	DMTRL	Ours
WTsl(mm)								
WT1	1.53 ± 1.73	1.70 ± 1.47	1.98 ± 1.58	1.78 ± 1.40	1.39 ± 1.13	1.32 ± 1.09	1.26 ± 1.04	1.21 ± 1.11
WT2	3.23 ± 2.83	1.71 ± 1.34	1.67 ± 1.40	1.68 ± 1.41	1.51 ± 1.21	1.38 ± 1.10	1.40 ± 1.10	1.25 ± 1.17
WT3	4.15 ± 3.17	1.97 ± 1.54	1.88 ± 1.63	1.92 ± 1.45	1.65 ± 1.36	1.57 ± 1.35	1.59 ± 1.29	1.50 ± 1.45
WT4	5.08 ± 3.95	1.82 ± 1.41	1.87 ± 1.55	1.66 ± 1.20	1.53 ± 1.25	1.60 ± 1.36	1.57 ± 1.34	1.45 ± 1.23
WT5	3.47 ± 3.25	1.55 ± 1.33	1.65 ± 1.45	1.20 ± 1.01	1.30 ± 1.12	1.34 ± 1.11	1.32 ± 1.10	1.34 ± 1.19
WT6	1.76 ± 1.80	1.68 ± 1.43	2.04 ± 1.59	1.63 ± 1.23	1.28 ± 1.00	1.26 ± 1.10	1.25 ± 1.01	1.23 ± 1.20
Average	3.21 ± 1.98	1.73 ± 0.97	1.85 ± 1.03	1.65 ± 0.77	1.44 ± 0.71	1.41 ± 0.72	1.39 ± 0.68	1.33 ± 0.82
Dim(mm)								
Dim1	2.81 ± 2.76	3.53 ± 2.77	2.99 ± 2.43	2.88 ± 2.48	/	2.62 ± 2.09	2.47 ± 1.95	2.35 ± 1.87
Dim2	2.60 ± 2.62	3.49 ± 2.87	2.55 ± 2.30	2.45 ± 2.01	/	2.64 ± 2.12	2.59 ± 2.07	2.27 ± 2.32
Dim3	2.49 ± 2.88	3.91 ± 3.23	3.10 ± 2.54	2.93 ± 2.49	/	2.77 ± 2.22	2.48 ± 2.34	2.46 ± 2.03
Average	2.65 ± 2.330	3.64 ± 2.61	2.88 ± 2.03	2.75 ± 1.90	/	2.68 ± 1.64	2.51 ± 1.58	2.36 ± 1.45
Area(mm^2)								
Area1	156 ± 193	231 ± 193	198 ± 169	208 ± 166	185 ± 162	181 ± 155	172 ± 148	160 ± 158
Area2	339 ± 272	291 ± 246	286 ± 242	269 ± 217	223 ± 193	199 ± 174	189 ± 159	176 ± 189
Average	247 ± 201	261 ± 165	242 ± 145	239 ± 135	204 ± 133	190 ± 128	180 ± 118	168 ± 130

TABLE II

OUR RESULTS OF EPILV, ENDOLV, AND MYO SEGMENTATION COMPARED WITH THE PERFORMANCE USING THE GAN [25], FCN [24], U-NET [58], AND V-NET [26]

	GAN	FCN	U-net	V-Net	Ours
EpiLV	0.952 ± 0.022	0.967 ± 0.080	0.964 ± 0.045	0.958 ± 0.051	0.976 ± 0.013
EndoLV	0.925 ± 0.014	0.889 ± 0.037	0.902 ± 0.026	0.911 ± 0.023	0.952 ± 0.021
Myo	0.912 ± 0.021	0.909 ± 0.043	0.906 ± 0.034	0.890 ± 0.038	0.916 ± 0.014

same intensity of papillary muscles and the myocardium at middle frames of cardiac systolic (ES), as shown in Fig. 5. Thus, the incorrect papillary muscles segmentation leads to bad results of multitype cardiac indices. In contrast to Max Flow, our method treats segmentation as a subtask, which allows us to simultaneously estimate segmentation with other cardiac indices. Compared with the best two-phase regression method MCDBN + RF, our method achieves average MAE reductions of 6.9%, 8.28%, and 5.1% for WTs, Dims, and Areas. These methods estimate cardiac indices in two phases based on handcrafted features, while our method employs an end-to-end framework that can extract task-aware features. Our method also achieves a better performance compare to deep multitask learning methods. Indices-Net can extract task-aware features based on the jointly learning of representation and regression. FullLVNet and DMTRL have the advantages of the combination of CNN and RNN, but they heavily dependent on a particular constraint for task dependencies learning. FullLVNet explores the task dependencies based on predefined assumptions. DMTRL explores task dependencies based on a carefully designed covariance matrix. Unlike these methods, our framework explores the task dependencies via further learning a reverse mapping from multitype cardiac indices to images. Compared with DMTRL, our method also achieves

average MAE reductions of 4.32%, 5.98%, and 6.67% for WTs, Dims, and Areas (see Table I).

For the segmentation of EpiLV, EndoLV, and Myo, our method achieves the best results in all compared approaches, as shown in Table II. GAN achieves a good segmentation of EpiLV, EndoLV, and Myo based on effective adversarial training. However, it still obtains an inferior performance compared with ours.

c) *Ablation studies*: To prove the benefits of our model, we first test the performance of our full framework that predicts different cardiac indices. Then, we also test the performance of our framework with different network architectures.

As shown in Table III, the (Independent) could not achieve satisfactory results. It increases the average MAE by 21.80%, 11.44%, and 16.07% for WTs, Dims, and Areas, while the reductions of average Dice is 1.43%, 1.50%, and 0.09% for EpiLV, EndoLV, and Myo. Moreover, we also cannot obtain good results when we calculate the 1-D cardiac indices from independently predicted segmentation, as shown in Fig. 7. When we perform the experiments that estimate two types of cardiac indices in each time as (Seg + Linear), (Seg + Planar), and (Planar + Linear), we also can obtain a significant boost of performance compared with (independent). This indicates that our full method (Ours) for joint estimation of multitype cardiac indices can effectively obtain task dependence and improve the overall performance.

As shown in Table III [($M + L_{DL}$) and ($M + L_{DL+I-GAN}$)], the frameworks could not achieve satisfactory results for 1-D cardiac indices and segmentation without a reverse mapping network, regardless of which loss function is used. In this case, the frameworks are using one-way mapping structure, which could not fully explore task dependencies. In contrast, the frameworks with the reverse mapping network [$(M + R + L_{M\&R})$, $(M + R + L_{M\&R+I-GAN})$, and (Ours)

TABLE III
RESULTS OF OUR ABLATION STUDIES

	Independent	Seg+Linear	Seg+Planar	Planar+Linear	$M+L_{DL}$	$M+L_{DL+I-GAN}$	$M+R+L_{M\&R}$	$M+R+L_{M\&R+I-GAN}$	Ours
WTs(mm)									
WT1	1.58 ± 1.53	1.49 ± 1.22	/	1.53 ± 1.44	1.55 ± 1.58	1.46 ± 1.23	1.35 ± 1.38	1.31 ± 1.35	1.21 ± 1.11
WT2	1.63 ± 1.46	1.51 ± 1.15	/	1.47 ± 1.35	1.62 ± 1.35	1.60 ± 1.27	1.45 ± 1.20	1.29 ± 1.19	1.25 ± 1.17
WT3	1.92 ± 1.98	1.65 ± 1.58	/	1.59 ± 1.87	1.70 ± 1.56	1.65 ± 1.49	1.56 ± 1.41	1.52 ± 1.37	1.50 ± 1.45
WT4	1.55 ± 1.66	1.35 ± 1.21	/	1.57 ± 1.74	1.49 ± 1.26	1.52 ± 1.33	1.48 ± 1.43	1.42 ± 1.18	1.45 ± 1.23
WT5	1.38 ± 1.51	1.34 ± 1.17	/	1.72 ± 1.41	1.42 ± 1.22	1.36 ± 1.26	1.33 ± 1.46	1.28 ± 1.10	1.34 ± 1.19
WT6	1.63 ± 1.93	1.30 ± 1.00	/	1.45 ± 1.52	1.32 ± 1.24	1.29 ± 1.17	1.30 ± 1.18	1.25 ± 1.12	1.23 ± 1.20
Average	1.62 ± 0.98	1.44 ± 0.91	/	1.56 ± 0.87	1.52 ± 1.15	1.48 ± 0.87	1.41 ± 0.93	1.35 ± 0.79	1.33 ± 0.82
Dim(mm)									
Dim1	2.81 ± 2.78	2.66 ± 2.37	/	2.57 ± 2.84	2.76 ± 2.55	2.65 ± 2.34	2.56 ± 2.22	2.50 ± 2.09	2.35 ± 1.87
Dim2	2.65 ± 2.81	2.47 ± 2.23	/	2.65 ± 2.67	2.73 ± 2.42	2.68 ± 2.28	2.33 ± 2.19	2.29 ± 2.35	2.27 ± 2.32
Dim3	2.43 ± 2.49	2.45 ± 2.13	/	2.48 ± 2.56	2.62 ± 2.59	2.57 ± 2.49	2.36 ± 2.33	2.42 ± 2.25	2.46 ± 2.03
Average	2.63 ± 2.01	2.53 ± 1.66	/	2.57 ± 1.71	2.70 ± 1.67	2.63 ± 1.62	2.42 ± 1.49	2.40 ± 1.43	2.36 ± 1.45
Area(mm^2)									
Area1	188 ± 187	/	181 ± 175	182 ± 176	190 ± 167	184 ± 171	175 ± 154	168 ± 144	160 ± 158
Area2	201 ± 223	/	192 ± 162	197 ± 219	204 ± 181	196 ± 166	188 ± 152	189 ± 150	176 ± 189
Average	195 ± 178	/	190 ± 133	200 ± 145	194 ± 134	190 ± 129	182 ± 140	179 ± 132	168 ± 130
Segmentation(Dice)									
EpiLV	0.962 ± 0.040	0.964 ± 0.054	0.970 ± 0.070	/	0.951 ± 0.051	0.955 ± 0.031	0.961 ± 0.038	0.968 ± 0.026	0.976 ± 0.013
EndoLV	0.938 ± 0.067	0.941 ± 0.044	0.944 ± 0.035	/	0.916 ± 0.025	0.934 ± 0.056	0.942 ± 0.038	0.945 ± 0.020	0.952 ± 0.021
Myo	0.908 ± 0.082	0.911 ± 0.043	0.912 ± 0.038	/	0.903 ± 0.039	0.895 ± 0.061	0.906 ± 0.043	0.910 ± 0.036	0.916 ± 0.014

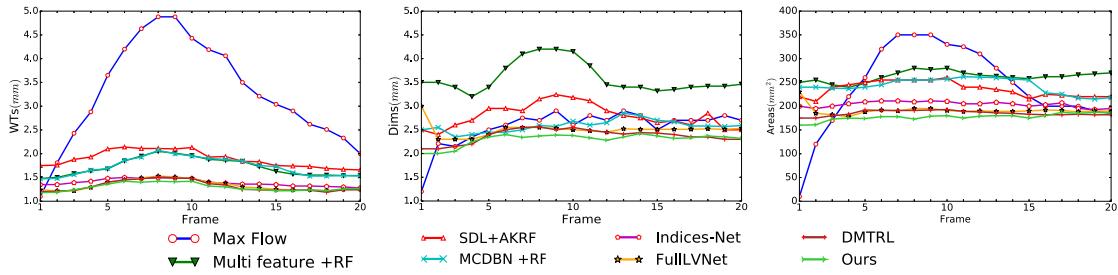


Fig. 5. Average framewise estimated errors of the WTs, Dims, and Areas indices using our proposed method and the compared methods. Our method delivers consistently lower estimated error across all cases for the corresponding frame than the compared methods.

in Tables I and II] achieved superior results. This indicates that the reverse mapping network can explore the task dependencies via generating cardiac images from multitype cardiac indices. It is also important to choose a proper strategy to train our framework. As shown in Table III [($M + R + L_{M\&R}$), ($M + R + L_{M\&R+I-GAN}$), and (Ours)], our method achieves the best performance. In contrast to ($M + R + L_{M\&R}$) and ($M + R + L_{M\&R+I-GAN}$) that only match the predicted I and real I , our method employs the our fully adversarial training approach that matches joint distribution $p_\theta(X, I)$ and $q_\phi(X, I)$. Relying on the ability of adversarial training to assess joint distributions of complex semantic variables, our method is more powerful to explore task dependencies.

As shown in Fig. 8, the model (Ours) with our adversarial training can effectively reconstruct the cardiac image and show the correct cardiac structure. Furthermore, (Ours, MAE = 0.20), (Seg + Linear, MAE = 0.23), (Seg + Planar, MAE = 0.25), and (Linear + Planar, MAE = 0.22) achieve smaller MAE than (Independent(Seg only), MAE = 0.26) of image reconstruction. This indicates that the learning of those highly abstract features of 1-D indices would further

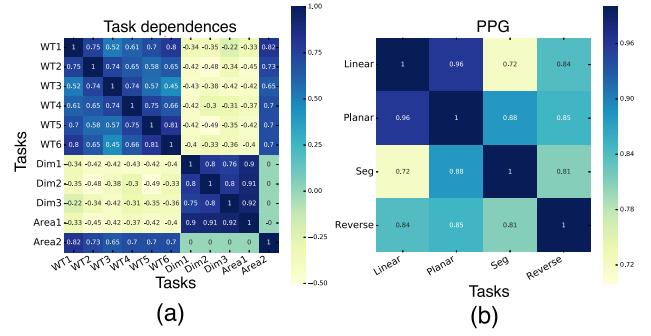


Fig. 6. (a) Multitask correlation matrix computed from the predicted results of different cardiac indices in our data set. (b) Visualization of pairwise performance gains among different tasks, i.e., linear cardiac indices prediction (Linear), planar cardiac indices prediction (Planar), cardiac segmentation (Seg), and reverse image generation (Reverse).

guide the model to find the proper low-level feature (z) for sharper image reconstruction. Those sharp images also show that our model stores the information of x into z and can, thus, reconstruct the details of the image. Combined with the good

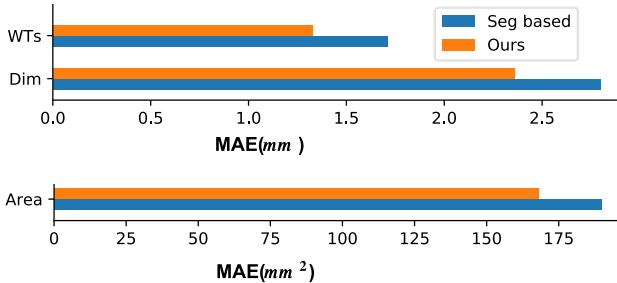


Fig. 7. Comparison of our joint estimation with segmentation-based 1-D cardiac indices calculation. (Our) denotes the results of the joint estimation. (Seg-based) denotes the results calculated via segmentation; we first employ our full model to predict the segmentation independently and then calculate the 1-D indices from the predicted segmentation.

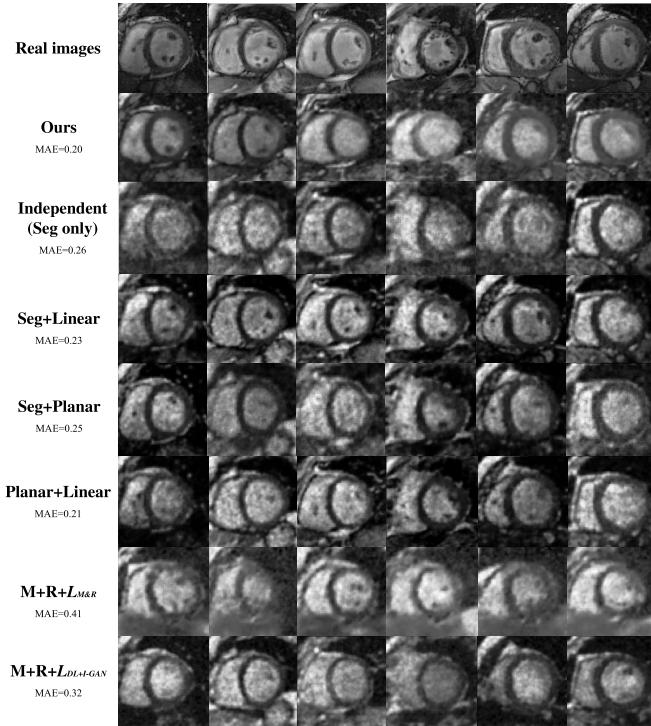


Fig. 8. Cardiac MRI reconstructed by different frameworks. On the first row are real images sampled from a test set. The rest rows are the corresponding images reconstructed by (Ours), [Independent(Seg only)], (Seg + Linear), (Seg + Planar), (Planar + Linear), ($M + L_{DL+I-GAN}$), and ($M + R + L_{M\&R+I-GAN}$), respectively. The MAE values of reconstructed results are also given for different frameworks.

performance of our model in the multitype cardiac indices estimation, our adversarial reverse mapping method may be an effective way to avoid overfitting or “steganography” effect when the domains have different entropy [55].

3) Bidirectional Parameter Sharing Scheme for Knowledge Transfer: As shown in Table IV, compared with (No sharing), the four methods with our bidirectional parameter sharing mechanism (Bi-para) obtain the best performance. In particular, they achieve average MAE reductions of 15.17%, 6.02%, and 7.34% for WTs, Dims, and Areas and Dice increase of 3.33%, 2.15%, and 2.66% for EpiLV, EndoLV, and

TABLE IV
TOP FOUR ROWS ARE TRANSFER LEARNING (FROM MRI TO CT)
RESULTS OF FOUR DEEP NETWORKS (2-D/3-D RESIDUAL NET AND 2-D/3-D DENSE NET) WITH DIFFERENT CONFIGURATIONS [(NO SHARING), (ONE-WAY-PARA), AND (BI-PARA)]. THEN, THE BOTTOM ROW IS THE TRANSFER LEARNING (FROM CT TO MRI) RESULTS OF THE 3-D DENSE NET

		No sharing	One-way-para	Bi-para
MRI to CT				
2D residual net	WT (mm)	1.91 \pm 1.23	1.82 \pm 1.17	1.63 \pm 0.98
	Dim (mm)	2.78 \pm 2.63	2.69 \pm 2.46	2.63 \pm 2.25
	Area (mm^2)	212 \pm 190	202 \pm 178	197 \pm 171
	EpiLV(Dice)	0.925 \pm 0.031	0.933 \pm 0.029	0.959 \pm 0.024
	EndoLV(Dice)	0.897 \pm 0.046	0.905 \pm 0.039	0.912 \pm 0.032
3D residual net	Myo(Dice)	0.834 \pm 0.098	0.836 \pm 0.113	0.857 \pm 0.053
	WT (mm)	1.84 \pm 1.12	1.72 \pm 1.09	1.60 \pm 1.12
	Dim (mm)	2.94 \pm 2.77	2.88 \pm 2.52	2.72 \pm 2.36
	Area (mm^2)	205 \pm 183	198 \pm 175	190 \pm 163
	EpiLV(Dice)	0.928 \pm 0.032	0.941 \pm 0.027	0.955 \pm 0.029
2D dense net	EndoLV(Dice)	0.901 \pm 0.039	0.908 \pm 0.043	0.913 \pm 0.020
	Myo(Dice)	0.832 \pm 0.119	0.846 \pm 0.110	0.854 \pm 0.080
	WT (mm)	1.90 \pm 1.10	1.67 \pm 0.79	1.58 \pm 0.72
	Dim (mm)	2.83 \pm 1.46	2.74 \pm 1.39	2.71 \pm 1.35
	Area (mm^2)	211 \pm 180	205 \pm 1.40	194 \pm 165
3D dense net	EpiLV(Dice)	0.931 \pm 0.029	0.946 \pm 0.031	0.962 \pm 0.021
	EndoLV(Dice)	0.891 \pm 0.034	0.914 \pm 0.026	0.921 \pm 0.019
	Myo(Dice)	0.840 \pm 0.109	0.847 \pm 0.102	0.861 \pm 0.089
	WT (mm)	1.86 \pm 1.12	1.70 \pm 1.17	1.56 \pm 0.89
	Dim (mm)	2.87 \pm 2.74	2.81 \pm 2.11	2.67 \pm 1.97
3D dense net	Area (mm^2)	202 \pm 183	194 \pm 175	188 \pm 144
	EpiLV(Dice)	0.934 \pm 0.025	0.948 \pm 0.018	0.966 \pm 0.015
	EndoLV(Dice)	0.903 \pm 0.033	0.911 \pm 0.024	0.923 \pm 0.012
	Myo(Dice)	0.837 \pm 0.120	0.841 \pm 0.110	0.860 \pm 0.057
CT to MRI				
3D dense net	WT (mm)	1.73 \pm 0.82	1.61 \pm 0.87	1.42 \pm 0.66
	Dim (mm)	3.12 \pm 1.45	2.64 \pm 1.33	2.52 \pm 1.20
	Area (mm^2)	195 \pm 157	175 \pm 140	178 \pm 119
	EpiLV(Dice)	0.926 \pm 0.038	0.948 \pm 0.023	0.962 \pm 0.012
	EndoLV(Dice)	0.912 \pm 0.057	0.942 \pm 0.019	0.951 \pm 0.021
3D dense net	Myo(Dice)	0.887 \pm 0.082	0.901 \pm 0.025	0.910 \pm 0.031

Myo. This indicates that our bidirectional parameter sharing mechanism can effectively transfer the learned task dependencies from the MRI to the CT. It is difficult to mine and represent complex task dependencies, which are high-level semantic information. Even so, compared with (One-way-para), the four methods with our bidirectional parameter sharing mechanism (Bi-para) can still achieve average MAE reductions of 7.87%, 3.46%, and 3.50% for WTs, Dims, and Areas and Dice increase of 3.45%, 1.70%, and 1.83% for EpiLV, EndoLV, and Myo. For (One-way-para) methods, even if they employ different networks, they are still trained under a one-way mapping framework from cardiac image to multitype cardiac indices regardless the learning and representation ability of the networks. However, our bidirectional parameter sharing mechanism provides a different perspective of the reverse mapping, which explores discriminative information that goes beyond the inherent learning and representation ability of the deep one-way network.

We also try an extended experiment, as shown at the bottom of Table IV, which transfers the parameters learned from the CT to the MRI. Both the (One-way-para) and (Bi-para) can enhance the performance of the MRI when there is knowledge transfer from the CT to MRI. This indicates that our method effectively meets clinical needs via transferring the proper task dependencies learned between the two different imaging modalities.

E. Limitations

This study has potential limitations. The study focuses on the 2-D mid-ventricular slices across one cardiac cycle. However, this might be limited to 3-D cardiac images in clinical application. Furthermore, the 3-D image-based cardiac functions, i.e., volume and ejection fraction, could not be obtained directly. In the future work, it is necessary to generalize the proposed method to 3-D and 3D + t cardiac images.

IV. CONCLUSION

In this study, we have developed a multitask learning framework based on adversarial reverse mapping for estimating multitype cardiac indices in different imaging modalities, e.g., MRI and CT. The proposed framework has an established reverse mapping network that explores task dependencies by learning the mapping from multitype cardiac indices to cardiac images via adversarial training. A bidirectional parameter sharing scheme then transfers the parameters from both the multitask learning network and the proposed reverse mapping network. Experimental results show that our method not only accurately estimates multitype cardiac indices in MRI but also performs well for knowledge transfer from MRI to CT.

REFERENCES

- [1] P. Peng, K. Lekadir, A. Gooya, L. Shao, S. E. Petersen, and A. F. Frangi, “A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging,” *Magn. Reson. Mater. Phys., Biol. Med.*, vol. 29, no. 2, pp. 155–195, Apr. 2016.
- [2] W. Xue, I. B. Nachum, S. Pandey, J. Warrington, S. Leung, and S. Li, “Direct estimation of regional wall thicknesses via residual recurrent neural network,” in *Proc. Int. Conf. Inf. Process. Med. Imag.*, 2017, pp. 505–516.
- [3] W. Xue, G. Brahm, S. Pandey, S. Leung, and S. Li, “Full left ventricle quantification via deep multitask relationships learning,” *Med. Image Anal.*, vol. 43, pp. 54–65, Jan. 2018.
- [4] M. Alsalamah, S. Amin, and V. Palade, “Heart disease diagnosis based on deep radial basis function kernel machine,” *ICST Trans. Ambient Syst.*, vol. 5, no. 17, 2018, Art. no. 154774.
- [5] A. K. Attili, A. Schuster, E. Nagel, H. C. Johan Reiber, and J. Rob van der Geest, “Quantification in cardiac MRI: Advances in image acquisition and processing,” *Int. J. Cardiovascular Imag.*, vol. 26, pp. 27–40, Feb. 2010.
- [6] M. Dewey, *Cardiac CT*. Berlin, Germany: Springer, 2011.
- [7] A. A. Hussein *et al.*, “Safety of computed tomography in patients with cardiac rhythm management devices: Assessment of the U.S. food and drug administration advisory in clinical practice,” *J. Amer. College Cardiol.*, vol. 63, no. 17, pp. 1769–1775, 2014.
- [8] G. Obozinski, B. Taskar, and M. I. Jordan, “Joint covariate selection and joint subspace selection for multiple classification problems,” *Statist. Comput.*, vol. 20, no. 2, pp. 231–252, Apr. 2010.
- [9] W. Cao, S. Wu, Z. Yu, and H.-S. Wong, “Exploring correlations among tasks, clusters, and features for multitask clustering,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 2, pp. 355–368, Feb. 2019.
- [10] Y. Li, X. Tian, T. Liu, and D. Tao, “On better exploring and exploiting task relationships in multitask learning: Joint model and feature learning,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 5, pp. 1975–1985, May 2018.
- [11] A. K. Bishwas, A. Mani, and V. Palade, “An all-pair quantum SVM approach for big data multiclass classification,” *Quantum Inf. Process.*, vol. 17, no. 10, p. 282, Oct. 2018.
- [12] Y. Zhang and Q. Yang, “A survey on multi-task learning,” 2017, *arXiv:1707.08114*. [Online]. Available: <https://arxiv.org/abs/1707.08114>
- [13] S. Liu, E. Johns, and A. J. Davison, “End-To-End multi-task learning with attention,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 1871–1880.
- [14] P. Yang, P. Zhao, J. Zhou, and X. Gao, “Confidence weighted multitask learning,” in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, Jul. 2019, pp. 5636–5643.
- [15] Y. Xia *et al.*, “Dual learning for machine translation,” in *Proc. Adv. Neural Inf. Process. Syst.*, Barcelona, Spain, 2016, pp. 820–828.
- [16] J. Ian Goodfellow *et al.*, “Generative adversarial nets,” in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
- [17] D. Nie, L. Wang, Y. Gao, J. Lian, and D. Shen, “STRAINet: Spatially varying stochastic residual Adversarial networks for MRI pelvic organ segmentation,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 5, pp. 1552–1564, May 2019.
- [18] I. Ben Ayed, H.-M. Chen, K. Punithakumar, I. Ross, and S. Li, “Max-flow segmentation of the left ventricle by recovering subject-specific distributions via a bound of the bhattacharyya measure,” *Med. Image Anal.*, vol. 16, no. 1, pp. 87–100, Jan. 2012.
- [19] M. A. Alattar, N. F. Osman, and A. S. Fahmy, “Segmentation of left ventricle in cardiac MRI images using adaptive multi-seeded region growing,” in *Proc. 5th Cairo Int. Biomed. Eng. Conf.*, Dec. 2010, pp. 25–28.
- [20] C.-Z. You, V. Palade, and X.-J. Wu, “Robust structured low-rank representation for image segmentation,” *Int. J. Artif. Intell. Tools*, vol. 27, no. 5, Aug. 2018, Art. no. 1850020.
- [21] X. Albà, R. M. F. i Ventura, K. Lekadir, C. Tobon-Gomez, C. Hoogendoorn, and A. F. Frangi, “Automatic cardiac LV segmentation in MRI using modified graph cuts with smoothness and interslice constraints,” *Magn. Reson. Med.*, vol. 72, no. 6, pp. 1775–1784, Dec. 2014.
- [22] X. Zhuang, “Multivariate mixture model for cardiac segmentation from multi-sequence MRI,” in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2016, pp. 581–588.
- [23] A. Suinesiaputra *et al.*, “Statistical shape modeling of the left ventricle: Myocardial infarct classification challenge,” *IEEE J. Biomed. Health Inform.*, vol. 22, no. 2, pp. 503–515, Mar. 2018.
- [24] E. Shelhamer, J. Long, and T. Darrell, “Fully convolutional networks for semantic segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017.
- [25] P. Luc, C. Couprie, S. Chintala, and J. Verbeek, “Semantic segmentation using adversarial networks,” 2016, *arXiv:1611.08408*. [Online]. Available: <https://arxiv.org/abs/1611.08408>
- [26] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-Net: Fully convolutional neural networks for volumetric medical image segmentation,” in *Proc. 4th Int. Conf. 3D Vis. (3DV)*, Oct. 2016, pp. 565–571.
- [27] J. Patravali, S. Jain, and S. Chilamkurthy, “2D-3D fully convolutional neural networks for cardiac MR segmentation,” in *Proc. Int. Workshop Stat. Atlases Comput. of the Heart*, Quebec, QC, Canada, Sep. 2017, pp. 130–139.
- [28] Q. Zheng, H. Delingette, N. Duchateau, and N. Ayache, “3-D consistent and robust segmentation of cardiac images by deep learning with spatial propagation,” *IEEE Trans. Med. Imag.*, vol. 37, no. 9, pp. 2137–2148, Sep. 2018.
- [29] P. K. Rudra Poudel, P. Lamata, and G. Montana, “Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation,” in *Proc. Int. Workshop Reconstruction Anal. Moving Body Organs*, Athens, Greece, Oct. 2016, pp. 83–94.
- [30] M. Abdulaziz Alayba, V. Palade, M. England, and R. Iqbal, “A combined CNN and LSTM model for arabic sentiment analysis,” 2018, *arXiv:1807.02911*. [Online]. Available: <https://arxiv.org/abs/1807.02911>
- [31] Z. Shi *et al.*, “Bayesian VoxDRN: A probabilistic deep Voxelwise dilated residual network for whole heart segmentation from 3D MR images,” in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Berlin, Germany: Springer, 2018, pp. 569–577.
- [32] R. S. Ahmed, J. Liu, Z. Fei, and M. Zahid, “Automated segmentation of whole cardiac CT images based on deep learning,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 4, pp. 1–8, 2018.
- [33] X. Zhen, Z. Wang, A. Islam, M. Bhaduri, I. Chan, and S. Li, “Multi-scale deep networks and regression forests for direct bi-ventricular volume estimation,” *Med. Image Anal.*, vol. 30, pp. 120–129, May 2016.
- [34] M. Afshin *et al.*, “Regional assessment of cardiac left ventricular myocardial function via MRI statistical features,” *IEEE Trans. Med. Imag.*, vol. 33, no. 2, pp. 481–494, Feb. 2014.
- [35] M. Yang, L. Zhang, S. C.-K. Shiu, and D. Zhang, “Robust Kernel representation with statistical local features for face recognition,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 6, pp. 900–912, Jun. 2013.
- [36] H. Zhang *et al.*, “A meshfree representation for cardiac medical image computing,” *IEEE J. Transl. Eng. Health & Med.*, vol. 6, 2018, Art. no. 1800212.

- [37] A. W. Kabani and R. Mahmoud El-Sakka, "Estimating ejection fraction and left ventricle volume using deep convolutional networks," in *Proc. Int. Conf. Image Anal. Recognit.*, 2016, pp. 678–686.
- [38] X. Zhen, Z. Wang, A. Islam, M. Bhaduri, I. Chan, and S. Li, "Direct estimation of cardiac Bi-ventricular volumes with regression forests," in *Proc. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*. Boston, MA, USA: Springer, 2014, vol. 17, no. 2, pp. 586–593.
- [39] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuring, and D. Comaniciu, "Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features," *IEEE Trans. Med. Imag.*, vol. 27, no. 11, pp. 1668–1681, Nov. 2008.
- [40] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit. (CVPR)*, vol. 1, Jun. 2005, pp. 886–893.
- [41] S. N. Tran and A. S. d'Avila Garcez, "Deep logic networks: Inserting and extracting knowledge from deep belief networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 2, pp. 246–258, Feb. 2018.
- [42] F. Liao, X. Chen, X. Hu, and S. Song, "Estimation of the volume of the left ventricle from MRI images using deep neural networks," *IEEE Trans. Cybern.*, vol. 49, no. 2, pp. 495–504, Feb. 2019.
- [43] X. Zhen, A. Islam, M. Bhaduri, I. Chan, and S. Li, "Direct and simultaneous four-chamber volume estimation by multi-output regression," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2015, pp. 669–676.
- [44] A. Ruiz-Garcia, M. Elshaw, A. Altahhan, and V. Palade, "A hybrid deep learning neural approach for emotion recognition from facial expressions for socially assistive robots," *Neural Comput. Appl.*, vol. 29, no. 7, pp. 359–373, Apr. 2018.
- [45] A. Ruiz-Garcia, M. Elshaw, A. Altahhan, and V. Palade, "Deep learning for emotion recognition in faces," in *Proc. Int. Conf. Artif. Neural Netw.*, Barcelona, Spain, Sep. 2016, pp. 38–46.
- [46] A. Cutler, D. Richard Cutler, and R. John Stevens, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2004.
- [47] K. Harza and R. Chellappa, "Growing regression forests by classification: Applications to object pose estimation," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 552–567.
- [48] H. Chen, X. Wang, and P. A. Heng, "Automated mitosis detection with deep regression networks," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2016, pp. 1204–1207.
- [49] W. Xue, A. Lum, A. Mercado, M. Landis, J. Warrington, and S. Li, "Full quantification of left ventricle via deep multitask learning network respecting intra- and inter-task relatedness," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Quebec, QC, Canada, Sep. 2017, pp. 276–284.
- [50] P. Luo, G. Wang, L. Lin, and X. Wang, "Deep dual learning for semantic image segmentation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2737–2745.
- [51] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 2242–2251.
- [52] T. Kim, M. Cha, H. Kim, J. K. Lee, and J. Kim, "Learning to discover cross-domain relations with generative adversarial networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Sydney, NSW, Australia, Aug. 2017, pp. 1857–1865.
- [53] Y. Xia, J. Bian, T. Qin, N. Yu, and T. Y. Liu, "Dual inference for machine learning," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, 2017, pp. 3112–3118.
- [54] Y. Pu *et al.*, "Adversarial symmetric variational autoencoder," in *Proc. Adv. Neural Inf. Process. Syst.*, Long Beach, CA, USA, Dec. 2017, pp. 4333–4342.
- [55] C. Chu, A. Zhmoginov, and M. Sandler, "CycleGAN, a master of steganography," 2017, *arXiv:1712.02950*. [Online]. Available: <http://arxiv.org/abs/1712.02950>
- [56] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.
- [57] H. Zou and T. Hastie, "Addendum: Regularization and variable selection via the elastic net," *J. Roy. Stat. Soc., Ser. B (Stat. Methodol.)*, vol. 67, no. 5, p. 768, Nov. 2005.
- [58] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Munich, Germany, Oct. 2015, pp. 234–241.
- [59] Y. Zhang and D.-Y. Yeung, "A regularization approach to learning task relationships in multitask learning," *ACM Trans. Knowl. Discovery from Data*, vol. 8, no. 3, pp. 1–31, Jun. 2014.
- [60] H. Zhang, L. Xiao, Y. Wang, and Y. Jin, "A generalized recurrent neural architecture for text classification with multi-task learning," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, Melbourne, Australia, Aug. 2017, pp. 3385–3391.
- [61] W. Zhang *et al.*, "Deep model based transfer and multi-task learning for biological image analysis," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2015, pp. 1475–1484.



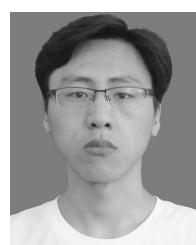
Chengjin Yu is currently pursuing the Ph.D. degree with the School of Computer Science and Technology, Anhui University, Hefei, China.

His research is in the interdisciplinary field of medical image analysis and artificial intelligence.



Zhifan Gao (Member, IEEE) received the B.S. and M.E. degrees in electronics and information engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2009 and 2011, respectively, and the Ph.D. degree in pattern recognition and intelligent systems from the University of Chinese Academy of Sciences, Beijing, China, in 2017.

He was a Post-Doctoral Fellow with Western University, London, ON, Canada, from 2018 to 2020. His research focuses on medical image processing, computer vision, and machine learning.



Weiwei Zhang is currently pursuing the Ph.D. degree with the School of Biomedical Engineering, Sun Yat-sen University, Guangzhou, China.

His research is in the interdisciplinary field of medical image analysis and artificial intelligence, for improving lesion detection, anatomical structure computation, and image quality promotion, with an impact to advance disease diagnosis and computer-assisted intervention via machine learning.



Guang Yang (Member, IEEE) received the M.Sc. degree in vision imaging and virtual environments from the Department of Computer Science, University College London, London, U.K., in 2006, and the Ph.D. degree in medical image analysis jointly from the CMIC and the Department of Computer Science and Medical Physics, University College London, in 2012.

He is currently an Honorary Lecturer with the Neuroscience Research Centre, Cardiovascular and Cell Sciences Institute, St George's, University of London, London. He is also an Image Processing Physicist and an Honorary Senior Research Fellow of the Cardiovascular Research Centre, Royal Brompton Hospital, London, and the National Heart and Lung Institute, Imperial College London, London.



Shu Zhao received the Ph.D. degree in computer science from Anhui University, Hefei, China, in 2007.

She is currently a Professor with the School of Computer Science and Technology, Anhui University. Her current research interests include quotient space theory, granular computing, social networks, and machine learning.



Yanping Zhang received the Ph.D. degree from Anhui University, Hefei, China, in 2005.

She is currently a Professor with Anhui University. She is a Principal Investigator of some 973 projects and the Leader of the National Natural Science Foundations of China and Anhui Province. Her main research interests include computational intelligence, quotient space theory, artificial neural networks, intelligent information processing, and machine learning.



Heye Zhang received the Ph.D. degree in electronic and computer engineering from The Hong Kong University of Science and Technology, Hong Kong, in 2007.

He is currently a Full Professor with the School of Biomedical Engineering, Sun Yat-sen University, Guangzhou, China. His research focuses on health informatics computing and machine learning.



Shuo Li received the Ph.D. degree in computer science and software engineering from Concordia University, Montreal, QC, Canada, in 2006.

He was with GE (Healthcare), Chicago, IL, USA, until 2015. He is an Academic Artificial Intelligence Scientist in both medical imaging and computer science & software engineering. He has been the Founder and the Scientific Director of the Digital Imaging Group of London (DIG), London, U.K., since 2007. He has been an Associate Professor with Western University, London, ON, Canada, since

2015. As a primary supervisor, he has mentored over 60 trainees, including post-doctoral fellows, Ph.D./M.Sc. students, undergraduates, and visiting scholars. Since 2015, he has published over 120 journal articles and conference papers, six books, and two awarded patents. His research focuses on the artificial intelligence systems solving the clinical and fundamental data analytics problems with an emphasis on the innovations of our learning schemes.

Dr. Li serves as a Board Member of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) and the General Chair of the MICCAI Conference 2022. He has been serving on the editorial board of multiple premier journals in the field of medical image analysis and pattern recognition. He has been frequently invited to review grant applications for several international funding agencies. He has been actively serving multiple professional societies in various leadership capabilities.