



HAL
open science

Unsupervised Three-Dimensional Image Registration Using a Cycle Convolutional Neural Network

Ziwei Lu, Guanyu Yang, Tiancong Hua, Liyu Hu, Youyong Kong, Lijun Tang,
Xiaomei Zhu, Jean-Louis Dillenseger, Huazhong Shu, Jean-Louis Coatrieux

► **To cite this version:**

Ziwei Lu, Guanyu Yang, Tiancong Hua, Liyu Hu, Youyong Kong, et al.. Unsupervised Three-Dimensional Image Registration Using a Cycle Convolutional Neural Network. 2019 IEEE International Conference on Image Processing (ICIP), Sep 2019, Taipei, Taiwan. pp.2174-2178, 10.1109/ICIP.2019.8803163 . hal-02281541

HAL Id: hal-02281541

<https://hal.science/hal-02281541>

Submitted on 9 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Unsupervised Three-Dimensional Image Registration Using a Cycle Convolutional Neural Network

Ziwei Lu, Guanyu Yang, Tiancong Hua, Liyu Hu, Youyong Kong, Lijun Tang,
Xiaomei Zhu, Jean-Louis Dillenseger, Huazhong Shu, Jean-Louis Coatrieux

► **To cite this version:**

Ziwei Lu, Guanyu Yang, Tiancong Hua, Liyu Hu, Youyong Kong, et al.. Unsupervised Three-Dimensional Image Registration Using a Cycle Convolutional Neural Network. 2019 IEEE International Conference on Image Processing (ICIP), Sep 2019, Taipei, Taiwan. pp.2174-2178, 10.1109/ICIP.2019.8803163 . hal-02281541

HAL Id: hal-02281541

<https://hal.archives-ouvertes.fr/hal-02281541>

Submitted on 9 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Unsupervised three-dimensional image registration using a Cycle Convolutional Neural Network

Ziwei Lu and Guanyu Yang and Tiancong Hua and Liyu Hu and Youyong Kong and Lijun tang and Xiaomei Zhu and Jean-Louis Dillenseger and Huazhon Shu and Jean-louis Coatrieux

Abstract—In this paper, an unsupervised cycle image registration convolutional neural network named CIRNet is developed for 3D medical image registration. Different from most deep learning based registration methods that require known spatial transforms, our proposed method is trained in an unsupervised way and predicts the dense displacement vector field. The CIRNet is composed by two image registration modules which have the same architecture and share the parameters. A cycle identical loss is designed in the CIRNet to provide additional constraints to ensure the accuracy of the predicted dense displacement vector field. The method is evaluated by the registration in 4D (3D+t) cardiac CT and MRI images respectively. Quantitative evaluation results demonstrate that our method performs better than the other two existing image registration algorithms. Especially, compared to the traditional image registration methods, our proposed network can finish 3D image registration in less than one second.

Index Term - Image registration, 3D, unsupervised, non-rigid, convolutional neural network

I. INTRODUCTION

Image registration is widely used in the medical image processing. Traditional registration methods are iteratively-optimized by parametric [1], [2] or nonparametric transformations [3], [4]. However, these traditional methods are limited by the large amount of computation in optimization.

With the success of deep neural network in image recognition, classification, segmentation and other computer vision tasks [5], deep neural networks have been used to solve the medical image registration. Recently, several supervised deep neural networks have been proposed to perform image registration [6], [7], [8]. These methods require supervised information to train the networks, such as dense displacement vector field (DVF) [6], [7], reference labels of segmentation [8]. However, these supervised methods could be influenced by the error or bias existed in the supervised information.

Z. Lu, G. Yang, T. Hua, L. Hu, Y. Kong and H. Shu are with LIST, Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, Nanjing, China.

J.-L. Dillenseger and J.-L. Coatrieux are with Univ Rennes, Inserm, LTSI - UMR1099, Rennes, F-35000, France.

L. Tang and X. Zhu are with the Dept. of Radiology, the First Affiliated Hospital of Nanjing Medical University, Nanjing, China.

G. Yang, H. Shu, Y. Kong, J.-L. Dillenseger and J.-L. Coatrieux are with Centre de Recherche en Information Biomédicale sino-français (CRIBS).

This research was supported by the National Key Research and Development Program of China (2017YFC0107903), the National Natural Science Foundation under grants (31571001, 61828101), the Short-Term Recruitment Program of Foreign Experts (WQ20163200398), and the Science Foundation for The Excellent Youth Scholars of Southeast University.

In the unsupervised image registration methods based on the deep learning [9], [10], [11], [12], supervised information, such as deformation displacement field or segmentation label, is not required in the process of training. Bob D et al. [9] developed a 2D end-to-end unsupervised registration network. Li H et al. [10] encoded the spatial transformations of images based on fully convolutional network (FCN) [13]. Balakrishnan et al. [11] implemented a registration network VoxelMorph combining the spatial transformer network (STN) [14]. Fan J et al. [12] proposed a patch-based method to achieve unsupervised registration using generative adversarial nets (GAN) [15]. Patch-based methods are fast to be trained, however, global information could be missed. The time required for image registration of these deep learning based unsupervised registration methods are much less than that of the traditional method, but some of these methods are inferior to traditional methods in terms of accuracy.

In this paper, we propose a convolutional neural network, i.e. CIRNet that can predict non-rigid spatial transform between the moving image and the fixed image. The proposed network is trained on the entire 3D image directly without any supervised information. The CIRNet consists of two registration modules that perform a cycle registration for the moving image. A cycle identical loss is introduced to improve the accuracy of the registration. Preliminary experiments on cardiac MR and CT images show that our method can generate accurate 3D registration results.

II. METHODS

The task of image registration is to find the DVF ϕ between the moving image M and the fixed images F , then deform the moving image with the DVF so that the two images are aligned in the spatial coordinate system. Image registration can be described as an optimization problem by minimizing an energy function P , The energy function P is described as below:

$$P = S(M(\phi), F) + Reg(\phi) \quad (1)$$

where $M(\phi)$ denotes the warped image obtained by deform the moving image M with the DVF ϕ , function $S(x, y)$ measures the dissimilarity between two images, and $Reg(\phi)$ is a smooth regularization constraint on the DVF.

A. Network structure

We proposed the CIRNet that learns details of the spatial transformation relationship between moving image and fixed

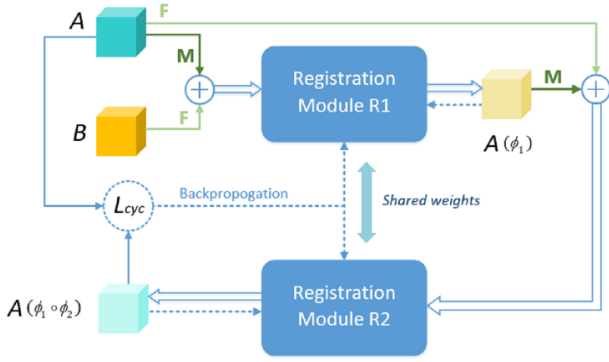


Fig. 1. Illustration of the CIRNet. The two 3D images A and B are inputs into the network. Dotted line represents back propagation. L_{cyc} is the cycle identical loss, measuring similarity between the image A and the warped image $A(\phi_1 \circ \phi_2)$.

image directly. The network takes 3D fixed and moving images as input and outputs the warped moving image.

As shown in Fig. 1, the CIRNet is composed of two cascaded registration modules R1, R2. The module R1 and R2 are identical in architecture. The parameters of the two modules are shared so that the number of total parameters in the whole network will not increase.

The moving image A and the fixed image B are concatenated into 2-channel input and fed into the module R1. The module R1 estimates the DVF ϕ to generate the warped image $A(\phi_1)$, which is similar to the image B in spatial coordinate domain. The module R2 takes the warped image $A(\phi_1)$ and the image A as input. In this stage, the difference to the module R1 is that the image $A(\phi_1)$ and the image A serve as moving and fixed images respectively. The DVF ϕ_2 and the warped image $A(\phi_1 \circ \phi_2)$ are the output of module R2. Thus, our network is able to perform a cycle registration by $R1 : \Omega_A \rightarrow \Omega_B$ and $R2 : \Omega_B \rightarrow \Omega_A$, where Ω_A and Ω_B refer to spatial coordinate domain of the image A and the image B respectively.

Assuming that the registration network is trained well, the warped image $A(\phi_1 \circ \phi_2)$ should be consistent with the original image A not only in spatial coordinate domain but also in gray-level distribution. The similarity measurement between the warped image $A(\phi_1 \circ \phi_2)$ and the image A is deduced as $A(\phi_1 \circ \phi_2) \approx A$, which helps to improve the training of the modules R1 and R2. Details about this similarity measurement L_{cyc} will be described in the section of loss function. It is worth to be mentioned that the training process of the cascaded network is performed end-to-end on the entire 3D image directly. Such a cycle registration network can provide additional constraints to give a more precise prediction of the local non-rigid deformation between two images. It is also helpful to the convergence of the network training for such an unsupervised machine learning method. In addition, the testing process will be done in one shot by the output of a registration module R1 or R2 without computing the auxiliary module.

The architecture of the modules R1 or R2 are mainly based

on the network presented in [11] as shown in Fig. 2. The module is composed of two parts. The first part is a convolutional neural network (ConvNet), which predicts the DVF ϕ . The second part performs tri-linear interpolation on image based on spatial transformation network [14]. ConvNet follows an encoder-decoder architecture which is similar to the UNet [16]. The encoder layers include four down-sampling operations which are applied by stride convolutions. The decoder layers include four up-sampling operations. There are four short-connections in the ConvNet module that allows concatenating the feature maps obtained by the encoder layers to the feature maps obtained by the corresponding decoder layers. Therefore, the low-level features and high-level features are integrated. We added batch normalization layer [17] and Leaky ReLU activation after each convolution layer. The last layer of the ConvNet is a 3D convolution layer with 3 convolutional kernels that parameterize the DVF ϕ in the X , Y , and Z axis.

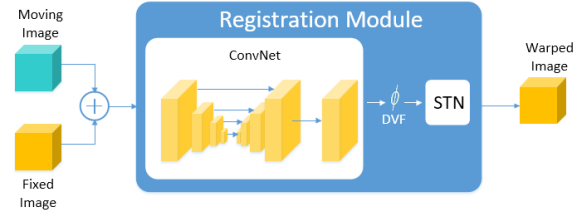


Fig. 2. Illustration of the architecture of the registration module R1 and R2.

B. Loss Function

The loss function designed for the CIRNet contains three terms: L_{NCC} , L_{cyc} and L_{reg} . The first two terms measure the image similarities and the last term penalizes the discontinuity of the estimated DVF.

In the CIRNet, cycle registrations were performed by the modules R1 and R2. The warped images generated by each registration module are consistent with the corresponding moving images in the spatial coordinate domain. As the other image registration methods, the normalized correlation coefficient (NCC) [18] is used in the L_{NCC} to measure the similarity between $A(\phi_1)$ and B . The NCC is calculated using a sliding window of $9*9*9$. If the LNCC equals to 0, it means that the two images are nearly uncorrelated. On the contrary, if L_{NCC} equals to 1, it means that the two images are highly positively correlated.

For the image A , the CIRNet should be able to deform A back to the original image A after the cycle registration performed by modules R1 and R2. Therefore, the cycle identical loss L_{cyc} enforces $A(\phi_1 \circ \phi_2)$ equivalent to the image A . Considering that two images should be identical in gray-level distribution and spatial coordinate domain, L2 regularization is used in the L_{cyc} , which is defined as follows,

$$L_{cyc}(A(\phi_1 \circ \phi_2), A) = \frac{1}{N} \sum_{i \in \Omega} (A(\phi_1 \circ \phi_2)(i) - A(i))^2 \quad (2)$$

where N is the number of all voxels, Ω refers to all the voxels in the image.

In order to ensure the continuities of the warped image, gradient optimization is used to smooth the DVF ϕ . For each voxel i , the smooth regularization L_{reg} is defined as follow:

$$L_{reg}(\phi) = \sum_{i=\Omega} \nabla \phi(i)^2 \quad (3)$$

Based on the above analysis, the loss functions of the two registration modules R1 and R2 are designed as below:

$$L_{R1}(A, B) = -L_{NCC}(A(\phi_1), B) + \alpha L_{reg}(\phi_1) + \beta L_{cyc}(A(\phi_1 \circ \phi_2), A) \quad (4)$$

$$L_{R2}(A, B) = -L_{NCC}(A(\phi_1 \circ \phi_2), A) + \alpha L_{reg}(\phi_1) + \beta L_{cyc}(A(\phi_1 \circ \phi_2), A) \quad (5)$$

Where $A(\phi_1)$ denotes the warped image produced by the module R1, and $A(\phi_1 \circ \phi_2)$ denotes the warped image obtained from the module R2, α and β controls the relative importance of the three terms. L_{cyc} is used for backpropagation of both R1 and R2 modules to enhance the training process.

III. EXPERIMENTS AND RESULTS

A. Dataset

We evaluated our method by inter-frame registration in 4D (3D+t) cardiac CT and MRI images respectively.

The first dataset includes the short-axis cine MRI images from 100 patients used in Automated Cardiac Diagnosis Challenge (ACDC) [19]. The pixel resolution is from 1.37 to 1.68 mm². Each 4D MRI image contains 28 to 40 volumetric frames in one heartbeat. The slice spacing is 5 mm or 8 mm. Manual segmentation labels of the left ventricular endocardium (LVE), pericardium (LVP) and the right ventricular endocardium (RVE) at end-diastolic (ED) and endsystole (ES) phases are provided. In the paper, we divided these 100 patients into 80 and 20 patients for training and testing respectively. The pairs of two frames with the time interval of 4 in one 4D sequence, i.e. (t0, t4), (t1, t5), (t2, t6), etc., were extract as the pair of the fixed and moving images for the registration. The total number of 472 and 58 image pairs were used for network training and testing respectively.

The second dataset contains cardiac CTA images of 20 patients, acquired on a Siemens Dual-source 64-detector CT scanner. Each 4D CT image consists of 10 volumetric frames in a cardiac cycle. The pixel size ranges from 0.314 to 0.4 mm² and the slice spacing is 0.5 mm. The dataset was split into 15 and 5 patients for network training and testing. Every two consecutive frames in one 4D image, i.e. (t0, t1), (t1, t2), (t2, t3), etc., were extracted as the fixed and moving image for the registration. The total number of 135 and 45 image pairs were used for network training and testing.

Considering of the limitation of GPU memory, the 3D ROIs of heart were cropped from the original images and

resampled to the same axial size of 128 × 128. The number of slices of the ROI extracted from MR and CT datasets are fixed to 10 and 112 respectively. The voxel intensity of all the images from these two datasets was normalized to [0, 1] before fed into the network. The segmentation labels at ED and ES phases in the first dataset were only used for computing quantitative measurements to evaluate the performance of the registration and not used for the training of our network respectively.

B. Training details

The proposed network was implemented in the TensorFlow framework on a workstation of Intel i5-7500 CPU, 8G RAM and NVIDIA GTX1080 GPU. We trained the network using Adam solver with a learning rate of 0.0001 for the first 100 epochs and then linearly declined over the next 50 epochs. The batch size equals to 8 for ACDC dataset and 1 for cardiac CT dataset due to the limitation of GPU memory. We set $\alpha = 1$ and $\beta = 5$ in Eq. (4) and Eq. (5). In addition, considering that the number of images in the Z -axis of the ACDC dataset is relatively small, only one down-sampling operation is performed in the Z -axis direction in ConvNets part, and four down-sampling operations are still performed in the X -axis and Y -axis direction.

C. Evaluation

For the test images in the first dataset, the reference segmentation labels at ED phase were deformed to generate the LVE, LVP and RVE labels at ES phase by a series of predicted DVFs ϕ . Dice coefficients [20] between the deformed segmentation labels and the reference segmentation labels were computed to evaluate the accuracy of the registration algorithm.

We compared our method with two existing methods. The first one is a traditional B-spline non-rigid registration algorithm with Elastix [21], advanced Mattes mutual information was used as similarity metric. The other one is a deep learning based unsupervised image registration algorithm VoxelMorph [11] implemented in Keras framework.

The dice coefficients computed in the ACDC dataset are summarized in Table I. We let $\alpha = 1$ and $\beta = 5$ in Eq. (4) and Eq. (5) to demonstrate the effectiveness of the cycle identical loss L_{cyc} . As shown in Table I, our method without cycle identical loss in CIRNet-1 outperforms VoxelMorph, which means that the cascaded network architecture of CIRNet can better predict the DVF. After combing cycle identical loss in CIRNet-2, our method has achieved the best dice coefficients of LVE, LVM and RVE, which are 83.56%, 80.81% and 72.58% respectively. Furthermore, all deep learning-based methods have large improvement of computation time compared to the traditional one, especially running on GPU. Compared to the VoxelMorph, our proposed method does not have significant increase of computational cost.

In the second dataset, considering that manual delineation of cardiac chambers is a large amount of labor work, we compared the image similarity of two images by subtracting

TABLE I

DICE COEFFICIENTS AND COMPUTATION TIME OF THE REGISTRATION IN THE ACDC DATASET .

Methods	Dice coefficients			GPU time(ms)	CPU time(ms)
	LVE(ES)	LVM(ES)	RVE(ES)		
Before registration	64.43%	53.25%	63.72%		
Elastix	79.43%	78.38%	70.31%	...	12396.97
Voxelmorph	80.86%	77.74%	71.29%	201.23	3347.03
CIRNet-1 (w/o L_{cyc})	81.24%	79.46%	71.79%	221.01	4502.78
CIRNet-2 (w/ L_{cyc})	83.56%	80.81%	72.58%	245.45	4535.84

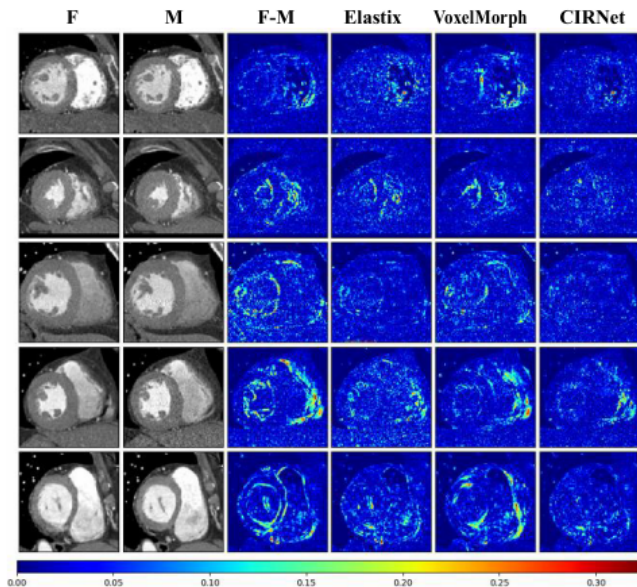


Fig. 3. Original images and difference images of five patients. The difference of image intensity in five testing CT images are displayed in color. The first two columns show fixed image and moving image respectively. The third column shows the difference between the moving image and the fixed image. The next three columns display the difference between the warped image and the fixed image obtained by the Elastix, the VoxelMorph and the CIRNet.

the warped image from the fixed image. Fig. 3 shows the fixed image, the moving image and the difference between the warped image and the fixed image in color for three different methods. As displayed in Fig. 3, the results of the CIRNet achieved the best results with the minimum mean and the variance of 0.022 ± 0.0007 . The average mean and the variance obtained by the Elastix and the Voxelmorph are 0.035 ± 0.0020 and 0.035 ± 0.0015 respectively. It indicates that the CIRNet with the loss L_{cyc} achieved the best performance among three image registration methods.

IV. CONCLUSION

In this paper, an end-to-end unsupervised convolutional neural network CIRNet for registration of 3D medical images is presented. The CIRNet adopted cascaded network architecture to estimate the DVF. The cycle identical loss introduced in this paper can further improve the accuracy of image registration. Our network CIRNet has been tested on the datasets of cardiac CT and MRI. Quantitative evaluation results show that our method achieves the best registration

accuracy than the other two registration methods. In addition, the CIRNet can perform image registration much faster than the traditional methods.

REFERENCES

- [1] Mark Jenkinson and Stephen Smith, "A global optimisation method for robust affine registration of brain images," *Medical image analysis*, vol. 5, no. 2, pp. 143–156, 2001.
- [2] Stefan Klein, Marius Staring, and Josien PW Pluim, "Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines," *IEEE transactions on image processing*, vol. 16, no. 12, pp. 2879–2890, 2007.
- [3] J-P Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Medical image analysis*, vol. 2, no. 3, pp. 243–260, 1998.
- [4] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," *NeuroImage*, vol. 45, no. 1, pp. S61–S72, 2009.
- [5] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436, 2015.
- [6] Hessam Sokooti, Bob de Vos, Floris Berendsen, Boudewijn PF Lelieveldt, Ivana Išgum, and Marius Staring, "Nonrigid image registration using multi-scale 3D convolutional neural networks," in *Medical Image Computing and Computer Assisted Intervention - MICCAI 2017*, 2017, pp. 232–239.
- [7] Marc-Michel Rohé, Manasi Datar, Tobias Heimann, Maxime Sermesant, and Xavier Pennec, "SVF-Net: Learning deformable image registration using shape matching," in *Medical Image Computing and Computer Assisted Intervention - MICCAI 2017*, 2017, pp. 266–274.
- [8] Yipeng Hu, Marc Modat, Eli Gibson, Nooshin Ghavami, Ester Bonmati, Caroline M Moore, Mark Emberton, J Alison Noble, Dean C Barratt, and Tom Vercauteren, "Label-driven weakly-supervised learning for multimodal deformable image registration," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 1070–1074.
- [9] Bob D de Vos, Floris F Berendsen, Max A Viergever, Marius Staring, and Ivana Išgum, "End-to-end unsupervised deformable image registration with a convolutional neural network," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 204–212, 2017.
- [10] Hongming Li and Yong Fan, "Non-rigid image registration using fully convolutional networks with deep self-supervision," *arXiv preprint arXiv:1709.00799*, 2017.
- [11] Guha Balakrishnan, Amy Zhao, Mert R Sabuncu, John Guttag, and Adrian V Dalca, "An unsupervised learning model for deformable medical image registration," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 9252–9260.
- [12] Jingfan Fan, Xiaohuan Cao, Zhong Xue, Pew-Thian Yap, and Ding-gang Shen, "Adversarial similarity network for evaluating image alignment in deep learning based registration," in *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018*, 2018, vol. 11070 of *Lecture Notes in Computer Science*, pp. 739–746.
- [13] Jonathan Long, Evan Shelhamer, and Trevor Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.
- [14] Max Jaderberg, Karen Simonyan, Andrew Zisserman, et al., "Spatial transformer networks," in *Advances in neural information processing systems*, 2015, pp. 2017–2025.

- [15] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems 27*, pp. 2672–2680. 2014.
- [16] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015*, Nassir Navab, Joachim Hornegger, William M. Wells III, and Alejandro F. Frangi, Eds., 2015, vol. 9351 of *Lecture Notes in Computer Science*, pp. 234–241.
- [17] Sergey Ioffe and Christian Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on Machine Learning, ICML*, Francis R. Bach and David M. Blei, Eds., 2015, vol. 37 of *JMLR Workshop and Conference Proceedings*, pp. 448–456.
- [18] J.N. Sarvaiya, S. Patnaik, and S. Bombaywala, "Image registration by template matching using normalized cross-correlation," in *2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies*, 2009.
- [19] O. Bernard, A. Lalande, C. Zotti, F. Cervenansky, X. Yang, P. Heng, I. Cetin, K. Lekadir, O. Camara, M. A. Gonzalez Ballester, G. Sanroma, S. Napel, S. Petersen, G. Tziritas, E. Grinias, M. Khened, V. A. Kollerathu, G. Krishnamurthi, M. Rohé, X. Pennec, M. Sermesant, F. Isensee, P. Jäger, K. H. Maier-Hein, P. M. Full, I. Wolf, S. Engelhardt, C. F. Baumgartner, L. M. Koch, J. M. Wolterink, I. Išgum, Y. Jang, Y. Hong, J. Patravali, S. Jain, O. Humbert, and P. Jodoin, "Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved?," *IEEE Transactions on Medical Imaging*, vol. 37, no. 11, pp. 2514–2525, Nov. 2018.
- [20] Lee R Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [21] Stefan Klein, Marius Staring, Keelin Murphy, Max A Viergever, and Josien PW Pluim, "Elastix: a toolbox for intensity-based medical image registration," *IEEE transactions on medical imaging*, vol. 29, no. 1, pp. 196–205, 2009.