**SPECIAL ISSUE PAPER**

# 3D shallow deep neural network for fast and precise segmentation of left atrium

Asma Kausar[1] · Imran Razzak[2] · Mohammad Ibrahim Shapiai[1] · Amin Beheshti[3]

**Abstract**

Knowledge of the underlying anatomy of the left atrium can promote improved diagnostic protocols and clinical interventions. Hence, an automatic segmentation of the left atrium on magnetic resonance imaging (MRI) can support diagnosis, treatment and surgery planning of heart. However, due to the small size of left atrium with respect to the whole MRI volume, accurate segmentation of left atrium is challenging. Most of the existing deep learning approaches are based on cropping or cascading networks. In this work, we present a novel deep learning architecture for the segmentation of left atrium from MRI volume which incorporates the residual learning based encoder-decoder network. We introduce a loss function and parameter adjustments to deal with the issue of class imbalance and unavailability of large medical imaging dataset. To facilitate the high quality segmentation, we present a three-dimensional multi-scale residual learning based architecture that maintains coarse and fine level features throughout the network. Experimental results have shown a considerable improvement in segmentation performance by surpassing the current benchmarks (especially the winner of Left Atrial Segmentation Challenge-2018) with fewer parameters compared to the state-of-the-art approaches, thus potentially supporting cardiac diagnosis and surgery without adding any extensive pre-processing of input volumes or any post-processing on the base network's output.

## 1 Introduction

Body imaging technology has become one of the most important revolutions in medicine for its contribution to medical diagnosis, prevention and prognosis during the past many years. It allows researchers and cardiologists to assess the size of heart, determine its dimensions and evaluate its mechanical work using X-Rays [19], magnetic resonance imaging (MRI) [9] and echocardiography imaging [23] respectively. MRI provides good quality imaging, excellent contrast of soft-tissue with the abstraction of ionizing radiation, therefore becomes the gold stranded modality to precisely identify patients cardiac etiology and structures, to guide therapy and diagnosis decisions [30]. Lately, the evolution of late gadolinium-enhanced MRI (LGE-MRI) has enabled the perception of scar tissue which is located inside myocardium. This technique helps understand and analyze the part of fibrosis and underlying structures that support atrial fibrillation (AF) which is the most common cardiac arrhythmia, expected to develop into a new epidemic in the decades ahead. An increase in the amount of fibrosis in the left atrial (LA) wall, correlated with poor atrial fibrillation ablation results, [28] is demonstrated as a result of atrial fibrillation. MRI has now become a widely accepted technique of choice to quantify and detect scar tissues located in the atrial wall (Fig. 1).

The medical management of AF patients is based on crucial task of atrial segmentation to perform structural analysis of the segmented 3D geometry. The most frequently used clinical practice for analyzing atrial structures and

✉ Asma Kausar
aftabasma@graduate.utm.my

Imran Razzak
imran.razzak@deakin.edu.au

Mohammad Ibrahim Shapiai
md_ibrahim83@utm.my

[1] MJIIT, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

[2] Deakin University, Geelong, Australia

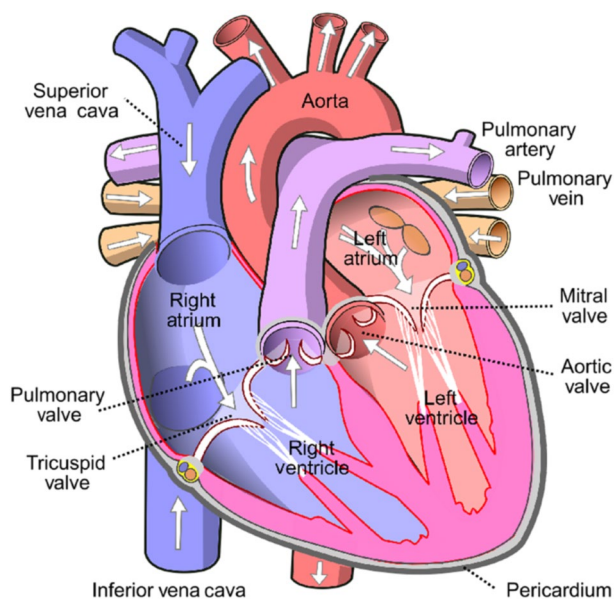[3] Macquire University, Sydney, Australia

**Fig. 1** Human heart illustration by Eric Pierce. Front view of heart shows upper chamber (Left atrium) through which blood enters the ventricles of the heart (en.wikipedia.org/wiki/Atrium_(heart))

determining fibrosis distribution is by performing manual segmentation of left atrium (LA) chamber from MRIs. The LA cavity is confined to thin atrium wall (2–3 mm) and represents a small volume ($73 \pm 14.9 cm^3$) having complex anatomy [27, 44]. Moreover, anatomical structures around the atrium can mislead some segmentation algorithms [41]. As a result, manual segmentation of atrium is a labor-intensive, error-prone and time-consuming process [5].

Most of the existing studies on structural analysis are based on manual segmentation approaches [1, 28] because of the limitations of current automated methodologies requiring often unavailable supporting information, such as additional magnetic resonance angiography (MRA) sequences to sustain the segmentation process [41] or shape priors for initialization [44]. The standard atrial anatomy is more impaired by the contrasting agent which makes it difficult for most of the algorithms based on conventional approaches, mentioned in the bench marking study [27], to be applied directly to the MRIs. There is an urgent need for an intelligent system that can perform automatic atrial segmentation directly from MRIs for accurate measurement and reconstruction of the atrial geometry for clinical usage.

Before the rise of deep learning, researchers tried to develop automated approaches like Random decision forests and Nearest neighbor pattern classification to reduce the burden of manual segmentation. Earlier algorithms required important manual tuning such as region growing approaches or thresholding methods. Other methods were developed later providing a higher degree of automation using classifiers or clustering approaches such as k-means clustering

[36] or k-nearest-neighbor [22]. More recent methods, using statistical classifiers like active shape model [14], multi-atlases [33] or support vector machine [39] approaches, gained higher interest for medical image analysis and cardiac segmentation. Though many of these approaches presented promising results, yet none showed sufficient consistency to be implemented widely in clinical practice.

The availability of large clinical databases and the development of more powerful computational hardware has now enabled deep learning to attain enormous progress in image segmentation and classification [38]. Deep learning has done remarkably well in image classification and processing tasks, mainly owing to convolutional neural networks (CNN) [11]. Despite the popularity of various traditional and machine learning techniques, the deep learning revolution has turned the tables so that many computer vision problems, mainly semantic segmentation are being dealt with deep learning architectures, mainly convolutional neural networks (CNNs) [7, 8, 7–8, 29], which are outperforming other approaches by a large margin in terms of accuracy and efficiency.

The existing convolutional neural networks (CNNs) are able to learn hierarchies of features making them powerful visual models benefiting in many advancements made in segmentation models. It is an undeniable fact by now that CNNs dominate the state of the art in biomedical image segmentation as well [16, 17, 16–17, 26, 32, 43]. Since the encoder-decoder based architectures for semantic segmentation, such as U-Net [34] and FCN [6, 35] proved to be successful for segmentation, certain similar deep learning architectures have also shown very effective results when applied to cardiac imaging. A three-stage approach proposed by Avendi et al. [2, 3] combines convolutional neural network (CNN), deformable models and stacked encoder to segment the left ventricle for a small MRI dataset of only 45 patients. On the other hand, a large MRI dataset was used by Bai et al. [4] that was provided to develop CNN for ventricular chamber assessment and segmentation by the UK Biobank database. This approach obtained the accuracy scores that competed the human-level precision.

Most of the current approaches use deep learning for the segmentation of left atrium where all promising ones are mainly based on encoder-decoder networks, yielding high accuracy and outperforming conventional segmentation approaches [31]. The leading approach for automatic left atrium segmentation directly from LGE-MRI dataset that has shown best benchmarked performance in 2018 challenge involves a two-stage 3D CNN method [45]. This approach is based on U-Net-like network with volumetric convolutions where the class imbalance is reduced effectively in the first network where the background isotropy is optimized with the help of dynamic cropping which provides the next network with a focused region for further localized segmentation. Extensive data augmentation is also employed to enhance

the generalization capability of this approach. Finally, a 3D approach is employed to reinforce the spatial representation of features. Another U-Net inspired architecture was trained on large input patches [18] to be employed in Medical Segmentation Decathlon 2018 allowing the network to capture as much contextual information as possible. One of the earlier architectures used residual connections [15] in the encoder only. Most of these networks were trained with deep supervision and multiclass Dice loss in order to improve the gradient flow.

Main hindrances faced in left atrium segmentation are its complex anatomy, blurry boundaries, noisy backgrounds and its small size. 3D CNNs attempt to fully utilize 3D image information but they have limited field of view, therefore most of the 3D U-Net based models apply two staged networks or hybrid approach. These models provide higher performance along with goof class imbalance management but they are all computationally as well as timely expensive. In this paper, we present a novel 3D shallow residual segmentation network (3D SR-Net) that takes an initial 3D volumetric image directly to the network for initial convolution layer. The proposed model is able to deal with the problem of class imbalance as well as the issue of unavailability of larger training dataset. We applied proposed network for the task of left atrium segmentation and in-depth analysis shows considerable improvement in segmentation results. The **key contributions** of this work are summarised as follows:

- We present a novel deep interleaved 3D Shallow Residual Network based on residual learning (3D SR-Net) for end to end volumetric segmentation of left atrium that consist of fewer parameters as compared to most of the state-of-the-art approaches especially the winner of Left Atrial Segmentation Challenge—LASC 2018.
- An interleaved residual network is designed to accomplish effective extraction low-level features and high-level features by using identity shortcuts within each stage. These skip connections are implemented through element-wise summation which adds up information in features block while resolving the issue of exponential growth of parameters and over-fitting.
- We have applied a loss function that is the based on $L_2$ regularisation loss with label-smoothed probabilistic Dice scores leveraged across subjects in each mini-batch for each class.
- We have applied the combination of instance level normalization in between each convolution layer that results in learning the better feature representation.

## 2 Methodology

The segmentation of Left Atrium from MRI volumes is considered an important step towards understanding the structure of heart which provides guidance towards therapy and

diagnosis decisions. Since the size of Atrial cavity is very small as compared to the whole image volume, it develops a huge class imbalance among the atrial structures and their background which severely impairs the learning process. Most of the deep learning techniques use cropping or cascaded network to reduce the background information which are mostly memory inefficient or computationally expensive. Furthermore, over-fitting is a very common issue of deep learning models when applied for Left Atrium segmentation due to the unavailability of large annotated dataset in heart MRI. Therefore, most of the models use as many types of augmentation techniques as available. The proposed model provides better results without adding any manual step or additional network pass to reduce the class imbalance. We have used only one augmentation technique based on affine perturbation to deal with small dataset. Our model is based on loosely connected convolutions in an encoder–decoder type segmentation model that enables high-resolution activation, mapping through feature reuse and dropout while keeping the network memory-efficient.

### 2.1 3D SR-Net segmentation

The proposed interleaved shallow residual network model, as shown in Fig. 2, takes an initial 3D image directly to the network for initial convolution layer. The standard input to each convolutional layer is added to its output creating a residual connection which is then followed by instance normalization and pooling creating the first downsampled output. The same way inputs are downsampled four times and the fifth time convolutional layer is applied without downsampling which keeps the same size of output. Afterwards, upward convolutions are applied in the same way following the same trend of adding skip connections and spatial information from same level of downsampling layer. This makes the network to consider both high-level features and low-level features extraction by using identity shortcuts within each stage.

Leaky ReLU (leakiness $10^{-2}$) is introduced as an activation function at each convolutional layer while the final layer is concluded using Softmax activation function to achieve final segmentation channels per class. The exponential moving averages of batch mean and variance obtained through batch normalization are unstable in the small batch of size 2, so they cannot reflect the activations of feature map very well during testing. We used instance normalization [42] to normalize all feature map activations (between non-linearity and convolution) to provide more consistent results.

### 2.1.1 Skip connections

Some of the information captured in initial layers of a segmentation task is beneficial which we would like to allow
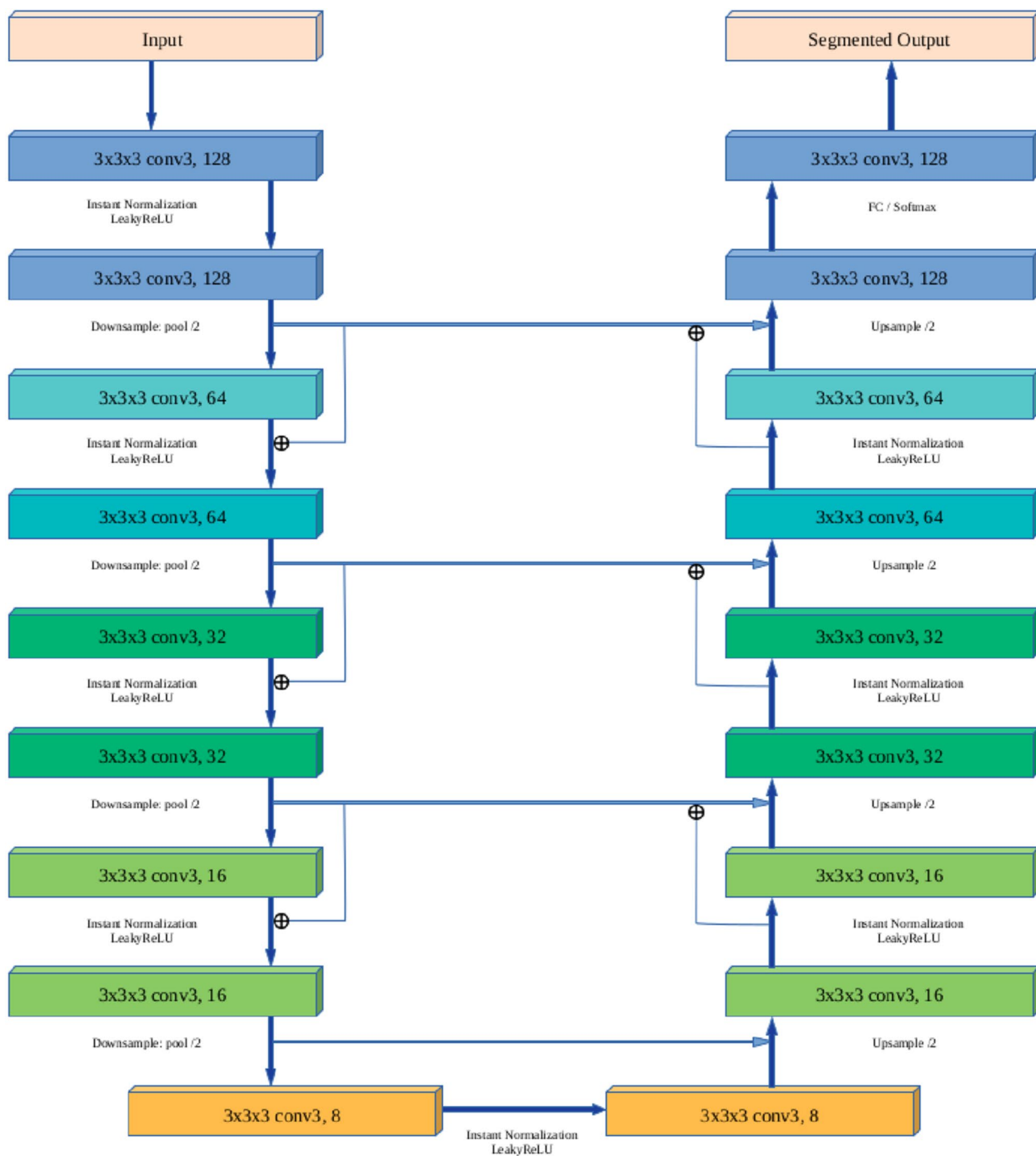
**Fig. 2** Detailed proposed 3D-SRNet model

for the later layers to also learn. In earlier layers of segmentation network, the learned features correspond to lower semantic information that is extracted from the input. Such information would be turned to abstract if we do not use the skip connections. Therefore, skip connections in residuals networks provide us a way to preserve lower semantic information learned at initial layers. The motivation behind skip connections is that they have an uninterrupted gradient

flow from the first layer to the last layer, which tackles the vanishing gradient problem.

Standard deep learning architectures with skip-connections use element-wise summation which can be viewed as an iterative estimation procedure where the features are refined through the various layers of the network. The main benefit of this choice is that it is a compact solution for keeping the number of features fixed across a block while adding information. It also resolves the issue of an exponential growth of the parameters due to the use of a lot of information being generated by concatenation in dense networks. It has also been experimentally validated [24] that the loss landscape changes significantly when introducing skip connections.

The proposed model introduces skip connections via addition (also known as residuals) after each block of convolution during downsampling as well as upsampling as shown in Fig. 3. This way the gradient is simply multiplied by one and its value will be maintained in the earlier layers instead of the gradient becoming very small as we approach the earlier layers in standard deep architecture. We use an identity function $F(x) = F(x) + x$ to preserve the gradient. The proposed model uses two types of skip connections. Firstly, we pass features from each level of the encoder path to corresponding level of the decoder path using long skip connections in order to recover spatial information which is lost during downsampling. Secondly, we use shorter skip connections to stabilize gradient updates in the architecture. Therefore, enabling feature reusability and stabilize training and convergence through skip connections.

### 2.1.2 Loss function

One of the major issue faced during left atrium segmentation is the overfitting of segmentation model because the size of required segment is substantially smaller as compared to the background of the image. Therefore, to resolve this huge class imbalance we need to use a loss function that specifically targets the foreground. We have applied a loss function that is the weighted sum of an $L_2$ regularisation loss with label-smoothed [10] probabilistic Dice scores for each class leveraged across subjects in each mini-batch.

$$\text{pDice}_l(L_l'', R_l) = \overline{\left( \frac{\min(L_l'', 0.9).R_l}{||R_l||_2 + ||\min(L_l'', 0.9)||_2} \right)}, \qquad (1)$$

where $L_l'' = \text{softmax}(L')_l$ and $R_l$ are the vectors of algorithm's probabilistic segmentation and the binary reference standard segmentation for organ $l$ of each subject, respectively. To further mitigate the extreme class imbalance, Dice scores below 0.01 and 0.10 were respectively introduced after warm up periods of 25 and 100 iterations. The loss function at iteration $i$ was

$$loss(L_l'', i) = \sum_{\forall W} \frac{\bar{w}^2}{40} - \frac{1}{8} \sum_{l=1}^{8} d(\text{pDice}(L_l'', R_l), i) \qquad (2)$$

$$d(l, i) = l + 100h(l, i, 0.01, 25) + 10h(l, i, 0.1, 100) \qquad (3)$$

$$h(l, i, v, t) = \text{sigmoid}(6(i - t)/t)(\max(v - l/v)^4), \qquad (4)$$

where $l$ is the Dice loss, $w \in W$ are the kernel values, $v$ is the threshold of hinge loss, and $t$ is the iteration's delay. Adam optimizer was used to train the network with $e = 0.001$ for 5000 iterations. Training of the proposed network took less than 17 hours using two GeForce GTX 1080 P8 GPUs.

### 2.1.3 Activation function

As we know that the training of neural network is based on updating of the weights and biases depending on the error received on the output during back-propagation. Activation functions make the back-propagation work by supplying gradients along with the error to update the biases and weights. Rectified Linear Activation Unit (ReLU) is linear for all positive values, and zero for all negative values. It is therefore cheap to compute as there is no involvement of complicated math computation and it is sparsely activated. The model would therefore take less time to train and converge faster. On the other had, once a neuron gets negative, it is unlikely for it to recover because the slope of ReLU in the negative range is always 0. Such neurons cannot play any role in discriminating the input and therefore become practically useless or dead. If we do not get rid of such neurons, with the passage of time a large part of our network might end up doing nothing. A large bias or too high learning rate accelerates the network to fall into this dying problem of
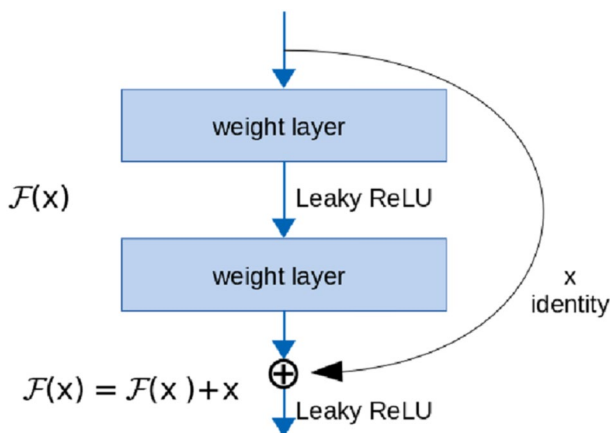


**Fig. 3** Residual skip connection

neurons. We can get a slope out of ReLU as long as there is still any live neuron present.

Leaky ReLU is a type of ReLU which has a small slope for negative values, instead of utterly zero. For example, leaky ReLU may have $y = 0.01x$ when $x < 0$. So the leaky ReLU fixes the dying ReLU problem as it does not have zero-slope part. Unlike ReLU, leaky ReLU is more balanced, and may therefore learn faster.

## 3 Experiments

In this section, we present comprehensive analysis of our proposed network for the task of left atrium segmentation. To generalize the performance, we have performed fivefold validation and compared the performance with state of the art methods of left atrial segmentation. In our experiment, we have used the NiftyNet repository [25] to build the proposed network model and to gather the results of other benchmark models for comparison. System parameters were kept constant to generate a fair comparison.

### 3.1 Dataset

One of the focal limitations of medical datasets is small training datasets as the annotation of data gathering is labor intensive and time consuming. We used twenty cases of heart dataset for this study which are provided by King's College London (London, United Kingdom) and were originally released through the Left Atrial Segmentation Challenge (LASC) [40]. This dataset includes MRIs of the entire heart that were acquired during a single cardiac phase. Images were obtained with the voxel resolution of $1.25 \times 1.25 \times 2.7mm^3$ on a 1.5T Achieva scanner. The left atrium appendage, mitral plane, and portal vein end points were segmented by an expert using an automated tool followed by manual correction [37].

### 3.2 Evaluation parameter

The best way to measure the accuracy of a segmentation depends on the consequences that any error in the segmentation might have and the variety of purposes segmentation could serve. If we are working on the segment which has to be monitored for changes in size then we need to apply an evaluation parameter that is sensitive to misestimations of the volume more than anything else. Therefore, in such situations we could use volume error: $VE = 2(V1 - V2)/(V1 + V2)$ and the absolute volume error if we want to average multiple measurements. On the other hand if we use the segmentation for the purpose interested in the shape fidelity of the segmentation outline of the true segment and any over-inclusion and under-inclusion

matters a lot in diagnosis. In such situation, surface distance measure like Hausdorff metric is more useful.

Overlap ratio measures are a compromise that applies to many situations. Most of the widely used overlap ratio measures range from 0 to 1 where 0 represents no overlap and 1 represents complete congruence. In contrast to volume error, they are sensitive to misplacement of the segmentation labels, but they are relatively insensitive to volumetric under-estimations and over-estimations. Shape infidelity is captured only if the deviation has volumetric impact.

A successful prediction is the one showing maximum overlap between the predicted and true volumes. Two most popular and related but different metrics for finding the maximum overlap are the Dice and Jaccard indices. The Dice similarity index is currently more popular than the Jaccard overlap ratio. Jaccard is more sensitive to mismatch numerically when there is reasonably strong overlap.

The Jaccard index is also known as Intersection over Union (IoU). In terms of the confusion matrix, the dice score metric can be rephrased in terms of true/false positives/negatives:

$$Dice(A, B) = \frac{2TP}{2TP + FP + FN}, \qquad (5)$$

$$Jaccard = \frac{Dice}{2 - Dice}, \qquad (6)$$

$$F\text{-score} = \frac{TP}{TP + \frac{1}{2}(FP + FN)}. \qquad (7)$$

For the evaluation of proposed model, we have applied Dice similarity using Eq. 5 and therefore calculated the respective Jaccard index based on Dice score as given in Eq. 6. We have also calculated the F-score based on the Eq. 7 for each model. F-score (F1 score) gives the harmonic mean of precision and recall. The results are recorded in Table 1 for the proposed model as well as the competitor models while keeping the parameters same in order to get fair comparison.

### 3.3 Model parameters

Hyper-parameters are very important as well as crucial to be tuned properly for better performance of deep learning

**Table 1** Evaluation of different models for Left Atrium Segmentation

| Model/evaluation | 3D SR-Net | nnUNet | Dense V-Net |
|---|---|---|---|
| Jaccard Index | **0.8743** | 0.8564 | 0.8630 |
| Dice similarity | **0.9329** | 0.9277 | 0.9265 |
| F1-score | **0.8237** | 0.8110 | 0.7389 |

**Table 2** Parameters of different UNet based models

| Model/parameter | Dense VNet | HighRes3DNet | ScaleNet | nnUNet | VNet | 3D SR-Net |
|---|---|---|---|---|---|---|
| Activation function | LeakyReLU | ReLU | PReLU | Leaky ReLU | PReLU | Leaky ReLU |
| Loss type | Dice Hinge | Dice | Dice | Dice | Dice | Dice Hinge |
| Normalization | Batch | Batch | Batch | Instance | Batch | Instance |
| Network depth | 12 | 20 | 20 | 10 | 10 | 10 |

networks. We fixed the size of spatial window as [128, 128, 128] with kernal size of $3 \times 3 \times 3$ and batch size of 2 with queue length of 36. We have used extent of the affine perturbation=0.1 (0.0 gives no perturbation and 1.0 gives the largest perturbation) as an additional data augmentation technique which is very common in medical imaging due to similar shape of organs and pose in medical images. Training is done using the ADAM optimizer with the learning rate $lr = 10^{-3}$.

Results are gathered on fivefold cross-validation over all subjects with 40% cases for training and 30% cases for both validation and testing. We compared the accuracy of the proposed model to each competitor as shown in Table 1 using Dice score which measures the relative volumetric overlap between segmentations. It is similar to F1-score which is the harmonic mean of precision and recall. These matrices provide fair comparison in background biased images of left atrium by targeting the foreground only.
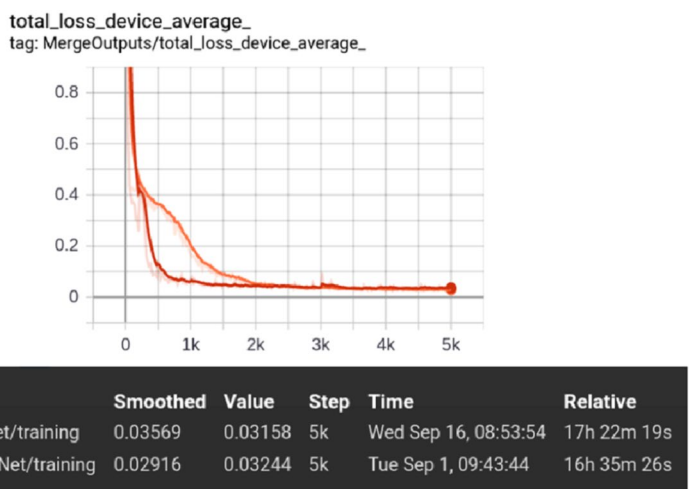
## 4 Results and discussion

The left atrium consists of an appendage, a venous component, and a vestibule. Anatomical variations in each of these components lead to the difficulty in performing safe and effective interventional cardiac procedures. An extremely small size of LA in comparison to the whole image is the reason of highly unbalanced label volumes segmentation,

therefore it can get stuck with very low probabilities very early in training, and never recover. The hinge loss weighs gradients for such segments with very low dice scores helps to more heavily recover these classes. Therefore, when the classes' dice scores reach around 0.1, the risk of stuck classes decreases. 3D SR-Net model shows fast and smooth learning without getting stuck in local optima as shown in Fig. 4, which also shows a comparison with the learning of nnUNet model which implements a different loss function based on cross entropy and dice loss. The segmentation models using the large size spatial window works well with smaller batch size only and therefore instance normalization and Leaky ReLU non-linearities reliably produce the desired results in proposed model.

We provided qualitative results of the model on 30% test set of heart dataset for left atrium segmentation. The Jaccard index, Dice score and F-score of the proposed model in comparison with two latest benchmark models are presented in Table 1. We can see that the proposed model has outperformed others by showing 93% performance gain. The 3D SR-Net has also shown faster convergence rate during the training phase than the other two competitors. Illustrations of sample outputs and their ground truth are provided in Fig. 5. These illustrations show the 3D axial, coronal and sagittal view of test case of heart MRI along with left atrium segmented in red colour. A clear higher performance can be seen in axial and coronal view where the proposed 3D

**Fig. 4** Training loss graph showing smoother learning of 3D SR-Net for Left Atrium Segmentation of Task02 Heart in comparison of No-New-Net

(a) Ground Truth - Axial View

(b) Ground Truth - Coronal View

(c) Ground Truth - Sagittal View

(d) No New Net - Axial View

(e) No New Net - Coronal View

(f) No New Net - Sagittal View

(g) Dense V-Net - Axial View

(h) Dense V-Net - Coronal View

(i) Dense V-Net - Sagittal View

(j) 3D SR-Net - Axial View

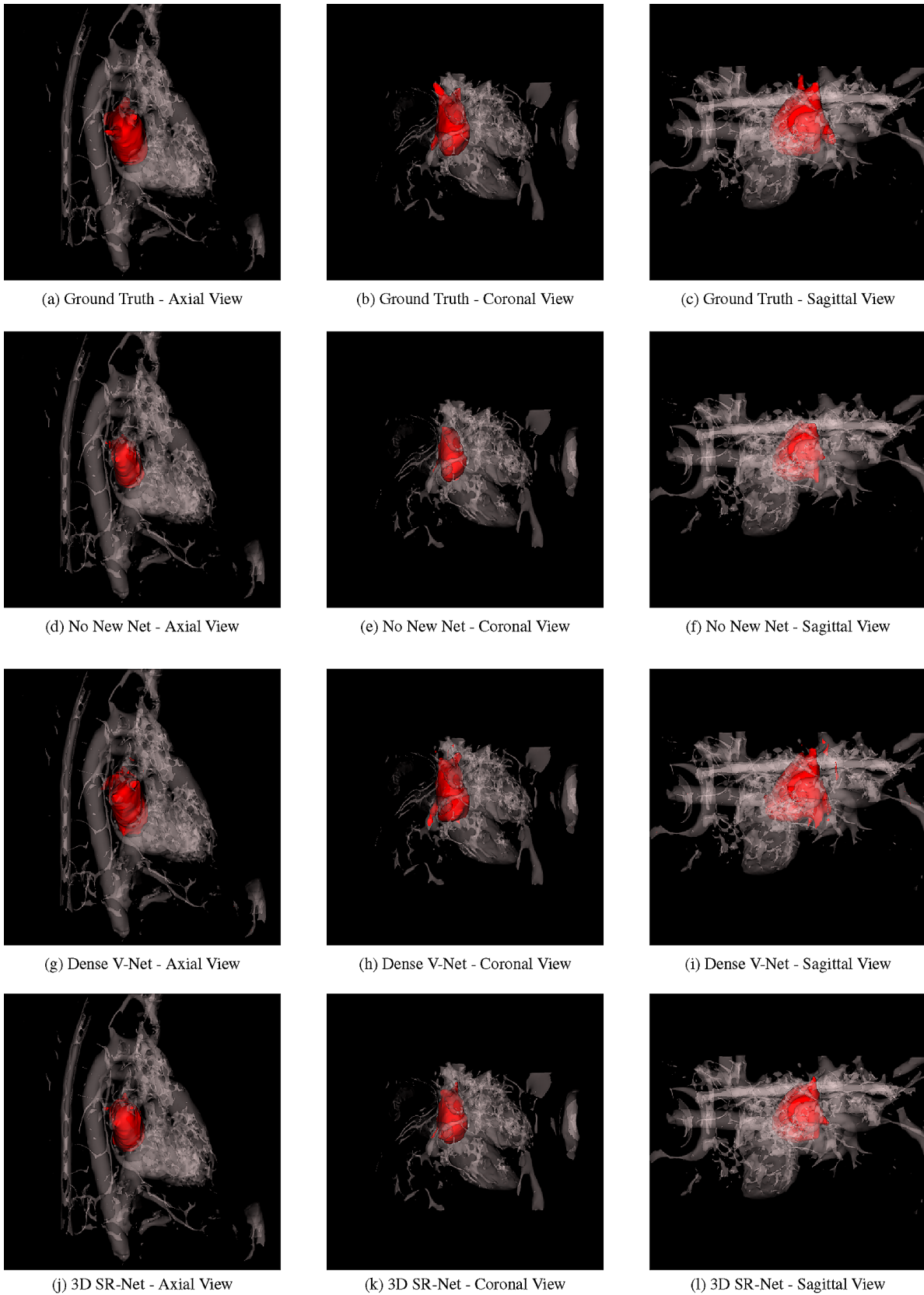(k) 3D SR-Net - Coronal View

(l) 3D SR-Net - Sagittal View

**Fig. 5** Comparison of 3D view of ground truth and predictions of left atrium segmentation in heart by nnUNet, Dense V-Network and proposed 3D SR-Net

SR-Net has extracted the left atrium segment closer to the ground truth more than the other networks.

We have also tested three other state of the art models; HighRes3DNet, ScaleNet and V-Net [25] which are deeper networks. ScaleNet implementation provided in niftynet repository, is based on HighRes3DNet which we kept as default and therefore, both models were having same depth. The backbone structure of ScaleNet is based on VNet model that uses PReLU as an activation function for training. All these models use a simple dice loss function as an activation function for training but they have shown over-fitting when implemented on small dataset of left atrium without any additional pre-processing or post-processing. A detailed comparison of default network architecture and parameters of these models is provided in Table 2. HighRes3DNet and ScaleNet compromise the depth and complexity with the fewer number of training parameters. On the other hand, Dense V-Network [10] compromise on training parameters while keeping less deeper network but involving complex pre-processing and post-processing steps along with higher computation cost due to the complexity of dense model network. No-New-Net (nnUNet) [18] is the model whose encoder-decoder architecture is as shallow as 3D SR-Net but it involves intense pre-processing based on cascading network for the segmentation of left atrium and various post-processing steps which increase its complexity. 3D SR-Net compromises only number of training parameters while keeping the network shallow and vivid; therefore making it suitable for smaller datasets besides potentially improving the performance. It has taken lesser learning time than all other networks except Dense V-Network while showing the performance gain of 1.79% from the benchmark model of No-New-Net.

## 5 Conclusion

In this paper, we have presented a residual interleaved deep 3D shallow network based on residual learning (3D SR-Net) for an end to end volumetric segmentation of left atrium which is designed to accomplish effective extraction of low-level and high-level features by using identity shortcuts within each stage. While our base model is quite strong, enhancing its training procedure by using a combination of Leaky ReLU and Dice Hinge loss, increases its performance significantly while managing the class imbalance issue very smoothly. Introducing skip connections to the proposed model ensures the transfer of generalised features resulting in the improvement of the segmentation results whereas the leaky rectified linear unit and instance normalization handle the neuron death of network while dealing with the overfitting in our base model. Experimental analysis showed that our deep model is able to achieve higher performance with

the depth of only nine levels which is quite less in comparison to most of the efficient deep learning techniques. We have finally concluded that shallow deep models are better than data hungry deep models and are eligible to potentially improve performance even without any involvement of complex pre-processing or post-processing besides.

## References

1. Akoum, N., Fernandez, G., Wilson, B., McGann, C.J., Kholmovski, E.G., Marrouche, N.F.: Association of atrial fibrosis quantified using LGE-MRI with atrial appendage thrombus and spontaneous contrast on transesophageal echocardiography in patients with atrial fibrillation. J. Cardiovasc. Electrophysiol. **24**(10), 1104–9 (2013)

2. Avendi, M.R., Kheradvar, A., Jafarkhani, H.: A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. Med. Image Anal. **30**, 108–119 (2016)

3. Avendi, M.R., Kheradvar, A., Jafarkhani, H.: Automatic segmentation of the right ventricle from cardiac MRI using a learning-based approach. Magn. Reson. Med. **78**(6), 2439–2448 (2017)

4. Bai, W., Sinclair, M., Tarroni, G., Oktay, O., Rajchl, M., Vaillant, G., Lee, A., Aung, N., Lukaschuk, E., Sanghvi, M., Zemrak, F., Fung, K., Paiva, J., Carapella, V., Kim, Y., Suzuki, H., Kainz, B., Matthews, P., Petersen, S., Piechnik, S., Neubauer, S., Glocker, B., and Rueckert, D.: Automated cardiovascular magnetic resonance image analysis with fully convolutional networks 08 Information and Computing Sciences 0801 Artificial Intelligence and Image Processing. J. Cardiovasc. Magn. Reson. **20**(1), (2018). ISSN 1097-6647. https://doi.org/10.1186/s12968-018-0471-x

5. Caudron, J., Fares, J., Lefebvre, V., Vivier, P.-H., Petitjean, C., Dacher, J.-N.: Cardiac MRI assessment of right ventricular function in acquired heart disease: factors of variability. Acad. Radiol. **19**(8), 991–1002 (2012)

6. Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: DeepLab: semantic image segmentation with deep convolutional Nets, Atrous convolution, and fully connected CRFs. IEEE Trans. Pattern Anal. Mach. Intell. **40**, 834–848 (2018)

7. Ciresan, D., Giusti, A., Gambardella, L.M., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. Adv. Neural Inf. Process. Syst. 2843–2851 (2012)

8. Farabet, C., Couprie, C., Najman, L., LeCun, Y.: Learning hierarchical features for scene labeling. IEEE Trans. Pattern Anal. Mach. Intell. **35**(8), 1915–1929 (2012)

9. Gerche, A.L., Claessen, G., Bruaene, A.V.D., Pattyn, N., van Cleemput, J., Gewillig, M., Bogaert, J., Dymarkowski, S., Claus, P., Heidbuchel, H.: Cardiac MRI: a new gold standard for ventricular volume quantification during high-intensity exercise. Circ. Cardiovasc. Imaging **6**, 329–338 (2013)

10. Gibson, E., Giganti, F., Hu, Y., Bonmati, E., Bandula, S., Gurusamy, K.S., Davidson, B.R., Pereira, S.P., Clarkson, M.J., Barratt, D.C.: Automatic multi-organ segmentation on abdominal CT with dense V-networks. IEEE Trans. Med. Imaging **37**, 1822–1834 (2018)

11. Goodfellow, I.G., Bengio, Y., Courville, A.C.: Deep learning. Nature **521**, 436–444 (2015)

12. Gupta, S., Girshick, R., Arbeláez, P., and Malik, J.: Learning rich features from RGB-D images for object detection and segmentation. In: European Conference on Computer Vision, 345–360. Springer (2014)

13. Hariharan, B., Arbeláez, P., Girshick, R., Malik, J.: Simultaneous detection and segmentation. In: European Conference on Computer Vision, 297–312. Springer (2014)

14. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. IEEE Conf. Comput. Vis. Pattern Recogn. (CVPR) **2016**, 770–778 (2016a)

15. He, K., Zhang, X., Ren, S., Sun, J.: Identity mappings in deep residual networks. ArXiv (2016b). arXiv:abs/1603.05027

16. Isensee, F., Jaeger, P. F., Full, P. M., Wolf, I., Engelhardt, S., Maier-Hein, K.: Automatic Cardiac Disease Assessment on cine-MRI via time-series segmentation and domain specific features. In: STACOM@MICCAI (2017a)

17. Isensee, F., Kickingereder, P., Wick, W., Bendszus, M., Maier-Hein, K.: Brain tumor segmentation and radiomics survival prediction: contribution to the BRATS 2017 challenge. In: BrainLes@MICCAI (2017b)

18. Isensee, F., Petersen, J., Klein, A., Zimmerer, D., Jaeger, P. F., Kohl, S., Wasserthal, J., Koehler, G., Norajitra, T., Wirkert, S. J., Maier-Hein, K.: Abstract: nnU-Net: self-adapting framework for U-Net-based medical image segmentation. In: Bildverarbeitung für die Medizin (2019)

19. Joo, H.S., Wong, J., Naik, V.N., Savoldelli, G.L.: The value of screening preoperative chest x-rays: a systematic review. Can. J. Anaesth. **52**, 568–574 (2005)

20. Kamnitsas, K., Bai, W., Ferrante, E., McDonagh, S. G., Sinclair, M., Pawlowski, N., Rajchl, M., Lee, M., Kainz, B., Rueckert, D., Glocker, B.: Ensembles of multiple models and architectures for robust brain tumour segmentation. In: BrainLes@MICCAI (2017a)

21. Kamnitsas, K., Ledig, C., Newcombe, V.F.J., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, B.: Efficient multiscale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. **36**, 61–78 (2017b)

22. Krizhevsky, A., Sutskever, I., Hinton, G. E.: ImageNet classification with deep convolutional neural networks. In: CACM (2017)

23. Lang, R., Badano, L., Mor-Avi, V., Afilalo, J., Armstrong, A., Ernande, L., Flachskampf, F., Foster, E., Goldstein, S., Kuznetsova, T., Lancellotti, P., Muraru, D., Picard, M.H., Rietzschel, E.R., Rudski, L., Spencer, K., Tsang, W., Voigt, J.: Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. Eur. Heart J. Cardiovasc. Imaging **16**(3), 233–70 (2015)

24. Li, H., Xu, Z., Taylor, G., Goldstein, T.: Visualizing the loss landscape of neural nets. In: NeurIPS (2018a)

25. Li, W., Wang, G., Fidon, L., Ourselin, S., Cardoso, M. J., and Vercauteren, T.: On the compactness, efficiency, and representation of 3D convolutional networks: brain parcellation as a pretext task. In: International Conference on Information Processing in Medical Imaging (IPMI) (2017)

26. Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.-W., Heng, P.-A.: H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. IEEE Trans. Med. Imaging **37**, 2663–2674 (2018b)

27. Maceira, A., Cosín-Sales, J., Roughton, M., Prasad, S., Pennell, D.: Reference left atrial dimensions and volumes by steady state free precession cardiovascular magnetic resonance. J. Cardiovasc. Magn. Reson. **12**, 65–65 (2010)

28. McGann, C., Akoum, N., Patel, A., Kholmovski, E., Revelo, P., Damal, K., Wilson, B., Cates, J., Harrison, A., Ranjan, R., Burgon, N., Greene, T., Kim, D.J., DiBella, E., Parker, D., Macleod, R., Marrouche, N.: Atrial fibrillation ablation outcome is predicted by left atrial remodeling on MRI. Circ.: Arrhythmia Electrophysiol. **7**, 23–30 (2014)

29. Ning, F., Delhomme, D., LeCun, Y., Piano, F., Bottou, L., Barbano, P.E.: Toward automatic phenotyping of developing embryos from videos. IEEE Trans. Image Process. **14**, 1360–1371 (2005)

30. Peterzan, M.A., Rider, O.J., Anderson, L.J.: The role of cardiovascular magnetic resonance imaging in heart failure. Cardiac Failure Rev. **2**(2), 115–122 (2016)

31. Pop, M. A., Sermesant, M., Zhao, J., Li, S., McLeod, K., Young, A., Rhode, K. S., Mansi, T.: Statistical atlases and computational models of the heart. Atrial segmentation and LV quantification challenges. In: Lecture Notes in Computer Science (2018)

32. Rehman A, Naz S, Razzak MI, Akram F, Imran M, (2020) A deep learning-based framework for automatic brain tumors classification using transfer learning. Circ Syst Signal PR 39(2):757–775

33. Ren, S., He, K., Girshick, R.B., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. **39**, 1137–1149 (2015)

34. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: MICCAI (2015)

35. Shelhamer, E., Long, J., Darrell, T.: Fully convolutional networks for semantic segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **39**, 640–651 (2017)

36. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. CoRR (2015). arXiv:abs/1409.1556

37. Simpson, A.L., Antonelli, M., Bakas, S., Bilello, M., Farahani, K., van Ginneken, B., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B.H., Ronneberger, O., Summers, R.M., Bilic, P., Christ, P.F., Do, R.K.G., Gollub, M., Golia-Pernicka, J., Heckers, S., Jarnagin, W.R., McHugo, M., Napel, S., Vorontsov, E., Maier-Hein, L., Cardoso, M.J.: A large annotated medical image dataset for the development and evaluation of segmentation algorithms. CoRR (2019). arXiv:abs/1902.09063

38. Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. **15**, 1929–1958 (2014)

39. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. IEEE Conf. Comput. Vis. Pattern Recogn. (CVPR) **2015**, 1–9 (2015)

40. Tobon-Gomez, C., Geers, A.J., Peters, J., Weese, J., Pinto, K., Karim, R., Ammar, M., Daoudi, A., Margeta, J., Sandoval, Z., Stender, B., Zheng, Y., Zuluaga, M.A., Betancur, J., Ayache, N., Chikh, M.A., Dillenseger, J., Kelm, B.M., Mahmoudi, S., Ourselin, S., Schlaefer, A., Schaeffter, T., Razavi, R., Rhode, K.S.: Benchmark for algorithms segmenting the left atrium from 3D CT and MRI datasets. IEEE Trans. Med. Imaging **34**(7), 1460–1473 (2015)

41. Tobon-Gomez, C., Geers, A.J., Peters, J., Weese, J., Pinto, K., Karim, R., Ammar, M., Daoudi, A., Margeta, J., Sandoval, Z.L., Stender, B., Zheng, Y., Zuluaga, M.A., Betancur, J., Ayache, N., Chikh, A., Dillenseger, J.-L., Kelm, B.M., Mahmoudi, S., Ourselin, S., Schlaefer, A., Schaeffter, T., Razavi, R., Rhode, K.S.: Benchmark for algorithms segmenting the left atrium from 3D CT and MRI datasets. IEEE Trans. Med. Imaging **34**, 1460–1473 (2015)

42. Ulyanov, D., Vedaldi, A., Lempitsky, V.: Instance normalization: the missing ingredient for fast stylization. ArXiv (2016). arXiv:abs/1607.08022

43. Wang, G., Li, W., Ourselin, S., Vercauteren, T.K.M.: Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. ArXiv (2017). arXiv:abs/1709.00382

44. Wang, Y., Xiong, Z., Nalar, A., Hansen, B.J., Kharche, S.R., Seemann, G., Loewe, A., Fedorov, V.V., Zhao, J.: A robust

computational framework for estimating 3D Bi-Atrial chamber wall thickness. Comput. Biol. Med. **114**, 103444 (2019)

45. Xia, Q., Yao, Y., Hu, Z., and Hao, A.: Automatic 3D atrial segmentation from GE-MRIs using volumetric fully convolutional networks. In: STACOM@MICCAI (2018)