Automatic 3D Segmentation of Closed Mitral Valve Leaflets on Transesophageal Echocardiogram

Maïlys Hau[1,2][0000−0002−3449−7268], Henrik Kildahl[2][0009−0000−0615−1908], Federico Veronesi[4][0000−0003−3774−206X], Frank Lindseth[1][0000−0002−4979−9218], Bjørnar Grenne[2,3][0000−0002−2984−6865], and Gabriel Kiss[1][0000−0001−5024−1548]

[1] Department of Computer Science, Norwegian University of Science and Technology, Gløshaugen, Trondheim, 7034, Norway
mailys.f.hau@ntnu.no

[2] Department of Circulation and Medical Imaging, Norwegian University of Science and Technology, Prinsesse Kristina, Trondheim, 7030, Norway

[3] Clinic of Cardiology, St. Olavs University hospital, Prinsesse Kristina, Trondheim, 7030, Norway

[4] GE Vingmed Ultrasound, Strandpromenaden, Horten, 3183, Norway

Abstract. Heart disease is a leading cause of death worldwide, with mitral valve (MV) disease being among the most prevalent valvular pathologies. The MV, constitutes a complex three-dimensional apparatus which makes clinical assessment challenging. Therefore, it would be highly desirable to have a patient-adapted model of the mitral annulus shape and its leaflets, both for diagnosis and intervention planning, as well as follow-up purposes.

The main objective of this work is two-fold: improve the valve segmentation's quality using modern architectures and extend it to a sequence of 3D ultrasound recordings for the entire systolic phase. For training purposes, we used a dataset consisting of 108 volumes that were semi-automatically segmented using a commercially available package. We tested several network architectures and loss functions available in the MONAI package to investigate which ones are best suited for the task at hand. We aimed for fast processing times that were usable in practice. Our method was evaluated on 30 recordings and compared to annotations made by two expert echocardiographers. The comparison metrics include Average Surface Distance (ASD), Hausdorff Distance 95% (HSD 95%), as well as standard classification metrics. Our results were a Dice score of 77.06 ± 13.18 % on the evaluation test and distance errors of 0.09 ± 0.12 mm for ASD and 0.49 ± 0.43 mm for HSD 95% and the segmentations were considered comparable to the ground truth provided by clinicians. The proposed annotation method was significantly faster than a manual one and yielded results comparable to the state-of-the-art using a noisier ground truth.

Keywords: Deep learning · Ultrasound · Mitral valve.

1 Introduction

Heart diseases are still a leading cause of death worldwide, and the prevalence is increasing [5]. Early detection is pivotal to decide on the appropriate treatment and follow-up for patients that are at risk. Among heart afflictions, valvular disease is a recurrent type of cardiac pathology that increase with age. Disorders of the mitral valve (MV) are common and may manifest as valve leakage or stenosis. Some disorders require surgery that must be planned in advance, monitored, and controlled after the intervention. Ultrasound (US) is the modality of preference to perform pre-operative, peri-operative and follow-up assessment of the MV and left ventricle, as it is non-invasive. A first diagnosis can be made with a transthoracic echocardiogram (TTE), but for detailed investigation of the MV, a transesophageal echocardiogram (TEE) is required [16].

The usual way of checking the condition of the MV is to measure the mitral valve area by interpolating several recordings of 2D transesophageal echocardiogram (TEE), each recording imaging the valve at a different cross-section. However, this measurement can easily be biased by the 2D plane used. To counter this problem, one can use 3D US. Unfortunately, due to the nature of ultrasound, a trade-off exists between spatial and temporal resolution in 3D, resulting in compromised image quality. Furthermore, ultrasound (US) is tedious to analyse, even for an experienced clinician, given its low image quality compared to other imaging techniques. It is time-consuming and has high interobserver variability. 3D data are even more cumbersome to evaluate. Moreover, different clinicians annotate differently [6], and even the same clinician can annotate the same image differently. This results in high inter- and intra-operator variability, significantly limiting the clinical yield.

This is why we wanted to automate the data analysis process. To do so, we planned on segmenting the mitral valve on 3D-TEE using machine learning (ML) techniques. We focus solely on the closed MV for this work, as other works exist on solving this problem in an open valve configuration, e.g. [3,14]. This yielded a stronger challenge, as the boundary between leaflets is unclear at this stage.

Convolutional neural networks (CNNs) were a breakthrough in image recognition and segmentation. However, they require vast amounts of annotated data, which is hard to acquire in the medical field due to privacy regulations; and near to impossible for 3D data, as annotating them is extremely time-consuming. In 2015, UNet [15] was presented as a solution to the amount of data needed for training. It yielded good results in image segmentation tasks, and its use is now democratized in the field of medical imaging to automate the analysis of data. Therefore, we used this method as our baseline for our work. We also added the nnUNet [8] way of parametrizing a UNet in our baseline as it has shown good results across multiple medical segmentation tasks.

Regarding the mitral valve segmentation on 3D-TEE, some preliminary work exists, such as [14], that uses a general model of the mitral valve, and the multi-atlas method to segment each patient's valve. Multi-atlas method, even though highly accurate, is however very time-consuming – several hours according to

[14]. Hence, some work has been done using machine learning (ML) as a segmentation algorithm, since it is fast to apply once the network is trained. Several UNet types of architectures to segment an *open or semi-open valve* were tested [3], and a Residual UNet achieved the best results at segmenting the leaflets [9]. It has also been shown that pre-training a network on classifying the state of the mitral valve – diastole or systole – helps achieve better Dice score and average surface distance on the mitral valve's leaflets 3D segmentation [4]. Finally, some more recent work has shown that using a multi-head type or architecture worked best when segmenting several classes [13], i.e., anterior and posterior leaflets of the mitral valve.

We base our work on semi-automatic ground truth, which is noisier than a regular, manual segmentation masks, but significantly faster to create. The mask generation is the basis of this work that aims, in addition to a new way of creating a ground truth, to: 1) compare the efficiency of modern architectures for segmenting the MV, 2) implement a multi-head decoder [13] that is able to differentiate between the anterior and posterior leaflets and 3) we prove that a network trained exclusively on the mid-systolic frame can segment the MV in all the systolic frames.

2 Methods

2.1 Dataset

Our dataset is an aggregation of 3D-TEE recordings obtained via General Electric HealthCare (GE HealthCare) Vingmed Ultrasound in Horten Norway and St Olavs hospital in Trondheim, Norway. Volumes were acquired using the Vivid E95 on 141 patients, who had provided informed consent. The probe was left floating and no attempts to improve image quality, other than the usual gain, angulation and depth adjustments, were performed. This way, our data is representative of real operating room images. Our dataset comprises 153 sequences that account for 1047 volumes with a closed MV. The dataset is divided into three collections across the sequences and patients. This means that no volumes from the same acquisition or patients can be found in two different sets. As the patients were all screened for investigating potential illnesses, some of those acquisitions depict MV afflictions. An overview of the data distribution between sets is available in Table 1.

The test set was entirely reviewed and amended by a clinician and a cardiologist, using our semi-automated pipeline (see Section 2.2).

2.2 Semi-automated labeling

Each 3D acquisition was semi-automatically annotated by trained engineers using the AutoMVQ program, a commercially available software from GE HealthCare. This tool exhibited good recognition of the MV geometry when compared

Table 1: Data division between train, validation and test

	Train	Validation	Test	**Total**
Sequence	108	15	30	**153**
Volumes	661	107	279	**1047**

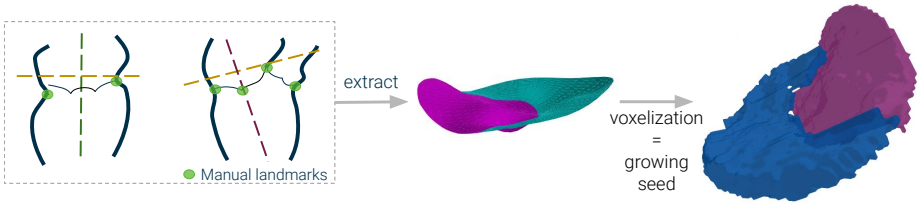


Fig. 1: Annotation process: using AutoMVQ to get the mitral leaflets plane followed by region growing algorithm to voxelize the surface into a bounding box. Pink represents the anterior leaflet and blue the posterior one.

to a more detailed type of imagery [7], and was previously used to quantify the geometry of patients with mitral regurgitation [1]. This program is on par with other vendors [17].

AutoMVQ detects when the valve is closed and allows the user to determine the mitral valve surface semi-automatically by defining six key landmarks as shown in green in the left part of Figure 1 — four on the mitral annulus, the coaptation and the base of the aortic valve. The user focused on the mid-systolic frame while placing these landmarks. The resulting meshes were then automatically propagated to the rest of the systole for further refinement. The surface was then extracted as two triangular meshes from the software, one for each leaflet. Since the meshes have no thickness they cannot be used directly as ground truth for the segmentation mask therefore we built a 3D annotation from there. We assumed the leaflets to have a 4 mm thickness and the mesh to be located in the centre. Subsequently, we produced a voxel grid of the same resolution as the input grid by expanding each triangle of the mesh along its normal vector direction until it reached 2 mm on both sides of the surface. At this point, we have a bounding box surrounding the leaflets. We then refined this bounding box with a growing seed algorithm,[5] to remove voxels with insufficient brightness. We used the US intensity of the voxels located on the *surface* as a seed for this algorithm. We found out that sub-dividing each mesh in two in each direction to initialize the seed worked best[6] — each sub-divided part of a mesh had its own seed.

[5] See this Wikipedia article: https://en.wikipedia.org/wiki/Region_growing
[6] Code to voxelize the surface is available at https://github.com/mailys-hau/echovox

Each 3D frame where the MV is considered closed by AutoMVQ is converted from DICOM to a grey-scale voxel grid of a fixed resolution of 0.5 mm. Some examples of our annotation are visible in Figure 3.

2.3 Model selection

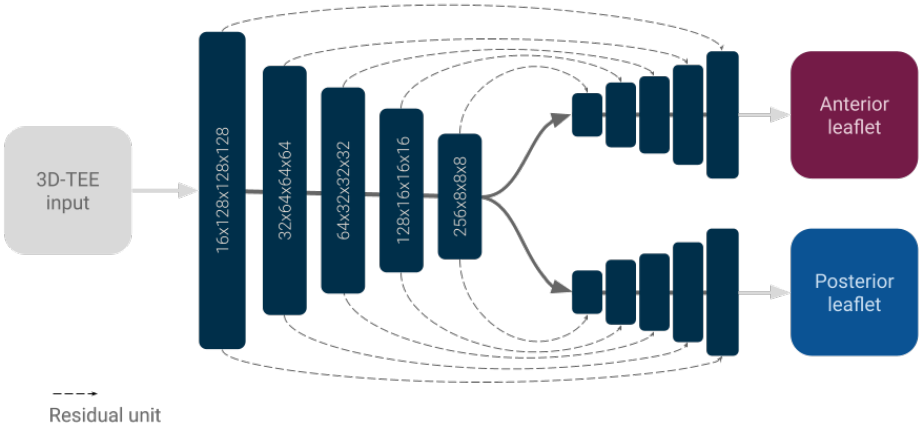


Fig. 2: Network architecture with the best results: multi decoder heads ResUNet[7]

We have considered the following architectures for our experiments: UNet, ResUNet, nnUNet, and multi-head ResUNet [15,9,8,13] and measured their performance on our validation set Table 1. Due to HSD 95% and ASD being computed on CPU, they are only computed for validation, and a validation pass is performed only every 3 epochs. We ranked them using Dice score, Hausdorff Distance 95% (HSD 95%) and Average Surface Distance (ASD).

Out of all the models, the multi-head ResUNet with two heads shown in Figure 2 — one for anterior leaflets and one for posterior leaflets — achieved the best results overall. Without the multi-head decoding part, networks failed to segment the posterior leaflet.

2.4 Implementation details

Our experiments were implemented using the PyTorch Lightning[8] framework, which provides an additional layer to PyTorch for ease of use. Our networks were implemented using MONAI[9] module [2] — or using MONAI as a basis. It is a PyTorch-based module with common network architectures for 2D and 3D data,

[7] Implementation of this network is available at here.
[8] https://www.pytorchlightning.ai/
[9] https://monai.io/

with an emphasis on medical imaging. Our implementation has been made open source on GitHub.[10]

The volumes are loaded and treated as grey-scale. The target is passed as a one-hot encoding of the volume. Each volume is cropped and/or padded to a volume of $128 \times 128 \times 128$. While training the crop is randomly centred. We also performed some data augmentation, including flipping axis, volume distortion, and Gaussian noise, as it helped the network to perform better. In addition, some of our validation metrics were computed on CPU, which increased the training time.

Table 2: Validation metrics for tested network architectures. Accuracies and Dice score are in %, distances are in **voxel**[11]. Each value is taken at the best trained epoch (out of 300) regarding the validation loss

Network	Best epoch	Leaflet	Accuracies			Dice	HSD 95%	ASD
			Accuracy	Precision	Recall			
UNet	185	anterior	77.89	97.77	77.89	86.70	NaN	$+\infty$
		posterior	0	0	0	0	57.48	**26.83**
ResUNet	170	anterior	80.64	98.38	**80.64**	**88.63**	46.01	22.22
		posterior	0	0	0	0	54.56	28.79
nnUNet	212	anterior	67.76	**98.74**	67.76	80.36	44.96	22.91
		posterior	0	0	0	0	**54.25**	30.65
Multi-head ResUNet	236	anterior	**99.56**	68.51	64.55	78.41	**44.59**	**21.76**
		posterior	**99.31**	**64.17**	**55.08**	**70.75**	56.14	29.17

We started by evaluating different training losses on our ResUNet architecture to see which would better suit the task at hand. Each training ran for 300 epochs with a batch size of 4. Based on accuracy results, Dice score, and surface distances, we decided to use a mix of weighted Dice and Focal loss [11] as this helped improve the distance metrics and the precision accuracy.

We tried different network architectures to create a baseline for our dataset. A summary with the best run of each network is available in Table 2 and shows that besides the multi-head ResUNet, no architectures were able to properly discriminate between the two leaflets. Each network was ranked according to the result of its best epoch in regard to its validation loss.

[10] https://github.com/mailys-hau/mitral-leaflets-3dseg

[11] Our direction matrix being different than I_3, we can't obtain the distances in mm simply by multiplying by the spacing. Instead of giving an approximate value in mm, we kept the precise value in voxel for HSD 95% and ASD.

2.5 Evaluation pipeline

Network The multi-head ResUNet Figure 2 was re-trained using both training and validation sets before being finally evaluated. We trained on one Nvidia RTX A6000, which has a graphic memory of 48GB. The batch size used was 4. We used the number of epochs yielding the best results in Table 2 (236) and added 10 to account for the appended data. The training was done using Adam optimizer [10] with a linear warm-up and a cosine with warm restart learning rate scheduler [12]. With those settings, training takes around seven hours to attain epoch 246 — there is CPU bound at validation, cf. Section 2.3. We evaluated our pipeline on a test set containing 30 volumes spread across 26 different patients (cf. Table 1).

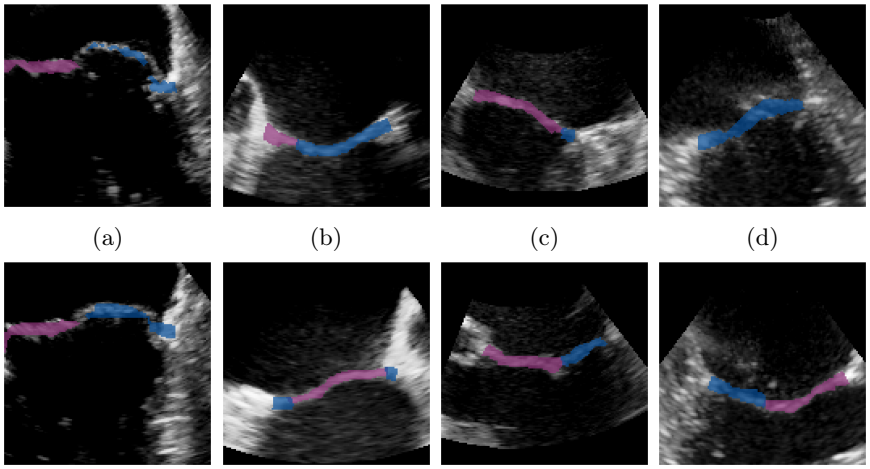
Clinical feedback Due to our annotation being semi-automatic, we consider it essential to present a clinical study of the results. As such, the same clinicians who annotated the evaluation set were presented with a pair of sliced views of a volume (similar to Figure 3). Each volume was presented several times to the clinicians, at various slice indices and along both long-axis (LAX) and its perpendicular. Images were randomly ordered, and the pair ground truth–predictions were arbitrarily placed left or right from one another. Clinicians were asked to choose which segmentations looked best — without knowing which one was the prediction or the ground truth — or "both segmentations look the same".

3 Results

Concerning the time spent annotating the datasets, we demonstrated a much faster process due to the use of semi-automated software. We aggregated the time spent annotating by the expert users over 85 sequences and a file takes on average $4m12s \pm 2m39s$, with only two cases over ten minutes — $21m06s$ and $13m58s$. The annotation time of a file includes: opening the sequence in AutoMVQ, aligning the views and placing the landmarks, reviewing the annotated frames, saving and quitting. A file contains an average of 6.84 systolic volumes that are all annotated during this workflow. In comparison, the manual annotation process described by [13] takes around 15 minutes to annotate one volume.

The evaluation metrics of our multi-head network are presented in Tables 3 and 4 as well as some example sliced views shown in Figure 4. Hausdorff Distance 95% (HSD 95%) represents the average distance between two segmentation after removing the 5% largest distances, while Signed Distance Function (SDF) is the average distance between the two boundaries of the segmentation mask. The SDF is negative when the prediction is outside of the annotation and positive otherwise. The Average Surface Distance (ASD) is the absolute value of the SDF.

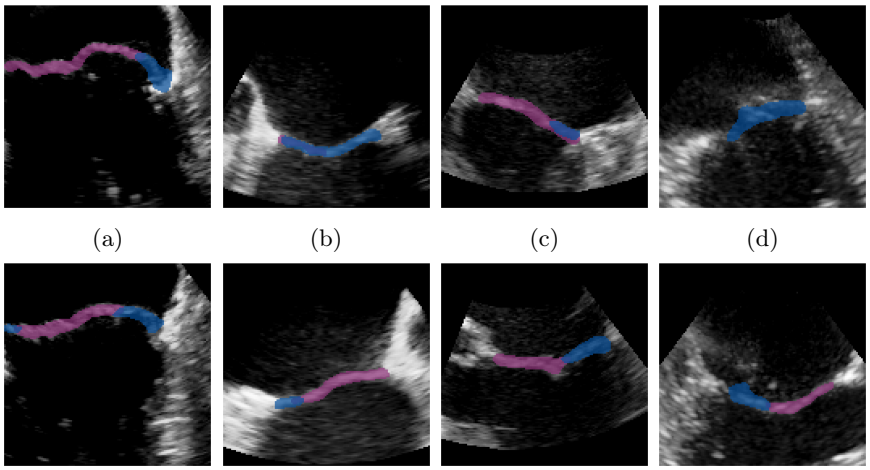
To have a better understanding of the prediction's shape in comparison to the ground truth, a distance comparison is also available in Figure 5 using SDF. Our worst case in shown in Figure 5a.



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 3: Example of **ground-truth** along LAX (top line), and their corresponding slice along the axis perpendicular to LAX (bottom line). Pink is the anterior leaflet, blue the posterior leaflet.



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 4: **Predictions** of our network on the four cases presented in Figure 3

The clinicians' study overview is described in Table 5. The "Issue with segmentation" column encloses two issues: 1) samples where the clinicians found it hard to differentiate with the leaflets or recognize the view due to poor image quality, 2) the segmentation failed to properly distinguish between leaflets — for both ground truth and prediction. The results of this study show that the

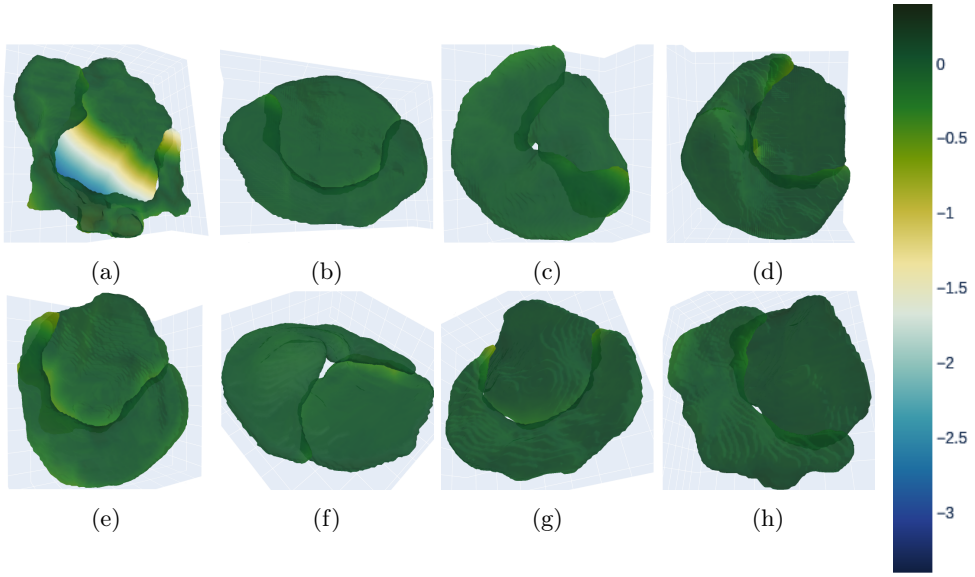


Fig. 5: Example of Signed Distance Function *in mm* between the ground-truth and the prediction results. Positive values indicate that the prediction is contained inside the ground truth. Examples on the top line are the same as the one shown in Figures 3 and 4

Table 3: Distance metrics in mm results for evaluation set at mid-systole frame (± std. dev.)

	Hausdorff distance	Average Surface Distance	Average Signed Distance
anterior	0.51 ± 0.46	0.10 ± 0.14	-0.07 ± 0.14
posterior	0.49 ± 0.41	0.09 ± 0.08	-0.06 ± 0.09

Multi-head ResUNet [13] architecture yielded results comparable to the clinical annotation using our semi-automated pipeline (cf Figure 1 and Section 2.2) as the difference in votes in favour of the ground truth and the predictions are small.

In addition, using the *same training* we also evaluated on all volumes with a closed valve, i.e., we segmented full systole sequences. The results displayed in Tables 6 and 7 show that given one training on just the mid-systole frame, our network, while being less accurate regarding the distance metrics, is still able to generalize to the rest of the systole sequence.

Table 4: Accuracy metrics in % results for evaluation set at mid-systole frame (± std. dev.)

	Accuracy	Precision	Recall	Dice
anterior	99.48 ± 0.20	72.04 ± 13.47	66.77 ± 10.50	79.60 ± 7.87
posterior	99.21 ± 0.41	64.13 ± 17.90	61.48 ± 16.58	74.53 ± 16.68

Table 5: Which segmentation was best according to clinicians when showing two slices side by side.

View	Ground truth	Prediction	Indistinguishable	Issue with segmentation
LAX	48	43	1	19
Perpendicular LAX	55	46		28
Total	**103**	**89**	**1**	**47**

4 Discussion

As detailed in Section 2.2, the automation of our annotation yielded a ground truth that is coherent though slightly imprecise, as shown in Figure 3. Nevertheless, it is a fast annotation process: on average, it takes no more than five minutes to annotate a recording, which contains several volumes — each file contains on average **6.84** annotated volumes — whereas a manual annotation done in 3D Slicer [13] takes fifteen minutes for a trained user and yields only one annotated volume.

Given our annotation, it seems obvious there is an ϵ error between ground truth and prediction that cannot be reduced. This is due to the annotation overestimating the leaflets (cf. Figure 3), and the network having trouble matching it due to the voxel intensity being bright only on the leaflets. As such it is easy to obtain a true positive result, but the network will naturally underestimate in comparison to the ground truth. This is why we focused more on the distance type of metrics found in Table 3 as well as the clinical study presented in Section 2.5 and Table 5.

Nevertheless, using a different ultrasound vendor, we achieved results comparable to previous research [13] in both mid-systole and systole sequence cases.

Concerning the extension of the mid-systole segmentation to the full systole sequence using the same training, we observe similar accuracies to the evaluation solely on the mid-systole frame (cf. Tables 4 and 7), but higher ASD (see Table 6). This is due to the lower quality of the frames' segmentation surrounding the mid-systole and the input volumes being slightly out of the training manifold. Because the position of the valve can be a bit open before or after the mid-systole, or appear more v-shaped. Some semi-supervised learning could be a lead to improve our results here.

Table 6: Distance metrics in mm results for evaluation set for the full systole sequence (± std. dev.)

	Hausdorff distance	Average Surface Distance	Average Signed Distance
anterior	0.64 ± 0.64	1.41 ± 0.89	-0.63 ± 0.64
posterior	0.75 ± 0.59	1.67 ± 0.87	-0.75 ± 0.59

Table 7: Accuracy metrics in % results for evaluation set for the full systole sequence (± std. dev.)

	Accuracy	Precision	Recall	Dice
anterior	99.48 ± 0.22	68.05 ± 15.67	65.88 ± 12.83	78.63 ± 10.60
posterior	99.26 ± 0.37	62.18 ± 13.64	62.89 ± 15.08	75.96 ± 13.89

As remarked by the clinicians during the qualitative assessment of results, there is still room for improvement for both the ground truth and the prediction to distinguish properly between the anterior and posterior leaflets. Refining the ground truth for part of the training dataset would help the network learn to discriminate better between the two leaflets.

5 Conclusion

We showed that given a less accurate but much faster annotation process, we can still obtain results of high quality, comparable to the previous state-of-the-art that are on pair with the clinical-made ground truth. We were also able to generalise to surrounding frames in the heart cycle — during systole here. This is advantageous as this annotation process requires significantly less time than a manual annotation.

Acknowledgments. We would like to acknowledge the involvement of Riccardo Munafò, Olivier Gérard and Emiliano Votta in the project's discussion, as well as the development contribution of Tollef Jørgensen and Sverre Herland.

Disclosure of Interests. We would like to acknowledge the financial support of the Norwegian University of Science and Technology's Department of Computer Science, as well as the Centre for Innovative Ultrasound Solutions. We would also like to thank General Electric HealthCare and the Department of Circulation and Medical Imaging of the Norwegian University of Science and Technology at St. Olavs Hospital for allowing us to use their data.

Patients who TEE are used in this study provided informed consent and the data was properly anonymized before being studied. Given consent did not include public sharing of the data.

As a contribution to open science, all the code that could be made public without conflict is available here and here. This includes the full training and evaluation

pipeline, as well as the process of creating a voxel grid from a DICOM and its corresponding surface meshes.

References

1. Aruta, P., Muraru, D., Guta, A.C., Mihaila, S., Ruozi, N., Palermo, C., Elnagar, B., Iliceto, S., Badano, L.P.: Comparison of mitral annulus geometry between patients with ischemic and non-ischemic functional mitral regurgitation: implications for transcatheter mitral valve implantation. Cardiovascular Ultrasound **16**(1), 27 (Oct 2018). https://doi.org/10.1186/s12947-018-0145-8, https://doi.org/10.1186/s12947-018-0145-8
2. Cardoso, M.J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., Nath, V., He, Y., Xu, Z., Hatamizadeh, A., Myronenko, A., Zhu, W., Liu, Y., Zheng, M., Tang, Y., Yang, I., Zephyr, M., Hashemian, B., Alle, S., Darestani, M.Z., Budd, C., Modat, M., Vercauteren, T., Wang, G., Li, Y., Hu, Y., Fu, Y., Gorman, B., Johnson, H., Genereaux, B., Erdal, B.S., Gupta, V., Diaz-Pinto, A., Dourson, A., Maier-Hein, L., Jaeger, P.F., Baumgartner, M., Kalpathy-Cramer, J., Flores, M., Kirby, J., Cooper, L.A.D., Roth, H.R., Xu, D., Bericat, D., Floca, R., Zhou, S.K., Shuaib, H., Farahani, K., Maier-Hein, K.H., Aylward, S., Dogra, P., Ourselin, S., Feng, A.: MONAI: An open-source framework for deep learning in healthcare (Nov 2022), http://arxiv.org/abs/2211.02701, arXiv:2211.02701 [cs]
3. Carnahan, P., Moore, J., Bainbridge, D., Eskandari, M., Chen, E.C.S., Peters, T.M.: DeepMitral: Fully Automatic 3D Echocardiography Segmentation for Patient Specific Mitral Valve Modelling. In: de Bruijne, M., Cattin, P.C., Cotin, S., Padoy, N., Speidel, S., Zheng, Y., Essert, C. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2021. pp. 459–468. Lecture Notes in Computer Science, Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-87240-3_44
4. Chen, J., Li, H., He, G., Yao, F., Lai, L., Yao, J., Xie, L.: Automatic 3D mitral valve leaflet segmentation and validation of quantitative measurement. Biomedical Signal Processing and Control **79**, 104166 (Jan 2023). https://doi.org/10.1016/j.bspc.2022.104166, https://www.sciencedirect.com/science/article/pii/S1746809422006206
5. European Heart Network, E.: 2022 Annual Report (Nov 2022), https://ehnheart.org/annual-reports.html
6. Faletti, R., Gatti, M., Salizzoni, S., Bergamasco, L., Bonamini, R., Garabello, D., Marra, W.G., La Torre, M., Morello, M., Veglia, S., Fonio, P., Rinaldi, M.: Cardiovascular magnetic resonance as a reliable alternative to cardiovascular computed tomography and transesophageal echocardiography for aortic annulus valve sizing. The International Journal of Cardiovascular Imaging **32**(8), 1255–1263 (Aug 2016). https://doi.org/10.1007/s10554-016-0899-8, https://doi.org/10.1007/s10554-016-0899-8
7. Hirasawa, K., Vo, N.M., Gegenava, T., Pio, S.M., van Wijngaarden, S.E., Ajmone Marsan, N., Bax, J.J., Delgado, V.: Mitral Valve Annulus Dimensions Assessment with Three-Dimensional Echocardiography Versus Computed Tomography: Implications for Transcatheter Interventions. Journal of Clinical Medicine **10**(4), 649 (Jan 2021). https://doi.org/10.3390/jcm10040649, https://www.mdpi.com/2077-0383/10/4/649, number: 4 Publisher: Multidisciplinary Digital Publishing Institute

8. Isensee, F., Jäger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: Automated Design of Deep Learning Methods for Biomedical Image Segmentation. Nature Methods **18**(2), 203–211 (Feb 2021). https://doi.org/10.1038/s41592-020-01008-z, http://arxiv.org/abs/1904.08128
9. Kerfoot, E., Clough, J., Oksuz, I., Lee, J., King, A.P., Schnabel, J.A.: Left-Ventricle Quantification Using Residual U-Net. In: Pop, M., Sermesant, M., Zhao, J., Li, S., McLeod, K., Young, A., Rhode, K., Mansi, T. (eds.) Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges. pp. 371–380. Lecture Notes in Computer Science, Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-12029-0_40
10. Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization (Jan 2017), http://arxiv.org/abs/1412.6980, arXiv:1412.6980 [cs]
11. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal Loss for Dense Object Detection (Feb 2018), http://arxiv.org/abs/1708.02002, arXiv:1708.02002 [cs]
12. Loshchilov, I., Hutter, F.: SGDR: Stochastic Gradient Descent with Warm Restarts (May 2017), http://arxiv.org/abs/1608.03983, arXiv:1608.03983 [cs, math]
13. Munafò, R., Saitta, S., Ingallina, G., Denti, P., Maisano, F., Agricola, E., Redaelli, A., Votta, E.: A Deep Learning-Based and Fully Automated Pipeline for Regurgitant Mitral Valve Anatomy Analysis from 3D Echocardiography (Feb 2023), http://arxiv.org/abs/2302.10634, arXiv:2302.10634 [cs, q-bio]
14. Pouch, A.M., Wang, H., Takabe, M., Jackson, B.M., Gorman, J.H., Gorman, R.C., Yushkevich, P.A., Sehgal, C.M.: Fully automatic segmentation of the mitral leaflets in 3D transesophageal echocardiographic images using multi-atlas joint label fusion and deformable medial modeling. Medical Image Analysis **18**(1), 118–129 (Jan 2014). https://doi.org/10.1016/j.media.2013.10.001, https://www.sciencedirect.com/science/article/pii/S1361841513001424
15. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation (May 2015). https://doi.org/10.48550/arXiv.1505.04597, http://arxiv.org/abs/1505.04597, arXiv:1505.04597 [cs]
16. Vahanian, A., Beyersdorf, F., Praz, F., Milojevic, M., Baldus, S., Bauersachs, J., Capodanno, D., Conradi, L., De Bonis, M., De Paulis, R., Delgado, V., Freemantle, N., Gilard, M., Haugaa, K.H., Jeppsson, A., Jüni, P., Pierard, L., Prendergast, B.D., Sádaba, J.R., Tribouilloy, C., Wojakowski, W., ESC/EACTS Scientific Document Group, ESC National Cardiac Societies: 2021 ESC/EACTS Guidelines for the management of valvular heart disease: Developed by the Task Force for the management of valvular heart disease of the European Society of Cardiology (ESC) and the European Association for Cardio-Thoracic Surgery (EACTS). European Heart Journal **43**(7), 561–632 (Feb 2022). https://doi.org/10.1093/eurheartj/ehab395, https://doi.org/10.1093/eurheartj/ehab395
17. Vo, N.M., van Wijngaarden, S.E., Marsan, N.A., Bax, J.J., Delgado, V.: Assessment of D-Shaped Annulus of Mitral Valve in Patients with Severe MR Using Semi-Automated 4-Dimensional Analysis: Implications for Transcatheter Interventions. Journal of Cardiovascular Development and Disease **7**(4), 48 (Dec 2020). https://doi.org/10.3390/jcdd7040048, https://www.mdpi.com/2308-3425/7/4/48, number: 4 Publisher: Multidisciplinary Digital Publishing Institute