

Article

Mitral Valve Segmentation Using Robust Nonnegative Matrix Factorization

Hannah Dröge ^{1,*}, Baichuan Yuan ², Rafael Llerena ³, Jesse T. Yen ⁴, Michael Moeller ¹ and Andrea L. Bertozzi ²

¹ Department of Electrical Engineering and Computer Science, University of Siegen, 57076 Siegen, Germany; michael.moeller@uni-siegen.de

² Department of Mathematics, University of California, Los Angeles, CA 90095, USA; ybcmath@gmail.com (B.Y.); bertozzi@math.ucla.edu (A.L.B.)

³ Non-Invasive Cardiology Department, Keck Medical Center of University of Southern California, Los Angeles, CA 90033, USA; rafael.llerena@med.usc.edu

⁴ Department of Biomedical Engineering, University of Southern California, Los Angeles, CA 90089, USA; jesse.yen@usc.edu

* Correspondence: hannah.droege@uni-siegen.de

Abstract: Analyzing and understanding the movement of the mitral valve is of vital importance in cardiology, as the treatment and prevention of several serious heart diseases depend on it. Unfortunately, large amounts of noise as well as a highly varying image quality make the automatic tracking and segmentation of the mitral valve in two-dimensional echocardiographic videos challenging. In this paper, we present a fully automatic and unsupervised method for segmentation of the mitral valve in two-dimensional echocardiographic videos, independently of the echocardiographic view. We propose a bias-free variant of the robust non-negative matrix factorization (RNMF) along with a window-based localization approach, that is able to identify the mitral valve in several challenging situations. We improve the average f1-score on our dataset of 10 echocardiographic videos by 0.18 to a f1-score of 0.56.

Keywords: mitral valve; segmentation; robust non-negative matrix factorization; echocardiography



Citation: Dröge, H.; Yuan, B.; Llerena, R.; Yen, J.T.; Moeller, M.; Bertozzi, A.L. Mitral Valve Segmentation Using Robust Nonnegative Matrix Factorization. *J. Imaging* **2021**, *7*, 213. <https://doi.org/10.3390/jimaging7100213>

Academic Editors: Fabiana Zama and Elena Loli Piccolomini

Received: 30 August 2021

Accepted: 8 October 2021

Published: 16 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The precise movement of the mitral valve is of crucial importance for a proper blood flow [1]. Therefore, accurate detection of mitral valve disease is of great importance for the treatment and prevention of several diseases. Especially in old age, a defect of a heart valve is a common disease. According to Mohty et al. [2], more than 12.5% of the elderly people of 75 years or over suffer from heart valve diseases. In a study in 2006, mitral regurgitation, which is a disease of the mitral valve that does not close tightly and thus partially allows the blood to flow back, was the most measured common kind of heart disease [3].

Very helpful for the recognition of the degree of mitral regurgitation and thus, the necessity of surgical intervention, is the diagnostic imaging procedure of echocardiography. Echocardiography uses ultrasound techniques to receive an image of the heart and has multiple qualities, such as the real-time displaying of heart motion as well as of the blood flow. It is the most common medical imaging technique despite other technologies, such as magnetic resonance and computer tomography [4,5] and benefits, compared to other diagnostic methods, from the portability of its required equipment, its easy implementation and the low price as well as the safety of this type of examination. Additionally, for the affected patient, the ultrasound examination is associated with minimal discomfort and stress [6,7].

The analysis of the echocardiographically recorded data is performed manually by trained specialists, which is slow and subjective and depends on the current stay of the specialist [8,9]. An automated, robust and repeatable method for mitral valve detection

would be beneficial in this regard, and the recording and analysis of cardiac ultrasound data by non-experts is of highest interest.

Many applications benefit from the successful automatic segmentation of the mitral valve of the heart in echocardiographic ultrasound videos. For example, in cardiac surgery, the automatic detection of the valve prior to or during surgery could be used as support or for robotic assistance [7]. Another field of application is diagnostics, where segmentation can lead to a better understanding of the behavior and geometry of the patient's mitral valve and thus of its various diseases [10]. Finally, the segmentation of the mitral valve can support the training of ultrasound diagnosticians [11].

Automatic and unsupervised segmentation of the mitral valve in echocardiographic ultrasound videos was first approached by Zhou et al. [10] by capturing the irregular motion of the mitral valve, using a matrix factorization method. This method was later taken up by Dukler et al. [9], who used robust non-negative matrix factorization (RNMF) for mitral valve segmentation, with the underlying basic concept of distinguishing the rigid motion of the myocardial muscle from the nonrigid and irregular motion performed by the heart valve and thus, being able to segment the latter. Consistent with this work, our method uses a bias-free version of RNMF to capture irregular motion of the mitral valve, followed by an additional variational segmentation technique to perform successful segmentation. Our paper's contributions are listed below:

1. We use Bregman iteration to avoid the inherent bias of ℓ^1 regularization in the RNMF model within the context of mitral valve segmentation.
2. As an additional variational segmentation technique, we use the Chan–Vese segmentation algorithm [12] to which we add a new regularization term so that it matches the problem of mitral valve segmentation.
3. We proposed a new segmentation refinement algorithm that takes into account the opening and closing motion of the heart valve and, in combination with the unbiased RNMF model and the regularized variational segmentation technique, allows us to perform fully automatic segmentation without any further knowledge about the heart valve.

Our paper is divided into six sections. Section 2 discusses related work dealing with general segmentation and previous semi-automatic and automatic mitral valve segmentation methods. Section 3 discusses our proposed mitral valve segmentation method and is divided into four subsections. The first three subsections cover the crucial steps for fully automatic valve segmentation, consisting of myocardial muscle detection, valve segmentation, and a refinement step. The fourth subsection deals with the optional step of additional myocardial valve detection, using a windowing method. In Section 4, we present the datasets we use to evaluate our method. Section 5 deals with the results of our experiments, which include the analysis of various hyperparameters, comparison with two related methods for cardiac valve segmentation, and discussion of error cases. In the conclusion in Section 6, we provide a brief summary of our results and present limitations for real-world applications as well as possible future novel approaches to improve mitral valve segmentation.

2. Related Work

Image segmentation is a commonly used technique with a wide range of applications, including medical imaging. The current state of the art in this field is largely dominated by deep learning approaches, which have also been successful in medical images; see [13]. However, the training of such neural networks usually requires large amounts of training data, which in medical applications, such as echocardiographic videos, are difficult to access in a sufficient amount together with their results.

Moreover, a very difficult to analyze and possibly hidden bias of a deep neural network toward the training data remains a risk in safety critical areas, such as medical image segmentation or disease classification.

While model-based (unsupervised) methods, such as the proposed approach, also contain a bias, this bias arises from the model itself (e.g., from an assumption of temporal or spatial coherence) such that the type of bias and uncertainty of the solution is known a priori.

So, instead of exploiting training data, classical techniques rely on cleverly designed methods, such as thresholding methods, clustering methods, region growing methods and more. An overview of some of these methods is given in [14,15]. More recent works consider variational methods (see, for example, [12,16]), or formulate similar energy minimization problems on suitable graphs, e.g., motivated by Markov random fields or conditional random fields; see [17–19] for examples.

The image formation of echocardiographic images differs from photographic images and thus, poses a special challenge in the field of segmentation. Cardiac ultrasound images are generated by measuring the time and amplitude of ultrasound waves sent in the direction of the heart and reflected back from its structures, whereby the intensities of the reflected echoes within an ultrasound image are converted into gray tones. When recording ultrasound videos, the frame rate, or temporal resolution, is determined by the desired field of view, line spacing, and speed of sound in the medium [20]. So, opposed to the segmentation of photographic images, the automatic segmentation of the mitral valve poses various additional challenges, as the mitral valve does not differ from the rest of the heart muscle in structure or intensity, and ultrasound videos often contain strong noise. In addition, the mitral valve does not have a smooth structure or hard edges in its ultrasound recordings; it may occur that in several frames, the mitral valve is hardly visible or not visible at all. Its form and structure also differ according to the considered echocardiographic view.

Such challenges are one of the reasons why many recent techniques have focused on semi-automatic methods that still rely on certain user inputs. Algorithmically, many of these methods use the active contours model of Kass et al. [16], for example, in [8,21], or variants thereof; see [22,23]. Other approaches, e.g., that of Shang et al. [24], exploit level set segmentation methods based on the intensity difference between the mitral valve and the background. The advantage of those semi-automatic methods over manual methods is that often, only one or few user inputs are required, such as drawing the initial contours around the mitral valve, and are, therefore, less time consuming than a manual segmentation of the mitral valve by specialists. However, specialists are still required to draw the initial contour. Important factors for the clinical application of those methods are their reliability and time consumption. According to [21], the method of Kass et al. [16] needs 20 min per ultrasound video, so the method can only be used in applications where no real-time application is required. Due to the speed of their real-time method, Martin et al. [21] suggests them for preoperative and intraoperative scenarios. In general, however, semi-automated and also automated heart valve recognition procedures should be critically reviewed for failure. In particular, [22,23] show that the simultaneous speed and reliability in a method is a great challenge. In [23], a more reliable method is proposed than in [22], but at the expense of speed.

The first automatic, unsupervised method for tracking and segmenting the mitral valve was proposed in 2012 by Zhou et al. [10] and takes advantage of the fast motion of the mitral valve and thus the spectral distribution of the ultrasound video that is strongest in the valve area. Inspired by this, a more robust method, dealing with the low-rank approximation of echocardiographic videos, was proposed by Liu et al. [7], called constrained outlier pursuit, to prevent possibly occurring drifts in tracking the mitral valve. In addition, they use region scalable active contours for segmenting the mitral valve. Based on the method in [10,25], Dukler et al. [9] developed a new method to detect the mitral valve by applying RNMF [26] on multiple areas of the echocardiographic video and measuring the residue at each position.

RNMF is a regularized version of nonnegative matrix factorization (NMF) on which, in 1999, Lee and Sung [27] conducted a study that attracted wide interest in several ap-

plication areas, including face recognition [28], medical applications [29,30] and general clustering approaches [31]. Various minimization algorithms were presented in the following years to optimize NMF, such as multiplicative updating rules [32] or hierarchical alternating least squares [33] for faster convergence. For an overview of the optimizations, we refer here to [34,35]. To improve the decomposition results of NMF, among others, orthogonality condition [36] or sparsity conditions [26,37,38] were added to NMF; in this paper, we consider the sparsity constraint of [26], known as RNMF, which is a special case of a method used by Corinzia et al. [39], who used neural network matrix factorization combined with a threshold network, which is trained in an unsupervised manner, on each echocardiographic video individually to capture the motion of the mitral valve. In order to define the region of the heart valve, they additionally localized the heart valve region, using optical flow calculated on the outcome of the neural network. In a subsequent paper on this topic, Corinzia et al. [40] extend this work and propose a generalised form of matrix factorisation together with a parameterised threshold operator. To find the region of the mitral valve, they use optical flow and the low-dimensional time series representation of the echocardiographic video. Also, they present post-processing algorithms for improvement of the segmentation.

In line with these prior works that successfully exploited RNMF to detect the position of the mitral valve in the echocardiographic video, we extend [9,40] in the following two ways: (1) We avoid the inherent bias of ℓ^1 regularization in the RNMF model by exploiting Bregman iterations [41], and (2) we utilize an additional variational segmentation technique (a modified version of the Chan–Vese algorithm [12]) on the result of the RNMF instead of relying on plain thresholding. We demonstrate in several numerical experiments that the proposed modifications result in higher quality results specifically on challenging examples, where the mitral valve is moving quickly and thus is difficult to detect. In contrast to [9,40], the segmentation technique allows us to run our algorithm, even without prior windowing techniques (which rely on the manually specified window size that contains the mitral valve) and therefore, represents advances toward fully-automatic mitral valve segmentation.

Table 1 shows a summary of our and previous mitral valve segmentation and tracking methods in echocardiographic recordings, sorts them into subgroups, and provides a listing of their respective advantages and disadvantages.

Table 1. Summary of mitral valve segmentation methods for echocardiographic recordings.

Method	Category	Prior Knowledge	Pros and Cons	Paper
Energy Minimization Method	Active Contours	Initial Contour	Reduces the effort of the manual segmentation of each frame to the segmentation of only one frame. However, the disadvantage is that there has to be prior knowledge of the position of the mitral valve and it has to be drawn in by an expert. In addition, a tuning of parameters is necessary.	[8,21–24]
			While this method has the advantage that no prior knowledge of mitral valve position is required, it has the disadvantage of requiring an extensive parameter adjustment.	[7]
	Matrix Factorization	Mitral Valve Size	These methods do not require a prior drawn contour of the mitral valve in the first frame, but the size of the mitral valve for a windowing method must be known in advance. In addition, an adjustment of parameters is also necessary here.	[9,10]
Matrix Factorization with Bias Avoiding		Optional: Mitral Valve Size	An advantage of this method is that the size of the mitral valve does not need to be known (optional), but the disadvantage is the required adjustment of parameters.	ours

Table 1. Cont.

Method	Category	Prior Knowledge	Pros and Cons	Paper
Machine Learning Method	Unsupervised, Videowise Training	Mitral Valve Size	These methods do not require a prior drawn contour of the mitral valve in the first frame, but the size of the mitral valve for a windowing method must be known in advance. In addition, an adjustment of parameters is also necessary here.	[39,40]
	Supervised	Training Data	An automatic segmentation method without parameter adjustment for each video. However, a drawback is that training data and annotations must be available, which is especially difficult for medical data. In addition, there is an unknown bias shift of segmentation toward the training data.	[13]

Related to mitral valve segmentation on echocardiographic images is the segmentation of the left ventricle, that is adjacent to the valve. In this area, there have been various approaches in recent years, including energy minimization methods as the method of Pedrosa et al. [42], who combine a shape-based deformable model with a statistical shape model, allowing the joint use of image information as well as shape-based cues for segmentation. A more general approach, which can be applied to ventricle segmentation, was presented by Ciecholewski in 2016 [43], presenting an edge-based active contour model that applies an inflation/deflation force for each node of an active contour, allowing it to shift active contour nodes and locate object boundaries, even in a noisy image while still approximating the rigor of the object shape. In addition, approaches with deep learning elements were studied, such as the method of Carneiro et al. [44], addressing the segmentation of the left ventricle by rigid and non-rigid decoupled segmentation using artificial neural networks, and Ali et al. [45] presenting a deep learning framework that combines ResNet and U-Net elements. Especially useful for comparison is the evaluation framework of Bernard et al. [46] for the comparison and evaluation of left ventricle segmentation algorithms.

3. Method

The discrete representation of an echocardiographic video with t frames is a three-dimensional tensor $X_{3D} \in \mathbb{R}_+^{s_x \times s_y \times t}$. By vectorizing the two spatial dimensions s_x and s_y , the video X_{3D} can be represented by a matrix $X \in \mathbb{R}_+^{s \times t}$ with $s = s_x s_y$, i.e., with each column of X representing a frame of the echocardiographic video. Our approach for segmenting the mitral valve in an echocardiographic video is to first apply RNMF to the matrix X to separate the structure of the mitral valve from the structure of the heart muscle. It follows a segmentation and a refinement step, as shown in Figure 1, to enclose the mitral valve structure at its edges. Windowing is an optional step that defines the area where the mitral valve can occur, thus increasing the accuracy of segmentation. However, this (optional) step requires knowledge of the approximate size of the mitral valve. In the following subsections, we will describe each of the steps illustrated in Figure 1 separately.

3.1. Muscle Detection

Similar to the classical RNMF, we model the video X to be a sum of the three following components:

$$X = WH + S + N, \quad (1)$$

where WH represents a low-rank part that captures the static background as well as simple rigid movements, S represents the non-rigidly moving parts (i.e., mainly the mitral valve), and N is the observation noise. The low-rank constraint is enforced by WH being a product of two matrices $W \in \mathbb{R}^{s \times r}$ and $H \in \mathbb{R}^{r \times t}$, where r is significantly smaller than s and t . Moreover, assuming S to occupy a comparably small spacial area but contain large values gives rise to penalizing the ℓ^1 norm of S (in the sense of the sum over the absolute values of S), and assuming N to be roughly Gaussian distributed gives rise to penalizing the squared ℓ^2 norm of N . Combining these modeling assumptions with (physically meaningful) non-

negativity constraints on all entries of W , H , and S gives rise to the usual RNMF model as follows:

$$\arg \min_{W, H, S \geq 0, N} \|N\|_2^2 + \lambda \|S\|_1, \quad \text{s.t. (1) holds,} \quad (2)$$

for a regularization parameter λ to enforce the sparsity of S . The more common way to write RNMF follows by using the constraint to replace N by $X - WH - S$ (which we do below).

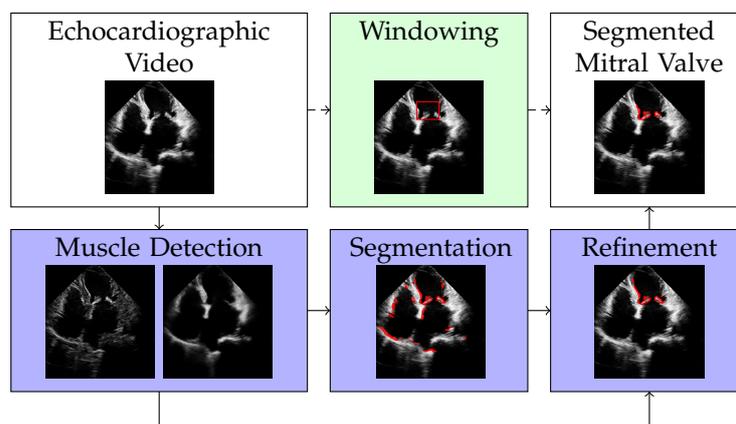


Figure 1. Overview of the proposed method. The white filling of the box shows the input and output of our method. The dark purple fill shows the mandatory steps, while the light green fill indicates the optional step of our algorithm.

During our initial experiments with Equation (2), we observed that the natural shrinking bias of the ℓ^1 norm (see [47]) significantly harms the performance of the RNMF model, even for small values of λ . On the other hand, $\lambda = 0$, leads to highly non-unique minimizers of Equation (2) with the undesired trivial solution $S = X, WH = 0$ being one of them.

Therefore, we exploit the idea of Bregman iterations [41] to iteratively replace the (constrained) regularization term $\delta_{\geq 0}(S) + \|S\|_1$ by $\delta_{\geq 0}(S) + \|S\|_1 - \langle p_S^k, S \rangle$ for a subgradient $p_S^k \in \partial(\delta_{\geq 0} + \|\cdot\|_1)(S^k)$, where $\delta_{\geq 0}$ denotes the characteristic function of the non-negativity constraint, S^k refers to a previous estimate of the mitral valve S , and the operation $\langle \cdot, \cdot \rangle$ denotes the dot product. More precisely, we alternate between updating WH, S and the subgradient p_S via the following:

$$(W^{k+1}, H^{k+1}) = \arg \min_{W, H \geq 0} \|X - WH - S^k\|_2^2 \quad (3)$$

$$S^{k+1} = \arg \min_{S \geq 0} \|X - W^{k+1}H^{k+1} - S\|_2^2 + \lambda (\|S\|_1 - \langle p_S^k, S \rangle). \quad (4)$$

$$p_S^{k+1} = p_S^k + \frac{1}{\lambda} (X - W^{k+1}H^{k+1} - S^{k+1}). \quad (5)$$

Note that if some component $(S^k)_{i,j}$ is positive, it follows that $(p_S^k)_{i,j} = 1$ such that the next regularization term $\|S\|_1 - \langle p_S^k, S \rangle$ no longer penalizes positive values of $(S^k)_{i,j}$ at all and therefore naturally removes the shrinking bias. We refer to [47] for details.

The alternations in Equations (3)–(5) are realized by the update steps shown in Algorithm 1, with σ being the thresholded pointwise proximal operator of the ℓ^1 norm:

$$\sigma(u)_\lambda = \begin{cases} u - \lambda & \text{if } u > \lambda \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Algorithm 1 Update step of W, H, S and p for muscle detection.

- 1: $S_{i,j} \leftarrow \sigma((X - W \cdot H + \lambda p)_{i,j})_\lambda$
 - 2: $p_{i,j} \leftarrow p_{i,j} + \frac{1}{\lambda}((X - W \cdot H)_{i,j} - S_{i,j})$
 - 3: $W_{i,j} \leftarrow W_{i,j} \left(\frac{|((S-X)H^T)_{i,j}| - ((S-X)H^T)_{i,j}}{(2WHH^T)_{i,j}} \right)$
 - 4: $H_{i,j} \leftarrow H_{i,j} \left(\frac{|(W^T(S-X))_{i,j}| - (W^T(S-X))_{i,j}}{2(W^TWH)_{i,j}} \right)$
-

The optimization steps are followed by a normalization step to acquire a unique solution. For more details, we refer to [26].

Opposed to [41], where Bregman iterations were considered in the context of convex energy minimization problems, the RNMF objective is non-convex such that one cannot expect to compute global minimizers (also not in update Equation (3)), which is why the initialization is important. We initialize $S = 0$ such that the first update of W and H solves a usual non-negative matrix factorization problem for which we use the multiplicative update algorithm [32].

We iterate Equations (3)–(5) for 100 iterations, with subproblem Equation (3) being approximated by a single step of the multiplicative algorithm. This way, we exploit early stopping before the method converges to a ℓ^1 -minimizing solution. We refer the reader to [41] (Theorem 3.5) for more properties of stopping the Bregman iteration early in the convex case and to [48] for more details on the analysis of Bregman iterations in the non-convex case.

3.2. Valve Segmentation

This section describes our method to compute the segmentation of the mitral valve as a binary mask $\tilde{B} \in \{0, 1\}^{s \times t}$ that is supposed to cover the mitral valve in X , i.e., the individual pixel values should behave as follows:

$$\tilde{B}_{i,j} = \begin{cases} 1 & \text{if the mitral valve appears in } X_{i,j} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Our segmentation model:

$$\hat{B} = \arg \min_{B, B_{ij} \in [0,1]} \frac{1}{2} \langle \Theta - \hat{S}, B \rangle + \lambda_1 TV(B) + \lambda_2 \langle \mathbb{1}, B \rangle + \lambda_3 \langle \hat{W}\hat{H}, B \rangle, \tag{8}$$

is based on the segmentation method by Chan and Vese [12] and contains modifications to adapt the problem formulation to our problem of mitral valve segmentation. Instead of minimizing over \tilde{B} , we use convex relaxation and minimize over all $B \in [0, 1]^{s \times t}$ to convexify our problem.

Here, \hat{W} , \hat{H} and \hat{S} are the results of the method in Section 3.1 and are kept fixed. The parameter $\Theta \in \mathbb{R}_{>0}$ is a threshold. Since the pixel values B_{ij} of B are enforced to grow, if $\hat{S}_{ij} > \Theta$, a small value of $\Theta = 0.01$ ensures that barely recognizable noise in \hat{S} does not cause the corresponding pixel values in B to increase. As a result, this noise is not segmented in B . The first regularization term applies (isotropic) Total Variation (TV) [49] on B for a regularization parameter λ_1 with the effect of penalizing the edge length of the segmented area in B . In the discrete setting, isotropic Total Variation (TV) is defined as $TV(B) = \sum_{i=1}^s \sqrt{((D_x b)_i)^2 + ((D_y b)_i)^2}$ (see [50]), where D_x and D_y are matrices that calculate the gradient on b by discretized forward differences in the spatial x - and y directions, and b the vectorized form of B .

The number of large pixels in B is reduced by the second regularization term, which is weighted with $\lambda_2 > 0$ with the intention to remove noise from B . The third regularization term $\langle \hat{W}\hat{H}, B \rangle$ penalizes large values in B , if there is correspondingly a large pixel value in $(\hat{W}\hat{H})$ at the same location. Since $(\hat{W}\hat{H})$ contains the rigid heart muscle, this means that an occurrence of muscle structure forces the values in B to be small since they are not supposed to be segmented. A reformulation of (8) is given in (9) and leads to another interpretation.

$$\hat{B} = \arg \min_{B, B_{ij} \in [0,1]} \frac{1}{2} \langle \Theta_2 - \hat{S}, B \rangle + \lambda_1 TV(B) + \lambda_2 \langle \mathbb{1}, B \rangle, \tag{9}$$

with $\Theta_2 = \Theta + 2\lambda_3 \hat{W}\hat{H}$

The threshold values of $\Theta_2 \in \mathbb{R}_{>0}^{s \times t}$ now depend on the entries in $(\hat{W}\hat{H})$, as large values of $(\hat{W}\hat{H})$ lead to large entries in the threshold Θ_2 . Thus, the values in B at positions with large values in Θ_2 only increase if accordingly large pixel values are present at the respective positions in S . Therefore, in this formulation, the threshold provides the information at which point the mitral valve occurs with which likelihood, while the third regularization term from Equation (8) is omitted.

The pixel values in B are typically close to 0 and 1 after minimizing the energy in (8) or (9). So, to finally obtain a binary segmentation mask \tilde{B} , we binarize \hat{B} by thresholding its pixel values at 0.5. To solve (9), we calculate the update steps via the primal–dual hybrid gradient (PDHG) algorithm by Chambolle and Pock [51] by solving the following:

$$\min_{b, b_i \in [0,1]} \max_p \langle Db, p \rangle + G(b) - F^*(p), \tag{10}$$

the primal dual form of (9), with the primal variable b , the dual variable p and the following:

$$G(b) = \frac{1}{2} \langle \Theta_{2,vec} - s, b \rangle + \lambda_2 \langle \mathbb{1}, b \rangle, \tag{11}$$

$$F(Db) = \lambda_1 TV(b).$$

The lowercase variables $s \in \mathbb{R}^{st}$ and $b \in [0, 1]^{st}$ represent the vectorized form of their corresponding uppercase matrices as well as $\Theta_{2,vec} = vec(\Theta_2)$. F^* is the convex conjugate form of F , and D represents the discrete derivative matrix. The update steps of (9) are realized as shown in Algorithm 2. Here $\iota_{\|\cdot\|_{2,\infty} \leq 1}$ is the convex conjugate of the $\ell^{2,1}$ -norm:

Algorithm 2 Update step of B for segmentation.

- 1: $p^{k+1} \leftarrow \iota_{\|\cdot\|_{2,\infty} \leq \lambda_1} (p^k + \sigma D \bar{b}^k)$
 - 2: $b^{k+1} \leftarrow b^k - \tau (D^T p^{k+1} + \lambda_1 (t - S) + \lambda_3 WH + \lambda_2)$
 - 3: $\bar{b}^{k+1} \leftarrow 2b^{k+1} - b^k$
-

3.3. Refinement

Due to a strong movement of the heart muscle, it is possible that certain structures that do not belong to the mitral valve are also segmented. In this phase, we refine the segmentation result \tilde{B} in two steps to extract the part of the segmentation mask that only covers the mitral valve.

3.3.1. Calculation of the Centroid

To detect the approximate location of the mitral valve, we consider the spatial centroid of the segmentation video with additional iterative steps alleviating the existence of outliers as shown in Algorithm 3. In the following, $\hat{S}_{3D} \in \mathbb{R}^{s_x \times s_y \times t}$ is the result of the previous method, introduced in Section 3.1.

At first, we combine the frames of video \hat{S}_{3D} and obtain a 2D map $\tilde{S}^1 \in \mathbb{R}_+^{s_x \times s_y}$ with high values in the places where bright structures in \hat{S}_{3D} appear most frequently (see Figure 2, left). In the first iteration of our algorithm, we calculate the centroid $C^1 = (C_x^1, C_y^1)$ of \tilde{S}^1 and consider it the position where the mitral valve is most likely to be located. To reduce the influence of the outlier structures on the calculation of the mitral valve location, we multiply \tilde{S}^1 pointwise with a 2D Gaussian function $g(\sigma, C_x, C_y)$ with

$$g_{x,y}(\sigma, C_x, C_y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x - C_x)^2}{(2\sigma^2)} - \frac{(y - C_y)^2}{(2\sigma^2)}\right),$$

which attenuates structures that are further away from the centroid. Here, $x \in \{1, \dots, s_x\}$ and $y \in \{1, \dots, s_y\}$. When the centroid of S^2 is subsequently recalculated, its position is less influenced by outlier structures.

By iterative continuation of these operations, with decreasing variance σ of the Gaussian function in each iteration, the centroid shifts further toward the mitral valve, as shown on the right in Figure 2. Here, each circle represents the position of the centroid in one iteration. Additionally, the right image in Figure 2 shows the result after eliminating the outlier structures after 11 iterations.

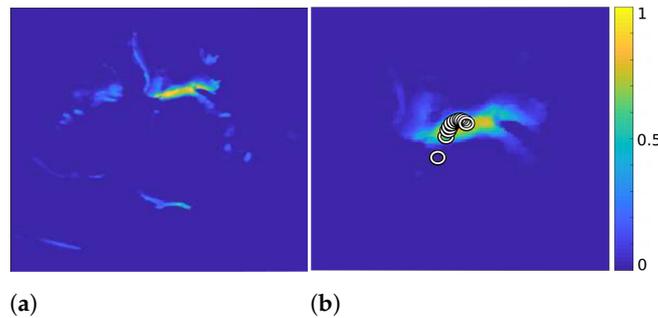


Figure 2. \tilde{S}^1 of Algorithm 3 (a) and \tilde{S}^{11} with the calculated centroids displayed by circles, after 11 iterations ((b), zoomed in).

Algorithm 3 Iterative calculation of the mitral valve position.

- 1: **for** $i = 1 : 11$ **do**
 - 2: $v = \sum_{x,y} \tilde{S}_{xy}^i$
 - 3: $C_x^i \leftarrow \frac{1}{v} \sum_{x,y} \tilde{S}_{xy}^i \cdot x$
 - 4: $C_y^i \leftarrow \frac{1}{v} \sum_{x,y} \tilde{S}_{xy}^i \cdot y$
 - 5: $\tilde{S}^{i+1} \leftarrow \tilde{S}^i \circ g(\sigma, C_x^i, C_y^i)$
 - 6: **end for**
 - 7: $C_{fin} = C^{11}$
-

3.3.2. Clustering

The idea is that with the information obtained from the last step on the approximate position of the mitral valve, only those segmentations should be retained that are considered likely to cover the mitral valve. For this purpose, the segmentation in \tilde{B} is interpreted as a set of connected components, in the following referred to as clusters $\{U_i\}$ with each $U_i = \{u_{1^i}, \dots, u_{n_i^i}\}$ being a set of connected pixels, where pixels of the binary video \tilde{B} that are equal 1 and connected in temporal and spatial dimensions form a cluster as shown in Figure 3 exemplarily in three successive frames and in Figure 4 in a three-dimensional visualization of the detected mitral valve.

In a purely spatial clustering, the alternately closing and opening valve leaflets do not touch each other and thus form two clusters in the open state, whereas in the closed state, only one cluster is formed due to the contacting valve leaflets. Because of this, it is particularly important to form the clusters not only in the spatial dimension, but also in

the temporal dimension to ensure that only a single cluster covers the mitral valve. To find the cluster $U_{\hat{i}}$ which is closest to the centroid C_{fin} , we measure the smallest distance of each cluster U_i to C_{fin} by first finding in each cluster the pixel $u_{\hat{p}^i}^i$ that is closest to the centroid with

$$\hat{p}^i = \arg \min_p \|u_p^i - C_{fin}\|^2, \tag{12}$$

measuring the distance and selecting the cluster $U_{\hat{i}}$ with the smallest distance with the following:

$$\hat{i} = \arg \min_i \|u_{\hat{p}^i}^i - C_{fin}\|^2. \tag{13}$$

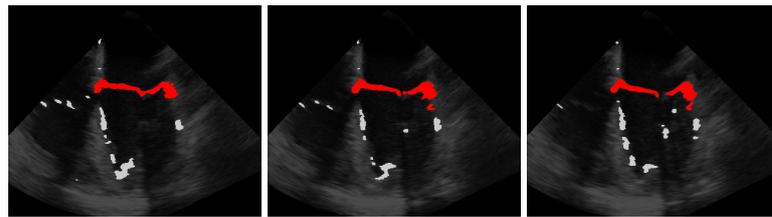


Figure 3. Three frames showing the extraction of the mitral valve (red) from the mis-segmented structures (light gray) by the refinement step.

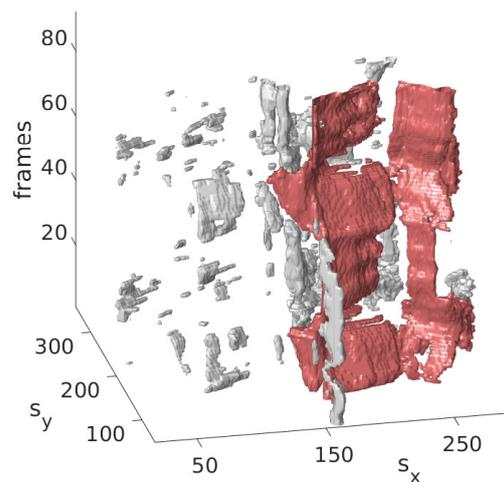


Figure 4. Three-dimensional view of the mitral valve (red), detected by the refinement step.

3.4. Windowing

A given mitral valve size can be used to restrict the segmentation of the mitral valve to a localized area of interest. To detect this area of interest, we first solve our model as follows:

$$\arg \min_{W,H,S \geq 0} \|X - WH - S\|_2^2 + \mu_1 \|S\|_1 + \mu_2 \langle WH, S \rangle \tag{14}$$

on the input echocardiographic video X ; Equation (14) is similar to Equation (2), but with an additional regularization to enforce exclusionary entries in (WH) and S for a regularization parameter $\mu_2 > 0$. A large value in (WH) is accompanied by a small value in S and vice versa. Then, we move a window $\mathcal{W} \in \mathbb{R}^{s_x^{\mathcal{W}} \times s_y^{\mathcal{W}} \times t}$ with $s_x^{\mathcal{W}} < s_x$ and $s_y^{\mathcal{W}} < s_y$ along the spatial dimensions of the minimizer $\hat{S} \in \mathbb{R}^{s_x \times s_y \times t}$ of Equation (14). For every window position, we calculate the squared Frobenius norm $c := \|\hat{S}\|_F^2$ of the values in \hat{S} that are covered by the window. The position of the window with the maximal resulting c is most likely to cover the mitral valve. Unlike [9], we do not solve Equation (14) on every window position, but only once on the whole video. This is computationally less expensive

and in our case, provides more precise window positions covering the mitral valve. To solve the minimization problem Equation (14), we alternate between updating W , H , S and the subgradient p_s in an iterative way. The update steps of W , H and S are similar to Equations (3) and (4), but this time, under the consideration of the new regularization term:

$$(W^{k+1}, H^{k+1}) = \arg \min_{W, H \geq 0} \|X - WH - S^k\|_2^2 + \mu_2 \langle WH, S^k \rangle, \tag{15}$$

$$S^{k+1} = \arg \min_{S \geq 0} \|X - W^{k+1}H^{k+1} - S\|_2^2 + \mu_1 (\|S\|_1 - \langle p_s^k, S \rangle) + \mu_2 \langle W^{k+1}H^{k+1}, S \rangle. \tag{16}$$

Similar to Algorithm 1, the updates are realized as shown in Algorithm 4, σ being the proximal operator of the ℓ^1 norm with the threshold (see Equation (6)):

Algorithm 4 Update step of W , H , S and p for windowing.

- 1: $S_{i,j} \leftarrow \sigma((X - WH + \mu_1 p - \mu_2 WH)_{i,j})_{\mu_1}$
 - 2: $p_{i,j} \leftarrow p_{i,j} + \frac{1}{\mu_1} ((X - WH)_{i,j} - S_{i,j} - \mu_2 (WH)_{i,j})$
 - 3: $W_{i,j} \leftarrow W_{i,j} \left(\frac{|((S - X + \frac{1}{2} \mu_2 S)H^T)_{i,j}| - ((S - X + \frac{1}{2} \mu_2 S)H^T)_{i,j}}{(2WHH^T)_{i,j}} \right)$
 - 4: $H_{i,j} \leftarrow H_{i,j} \left(\frac{|(W^T(S - X + \frac{1}{2} \mu_2 S))_{i,j}| - (W^T(S - X + \frac{1}{2} \mu_2 S))_{i,j}}{2(W^TWH)_{i,j}} \right)$
-

4. Dataset

The echocardiographic videos used for the evaluation were obtained from the Keck Medical Center of the University of Southern California. This study was approved by the institutional review board at the University of Southern California (Protocol ID: HS-15-00258). The dataset contains 10 different recordings of the heart in multiple views as shown in Figure 5. The analysis of the mitral valve segmentation was performed from these views, and the sample included the most common types of mitral valve pathology in addition to normal valve anatomy. Three cases show an abnormality of the mitral valve, specifically of the posterior mitral valve leaflet called flail mitral valve. This is a variation of a mitral valve prolapse category. The other videos do not show clear evidence of mitral valve pathology. The ultrasound videos differ also in length (from 17 to 147 frames) and in quality, with side lengths from 304 to 451 pixels, and have a frame rate of 30 frames per second. The mitral valve appears differently pronounced in the videos, which causes different results also regarding the detection as well as the segmentation. For each video frame, the ground truth of the segmentation consists of a contour that encloses the edges of the mitral valve and was manually created by non-medical experts and approved by an echocardiographer expert.

In addition, we evaluate our method on 46 videos from the EchoNet Dynamic dataset (see [52]), selected in [40]. The videos have a resolution of 112×112 pixels, which is lower than the resolution of our previously introduced dataset. For evaluation, we use the ground truth labeling provided by [40].

For the results on this dataset, we refer to Section 5.5.

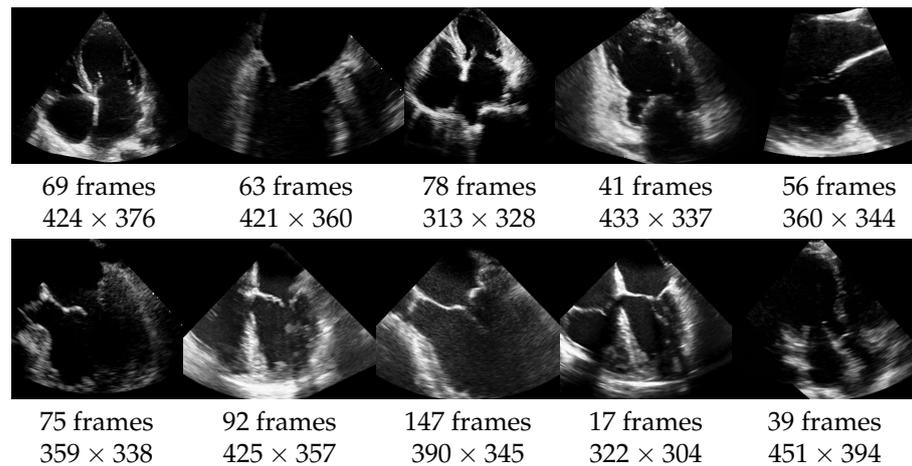


Figure 5. Dataset of 10 echocardiographic videos, with their number of frames and their spatial resolution.

5. Results

In the following, we compare our method with [9,40] according to the f1-score. We distinguish between the automatic method, in which no windowing is applied, and the method in which additionally a window slides over the video. Furthermore, we consider the recall regarding the windowing method, which is the number of ground truth pixels that are covered by the window with respect to the whole number of ground truth pixels. This gives us a measure of how much of the mitral valve is covered by the window. The hyperparameters used for the experiments are given in Table 2 for the muscle detection and valve segmentation algorithms and in Table 3 for the window detection algorithm. The hyperparameters for [40] are set as indicated in [40]. In agreement with the authors, we use for the experiments with [9] rank 2, a sparseness coefficient of 0.2, and for the threshold function, the top 1% pixels in terms of intensity.

Table 2. Hyperparameters of the proposed methods automatic segmentation (a), segmentation with windowing (b) for muscle detection and valve segmentation algorithms.

	λ	λ_1	λ_2	λ_3	Rank(WH)
(a)	0.1	0.04	0.075	1.0	2
(b)	0.1	0.04	0.05	1.0	2

Table 3. Hyperparameters of the Windowing method.

μ_1	μ_2	Rank(WH)
1	0.4	5

5.1. On Bregman Iteration and RNMF

To find the heart muscle structures, we use RNMF with Bregman iteration (see Section 3.1). For the subsequent segmentation, it is of great importance that no heart valve structures are visible in $(\hat{W}\hat{H})$ because structures that appear in $(\hat{W}\hat{H})$ are penalized, so they do not appear in the segmentation result. In the following, λ_{rnmf} refers to the hyperparameter λ of RNMF without Bregman iteration. As shown in Figure 6, increasing λ_{rnmf} pushes parts of the mitral valve into the product $(\hat{W}\hat{H})$, which our segmentation algorithm assumes not to contain the mitral valve. In addition, we observe in videos with a low contrast mitral valve that the mitral valve is partly removed from the result \hat{S} with increasing λ_{rnmf} . The f1-score for *automatic segmentation* decreases when increasing λ_{rnmf} as shown in Figure 7: with Bregman iteration, the f1-score is 0.45; for $\lambda_{rnmf} = 0.01$, the f1-score is 0.314; and for $\lambda_{rnmf} = 0.1$, the f1-score is 0.066.

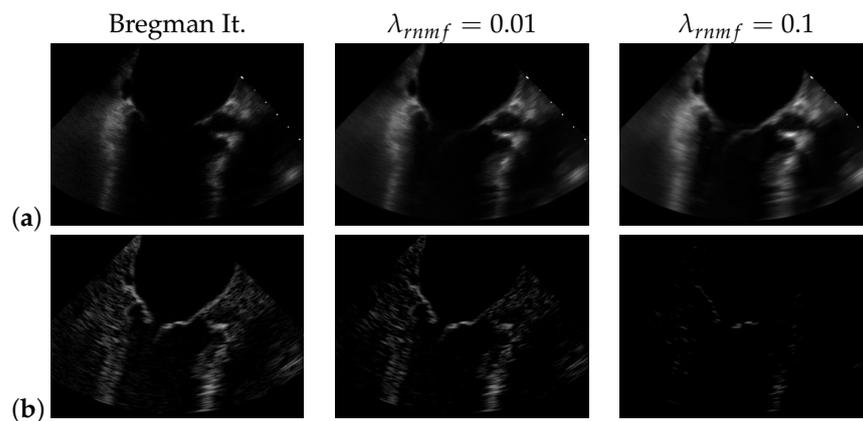


Figure 6. $(\hat{W}\hat{H})$ (a) and \hat{S} (b) in case of using Bregman iteration and of standard robust non-negative matrix factorization (RNMF) for different values of the regularization parameter λ_{rnmf} .

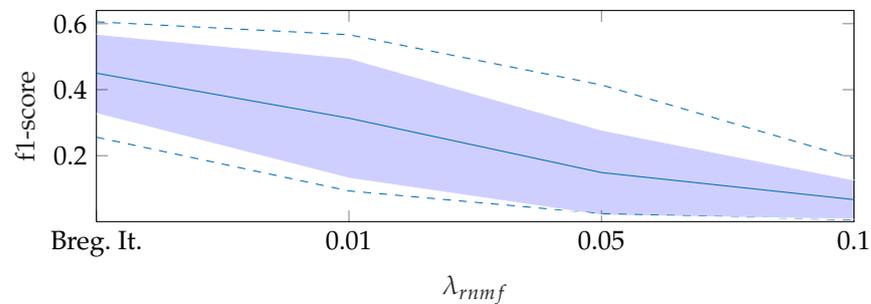


Figure 7. F1-score for automatic segmentation depending on λ_{rnmf} , where the dashed lines show the minimum and maximum f1-score values, and the standard deviation is indicated by the filled range.

5.2. Automatic Segmentation

In our mitral valve segmentation method that does not include the windowing step (see Section 3.4), we first split the muscle and the mitral valve structure as described in Section 3.1 by solving Equation (2), followed by the segmentation step of Section 3.2. Here, a strong weighting of the third regularization term in Equation (8) has a positive effect on the segmentation of the valve, as structures that are not located in the area of the muscle structure are preferably segmented. This is shown in Figure 8, by showing the behavior of the segmentation depending on λ_2 and λ_3 , whereby a large value in λ_3 provides better capturing of the mitral valve than an increase of λ_2 .

As a third step, we remove the segmented noise with the refinement that is described in Section 3.3. The resulting f1-scores compared to the ground truth are visualized in Figure 9 in the form of boxplot diagrams for segmentation, using the *windowing* approach in (a), for the method of Dukler et al. [9] in (c) and for the method of Corinzia et al. [40] in (d). The f1-score for *automatic segmentation* is shown in (b) and shows a slightly lower median f1-score than the one in (a), but without conspicuous negative outliers regarding the segmentation accuracy. Outliers indicate a complete missing of the mitral valve, which is not the case here. In [40], there are shown outliers, whereas [9] has a slightly lower f1-score with a larger variation. In addition to the boxplot diagrams, the average recall, precision, and f1-score values are given in Table 4.

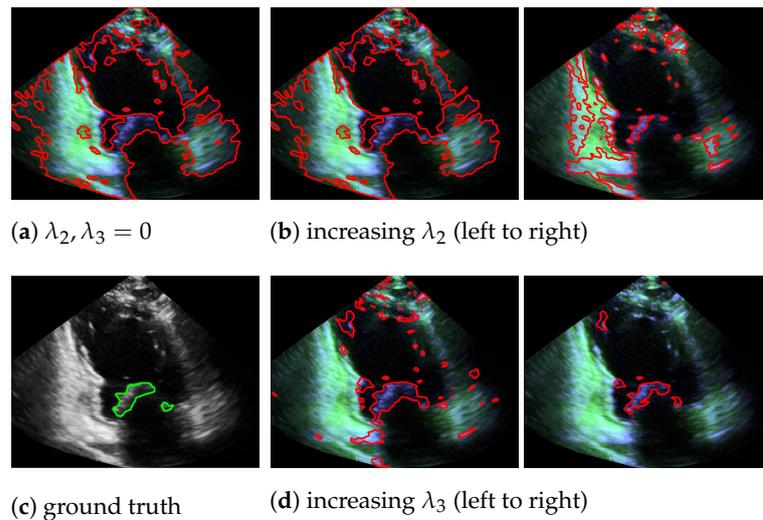


Figure 8. Segmentation of the mitral valve (red contour), depending on λ_2 and λ_3 of Equation (8), with the entries in $\hat{W}\hat{H}$ appear in green and the entries in \hat{S} in blue; the ground truth is highlighted with a green contour.

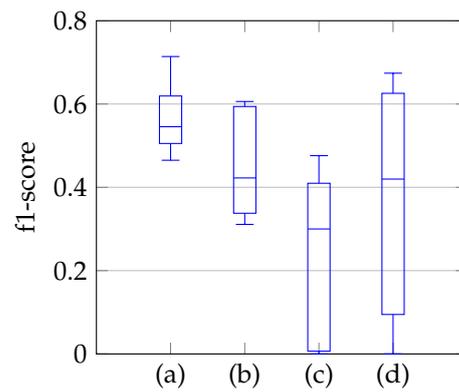
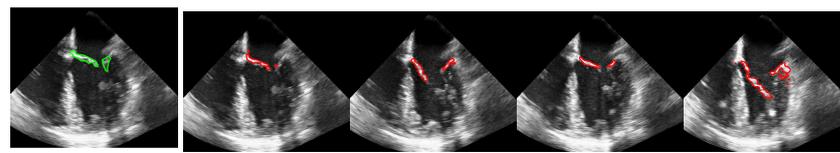


Figure 9. F1-score of segmentation result over ten videos for segmentation with windowing (a), automatic segmentation (b), segmentation by Dukler et al. (c) and segmentation by Corinzia et al. (d).

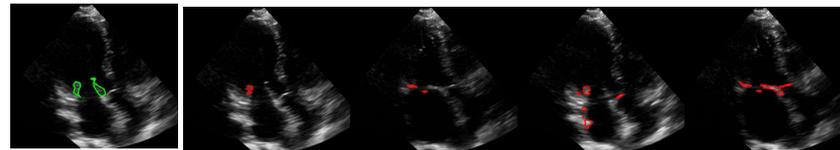
Table 4. Average recall, precision and f1-score using segmentation with windowing (a), automatic segmentation (b), segmentation by Dukler et al. (c) and segmentation by Corinzia et al. (d).

	Recall	Precision	f1-Score
(a)	0.494	0.692	0.565
(b)	0.44	0.558	0.45
(c)	0.165	0.551	0.244
(d)	0.378	0.43	0.377

The advantage of this method is that the mitral valve segmentation is performed without specifying a window size. In order to obtain a visual impression of the results, one of the best and the worst segmentations with respect to the f1-score are shown in Figure 10. Here, the example in the last row is particularly challenging because the mitral valve moves very quickly, as well as the heart muscle, and partially disappears in some frames. Nevertheless, the correct position of the valve is detected.



(a) Ground truth and one of the best segmentation results by automatic segmentation.



(b) Ground truth and one of the worst segmentation results by automatic segmentation.

Figure 10. One of the best (a) and the worst (b) segmentation results by automatic segmentation (red), with the corresponding ground truth (green).

The robustness of the automatic segmentation method against variations in the hyperparameters is shown in Figure 11, which shows the changes in the f1-score when changing the hyperparameter values in Equation (8). In each of these graphs, one parameter was varied and the others were fixed as specified in Table 2.

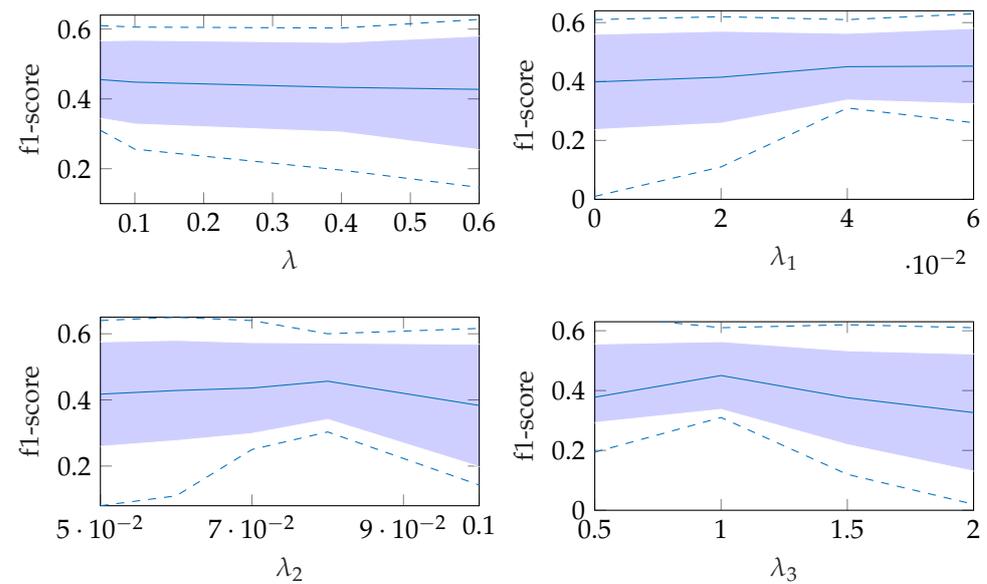


Figure 11. Effect of changing segmentation parameter for automatic segmentation, where the dashed lines show the minimum and maximum f1-score values and the standard deviation is indicated by the filled range.

5.3. Windowing

In window search, it is important that as much of the mitral valve structure as possible is covered by the window. Different views of the heart are accompanied by different sizes of the mitral valve appearance in the echocardiographic video. Therefore, it is necessary to know the required size of the window to cover the valve in the respective video if many different views of the heart should be supported by the method. To determine the amount of mitral valve structure that is covered by the window, we measure the recall of the ground truth and the detect window position (see Section 3.4). These values are compared to the recall values obtained by applying the method of Dukler et al. [9] and the method of Corinzia et al. [40] and visualized in the form of boxplot diagrams in Figure 12. We see that our method detects the mitral valve in the median more accurately and without negative outliers. Again, these outliers with a recall near zero indicate a missed mitral valve. The

robustness of the *windowing* method against variations in the hyperparameters is shown in Figure 13, which show the changes in the mean recall when changing the hyperparameter values of the method. In each of these graphs, one parameter was varied and the others were fixed as specified in Table 3.

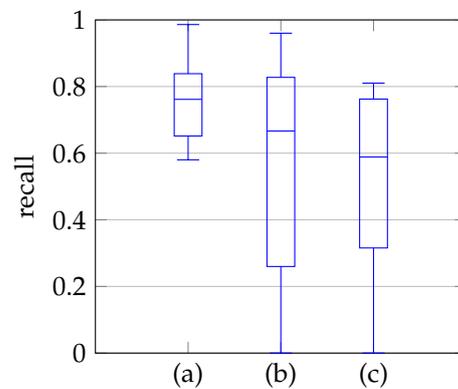


Figure 12. Recall of mitral valve detection over ten videos for Windowing (a), detection by Dukler et al. (b) and detection by Corinzia et al. (c).

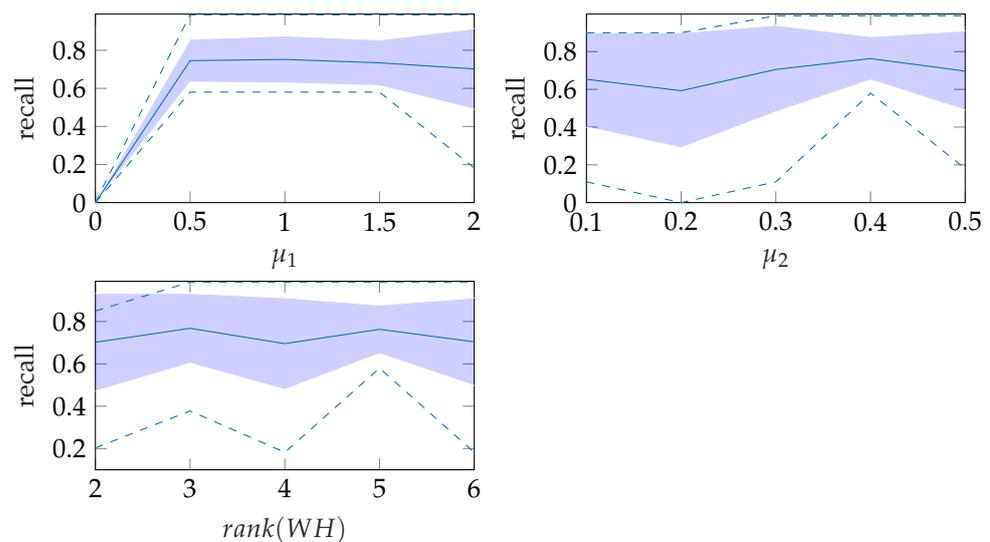


Figure 13. Robustness of windowing parameters, where the dashed lines show the minimum and maximum recall values and the standard deviation is indicated by the filled range.

5.4. Segmentation with Windowing

To segment the mitral valve, we pass the steps of Sections 3.1–3.3 with almost the same parameter settings as the method that does not include the windowing step and limit additionally the area in which the mitral valve may appear as described in Section 3.4. The effect of windowing is illustrated in Figure 14, which shows a segmentation with and without windowing. Without windowing, structures that are close to the mitral valve and move strongly may be segmented. These cases are handled by additional windowing. Since the windowing step limits the spatial range of the segmentation, here the weight λ_2 in (8) is set to a slightly lower value, compared to the segmentation without the windowing step, with the effect that more structure is detected as the mitral valve structure, and thus, the chance of potentially not capturing the whole valve is reduced.

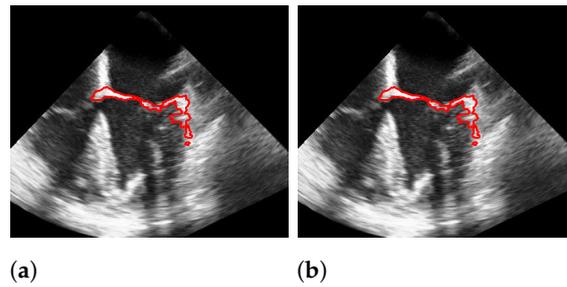


Figure 14. Effect of applying windowing on segmentation. The red marks indicate the segmentation. (a) without windowing; (b) with windowing.

The f1-score of the segmentation is in general larger with applying *windowing* instead of *automatic segmentation*, as shown in Figure 9. In addition, our segmentation achieves better results than [9,40]. By subtracting the resulting f1-scores of Dukler et al. [9] and Corinzia et al. [40] from the f1-scores of *segmentation with windowing* for each video, we obtain a comparison of the accuracy of these two methods for multiple echocardiographic videos as shown in Figure 15.

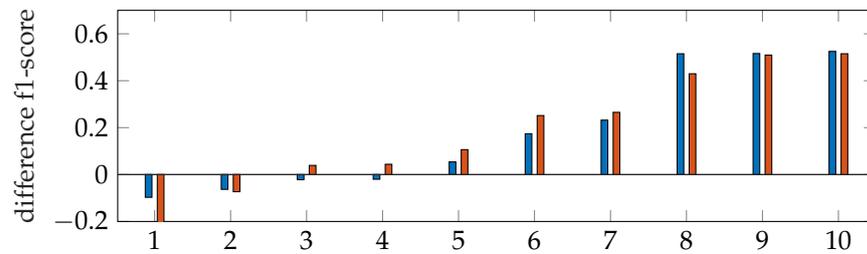


Figure 15. Ascending ordered difference of the f1-score for 10 videos: f1-score of segmentation with windowing subtracted by the f1-score of Dukler et al. (blue) and f1-score of segmentation with windowing subtracted by the f1-score of Corinzia et al. (red).

Here, the blue bars correspond to the measured differences with [9], and the red bars correspond to the differences with [40]. For both methods, it can be seen that the segmentation accuracy deviates less from ours in cases where the methods [9,40] are better (left bars) than in cases where our method has better accuracy (right bars).

For Dukler et al. [9], the segmentation results that correspond to the leftmost and the rightmost bars in Figure 15 are shown in Figure 16. The corresponding ground truths of the videos are shown in Figures 8 and 17. For Corinzia et al. [40], the segmentation results that correspond to the leftmost and the rightmost bars in Figure 15 are shown in Figure 18.

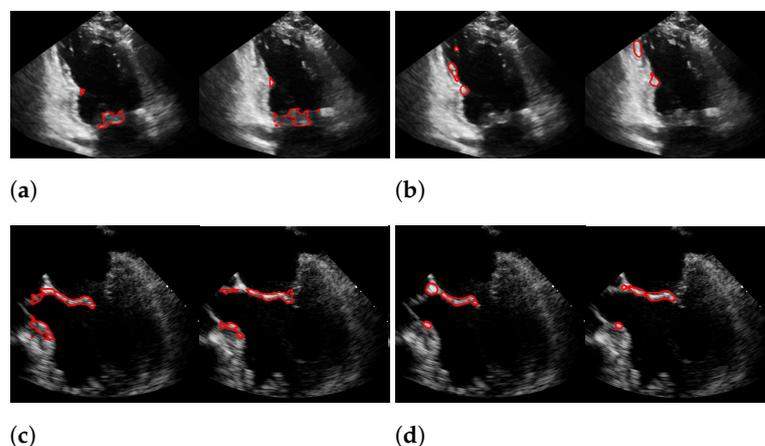


Figure 16. Segmentation in the cases of the largest differences between the f1-scores of our approach (segmentation with windowing) (a,c) and the one of Dukler et al. (b,d). The results in (a,b) belong to the blue rightmost bar of Figure 15, the results in (c,d) belong to the leftmost bar.

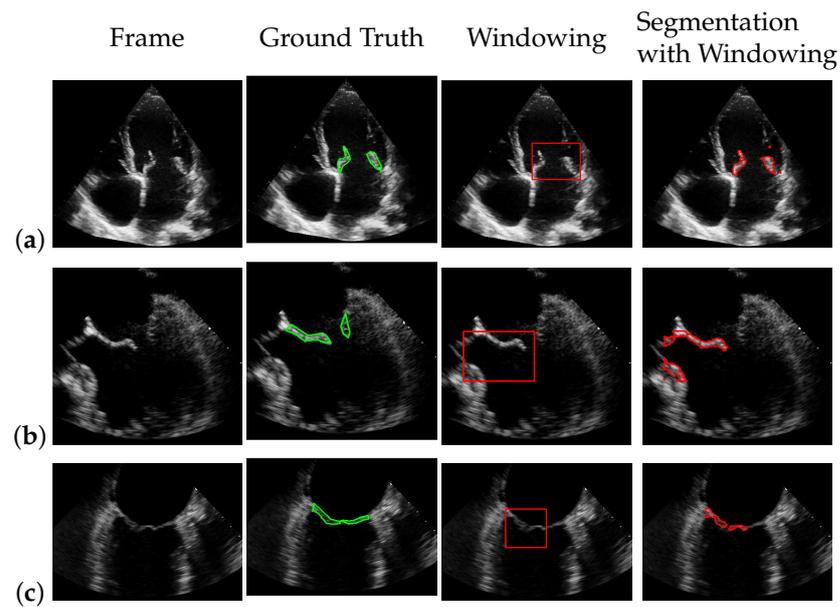


Figure 17. Best (a), worst (b) and an average (c) detection and segmentation result in terms of the recall and f1-score. The green contour indicates the ground truth segmentation, the red contour indicates the detected window location and the calculated segmentation.

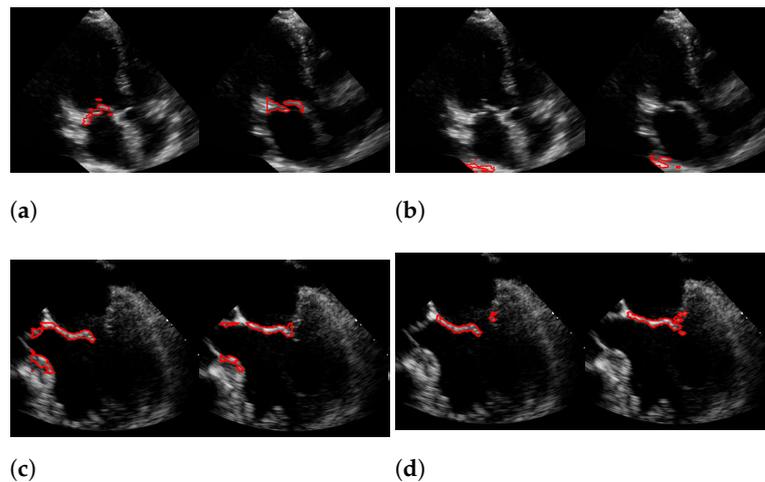


Figure 18. Segmentation in the cases of the largest differences between the f1-scores of our approach (segmentation with windowing) (a,c) and the one of Corinzia et al. (b,d). The results in (a,b) belong to the red rightmost bar of Figure 15, the results in (c,d) belong to the leftmost bar.

In both figures (Figures 16 and 18), the upper row of the segmentation by [9,40] gives an example of a missed mitral valve using a windowing technique and underlines the importance of correct valve detection. The upper row videos are especially challenging because of a partly disappearing mitral valve and weak contrast.

The bottom rows show the result on a video, including the mitral valve and the aortic valve on the left, which is also segmented. Since the final segmentation accuracy strongly depends on the windowing result the best/worst f1-score of the *segmentation with windowing* and the best/worst recall value of the *windowing* method on our evaluation dataset refer to the same videos as expected. These are shown in Figure 17. Here, the first row shows the best result and the second row, the worst. The latter is, again, caused by the second valve on left, which is also moving fast. The last row shows an average result of the mitral valve detection and segmentation.

During our experiments, we observed different kinds of failures in the segmentation with windowing, which we illustrated in Figure 19 with three example frames. The most

far-reaching failure, which appears in the form of a very low f1-score, is the complete missing of the mitral valve during windowing, as shown in the upper left image, and can occur in methods with windowing, as here in the method of Corinzia et al. [40]. The two images in the middle show failures that can occur due to the rectangular shape of the windows, where in some frames, parts of the heart muscle are segmented, as it is located in the area of the window during the contraction of the heart. The third type of failure is a slightly shifted window, as seen in the upper right image, which can occur when one valve leaflet stands out, due to strong pixel intensity and strong surrounding muscle movement in contrast to the other leaflet as well as when a second valve near the mitral valve appears in the video. This failure can also be seen in Figure 17 in the middle image in our worst result. Failures, which can occur in our method, due to too strong regularization of λ_1 or λ_2 in (8), are partially too narrow, perforated segmentations as shown in the lower right image. Here, the valve is correctly localized, but is not completely captured in some frames of the video.

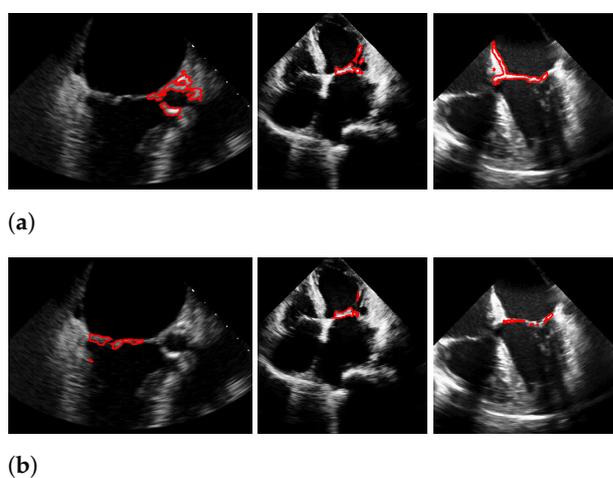


Figure 19. Types of segmentation failures. (a) segmentation of Corinzia et al. ; (b) segmentation with windowing.

5.5. Results on the EchoNet-Dynamic Dataset

For evaluation of the EchoNet-Dynamic dataset [52], we used the same hyperparameters as listed in Tables 2 and 3 for 46 selected videos, which almost all have the characteristic of showing a second heart valve besides the mitral valve. We measured an f1-score of 0.42 for *segmentation with windowing* and 0.35 for *automatic segmentation*, despite the presence of multiple valves in the videos. On the EchoNet-Dynamic dataset, [40] achieved an f1-score of 0.3 with their proposed method in combination with an optical flow approach. By omitting the optical flow, which is not optimal for the given data, due to its low resolution, they were able to increase the f1-score to 0.53. For the method of [9], we measured an f1-score of 0.1, due to shifts in their windowing approach.

As our method is not designed to distinguish between two valves, we performed additional experiments after masking the second heart valve from the videos and were thus able to increase the segmentation accuracy to an f1-score of 0.51 for *segmentation with windowing* and 0.43 for *automatic segmentation*. At this point, we refer to the conclusion for discussion to increase the robustness of our method against the presence of multiple valves in echocardiographic videos in future work.

6. Discussion

In this work, we propose a fully automatic and unsupervised method to detect and segment the mitral valve from echocardiographic videos by using a bias-free version of the RNMF model. In combination with a window-based localization approach and a modified version of the Chan–Vese segmentation method, we could improve the average f1-score value of the resulting segmentation by 0.18, compared to [40] on a dataset with

10 echocardiographic videos to a f1-score of 0.56. Furthermore, a possibility to segment the mitral valve without knowing its spatial size is introduced. The used technique, which employs a newly developed refinement method, is accompanied by a slight loss in accuracy.

During the experiments, we observed that failing cases included mitral valves with low contrast, fast movement of the heart muscle and recordings including more than one heart valve. However, we discovered that even in the video cases where the mitral valve is difficult to detect and segment because it partially disappears or is only in a slight contrast to the background, the correct position of the valve is detected. The method introduced in this paper represents a step toward fully automated mitral valve segmentation. However, especially when segmenting an anatomy as challenging as the mitral valve, it is essential that the quality of the results are critically controlled and not blindly trusted.

Our method offers some advantages, but also has weaknesses that need to be addressed in future work if applied in real medical settings. A limitation of our method is its time consumption since it takes a little over 8 min to calculate the segmentation of an ultrasound video with about 40 frames on an Intel(R) Core(TM) i7-7700HQ CPU @ 2.80 GHz. Therefore, the method is not suitable for time critical scenarios, such as intra-operative guidance, but could be sped up by implementing it on GPU. In addition, the method in [9] is not a real-time method, nor is the method in [40], which requires 13 ± 8 min per video for mitral valve segmentation on a GeForce GTX 1060 GPU but also offers a possibility for time-critical scenarios for which viability was shown in [40]. Furthermore, it should be noted as a weakness that the accuracy of our method, as well as that of the methods of [9,40], is not sufficient to be actually applied in real-life scenarios.

A strength of our method is that we avoid the inherent bias of ℓ^1 -regularization in solving the RNMF-model by using Bregman iterations, which gives us the promising advantage of balancing bias avoidance and uniqueness, unlike the method of [9], who uses classical RNMF and unlike the method in [40]. Corinzia et al. [40] compute first the cardiac muscle structures by the neural matrix factorization model and, in a subsequent step, the mitral valve structures by a thresholding network, which predicts a sparse signal, similar to the classical RNMF approach, without avoiding the inherent bias of the ℓ^1 -regularization. Another advantage of our method is that we can obtain valid segmentation results that are independent of a windowing method, as we use variational segmentation [12] and a refinement step instead of a simple subtraction of the cardiac muscle structures, as is the case in [9]. Dukler et al. [9] perform RNMF twice to extract the muscle structure from the data. The mitral valve is segmented by subtracting the muscle structure from the original video, which makes a windowing step crucial since no mechanism captures mis-segmented structures. Additionally, [40] need a windowing for their segmentation step since they obtain their resulting segmentation of the mitral valve by morphological opening followed by a connected component approach on the trained sparse signal in the windowed region of interest. Both our method and that of [9] use window detection (which is optional in our case); however, our method includes additional regularization that has the benefit of eliminating the co-occurrence of the muscle and valve on the same pixel. Both of these methods have the disadvantage that they cannot distinguish between the two heart valves. A solution to this problem is introduced in [40] in form of a time-varying weighting during windowing. The weighting depends on the phase of the cardiac cycle and thus allows to distinguish between right and left heart valves since the valves open and close at different time intervals.

In the future, it would be interesting to combine our methods with those in [40] to increase robustness. In particular, the time-variant weighting presented in [40] improves the detection accuracy of the mitral valve for cases where more than one valve is shown in the echocardiographic video. Since our windowing method has no mechanism to distinguish between the two valves, the introduction of time-varying weighting would likely increase segmentation accuracy. Additionally, for the fully automated method, time-variant weighting could be useful as a certainty measurement to distinguish low certain segmentation from high certain segmentation, and to improve the first mentioned by, for

example, interpolating the high certain segmentation. This would be of particular interest for fully removing the windowing step from the mitral valve segmentation pipeline to avoid failures and remove the dependency on window sizes.

Author Contributions: Conceptualization and methodology, H.D., B.Y., A.L.B. and M.M.; software, H.D.; validation, H.D.; visualization, H.D.; resources, R.L. and J.T.Y.; writing—original draft preparation, M.M. and H.D.; writing—review and editing, B.Y., A.L.B., M.M., R.L., J.T.Y. and H.D.; supervision, M.M., A.L.B. and B.Y.; data curation R.L. and H.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the institutional review board at the University of Southern California (Protocol ID:HS-15-00258).

Informed Consent Statement: The requirement of informed consent was waived by the IRB at the University of Southern California since this study performed secondary analysis on existing data and does not involve personal contact with participants.

Data Availability Statement: Restrictions apply to the availability of these data. The dataset is provided to the authors by the Keck Medical Center of the University of Southern California.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Pocock, G.; Richards, C.D.; Richards, D.A. *Human Physiology*; Oxford University Press: Oxford, UK, 2018; pp. 419–420.
- Mohty, D.; Pislaru, S. Valvular Heart Disease in Older Adults. 2019. Available online: <https://www.uptodate.com/contents/valvular-heart-disease-in-older-adults> (accessed on 5 December 2019).
- Nkomo, V.T.; Gardin, J.M.; Skelton, T.N.; Gottdiener, J.S.; Scott, C.G.; Enriquez-Sarano, M. Burden of valvular heart diseases: A population-based study. *Lancet* **2006**, *368*, 1005–1011. [[CrossRef](#)]
- Leech, G.J.; McCulloch, M.L.; Adams, D. Physical principles of ultrasound. In *Echocardiography*; Springer International Publishing: Cham, Switzerland, 2018; pp. 3–31. [[CrossRef](#)]
- Abraham, T. *Case Based Echocardiography: Fundamentals and Clinical Practice*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2011; p. V.
- Hauff, P.; Reinhardt, M.; Foster, S. Ultrasound basics. In *Molecular Imaging I*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 91–107. [[CrossRef](#)]
- Liu, X.; Cheung, Y.M.; Peng, S.J.; Peng, Q. Automatic mitral valve leaflet tracking in Echocardiography via constrained outlier pursuit and region-scalable active contours. *Neurocomputing* **2014**, *144*, 47–57. [[CrossRef](#)]
- Mikic, I.; Krucinski, S.; Thomas, J.D. Segmentation and tracking in echocardiographic sequences: Active contours guided by optical flow estimates. *IEEE Trans. Med. Imaging* **1998**, *17*, 274–284. [[CrossRef](#)]
- Dukler, Y.; Ge, Y.; Qian, Y.; Yamamoto, S.; Yuan, B.; Zhao, L.; Bertozzi, A.L.; Hunter, B.; Llerena, R.; Yen, J.T. Automatic valve segmentation in cardiac ultrasound time series data. In Proceedings of the Medical Imaging 2018: Image Processing, Houston, TX, USA, 11–13 February 2018; Volume 10574, p. 105741Y.
- Zhou, X.; Yang, C.; Yu, W. Automatic mitral leaflet tracking in echocardiography by outlier detection in the low-rank representation. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 972–979.
- Burlina, P.; Sprouse, C.; DeMenthon, D.; Jorstad, A.; Juang, R.; Contijoch, F.; Abraham, T.; Yuh, D.; McVeigh, E. Patient-specific modeling and analysis of the mitral valve using 3D-TEE. In Proceedings of the International Conference on Information Processing in Computer-Assisted Interventions, Geneva, Switzerland, 23 June 2010; Springer: Berlin/Heidelberg, Germany, 2010; pp. 135–146.
- Chan, T.F.; Vese, L.A. Active contours without edges. *IEEE Trans. Image Process.* **2001**, *10*, 266–277. [[CrossRef](#)] [[PubMed](#)]
- Costa, E.; Martins, N.; Sultan, M.S.; Veiga, D.; Ferreira, M.; Mattos, S.; Coimbra, M. Mitral Valve Leaflets Segmentation in Echocardiography using Convolutional Neural Networks. In Proceedings of the 2019 IEEE 6th Portuguese Meeting on Bioengineering (ENBENG), Lisbon, Portugal, 22–23 February 2019; pp. 1–4.
- Pal, N.R.; Pal, S.K. A review on image segmentation techniques. *Pattern Recognit.* **1993**, *26*, 1277–1294. [[CrossRef](#)]
- Fu, K.S.; Mui, J. A survey on image segmentation. *Pattern Recognit.* **1981**, *13*, 3–16.
- Kass, M.; Witkin, A.; Terzopoulos, D. Snakes: Active contour models. *Int. J. Comput. Vis.* **1988**, *1*, 321–331. [[CrossRef](#)]
- Boykov, Y.Y.; Jolly, M.P. Interactive graph cuts for optimal boundary region segmentation of objects in N-D images. In Proceedings of the Eighth IEEE International Conference on Computer Vision, Vancouver, BC, Canada, 7–14 July 2001; Volume 1, pp. 105–112.

18. Rother, C.; Kolmogorov, V.; Blake, A. Grabcut: Interactive Foreground Extraction Using Iterated Graph Cuts. *ACM Trans. Graph.* **2004**, *23*, 309–314.
19. Boykov, Y.; Funka-Lea, G. Graph cuts and efficient ND image segmentation. *Int. J. Comput. Vis.* **2006**, *70*, 109–131. [[CrossRef](#)]
20. Mohamed, A.A.; Arifi, A.A.; Omran, A. The basics of echocardiography. *J. Saudi Heart Assoc.* **2010**, *22*, 71–76. [[CrossRef](#)]
21. Martin, S.; Daanen, V.; Chavanon, O.; Troccaz, J. Fast segmentation of the mitral valve leaflet in echocardiography. In *International Workshop on Computer Vision Approaches to Medical Image Analysis*; Springer: Berlin, Germany, 2006; pp. 225–235.
22. Sultan, M.S.; Martins, N.; Veiga, D.; Ferreira, M.J.; Coimbra, M.T. Tracking of the anterior mitral leaflet in echocardiographic sequences using active contours. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 1074–1077.
23. Sultan, M.S.; Martins, N.; Costa, E.; Veiga, D.; Ferreira, M.J.; Mattos, S.; Coimbra, M.T. Tracking large anterior mitral leaflet displacements by incorporating optical flow in an active contours framework. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju Island, Korea, 11–15 July 2017; pp. 3244–3247.
24. Shang, Y.; Yang, X.; Zhu, L.; Deklerck, R.; Nyssen, E. Region competition based active contour for medical object extraction. *Comput. Med. Imaging Graph.* **2008**, *32*, 109–117. [[CrossRef](#)]
25. Yuan, B.; Chitturi, S.R.; Iyer, G.; Li, N.; Xu, X.; Zhan, R.; Llerena, R.; Yen, J.T.; Bertozzi, A.L. Machine learning for cardiac ultrasound time series data. In *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*; International Society for Optics and Photonics: Bellingham, WA, USA 2017; Volume 10137, p. 101372D.
26. Zhang, L.; Chen, Z.; Zheng, M.; He, X. Robust non-negative matrix factorization. *Front. Electr. Electron. Eng. China* **2011**, *6*, 192–200. [[CrossRef](#)]
27. Lee, D.D.; Seung, H.S. Learning the parts of objects by non-negative matrix factorization. *Nature* **1999**, *401*, 788–791. [[CrossRef](#)] [[PubMed](#)]
28. Guillaumet, D.; Vitria, J. Non-negative matrix factorization for face recognition. In *Catalonian Conference on Artificial Intelligence*; Springer: Berlin, Germany, 2002; pp. 336–344.
29. Lee, J.S.; Lee, D.D.; Choi, S.; Lee, D.S. Application of nonnegative matrix factorization to dynamic positron emission tomography. In Proceedings of the 3rd International Conference on Independent Component Analysis and Blind Signal Separation, San Diego, CA, USA, 9–13 December 2001; pp. 556–562.
30. Ji, Z.; Meng, G.; Huang, D.; Yue, X.; Wang, B. NMFBS: A NMF-based feature selection method in identifying pivotal clinical symptoms of hepatocellular carcinoma. *Comput. Math. Methods Med.* **2015**, *2015*, 846942. doi: 10.1155/2015/846942. [[CrossRef](#)] [[PubMed](#)]
31. Li, T.; Ding, C.C. Nonnegative matrix factorizations for clustering: A survey. In *Data Clustering*; Chapman and Hall/CRC: Boca Raton, FL, USA 2018; pp. 149–176.
32. Lee, D.D.; Seung, H.S. Algorithms for non-negative matrix factorization. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2000; pp. 556–562.
33. Cichocki, A.; Zdunek, R.; Amari, S.I. Hierarchical ALS algorithms for nonnegative matrix and 3D tensor factorization. In Proceedings of the International Conference on Independent Component Analysis and Signal Separation, London, UK, 9–12 September, 2007; Springer: Berlin, Germany, 2007; pp. 169–176.
34. Gillis, N. The why and how of nonnegative matrix factorization. *Connections* **2014**, *12*, 2.
35. Berry, M.W.; Browne, M.; Langville, A.N.; Pauca, V.P.; Plemmons, R.J. Algorithms and applications for approximate nonnegative matrix factorization. *Comput. Stat. Data Anal.* **2007**, *52*, 155–173. [[CrossRef](#)]
36. Ding, C.; Li, T.; Peng, W.; Park, H. Orthogonal nonnegative matrix t-factorizations for clustering. In Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 20–23 August 2006; pp. 126–135.
37. Hoyer, P.O. Non-negative matrix factorization with sparseness constraints. *J. Mach. Learn. Res.* **2004**, *5*, 1457–1469.
38. Pascual-Montano, A.; Carazo, J.M.; Kochi, K.; Lehmann, D.; Pascual-Marqui, R.D. Nonsmooth nonnegative matrix factorization (nsNMF). *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 403–415. [[CrossRef](#)]
39. Corinzia, L.; Provost, J.; Candreva, A.; Tamarasso, M.; Maisano, F.; Buhmann, J.M. Unsupervised Mitral Valve Segmentation in Echocardiography with Neural Network Matrix Factorization. In Proceedings of the Conference on Artificial Intelligence in Medicine in Europe, Poznan, Poland, 26–29 June 2019; Springer: Berlin, Germany, 2019; pp. 410–419.
40. Corinzia, L.; Laumer, F.; Candreva, A.; Taramasso, M.; Maisano, F.; Buhmann, J.M. Neural collaborative filtering for unsupervised mitral valve segmentation in echocardiography. *Artif. Intell. Med.* **2020**, *110*, 101975. Available online: <https://github.com/lucori/NN-MitralSeg> (accessed on 2 March 2021). [[CrossRef](#)] [[PubMed](#)]
41. Osher, S.; Burger, M.; Goldfarb, D.; Xu, J.; Yin, W. An iterative regularization method for total variation-based image restoration. *Multiscale Model. Simul.* **2005**, *4*, 460–489. [[CrossRef](#)]
42. Pedrosa, J.; Queirós, S.; Bernard, O.; Engvall, J.; Edvardsen, T.; Nagel, E.; D’hooge, J. Fast and fully automatic left ventricular segmentation and tracking in echocardiography using shape-based b-spline explicit active surfaces. *IEEE Trans. Med. Imaging* **2017**, *36*, 2287–2296. [[CrossRef](#)]
43. Ciecholewski, M. An edge-based active contour model using an inflation/deflation force with a damping coefficient. *Expert Syst. Appl.* **2016**, *44*, 22–36. [[CrossRef](#)]

44. Carneiro, G.; Nascimento, J.C.; Freitas, A. The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods. *IEEE Trans. Image Process.* **2011**, *21*, 968–982. [[CrossRef](#)] [[PubMed](#)]
45. Ali, Y.; Janabi-Sharifi, F.; Beheshti, S. Echocardiographic image segmentation using deep Res-U network. *Biomed. Signal Process. Control* **2021**, *64*, 102248. [[CrossRef](#)]
46. Bernard, O.; Bosch, J.G.; Heyde, B.; Alessandrini, M.; Barbosa, D.; Camarasu-Pop, S.; Cervenansky, F.; Valette, S.; Mirea, O.; Bernier, M.; et al. Standardized evaluation system for left ventricular segmentation algorithms in 3D echocardiography. *IEEE Trans. Med. Imaging* **2015**, *35*, 967–977. [[CrossRef](#)]
47. Osher, S.; Ruan, F.; Xiong, J.; Yao, Y.; Yin, W. Sparse recovery via differential inclusions. *Appl. Comput. Harmon. Anal.* **2016**, *41*, 436–469. [[CrossRef](#)]
48. Bachmayr, M.; Burger, M. Iterative total variation schemes for nonlinear inverse problems. *Inverse Probl.* **2009**, *25*, 105004. [[CrossRef](#)]
49. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Phys. D Nonlinear Phenom.* **1992**, *60*, 259–268. [[CrossRef](#)]
50. Condat, L. Discrete total variation: New definition and minimization. *SIAM J. Imaging Sci.* **2017**, *10*, 1258–1290. [[CrossRef](#)]
51. Chambolle, A.; Pock, T. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* **2011**, *40*, 120–145. [[CrossRef](#)]
52. Ouyang, D.; He, B.; Ghorbani, A.; Lungren, M.P.; Ashley, E.A.; Liang, D.H.; Zou, J.Y. *EchoNet-Dynamic: A Large New Cardiac Motion Video Data Resource for Medical Machine Learning*; NeurIPS ML4H Workshop: Vancouver, BC, Canada, 2019.