Image-based Incision Detection for Topological Intraoperative 3D Model Update in Augmented Reality Assisted Laparoscopic Surgery

Tom François[1,2], Lilian Calvet[1], Callyane Sève-d'Erceville[1], Nicolas Bourdel[1], Adrien Bartoli[1]

[1] Université Clermont Auvergne, CHU Clermont-Ferrand, CNRS, SIGMA Clermont, Institut Pascal.
[2] Be-Ys Research, 123 route de Meyrin, 1219 Châtelaine, Suisse.
`tom.francois@etu.uca.fr`

Abstract. Augmented Reality (AR) is a promising way to precisely locate the internal structures of an organ in laparoscopy. Several methods have been proposed to register a preoperative 3D model reconstructed from MRI or CT to the intraoperative laparoscopy 2D images. These methods assume a fixed topology of the 3D model. They thus quickly fail once the organ is cut to remove pathological internal structures. We propose to add image-based incision detection in the registration pipeline, in order to update the topology of the organ model. Whenever an incision is detected, it is transferred to the 3D model, whose topology is then updated accordingly, and registration started. We trained a UNet as incision detector from 181 labelled incision images, collected from 10 myomectomy procedures. It obtains a mean precision, recall and f1 score of 0.05, 0.36, and 0.08 from 10-fold cross-validation. Overall, topology updating improves 3D registration accuracy by 5% on average.

Keywords: Incision Detection · Laparoscopic Surgery · Registration.

1 Introduction

A large number of procedures are nowadays performed by laparoscopy, during which localising the internal anatomical structures such as tumours is a major challenge. An active research subject in Computer Assisted Intervention (CAI) is hence to develop laparoscopic surgery guidance systems using Augmented Reality (AR). AR is achieved by overlaying the laparoscopic video stream in realtime with the internal anatomical structures extracted as a preoperative 3D model from MRI or CT. AR thus relies on the ability to compute the geometric transformation between the preoperative 3D model and the laparoscopic video stream. This is the *registration* problem, which represents a strong technical challenge because of the organ deformation. The current realtime registration methods work in two steps [7, 14, 12]. First, they deform the preoperative 3D model to fit the intraoperative deformation state of the organ. This step uses an intraoperative 3D model reconstructed by means of Structure-from-Motion (SfM),
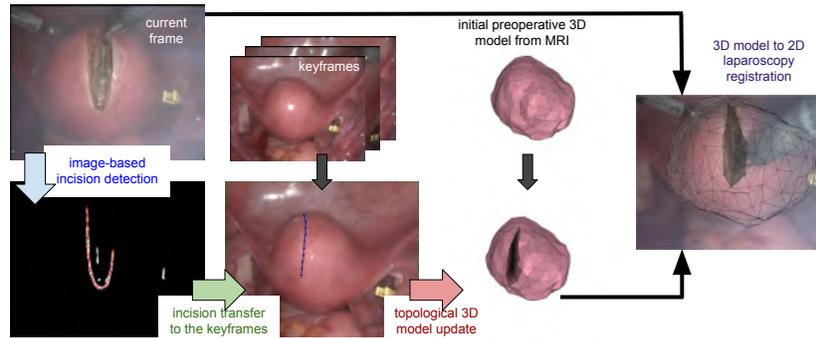
Fig. 1. Description of the complete pipeline. (column 1) Image-based incision detection. (column 2) Incision transfer to the keyframes. (column 3) Topological 3D model update. (column 4) Registered 3D model projected onto the current frame.

SLAM or stereovision. Second, they track the organ using the intraoperative 3D model. This process works well as long as the organ is mobilised but does not change its shape too much. However, when the surgeon starts incising the organ, it obviously affects its shape in much more dramatic ways than mobilisation, causing existing registration methods to break down. Consequently, AR is only available at the very early steps of surgery and stops as soon as organ incision starts. Our main objective is to develop a solution to this major limitation.

We propose a framework which both tracks the organ and adapts the intraoperative 3D organ model to the transformations undergone by the organ during surgery. Concretely, this means that we update the intraoperative 3D organ model. The adaptation handles changes in the 3D model shape due to deformation of the organ and changes in the 3D model topology due to incision of the organ. A closely related work to ours is [16], which adapts the 3D model topology using a *geometric criterion*. Their key idea is to compute the registration and then detect the incisions. Concretely, an excess of extension of a meshs edge triggers the deletion of the edge. This is an interesting idea but requires the registration to be highly accurate: noise in the registration, even temporary and mild, will cause spurious edge deletion, with no possibility of later recovery. Because deformable registration is a difficult problem, and because the model topology is wrong before registration, we cannot expect registration to be always highly accurate.

Our framework introduces a novel key idea in intraoperative 3D model update: *image-based incision detection*. Concretely, we detect the presence of an incision and its visible boundaries in the laparoscopy image. Importantly, this detection is independent of the registration. As opposed to the geometric criterion of [16], which depends on the registration, our image-based incision detection can be exploited to strengthen registration computation. Concretely, we use an incision detection DNN to obtain the image boundaries of the incision, which we transfer to the intraoperative 3D model, and update the model topology prior

to solving for the non-rigid registration. Transferring the incision boundaries to the intraoperative 3D model represents a key step, which we achieve by means of special deformable image warps exploiting the incision boundaries explicitly.

To summarise, our contributions are three-fold. Our first contribution is a novel framework for topological intraoperative 3D model update, in order to facilitate AR in the presence of organ incisions by exploiting the geometric and image-based incision criteria. This framework has two new core components. Our second contribution is a labelled database and an incision detection DNN. Our third contribution is an incision transfer method, from the laparoscopy image to the intraoperative 3D model, prior to registration. We evaluate our framework in three parts. First, we evaluate our incision detector using edge detection metrics. Second, we evaluate the benefit of updating the 3D model topology on an ex-vivo experiment using 4 ex-vivo pig kidneys. We perform an ablation study for our framework and compare it with previous work. Third, we evaluate on postoperative surgical patient images.

2   Related Work

Registration of deformable organs. The registration of deformable organs uses priors including deformation smoothness [4, 7] and bio-mechanics [3, 13]. Recent works based on deep learning [19] are yet inapplicable to laparoscopy. These methods all assume that the model topology is given and fixed.
Image-based detection. Image-based incision detection has not been specifically addressed in the literature, but many other detection and segmentation tasks have been explored [15]. In laparoscopy, tool detection and segmentation [9, 11] and organ-specific contour detection [14, 8] are active subjects. Existing works train an encoder-decoder derived from the U-Net [18] with dedicated datasets.
Incision simulation. Cutting a virtual 3D model can be achieved in several ways [20]. We have chosen *element deleting*, which is simple and runs fast.
Geometric incision detection. [16] uses a Finite Element Method to simulate the strain energy of the 3D model and visually controls it using SURF point correspondences. It uses a metric to measure how much has the distance changed between the deformed points. If some point pairs move apart significantly, a cut point is detected. The ex-vivo experiments described in [16] present large cuts with limited deformation but large piecewise rigid motion of the incised organ. The method requires that the registration is solved perfectly from the SURF points, which is very unlikely in practice.

3   Methodology

We base our pipeline on the uterus registration method [7]. We consider that the intraoperative 3D model has been reconstructed successfully using SfM [1, 2]. We denote the input data as $\{\mathcal{T}_i, \mathbf{p}_i, \mathbf{S}_0, \mathcal{I}, \mathbf{q}\}$, with $\mathcal{T}_i$ one keyframe, $\mathbf{p}_i$ the feature points in $\mathcal{T}_i$ corresponding to the feature points $\mathbf{q}$ in the current frame $\mathcal{I}$, and $\mathbf{S}_0$ the 3D model reconstructed from the N keyframes. In particular, $\mathbf{S}_0$ represents

the reconstructed *intact model*, as a closed surface. The output is $\{\mathbf{S}_t, \mathbf{Q}\}$, with $\mathbf{S}_t$ the incised 3D model and $\mathbf{Q}$ the registered model in the target frame.

3.1   Overview

Updating the 3D model according to the laparoscopic video stream involves to detect the incision both spatially and temporally. To simplify the problem, we make the following hypotheses. *(i)* A single incision is observed during the video sequence. *(ii)* The length of the incision only grows in time, which is always true. *(iii)* There is at least one frame where the incision is visible entirely, which is realistic as the surgeon checks the incision once done. These assumptions allow us to only detect the incision in the current frame and update the 3D model when the length of the detected incision has grown. Our pipeline can be summarized in 4 steps. (1) We detect the location of the incision in the current frame. (2) The detected incision location is transferred independently to multiple keyframes $\mathcal{T}_i$, where the organ is still intact. These keyframes have been used in the SfM reconstruction beforehand. This provides the transformation to both merge all transferred incisions in one reference keyframe, and to backproject the incision onto the reconstructed 3D model. (3) The 3D model is virtually cut according to the detected incision. (4) The registration is estimated between the updated 3D model and the current frame. In this pipeline, incision under-detection is very well-handled and hence preferred to over-detection.

3.2   Image-based Incision Detection

We collected and annotated a total of 181 images from 10 myomectomy procedure videos. All enrolled participants gave their written informed consent following the IRB approval 2018-A03130-55. We trained a UNet [18] architecture to predict the visible boundaries of the incision. Our network produces an output $P(x,y)$ representing the probability of having an incision boundary at (x,y). The actual detected incision is obtained by selecting the main connected component, which is thinned and converted to a polyline. This post-processing removes many false positive pixels, contributing to the methods accuracy. The dataset contains various incision techniques that used different types of tools. The first type is the regular scissors. The boundaries of the incision made by these are difficult to perceive if the incision is not bleeding. The second type of tools are the heat cutting tools. The boundaries of the incision are burned while cutting which makes them easier to perceive but the cutting may also produce smoke.

3.3   Image to Model Incision Transfer

Image deformation model. The image deformation model we use is a parametric warp. There is a large variety of such functions in the literature. We use the generic model [5]. The warp $\mathcal{W} : \mathbb{R}^2 \times \mathbb{R}^{l \times 2} \mapsto \mathbb{R}^2$ maps 2D points from the current image to the reference image and depends on a set of l 2D control

Fig. 2. Image-based incision detection on test images. The groundtruth is in blue, the detection in red and the overlap in green.

points $c_1, \ldots, c_l$ stacked in the parameter matrix $L \in \mathbb{R}^{l \times 2}$. The parameters are chosen by minimizing a cost function ε, composed of a data term ε_d , based on the average distance between the warped points in $\mathbf{q}$ and the matched points in $\mathbf{p}_i$, a smoothing term ε_s that controls the smoothness of the motion field, and a shrinking term ε_f that forces the warp to close the incision area. The first two terms are described in [17]:

$$\varepsilon_d(L) = \frac{1}{n_c} \sum_{j=1}^{n_c} \|\mathcal{W}(\mathbf{q}_j, L) - \mathbf{p}_{i,j}\|_2^2 \qquad \varepsilon_s(L) = \frac{1}{m_2} \|ZL\|_{\mathcal{F}}^2, \qquad (1)$$

with n_c the number of correspondences, Z the second derivatives of $\mathcal{W}$ stacked, evaluated at m_2 points. Minimizing the data and smoothing terms has a closed-form solution [5], which we use to initialise our solution.

Closing the incision. In the keyframe $\mathcal{T}_i$, the boundaries of the incision should be superimposed to each other, as the organ is still intact. Thus, the estimated warp should close the detected incision area, denoted $\mathcal{H}$. Several works on warp estimation [17, 5, 10] have tackled this issue for self-occlusions, which occur when some parts of an observed surface are occluded by itself. In a self-occluded area, the data term gives absolutely no constraint. The best method from [10] uses a shrinking term:

$$\varepsilon_f(L) = \frac{1}{m_1} \sum_{\mathbf{a} \in \mathcal{H}} \min_{d \in \mathbb{S}^1} \left\| \frac{\partial_d \mathcal{W}(\mathbf{a}, L)}{\partial \mathbf{a}} \right\|_{\mathcal{F}}^2, \qquad (2)$$

with $\mathbb{S}^1$ the unit circle, and $\frac{\partial_d \cdot}{\partial \mathbf{a}}$ the directional derivatives along direction d, evaluated at m_1 points. The incision area is geometrically similar to a self-occlusion of the target image $\mathcal{I}$. We thus use Levenberg-Marquardt to minimise:

$$\varepsilon(L) = \varepsilon_d(L) + \lambda_s \varepsilon_s(L) + \lambda_f \varepsilon_f(L). \qquad (3)$$

Keyframes mutual agreement. The image to keyframe transfer is computed for every keyframe $\mathcal{T}_i$ that has been used in SfM to reconstruct $\mathbf{S}_0$. SfM provides the relative pose of every keyframe, which allows us to accurately transfer any point of the organ across the keyframes. For every estimated warp $\mathcal{W}_i$ from the set of correspondences $(\mathcal{I} \leftrightarrow \mathcal{T}_i)$, we can thus transfer the warped incision boundaries in a reference keyframe $\mathcal{T}_r$. As all the resulting warped incisions should align in the reference frame $\mathcal{T}_r$, we form a robust average using the median of the transferred incisions, see figure 3. The final transferred incision is obtained by thinning.

Fig. 3. Incision transfer to a reference keyframe. Left: current frame with the detected incision boundaries. Middle: incision boundaries transferred to the reference keyframe. Right: median warp (yellow) and final transferred incision (green).

3.4   Topological Model Update

Whenever the detected incision length increases, we create a 3D model that represents the incision. Once the incision is backprojected on the intact model surface, we create the outline of the incision. The bottom of the incision is defined by moving at a fixed depth along the surface normal. The depth and width of the incision are manually set for each procedure.

3.5   3D-2D Registration

We formulate the problem as a non-linear energy-minimisation problem, with energies coming from prior and data terms. The prior term encodes the model's internal energy, which is used to regularise the problem. The data term enforces the projected registered 3D model to match the 2D correspondences $\mathbf{q}$ in the target frame. The energy function $E \in \mathbb{R}^+$ is given by:

$$E(x) = E_c(x, \mathbf{p}, \mathbf{q}) + \lambda_i E_i(x), \tag{4}$$

where λ_i is a scalar weight to control the non-rigid deformation. We optimize E by iterative non-linear optimization using a *stiff-to-flexible* strategy [6].

4   Results

Acquiring valid and precise groundtruth is a common challenge in surgical AR. The image-based incision detection can be evaluated very efficiently on a test set of labelled clinical images. The rest of the pipeline is evaluated on ex-vivo organs, for which the image-based detection is controlled. Qualitative results of the complete pipeline on patient data are shown in figure 4.

4.1   Image-based Incision Detection

For evaluation, we applied 10-fold cross-validation using the 10 procedures forming 10 image sets. For every permutation, the trained model is applied on the left-out images, considered as test set. The mean results for the concatenated test sets are 0.049, 0.356 and 0.084 for precision, recall and f1 score respectively.
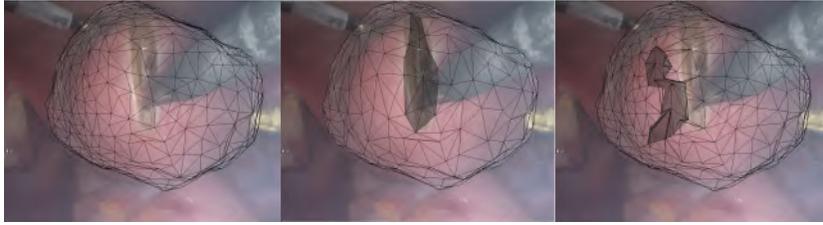
Fig. 4. Visual comparison on patient data: from left to right, Intact (baseline), ImageDet (ours), and GeoDet [16]. The detected incision is colored in black on the mesh.

We notice that the quality of the predictions differs greatly between procedures (see figure 2). We obtained the best results with the regular Weighted Cross Entropy (WCE), compared to training strategies with the combination of WCE with Tversky Loss. We initialize the UNet with weights pretrained for occluding contours detection of the uterus [8] and train using a step learning rate strategy for 80 epochs, scaling down every 10 epochs by 0.1. The initial learning rate is set to 10^{-4} and the model stops improving after 50 epochs. The class weights for WCE are set to 10 and 0.1 for incision and non-incision respectively.

4.2   Ex-vivo Registration

Description. We conducted an ablation study over 4 ex-vivo pig kidneys. We reconstruct the 3D scene before and after the incision using SfM, obtaining respectively the intact and groundtruth incised models. We recorded the video while the kidney was incised using a scalpel. Kidneys 1-3 are deeply incised, namely throughout the organ, while kidney 4 undergoes a shallow incision (6-10 mm). The average kidney dimensions are $140 \times 75 \times 25$ mm. Our experiments are similar to [16] but with less displacement, which is more realistic. For warp estimation, we used $\lambda_s = 0.5, \lambda_f = 5$ and a 5×5 grid of control points. For the registration, λ_i is set to 10^{-10}.

Image to model incision transfer. The transferred incision has been evaluated on its own with a manually annotated ground-truth. We obtain an Hausdorff distance and symmetric distance of 15.8px and 6.1px respectively. The average RMSR over the different kidneys is 8.5px.

Geometric incision detection. We have reimplemented [16], which we call GeoDet. The main difference is that we use a different registration method. For our two last experiments, we have not been able to make GeoDet work. The method either detects many false cut points or detects no incision. We use GeoDet with $r_{P_F^0} = 30$ mm, $\tau = 4.5$, $n = 7$.

Evaluation. In the ablation study, we compare the methods Intact [7], ImageDet (ours), ImageDet-50% (ours), and GeoDet [16]. ImageDet-50% simply uses half of the incision to simulate misdetection. In table 1, we compare the mean reprojection error between the registered model and the manual correspondences. In

		2D (px)	A (iGT)	A (GTi)	ROI (iGT)	ROI (GTi)
K1	Intact	8.92±7,05	5±2.11	5.14±2.7	4.95±2.06	5.42±2.99
	ImageDet	**8.85±7.23**	4.73±2.89	4.11±2.14	4.95±3.56	**3.57±1.92**
	ImageDet-50%	9.02±7.45	**4.15±1.85**	**4.11±2.11**	**4.02±1.95**	3.8±2.05
	GeoDet	8.88±7.03	5.12±2.88	4.56±2.48	5.12±3.34	3.88±2.32
K2	Intact	6.09±4.03	10.82±5.6	9.23±5.48	14.47±5.02	12.1±5.66
	ImageDet	6.07±3.97	**10.7±5.34**	**8.6±4.84**	13.62±4.65	**10.41±4.59**
	ImageDet-50%	**6.07±3.98**	10.8±5.43	9.03±5.15	14.08±4.73	11.55±5.06
	GeoDet	6.1±4.04	10.81±5.5	8.89±5.18	**13.44±5.44**	11.02±5.51
K3	Intact	8.84±7.96	13.45±6.51	11.56±6.21	15.26±6.52	14.13±6.32
	ImageDet	8.5±8.42	**11.44±5.77**	**9.35±5.33**	**11.46±5.95**	**10.07±5.62**
	ImageDet-50%	**8.21±7.6**	12.85±6.36	10.79±5.94	13.04±6.36	11.97±6.27
	GeoDet	-	-	-	-	-
K4	Intact	5.69±5.68	2.4±1.9	2.05±1.27	1.65±1.33	1.63±1.2
	ImageDet	5.68±5.62	2.4±1.91	1.99±1.25	1.81±1.48	1.56±1.17
	ImageDet-50%	**5.68±5.61**	**2.32±1.88**	**1.97±1.23**	**1.61±1.28**	**1.54±1.14**
	GeoDet	-	-	-	-	-

Table 1. Evaluation of ex-vivo kidney data for the tested approaches . '-' in GeoDet means that no incision was detected. '2D' is the reprojection distance error (in px). 3D distances are presented for the external registered surface (A) and a cropped region centered on the incision (ROI). Results are expressed with mean and standard deviation (the lower, the better; best in bold, second best underlined).

order to compare the different methods, we measure the sampled distance between the external surface of the registered model with the groundtruth model. We also provide the same distances on a Region of Interest centered on the incision location. the registration examples from figures 4 and 5 show how ImageDet manages to consistently update and register the 3D model, and outperforms GeoDet.

5   Discussion and Future Work

We have given the first pipeline which detects incisions from the images in endoscopic surgery and uses the information to perform topological preoperative 3D model registration. Our dataset for incision detection is still small. We expect a significant performance boost upon expanding it. We expect a combination of the ImageDet (ours) and GeoDet [16] methods to improve precision and robustness. Finally, we plan to include new terms in the registration loss, to associate the image detected incision boundaries to the 3D model.

References

1. Meshroom, https://github.com/alicevision/meshroom
2. Photoscan, https://www.agisoft.com
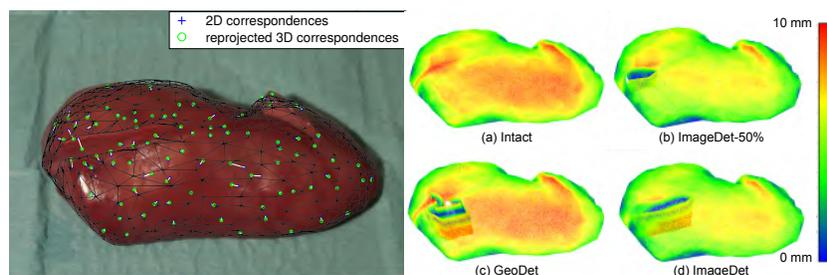
Fig. 5. Visual results for registration on kidney 1. Left: Reprojection error between correspondences (white line) for ImageDet. Right: Distance from registered model to groundtruth with methods Intact (a), ImageDet-50% (b), GeoDet (c), ImageDet (d).

3. Adagolodjo, Y., Trivisonne, R., Haouchine, N., Cotin, S., Courtecuisse, H.: Silhouette-based pose estimation for deformable organs application to surgical augmented reality. In: IROS. pp. 539–544. IEEE (2017)

4. Amir-Khalili, A., Nosrati, M.S., Peyrat, J.M., Hamarneh, G., Abugharbieh, R.: Uncertainty-encoded augmented reality for robot-assisted partial nephrectomy: A phantom study. In: AE-CAI, pp. 182–191. Springer (2013)

5. Bartoli, A.: Maximizing the predictivity of smooth deformable image warps through cross-validation. JMIV **31**(2-3), 133–145 (2008)

6. Collins, T., Chauvet, P., Debize, C., Pizarro, D., Bartoli, A., Canis, M., Bourdel, N.: A system for augmented reality guided laparoscopic tumour resection with quantitative ex-vivo user evaluation. In: CARE@MICCAI. pp. 114–126 (2016)

7. Collins, T., Pizarro, D., Gasparini, S., Bourdel, N., Chauvet, P., Canis, M., Calvet, L., Bartoli, A.: Augmented reality guided laparoscopic surgery of the uterus. TMI (2020)

8. François, T., Calvet, L., Zadeh, S.M., Saboul, D., Gasparini, S., Samarakoon, P., Bourdel, N., Bartoli, A.: Detecting the occluding contours of the uterus to automatise augmented laparoscopy: score, loss, dataset, evaluation and user study. IJCARS **15**(7), 1177–1186 (2020)

9. Garcia-Peraza-Herrera, L.C., Li, W., Fidon, L., Gruijthuijsen, C., Devreker, A., Attilakos, G., Deprest, J., Vander Poorten, E., Stoyanov, D., Vercauteren, T., et al.: Toolnet: holistically-nested real-time segmentation of robotic surgical tools. In: IROS. pp. 5717–5722. IEEE (2017)

10. Gay-Bellile, V., Bartoli, A., Sayd, P.: Direct estimation of nonrigid registrations with image-based self-occlusion reasoning. TPAMI **32**(1), 87–104 (2008)

11. Han, L., Wang, H., Liu, Z., Chen, W., Zhang, X.: Vision-based cutting control of deformable objects with surface tracking. IEEE/ASME Transactions on Mechatronics (2020)

12. Haouchine, N., Dequidt, J., Berger, M.O., Cotin, S.: Monocular 3d reconstruction and augmentation of elastic surfaces with self-occlusion handling. IEEE transactions on visualization and computer graphics **21**(12), 1363–1376 (2015)

13. Haouchine, N., Dequidt, J., Peterlik, I., Kerrien, E., Berger, M.O., Cotin, S.: Image-guided simulation of heterogeneous tissue deformation for augmented reality during hepatic surgery. In: ISMAR. pp. 199–208. IEEE (2013)

14. Hattab, G., Arnold, M., Strenger, L., Allan, M., Arsentjeva, D., Gold, O., Simpfendörfer, T., Maier-Hein, L., Speidel, S.: Kidney edge detection in laparoscopic image data for computer-assisted surgery. IJCARS **15**(3), 379–387 (2020)
15. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis. Medical image analysis **42**, 60–88 (2017)
16. Paulus, C.J., Haouchine, N., Kong, S., Soares, R.V., Cazier, D., Cotin, S.: Handling topological changes during elastic registration. IJCARS **12**(3), 461–470 (2017)
17. Pizarro, D., Bartoli, A.: Feature-based deformable surface detection with self-occlusion reasoning. IJCV **97**(1), 54–70 (2012)
18. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: MICCAI. pp. 234–241 (2015)
19. de Vos, B.D., Berendsen, F.F., Viergever, M.A., Sokooti, H., Staring, M., Išgum, I.: A deep learning framework for unsupervised affine and deformable image registration. Medical image analysis **52**, 128–143 (2019)
20. Wu, J., Westermann, R., Dick, C.: A survey of physically based simulation of cuts in deformable bodies. Comput. Graph. Forum **34**(6), 161–187 (2015)