Wei Wei, Haishan Xu, Julian Alpers, Zhang Tianbao,
Lei Wang, Marko Rak, Christian Hansen

Fast Registration for Liver Motion Compensation in Ultrasound-guided Navigation

Pre-print version

Haishan Xu, Zhang Tianbao, Lei Wang,

Department of Diagnostic Ultrasound

Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, China

Wei Wei, Julian Alpers, Marko Rak,
Christian Hansen
Faculty of Computer Science,
Otto-von-Guericke University Magdeburg, Germany

christian.hansen@ovgu.de

This is a pre-print of an article published in the IEEE International Symposium on Biomedical Imaging (ISBI 2019).

FAST REGISTRATION FOR LIVER MOTION COMPENSATION IN ULTRASOUND-GUIDED NAVIGATION

Wei Wei[a,*], *Haishan Xu*[b,*], *Julian Alpers*[a], *Zhang Tianbao*[b], *Lei Wang*[b], *Marko Rak*[a], *Christian Hansen*[a]

[a] Faculty of Computer Science & Research Campus STIMULATE, University of Magdeburg, Germany
[b] Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, China

ABSTRACT

In recent years, image-guided thermal ablations have become a considerable treatment method for cancer patients, including support through navigational systems. One of the most critical challenges in these systems is the registration between the intraoperative images and the preoperative volume. The motion secondary to inspiration makes registration even more difficult. In this work, we propose a coarse-fine fast patient registration technique to solve the problem of motion compensation. In contrast to other state-of-the-art methods, we focus on improving the convergence range of registration. To this end, we make use of a Deep Learning 2D U-Net framework to extract the vessels and liver borders from intraoperative ultrasound images and employ the segmentation results as regions of interest in the registration. After an initial 3D-3D registration during breath hold, the following motion compensation is achieved using a 2D-3D registration. Our approach yields a convergence rate of over 70% with an accuracy of 1.97 ± 1.07 mm regarding the target registration error. The 2D-3D registration is GPU-accelerated with a time cost of less than 200 ms.

Index Terms— U-Net, CMA-ES, CUDA, Registration

1. INTRODUCTION

Liver tumor ablation techniques such as microwave or radiofrequency ablation have become relevant treatment options to handle metastases in recent years [1, 2]. A number of intraoperative imaging modalities like ultrasound (US) or computed tomography (CT) can be used to guide the performing radiologist during needle insertion. Compared to CT, US devices have the advantage of real-time image acquisition which is necessary to handle the motion of the internal organs without breath control devices. Therefore, radiologists may prefer using US in liver ablation procedures. Unfortunately, due to its low image quality US images rarely show tissue details, especially small tumors or other small risk structures. To overcome the lack of visibility, interventional navigation systems have been developed in recent years to guide the radiologist during needle insertion [3, 4]. Regarding navigation accuracy, the most critical part is the registration of the preoperative volume to the intraoperative data.

One type of method covers the static 3D-3D registration. These approaches usually yield a higher global accuracy during breath holding intervals but often lack the ability for small local adjustments within a Region-Of-Interest (ROI) around the target structure. Penny et al.[5] have introduced a navigation system where registration relies on manual annotation and a simple Iterative-Closest-Points algorithm. In [6], a 3D US volume is reconstructed from 2D laparoscopic US images and aligned with a CT volume using a stochastic optimizer. Lange et al. [7] propose a method of combining anatomical landmark information with TPS non-rigid wrapping method which yields a promising registration result. Haque et al.[8] have proposed a coarse-to-fine registration method which achieves high accuracy on both MR and CT data with a tolerable computation time of 40s. These approaches require breath holding during the scan. Once the breath holding machine is turned off, the registration is disturbed by breath motion up to several centimeters. Banerjee [9] takes the advantage of the 3D ultrasound device and introduces a registration with block-matching technique. His method yields a high registration accuracy 1.8 mm with low time cost of 125 ms.

Another approach is the registration of 2D US images to a 3D volume without reconstruction. These approaches usually do not provide a high global accuracy but focus on minimization of the target registration error (TRE) in a specific ROI. Wein et al. [10] apply a method to generate synthetic US images out of a resliced CT volume. Afterwards, the synthetic US image is aligned with the real US image using the correlation coefficient as similarity measurement. Xu et al. [11] extend mutual information similarity with additional spatial mutual information and show a promising registration result. Weon et al. [12] propose a real-time registration method between 4D preoperative image and intraoperative 2D ultrasound image. They take the advantages of acquired 4D preoperative images to demonstrate breath motion and a fixed 3D US probe which can capture 3D US volume for the rough pose estimation and 2D US image for the motion compensation. Although the registration method outperforms

*These authors contributed equally to this work.

other state-of-the-art methods, their clinic setups are significantly different to the common clinical workflow of the liver tumor ablation procedure. In this work, we propose a fast registration pipeline to compensate translational liver movement caused by patient breathing. In contrast to the other state-of-the-art methods, our goal was to speed up and improve the convergence range of image registration, which enables a fully automatic registration without any manual initial pose placement and handles large motion secondary to deep inspiration. The main contribution to achieve this is two stage coarse-fine approach combining 3D-3D registration to find the initial alignment and a fast 2D-3D registration to optimize the accuracy while ensuring a rather wide convergence range.

Authors	Type	Accuracy [mm]	Time [s]
Penny et al. [5]	3D-3D	$10 \pm ?$	300
Fusaglia et al. [6]	3D-3D	$5 \pm ?$	720
Banerjee et al. [9]	3D-3D	$1.8 \pm ?$	0.125
Haque et al. [8]	3D-3D	3.22 ± 1.91	40
Wein et al. [10]	2D-3D	$8.1 \pm ?$	40
Xu et al. [11]	2D-3D	4.13 ± 1.27	76
Weon et al. [12]	2D-4D	2.42 ± 0.87	0.06
Our method	2D-3D	1.97 ± 1.07	0.16

Table 1. Overview of the related work regarding the patient registration compared to our approach.

2. MATERIALS AND METHODS

2.1. System Setup

In this setup, we used a Northern Digital Inc. Polaris Vicra camera to track passive optical tracking markers, which are mounted onto the US probe. As long as the line of sight is not interrupted, this device allows an accurate tracking of the probe providing the transformation ${}^{Cam}T_{US\_P}$. The relation between the probe and the US image ${}^{US\_P}T_{US\_I}$ is given by calibrating the probe once in advance using a specially designed calibrator. The US image coordinate system is defined in 3D space, which has the same origin and aligned axes as the 2D US image coordinate system. The missing transformation from the preoperative CT volume to the intraoperative US image ${}^{CT}T_{US\_I}$ is initially computed using the coarse registration approach described in the Section 2.2 and then improved by 2D-3D fine registration discussed in the following Sections 2.2 and 2.3. The final visualization of the navigation system is based on the preoperative CT volume data. Therefore, the final transformation from the NDI tracking camera to the preoperative volume ${}^{Cam}T_{CT}$ is given by Equation 1.

$$ {}^{Cam}T_{CT} = {}^{Cam}T_{US\_P} * {}^{US\_P}T_{US\_I} * ({}^{CT}T_{US\_I})^{-1} \quad (1) $$

2.2. Preprocessing

As depicted in Figure 2, US images mainly reveal the anatomy of vessels and the boundary of the liver surface. On the other hand, a CT volume provides a wide capture range and shows more organs and details of tissues. To handle image variability between CT and US images, our registration method is focused on both, the vessel structures and the boundary of the liver. Thus a 2D U-Net supporting multiple classes was implemented to extract the necessary information from the US images as a variation of the original U-Net introduced in [13]. In this network, stochastic gradient descent was used as the optimizer with a momentum of 0.85 with a learning rate of 0.05. Dropout layers were used in order to improve the generalization capability of the network with a drop-rate of 0.5. The network has been extended with multiple classes including background, liver boundary and vessel. A cost function was defined as a weighted summary of Dice similarity coefficient of the three components using the coefficient 0.001, 0.5 and 1.0 for background, liver boundary and vessel, respectively. The network is trained with 262 US images (with ground truth segmentation) from seven different patients. Eighty percent of the images were used for training and 20 percent for validation. To avoid the correlation of training and test datasets, the leave-one-patient-out strategy is applied with seven folds. The prediction results show a good segmentation with mean Dice scores of 0.5 for vessel structures and mean Dice scores of 0.6 for liver boundary. This segmentation provides an adequate quality of mask which is used in image registration. The result of prediction is used to define a mask (Figure 2) for the similarity measurement where Φ_v and Φ_b correspond to the vessel and liver boundary respectively.

During the perioperative phase, tracked US images are acquired by swiping the patients while breath holding. The acquired images are processed in real-time using the trained 2D U-Net. Afterwards the segmented 2D vessel structures are reconstructed into 3D space. A rigid registration is performed to align the preoperative CT volume and the reconstructed vessel structures using a cost function calculating the overlap of the vessel structures and an optimizer: evolutionary strategy of a co-variance matrix adaptation (CMA-ES) [14] which can maximize the overlap.

2.3. 2D-3D image registration

Due to the variability of liver tissue between US images and CT images the detected masks are applied on US images to minimize the possible ambiguities and to reduce the computational effort. Furthermore, gray values of US images are inverted to reduce the variability of vessel appearance between the US and enhanced CT image. In addition, we make use of the gradient orientation (GO) to measure the similarity of vessel structures.

$$ GO = \begin{cases} \frac{1}{N}\sum_{i=1}^{m} \frac{2-ln(|arccos(cos(\theta_i))|+1)}{2}, & \Phi_b \cap \Psi_b > 0 \\ \frac{2-ln(\pi+1)}{2}, & \Phi_b \cap \Psi_b = 0 \end{cases} $$

$$ cos(\theta_i) = \frac{\nabla I_{CT\_r}(i) * \nabla I_{US}(i)}{\|\nabla I_{CT\_r}(i)\| * \|\nabla I_{US}(i)\|} \quad (2) $$

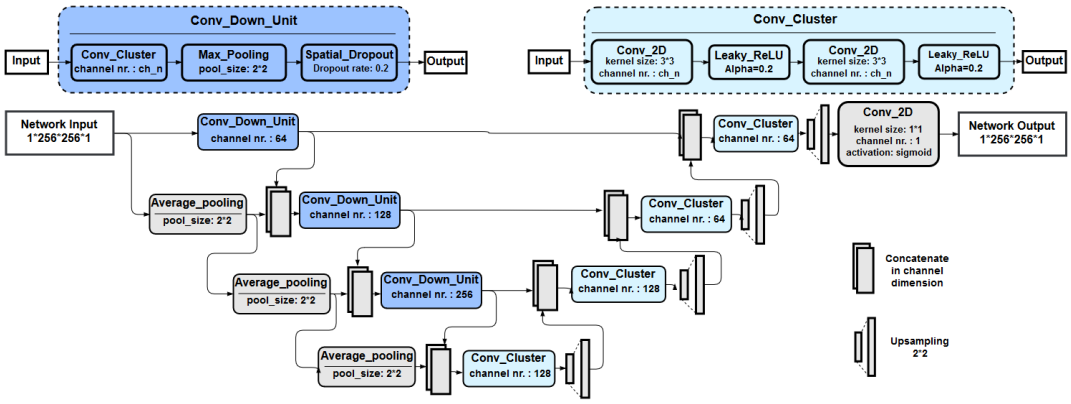

Fig. 1. The diagram of the 2D U-Net which was trained to extract information about the vessel and liver boundary masks.

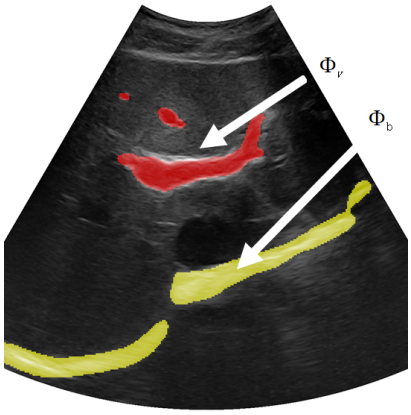

Fig. 2. Example segmentation result for a single slice US image using the trained 2D U-Net. Φ_v (red) shows the vessel segmentation, and Φ_b (yellow) shows the mask of the liver surface.

The GO compares gradient vectors on vessel borders from both US and resliced images. The intraoperative detected vessel mask and the resliced vessel mask derived from preoperative vessel segmentation are applied to filter out gradient vectors out of vessel regions. To increase the robustness, we only observe gradient vectors on the vessel outline, which show a gradient magnitude greater than the median of the sorted overall vectors. Equation 2 shows the similarity measurement of the inverted US image I_{US} and the resliced CT image I_{CT_r}, where $\nabla I_{CT\_r}(i)$ and $\nabla I_{US}(i)$ denote the gradient fields of resliced CT image and US image respectively.

Here, Φ_b and Ψ_b stand for liver boundary mask in US image and resliced CT image, respectively. To keep both masks being overlapped, a constraint $\Phi_b \cap \Psi_b > 0$ is defined, which can limit the search range and hence improve the convergence rate. The similarity of the resliced CT image and the US image is optimized using CMA-ES (see Equation 3). Based on the previous study [15] which shows that the hepatic motion secondary to respiration causes mainly liver movement along the patient Cranio-Caudal axis, the optimization parameter ${}^{CT}T_{US}$ is limited to the translations along X, Y and Z axes of the volume coordinate system.

$$\underset{{}^{CT}T_{US}}{\arg\min}(-GO(I_{US}, I_{CT\_r}({}^{CT}T_{US}))) \tag{3}$$

3. EVALUATION

To evaluate the performance of our liver registration method, 12 patients were involved in the study. To avoid the correlation between training and test dataset the evaluation has been made with leave-one-patient-out strategy. The preoperative CT images were captured with GE LightSpeed VCT (image dimension 512x512x400 voxels and voxel size 0.68x0.68x1.25 mm) after contrast enhancement. Intraoperative images were acquired using the low-end Telemed Ultrasound System with 3.5MHz abdominal probe and harmonic scan protocol. The ultrasound probe was modified with tracking markers and calibrated with a calibration phantom. In total, 99 US images were selected for the evaluation, which provides a segmented vessel-to-image area ratio greater than 1%. For each US image, the corresponding ground truth of the registration was established manually by one field expert and later verified by one clinician with more than ten years of experience. To evaluate the quality of GO similarity measurement, the search space is sampled in three dimensions

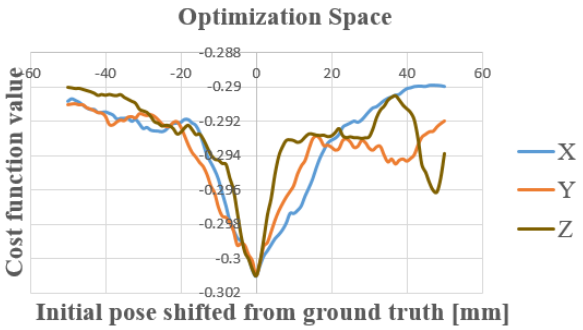

Fig. 3. Example search space. The plotted lines correspond to the translation search space in X (blue), Y (orange), and Z (brown) direction taken from one US image.

and plotted around their ground truth. As shown in Figure 3, the search space is sampled with a wide scope: translational shift up to ± 50 mm. Despite several local minima, a global minimum on the ground truth always exists. A robust optimizer like CMA-ES is chosen to lead optimization to the global minimum. Considering the run-time performance of registration, the optimizer is set with a population of 100 and maximal iteration of 100. To evaluate the robustness and accuracy of the proposed method, the registration was repeated with initial poses shifted away from their ground truth in X, Y and Z axes. A threshold of 5 mm is applied to differentiate successful and failed registration. This threshold is chosen according to the common minimum ablation safety margin [16]. As shown in Figure 4, the registration has a high convergence rate beyond 80% within slight patient motions up to 30 mm. The robustness is decreased to 70% under a large amplitude patient motion of 31 mm to 80 mm. Even though 20% - 30% of the registration fails this rate is negligible because this error is dissolved within the next frames. To evaluate

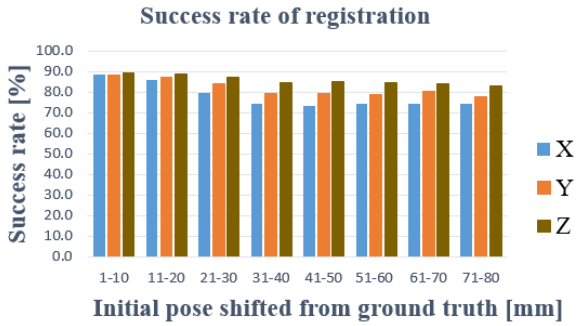

Fig. 4. Success rate of the registration. The initial pose shifted from the ground truth shows an interval of ten sample steps. Each interval was separated into X (blue), Y (orange) and Z (brown) direction.

the accuracy of 2D-3D registration, a target region is defined by clinic experts for each patient data. After registration, the point cloud in the target region is transformed into the US image coordinate system using ground truth matrix and the registration matrix, respectively. The TRE is derived by calculating the root-mean-square (RMS) residual distances of those two transformed point clouds. An example of successful registration is shown in Figure 5. The results show that the registration has a high accuracy of 1.97 ± 1.06 mm, 1.97 ± 1.07 mm and 1.97 ± 1.07 mm with initial translation along X, Y and Z axes of CT volume coordinate system respectively. The registration method is implemented with CUDA and evaluated on a laptop with Intel Core i7 7700HQ, 16 GB RAM and NVIDIA Geforce GTX 1050 (4 GB VRAM). The registration time cost is 420 ms on average. We also tested the computational effort with better hardware by upgrading the graphics card to an NVIDIA Geforce GTX 1080 Ti (11 GB VRAM), which provides six times as many cores and a higher clock rate. This results in an average computation time of 160 ms, which matches the frame rate of the US device and therefore can satisfy the requirement of speed for the needle guidance.

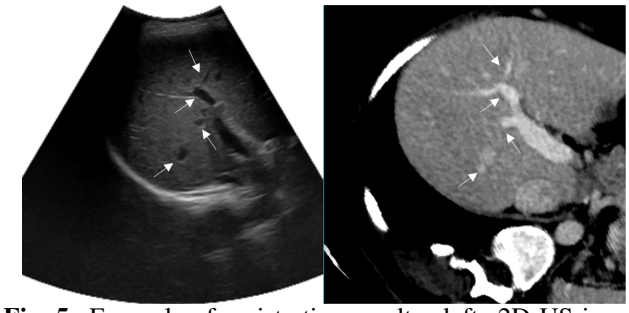

Fig. 5. Example of registration results. left: 2D US image, right: registered preoperative CT image.

4. CONCLUSION

In this work, we have presented an automatic fast registration method for motion compensation in the procedure of liver tumor ablation. In contrast to other state-of-the-art methods, we focus on improving convergence range of registration. To this end, a 2D U-Net is proposed which can provide mask of vessel and liver boundary to stabilize similarity calculation and enlarge the convergence range of registration. Evaluation has been conducted with ultrasound images captured from 12 real patients. To perform the accuracy test, 99 ultrasound images were selected, where segmented vessel-to-image area ratio is higher than 1%. The result shows that the method has robustness of >80% success rate in near range [1mm-30mm] and >70% success rate in wide range [31mm-80mm] to the ground truth. The overall accuracy of registration is 1.97 ± 1.07 mm. Furthermore, thanks to the hardware acceleration the registration for each image can be finished in under 500 ms on a low-end device and in under 200 ms on a high-end device. This preliminary result shows the possibility of fast patient motion compensation during navigation of liver tumor ablation. Despite showing the promising registration results, the performance of our method still depends on visibility and mass of vessel structures in US images. The method fails when too few vessel structures are available in captured US images. Recent research works show the power of using generative adversarial networks (GAN)[17, 18] to address registration task, where the similarity measurement can be derived from the discrimination network, and the registration might be improved by using GAN in future work.

Acknowledgement

This work has been funded by the EU and the federal state of Saxony-Anhalt, Germany under grant number ZS/2016/10/81684 and ZS/2016/04/78123 as part of the initiative Sachsen-Anhalt WISSENSCHAFT Schwerpunkte, and the Zhejiang Provincial Natural Science Foundation of China (SY19H180037).

5. REFERENCES

[1] G Mauri, L Cova, S De Beni, T Ierace, T Tondolo, A Cerri, S N Goldberg, and L Solbiati, "Real-time us-ct/mri image fusion for guidance of thermal ablation of liver tumors undetectable with us: results in 295 cases," *Cardiovascular and Interventional Radiology*, vol. 38, no. 1, pp. 143–151, 2015.

[2] L. Solbiati, T. Ierace, S. N. Goldberg, S. Sironi, T. Livraghi, R. Fiocca, G. Servadio, G. Rizzatto, P. R. Mueller, A. Del Maschio, et al., "Percutaneous us-guided radio-frequency tissue ablation of liver metastases: treatment and follow-up in 16 patients.," *Radiology*, vol. 202, no. 1, pp. 195–203, 1997.

[3] L. Crocetti, R. Lencioni, S. DeBeni, T. C. See, C.e Della Pina, and C. Bartolozzi, "Targeting liver lesions for radiofrequency ablation: an experimental feasibility study using a ct–us fusion imaging system," *Investigative Radiology*, vol. 43, no. 1, pp. 33–39, 2008.

[4] M. Peterhans, A. vom Berg, B. Dagon, D. Inderbitzin, C. Baur, D. Candinas, and S. Weber, "A navigation system for open liver surgery: design, workflow and first clinical applications," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 7, no. 1, pp. 7–16, 2011.

[5] G. P. Penney, J. M. Blackall, D. Hayashi, T. Sabharwal, A. Adam, D. J. Hawkes, et al., "Overview of an ultrasound to ct or mr registration system for use in thermal ablation of liver metastases," in *Proceedings Medical Image Understanding and Analysis*, 2001, vol. 1, p. 6568.

[6] M. Fusaglia, P. Tinguely, V. Banz, S. Weber, and H. Lu, "A novel ultrasound-based registration for image-guided laparoscopic liver ablation," *Surgical Innovation*, vol. 23, no. 4, pp. 397–406, 2016.

[7] T Lange, N Papenberg, S Heldmann, J Modersitzki, B Fischer, H Lamecker, and P M Schlag, "3d ultrasound-ct registration of the liver using combined landmark-intensity information," *International journal of computer assisted radiology and surgery*, vol. 4, no. 1, pp. 79–88, 2009.

[8] Haque H., Omi Y., Rusk L.and Annangi P., and Kazuyuki O., "Automated registration of 3d ultrasound and ct/mr images for liver," in *2016 IEEE International Ultrasonics Symposium (IUS)*, Sep. 2016, pp. 1–4.

[9] J Banerjee, C Klink, E D Peters, W J Niessen, A Moelker, and T van Walsum, "Fast and robust 3d ultrasound registration–block and game theoretic matching," *Medical Image Analysis*, vol. 20, no. 1, pp. 173–183, 2015.

[10] W. Wein, S. Brunke, A. Khamene, M. R. Callstrom, and N. Navab, "Automatic ct-ultrasound registration for diagnostic imaging and image-guided intervention," *Medical Image Analysis*, vol. 12, no. 5, pp. 577–585, 2008.

[11] L. Xu, J. Liu, W. Zhan, and L. Gu, "A novel algorithm for ct-ultrasound registration," in *Point-of-Care Healthcare Technologies, 2013 IEEE*, 2013, pp. 101–104.

[12] CJ Weon, W Hyun Nam, D Lee, JY Lee, and JB Ra, "Position tracking of moving liver lesion based on real-time registration between 2d ultrasound and 3d preoperative images," *Medical physics*, vol. 42, no. 1, pp. 335–347, 2015.

[13] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-assisted Intervention*. Springer, 2015, pp. 234–241.

[14] N. Hansen, "The cma evolution strategy: a comparing review," in *Towards a new Evolutionary Computation*, pp. 75–102. Springer, 2006.

[15] MA Clifford, F Banovac, E Levy, and K Cleary, "Assessment of hepatic motion secondary to respiration for computer assisted interventions," *Computer Aided Surgery: Official Journal of the International Society for Computer Aided Surgery (ISCAS)*, vol. 7, no. 5, pp. 291–299, 2002.

[16] J. Iwazawa, S. Ohue, N. Hashimoto, and T. Mitani, "Ablation margin assessment of liver tumors with intravenous contrast-enhanced c-arm computed tomography," *World Journal of Radiology*, vol. 4, no. 3, pp. 109, 2012.

[17] J Fan, X Cao, Z Xue, P Yap, and DG Shen, "Adversarial similarity network for evaluating image alignment in deep learning based registration," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 739–746.

[18] I Goodfellow, J Pouget-Abadie, M Mirza, B Xu, D Warde-Farley, S Ozair, A Courville, and Y Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, 2014, pp. 2672–2680.