A Deep Learning-based Interactive Medical Image Segmentation Framework

I. Mikhailov[1,2], B. Chauveau[2,3], N. Bourdel[2,3], and A. Bartoli[1,2]

[1] EnCoV, Institut Pascal, Université Clermont Auvergne, Clermont-Ferrand, France
[2] SurgAR, Clermont-Ferrand, France
[3] CHU de Clermont-Ferrand, France

Abstract. Image segmentation is an essential component in medical image analysis. The case of 3D images such as MRI is particularly challenging and time consuming. Interactive or semi-automatic methods are thus highly desirable. While deep learning outperforms classical methods in automatic segmentation, its use in interactive frameworks is still limited. The main reason is that most neural networks do not lend themselves well to the required user interaction loop. We propose a general deep learning-based interactive framework for image segmentation, which embeds a base network in a user interaction loop with a user feedback memory. We propose to model the memory explicitly as a sequence of consecutive framework states, from which the features can be learned. A major difficulty is related to training, as the network inputs include the user feedback and thus depend on the network's previous output. We propose to introduce a virtual user in the training process, modelled by simulating the user feedback from the current segmentation. We demonstrate our framework on the task of female pelvis MRI segmentation, using a new dataset. We evaluate our framework against existing work with the standard metrics and conduct a user evaluation. Our framework outperforms existing systems.

Keywords: Interactive segmentation · Deep Learning · MRI · RNN.

1 Introduction

Image segmentation is an essential component of many visual processing systems, which involves classifying each pixel or, equivalently, delineating the regions containing pixels of the same class. In medical image analysis, the images are often patient scans from modalities such as MRI (Magnetic Resonance Imaging) or CT (Computed Tomography). MRI segmentation is a tremendously difficult task, owing to it being 3D, low contrast, noisy, low resolution and artifacted. Existing segmentation approaches can be divided into three settings based on user involvement: manual, automatic and interactive. The manual approach is the most time-consuming, as each pixel has to be attributed a label independently, which may require hours for a single MRI. It is error-prone and infeasible in the clinical environment. At the other extreme lies the automatic approach,

which works without user involvement. This strongly limits its applicability, as a clinician operator shall validate and possibly edit the result before its use in a therapeutic act. The interactive approach trades-off manual and automatic features: it typically involves an automatic part with an extent of user control. The interactive approach attempts at minimising the amount of required user interaction whilst maximising the quality of the result. It is the most adapted approach to the clinical environment.

The automatic approach is largely dominated by deep learning, which overturned classical methods over the last decade. In contrast, interactive deep learning methods present specific difficulties and have yet received limited attention. Concretely, deep learning interactive segmentation requires embedding a network in a loopy system allowing the user to interact. Indeed, the network inputs must include the user feedback, which depends on the network outputs. This creates a dependency between the inputs and outputs of the network, which is poorly resolved by a regular training process from static data. Creating such interactive systems exploiting deep learning is nonetheless fundamental to simplify, speed up and secure the performance of segmentation in the clinical environment.

We propose a general multi-class deep learning interactive segmentation framework and training methods. Our system consists of an embedded network, a user interaction loop and an interaction memory. First, the user reviews the current segmentation result and, if satisfied, accepts. Otherwise, the user may quickly make simple corrections by placing points and strokes to refine the segmentation, which is achieved by a special input configuration of the embedded network. Indeed, this network inputs the image, user correction masks, and possibly other memorised parameters, and outputs the segmentation probability maps. The system then loops back to the user review step, whilst updating the interaction memory to keep track of the user corrections through the interactions. Our

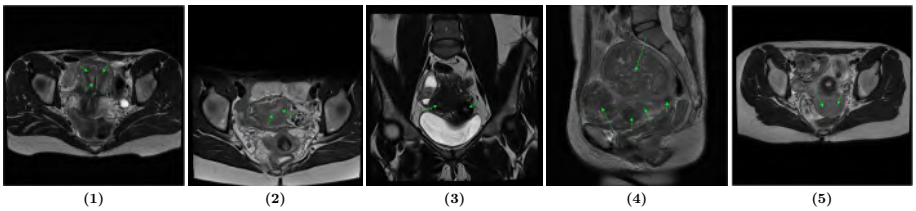

Fig. 1. MRI female pelvis dataset samples, main difficulties indicated with green arrows, series 1 to 5: **(1)** presence of an IUD, not seen in the training set; **(2,5)** unclear contours, blurriness of the uterine cavity; **(3)** similarity of the uterine (left) and cervix cavities (right); **(4)** strong uterus deformation due to tumours, with here five tumours.

contributions are two-fold. First, a general deep learning-based interactive multi-class image segmentation framework, with a user interaction loop, dynamic data training process and an interaction memory. Second, a sequential interaction memory, which keeps track of the segmentation results and user corrections,

maintaining sequentiality within the system. We demonstrate our framework in semantic multi-class MRI segmentation of the female pelvis for which we introduce a new dataset. For this task, we instantiate our system with an existing encoder-decoder architecture optionally featuring RNN [38] modules. We validate the results against automatic and existing interactive systems with the standard metrics and perform an ablation study of our system's components. We report results of a preliminary user study conducted with medical users, using a specifically developed graphical user interface connected to our system.

2 Related Work

We review classical and deep learning approaches to medical image segmentation, distinguishing automatic and interactive approaches for each.

Classical automatic segmentation encompasses a wide variety of methods [30]. Their performance are usually insufficient to achieve clinically-acceptable accuracy and they have been largely taken over by deep learning. In contrast, classical interactive methods are still widely used. The most well known ones are probably the Graph Cuts [2], Random Walker [9] and Geodesic Image Segmentation (GeoS) [4]. They achieve acceptable performance for simple cases. However, medical data often features structures with complex shapes and poorly defined contours, noise and artefacts. This results in a substantial increase of user time required to perform segmentation and limited achievable accuracy. Deep learning-based automatic segmentation includes a multitude of methods.

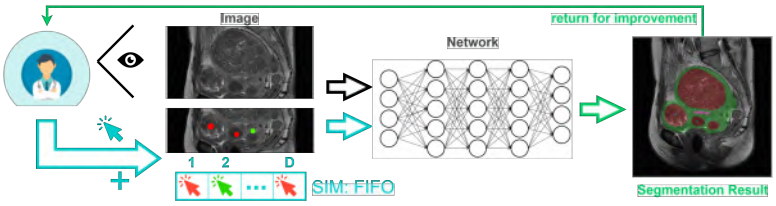

Fig. 2. Proposed interactive system.

A review and evaluation of over 100 methods [19] was conducted with ResNet [12] extensively used as a backbone, represented by EMANet [16] with top scores on the PASCAL VOC dataset. Most of the models use an encoder-decoder architecture [19]. This includes the U-Net [20], with a wide spectrum of applications [23], and recent variants [7, 24] reaching top positions in the BraTS challenge 2021. Automatic MRI segmentation was attempted for various targets, including the kidney [14], the prostate [10] and brain tumours [11]. These methods demonstrate state-of-the-art performance in their respective tasks. However, they are automatic and do not allow the user to interact. Automatic segmentation is highly appropriate in applications which can not involve user interactions, such

as real-time organ tracking. In contrast, many applications require validation and corrections from a certified user. For such applications, the direct use of automatic deep learning methods is inappropriate.

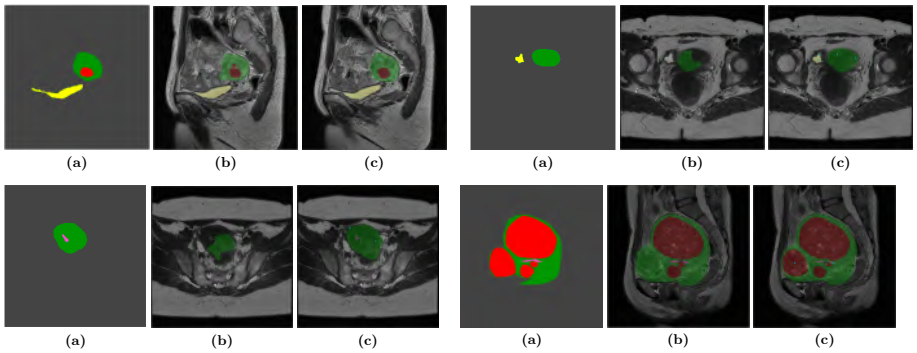

Fig. 3. Segmentation results, where uterus - green, bladder - yellow, tumour - red, cavity - pink and user clicks - cyan: **(a)** ground truth; **(b)** auto; **(c)** human user-controlled DDG-SIM.

The integration of deep learning within interactive segmentation systems is a major challenge. A simple approach is to use a classical interactive method to post-process the result from an automatic deep learning method [26] or correct it manually [22]. Such systems inherit the intrinsic limitations of the chosen classical method. A more advanced approach is to use a neural network to process user feedback in a loopy structure [1, 27–29, 17, 21]. These methods use a network which takes the image and user interaction masks as inputs. Training is challenging owing to the loop. Existing approaches generate user interaction masks from labelled data, either statically before training or dynamically during training. Static data training methods [27–29] limit the system's generalisation and interaction effectiveness. Dynamic data training methods [1, 17] improve performance. They simulate user interaction by randomly sampling mis-segmented regions. This is done once from a single prediction [1] or from the latest segmentation result [17]. These methods diversify the training data, but do not reproduce the typical sequentiality of real user interactions. The lack of sequentiality is also a consequence of the interaction memory used in these systems, which simply accumulates the user corrections, discarding ordering. A sequential memory was used in [29] to 'transfer' the user interaction recorded on one slice to the other slices, but was not used to exploit sequentiality during slice segmentation. An open-source interactive segmentation platform [6] was recently made available, which offers both deep learning-based [26, 21] and classical methods [2], inheriting their limitations.

In contrast to existing work, our framework uses a sequential interaction memory which captures the sequentiality of user interactions at training and inference times. It is also the first framework to cope with multiple classes.

Table 1. Experimental evaluation results where bold means best.

Method	BGD		Uterus		Bladder		Tumours		Cavity	
	IoU	Dice	IoU	Dice	IoU	Dice	IoU	Dice	IoU	Dice
Auto	99.2	99.6	64.7	78.6	71.9	83.6	60.4	75.3	40.4	57.6
SDG-base	99.1	99.6	61.7	76.3	70.1	82.4	62.5	76.9	21.1	34.9
SDG-CIM	99.3	99.7	66.5	79.9	83.9	91.2	72.8	84.3	29.0	44.9
DDG-CIM	**99.6**	**99.8**	77.4	87.3	**87.4**	**93.3**	77.7	87.4	39.6	56.7
DDG-SIM	**99.6**	**99.8**	**79.8**	**88.7**	87.0	93.0	**79.0**	**88.3**	**57.8**	**73.3**

3 Applicative Scope

While our framework may be applied to numerous segmentation problems, we focus on the interactive female pelvis MRI segmentation, involving five classes: uterus, bladder, uterine cavity, tumours and background. The intended use is surgical planning and surgical augmented reality [3]. We create public female pelvis MRI dataset, consisting of 97 MRI series with 3066 slices in total, manually annotated in 3D Slicer [13] and in MITK [8] by expert radiologists. This took from 10' to 50' per series with 25' on average, which is clearly infeasible in the clinical setting. The segmentation of anatomical structures of the female pelvis is particularly challenging due to a large variance in their representation, including shape, size, position, orientation and texture among the patients, with and without pathologies. Difficult samples can be seen in figure 1. Moreover, the target anatomical structures form a naturally imbalanced dataset, where background takes 96.15%, uterus 2.11%, bladder 1.02%, tumours 0.67% and uterine cavity 0.05%. The strongest imbalance is observed for the uterine cavity and the background, whose average ratio of volumes is 0.057%. Our objective is to develop a segmentation system which minimises the time required to complete the segmentation with acceptable accuracy, while allowing an expert reviewer to have control and guide the segmentation, as and when necessary.

4 Methodology

4.1 System

Structure. We build the proposed system shown in figure 2 starting with a basic interactive segmentation system named base, featuring an interaction loop. This system does not have a memory of user corrections or previous segmentation results and processes each set of user corrections in isolation. The interaction loop allows iterative refinement by forming new inputs through a combination of network outputs and user corrections. The system is generic as it does not depend on a specific network architecture, as long as the network takes both the image and the user corrections as inputs. The user corrections are represented by N binary masks, where N is the number of classes. The network inputs are concatenated to a single tensor of size $H \times W \times C$, where $H \times W$ is the image size and C is the number of channels, varying depending on the system. For the

base system $C_{\text{base}} = 1 + N$. Indeed, as there is no memory, the network takes the image as the first channel and the binary masks of the user corrections for the N classes for next N channels. This strongly harms user experience as the past user corrections would be forgotten by the system at the next interaction [27, 26].

Sequential Interaction Memory. Existing works use the type of interaction memory, which aggregates the raw system states by merging the successive interaction masks [1, 28, 17]. We call this a cumulative interaction memory (CIM). The network takes the image and the merged user correction masks, and its input tensor thus has $C_{\text{cim}} = C_{\text{base}} = 1 + N$ channels. This type of memory discards the ordering of interactions - the sequentiality, typical of user corrections. We introduce a second type of interaction memory which, in contrast to CIM, preserves the past D system states, hence the user's sequential behaviour. We call this a sequential interaction memory (SIM). We call the number of states D the SIM's size or depth. For the task of multi-class segmentation, a single state consists of a probability map for the network outputs and a binary mask for the user corrections, for each of the N classes. The network takes an image and the SIM, which are automatically combined to form the input tensor with $C_{\text{sim}} = 1+2DN$. It is important to make a distinction between the proposed SIM and internal RNN memory. SIM tracks and stores system states, represented by inputs and outputs of the network. Indeed, SIM is external to the network and does not depend on a specific network architecture. The RNN memory, however, is specific to the network architecture, enabled by passing hidden state from step to step and represented by weights. In our ablation study we show that RNN's suitability for sequential data may further reinforce the proposed framework.

4.2 Training with Dynamic Data Generation

A regular training process from static data will poorly reproduce the real system usage at test time and hence limit the achievable accuracy and user interaction efficiency. To resolve this, we propose a dynamic training approach, where the training data is generated from the labelled dataset during training by a virtual user. The basic idea of the virtual user is to generate corrections similarly to a real user, whose involvement in training is not feasible. These corrections are represented by one binary mask per class, populated by foreground clicks for each class. The click is handled by an interaction-control process, which exploits the difference image between the previous network output and the ground truth. Training with the proposed SIM means filling its D states with realistic values produced by the virtual user. We thus run the system for D iterations with fixed weights to populate the SIM with simulated user input data prior to backpropagation. We choose D experimentally with the goal of maximizing the performance with the minimum number of interactions.

Table 2. User evaluation results given as Time, IoU and number of interactions per class as 'Int.' ;'A' - gynecologic surgeon; 'B', 'C' - certified radiologists. Tumour presence is in proportion to that of the whole dataset.

Series	Time	BGD		Uterus		Bladder		Tumours		Cavity		mIoU	Total Int.
		IoU	Int.	IoU	Int.	IoU	Int.	IoU	Int.	IoU	Int.		
1-A	1'44"	99.7	-	69.5	23	92.0	5	-	-	44.0	-	76.3	28
1-B	3'00"	99.6	-	64.1	12	93.6	6	-	-	41.4	2	74.7	20
1-C	3'19"	99.6	8	67.6	30	93.5	6	-	-	37.7	9	74.6	53
2-A	3'10"	99.3	-	67.0	21	79.4	9	71.7	-	35.0	10	70.5	40
2-B	2'42"	99.3	-	69.9	20	78.4	7	71.7	-	42.6	1	72.4	28
2-C	5'04"	98.2	11	51.9	29	52.3	12	71.8	-	46.6	13	64.1	65
3-A	3'10"	99.6	-	70.3	7	76.1	15	-	-	38.5	6	71.1	28
3-B	2'50"	99.6	1	72.2	5	76.8	14	-	-	42.6	5	72.8	25
3-C	4'13"	99.6	6	70.7	12	77.8	11	-	-	21.8	9	67.4	38
4-A	7'08"	98.1	-	63.0	3	71.9	6	81.0	137	44.0	27	71.6	173
4-B	8'30"	98.2	8	66.8	21	76.6	9	58.6	27	41.9	-	68.4	65
4-C	8'47"	98.0	13	53.0	-	80.0	11	72.7	96	34.7	-	67.6	120
5-A	2'38"	99.8	-	61.0	17	93.2	14	-	-	24.9	1	69.7	32
5-B	4'17"	99.9	11	68.0	5	93.5	11	-	-	30.6	6	73.0	33
5-C	3'51"	99.8	13	66.9	29	93.2	12	-	-	29.3	13	72.3	67

5 Experimental Results

5.1 Setup

We instantiate our system with an existing encoder-decoder architecture featuring RNN modules. Namely, we use a ResNet34 [12] encoder pre-trained on ImageNet [5] and a decoder equipped with a pair of standard convolutional layer and a matching convolutional LSTM layer at every step of the upsampling path. To counter the dataset imbalance, we use the focal loss [18] and dataset-wide precalculated per-class weights. The dataset split is as follows: training set - 77 series/2449 slices, validation set - 10 series/308 slices and test set - 10 series/309 slices. We preprocessed all data via normalisation, standardisation and N4BFC [25] and performed random data augmentation: vertical and horizontal flipping, intensity shifting, gamma correction, blurring and unsharp masking.

5.2 Automated Evaluation

We compare one automatic method and four interactive methods, where SDG is Static Data Generation and DDG is Dynamic Data Generation: 1) Auto: U-Net with ResNet34 encoder [15]; 2) SDG-base: memory-less system trained with SDG, as described in [1]; 3) SDG-CIM: network from SDG-base used with a CIM overlay; 4) DDG-CIM: system with CIM trained with DDG; 5) DDG-SIM: complete proposed system with SIM trained with DDG. The evaluation setup is a ResNet34 encoder with (1-4) a generic decoder or (5) an LSTM-decoder as in section 5.1. At test time, clicks are generated via the virtual user. The

metrics are reported in table 1. We observe that DDG-SIM outperforms the other methods with a minor disadvantage for bladder, for which DDG-CIM is slightly better with 87.4% against 87.0% IoU. The ablation study shows a steady increase in performance, starting with SDG-base and adding the proposed components towards DDG-SIM. Auto outperforms both SDG-base on uterus, bladder and cavity, and SDG-CIM on cavity. This can be attributed to static data generation, which does not perform well for smaller numbers of interactions. In our experience, the higher the number of interactions at training, the lower the effectiveness of individual interactions at test time. While the opposite is also true, it can be observed from the results that certain systems may not be able to learn efficiently from a small number of interactions at training. We observe a comparatively lower accuracy for cavity, whose IoU lies between 21.1% and 57.8%. We explain this with its low volume, which accounts for only 0.054% of the dataset.

5.3 User Evaluation

We performed a preliminary user study with DDG-SIM involving three senior medical experts, using a specifically developed graphical user interface. We randomly selected 6 test series containing 144 slices in total, where 1 series is used to familiarise the users with the graphical user interface and 5 series are used in a random order for user evaluation. MRI image samples from each of the series can be seen in figure 1. We evaluate the user performance in table 2 using elapsed time, IoU and the number of interactions employed per class. The segmentation result is compared with Auto method in figure 3.

We note that the time is low enough to be clinically feasible, even if the users are barely acquainted with the system. Indeed, the average elapsed time for all series is 4'18", which is largely below the reported average of 25' for existing systems. Series 4 was a complex case with 11 tumours and a heavy deformation of the uterus shape, taking 8'08" on average for our system and 40' for existing systems.

6 Conclusion

We have proposed a general deep learning-based interactive multi-class image segmentation framework, with a user interaction loop and a sequential interaction memory. We have demonstrated our framework in female pelvis MRI segmentation, using a new dataset. We have evaluated our framework against existing work with the standard metrics and conducted a user evaluation. This shows that our framework largely outperforms existing systems in accuracy and drastically reduces the average user segmentation time from 25' to 4'18".

We plan to further improve the proposed solution towards its clinical usage. First, through application to other segmentation tasks and expansion of the user study. Second, by using it to aid annotation, reducing interaction demand through SIM initialization with automatic segmentation. Third, by directly applying it to 3D images, to further shorten the segmentation time.

References

1. Amrehn, M., Gaube, S., Unberath, M., Schebesch, F., Horz, T., Strumia, M., Steidl, S., Kowarschik, M., Maier, A.: UI-Net: Interactive Artificial Neural Networks for Iterative Image Segmentation Based on a User Model. In: Eurographics Workshop on Visual Computing for Biology and Medicine. The Eurographics Association (2017)
2. Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary amp; region segmentation of objects in n-d images. In: Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001. vol. 1, pp. 105–112 vol.1 (2001)
3. Collins, T., Pizarro, D., Gasparini, S., Bourdel, N., Chauvet, P., Canis, M., Calvet, L., Bartoli, A.: Augmented reality guided laparoscopic surgery of the uterus. IEEE Transactions on Medical Imaging **40**(1), 371–380 (2021)
4. Criminisi, A., Sharp, T., Blake, A.: Geos: Geodesic image segmentation. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) Computer Vision – ECCV 2008. pp. 99–112. Springer Berlin Heidelberg, Berlin, Heidelberg (2008)
5. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. pp. 248–255 (2009)
6. Diaz-Pinto, A., Alle, S., Ihsani, A., Asad, M.H., Nath, V., P'erez-Garc'ia, F., Mehta, P., Li, W., Roth, H.R., Vercauteren, T.K.M., Xu, D., Dogra, P., Ourselin, S., Feng, A., Cardoso, M.J.: Monai label: A framework for ai-assisted interactive labeling of 3d medical images. ArXiv **abs/2203.12362** (2022)
7. Futrega, M., Milesi, A., Marcinkiewicz, M., Ribalta, P.: Optimized u-net for brain tumor segmentation. ArXiv **abs/2110.03352** (2021)
8. Goch, C.J., Metzger, J., Nolden, M.: Abstract: Medical research data management using mitk and xnat. In: Maier-Hein, geb. Fritzsche, K.H., Deserno, geb. Lehmann, T.M., Handels, H., Tolxdorff, T. (eds.) Bildverarbeitung für die Medizin 2017. pp. 305–305. Springer Berlin Heidelberg, Berlin, Heidelberg (2017)
9. Grady, L.: Random walks for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence **28**(11), 1768–1783 (2006)
10. Guo, Y., Gao, Y., Shen, D.: Deformable mr prostate segmentation via deep feature learning and sparse patch matching. IEEE Transactions on Medical Imaging **35**(4), 1077–1089 (2016)
11. Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.M., Larochelle, H.: Brain tumor segmentation with deep neural networks. Medical Image Analysis **35**, 18–31 (2017)
12. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) pp. 770–778 (2016)
13. Kikinis, R., Pieper, S.D., Vosburgh, K.G.: 3D Slicer: A Platform for Subject-Specific Image Analysis, Visualization, and Clinical Support, pp. 277–289. Springer New York, New York, NY (2014)
14. Kline, T.L., Korfiatis, P., Edwards, M.E., Blais, J.D., Czerwiec, F.S., Harris, P.C., King, B.F., Torres, V.E., Erickson, B.J.: Performance of an artificial multi-observer deep neural network for fully automated segmentation of polycystic kidneys. Journal of Digital Imaging **30**(4), 442–448 (Aug 2017)
15. Le'Clerc Arrastia, J., Heilenkötter, N., Otero Baguer, D., Hauberg-Lotte, L., Boskamp, T., Hetzer, S., Duschner, N., Schaller, J., Maass, P.: Deeply Supervised UNet for Semantic Segmentation to Assist Dermatopathological Assessment of Basal Cell Carcinoma. J Imaging **7**(4) (Apr 2021)

16. Li, X., Zhong, Z., Wu, J., Yang, Y., Lin, Z., Liu, H.: Expectation-maximization attention networks for semantic segmentation. 2019 IEEE/CVF International Conference on Computer Vision (ICCV) pp. 9166–9175 (2019)
17. Liao, X., Li, W., Xu, Q., Wang, X., Jin, B., Zhang, X., Zhang, Y., Wang, Y.: Iteratively-refined interactive 3d medical image segmentation with multi-agent reinforcement learning. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) pp. 9391–9399 (2020)
18. Lin, T.Y., Goyal, P., Girshick, R.B., He, K., Dollár, P.: Focal loss for dense object detection. 2017 IEEE International Conference on Computer Vision (ICCV) pp. 2999–3007 (2017)
19. Minaee, S., Boykov, Y.Y., Porikli, F., Plaza, A.J., Kehtarnavaz, N., Terzopoulos, D.: Image segmentation using deep learning: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence pp. 1–1 (2021)
20. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. pp. 234–241. Springer International Publishing, Cham (2015)
21. Sakinis, T., Milletari, F., Roth, H.R., Korfiatis, P., Kostandy, P.M., Philbrick, K.A., Akkus, Z., Xu, Z., Xu, D., Erickson, B.J.: Interactive segmentation of medical images through fully convolutional neural networks. ArXiv **abs/1903.08205** (2019)
22. Shan, F., Gao, Y., Wang, J., Shi, W., Shi, N., Han, M., Xue, Z., Shen, D., Shi, Y.: Lung infection quantification of covid-19 in ct images with deep learning. ArXiv (2020)
23. Siddique, N., Paheding, S., Elkin, C.P., Devabhaktuni, V.: U-net and its variants for medical image segmentation: A review of theory and applications. IEEE Access **9**, 82031–82057 (2021)
24. Siddiquee, M.M.R., Myronenko, A.: Redundancy reduction in semantic segmentation of 3d brain tumor mris. ArXiv **abs/2111.00742** (2021)
25. Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C.: N4itk: Improved n3 bias correction. IEEE Transactions on Medical Imaging **29**(6), 1310–1320 (2010)
26. Wang, G., Li, W., Zuluaga, M.A., Pratt, R., Patel, P.A., Aertsen, M., Doel, T., David, A.L., Deprest, J.A., Ourselin, S., Vercauteren, T.K.M.: Interactive medical image segmentation using deep learning with image-specific fine tuning. IEEE Transactions on Medical Imaging **37**, 1562–1573 (2018)
27. Wang, G., Zuluaga, M.A., Li, W., Pratt, R., Patel, P.A., Aertsen, M., Doel, T., David, A.L., Deprest, J.A., Ourselin, S., Vercauteren, T.K.M.: Deepigeos: A deep interactive geodesic framework for medical image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence **41**, 1559–1572 (2019)
28. Zhou, B., Chen, L., Wang, Z.: Interactive deep editing framework for medical image segmentation. In: Shen, D., Liu, T., Peters, T.M., Staib, L.H., Essert, C., Zhou, S., Yap, P.T., Khan, A. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. pp. 329–337. Springer International Publishing, Cham (2019)
29. Zhou, T., Li, L., Bredell, G., Li, J., Konukoglu, E.: Quality-aware memory network for interactive volumetric image segmentation. In: MICCAI (2021)
30. Zhu, H., Meng, F., Cai, J., Lu, S.: Beyond pixels: A comprehensive survey from bottom-up to semantic image segmentation and cosegmentation. Journal of Visual Communication and Image Representation **34**, 12–27 (2016)