

PSEUDO-LABEL REFINEMENT USING SUPERPIXELS FOR SEMI-SUPERVISED BRAIN TUMOUR SEGMENTATION

Bethany H. Thompson ^{†*} Gaetano Di Caterina ^{*} Jeremy P. Voisey [†]

[†] Canon Medical Research Europe Ltd., AI Research, Edinburgh, UK

^{*} University of Strathclyde, Electronic & Electrical Engineering, Glasgow, UK

ABSTRACT

Training neural networks using limited annotations is an important problem in the medical domain. Deep Neural Networks (DNNs) typically require large, annotated datasets to achieve acceptable performance which, in the medical domain, are especially difficult to obtain as they require significant time from expert radiologists. Semi-supervised learning aims to overcome this problem by learning segmentations with very little annotated data, whilst exploiting large amounts of unlabelled data. However, the best-known technique, which utilises inferred pseudo-labels, is vulnerable to inaccurate pseudo-labels degrading the performance. We propose a framework based on superpixels - meaningful clusters of adjacent pixels - to improve the accuracy of the pseudo labels and address this issue. Our framework combines superpixels with semi-supervised learning, refining the pseudo-labels during training using the features and edges of the superpixel maps. This method is evaluated on a multimodal magnetic resonance imaging (MRI) dataset for the task of brain tumour region segmentation. Our method demonstrates improved performance over the standard semi-supervised pseudo-labelling baseline when there is a reduced annotator burden and only 5 annotated patients are available. We report DSC=0.824 and DSC=0.707 for the test set whole tumour and tumour core regions respectively.

Index Terms— Segmentation, Tumour, Brain, Superpixels, Semi-supervised.

1. INTRODUCTION

In medical imaging, segmentation of pathology helps clinicians to diagnose the severity of disease, make a recommendation for treatment and monitor the response to treatment over time [1, 2]. However, training accurate medical imaging models requires vast amounts of expertly annotated data which can be challenging, time-consuming and prone to error and user bias, with annotation taking approximately 1 hour per patient for the 2020 Brain Tumour Segmentation dataset (BraTS 2020) [3]. Developing automatic segmentation tools trained with limited labelled data is a challenging task, especially in pathology segmentation where the object varies notably between patients and there is low contrast between the object and the background.

Semi-supervised learning, which utilises both labelled and unlabelled data, is a popular way to take advantage of large amounts of unlabelled data and maintain a low annotator burden. Existing pseudo-label-based methods obtain pseudo-labels by first training a segmentation model with limited labelled data and then inferring pseudo-labels for the unlabelled data before subsequent training on both [4, 5]. However, these methods may generate inaccurate

pseudo-labels and degrade the subsequent training process and can be sensitive to the mask threshold with pseudo-labels being either over or under confident [6, 7].

To overcome this problem, some previous works have suggested combining active contours, a classic unsupervised segmentation technique, with semi-supervised deep learning to refine the pseudo-labels inferred during semi-supervised training [8]. The general idea is that the pseudo-labels provide the seed for region growing before using the updated pseudo labels for subsequent semi-supervised training. However, traditional boundary-based methods which may be used for region growing (e.g. active contours/snakes [9]) can be time consuming, with algorithms taking many iterations before convergence [10], especially in 3D. To overcome this problem, region growing on the superpixel level rather than on the pixel level is proposed.

Superpixels are the result of iterative grouping of pixels into an irregular grid of regions that mould well to salient features and edges within the image [11]. The most popular algorithm for generating superpixels is the Simple Linear Iterative Clustering (SLIC) algorithm [11], which adapts a k-means clustering approach to generate superpixels. Although originally developed for natural images, SLIC has been shown to be effective in MRI [12], CT [13] and Ultrasound [14] and is therefore able to detect the, sometimes subtle, region boundaries found in medical images. Previous work done by Borovec et al. (2017) [15] developed a superpixel region growing algorithm with a learned shape prior for segmenting individual eggs in microscopy images of *Drosophila* ovaries. However, in the case of pathology, such as brain tumour segmentation, a circular shape prior is not suitable as there is a large variance in topology and extent of disease between patients. Chaibou et al. (2018) [16] developed a strategy for unsupervised segmentation of natural images based on superpixel clustering. Superpixels to be merged are assumed to satisfy two important criteria: spatial adjacency and perceptual similarity. This suggests that an efficient approach should be able to pick the most visually similar spatial neighbour of the region of interest. They argue that taking advantage of the initial clustering performed by SLIC before further clustering of those superpixels into object regions allows more efficient semantic feature extraction compared to their pixel-level counterparts.

Another related work was presented by Wang et al. (2021) [17], where a filtering mechanism was developed to remove the pseudo-labels with the lowest confidence. The work presented in our paper takes a different approach and refines the predicted pseudo-labels, making them more robust rather than reducing the dataset size through filtering.

This paper presents a modified version of the pseudo label-based semi-supervised framework combined with a superpixel-based region growing algorithm - inspired by the work of Chaibou et al.

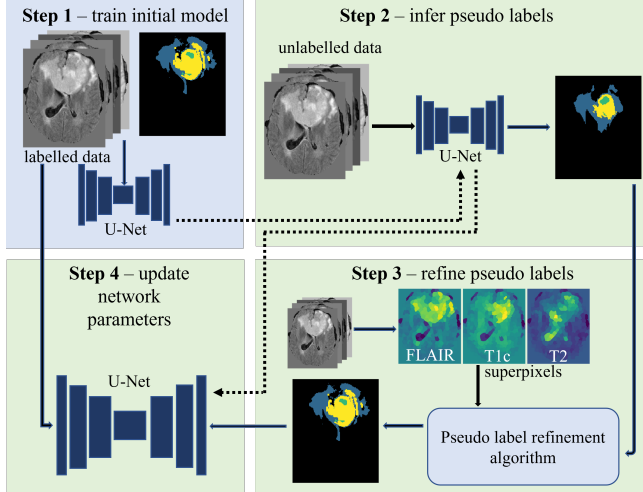


Fig. 1: Semi-supervised pseudo-labelling segmentation pipeline with the proposed superpixel-based pseudo-label refinement step

(2018) - to refine the network predicted pseudo labels and aid subsequent semi-supervised training. The proposed method is demonstrated for the task of brain tumour segmentation in multimodal MRI. Fig. 1 shows the proposed 4-step pipeline which incorporates superpixel region refinement after the pseudo-label update step.

The contributions of this work are summarised as follows:

1. A simple 3D in-training algorithm to refine network-inferred pseudo-labels using superpixel information
2. A similarity measure for sub-region refinement based on comparison of different MRI channels (Fig. 3) which mimics how expert annotators determine region boundaries
3. Demonstration of pseudo-label refinement for brain tumour segmentation which have varying topology and low contrast

2. METHOD

2.1. Semi-supervised with Pseudo-labels

We implement a modified version of the original pseudo-label method proposed by Lee (2013) [4] as follows:

1. The network is trained on p_1 ground truth (GT) labelled patients for 200 epochs
2. This model is then used to infer the pseudo labels for the p_2 unlabelled patients
3. The network is then trained in a supervised manner on both GT labelled patients and pseudo-labelled patients for the remaining 800 epochs

For unlabelled data, pseudo-labels are only re-calculated every 200 epochs to avoid long training times. As training progresses, more pseudo-labelled batches are introduced into the training, according to a dynamic weighting factor $\alpha(e)$ as described by (1) [4]. The number of pseudo-labelled batches per epoch as training progresses is described by (2), where N_e is the number of pseudo-labelled batches in epoch e and N_T is the total number of batches in an epoch.

$$\alpha(e) = \begin{cases} 0 & e < T_1 \\ \frac{e-T_1}{T_2-T_1} \alpha_f & T_1 \leq e < T_2 \\ \alpha_f & e \geq T_2 \end{cases} \quad (1)$$

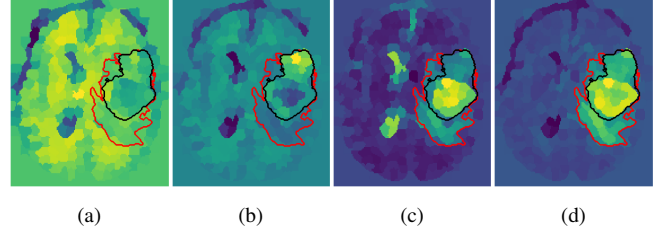


Fig. 2: Superpixel maps with $\sigma=1$, $\text{compactness}=0.01$, $n_{\text{segments}} = 350$ for patient 224. Red contour: WT GT boundary, Black contour: TC GT boundary (a) T1, (b) T1Gd, (c) T2 and (d) T2-FLAIR

$$N_e = N_T \frac{\alpha(e)}{\alpha(e) + 1} \quad (2)$$

2.2. Superpixels

Using superpixels allows feature statistics to be measured on a naturally adaptive domain rather than on a fixed window. Since superpixels tend to preserve boundaries, there is an opportunity to create an accurate segmentation by simply finding the superpixels which are part of the tumour region.

The superpixel masks were generated for each unlabelled training volume using the SLIC algorithm [11]. Fig. 2 shows the superpixel maps with $\sigma=1$, $\text{compactness}=0.01$, $n_{\text{segments}} = 350$ for the four different MRI modalities used in this work. The overlaid label boundaries for the two regions of interest - Whole tumour (WT) and Tumour Core (TC) - are shown in red and black respectively. The annotator protocol used for the original BraTS dataset [3, 18, 19] states that different regions are more easily distinguishable by using different MRI modalities as shown in both Fig. 2 and Fig. 3. To mimic this, to refine the WT region, a 3D superpixel map generated from the T2-FLAIR scan should be used. Similarly, if interested in refining pseudo-labels for the tumour core region then the superpixel maps generated from the T2 and T1Gd scans should be used.

2.3. Pseudo-Label Refinement

The proposed pipeline consists of the semi-supervised pseudo-labelling method described in subsection 2.1, with the addition of a pseudo-label seeded region refinement algorithm as shown in Fig. 1.

A feature vector was calculated for each superpixel which comprises 9 intensity, texture and gradient based statistical features (mean, variance, skewness, 10-bin intensity histogram, contrast, energy, entropy, 10-bin histogram of gradient orientation and 10-bin histogram of gradient magnitude). Inspired by Chaibou et al. (2018) [16], the proposed merging procedure is based on the agglomerative clustering algorithm via the Ward method [20]. The superpixel similarity measure is defined in the same way as Chaibou et al. (2018), where both content similarity and border similarity make up the overall similarity measure. The important difference in our work is that the region refinement algorithm is initially seeded with the network-inferred pseudo-labels as a starting point, where clustering is constrained to the pseudo-label region rather than general unsupervised clustering throughout the image. Furthermore, our work is in 3D rather than 2D and utilises 3-channels of low contrast MRI images rather than natural images. The proposed superpixel region

Algorithm 1: Pseudo-label seeded superpixel region refinement

Data: $Image, sim_0, p_0, n_c$
Result: p_r
 $s_p \leftarrow \text{Slic}(Image)$
 $p_r \leftarrow \text{FitPseudolabelToSuperPixels}(s_p, p_0)$
repeat
 $N_p \leftarrow \text{FindNeighbours}(s_p, p_r)$
 $merged \leftarrow false$
 $count \leftarrow 0$
 repeat
 $count \leftarrow count + 1$
 $sim_p \leftarrow \{Sim(p_r, n) \forall n : n \in N_p\}$
 $n_c^* \leftarrow \text{argmax}(sim_p)$
 $N_c \leftarrow \text{FindNeighbours}(s_p, n_c^*)$
 $sim_c \leftarrow \{Sim(n_c^*, n) \forall n : n \in N_c\}$
 $n_q^* \leftarrow \text{argmax}(sim_c)$
 if n_q^* **in** p_r **then**
 if $Sim(p_r, n_c^*) > sim_0$ **then**
 $merged \leftarrow true$
 $p_r \leftarrow \text{Merge}(p_r, n_c^*)$
 end
 end
 $N_p \leftarrow N_p \setminus n_c^*$
 until $merged$ **or** $count > n_c$;
until $merged = false$;

refinement algorithm is shown in algorithm 1. As the pseudo-label region grows, the similarity between the region and its neighbouring superpixels is recalculated after each merging operation. This is in recognition that as the region grows, the similarity between the region and its neighbours will also change. The refinement of a pseudo-label stops when there are no more superpixel neighbours which satisfy the merging conditions: the similarity must be above the stopping similarity sim_0 and the region neighbour must mutually choose the pseudo-label region as its most similar neighbour.

The variables in algorithm 1 are defined as follows: $Image$: 3D MRI patient scan, sim_0 : user-defined stopping similarity, p_0 : initial 3D pseudo-label predicted by the network, n_c : hyperparameter which defines the number of candidate neighbours to check for mutual similarity with the pseudo-label region in a given iteration, s_p : superpixel map, p_r : refined pseudo-label, N_p : the superpixel neighbours of the pseudo-label region, sim_p : similarities between a merging candidate and the pseudo-label region, n_c^* : the candidate with the highest similarity to the pseudo-label region, N_c : the superpixel neighbours of a given merging candidate, sim_c : similarities between a merging candidate and each of its neighbours, n_q^* : the neighbour q with the highest similarity to a given merging candidate.

3. EXPERIMENT SETUP

3.1. Dataset

The data used in this paper is from the BraTS 2020 dataset [3, 18, 19], comprising 369 patients of multimodal MRI scans of glioblastoma (GBM) and lower grade glioma (LGG) with accompanying ground truth labels.

The data was split between training and testing with a ratio of 9:1, resulting in a training set of 331 patients and a hold-out test set of 38 patients. The data contains four different MRI modalities

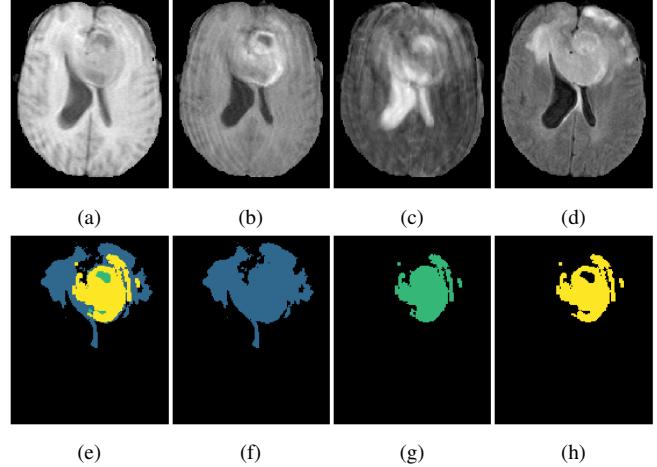


Fig. 3: The different MRI modalities present in BraTS. (a) native (T1), (b) post-contrast T1-weighted (T1Gd), (c) T2-weighted (T2), (d) T2 Fluid Attenuated Inversion Recovery (T2-FLAIR), (e) original BraTS annotations, (f)-(h) region-based annotations: whole tumour (blue), tumour core (green), enhancing tumour (yellow)

(native (T1), post-contrast T1-weighted (T1Gd), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (T2-FLAIR)) for each patient as shown in Fig. 3. It should be noted that the visibility of the different tumour regions depends on the MRI modality used.

Fig. 3e shows the original tumour regions annotated in the BraTS 2020 dataset: Background, Edema, Active Enhancing and Non-enhancing. Fig. 3f and 3g show the regions of interest for this work: whole tumour (WT) and tumour core (TC) respectively.

3.2. Implementation Details

3.2.1. U-Net Training Framework

The nnU-Net framework [21, 22] was used in this work and its trainer was modified for semi-supervised training. The U-Net makes a prediction for the WT and TC regions for each pixel. The default training protocols are used: an initial learning rate of 0.01 updated by the ‘polyLR’ learning rate schedule [23], stochastic gradient descent optimiser with momentum (0.99), all models were trained with a fixed length of 1000 epochs and batch size 2. Patch-based training was performed with a patch size of $128 \times 128 \times 128$ and training was performed for 125,000 iterations.

Data augmentation, as in [21, 22], was applied during training and patch oversampling of patches containing at least one of the foreground classes was implemented, so that 33.3% of patches were guaranteed to contain one of the foreground classes present in the selected training sample. The loss function was an equally-weighted combination of binary cross-entropy and dice loss.

3.2.2. Pseudo-label Baseline

The pseudo-labelling method is described in subsection 2.1 where the number of labelled patients is 5 and the number of unlabelled patients is 259. As training progresses, more pseudo-labelled batches are introduced into the training as described by the dynamic weighting factor $\alpha(e)$ in (1). The hyperparameters were set to be $\alpha_f = 3$, $T_1 = 200$, $T_2 = 700$. The total number of batches in an epoch, N_T , is defined as 250 in this work. The hyperparameter N_e , the number

Table 1: Results on the test set of 38 patients. nnUNet fully-supervised ceiling, nnUNet fully-supervised baseline, semi-supervised pseudo-label baseline and the proposed method in the format average \pm std. The best values are highlighted in bold.

Method	WT DSC	TC DSC	WT HD-95	TC HD-95	WT mIoU	TC mIoU
Fully supervised ceiling (331 labelled)	0.904 ± 0.006	0.869 ± 0.004	6.099 ± 0.496	4.128 ± 0.371	0.836 ± 0.006	0.802 ± 0.004
Fully supervised baseline (5 labelled)	0.794 ± 0.006	0.624 ± 0.045	14.588 ± 1.386	17.671 ± 2.559	0.682 ± 0.011	0.529 ± 0.050
Semi-supervised baseline (5 labelled, 259 unlabelled)	0.799 ± 0.025	0.636 ± 0.070	14.244 ± 1.873	15.164 ± 4.782	0.686 ± 0.033	0.548 ± 0.076
Our method (5 labelled, 259 unlabelled)	0.824 ± 0.023	0.707 ± 0.027	14.295 ± 3.094	14.735 ± 3.385	0.722 ± 0.025	0.621 ± 0.024

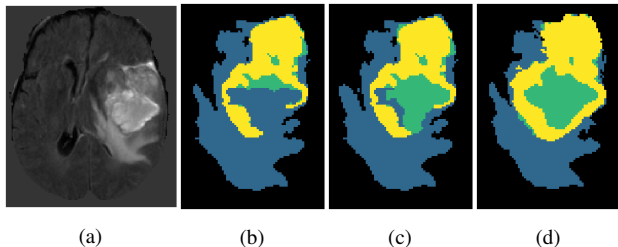


Fig. 4: Pseudo-label refinement for patient 224 after 200 epochs training (a) FLAIR image (b) network-inferred pseudo-label (c) pseudo-label post 3-D superpixel refinement and (d) GT label

of pseudo-labelled batches in epoch e as defined in (2), translates to there being 2 pseudo-labelled batches at epoch 200 (1% of the epoch) to 188 unlabelled batches from epoch 700 onwards (75% of the epoch).

3.2.3. Proposed Pipeline

The proposed pipeline is shown in Fig. 1, which incorporates superpixel region refinement after the pseudo-label update step. In step 1, the network is trained in a fully-supervised manner on only the labelled data. In step 2, the trained network is used to infer pseudo-labels for the unlabelled data. In step 3, the pseudo-labels are refined by the superpixel region refinement algorithm as described in subsection 2.3. In step 4, the network is trained on both the labelled and pseudo-labelled data. Steps 2-4 are repeated as described in 2.1.

The U-Net training framework was a modified version of nnUNet for semi-supervised training as described in subsection 3.2.1, subsection 2.1 and subsection 3.2.2. The region refinement algorithm shown in algorithm 1 was run after every pseudo-label update step (every 200 epochs). The stopping similarity was empirically set to $sim_0=0.1$ and $n_c = 30$ in this work. Experiments were run as 5-fold cross-validation before evaluating the resulting models on the 38-patient test set.

4. RESULTS & DISCUSSION

Table 1 shows the results for the 38-patient hold-out test set. To evaluate the performance of the proposed method, it was compared with a fully-supervised method for all the training data (331 patients: 264 training, 67 validation) which gives a performance ceiling, a fully-supervised baseline (72 patients: 5 training, 67 validation) and a semi-supervised pseudo-labelling baseline method inspired by Lee

(2013) [4] (331 patients: 264 training (5 labelled, 259 unlabelled, 67 validation). The proposed method used the same setup as the pseudo-label-based semi-supervised method, with the addition of the superpixel region refinement algorithm applied to the pseudo-labels after every update step. We report three different metrics to evaluate the segmentation quality: Dice (DSC), 95% Hausdorff Distance (HD-95) and the mean Intersection-over-Union (mIoU). The best results in Table 1 are highlighted in bold. In general, Table 1 shows that our proposed method achieves a consistent improvement in performance across the three metrics compared to the baseline methods. We report DSC=0.824 and DSC=0.707, HD-95=14.295 and HD-95=14.735, mIoU=0.722 and mIoU=0.621 for the WT and TC regions respectively, when trained on only 5 labelled patients and 259 unlabelled patients. Fig. 4 shows the refinement of the pseudo-label for patient 224 in the training set. Fig. 4 (c) clearly shows that the superpixel refinement has improved the inferred pseudo-label shown in Fig. 4 (b) and is now closer to the ground truth label shown in Fig. 4 (d).

These results demonstrate the ability of the proposed method, in particular the use of information from multiple modalities, in merging visually non-homogeneous superpixels which belong to the same semantic region, thus overcoming a problem known as “semantic gap”. Further, this method is able to deal with the irregular shapes typical of pathology, such as brain tumours, something superpixels are well suited for. In general, we show that refined pseudo labels result in a better trained model. Given the promising initial results reported in Table 1, we plan, in future work, to vary the number of labelled patients as well as apply our method to some other medical imaging datasets to determine its robustness.

5. CONCLUSIONS

This work investigated the problem of inaccurate pseudo-labels in semi-supervised learning. To overcome this, a simple 3D superpixel region growing algorithm has been developed which refines the pseudo-labels as an update step during training. The results show that, by utilising the features and edges predefined by superpixels, the pseudo-labels, and consequently training, can be improved as shown by the stated performance boost. Although this method has been demonstrated for the task of brain tumour segmentation in MRI, this proof of concept demonstrates the robustness of the proposed method to objects with varying topology and low contrast, which is of use to a wide variety of problems in the medical imaging domain. Furthermore, this method may be used to determine the level of supervision required in future biomedical projects, thus alleviating the annotator burden.

6. COMPLIANCE WITH ETHICAL STANDARDS

This research study was conducted retrospectively using human subject data made available in open access by the BraTS 2020 challenge. Ethical approval was not required as confirmed by the license attached with the open access data.

7. REFERENCES

- [1] P. Suetens, E. Bellon, D. Vandermeulen, M. Smet, G. Marchal, J. Nuyts, and L. Mortelmans, "Image segmentation: methods and applications in diagnostic radiology and nuclear medicine," *Eur J Radiol*, vol. 17, no. 1, pp. 14–21, 1993.
- [2] A. Hosny, C. Parmar, J. Quackenbush, L. Schwartz, and H. Aerts, "Artificial intelligence in radiology," *Nature Reviews Cancer*, vol. 18, pp. 500–510, 2018.
- [3] B.H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, and J. Kirby et al., "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015.
- [4] D.H. Lee, "Pseudo-label : The simple and efficient semi-supervised learning method for deep neural networks," *ICML 2013 Workshop: Challenges in Representation Learning*, pp. 1–6, 2013.
- [5] L.C. Chen, R.G. Lopes, B. Cheng, M.D. Collins, E.D. Cubuk, B. Zoph, H. Adam, and J. Shlens, "Naive-student: Leveraging semi-supervised learning in video sequences for urban scene segmentation," vol. 12354, pp. 695–714, 2020.
- [6] S. Li, Z. Wei, J. Zhang, and L. Xiao, "Pseudo-label selection for deep semi-supervised learning," pp. 1–5, 2020.
- [7] O. Chapelle, B. Scholkopf, and A. Zien, "Semi-supervised learning," *IEEE Transactions on Neural Networks*, vol. 20, no. 3, pp. 542, 2009.
- [8] J. Ma, Z. Nie, C. Wang, G. Dong, Q. Zhu, J. He, L. Gui, and X. Yang, "Active contour regularized semi-supervised learning for covid-19 ct infection segmentation with limited annotations," *Physics in Medicine & Biology*, vol. 65, no. 22, pp. 225034, 2020.
- [9] T.F. Chan and L.A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
- [10] J. Tavares and J. Renato, *Topics in Medical Image Processing and Computational Vision*, vol. 8, Springer Netherlands, 2013.
- [11] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.
- [12] Z. Tian and L. Liu, Z. Zhang, and B. Fei, "Superpixel-based segmentation for 3d prostate mr images," *IEEE transactions on medical imaging*, vol. 35, no. 3, pp. 791–801, 2016.
- [13] W. Qin, J. Wu, F. Han, Y. Yuan, W. Zhao, B. Ibragimov, J. Gu, and L. Xing, "Superpixel-based and boundary-sensitive convolutional neural network for automated liver segmentation," *Physics in Medicine & Biology*, vol. 63, no. 9, pp. 095017, 2018.
- [14] M. Daoud, A. Atallah, F. Awwad, M. Al najar, and R. Alazrai, "Automatic superpixel-based segmentation method for breast ultrasound images," *Expert Systems with Applications*, vol. 121, pp. 78–96, 2019.
- [15] J. Borovec, J. Kybic, and A. Sugimoto, "Region growing using superpixels with learned shape prior," *Journal of Electronic Imaging*, vol. 26, no. 6, pp. 1–13, 2017.
- [16] M. Chaibou, P.H. Conze, K. Kalti, B. Solaiman, and M. Mahjoub, "Adaptive strategy for superpixel-based region-growing image segmentation," *Journal of Electronic Imaging*, vol. 26, 2017.
- [17] H. Wang, Y. Cong, O. Litany, Y. Gao, and L. Guibas, "3diomatch: Leveraging iou prediction for semi-supervised 3d object detection," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14615–14624, 2021.
- [18] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, and et al., "Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features," *Nature Scientific Data*, vol. 4, pp. 170117, 2017.
- [19] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, R.T. Shinohara, C. Berger, and et al., "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge," p. 38, 2019.
- [20] J.H. Ward Jr., "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236–244, 1963.
- [21] F. Isensee, J. Petersen, S.A. Kohl, P.F. Jäger, and K.H. Maier-Hein, "nnu-net: a self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, pp. 203–211, 2021.
- [22] F. Isensee, P.M. Jäger, P.F. and Full, P. Vollmuth, and K.H. Maier-Hein, "nnu-net for brain tumor segmentation," pp. 118–132, 2021.
- [23] L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deepplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.