Model-Agnostic Post-Processing based on Recursive Feedback for Medical Image Segmentation

JAEO KIM1, SEOKHO KANG1

1Department of Industrial Engineering, Sungkyunkwan University, Suwon 16419, South Korea
Corresponding author: Seokho Kang (s.kang@skku.edu).

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT; Ministry of Science and ICT) (No. NRF-2020R1C1C1003232).

ABSTRACT In medical image segmentation, post-processing can effectively improve the performance of a segmentation model. Existing post-processing methods generally require additional training of a post-processing model using training data or designing a post-processing procedure based on a high level of domain knowledge. Their application is limited in many real-world situations due to the lack of prerequisites. In this study, we present a post-processing method that can be applied to any existing segmentation model without requiring the use of training data or domain knowledge. Given a segmentation model of any type, the proposed method improves its prediction based on the recursive feedback mechanism. For a query image, we first obtain its prediction mask by using the segmentation model. Based on the prediction mask, we modify the original image by selectively blurring the area in which the target object is expected to be absent. Subsequently, the modified image is fed again into the model to acquire a refined prediction mask. We repeat this process to obtain multiple prediction masks, which are then combined to yield the final prediction. We verified the effectiveness of the proposed method through experiments using real-world medical datasets.

INDEX TERMS image segmentation, post-processing, recursive feedback, convolutional neural network

I. INTRODUCTION

Image segmentation is a task of classifying every pixel in an image to identify a set of pixels to which the target object belongs. Segmenting numerous images manually is time-consuming and tedious. Hence, the demand for the automation of image segmentation is high [1]. Especially, image segmentation has been actively studied in the medical domain because it plays an important role in identifying the objects of interest from medical images, such as polyps, tumors, and skin lesions [1-3].

The image segmentation task can be mathematically formulated as follows. A query image of $m \times n$ size is given in the form of $X \in \mathbb{R}^{m \times n \times 3}$. Based on its object region, image $X$ is annotated at the pixel level with the segmentation mask $Y \in \{0, 1\}^{m \times n}$, in which the value of each element is 1 if the corresponding pixel in $X$ belongs to the target object region, and 0 otherwise. In the absence of the actual mask $Y$, segmentation model $f$ seeks to segment $X$ by predicting $Y$ as a function of $X$, thereby yielding the prediction mask $\hat{Y} = f(X) \in [0, 1]^{m \times n}$. For prediction mask $\hat{Y}$, each element is a probabilistic prediction of the corresponding element in $Y$, ranging from 0 to 1.

Considerable research efforts have been made to achieve high-performance segmentation models. Conventional approaches are based on Markov random fields [4, 5], conditional random fields (CRFs) [6, 7], support vector machines [8, 9] and random forests [10, 11]. In addition, unsupervised learning algorithms have been applied to image segmentation, such as k-means clustering [12, 13], region growing [15-17], and watershed transformation [18, 19]. With the advance of deep learning in recent years, convolutional neural networks (CNNs) have shown dramatic performance improvement in image segmentation [20]. Given a training dataset of $N$ images and their segmentation masks, denoted by $D = \{(X_i, Y_i)\}_{i=1}^N$, a CNN is trained in an end-to-end fashion as a segmentation model. The primary research direction for improving image segmentation has focused on designing the CNN architectures for segmentation models.
Another important research direction on improving image segmentation is the post-processing of the output from the segmentation model for better predicting the segmentation mask \[21, 22\]. Existing post-processing methods can be categorized into two main approaches. The first approach is to build a post-processing model that refines the output from the segmentation model. The requirements of this approach are access to the original training dataset, and extra time and resources for additional training. The second approach is to devise a post-processing procedure manually, for which a high level of domain knowledge is necessary in general. Hence, the existing methods may not be applicable in situations where their prerequisites are not satisfied due to some practical constraints.

In this study, we present a model-agnostic post-processing method that can be applied to any segmentation model without requiring any specific prerequisites. The key idea of the proposed method is to make the segmentation model focus more on its input region, where the target object is expected to be present. Given an existing segmentation model of any type, the proposed method improves its prediction by introducing a post-processing procedure based on the recursive feedback mechanism. For a query image, we feed it into the segmentation model to obtain the prediction mask. Based on the prediction mask, we modify the original input image by selectively blurring its region where the target object is expected to be absent. Subsequently, we obtain an improved prediction mask by re-entering the modified image into the segmentation model. We perform the process repeatedly to obtain several prediction masks, which are then aggregated to obtain the final prediction mask. To validate the effectiveness of the proposed method, we conducted experiments with four real-world medical datasets.

The remainder of this paper is organized as follows. In Section II, we review the related work. In Section III, we introduce the proposed post-processing method. Section IV shows the experimental results. Finally, the conclusion and future work are given in Section V.

II. RELATED WORK

A. POST-PROCESSING METHOD FOR IMAGE SEGMENTATION

Various post-processing methods have been developed to improve the performance of existing segmentation models for image segmentation tasks. Given a query image, these methods post-process the prediction mask obtained from the model to predict the actual segmentation mask more accurately. Existing methods for medical image segmentation can be categorized into two main approaches.

The first approach is to build an additional post-processing model using the training data. In general, this approach has no restrictions on the domain to which it applies. The post-processing model uses the prediction mask by the segmentation model as well as additional inputs, such as the original image and handcrafted features, to obtain a refined prediction mask. Many studies utilized CRFs as a post-processing model \[25\]. CRFs help correcting noise and inconsistency in the prediction mask by considering the geometric relations between pixels \[26, 27\]. To alleviate the high computational cost of CRFs, dense-CRFs \[28, 31\] and convolutional-CRFs \[32\] have been proposed. In other studies, non-CRF post-processing models have been used. Chlebus et al. \[33\] used a random forest with 36 handcrafted features regarding the shape of object and location in the image. Larrazabal et al. \[34\] used a denoising autoencoder to reduce noise in the prediction mask by considering topological restrictions and convexity.

The second approach is to manually design a post-processing procedure specialized to the task based on a high-level of domain knowledge for a specific domain. In the medical domain, this approach works by imposing constraints on the number, volume, shape, or location of objects on the prediction mask. The methods based on this approach are generally applicable to any segmentation model without requiring the use of training data. Zhao et al. \[35\] eliminated isolated small blocks in the prediction mask to reduce noise to perform brain tumor segmentation using dermoscopy images. Feng et al. \[36\] retained the largest interconnected block and removed other parts in the prediction mask to identify an organ to perform multi-organ segmentation from CT images. Groza and Kuzin \[37\] reregarded small blocks in the prediction mask as noise to improve pneumothorax segmentation from chest X-ray images. Birenbaum and Greenspan \[38\] set a threshold of lesion size to perform multiple sclerosis lesion segmentation using MRI images. Wang et al. \[39\] merged neighboring blocks with border smoothing in the prediction mask to perform skin lesion segmentation using dermoscopy images.

Both approaches effectively improve the prediction mask from an existing segmentation model when specific prerequisites are satisfied. The first approach involves training of an additional model using training data. The second approach utilizes high-level domain knowledge. Despite their effectiveness, their prerequisites may not be fulfilled in many real-world applications. To overcome the limitation, this study aims to present a post-processing method that does not require any specific prerequisites, such as original training data and domain knowledge.

B. FEEDBACK MECHANISM FOR IMAGE SEGMENTATION

The feedback mechanism is a process wherein the output of a model is recursively routed back to the input of the model. This makes the input and output of the model to affect each other consecutively \[40\]. Recently, several studies have adapted the feedback mechanism to improve image segmentation. In existing methods, a segmentation model is augmented to incorporate the feedback mechanism and train a model using a specialized learning objective.

An and Liu \[41\] augmented a CNN to incorporate feedback connections from the output layer to hidden layers based...
on feedback recovery and feedback selective algorithms. Shen et al. [42] used multi-stage multi-recursive-input CNNs as a segmentation model. In terms of the sequence of the CNNs, the side outputs in each stage were concatenated with the original input image to serve as the input to the CNN in the next stage. Shibuya and Hotta [43] used a CNN with LSTM layers as a segmentation model. The first output of the model is fed into the model as input to obtain the second output as the final prediction mask. Girum et al. [44] used an encoder–decoder CNN as a segmentation model. The first output of the model was processed with another CNN to extract features, that were reintegrated into the intermediate layer of the model to obtain the second output.

Unlike existing methods, this study aims to make use of the feedback mechanism as the post-processing for a given segmentation model of any type. The proposed post-processing method does not require any architectural modifications or additional training.

III. PROPOSED METHOD

A. OVERVIEW

In this section, we present the proposed method. The problem situation we address in this study is as follows. We are given a segmentation model \( f \) of any type, which predicts the segmentation mask \( Y \in \{0, 1\}^{m \times n} \) for an input image \( X \in \mathbb{R}^{m \times n \times 3} \) as \( Y = f(X) \). The training data or explicit domain knowledge are not available for use.

For a query image \( X \), we wish to improve its prediction mask \( \hat{Y} \) under this situation. The proposed method introduces a model-agnostic post-processing procedure that can be applied to any segmentation model without requiring any specific prerequisites. The key idea of the proposed method is the iterative updating of prediction mask \( \hat{Y} \) based on the recursive feedback mechanism. At each update, the prediction mask \( \hat{Y} \) is utilized as feedback information to modify the input image \( X \) by selectively blurring the area in which the target object is expected to be absent.

Algorithm 1 shows the pseudocode of the post-processing procedure for obtaining the predicted mask \( \hat{Y} \). The schematic diagram is illustrated in Fig. 1.

Algorithm 1: Post-processing procedure based on recursive feedback

| input: query image \( X \in \mathbb{R}^{m \times n \times 3} \), segmentation model \( f \) |
| output: prediction mask \( \hat{Y} \in \{0, 1\}^{m \times n} \) |

1. procedure POST-PROCESSING
2. \( X^{(0)} \leftarrow X \)
3. \( \hat{Y}^{(0)} \leftarrow f(X^{(0)}) \)
4. for \( t = 1 \) to \( T \) do
5. \( X^{(t)} \leftarrow \hat{Y}^{(t-1)} \odot X^{(0)} + \left(1 - \hat{Y}^{(t-1)}\right) \odot \phi_{\text{blur}}(X^{(t-1)}) \)
6. \( \hat{Y}^{(t)} \leftarrow f(X^{(t)}) \)
7. end for
8. \( \hat{Y} \leftarrow \frac{\sum_{t=0}^{T} \gamma^{t} \hat{Y}^{(t)}}{\sum_{t=0}^{T} \gamma^{t}} \)
9. end procedure

B. POST-PROCESSING PROCEDURE

The post-processing procedure consists of three steps: (1) initial prediction; (2) iterative refinement based on recursive feedback; (3) aggregation for final prediction. We describe each step below.

1) Initial prediction

Given a query image \( X \) and a segmentation model \( f \), it is used as the initial input \( X^{(0)} \), i.e., \( X^{(0)} = X \). Using the segmentation model \( f \), we obtain the initial predicted mask as:

\[
\hat{Y}^{(0)} = f(X^{(0)}). \tag{1}
\]

2) Iterative refinement based on recursive feedback

After initialization, we repeat the recursive feedback process \( T \) times to obtain refined predicted masks. At the \( t \)-th iteration, \( t = 1, \ldots, T \), the previously predicted mask \( \hat{Y}^{(t-1)} \) is used to produce the modified image \( X^{(t)} \) as follows:

\[
X^{(t)} = \hat{Y}^{(t-1)} \odot X^{(0)} + \left(1 - \hat{Y}^{(t-1)}\right) \odot \phi_{\text{blur}}(X^{(t-1)}) \tag{2}
\]

where \( \odot \) is the element-wise product and \( \phi_{\text{blur}} \) is a blur operator. Each pixel in \( X^{(t)} \) is a convex combination of the corresponding pixels in the original image \( X^{(0)} \) and the blurred version of the previous image \( \phi_{\text{blur}}(X^{(t-1)}) \), for which the prediction mask \( \hat{Y}^{(t-1)} \) weights for their contributions. If an element in \( \hat{Y}^{(t-1)} \) is closer to 1 and 0, then the contribution of the corresponding pixels in \( X^{(0)} \) and \( \phi_{\text{blur}}(X^{(t-1)}) \) are high, respectively. The pixels in \( X^{(t)} \) that are expected to be devoid of the target object are blurred, while the remaining pixels are preserved. Consequently, the modified image \( X^{(t)} \) maintains the target object context whereas irrelevant aspects are suppressed. The modified image \( X^{(t)} \) is fed into the segmentation model \( f \) to obtain the updated predicted mask as \( \hat{Y}^{(t)} \), which is expressed as:

\[
\hat{Y}^{(t)} = f(X^{(t)}). \tag{3}
\]

3) Aggregation for final prediction

After the repetition of the recursive feedback process, we finally obtain \( T + 1 \) prediction masks \( \hat{Y}^{(0)}, \hat{Y}^{(1)}, \ldots, \hat{Y}^{(T)} \). We aggregate them with an exponential decay with respect to \( t \) to obtain the final prediction mask \( \hat{Y} \), as follows:

\[
\hat{Y} = \frac{\sum_{t=0}^{T} \gamma^{t} \hat{Y}^{(t)}}{\sum_{t=0}^{T} \gamma^{t}} \tag{4}
\]

where \( \gamma \) is a decay factor whose value is in the range of \((0, 1]\). Owing to the decay, earlier prediction masks with a smaller \( t \) are weighted more than the latter ones with larger \( t \), hence the effect on the final prediction mask \( \hat{Y} \) is higher. Normalization by the denominator makes every element of \( \hat{Y} \) to be within the range of \([0, 1]\).

C. DISCUSSION

The proposed method utilizes the prediction mask of the segmentation model to determine where to blur. In the input
image, the region predicted as the target class is regarded as important context, whereas the remaining region is regarded as irrelevant context. The image is repetitively modified by selectively blurring irrelevant contexts while preserving important contexts. Accordingly, the modified image contains fewer irrelevant aspects. This enables the segmentation model to focus more on important contexts while suppressing irrelevant contexts to predict the segmentation mask more accurately. Consequently, the final prediction is enhanced toward improving the segmentation performance.

IV. EXPERIMENTS

A. DATA DESCRIPTION

We investigated the effectiveness of the proposed method using four medical image datasets: CVC-ClinicDB, Kvasir-SEG, Kvasir-Instrument, and ISIC-2017. The datasets correspond to different image segmentation tasks. Each dataset contains a number of images and their corresponding segmentation masks, which are annotated and verified by domain experts. Table 1 summarizes the characteristics of the datasets. Since the original images have various resolutions, we resized every image and segmentation mask to 384 × 384. In the experiments, we randomly partitioned each dataset into 80% as the training set for training the segmentation model and the remaining 20% as the test set to evaluate the performance of the proposed method.

B. SEGMENTATION MODEL

To build segmentation models, we employed U-Net [49] which is the most widely used CNN architecture for medical image segmentation tasks [50, 51]. U-Net consists of encoder–decoder like architecture with skip connections. It has ability to extract global and local features simultaneously via feature map concatenation through skip connections. Consequently, it affords high localization performance. We used VGG16 [52] as the encoder backbone of the architecture. Segmentation models were implemented based on the Segmentation_Models library.

For the experiments, each segmentation model was trained as follows. The loss function was set to binary cross-entropy. The parameters of the encoder in the model were pre-trained on ImageNet [53]. Subsequently, the parameters of the model were updated using 75% of the training dataset, for which we used the Adam optimizer [54] with a learning rate of 10^{-4} and a mini-batch size of 10. We applied online data augmentation with rotation, flip, shift, zoom, and brightness operations. The remaining 25% was used to validate the model for early-stopping. The training was terminated when the validation loss failed to decrease over 10 successive epochs, or when the number of epochs reached 500.

C. EXPERIMENT SETTINGS

The proposed method involves three hyperparameters: the number of iterations \( T \), blur operator \( \phi_{\text{blur}} \), and decay factor.
In the experiments, we used the following settings for the hyperparameters. The number of iterations $T$ was varied from 1 to 10 to investigate its effect on the performance. As a baseline, we evaluated $T = 0$, for which the vanilla prediction mask of the segmentation model was used without applying recursive feedback. For the blur operator $\phi_{\text{blur}}$, we compared the four algorithms implemented in the OpenCV library\textsuperscript{[4]} to evaluate their effect on the performance: average blur, median blur, Gaussian blur, and bilateral filter. For all blur operators, we set the kernel size to 3. For Gaussian blur, the standard deviation of the Gaussian kernel was automatically set based on the default setting in the library. For bilateral filter, the standard deviation of the filter was set to 75. The decay factor $\gamma$ was set to 0.5 for all experiments.

The performance of the proposed method was evaluated in terms of intersection over union (IoU) and area under the receiver operating characteristic curve (AUROC), which are typically used metrics for image segmentation. Given a ground-truth segmentation mask $Y$ and its prediction $\hat{Y}$ for an image, IoU is defined as the number of target object pixels common in the masks divided by the number of target object pixels present across the masks. For IoU calculation, we specified the decision threshold to identify the target object pixels in $\hat{Y}$ as 0.5. The ROC curve plots the true positive rate against the false positive rate of pixel-level predictions for an image by varying the decision threshold. AUROC evaluates the overall performance of an image in a threshold-independent manner by integrating the performance over all possible settings of the threshold.

All experiments were performed 10 times independently with different random seeds. We report the mean and standard deviation of each result over 10 replications.

### D. EXPERIMENTAL RESULTS

Tables\textsuperscript{2} and \textsuperscript{3} show a comparison of the performance of the proposed method in terms of IoU and AUROC, respectively, with various settings of the number of iterations $T$ and blur operator $\phi_{\text{blur}}$. In each row, the results better than the baseline ($T = 0$) are shown in bold. The single and double asterisks (*) and **) indicate that the proposed method significantly outperformed the baseline at significance levels of 0.05 and 0.01, respectively, based on the paired t-test.

As shown in the tables, the proposed method improved the segmentation performance over the baseline in terms of IoU and AUROC with statistical significance for CVC-ClinicDB, Kvasir-SEG, and Kvasir-Instrument datasets. In these datasets, IoU continued to increase with $T$ and saturated at approximately $T=5$, whereas AUROC was the highest at approximately $T=1$. On ISIC-2017 dataset, the proposed method did not indicate effectiveness in terms of IoU and slightly improved the performance in terms of AUROC. No significant performance differences were observed among the blur operators, whereas median blur consistently showed favorable performance over the experiments.

Fig. 2 shows the images and their prediction masks, in which the proposed method with the median blur was the most effective. We observed that the prediction mask improved by expanding target object region while removing noise as the number of iterations $T$ increased.

We examined the characteristics of the images in the benchmark datasets to determine the case in which the proposed method performed better. For the first three datasets, the images exhibited distinct features in both the target object and background regions. Hence, blurring operations helped smoothing the features on the background region. By contrast, the images in ISIC-2017 dataset exhibited a monotonous background region, therefore blurring operations gave little effect on the background region.

Based on the findings from the experimental results, we suggest the following settings for the proposed method. For the number of iterations $T$, we recommend setting $T = 5$ when the goal is to improve IoU and $T = 1$ when AUROC is the target performance indicator. For the blur operator $\phi_{\text{blur}}$, we recommend using the median blur owing to its robustness and low computational cost compared with the other blur operators.

### V. CONCLUSION

We presented a model-agnostic post-processing method based on the recursive feedback mechanism to improve image segmentation. Given a segmentation model and a query image, the proposed method worked by inputting the query image into the segmentation model to obtain the prediction mask. Subsequently, the image is modified by selectively blurring the region where the target object is not expected to be present. The modified image is re-entered into the segmentation model to obtain a refined prediction mask. Multiple prediction masks are obtained by repeating the process, and they are then aggregated with an exponential decay to obtain the final prediction mask. The proposed method can be applied to any existing segmentation model without requiring additional training.

We demonstrated the effectiveness of the proposed method using four real-world medical datasets. Experimental results


\begin{table}[h]
\centering
\caption{Dataset description}
\begin{tabular}{|c|c|c|c|}
\hline
Dataset & No. Images & Resolution & Target Object & Domain \\
\hline
CVC-ClinicDB & 612 & 574×500 & colorectal poly & colonoscopy \\
Kvasir-SEG & 1,000 & 332×487 - 1920×1072 & gastrointestinal poly & gastrointestinal endoscopy \\
Kvasir-Instrument & 590 & 720×487 - 1280×1024 & diagnostic and therapeutic tool & gastrointestinal endoscopy \\
ISIC-2017 & 2,000 & 1022×767 - 6688×4439 & skin lesion & dermatoscopy \\
\hline
\end{tabular}
\label{table:dataset}
\end{table}
showed that the proposed method improved the segmentation performance with statistical significance on the datasets on benchmark datasets. We expect that the proposed method will contribute to additional improvement of segmentation models in real-world situations where existing post-processing methods are not adaptable owing to practical constraints. In fact, the proposed method can also be employed along with other post-processing methods. For future work, we will further enhance the proposed method by utilizing other image-smoothing operators. In addition, we will investigate the characteristics of query images for which the application of the proposed method is more beneficial, such that the hyperparameter \( T \) can be adjusted adaptively for different query images.

**REFERENCES**

[1] D. D. Patil and S. G. Deore, “Medical image segmentation: a review,” International Journal of Computer Science and Mobile Computing, vol. 2, no. 1, pp. 22–27, 2013.

[2] M. E. Celebi, N. Codella, and A. Halpern, “Dermoscopy image analysis: Overview and future directions,” IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 2, pp. 474–478, 2019.

[3] I. R. I. Haque and J. Neubert, “Deep learning approaches to biomedical image segmentation,” Informatics in Medicine Unlocked, vol. 18, p. 100297, 2020.

[4] Y. Zhang, M. Brady, and S. Smith, “Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm,” IEEE Transactions on Medical Imaging, vol. 20, no. 1, pp. 45–57, 2001.

[5] G. Moser, S. B. Serpico, and J. A. Benediktsson, “Markov random field models for supervised land cover classification from very high resolution multispectral remote sensing images,” in Proceedings of the Tyrrhenian Workshop on Advances in Radar and Remote Sensing. IEEE, 2012, pp. 235–242.

[6] L. Ladicky, C. Russell, P. Kohli, and P. H. S. Torr, “Associative hierarchical CRFs for object class image segmentation,” in Proceedings of the IEEE 12th International Conference on Computer Vision. IEEE, 2009, pp. 739–746.

[7] J. Shotton, J. Winn, C. Rother, and A. Criminisi, “Texton-Robe: Joint appearance, shape and context modeling for multi-class object recognition and segmentation,” in Proceedings of the European Conference on Computer Vision. Springer, 2006, pp. 1–15.

[8] Y. Yang, S. Hallman, D. Ramanan, and C. C. Fowlkes, “Layered object models for image segmentation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 9, pp. 1731–1743, 2011.

[9] P. F. Felzenszwalb, B. R. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part-based models,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 9, pp. 1627–1645, 2010.

[10] F. Schroff, A. Criminisi, and A. Zisserman, “Object class segmentation using random forests,” in Proceedings of the British Machine Vision Conference, 2008, pp. 1–10.

[11] J. Shotton, M. Johnson, and R. Cipolla, “Semantic texture forests for image categorization and segmentation,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2008, pp. 1–8.

[12] M.-N. Wu, C.-C. Lin, and C.-C. Chang, “Brain tumor detection using color-based K-means clustering segmentation,” in Proceedings of the 3rd International Conference on Intelligent Information Hiding and Multimedia Signal Processing, vol. 2. IEEE, 2007, pp. 245–250.
(a) Examples with the largest IoU increase by the proposed method

FIGURE 2: Examples of images and their prediction masks that improved the most by proposed method
[13] N. Dhanachandra, K. Manglem, and Y. J. Chanu, “Image segmentation using K-means clustering algorithm and subtractive clustering algorithm,” Procedia Computer Science, vol. 54, pp. 764–771, 2015.

[14] H. P. Ng, S.-H. Ong, K. W. C. Foong, P.-S. Goh, and W. L. Nowinski, “Medical image segmentation using K-means clustering and improved watershed algorithm,” in Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation. IEEE, 2006, pp. 61–65.

[15] J. Tang, “A color image segmentation algorithm based on region growing,” in Proceedings of the 2nd International Conference on Computer Engineering and Technology, vol. 6. IEEE, 2010, pp. V6–634.

[16] N. Mesanovic, M. Grigic, H. Huseinagic, M. Males, E. Skejic, and M. Smajlovic, “Automatic CT image segmentation of the lungs with region growing algorithm,” in Proceedings of the 18th International Conference on Systems, Signals and Image Processing, 2011, pp. 395–400.

[17] F. Y. Shih and S. Cheng, “Automatic seeded region growing for color image segmentation,” Image and Vision Computing, vol. 23, no. 10, pp. 877–886, 2005.

[18] L. J. Belaid and W. Mourou, “Image segmentation: A watershed transformation algorithm,” Image Analysis & Stereology, vol. 28, no. 2, pp. 93–102, 2009.

[19] M. Kaur and G. Jindal, “Medical image segmentation using marker controlled watershed transformation,” International Journal of Computer Science and Technology, vol. 2, no. 4, pp. 548–551, 2011.

[20] J. Kuruvilla, D. Sukumaran, A. Sankar, and S. P. Joy, “A review on image processing and image segmentation,” in Proceedings of the International Conference on Data Mining and Advanced Computing. IEEE, 2016, pp. 198–203.

[21] S. Minaee, Y. B. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, “Image segmentation using deep learning: A survey,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.

[22] F. Sultana, A. Sufian, and P. Dutta, “Evolution of image segmentation using deep convolutional neural network: A survey,” Knowledge-Based Systems, vol. 201-202, pp. 106062, 2020.

[23] P. Shivhare and V. Gupta, “Review of image segmentation techniques including pre & post processing operations,” International Journal of Engineering and Advanced Technology, vol. 4, no. 3, pp. 153–157, 2015.

[24] B. Li, Y. Shi, Z. Qi, and Z. Chen, “A survey on semantic segmentation,” in Proceedings of the International Conference on Data Mining Workshops. IEEE, 2018, pp. 1233–1240.

[25] N. Atif, M. Bhuyan, and S. Ahamed, “A review on semantic segmentation from a modern perspective,” in Proceedings of the International Conference on Electrical, Electronics and Computer Engineering. IEEE, 2019, pp. 1–6.

[26] X. He, R. S. Zemel, and M. A. Carreira-Perpinán, “Multiscale conditional random fields for image labeling,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2. IEEE, 2004, pp. 695–703.

[27] P. Kohli and P. H. S. Torr, “Robust higher order potentials for enforcing label consistency,” International Journal of Computer Vision, vol. 82, no. 3, pp. 302–324, 2009.

[28] Philipp Krähenbühl and V. Koltun, “Efficient inference in fully connected CRFs with Gaussian edge potentials,” Advances in Neural Information Processing Systems, vol. 24, pp. 109–117, 2011.

[29] Y. Zhang, Z. He, C. Zhong, Y. Zhang, and Z. Shi, “Fully convolutional neural network with post-processing methods for automatic liver segmentation from CT,” in Proceedings of the Chinese Automation Congress. IEEE, 2017, pp. 3864–3869.

[30] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “DeepLab: 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, 2018.

[31] D. Jha, P. H. Smetsrud, D. Johansen, T. de Lange, H. D. Johansen, P. Halvorsen, and M. A. Riegler, “A comprehensive study on colorectal polyp segmentation with ResUNet++, conditional random field and test-time augmentation,” IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 6, pp. 2029–2040, 2021.

[32] A. Arnab, S. Zheng, S. Jayasumana, B. Romera-Paredes, M. Larsson, A. Kirill, B. Savchynsky, C. Roher, F. Kahl, and P. H. S. Torr, “Conditional random fields meet deep neural networks for semantic segmentation: Combining probabilistic graphical models with deep learning for structured prediction,” IEEE Signal Processing Magazine, vol. 35, no. 1, pp. 37–52, 2018.

[33] G. Chlebus, A. Schenk, J. H. Moltz, B. van Ginneken, H. K. Hahn, and H. Meine, “Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing,” Scientific Reports, vol. 8, no. 1, p. 15497, 2018.

[34] A. J. Larrazabal, C. Martinez, and E. Ferrante, “Anatomical priors for image segmentation via post-processing with denoising autoencoders,” in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2019, pp. 585–593.

[35] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan, “A deep learning model integrating FCNNs and CRFs for brain tumor segmentation,” Medical Image Analysis, vol. 43, pp. 98–111, 2018.

[36] M. Feng, W. Huang, Y. Wang, and Y. Xie, “Multi-organ segmentation using simplified dense V-Net with post-processing,” in Proceedings of the Challenge Segmentation THoracic Organ Risk CT Images, 2019.

[37] V. Groza and A. Kuzin, “Pneumothorax segmentation with effective conditioned post-processing in chest X-ray,” in Proceedings of the IEEE International Symposium on Biomedical Imaging Workshops. IEEE, 2020.

[38] A. Birenbaum and H. Greenspan, “Multi-view longitudinal CNN for multiple sclerosis lesion segmentation,” Engineering Applications of Artificial Intelligence, vol. 65, pp. 111–118, 2017.

[39] H. Wang, R. H. Moss, X. Chen, R. J. Stanley, W. V. Stoecker, M. E. Celebi, J. M. Malters, J. M. Grichnik, A. A. Marghoob, H. S. Rabinovitz, S. W. Menzies, and T. M. Szalapski, “Modified watershed technique and post-processing for segmentation of skin lesions in dermoscopy images,” Computerized Medical Imaging and Graphics, vol. 35, no. 2, pp. 116–120, 2011.

[40] R. M. Cichy, D. Pantazis, and A. Oliva, “Resolving human object recognition in space and time,” Nature Neuroscience, vol. 17, no. 3, pp. 455–462, 2014.

[41] F.-P. An and Z.-W. Liu, “Medical image segmentation algorithm based on feedback mechanism convolutional neural network,” Biomedical Signal Processing and Control, vol. 53, pp. 101589, 2019.

[42] W. Shen, B. Wang, J. Yuan, Y. Wang, and A. Yuille, “Multi-stage multi-recursive-input fully convolutional networks for neuronal boundary detection,” in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2391–2400.

[43] E. Shibuya and K. Hotta, “Feedback U-Net for cell image segmentation,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020.
JAEHO KIM received the B.S. degrees in systems management engineering and in mechanical engineering from Sungkyunkwan University in 2020. He is currently a M.S. student at the Department of Industrial Engineering, Sungkyunkwan University. His research interests include machine learning applications in uncertainty quantification and post-processing methods for image segmentation.

SEOKHO KANG is an assistant professor of systems management engineering (industrial engineering) at Sungkyunkwan University. He received the B.S. and Ph.D. degrees in industrial engineering from Seoul National University in 2011 and 2015, respectively, and was a research staff member with Samsung Advanced Institute of Technology. His research focuses mainly on developing learning algorithms for efficient data-driven modeling and their applications to real-world data mining problems. He has published a number of papers in refereed journals and conference proceedings related to these areas.

---

J. Bernal, F. J. Sánchez, G. Fernández-Esparrach, D. Gil, C. Rodríguez, and F. Vilariño, “WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians,” Computerized Medical Imaging and Graphics, vol. 43, pp. 99–111, 2015.

D. Jha, P. H. Smedsrud, M. A. Riegler, P. Halvorsen, T. de Lange, D. Johansen, and H. D. Johansen, “Kvasir-SEG: A segmented polyp dataset,” in Proceedings of the International Conference on Multimedia Modeling. Springer, 2020, pp. 451–462.

D. Jha, S. Ali, K. Emanuelsen, S. A. H. A. V. Thambawita, E. Garcia-Ceja, M. A. Riegler, T. de Lange, P. T. Schmidt, H. D. Jogansen, D. Johansen, and P. Halvorsen, “Kvasir-Instrument: Diagnostic and therapeutic tool segmentation dataset in gastrointestinal endoscopy,” in Proceedings of the International Conference on Multimedia Modeling. Springer, 2021, pp. 218–229.

N. C. F. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittler, and A. Halpern, “Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC),” in Proceedings of the International Symposium on Biomedical Imaging. IEEE, 2018, pp. 168–172.

O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2015, pp. 234–241.

G. Du, X. Cao, J. Liang, X. Chen, and Y. Zhan, “Medical image segmentation based on U-Net: A review,” Journal of Imaging Science and Technology, vol. 64, no. 2, p. 20508, 2020.

N. Siddique, S. Phadegan, C. P. Elkin, and V. Devabhaktuni, “U-Net and its variants for medical image segmentation: A review of theory and applications,” IEEE Access, vol. 9, pp. 82 031–82 057, 2021.

K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in Proceedings of the 3rd International Conference on Learning Representations, 2015.

J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2009, pp. 248–255.

D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in Proceedings of the 3rd International Conference on Learning Representations, 2015.