Network Modeling Analysis in Health Informatics and Bioinformatics (2024) 13:47
https://doi.org/10.1007/s13721-024-00483-0

ORIGINAL ARTICLE

One-click annotation to improve segmentation by a convolutional neural network for PET images of head and neck cancer patients

Oona Rainio[1] · Joonas Liedes[1] · Sarita Murtojärvi[1] · Simona Malaspina[1] · Jukka Kemppainen[1] · Riku Klén[1]

Received: 26 March 2024 / Revised: 20 July 2024 / Accepted: 13 August 2024

Abstract
A convolutional neural network (CNN) can be used to perform fully automatic tumor segmentation from the positron emission tomography (PET) images of head and neck cancer patients but the predictions often contain false positive segmentation caused by the high concentration of the tracer substance in the human brain. A potential solution would be a one-click annotation in which a user points the location of the tumor by clicking the image. This information can then be given either directly to a CNN or an algorithm that fixes its predictions. In this article, we compare the fully automatic segmentation to four semi-automatic approaches by using 962 transaxial slices collected from the PET images of 100 head and neck cancer patients. According to our results, a semi-automatic segmentation method with information about the center of the tumor performs the best with a median Dice score of 0.708.

Keywords Convolutional neural network · Head and neck squamous cell carcinoma · Positron emission tomography · Semi-automatic segmentation · Tumor segmentation

1 Introduction

Head and neck squamous cell carcinoma (HNSCC) is the sixth most common cancer type in the world, with closer to a million new cases estimated yearly (GCO 2022). It develops on the squamous cells of the mucosal surfaces in the human head and neck area and the cancer tumor can be found in lips, oral and nasal cavities, sinuses, pharynx, larynx, salivatory glands, and vocal cords (Johnson et al. 2020). HNSCC patients are typically treated with surgery or chemoradiotherapy and, while their prognoses greatly depend on the primary tumor site, post-treatment surveillance is crucial for all patients to detect a recurrent tumor at an early stage (De Felice et al. 2015).

To monitor the patients, different imaging methods are used, including a nuclear medicine imaging method called positron emission tomography (PET). During a PET scan, a patient is first injected with a positron emitting radiotracer and the PET scanner records gamma rays produced by the annihilation of the positrons with the electrons of the human body (Townsend 2004). ^{18}F-fluorodeoxyglucose (FDG) is the most common PET tracer used in clinical oncology, as most cancers exhibit increased glucose metabolism. PET images are fused with anatomic information from magnetic resonance imaging (MRI) or computed tomography (CT), depending on whether the scan is performed with a dedicated PET/CT or PET/MRI device.

Segmentation is one of the routine tasks required in the clinical treatment after the PET scan of a cancer patient. It means denoting all the cancerous segments in the three-dimensional PET image by creating a three-dimensional binary mask corresponding with the original medical image so that one can see the location of tumor, its size, and the possible metastases.

✉ Oona Rainio
ormrai@utu.fi

Joonas Liedes
joolie@utu.fi

Sarita Murtojärvi
sarita.murtojarvi@tyks.fi

Simona Malaspina
simona.malaspina@tyks.fi

Jukka Kemppainen
Jukka.Kemppainen@tyks.fi

Riku Klén
riku.klen@utu.fi

[1] Turku PET Centre, University of Turku and Turku University Hospital, Turku, Finland

Published online: 28 August 2024

This task requires a lot of time and effort for even a singular PET image but a convolutional neural network (CNN) has been noted to be a useful tool in the recent research (Liedes et al. 2023; Ren et al. 2021; Yuan 2021; Xie and Peng 2021; Rainio et al. 2023b) because it can be trained to recognize cancer tumors so that the segmentation can be performed automatically.

However, given limited access to large amounts of sensitive patient data, it is possible that a CNN only learns to recognize the regions with high amounts of FDG but not to separate which of these areas actually contain cancerous tissue. In particular, the human brain and Waldeyer's tonsillar ring could potentially cause false positive segmentation. A potential solution would be for a medical expert to click one pixel in the correct tumor location on the computer screen and give this information to either directly to a CNN or an algorithm that processes the results of the CNN. In practice, this would mean that either the final application first performs automatic segmentation and then asks the user to choose the correctly segmented components by clicking them or it performs semi-automatic segmentation after the user has clicked all the visible tumors. Since clicking a few times is considerably faster than creating the whole mask, this type of a semi-automatic approach would still be very useful. During the recent years, similar one-click methods have been also been developed, examples of them including Segment Anything by FocalClick Chen et al. (2022), Kirillov et al. (2023), One Thing One Click Liu et al. (22021), SimpleClick by Liu et al. (2023), One-Click-Based Perception by Wang et al. (2023), and One Class One Click by Wang et al. (2023) but, as far as we are aware, there is no research considering how the location of the click affects to the results or whether it is better to give this information to a semi-automatic CNN or use it in post-processing.

In this article, we study the one-click annotation method compared to the fully automatic segmentation for a CNN used to process PET images of HNSCC patients with recurrent tumors. Here, the pixel annotated with one click is either a fully random choice of all the ground-truth positive pixels or the center of mass (CM) of all these positive pixels. We study both semi-automatic one-click CNNs and such a post-processing algorithm that, when given results of the fully automatic segmentation and information about the so-called one-click pixels, it classifies only the area containing the chosen pixels as positive and removes the segmentation elsewhere in the image.

2 Materials and methods

2.1 Software requirements

The experiments were done in Python (version: 3.9.9) (van Rossum and Drake 2009) by using the packages TensorFlow (version: 2.7.0) (Abadi et al. 2015), Keras (version: 2.7.0) (Chollet et al. 2015), and SciPy (version: 1.7.3) (Virtanen et al. 2020), and the medical images were inspected by using Carimas (version: 2.10) (Rainio et al. 2023a).

2.2 Data

The data in this study was retrospectively collected from 100 head and neck patients, 89 of whom had been diagnosed with HNSCC and 11 with other head and neck cancer diseases such as adenocystic carcinoma of oral cavity, chondrosarcoma of neck, and papillary or follicular carcinoma of thyroid gland. During their treatment at Turku University Hospital, they had been referred for a PET/MRI treatment response assessment in Turku PET Centre, Turku, Finland, between years 2014–2022. The presence of tumors in their scans was confirmed with either histopathological sampling or follow-up imaging. The mean age of the patients was 62 years with standard deviation of 12 years, and their male–female sex ratio was 2.1. All the patients were at least 18 years of age, consented to research use of their data, and this research was approved by Ethics Committee of the Hospital District of Southwest Finland. The same data has been used earlier in Hellström et al. (2023) where the readers can also find more details and examples about the images.

The patients were imaged with either 3T Philips Ingenuity TF PET/MRI scanner (Philips Health Care) or SIGNA^{TM} PET/MRI scanner with QuantWorks (GE Healthcare) by using ^{18}F-FDG as tracer substance. While this camera also simultaneously performed MRI, only the PET images were given to the CNNs in this study. All the PET images depicted the head and neck area of the patients, and each of them consisted of 32–66 transaxial slices of the size of 512×512 pixels on top of each other. A medical doctor created binary masks for the PET images with aid of an experienced nuclear medicine physician by labelling the voxels with cancer as positive and the rest as negative, and both the PET image and these masks were stored in files of NIfTI format.

Since this study is focused on the two-dimensional segmentation, every transaxial slice of a three-dimensional PET image is considered an individual image. During the pre-processing of the slices, they were converted into size of 128×128 pixels and the pixel values were scaled onto the interval [0, 1] for each slice separately. Similarly, the binary masks were resized into the same size as the image slices so that one pixel in the resized mask was labelled as positive if and only if at least 5 of the corresponding 16 pixels in the original 512×512 mask were positive. After this scaling, the slices that contained less than 6 positive pixels were removed from the data, which left us 962 slices in total from 89 different patients (79 HNSCC patients and 10 patients with other head and neck cancer types). The slices were divided into five sets, each of

which contained 191–194 slices or, equivalently, 19.9–20.1% of the total data, and the CNNs were trained by using five-fold cross-validation based on these five sets. Since this split into five sets was done patient-wise and the CNNs are always initialized before training them with new data, the CNNs have no such information about the patients of the test set that could corrupt the predictions.

For most of the slices, there were only one cancerous component visible, but there were some slices where either the primary tumor was divided into a few separate parts or there were additional metastases. Consequently, one random positive pixel was chosen from each separate positive component of every binary mask and the pixel closest to the CM of every component was computed. Note that the CM of the positive pixels is not necessarily positive but, given the usual size and shape of cancer tumors, it should be very close to the positive pixels. These chosen one-click pixels were saved in two-dimensional 128×128 binary matrices with all elements zeros except the ones identifying the specific pixels. Depending on the method studied, the CNN is given real binary masks as dependable variable and either just the PET slices as two-dimensional matrices or such matrices of the size $128 \times 128 \times 2$ that contain both the PET slice and a matrix with the additional information about the one-click pixels.

2.3 Convolutional neural networks

Our CNN follows the U-Net architecture originally introduced by Ronneberger et al. in 2015 Ronneberger et al. (2015). The idea of the U-Net is that a CNN consists of a contracting path and a symmetric expanding path so that it first understands the context of an image and can then focus on its details to perform segmentation. We chose the U-Net design given its popularity in research about automatic tumor segmentation. Our CNN is similar to the ones in Liedes et al. (2023),[13] but designed for images of 128×128 pixels. The usual sigmoid function on the activation layer was replaced with a linear function because it was noticed on a few trial runs that the CNN with the sigmoid function has a tendency to stop from converging and predict all pixels as negative. To obtain binary predictions from the outcome of the model with a linear activation, a suitable threshold value needs to be chosen from the predictions of the training set to convert the test set predictions. For training of the CNNs, we used 50 epochs, Adam from the Keras package as the optimizer, and the binary cross entropy as the loss function with a learning rate of 0.001. The order of data was shuffled and 30% of it was used for validation. The convergence was checked visually by plotting the values of the loss function.

2.4 Algorithm for removing false positive segmentation

Since the fully automatic segmentation sometimes creates a binary mask with multiple separate areas of pixels labelled as positive and, typically, only one of these areas is a tumor and the rest some other regions with high FDG uptake, a post-processing algorithm was designed to included only some of these areas into the final segmentation. This algorithm receives the predicted binary masks and the locations the annotated one-click pixels. If there is more than one positive segment in the binary mask, the algorithm re-labels all such components that do not contain a one-click point as negative. This is accomplished with help of the function *label* in *scipy.ndimage.measurements* (Virtanen et al. 2020).

2.5 Evaluation of the models

To evaluate the performance of different model options, we use the Dice score (D) defined as

$$D = \frac{2\mathrm{TP}}{2\mathrm{TP} + \mathrm{FP} + \mathrm{FN}},$$

where TP is the number of correctly predicted positive pixels, FP is the number of negative pixels predicted as positive, and FN is the number of positive pixels predicted as negative. The Dice score is computed separately for every predicted binary mask by comparing it to the corresponding mask created by a medical doctor. Clearly, the greater the Dice score is the better the predicted mask is, and the number of correctly predicted negative pixels (TN) does not affect the Dice score. To estimate whether the differences in the Dice scores from different model options are statistically significant or not, we use the Wilcoxon signed-rank test, which is a non-parametric alternative to the better-known paired t-test and can be used for non-normally distributed data. This test is also not sensitive to error caused by potential outliers. The F-test of equality of variances is also applied to compare the methods.

2.6 Structure of the experiment

Since the data was split into different training and test sets with five-fold cross-validation, the CNNs were initialized and trained five times to obtain the predictions for each PET slice. During each iteration round, three versions of the same CNN was run to obtain results of fully automatic segmentation, segmentation with information about one random positive pixel, and segmentation with information about the CM of the positive pixels. Afterwards, the results of the fully automatic segmentation were

given to the algorithm with information about one-click pixels. These methods are named here so that CNNonly is the fully automatic segmentation with no post-processing algorithm, CNNrand is the fully automatic segmentation followed by an algorithm that uses the information about the random positive pixels, CNNcm is the fully automatic segmentation with an algorithm using the CMs, rand1CNN is the semi-automatic one-click segmentation based on the random positive pixels, and cm1CNN is the one-click segmentation based on the CMs instead. See the flowchart of Fig. 1.

3 Results

Table 1 contains the median Dices scores of the methods CNNonly, CNNrand, CNNcm, rand1CNN, and cm1CNN, while Table 2 contains the p-values of the Wilcoxon tests comparing the five methods. According to Table 1, the methods are in order cm1CNN, rand1CNN, CNNcm, CNNrand, CNNonly when listed from the best to the worst. This can be also seen from Fig. 2, which contains the boxplots for the Dice scores computed from the binary masks predicted with these methods. It can be computed that the Dice scores of the best method, cm1CNN, are 32.3% higher than those of

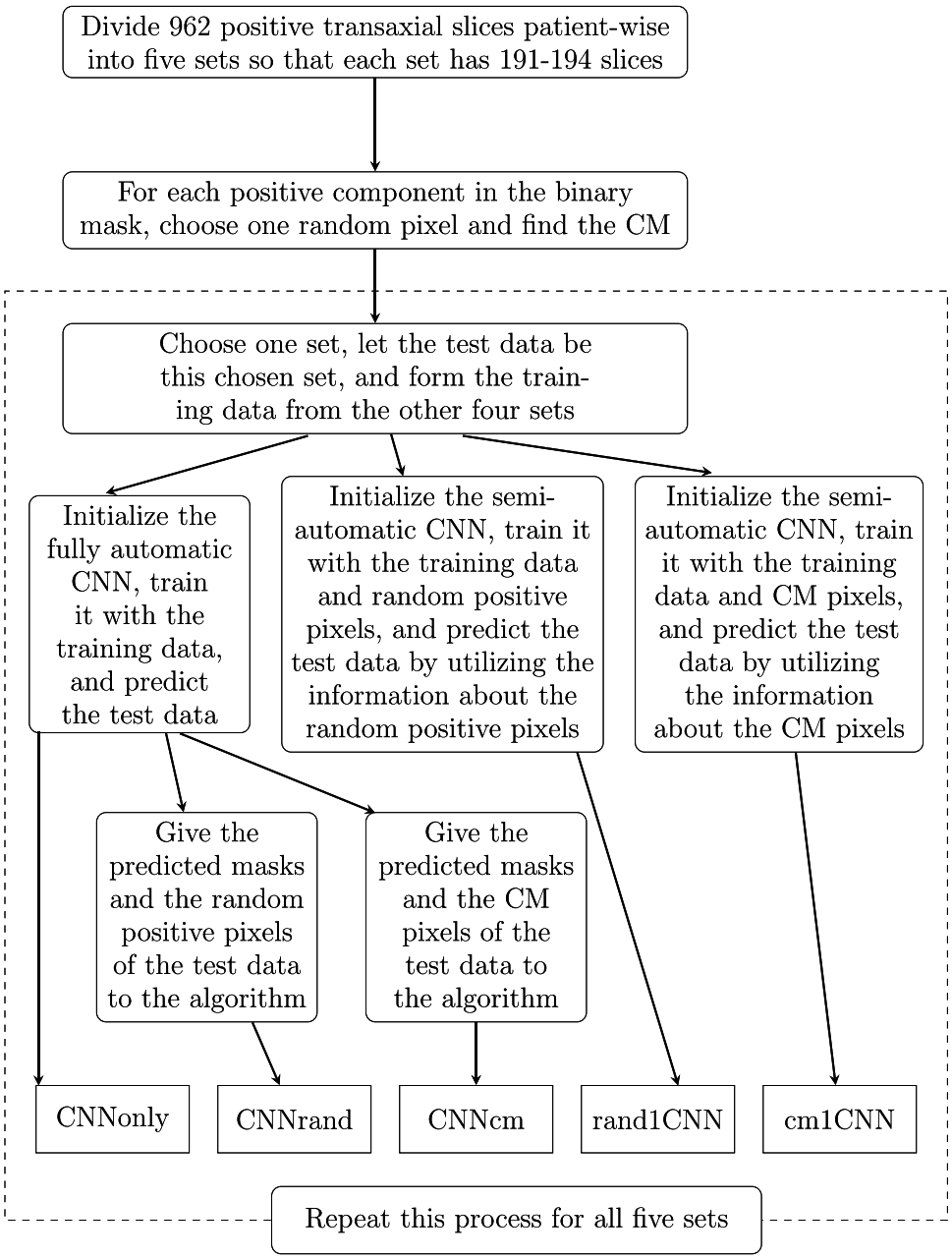


Fig. 1 A flowchart explaining how the predicted binary masks are obtained with our five methods

Table 1 Median Dice scores computed from the predictions of each test set separately and then over all the test sets for each of the five methods. The maximum of each row is in bold

Test set	CNNonly	CNNrand	CNNcm	rand1CNN	cm1CNN
1	0.571	0.593	0.604	0.659	**0.739**
2	0.431	0.484	0.500	0.571	**0.650**
3	0.480	0.525	0.568	0.631	**0.715**
4	0.492	0.525	0.533	0.516	**0.780**
5	0.656	0.681	**0.698**	0.695	0.654
All test sets	0.535	0.563	0.593	0.615	**0.708**

Table 2 The p-values of the Wilcoxon tests comparing the Dice scores computed from predictions of all the five test set with the five methods. All values are statistically significant with 5% level of significance

Methods	CNNonly	CNNrand	CNNcm	rand1CNN
cm1CNN	8.44e−58	5.64e−43	3.52e−35	5.13e−26
rand1CNN	1.82e−28	5.44e−19	2.50e−12	
CNNcm	0.00612	9.09e−07		
CNNrand	0.0487			

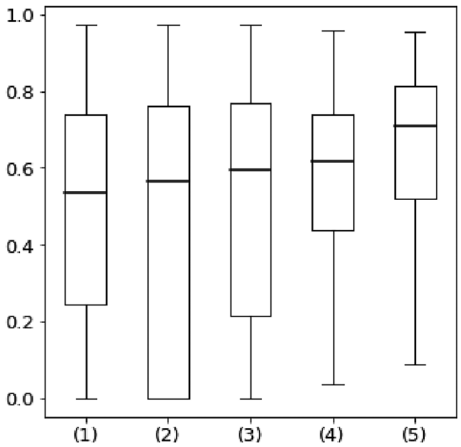


Fig. 2 Boxplots of the Dice scores computed from the predicted binary masks obtained with different methods, including (1) CCNonly, (2) CNNrand, (3) CNNcm, (4) rand1CNN, and (5) cm1CNN

the fully automatic segmentation. Table 2 also reveals that all the differences between the Dice scores are statistically significant.

The standard deviations of CNNonly, CNNrand, CNNcm, rand1CNN, and cm1CNN are 0.293, 0.331, 0.320, 0.196, and 0.205, respectively. The p-values of F-tests of equality of variances for the five methods can be found in Table 3. We see that there is significantly less variation in the Dice scores obtained with cm1CNN and rand1CNN than the other three methods and significantly more variation with methods CNNrand and CNNcm. There are no significant differences between cm1CNN and rand1CNN nor CNNrand and CNNcm.

Table 3 The p-values of the F-tests comparing the Dice scores computed from predictions of all the five test set with the five methods. Statistically non-significant values with 5% level of significance are italicized

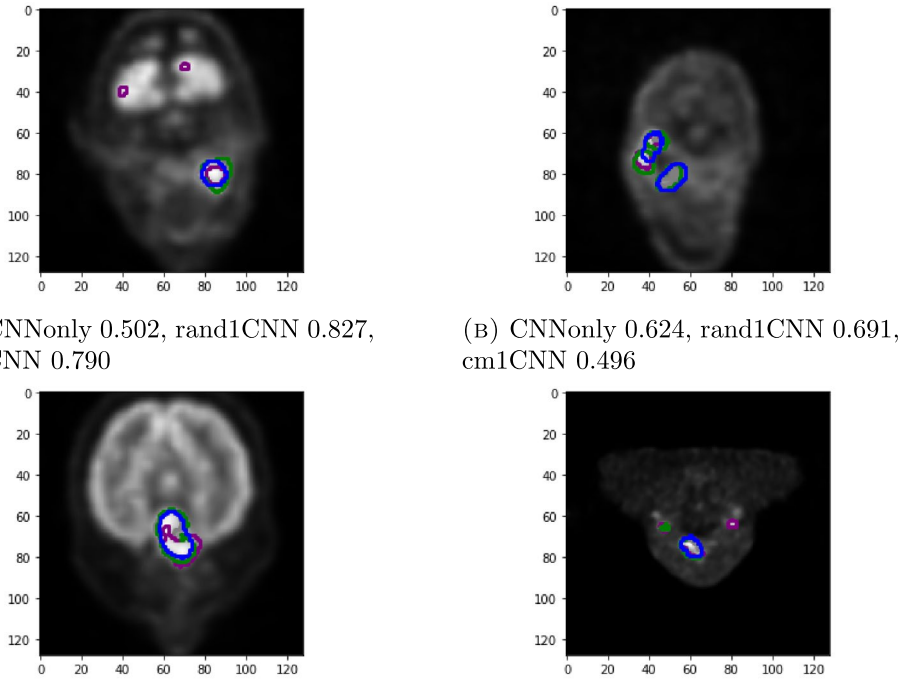
Methods	CNNonly	CNNrand	CNNcm	rand1CNN
cm1CNN	2.22e−16	2.22e−16	2.22e−16	*0.167*
rand1CNN	2.22e−16	2.22e−16	2.22e−16	
CNNcm	0.00572	*0.302*		
CNNrand	0.000149			

To represent the results visually, Fig. 3 has four examples of a transaxial PET slice showing a cancer tumor with the correct outline and the outlines predicted by the methods CNNonly, rand1CNN, and cm1CNN. The method CNNonly produces false positive segmentation caused for the PET slice in 3(A) because of the greater amount of the FDG tracer in the patient's brains than the surrounding image. According to our results, the use of the method CNNcm fixes this issue resulting in a Dice score of 0.541 while CNNrand removes all segmentation (Dice score: 0.00). For the other three examples, there are no differences between these two methods: The Dice scores of both CNNrand and CNNcm are 0.00, 0.618, and 0.838 for 3(B), 3(C), and 3(D), respectively.

4 Discussion

Our results suggest that, by applying the one-click annotation method to denote the locations of the cancer tumors and then giving this information to a semi-automatic CNN, we obtain significantly better predictions of the binary masks compared to a CNN performing fully automatic segmentation. Interestingly, this happens regardless of whether the results of the automatic CNN are fixed afterwards with an algorithm choosing only correct components. The difference between the semi-automatic CNN and the use of algorithm might be caused by the fact that the automatic CNN classifies a smaller number of pixels as positive in the correct region to avoid a large number of FP observations.

According to our results, both the semi-automatic CNNs and the algorithms also perform the better when they have the information about the CMs of all the positive components rather than the random positive pixels. It is likely that the semi-automatic CNN labels non-cancerous pixels as positive if the one-click pixel is a random pixel close to the border of the tumor near those negative pixels. Similarly, our algorithm removes some of the mostly correct components



(A) CNNonly 0.502, rand1CNN 0.827, cm1CNN 0.790

(B) CNNonly 0.624, rand1CNN 0.691, cm1CNN 0.496

(C) CNNonly 0.618, rand1CNN 0.858, cm1CNN 0.853

(D) CNNonly 0.681, rand1CNN 0.739, cm1CNN 0.800

Fig. 3 Four transaxial PET slices with the correct tumor outline in gray and the predicted outlines by the methods CNNonly, rand1CNN, and cm1CNN in purple, green, and blue, respectively. The related Dice scores of these three methods are in the sub-captions, and the Dice scores of CNNrand and CNNcm for these PET slices are in the text

whenever its information is based on a random positive pixel that is near these components but still outside of them.

Compared to similar published studies about tumor segmentation, the median Dice score of 0.708 given by our best method is close to their results. For instance, Ren et al. (2021) obtained average Dice scores ranging 0.72–0.74, Yuan (2021) an average Dice score of 0.732, and Xie and Peng (2021) a Dice score of 0.735 in their research about segmentation of head and neck cancer. While the Dice scores of the four other methods studied here are quite low in comparison, it should be noted that the aim of this study was to compare to different segmentation methods instead of obtaining the best possible predictions. Our methods would very likely perform better if the CNNs could be trained by using PET images of more cancer patients. Namely, our data is very heterogeneous because the positive slices look very different depending on the location of the tumor. Alternatively, we could train the CNNs numerous times and only choose the iteration rounds with the best results in the training set to predict the test set data.

Our research differs from the earlier one-click type approaches presented in Chen et al. (2022); Kirillov et al. (2023); Liu et al. (22021, 2023); Wang et al. (2023, 2023). Namely, they mostly used actual clicks chosen by the user whereas we compared the automatically found one-clicks to evaluate the difference between random choices and choices closer to the center of the object. Additionally, the earlier methods were designed for either typical photographs or medical modalities other than PET and therefore might not be suited for PET images which look blurry in comparison and lack clear boundaries between different areas of the image. Furthermore, fair comparison is difficult as the other methods have been trained for much greater datasets.

Our methods should be studied further. In particular, both the algorithm and the CNNs used here can be easily modified to suit for three-dimensional data. While the three-dimensional segmentation requires significantly more training data, it would offer a more useful application. Also, our one-click annotation technique could be combined to the new segmentation technique in [13], which uses a CNN to choose threshold values for PET slices to classify their pixels. Additionally, while we studied only the standard U-Net here as it is the most common CNN design for tumor segmentation, it could be investigated whether there are any differences in the results between different segmentation CNNs.

5 Conclusion

An semi-automatic CNN introduced here significantly improves the results of of tumor segmentation from the PET images of head and neck cancer patients compared to a fully automatic CNN. Using this approach requires collecting information with a so-called one-click annotation method, in which a user clicks the location of a tumor in

a few transaxial slices of a PET image. With this information, the CNN can then avoid the false positive segmentation outside of the tumor, so that the final mask will not show the tracer substance accumulation in the human brain or other non-cancerous regions. Notably, giving the results of one-click annotation directly to a semi-automatic CNN works significantly better than using a post-processing algorithm with this same information for correcting the binary masks predicted by a fully automatic CNN. In the future research, this topic could be studied for segmentation CNNs other than the U-Net, especially ones processing three-dimensional data.

Acknowledgements The authors are grateful to the referees for their detailed and constructive suggestions.

Funding Open Access funding provided by University of Turku (including Turku University Central Hospital). The first author was financially supported by the Finnish Culture Foundation.

Data availibility Patient data not available due to ethical restrictions.

Code availability Available at https://github.com/rklen/One_Click_CNN_for_HNSCC.

Declarations

Conflict of interest On the behalf of all authors, the corresponding author states that there is no Conflict of interest.

Ethical approval The study was approved by Ethics Committee of the Hospital District of Southwest Finland and the research was performed in accordance with the Declaration of Helsinki.

Informed consent All participants were at least 18 years of age and consented to their data to be used in this research.

References

Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, ..., Zheng X (2015) TensorFlow: Large-scale machine learning on heterogeneous systems

Chen X, Zhao Z, Zhang Y, Duan M, Qi D, Zhao H (2022) Focalclick: Towards practical interactive image segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1300-1309)

Chollet F et al (2015) Keras. GitHub

De Felice F, Musio D, Tombolini V (2015) Follow-Up in Head and Neck Cancer: A Management Dilemma, Advances in Otolaryngology, vol. 2015, Article ID 703450

Global Cancer Observatory (GCO) (2022) Cancer today [Online analysis table]

Hellström H, Liedes J, Rainio O, Malaspina S, Kemppainen J, Klén R (2023) Classification of head and neck cancer from PET images using convolutional neural networks. Sci Rep 13:10528

Johnson DE, Burtness B, Leemans CR, Wai Yan Lui V, Bauman JE, Grandis JR (2020) Head and neck squamous cell carcinoma. Nat Rev Dis Primers 6:92

Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L et al (2023) Segment anything. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 4015-4026)

Liedes J, Hellström H, Rainio O, Murtojärvi S, Malaspina S, Hirvonen J, Klén R, Kemppainen J (2023) Automatic segmentation of head and neck cancer from PET-MRI data using deep learning. Journal of Medical and Biological Engineering (to appear)

Liu Z, Qi X, Fu CW (2021) One thing one click: A self-training approach for weakly supervised 3d semantic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1726-1736)

Liu Q, Xu Z, Bertasius G, Niethammer M (2023) Simpleclick: Interactive image segmentation with simple vision transformers. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 22290-22300)

Rainio O, Han C, Teuho J, Nesterov SV, Oikonen V, Piirola S, Laitinen T, Tättäläinen M, Knuuti J, Klén R (2023a) Carimas: An extensive medical imaging data processing tool for research. J Digit Imaging 36:1885–1893

Rainio O, Lahti J, Anttinen M, Ettala O, Seppänen M, Boström P, Kemppainen J, Klén R (2023b) New method of using a convolutional neural network for 2D intraprostatic tumor segmentation from PET images. Biomed Eng Res 39:905–913

Ren J, Eriksen JG, Nijkamp J, Korreman SS (2021) Comparing different CT, PET and MRI multi-modality image combinations for deep learning-based head and neck tumor segmentation. Acta Oncol 60(11):1399–1406

Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation (pp. 234-241). In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham

van Rossum G, Drake FL (2009) Python 3 Reference Manual. CreateSpace

Townsend DW (2004) Physical principles and technology of clinical PET imaging. Ann Acad Med Singap. 33(2):133–45

Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D et al (2020) SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, 17(3):261-272

Wang T, Li H, Zheng Y, Sun Q (2023) One-Click-Based Perception for Interactive Image Segmentation. IEEE Transactions on Neural Networks and Learning Systems

Wang P, Yao W, Shao J (2023) One Class One Click: Quasi scene-level weakly supervised point cloud semantic segmentation with active learning. ISPRS J Photogramm Remote Sens 204:89–104

Yuan Y (2021) Automatic Head and Neck Tumor Segmentation in PET/CT with Scale Attention Network. In: V. Andrearczyk, V. Oreiller, A. Depeursinge (eds) Head and Neck Tumor Segmentation. HECKTOR 2020. Lecture Notes in Computer Science, vol 12603. Springer, Cham

Xie J, Peng Y (2021) The Head and Neck Tumor Segmentation Using nnU-Net with Spatial and Channel 'Squeeze & Excitation' Blocks. V. Andrearczyk, V. Oreiller, A. Depeursinge (eds) Head and Neck Tumor Segmentation. HECKTOR 2020. Lecture Notes in Computer Science, vol 12603. Springer, Cham