ORIGINAL ARTICLE

Can 3D artificial intelligence models outshine 2D ones in the detection of intracranial metastatic tumors on magnetic resonance images?

Ying-Chou Sun[a,b], Ang-Ting Hsieh[c], Ssu-Ting Fang[d], Hsiu-Mei Wu[a,b], Liang-Wei Kao[d], Wen-Yuh Chung[e,f], Hung-Hsun Chen[g], Kang-Du Liou[e,f], Yu-Shiou Lin[c], Wan-Yuo Guo[a,b], Henry Horng-Shing Lu[c,d,h,*]

[a]Department of Radiology, Taipei Veterans General Hospital, Taipei, Taiwan, ROC; [b]School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan, ROC; [c]Institute of Data Science and Engineering, National Yang Ming Chiao Tung University, Hsinchu, Taiwan, ROC; [d]Institute of Statistics, National Yang Ming Chiao Tung University, Hsinchu, Taiwan, ROC; [e]Division of Functional Neurosurgery, Taipei Veterans General Hospital, Taipei, Taiwan, ROC; [f]Institute of Neurological, Yang Ming Chiao Tung University, Taipei, Taiwan, ROC; [g]Center of Teaching and Learning Development, National Yang Ming Chiao Tung University, Hsinchu, Taiwan, ROC; [h]Department of Medical Research, Taipei Veterans General Hospital, Taipei, Taiwan, ROC

Abstract

Background: This study aimed to compare the prediction performance of two-dimensional (2D) and three-dimensional (3D) semantic segmentation models for intracranial metastatic tumors with a volume ≥ 0.3 mL.

Methods: We used postcontrast T1 whole-brain magnetic resonance (MR), which was collected from Taipei Veterans General Hospital (TVGH). Also, the study was approved by the institutional review board (IRB) of TVGH. The 2D image segmentation model does not fully use the spatial information between neighboring slices, whereas the 3D segmentation model does. We treated the U-Net as the basic model for 2D and 3D architectures.

Results: For the prediction of intracranial metastatic tumors, the area under the curve (AUC) of the 3D model was 87.6% and that of the 2D model was 81.5%.

Conclusion: Building a semantic segmentation model based on 3D deep convolutional neural networks might be crucial to achieve a high detection rate in clinical applications for intracranial metastatic tumors.

Keywords: Area under the curve; Deep learning; Neural networks; Semantic

1. INTRODUCTION

Magnetic resonance imaging (MRI) is currently the most valuable diagnostic imaging tool in hospital settings. Its application for the central nervous system imaging (ie, the brain and spine) is of particular importance. One of the most common clinical applications of MRI is imaging of the brain for detecting metastatic tumor after a patient is diagnosed with a primary cancer (eg, lung cancer). The presence or absence of brain metastatic tumor is essential for defining cancer extension and staging, which are used for determining the subsequent treatment strategy. However, interpreting magnetic resonance (MR) images of the brain is a time- and experience-demanding task. Once brain metastasis occurs in a cancer patient, radiation treatment of the brain lesion (radiosurgery or whole brain radiotherapy) is usually considered as treatment-of-recommendation. Therefore, we conducted this study to meet the clinical need of cancer staging and treatment strategy making. Cutting-edge deep convolutional neural network (CNN) segmentation was used in this study to assist physicians in detecting brain metastatic tumors.

In recent research, a multimodal brain tumor image segmentation benchmark was presented.[1] Pereira et al[2] proposed a novel CNN-based method for segmenting brain tumors in MR images. Mohseni Salehi et al[3] used a fully convolutional network architecture based on the U-Net architecture, which was the inspiration for this study.

Three-dimensional (3D) deep learning models have been widely discussed. The segmentation performance of two deep learning methods that used two-dimensional (2D) and 3D CNNs were compared on computed tomography (CT) images.[4] Tran et al[5] demonstrated the advantages of 3D CNNs over 2D CNNs with residual learning. They also observed that factorizing the 3D convolutional filters into separate spatial and temporal components yielded significantly increased accuracy. The robustness of the application of 3D models was also proved. A 3D multiview CNN with both the chain and directed acyclic graph architectures was introduced by Kang et al.[6]

Accurate and automatic brain metastases segmentation is a key step for efficient and effective stereotactic radiosurgery treatment planning. Liu et al[7] developed a deep learning CNN algorithm for segmenting brain metastases on contrast-enhanced T1-weighted MR imaging datasets. For automatic segmentation tool, a study presents a computational tool for

*Address correspondence. Dr. Henry Horng-Shing Lu, Institute of Statistics, National Yang Ming Chiao Tung University, 4F, Assembly Building 1, 1001, University Road, Hsinchu 300, Taiwan, ROC. E-mail address: hslu@stat.nycu.edu.tw (H. H.-S. Lu).

Conflicts of interest: The authors declare that they have no conflicts of interest related to the subject matter or materials discussed in this article.

Journal of Chinese Medical Association. (2021) 84: 956-962.

Received October 30, 2020; accepted March 30, 2021.

doi: 10.1097/JCMA.0000000000000614.

autosegmenting the distribution of brain infusions observed by MR imaging using a Gaussian Mixture Model to efficiently classify pixels as belonging to either the high-intensity infusate or low-intensity background.[8] Automatic classification also plays an important role in computer-aided diagnosis. Cheng et al[9] introduced a study focusing on the classification of three types of brain tumors (ie, meningioma, glioma, and pituitary tumor) in T1-weighted contrast-enhanced MR images.

Numerous studies have used 2D images to construct semantic segmentation models.[10,11] However, 2D image segmentation models do not entirely use the relevance and spatial information between neighboring slices, whereas 3D models do. In general, 2D models extract less helpful information than 3D models do. Three-dimensional models enable a higher detection rate in clinical application than 2D models do.

In this study, we concentrated on the semantic segmentation of clinical images and used both 2D and 3D U-net. The 2D U-Net[11] is a fully convolutional semantic segmentation model. The model consists of two phases, namely the contracting and expansive phases. In the contracting phase, the model extracts high-level representation and denoises the image features by different convolutional filters and down-sampling mechanisms. In the expansive phase, the model considers low-level high-resolution feature maps to perform up-sampling. The 3D U-Net[12] is an expansion of 2D U-Net to consider the 3D spatial correlation. We took advantage of 2D and 3D U-net to construct the 2D and 3D intracranial metastasis segmentation systems by implementing various preprocessing and data augmentation techniques. We compared the experimental results of the 3D U-Net, 3D Regularization U-Net, 2D U-Net, and 2D Regularization U-Net by considering the area under the curve (AUC) of the precision-recall (PR) curve.

2. METHODS

We separated the entire procedure into three parts: dataset partition and description, image preprocessing and data augmentation, and deep learning model training. For the first part, the entire dataset was partitioned into the training, validation, and testing sets through stratified sampling. Stratified sampling is based on the tumor volume in each patient, therefore the tumor size distribution between training, validation, testing sets will be the same. Next, we pruned the outliers to generate suitable grayscale images. Thus, we could use suitable enhancements to crop out meaningful information from the images. Finally, we treated the U-Net as the basic model in this study. The overall flowchart is shown in Fig. 1.

2.1. Dataset Description

This study was approved by the institutional review board (IRB) of Taipei Veterans General Hospital (TVGH). Informed consent was waived by the IRB because in this study, secondary use of existing MRI was made with suitable deidentification. The MRIs were originally acquired for guiding radiosurgical treatment using the Gamma Knife. For this therapeutic need, stereotactic frame was applied and fixed on patient's skull throughout the whole procedures, from imaging acquiring to treatment completeness. The stereotactic manner of imaging data acquisition prevented patient from head movement during MR scanning and made the stereotactic MR imaging superior to nonstereotactic ones. The MRIs were performed on 1.5T MR scanners. The imaging protocol included precontrast agent intravenous administration axial T1-weighted (T1W), T2-weighted (T2W), and coronal T1-weighted, and postcontrast agent intravenous administration coronal T1-weighted and axial T1-weighted imaging (T1W + C). Each image had a size of 512 × 512 pixels within a field-of-view of 260 × 260 mm; each pixel covered approximately 0.5 × 0.5 mm; and each slice had a thickness of 3 mm. Tumor annotation of the metastases has been based on information integration of multiparametric MRI (ie, T1W, T1W + C, and T2W). The precise tumor locations and contours were cautiously marked on T1W + C MRI based on the integration and verified by the neuroradiologists, neurosurgeons, and radiation oncologists involved in conducting the radiosurgery. The diagnosis of metastases is based on imaging diagnosis. In addition to the tumor contours marked by the doctors, there are also skull contours (a rough shape of skull) generated through proprietary software and methods. The skull contours were generated by the built-in function of Leksell GammaPlan, relying

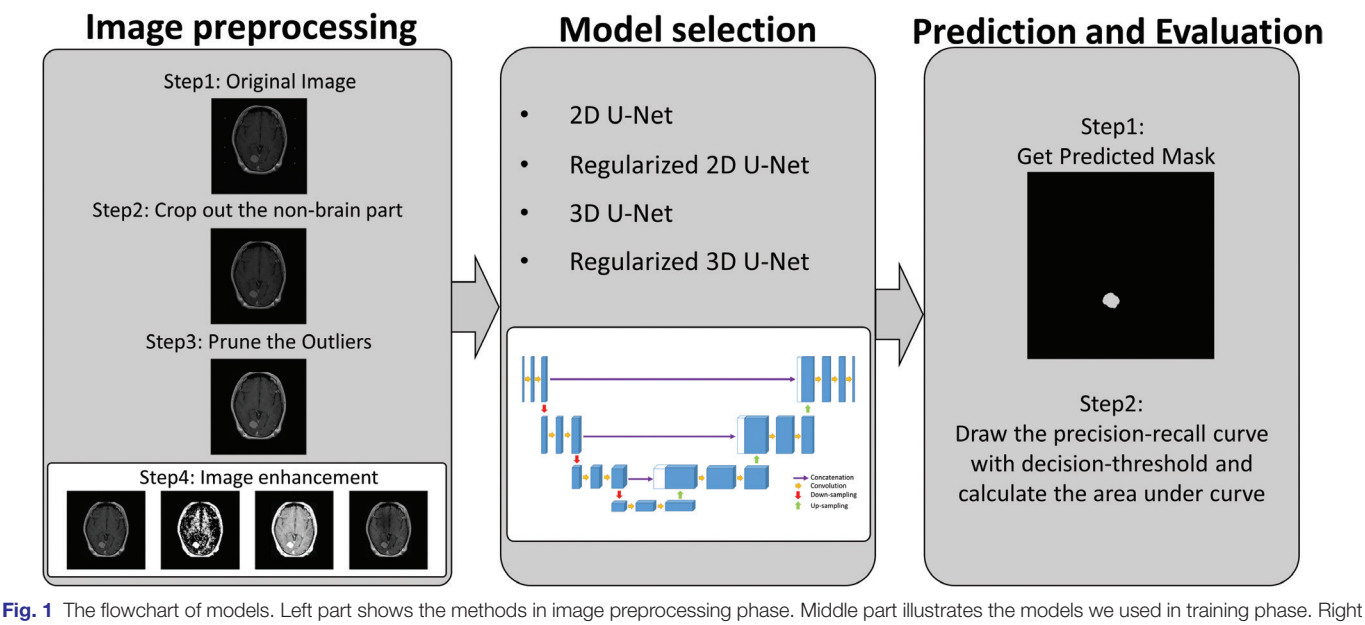

Fig. 1 The flowchart of models. Left part shows the methods in image preprocessing phase. Middle part illustrates the models we used in training phase. Right part shows the schematic diagram of prediction and verification method. 2D = two-dimensional; 3D = three-dimensional.

on manual measurement of multiple measuring points using a dedicated skull measurement sphere and a special ruler that measures the distance to the scalp.[13] These images have been stored in the digital imaging and communications in medicine (DICOM) format. The contours with annotations were all stored in Radiotherapy Structure Set (RTSS) format of DICOM standard.

In this study, we only recruited MR images of patients who suffered from naive brain metastasis, 450 patients, including 285 patients in training set, 72 patients in validation set, and 93 patients in testing set. We only used axial T1W + C whole-brain MR for deep learning training. The size of each tumor was ≥0.3 mL. Doctors can recognize tumors whose sizes are equal to or larger than 1 cm^2 on MR images without missing, and the thickness of each MR image is 0.3 cm. Therefore, we set 0.3 mL as the base point in the size comparison experiment. According to the tumor size of testing set patients, it is divided into four categories, by using the quartile division method. The tumor size quartile of testing data is 0.62, 1.37, and 4.55 mL. As a result, the four ranges are (1) ≥0.3 mL, (2) ≥0.62 mL, (3) ≥1.37 mL, and (4) ≥4.55 mL. We named these ranges as groups I to IV, respectively, for further presentation. Group I includes group II to IV, group II includes group III to IV, and group III includes group IV. This study used these four ranges of tumor volume to verify the precision and recall of each category.

2.2. Image Preprocessing and Data Augmentation

In this research, we used the following image preprocessing methods, including pruning the outliers, cropping, enhancement, contrast-limited adaptive histogram equalization (CLAHE), and normalization. On the other hand, we adjusted the image size after cropping by resize and interpolation method to set the image size to 256 * 256 and 128 * 128 * 128 for 2D and 3D models.

First, we pruned the outliers based on the apparent brightness variations in some images. Balan et al[14] clustered histograms into several groups. The different groups represented different tissues. For this reason, we considered the extremely large grayscale levels in the histogram as outliers. To make a robust estimation and inference, the outlier pixels were removed. In Fig. 2, the upper two rows illustrate the grayscale image of the original MR images and the corresponding histograms. The lower two rows illustrate the modified MR images and the corresponding histograms after modification. The brightness variation between different slices in the modified images was considerably lower than that in the original MR images.

Second, due to the nonstandardize brain sizes, we had to crop and position the images in a specific location and create a standardize brain images. Therefore, we used the contours with annotation of "*Skull" in the RTSS files to crop out the nonbrain part of the MR images, for example, the imaging signal of the skull positioning frame.

Next, we carried out the process of enhancing the images. The main aim of enhancement is to select distinct observation windows. We can specify a specific observation window by defining the level and window. The grayscale levels outside the observation window are set to the upper bound or lower bound, whichever is closer to them. In this study, two enhancement techniques, namely "Window-Level I" and "Window-Level II," produce the enhanced image by normalization in Equation (1) and Look Up Table (LUT) Value in Equation (2).[15] The value of w/l are 0.7/1.8 and 3.7/1 for Window-Level I and Window-Level II, respectively. The results including original image are presented in Fig. 3A–C.

$$z = \frac{x - \bar{x}}{S} \tag{1}$$

$$\text{LUT Value} = \begin{cases} 0, & \text{if } z \le (\ell - 0.5 - (w-1)/2) \\ 255, & \text{if } z > (\ell - 0.5 + (w-1)/2) \\ \left(\dfrac{z - (\ell - 0.5)}{w - 1} + 0.5\right) \times 255, & \text{others} \end{cases} \tag{2}$$

where x is the pixel array for a slice image, $\bar{x}$ is the mean of x, S is the standard deviation of x, l, and w are constant that determinate brightness of the image, and LUTValue is a grayscale Look-Up Table for the image.

In addition to adjusting the brightness of the image through the enhancement methods introduced above, we also perform CLAHE on the image. Because of the inevitable skew in the pruned grayscale histograms (the bottom two rows of Fig. 2), we used histogram equalization (HE) to increase the global contrast of the images for improving the distribution of the grayscale levels. In adaptive HE (AHE), a square window is used to increase the local contrast. However, the problem of "noise amplification in near-constant regions" appears when we limit HE to a specific location. To deal with this issue, CLAHE[16] is utilized to clip and redistribute the histogram iteratively. Thus, we can obtain images with high contrast by performing HE on the modified histograms. Fig. 3D displays the result after performing CLAHE on the original images in Fig. 3A.

At last, we normalized the images. When training the model, normalization considerably influences the overall performance.[17] In our data, the grayscale range was set by the DICOM "Bits Stored" tag to represent the range from 0 to 2^(Bits Stored) – 1, where "Bits Stored" is usually set as 16. For convenience, we normalized the grayscale range as –1 to +1.

For the training of 2D and 3D models, we added some augmentation techniques such as zooming, rotation, and panning in training data.[18] Also, we erased the selected cuboids randomly to create new input data and added random noise to make the model invariant to external noises.[19]

2.3. Deep Learning Model and Training

Semantic segmentation is involved in numerous computer vision tasks. Olafm et al[11] introduced a U-shaped architecture to construct a 2D model for the semantic segmentation of biomedical images. A fully CNN was designed for the segmentation of entire images instead of using the sliding window convolutional method.[20] The model performed well and was sufficiently fast to win a competition held at the 2015 International Symposium on Biomedical Imaging.

In this study, 2D U-Net model and 3D U-Net model for segmentation were constructed. The training of 2D model only employed the slices with lesions. Therefore, the 2D model treated all images independently. Because the model ignored the correlation between two continuous slices, it could not reflect the spatial relationship of adjacent images. In contrast, the 3D model took the images of whole brain and extracted the correlated information between slices. To build an isotropic brain for 3D segmentation, linear interpolation is used. We performed data augmentation and L2 regularization in 2D U-Net and 3D U-Net to assure the generalizability of the model and prevent overfitting.

The expansive part combines the high-level representation with low-level high-resolution feature maps to create overall consideration. We implemented up-sampling blocks

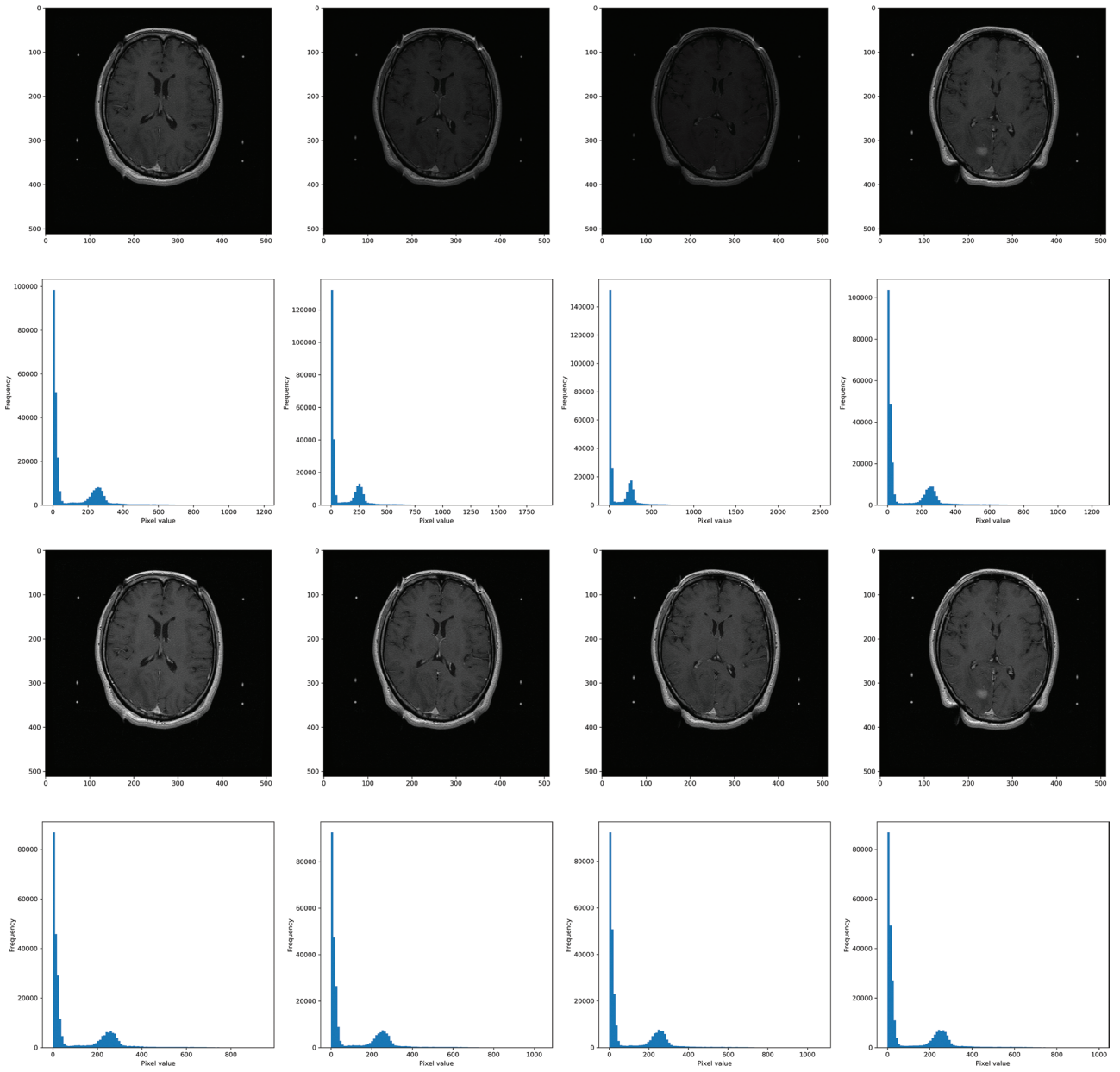

Fig. 2 Comparison between the original and modified MR images. MR = magnetic resonance.

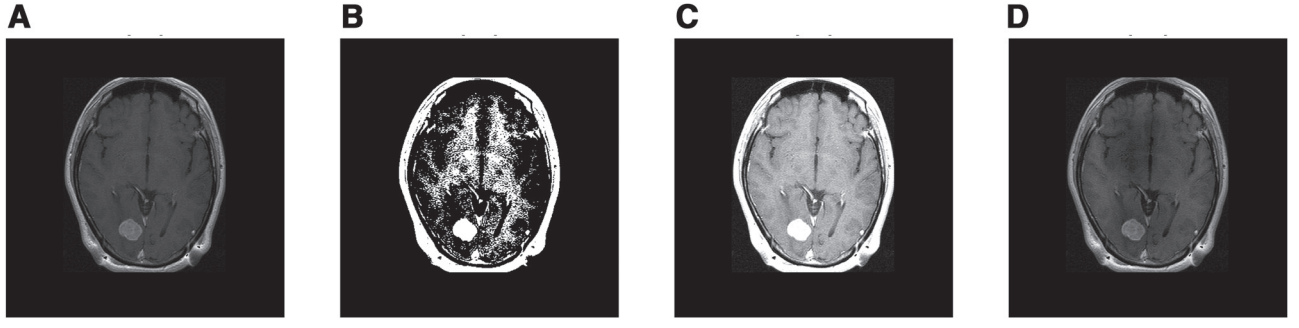

Fig. 3 Illustration of the brain images by different enhancement techniques. A, Brain images with "none" enhancement; (B) brain images with "Window-Level I" enhancement; (C) brain images with "Window-Level II" enhancement; and (D) brain images with "CLAHE" preprocessing. CLAHE = contrast-limited adaptive histogram equalization.

to reconstruct feature maps that match to the contracting part. Finally, we created a probabilistic mask by applying the sigmoid function to all the pixel values in the last layer. Experiment was run in an NVIDIA GEFORCE GTX 1080 Ti GPU environment.

2.4. Evaluation

This article focuses on detection rate. Hence, we use AUC produced by PR curve as the standard for comparing the quality of the model. The predicted masks are obtained from the model prediction results which values are between 0 and 1. Therefore, we set a "decision-threshold" and transform the image pixel value to 0 or 1. The tumor has been detected as long as an intersection exists between the prediction pixels/voxels and the ground truth. On the other hand, "True Negative" is not defined in the segmentation task. Final comparisons are conducted by monitoring precision and recall.

3. RESULTS

There were two main results in this study, one for the presentation of the curve results of the 3D model training, and the other for the comparison of the ability to predict the tumors between different models.

The AUC of the PR curve considers the precision and recall under different decision-threshold according to the existing training differences between 2D and 3D models (3D models focus on the entire brain, whereas 2D models only focus on the slices with tumors.). As displayed in Fig. 4, four curves were obtained, namely the yellow (2D U-Net), pink (regularized 2D

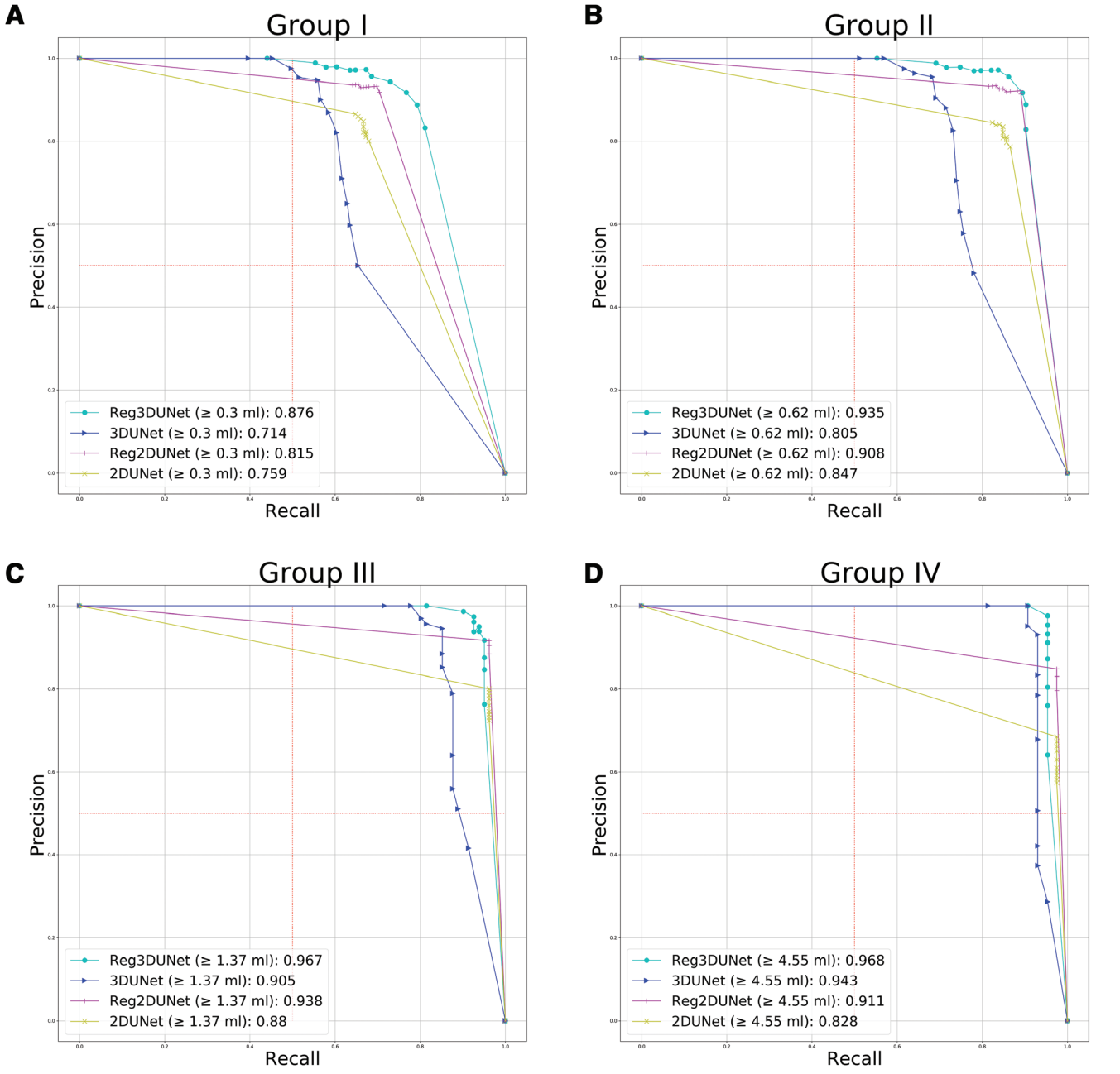

Fig. 4 AUC for different models and different tumor sizes. A, Group I (≥0.3 mL), (B) group II (≥0.62 mL), (C) group III (≥1.37 mL), (D) group IV (≥4.55 mL). AUC = area under the curve.

Table 1

AUC ratio for each model with different tumor size ranges

	Model			
Interval	3D U-Net with regularization (%)	3D U-Net (%)	2D U-Net with regularization (%)	2D U-Net (%)
Group I: >0.3 cc	87.6	71.4	81.5	75.9
Group II: >0.62 cc	93.5	80.5	90.8	84.7
Group III: >1.37 cc	96.	90.5	93.8	88.0
Group IV: >4.55 cc	96.8	94.3	91.1	82.8

2D = two-dimensional; 3D = three-dimensional; AUC = area under the curve.

U-Net), blue (3D U-Net), and light blue (regularized 3D U-Net) curves. The AUC ratio for each model with different tumor size ranges are displayed in Table 1.

Fig. 4A indicates that the AUC of the regularized 3D U-Net (0.876) is higher than the 3D U-Net (0.714), the regularized 2D U-Net (0.815), and the 2D U-Net (0.759). Compared to other three PR curves, the PR curve of the regularized 3D U-Net shows the dense concentration points on the top right corner which implies the stable prediction characteristic of the regularized 3D U-Net.

Fig. 4B–D displays the results for the other three experiments with different tumor sizes. Each part of the figure comprises four curves. The regularized 3D U-Net model has best performance in each tumor size group. Another general trend indicates better predictions in larger tumors for all models.

By the results above, in the case that tumor size is larger than 0.3 mL, we conclude that the prediction ability of the 3D model is better than 2D model and the regularized 3D U-Net has the best performance.

4. DISCUSSION

First, the different models were compared in terms of AUC of PR curve. Fig. 4A indicates that, with regularization, the AUC of the 3D U-Net (87.6%) is higher than that of the 2D U-Net (81.5%). Therefore, a 3D representation of an individual's brain enables a higher tumor detection rate than 2D image slices do.

The regularized and original 3D models are compared in this paragraph. The PR curves shift from left to right. (Blue and light blue curves displayed in Fig. 4A as the AUC increases from 0.714 to 0.876.) By comparing the PR curves, the AUC of the regularized 3D U-Net shows the dense concentration points on the top right corner. Thus, the regularized 3D U-Net is a more reliable model than the original 3D U-Net. To conclude, the experimental results indicate that among the models considered, the regularized 3D U-Net is the optimal one.

Next, we compared the performance of the 2D U-Net, regularized 2D U-Net, 3D U-Net, and regularized 3D U-Net models for different tumor sizes. Due to the intrinsic training differences between the 2D and 3D models, we compared the AUC of the PR curve, which indicates the precision and recall under different decision-threshold.

Fig. 4 indicates that our models have favorable AUC scores for large tumors. Moreover, the regularization term improves the performance of the 3D/2D model for all tumor sizes. The experimental results indicate that the AUC scores of group III and group IV are very close. However, compared with the results for group II and group I, the AUC score of the regularized 3D model improved significantly. Consequently, our model exhibited the best performance for all tumor sizes in the conducted experiments.

Last, this study has limitations. To develop the parsimonious model with the limited amount of data, we downsized the original 512*515 planar resolution to 256*256 for 2D models, and to 128*128 for 3D models; besides, we used only T1W + C images of the MR dataset, instead of using all series of images (T1W + C, T1W, and T2W). Further research could investigate the appropriate model to utilize multiparametric images with the original resolution to improve the performance of tumor detection.

In conclusion, our results indicate that the 3D U-Net outperforms the 2D U-Net in all groups in terms of the AUC of the PR curve with regularization. Moreover, a regularized model is more stable than a model without the regularization term in both 2D and 3D model. Comparing the PR curves, the AUC of the regularized 3D U-Net shows the dense concentration points on the top right corner (Fig. 4A). Thus, the 3D model appears more reliable than the 2D model. Finally, the highest AUC of the PR curve in our experiments was 0.876, for all tumors larger than 0.3 mL.

To conclude, with the help of deep learning-based architecture, we generated a regularized 3D U-Net model which has the best performance comparing to the other three models discussed above. It has reliable ability in the detection of the tumors larger than 0.3 mL.

ACKNOWLEDGMENTS

The brain tumor images used in this study were provided by TVGH. This study is primarily supported by the Ministry of Science and Technology, Taiwan, ROC (grant numbers MOST 106-2634-F-075-001, MOST 107-2634-F-075-001, MOST 108-3011-F-075-001, MOST 107-2118-M-009-006-MY3) as well as TVGH (grant number V109E-008-5_110).

We are grateful to the National Center for High-performance Computing for computer time and facilities.

REFERENCES

1. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans Med Imaging* 2015;**34**:1993–2024.
2. Pereira S, Pinto A, Alves V, Silva CA. Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Trans Med Imaging* 2016;**35**:1240–51.
3. Mohseni Salehi SS, Erdogmus D, Gholipour A. Auto-context convolutional neural network (Auto-Net) for brain extraction in magnetic resonance imaging. *IEEE Trans Med Imaging* 2017;**36**:2319–30.
4. Zhou X, Yamada K, Kojima T, Takayama R, Wang S, Zhou X, et al. Performance evaluation of 2D and 3D deep learning approaches for automatic segmentation of multiple organs on CT images. Proc. SPIE 10575, Medical Imaging, February 27, 2018; Houston, TX: Computer-Aided Diagnosis, 105752C. https://doi.org/10.1117/12.2295178.
5. Tran D, Wang H, Torresani L, Ray J, LeCun Y, Paluri M. A closer look at spatiotemporal convolutions for action recognition. *CVPR* 2018;**1**:2.
6. Kang G, Liu K, Hou B, Zhang N. 3D multi-view convolutional neural networks for lung nodule classification. *PLoS One* 2017;**12**:e0188290.
7. Liu Y, Stojadinovic S, Hrycushko B, Wardak Z, Lau S, Lu W, et al. A deep convolutional neural network-based automatic delineation strategy for multiple brain metastases stereotactic radiosurgery. *PLoS One* 2017;**12**:e0185844.

8. Rosenbluth KH, Gimenez F, Kells AP, Salegio EA, Mittermeyer GM, Modera K, et al. Automated segmentation tool for brain infusions. *PLoS One* 2013;**8**:e64452.
9. Cheng J, Huang W, Cao S, Yang R, Yang W, Yun Z, et al. Enhanced performance of brain tumor classification via tumor region augmentation and partition. *PLoS One* 2015;**10**:e0140381.
10. He K, Gkioxari G, Dollar P, Girshick R. Mask R-CNN. IEEE International Conference on Computer Vision, October 22–29, 2017; Venice, Italy: 2980–8.
11. Olafm R, Philipp F, Thomas B. U-net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, editors. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Cham: Springer. https://doi.org/10.1007/978-3-319-24574-4_28.
12. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-net: learning dense volumetric segmentation from sparse annotation. *Med Image Comput Comput Assist Interv* 2016:424–32.
13. Nakazawa H, Komori M, Mori Y, Hagiwara M, Shibamoto Y, Tsugawa T, et al. Effect of skull contours on dose calculations in Gamma Knife Perfexion stereotactic radiosurgery. *J Appl Clin Med Phys* 2014;**15**:4603.
14. Balan AG, Traina AJ, Ribeiro MX, Marques PM, Traina C Jr. Smart histogram analysis applied to the skull-stripping problem in T1-weighted MRI. *Comput Biol Med* 2012;**42**:509–22.
15. Panchal A, Couture G, Galler N, Nakamoto H, et al. dicompyler/dicompyler-core v0.5.5 (Version v0.5.5). Zenodo; 2019. http://doi.org/10.5281/zenodo.3236628.
16. Reza AM. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *J VLSI Signal Process Syst Signal Image Video Technol* 2004;**38**:35–44.
17. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: *International Conference on Machine Learning*, July 6–11, 2015; Lille, France: 448–56.
18. Schaefer S, McPhail T, Warren J. Image deformation using moving least squares. *ACM TOG* 2006; **25**:533–40.
19. Zhong Z, Zheng L, Kang G, Li S, Yang Y. Random erasing data augmentation. 2017. https://arxiv.org/abs/1708.04896.
20. Szegedy C, Toshev A, Erhan D. Deep neural networks for object detection. *Adv Neural Inf Process Syst* 2013:2553–61. Available at https://papers.nips.cc/paper/2013/hash/f7cade80b7cc92b991cf4d2806d6bd78-Abstract.html.