

Article

Segmentation Method of Cerebral Aneurysms Based on Entropy Selection Strategy

Tingting Li¹, Xingwei An^{1,2,*} , Yang Di¹, Jiaqian He¹, Shuang Liu^{1,2} and Dong Ming^{1,2,3}

¹ Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin 300110, China; ltt1999@tju.edu.cn (T.L.); dy_aquarius@tju.edu.cn (Y.D.); hhhhhjq@tju.edu.cn (J.H.); shuangliu@tju.edu.cn (S.L.); richardming@tju.edu.cn (D.M.)

² Tianjin Center for Brain Science, Tianjin 300110, China

³ Department of Biomedical Engineering, School of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin 300110, China

* Correspondence: anxingwei@tju.edu.cn

Abstract: The segmentation of cerebral aneurysms is a challenging task because of their similar imaging features to blood vessels and the great imbalance between the foreground and background. However, the existing 2D segmentation methods do not make full use of 3D information and ignore the influence of global features. In this study, we propose an automatic solution for the segmentation of cerebral aneurysms. The proposed method relies on the 2D U-Net as the backbone and adds a Transformer block to capture remote information. Additionally, through the new entropy selection strategy, the network pays more attention to the indistinguishable blood vessels and aneurysms, so as to reduce the influence of class imbalance. In order to introduce global features, three continuous patches are taken as inputs, and a segmentation map corresponding to the central patch is generated. In the inference phase, using the proposed recombination strategy, the segmentation map was generated, and we verified the proposed method on the CADA dataset. We achieved a Dice coefficient (DSC) of 0.944, an IOU score of 0.941, recall of 0.946, an F2 score of 0.942, a mAP of 0.896 and a Hausdorff distance of 3.12 mm.

Keywords: segmentation; cerebral aneurysm; Transformer; 2D CNN; entropy



Citation: Li, T.; An, X.; Di, Y.; He, J.; Liu, S.; Ming, D. Segmentation Method of Cerebral Aneurysms Based on Entropy Selection Strategy. *Entropy* **2022**, *24*, 1062. <https://doi.org/10.3390/e24081062>

Academic Editor: Francesco Carlo Morabito

Received: 18 June 2022

Accepted: 28 July 2022

Published: 1 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cerebral aneurysms occur in about 3% of the general population. With the development of neuroimaging, an increasing number of cerebral aneurysms are incidentally discovered [1]. A cerebral aneurysm is a pathological dilation of an intracranial blood vessel whose walls may be abnormally weak and prone to rupture. The rupture of aneurysms causes hemorrhage to the subarachnoid space surrounding the brain, and sometimes in the brain parenchyma [2]. Aneurysm size, shape and location are important factors of rupture [3]. Some traditional methods for cerebral aneurysms are based on statistical thresholding [4] and deformable models [5]. Linear convolution is applied to image processing [6]. The use of geometrically deformable models within a level-set framework is an automated segmentation technique for cerebral aneurysms, and these models' ability to handle topological changes and adapt to complex structural shapes makes them well suited to automated segmentation of complex vascular structures [7]. These methods take lots of time and effort. Therefore, we need accurate and rapid automatic algorithms for the segmentation of aneurysms.

The development of artificial intelligence (AI)-based technologies in medicine is advancing rapidly, and AI has recently experienced an era of explosive growth across many industries—the healthcare industry is no exception. Research in multiple medical specialties has used AI to mimic the diagnostic capabilities of doctors [8–11]. Recent advances in deep learning [12] have made it possible to realize this idea. In this regard, convolutional

neural networks (CNNs) [13] have been the most ground-breaking addition, which are dominating the field of computer vision. CNNs have also revolutionized semantic segmentation tasks. In medical image analysis, the novel CNN architecture is U-Net. Other authors have also built derivatives of the U-Net architecture [14]. U-Net comprises an encoder and a decoder. U-Net has shown impressive potential in segmenting medical images, even with a lack of labeled training data, to the extent that it has become the de facto standard in medical image segmentation [14]. Wasiq et al. proposed a coarse-to-fine method for locating pupils and eye center estimation by combining machine learning and image processing [15].

U-Net-based networks have become popular in medical image segmentation. MA-Unet [16] extracts multiscale features and combines local features with their corresponding global dependencies by attention mechanisms. Isensee et al. proposed nnU-Net, a deep learning framework that can automatically adjust the necessary relevant parameters according to the characteristics of the dataset [17]. Milletari et al. proposed a 3D variant of the U-Net architecture called V-Net, a fully convolutional neural network based on volumetrics [18]. Despite the inspiring results achieved, several issues exist in the developed approaches. For 2D networks, 2D inputs do not fully exploit the 3D image information [19–21]. However, 3D convolutions do not focus on the different in-plane and depth resolutions [21–23]. Therefore, in order to obtain the dependencies between channels, balancing between 2D and 3D is the key to further improving network performance. To solve this problem, H-DenseUNet [24] was proposed. It consists of a 2D DenseUNet for extracting intra-slice features and a 3D counterpart for hierarchically aggregating volumetric contexts for liver and tumor segmentation. Although it combines the advantages of 2D and 3D, it is not suitable for aneurysm segmentation with class imbalance, and separate modeling of intra-slice and inter-slice features will exacerbate class imbalance.

Attention mechanisms have recently become popular in computer vision. Instead of compressing the entire image or sequence into a static representation, attention allows the model to focus on the most relevant features as needed. Transformers are the focus of natural language processing. Although their impact has been limited in vision applications, an increasing number of methods with attention mechanisms are being proposed. Due to the limitations of convolutions, researchers tried to introduce Transformers in both the encoder and decoder. Xu et al. used LeViT as the encoder and passed the multiscale feature map to the decoder through skip connections, which achieved better performance in medical image segmentation [25]. Transformer networks in computer vision can be found in [26]. The Vision Transformer (ViT) [27] adapts Transformer models for computer vision applications.

Deep learning methods with good performance have recently been proposed to segment cerebral aneurysms [28,29]. Due to the class imbalance of medical images, the U-Net framework will cause false negative predictions. The lack of labeled medical images is also a big challenge. Feng et al. proposed a patch-based fully CNN architecture in retinal blood vessel segmentation tasks and used a patch selection based on entropy to ensure the retinal blood vessels were contained in the patches [30]. The entropy of images indicates the richness of information, where images with higher entropy will contain more foreground class objects. This approach will alleviate class imbalance.

To sum up, both 2D networks and 3D networks have their limitations. Although cascading two networks can achieve better results, it also increases the number of parameters and complexity of the network. In recent work, no novel network structure has been proposed for cerebral aneurysm segmentation, and no method to overcome class imbalance has been proposed for the characteristics of cerebral aneurysm data. Therefore, we improved the network structure according to the characteristics of the data and relieved the class imbalance of the data through the gradient entropy strategy.

This work was inspired by the successful application of Transformers in 3D CNNs in the field of brain tumor segmentation [31]. Due to the class imbalance of the CADA dataset, we introduce a patch-based architecture that relies on the 2D U-Net as the backbone and

adds a Transformer block to capture remote information. We propose a patch selection strategy based on entropy to make the training data more sufficient. Then, three continuous patches are taken as inputs and a segmentation map corresponding to the central patch is generated.

As illustrated above, in this paper, the main contributions to aneurysm segmentation are as follows:

- (1) In order to obtain more sufficient training data, we used a new patch selection strategy. More training data with aneurysms will alleviate class imbalance.
- (2) We used three channels as inputs, which represents an approach between 2D and 3D. This approach can use 3D information and pay attention to the in-plane resolution.
- (3) We improved the recombination strategy. This will make the boundary of the segmentation target clearer.

The rest of this paper is organized as follows: Section 2 describes the proposed methodology in detail. Section 3 shows the experimental results. Finally, we discuss and conclude our paper in Sections 4 and 5.

2. Materials and Methods

2.1. Dataset and Preprocessing

The MICCAI 2020 CADA challenge provided 109 cases. Image data of patients with cerebral aneurysms without vasospasm were collected for the purpose of assisting diagnosis and treatment [32]. The image data were acquired utilizing the digital subtraction AXIOM Artis C-arm system. Post-processing was performed using LEONARDO InSpace 3D (Siemens, Forchheim, Germany). After implementation of the contrast agent, a reconstruction of a volume of interest was selected by a neurosurgeon. The reconstructed images generally consist of 220 contiguous slices. The imaging parameters were as follows: in-plane size of 256×256 ; iso-voxel size of 0.5 mm. Patients were of different ages and genders, making the samples diverse.

2.1.1. Slice Selection Strategy

In cerebral aneurysm images, approximately 98% of the pixels belong to the background, with the remaining 2% of pixels belonging to the foreground class. We selected the slices using the range entropy strategy proposed in [33]. For a given sample $X \in \mathbb{R}^{D \times H \times W}$, its spatial resolution is $H \times W$, and its depth dimension is D (# of slices), normalizing the images using the following formula:

$$X_i^{norm} = \frac{X_i - \min(X_i)}{\max(X_i) - \min(X_i)} * 255, \quad i = [1, 2, \dots, D] \tag{1}$$

where $\min(X_i)$ and $\max(X_i)$ denote the minimum and maximum values of the i -th slice in X . For every ten continuous slices, seven slices with the highest RH (range entropy) values were selected as the final slices annotated as $X_S \in \mathbb{R}^{7 \times \frac{D}{10} \times H \times W}$. The calculation formula is as follows:

$$H(S) = - \sum_{i=0}^{255} p_i \log_2 p_i \tag{2}$$

$$RH(S) = H(S) + w * R(S) \tag{3}$$

$$R(S) = \frac{1}{b} (\sum \max_b(S) - \sum \min_b(S)) \tag{4}$$

where $H(S)$ is the entropy of each slice, and p_i is the probability of the pixel value i that exists in S ; $R(S)$ is a generalized range of the slice S , and w depicts the weight of the range $R(S)$; $\max_b(S)$ denotes the top b maximum gray values in the slice S , and $\min_b(S)$ denotes the last b minimum gray values in the slice S , where b denotes the number of selected pixels [33].

2.1.2. Sliding Window Strategy

In the cerebral aneurysm datasets, the number of available annotated images is not large enough for a good training model. Thus, for cerebral aneurysm segmentation, we used the patch method with the sliding window strategy.

From Figure 1a, under the same stride size, the proportion of patches containing aneurysms is large when the patch size is 96×96 pixel². Although the proportions of 108×108 pixel² and 128×128 pixel² are larger, they will contain more noise due to the larger size. Additionally, in the segmentation of cerebral blood vessels [33], a patch size of 96×96 pixel² was chosen, so we chose 96×96 pixel² as the patch size.

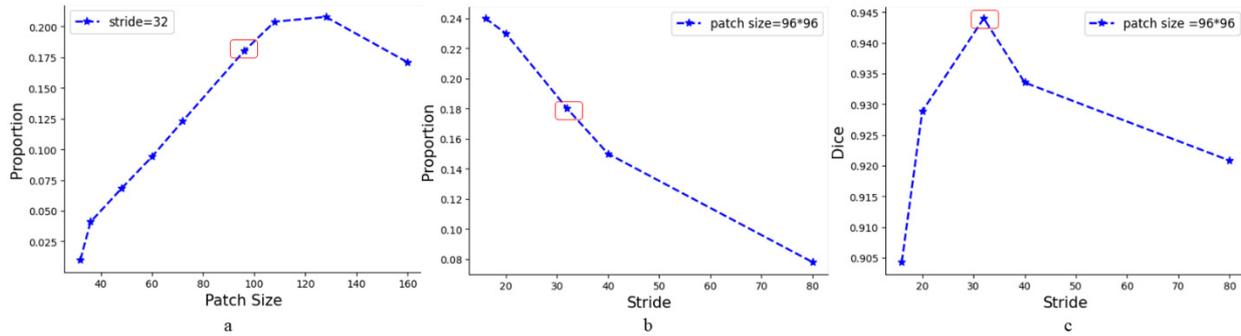


Figure 1. The choice of patch size and stride size. (a) The proportion of patches containing aneurysms with different patch sizes when the stride size is 32. (b) The proportion of patches containing aneurysms with different stride sizes when the patch size is 96×96 pixel². (c) The Dice at different stride sizes.

The original resolution of the CADA dataset is 256×256 pixel². From Figure 1b,c, although smaller strides produce more patches containing aneurysms, their performances are worse. This may be because the selected patches contain duplicate positive samples, resulting in information redundancy. When the stride is 32, the Dice is the best, and 32 is the largest common factor of 96 and 160. Considering the accuracy and computational complexity, we chose 32 as the moving stride of the sliding window. That is, a 96×96 pixel² sliding window starts from the upper left corner of the slice with a moving stride of 32 pixels. Each step of moving the sliding window yields a corresponding patch. For a whole 256×256 pixel² slice, we could acquire 36 patches.

2.1.3. Patch Selection by Gradient Entropy Sampling

The generated patch is denoted as $P \in \mathbb{R}^{H_p \times W_p}$. Owing to the similarity between blood vessels and aneurysms, selecting patches only through information entropy will include a lot of noise. Since aneurysms have higher gradients than vessels, we combined information entropy with a gradient strategy for patch selection called *GH*, which is proposed as the gradient entropy strategy. The calculation formula is as follows:

$$GH(P) = H(P) + \gamma * G_y(P) \tag{5}$$

$$G_y(P) = \frac{1}{c} \sum \max_c(P(i, j) - P(i, j - 1)) \tag{6}$$

$[i = 1, 2, 3, \dots, H_p, j = 2, 3, \dots, W_p + 1]$, where $GH(P)$ denotes the gradient entropy value of each patch, $P(i, j)$ is the pixel value of the index (i, j) in P , $G_y(P)$ represents the gradient in the y-direction and $\max_c(\cdot)$ denotes the top c maximum G_y in the patch P . Patches with a higher GH are selected in the training set.

Applying the sliding window strategy, patch selection by information entropy sampling showed that the proportion of patches containing aneurysms was 12%, i.e., 12% of the selected patches included aneurysms. Additionally, through patch selection by gradient entropy sampling, the proportion was increased to 16%. The patches selected by

the gradient entropy sampling strategy are shown in Figure 2. This shows that most of the selected patches were aneurysms or blood vessels.

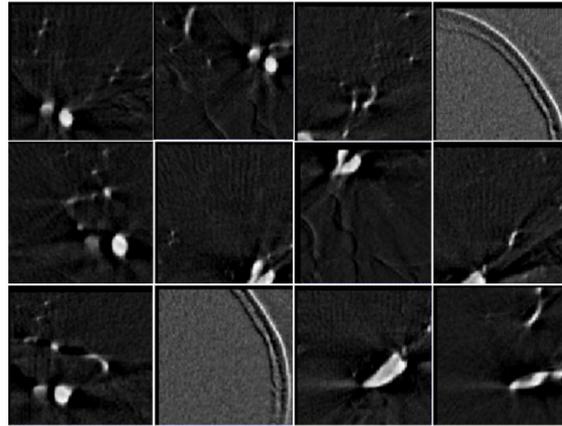


Figure 2. Selected patches based on the gradient entropy sampling strategy.

2.2. Network Architecture

We chose the structure of U-Net as the backbone. Figure 3 presents the architecture of our network that consists of a CNN encoder block, a Transformer block and a decoder block with a shortcut connection at each resolution level. The encoder obtains high-dimensional features, and the decoder utilizes these encoded features to recover the segmentation target. The spatial attention is utilized to strengthen the region of interest on the feature maps while suppressing the potential background or irrelevant parts. Hence, we propose a Transformer block that shares space and learns the relation between these feature map embeddings using self-attention modules. The network uses the image information in three consecutive patches to predict the segmentation for the center patch, and adjacent slices provide rich spatial information.

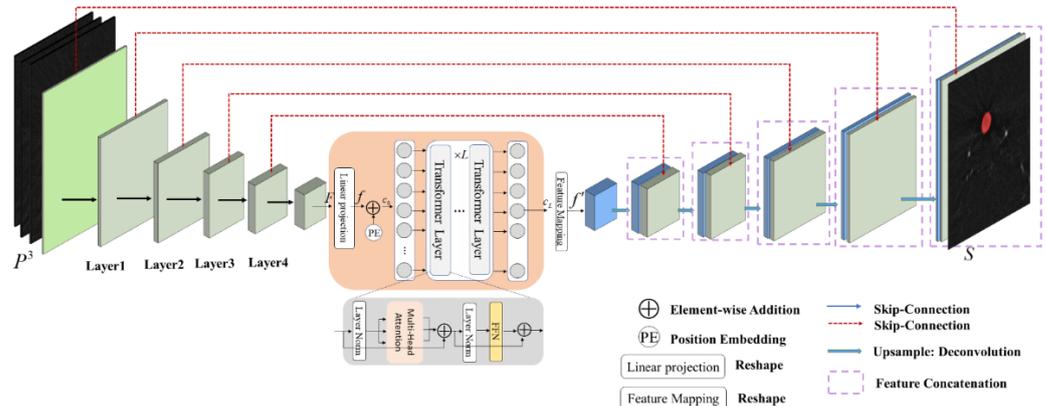


Figure 3. The overall network. Three consecutive patches are used as inputs. The left part of the network is the encoder based on ResNet34, and each green block corresponds to the layer of ResNet34. The middle part of the network is the Transformer block. The right part of the network is the decoder, and each blue block corresponds to the upsampling block.

2.2.1. Network Encoder

The encoder blocks are composed of ResNet34 and the Transformer block. ResNet34 is mainly composed of a Bottleneck. In order to prevent overfitting, we added a dropout block to the original Bottleneck (see Figure 4). ResNet34 is a type of neural network that captures more deeper features by using skip connections to “skip” a number of convolutional layers in every Bottleneck in the network. The structure of the ResNet34 encoder block is shown in Figure 5.

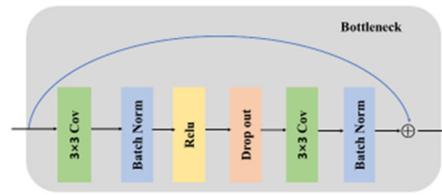


Figure 4. The visual of the Bottleneck.

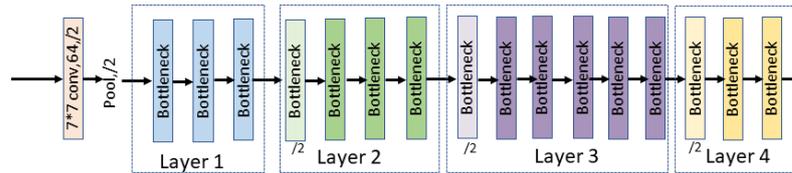


Figure 5. The structure of the ResNet34 encoder block.

There are 4 layers in the ResNet34 encoder block, and the numbers of Bottlenecks in each layer are 3, 4, 6 and 3. The final output of ResNet34 can be written as $F \in \mathbb{R}^{C \times H' \times W'}$.

Next, we present a Transformer block comprising L repeated Transformer layers to achieve a global context using an attention mechanism. For a given F , it is flattened into a vector of size $\mathbb{R}^{d \times N}$, by a linear projection operation W_p , resulting in $f \in \mathbb{R}^{d \times N}$, $d = 512$, $N = 3 \times 3$. $c_0 = f + PE \in \mathbb{R}^{d \times N}$ constitutes the input of the Transformer block, and PE is the learnable position embedding.

Each Transformer layer has a Multi-Head Attention (MHA) block and a feed-forward neural network (FFN), and the output of each layer can be calculated by the following formula:

$$c'_i = MHA(LN(c_{i-1})) + c_{i-1} \tag{7}$$

$$c_i = FFN(LN(c'_i)) + c'_i \tag{8}$$

where $LN(*)$ denotes the layer normalization, and c_i is the output of the i th ($i \in [1, 2, \dots, L]$) Transformer layer.

2.2.2. Network Decoder

To fit the input dimension of the 2D CNN decoder, f is then reshaped to $f' \in \mathbb{R}^{d \times H'' \times W''}$ by a feature mapping module. The decoding process corresponds to the encoding process, which combines local and global features until the original resolution is restored and pays more attention to the local context to obtain edge and semantic information. Additionally, through cascaded upsampling operations and convolution blocks, the final segmentation map $S \in \mathbb{R}^{H_p \times W_p}$ is generated.

2.3. Prediction

After the training phase, the trained model can only segment 96×96 images. In the prediction phase (see Figure 6), the input samples were segmented into 96×96 patches using the sliding window strategy. Then, the trained model generated segmentations of patches. The segmentation maps were recombined to obtain the final segmentation map. The stride was 32, and adjacent patches had overlapping pixels. For each pixel, the average strategy [33] is where the aneurysm probability of each pixel is obtained by averaging possibilities over all the predicted patches covering the pixel. This solution will lose a lot of detail and is time-consuming. We propose a new recombination strategy.

For each image, each row of the sliding window produced six overlapping patches, and a total of 36 patches were generated. There were $C_4^2 + 1 = 7$ cases for recombining each row. Similarly, each column also had 7 cases. There was a total of $6 \times 7^4 + 7^3 = 14,749$ recombination situations. Due to the robustness of the model, we randomly selected only 49 of these cases, and then 49 segmentation maps were obtained. We averaged these

49 maps to obtain the $S_{Aver} \in \mathbb{R}^{H \times W}$; the $S_{Aver} \in \mathbb{R}^{H \times W}$ had obvious clipped parts. We added the self-attention of each patch to the segmentation result $S_{candidate} \in \mathbb{R}^{36 \times H \times W}$ to ensure the integrity of the segmented aneurysms by replacing the corresponding position of S_{Aver} with each patch segmentation map. If S_{Aver} contained aneurysms and the number of aneurysms contained in $S_{candidate}$ was more than the threshold η , and if S_{Aver} did not contain an aneurysm, but the number of aneurysms contained in $S_{candidate}$ was more than the threshold $36 - \eta$, then we selected the one with the largest aneurysm size among the 36 results as the final segmentation result. Otherwise, we considered the image to be without an aneurysm.

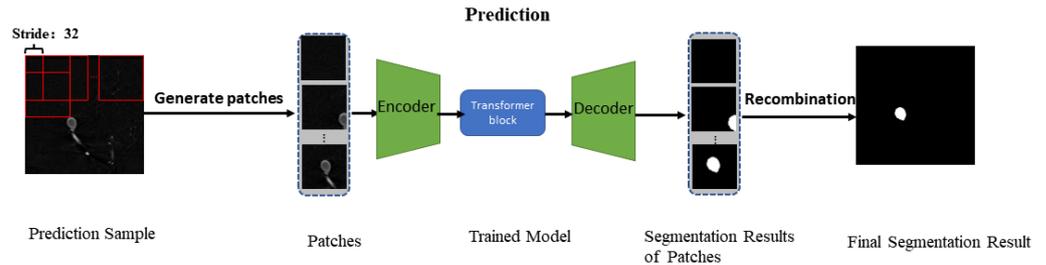


Figure 6. The pipeline of prediction. The test set samples first generate patches through the sliding window, then input them into the trained network to generate the corresponding prediction and finally splice the final results.

3. Experiment

3.1. Data Augmentation

Most selected patches contained aneurysms or vessels. For data augmentation, the following data augmentation techniques were applied: (1) random cropping; (2) horizontal flipping; (3) 45° rotation. We only used data augmentation on the training data. After applying patch selection, each sample $X \in \mathbb{R}^{D \times H \times W}$ was cut into 200 inputs consisting of three consecutive patches. In terms of the specific process of random clipping, for every three consecutive slices, we chose an identical random position to crop, generating three consecutive patches, with a total of $D/3$ inputs. Then, we applied horizontal flipping and 45° rotation on patches generated by random clipping. To sum up, after data augmentation, D additional inputs were generated.

3.2. Evaluation Metrics

The metrics of evaluation were the Dice score, recall, Hausdorff distance (95%), F2 score, mean average precision (mAP) and Intersection over Union (IOU) score.

Dice similarity coefficient: The Dice similarity coefficient (Dice) [34] is a metric used for assessing the quality of segmentation. It measures the similarity between the predicted label and ground truth.

$$Dice = \frac{2|Pre \cap G|}{|Pre| + |G|}, \text{ where } Pre \text{ is the predicted segmentation map, and } G \text{ is the ground truth.}$$

Hausdorff distance: A high Hausdorff distance value implies that the two contours do not closely match. It is a symmetric measure of distance between two contours and is defined as [35]

$$H(Pre, G) = \max(h(Pre, G), h(G, Pre)), h(Pre, G) = \max_{p_i \in Pre} \min_{g_j \in G} \|p_i - g_j\|$$

Recall: Recall is the ability to segment the region of interest in the segmentation experiment. It indicates the proportion of all true positives that are correctly predicted.

$recall = \frac{TP}{TP + FN}$, where TP means true positive predictions, and FN means false negative predictions.

$$F2 \text{ score: } F2 = 5 \cdot \frac{precision \cdot recall}{4 \cdot precision + recall}, \text{ where } precision = \frac{TP}{TP + FP}.$$

mAP: The mAP is the average precision of all categories detected, that is, the average precision of segmenting the foreground and background.

$$IOU: IOU = \frac{Pre \cap G}{Pre \cup G}, \text{ the closer the } IOU \text{ is to } 1, \text{ the better the segmentation result.}$$

3.3. Implementation Details

We used a random split (70% training, 10% validation and 20% test) at the patient level and conducted a five-fold cross-validation evaluation. All experiments were implemented in Pytorch.

In the data preprocessing stage, the parameters should be determined. w in Equation (3), b in Equation (4), γ in Equation (5) and c in Equation (6) were set to 0.05, 10, 0.1 and 20, respectively. Because the patch size is 96×96 pixel² and the stride of the sliding window is 32, a patch can be composed of 1/3 of three adjacent patches. η in Section 2.3 was set to $36 - 3 = 33$. The dropout regularization with $p = 0.2$ was used.

We used Resnet34 pretrained on ImageNet as the CNN encoder block. We set the training of our model on Pytorch with an initial learning rate of 1×10^{-4} . If the Dice score on the validation dataset did not improve for 15 epochs, the learning rate was reduced to half of the original rate. It was optimized by Adam with a batch size of 32. The training iterated 50 epochs on a single NVIDIA GeForce RTX 3090 GPU with 24 GB memory. The softmax Dice loss was employed to train the network.

To demonstrate the advantages of our work, we compared it with other methods (3D U-Net, DeepLabV3+, DeepLabV3, Linknet, FPN, UNet++). (1) The 3D U-Net took a learning rate of 1×10^{-4} , and it was optimized by Adam and trained with an NVIDIA GeForce RTX 3090 for 500 epochs from scratch using a batch size of 4. (2) DeepLabV3+ was pretrained on ImageNet, and then it was trained on our dataset with an initial learning rate of 1×10^{-4} . It was optimized by Adam with a batch size of 16. (3) DeepLabV3 implemented the same setup as DeepLabV3+. (4) Linknet was pretrained on ImageNet. Its initial learning rate was 1×10^{-5} , it was optimized by Adam and the weight decay was 1×10^{-4} . It was trained with an NVIDIA GeForce RTX 3090 for 50 epochs using a batch size of 16. (5) FPN implemented the same setup as Linknet. (6) UNet++ implemented the same setup as Linknet.

3.4. Segmentation Result and Comparisons

We conducted a five-fold cross-validation evaluation on the training set, and our method achieved an average Dice score of 0.944, an IOU score of 0.941, recall of 0.946, an F2 score of 0.942, a mAP of 0.896 and a Hausdorff distance of 3.12mm, which are comparable or higher results than those of previous state-of-the-art (SOTA) methods presented in Table 1. Compared with the 3D U-Net, our method showed superiority in four metrics, with a significant improvement.

Table 1. Comparison of SOTA methods.

| Model | Dice | IOU | Recall | Hausdorff_95 | mAP | F2 Score |
|-----------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 3D U-Net | 0.631 | 0.521 | 0.690 | 19.1 | 0.857 | 0.653 |
| Linknet | 0.867 | 0.856 | 0.952 | 19.85 | 0.893 | 0.859 |
| DeepLabV3 | 0.916 | 0.912 | 0.936 | 10.22 | 0.632 | 0.897 |
| FPN | 0.929 | 0.925 | 0.925 | 8.40 | 0.838 | 0.936 |
| DeepLabV3+ | 0.937 | 0.934 | 0.939 | 6.36 | 0.835 | 0.936 |
| UNet++ | 0.939 | 0.935 | 0.945 | 10.28 | 0.874 | 0.937 |
| Proposed | 0.944 | 0.941 | 0.946 | 3.12 | 0.896 | 0.942 |

Figure 7 shows example slices from test images in the dataset and the segmentations predicted by the proposed method. We observed that the proposed method could accurately segment fine or large cerebral aneurysms.

Figure 8 shows that the predicted segmentations of DeepLabV3+ mostly contained holes for larger aneurysms, while for smaller aneurysms, it was prone to false negative predictions.

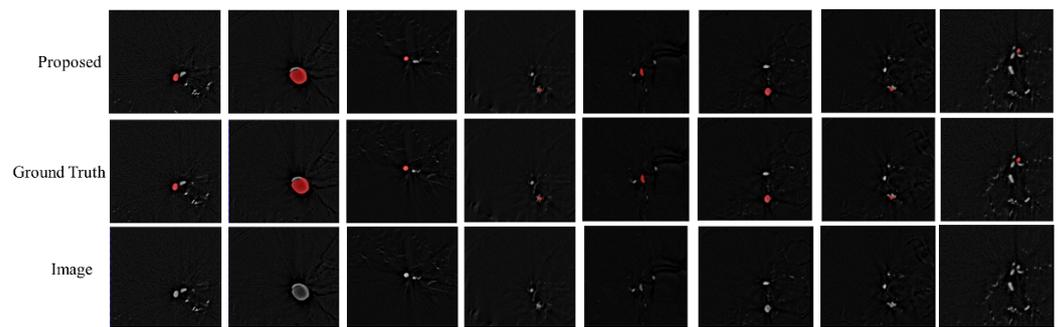


Figure 7. The visual segmentation result of our method.

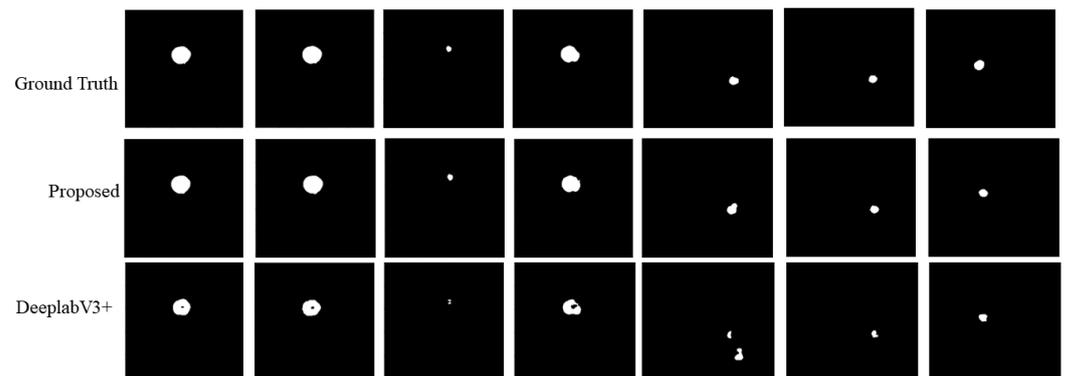


Figure 8. The visual results of our method and DeepLabV3+.

3.5. Ablation Study

We designed different ablation studies to evaluate the contribution of the gradient entropy sampling strategy and the three-channel input, based on the patch and recombination strategies.

In the data preprocessing, the gradient entropy sampling strategy was implemented to generate a sufficient number of training patches from the limited images. We compared the traditional entropy sampling strategy with our strategy, and the result is shown in Table 2. The visual result is shown in Figure 9. The last row shows the segmentation of our method, where the segmentation is almost the same as the ground truth. As shown in Figure 9, our gradient entropy sampling strategy can better distinguish vessels and aneurysms. It is easy for information entropy sampling selection to produce false positives, and it has poor segmentation performance concerning smaller aneurysms.

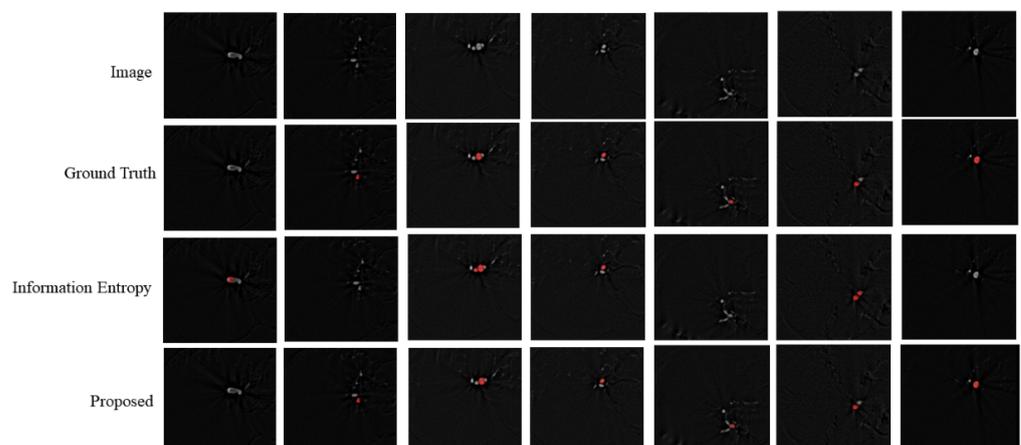


Figure 9. The visual results of different patch selection strategies.

Table 2. Ablation study on the gradient entropy sampling strategy.

| Model | Dice | IOU | Recall | Hausdorff_95 |
|---------------------|--------------|--------------|--------------|--------------|
| Information entropy | 0.928 | 0.925 | 0.945 | 4.42 |
| Proposed | 0.944 | 0.941 | 0.946 | 3.12 |

As shown in Table 2, our model implementing the proposed gradient entropy sampling strategy achieved a better segmentation performance than the traditional entropy sampling strategy. This demonstrates that the quality of the training dataset can also improve the performance of segmentation, and that the gradient entropy sampling strategy provided more patches that contained aneurysms.

Table 3 shows that our recombination strategy achieved a better performance than the average strategy. Our recombination strategy combines the global attention and self-attention of patches, which not only ensures the overall performance of segmentation but also ensures the integrity of the foreground.

Table 3. Ablation study on the recombination strategy.

| Model | Dice | IOU | Recall | Hausdorff_95 |
|-----------------|--------------|--------------|--------------|--------------|
| Ours w/o post | 0.942 | 0.938 | 0.942 | 4.00 |
| Proposed | 0.944 | 0.941 | 0.946 | 3.12 |

A characteristic of cerebral aneurysm image data is class imbalance. Furthermore, the patch-based approach ensures that the classes of the input data are balanced and the patches give the network access to local information about the pixels, which has an impact on the overall prediction. Table 4 shows the ablation experimental results of one patch as an input, three continuous slices as an input and our proposed three continuous patches as an input. The three continuous patches as an input achieved the best results, the three continuous slices as an input achieved the worst Hausdorff distance and the one patch as an input achieved the worst result for the other three metrics. This indicates that the multichannel input provided more 3D spatial information. The Hausdorff distance implies the degree to which two contours closely match. The three continuous slices as an input lost detail of the local information, resulting in a poor effect of aneurysm contours.

Table 4. Ablation study on the three-channel input and resolution.

| Model | Dice | IOU | Recall | Hausdorff_95 |
|-----------------|--------------|--------------|--------------|--------------|
| One-patch input | 0.931 | 0.919 | 0.932 | 6.40 |
| Based on slices | 0.939 | 0.934 | 0.946 | 7.42 |
| Proposed | 0.944 | 0.941 | 0.946 | 3.12 |

4. Discussion

In the experiment, we found that the proposed segmentation network architecture had a great advantage over the previous algorithms. In the ablation study, we verified the validity of the gradient entropy selection strategy, and the result shows that it had a better performance in aneurysm segmentation than traditional information entropy selection. This provides ideas for the development of small medical area segmentation fields. It reduces false positive and false negative predictions. Three channels provided more 3D information and had a significant effect on extracting features. The results in Table 1 show that our network was optimal in all six metrics. Additionally, we can see from Figure 8 that, whether it is a large aneurysm or a small aneurysm, our model's performance was better. Due to the size of the training set being small, the segmentation performance of the 3D U-Net was not ideal. Compared with the remaining 2D SOTA segmentation networks, our

input provided additional 3D information, so the segmentation results were better than those of other networks. Although the DeepLabv3+ network is better than the remaining networks, there are still holes in the segmentation of aneurysms of a large size.

For our method, although the average performance was better than the other models, there were still a small number of samples whose segmentation was not ideal, which may be due to the excellent threshold selection when selecting patches. As a result, the network did not learn the information specific to the sample, and thus the hyperparameter selection should be improved.

In the clinical field, aneurysm screening is essential, as early detection can prevent stroke. Relying on a doctor's manual observation is inefficient, and different doctors have different evaluation criteria. In this paper, we provided a new efficient aneurysm segmentation algorithm, which facilitates rapid diagnosis and unified evaluation criteria.

5. Conclusions

Cerebral aneurysm is one of the most common cerebrovascular diseases, and its rupture has a high mortality rate from subarachnoid hemorrhage (SAH). Due to the limited training data, existing automatic segmentation methods cannot segment aneurysms sufficiently, so we adopted an entropy selection strategy to provide informative training data. Specifically, we proposed a patch-based segmentation model. Compared with full resolution inputs, the selected patches are only part of the training data, thus preventing overfitting. We used the gradient entropy strategy to select patches that may contain aneurysms, improving and speeding up the network. The experimental results show that better training data can also improve the network's performance. We proposed a new application scenario for entropy. In future work, we will continue to design deep architectures for small datasets in medical image processing.

Author Contributions: Conceptualization, X.A. and S.L.; Investigation, X.A., J.H. and D.M.; Methodology, T.L. and Y.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China grant number 81925020 and 81630051, and Tianjin Science and Technology Program grant number 20JCZDJC00620.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Greving, J.P.; Wermer, M.J.; Brown, R.D., Jr.; Morita, A.; Juvela, S.; Yonekura, M.; Ishibashi, T.; Torner, J.C.; Nakayama, T.; Algra, A.; et al. Development of the PHASES score for prediction of risk of rupture of intracranial aneurysms: A pooled analysis of six prospective cohort studies. *Lancet Neurol.* **2014**, *13*, 59–66. [[CrossRef](#)]
2. Frösen, J.; Tulamo, R.; Paetau, A.; Laaksamo, E.; Korja, M.; Laakso, A.; Niemelä, M.; Hernesniemi, J. Saccular intracranial aneurysm: Pathology and mechanisms. *Acta Neuropathol.* **2012**, *123*, 773–786. [[CrossRef](#)] [[PubMed](#)]
3. The UCAS Japan Investigators. The natural course of unruptured cerebral aneurysms in a Japanese cohort. *N. Engl. J. Med.* **2012**, *366*, 2474–2482. [[CrossRef](#)]
4. Chung, A.C.S.; Noble, J.A.; Summers, P. Vascular segmentation of phase contrast magnetic resonance angiograms based on statistical mixture modeling and local phase coherence. *IEEE Trans. Med. Imaging* **2004**, *23*, 1490–1507. [[CrossRef](#)] [[PubMed](#)]
5. Volkau, I.; Zheng, W.; Baimouratov, R.; Aziz, A.; Nowinski, W.L. Geometric modeling of the human normal cerebral arterial system. *IEEE Trans. Med. Imaging* **2005**, *24*, 529–539. [[CrossRef](#)]
6. Khan, W.; Ansell, D.; Kuru, K.; Amina, M. Automated aircraft instrument reading using real time video analysis. In Proceedings of the 2016 IEEE 8th International Conference on Intelligent Systems (IS), Sofia, Bulgaria, 4–6 September 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 416–420.
7. Osher, S.; Sethian, J.A. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.* **1988**, *79*, 12–49. [[CrossRef](#)]
8. Gulshan, V.; Peng, L.; Coram, M.; Stumpe, M.C.; Wu, D.; Narayanaswamy, A.; Venugopalan, S.; Widner, K.; Madams, T.; Cuadros, J.; et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **2016**, *316*, 2402–2410. [[CrossRef](#)] [[PubMed](#)]
9. Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* **2018**, *172*, 1122–1131. [[CrossRef](#)]
10. Esteva, A.; Kuprel, B.; Novoa, R.A.; Ko, J.; Swetter, S.M.; Blau, H.M.; Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **2017**, *542*, 115–118. [[CrossRef](#)]

11. Cheng, J.Z.; Ni, D.; Chou, Y.H.; Qin, J.; Tiu, C.M.; Chang, Y.C.; Huang, C.S.; Shen, D.; Chen, C.M. Computer-aided diagnosis with deep learning architecture: Applications to breast lesions in US images and pulmonary nodules in CT scans. *Sci. Rep.* **2016**, *6*, 24454. [[CrossRef](#)] [[PubMed](#)]
12. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436. [[CrossRef](#)] [[PubMed](#)]
13. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [[CrossRef](#)]
14. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; Van Der Laak, J.A.; Van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [[CrossRef](#)] [[PubMed](#)]
15. Khan, W.; Hussain, A.; Kuru, K.; Al-Askar, H. Pupil localisation and eye centre estimation using machine learning and computer vision. *Sensors* **2020**, *20*, 3785. [[CrossRef](#)]
16. Cai, Y.; Wang, Y. Maunet: An improved version of unet based on multi-scale and attention mechanism for medical image segmentation. In Proceedings of the Third International Conference on Electronics and Communication, Network and Computer Technology (ECNCT 2021), Harbin, China, 3–5 December 2022; SPIE: Bellingham, DC, USA, 2022; Volume 12167, pp. 205–211.
17. Isensee, F.; Jaeger, P.F.; Kohl, S.A.; Petersen, J.; Maier-Hein, K.H. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **2020**, *18*, 203–211. [[CrossRef](#)]
18. Milletari, F.; Navab, N.; Ahmadi, S.A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 565–571.
19. Ibragimov, B.; Xing, L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. *Med. Phys.* **2016**, *44*, 547–557. [[CrossRef](#)]
20. Liang, S.; Thung, K.-H.; Nie, D.; Zhang, Y.; Shen, D. Multi-View Spatial Aggregation Framework for Joint Localization and Segmentation of Organs at Risk in Head and Neck CT Images. *IEEE Trans. Med. Imaging* **2020**, *39*, 2794–2805. [[CrossRef](#)]
21. Men, K.; Geng, H.; Cheng, C.; Zhong, H.; Huang, M.; Fan, Y.; Plastaras, J.P.; Lin, A.; Xiao, Y. More accurate and efficient segmentation of organs-at-risk in radiotherapy with convolutional neural networks cascades. *Med. Phys.* **2019**, *46*, 286–292.
22. Gao, Y.; Huang, R.; Chen, M.; Wang, Z.; Deng, J.; Chen, Y.; Yang, Y.; Zhang, J.; Tao, C.; Li, H. FocusNet: Imbalanced large and small organ segmentation with an end-to-end deep neural network for head and neck CT images. In *MICCAI 2019, LNCS*; Shen, D., Ed.; Springer: Cham, Switzerland, 2019; Volume 11766, pp. 829–838.
23. Ren, X.; Xiang, L.; Nie, D.; Shao, Y.; Zhang, H.; Shen, D.; Wang, Q. Interleaved 3D-CNNs for joint segmentation of small-volume structures in head and neck CT images. *Med. Phys.* **2018**, *45*, 2063–2075. [[CrossRef](#)]
24. Li, X.; Chen, H.; Qi, X.; Dou, Q.; Fu, C.W.; Heng, P.A. H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *IEEE Trans. Med. Imaging* **2018**, *37*, 2663–2674. [[CrossRef](#)]
25. Xu, G.; Wu, X.; Zhang, X.; He, X. Levit-unet: Make faster encoders with transformer for medical image segmentation. *arXiv Prepr.* **2021**, arXiv:2107.08623. [[CrossRef](#)]
26. Khan, S.; Naseer, M.; Hayat, M.; Zamir, S.W.; Khan, F.S.; Shah, M. Transformers in vision: A survey. *arXiv Prepr.* **2021**, arXiv:2101.01169. [[CrossRef](#)]
27. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth 16×16 words: Transformers for image recognition at scale. *arXiv Prepr.* **2020**, arXiv:2010.11929.
28. Duan, Z.; Montes, D.; Huang, Y.; Wu, D.; Romero, J.M.; Gonzalez, R.G.; Li, Q. Deep Learning Based Detection and Localization of Cerebral Aneurysms in Computed Tomography Angiography. *arXiv Prepr.* **2020**, arXiv:2005.11098.
29. Shahzad, R.; Pennig, L.; Goertz, L.; Thiele, F.; Kabbasch, C.; Schlamann, M.; Kriscsek, B.; Maintz, D.; Perkuhn, M.; Borggrefe, J. Fully automated detection and segmentation of intracranial aneurysms in subarachnoid hemorrhage on CTA using deep learning. *Sci. Rep.* **2020**, *10*, 21799. [[CrossRef](#)] [[PubMed](#)]
30. Feng, Z.; Yang, J.; Yao, L. Patch-based fully convolutional neural network with skip connections for retinal blood vessel segmentation. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1742–1746.
31. Wang, W.; Chen, C.; Ding, M.; Yu, H.; Zha, S.; Li, J. Transbts: Multimodal brain tumor segmentation using transformer. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Strasbourg, France, 27 September–1 October 2021; Springer: Cham, Switzerland, 2021; pp. 109–119.
32. Ivantsits, M.; Goubergrits, L.; Kuhnigk, J.-M.; Huellebrand, M.; Brüning, J.; Kossen, T.; Pfahringer, B.; Schaller, J.; Spuler, A.; Kuehne, T.; et al. Cerebral Aneurysm Detection and Analysis Challenge 2020 (CADA). In *Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2021; pp. 3–17. [[CrossRef](#)]
33. Meng, C.; Sun, K.; Guan, S.; Wang, Q.; Zong, R.; Liu, L. Multiscale dense convolutional neural network for DSA cerebrovascular segmentation. *Neurocomputing* **2020**, *373*, 123–134. [[CrossRef](#)]
34. Dice, L.R. Measures of the Amount of Ecologic Association between Species. *Ecology* **1945**, *26*, 297–302. [[CrossRef](#)]
35. Taha, A.A.; Hanbury, A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Med. Imaging* **2015**, *15*, 29. [[CrossRef](#)] [[PubMed](#)]