N-PointNet: A multi-layer embedded deep learning model for 3D intracranial aneurysm classification and segmentation

Jiaqi Wang
  Wannan Medical College

Juntong Liu
  Wannan Medical College

Zhengyuan Xu
  Wannan Medical College

Pengzhan Yin
  First Affiliated Hospital of Wannan Medical College

Jinlong Yuan
  First Affiliated Hospital of Wannan Medical College

Yunfeng Zhou
  First Affiliated Hospital of Wannan Medical College

Mingquan Ye (✉ ymq@wnmc.edu.cn)
  Wannan Medical College

Research Article

Abstract

Background

In computer-aided intracranial aneurysm (IA) classification and segmentation, applications of 3D point cloud algorithms are increasingly widespread. However, the traditional point-based deep learning algorithm has the problem of poor segmentation effect.

Methods

An improved end-to-end depth network structure (N-PointNet) is proposed for IA classification and segmentation. First, the point cloud data of the IA are preprocessed. Then, the PointNet + + network structure is used as a backbone with learned hierarchical properties. After that, the preprocessed and resampled data produce multiple layers of information embedded in the original network input to further enhance its characteristics. Finally, a side output block is added, and the loss function of the corresponding layer is calculated. The multi-loss function facilitates fast convergence and improves model performance.

Conclusion

An experiment on the IntrA dataset proved the superiority of N-PointNet and obtained the best classification and segmentation results among the models tested. In addition, the proposed method has good generalization ability and has been verified on the common ModelNet40 dataset.

Introduction And Background

An IA is an abnormal cystic protrusion caused by structural changes in the blood vessels of the brain [1] [2], usually in the form of IA, and is one of the most dangerous cerebrovascular diseases. Prompt diagnosis and preoperative examination are essential treatment strategies. The main current treatment is neck clamping to prevent aneurysmal rupture. Decisions about the location and position of the clamp remain heavily dependent on clinical judgment based on the physician's experience. As part of an IA surgical support system that simulates true neurosurgery [3] [4], accurate IA segmentation is the most critical part because it is used to extract the aneurysm neck, i.e., the aneurysm boundary.

With the development of medical imaging technologies such as computed tomography (CT) and magnetic resonance angiography (MRA), more methods for detecting IAs are gradually becoming available. However, doctors have strong subjectivity in diagnosing IAs by observing angiographic results from imaging equipment. Studies have shown [5]that compared with manually annotated 2D planar images, 3D morphological analysis for computer-aided diagnosis improves the accuracy and consistency of parameter measurements and also quantifies shape irregularities more reliably. Computer-aided

analysis tools can help clinicians standardize the morphological assessment of lesion areas, resulting in more consistent and accurate assessment of IAs. Studies on 2D MRA images have been limited to pixel- and voxel-based 3D neural networks, which ignore the information in the manifold. Considering these issues, an open 3D point-cloud-based IA dataset (IntrA) [6] was used as training data.

Although existing models perform well on general 3D datasets, they sometimes fail on medical data due to domain gaps. Accurate pathological sections are of great significance to disease diagnosis and treatment. However, 3D medical data may contain incomplete pathological structures, which are difficult to distinguish from healthy parts within an object [6]. Furthermore, for non-medical datasets, different types of the same object class often share the same schema. In contrast, medical point clouds of the same object class are more diverse and complex in shape geometrically and topologically [7], making object classification or segmentation difficult. Insufficient medical data samples also impede the learning of feature shape descriptors. Therefore, it is crucial to design an effective deep learning method that can both exhibit good performance on medical data and generalize well to non-medical data. The task is therefore the following:

- Use a PointNet++ [8] network structure as a backbone for pre-processed IA data with hierarchical learning characteristic.
- Generate the preprocessed and resampled data as a multi-layer embedding into the original network input, further strengthening the input characteristics.
- Add a side output block and calculate the loss function of the corresponding layer. A multi-loss function is conducive to rapid model convergence and improves performance [9].

The experiment on the IntrA dataset proved the superiority of N-PointNet and obtained the best classification and segmentation results. In addition, the proposed method has good generalization ability and has been verified on the common ModelNet40 dataset [10]. The workflow is shown in Fig. 1.

Related Works

Deep learning on intracranial aneurysms

With the integration of computer intelligence technology and medical imaging technology, deep learning (DL) has become a hot research topic in the field of computer-aided diagnosis (CAD) in recent years [11]. It is useful in medical image detection, classification, segmentation, alignment, image retrieval, image generation, and image enhancement. Deep learning can help doctors find the common features of disease diagnosis from the Big Data of medical images and provide scientific methods for disease screening and treatment in clinical medicine.

In recent years, many researchers have applied deep learning algorithms to aneurysm detection and measurement to assist physicians in detecting lesions before they rupture [12]. Stember *et al*. [13] first used convolutional neural networks to automatically detect cerebral aneurysms from MRA and

successfully detected 85 of 86 test set aneurysms; second, the *u*-net network was used to predict the size of 14 basilar aneurysms, with results that differed from radiologist-labeled aneurysms by an average size of 2.01 mm and an average area of 8.1 mm^2. Daju *et al.* [14] used a ResNet-18 structured neural network for automatic detection of aneurysms on TOF-MR angiography images. To detect aneurysms that were overlooked in the initial report, two radiologists were blinded to the algorithm's detection results. The results showed that the algorithm improved the detection of aneurysms in the internal and external test sets by 4.8% and 13% respectively over the original data set. Sichterman *et al.* [15] used the CNN-based DeepMedic framework for aneurysm detection on 3D TOF-MRA images with datasets collected from different magnetic fields as well as different hospitals and showed that the algorithm has a sensitivity of 96% for aneurysms of 3–7 mm. In summary, almost all existing work focuses on medical images rather than 3D geometric point clouds.

Point-based deep learning

A 3D point cloud is characterized by a massive collection of points representing target surface properties [16] [17] [18], which is usually obtained by laser measurements or photogrammetry. This collection can reflect the real situation on the surface, such as the ground state, feature reflection characteristics, etc., with high accuracy. However, 3D points are disordered and irregular, which makes it difficult to extract general features. Traditional methods focus on converting 3D points into structured data to solve this problem. Methods such as [10] [19] [20] [21] [22] [23] regularize the structure of 3D points by voxelization or projection, which leads to the loss of intrinsic geometric information after data transformation. Therefore, direct point-based methods began to be investigated [24] [8] [25] [26] [27] [28] [29] [30] [31] [32], and others attempted to learn spatial point features directly from 3D Euclidean space.

PointNet [24], proposed by Charles R. Qi *et al.*, is the origin of point cloud learning. This type of learning can directly process the point cloud by learning the corresponding spatial encoding for each point in the input point cloud and then using the features of all points to obtain a global point cloud feature, ignoring the local relations of geometric shapes. Subsequently, Qi *et al.* proposed PointNet++, which uses query sphere grouping and hierarchical point networks to capture local structure and substantially improve the performance of the original network. However, PointNet + + did not address the spatial distribution of the input points. Tatzmon *et al.* [25] proposed a PCNN framework with extended and restricted operators to map between point-based and voxel-based representations. Volume convolution is performed on voxels for point feature extraction. SO-Net [26] selects representative points by a self-organizing map (SOM) [33] to build the spatial distribution of the point cloud. The MCCNN developed by Hermosilla *et al.* [27] can accommodate non-uniformly sampled point clouds; convolution is considered as a Monte Carlo integration problem. Similarly, PointConv [28] designs an efficient convolution with heavy density weighting that approximates convolutional networks on arbitrary 3D point sets and extends the network to a PointDeconv version for better segmentation. A more prominent element of this network is the inclusion of kernel density estimation to solve the density inhomogeneity problem. Li *et al.* [29] proposed a point cloud convolutional network, PointCNN, where they redefined convolution to perform operations on irregular point cloud data. The approach involves learning a parameter that solves the ordering

problem of point clouds to make them orderly. Tatarchenko *et al.* [30] proposed the idea of tangent line convolution, which was experimentally shown to learn surface geometric features from projected virtual tangent line images. Xu *et al.* [31] proposed SpiderCNN consisting of units called SpiderConv, which parameterizes a series of convolution filters by convolutional operations from a regular raster to an embeddable set of irregular points. DGCNN [32] addresses the problem that previous work focused only on point features and ignored the relationship between points. EdgeConv was proposed to incorporate the relationships between points into point cloud processing.

Methods

An improved end-to-end depth network structure is provided by N-PointNet (as shown in Fig. 2), which is proposed for aneurysm classification and segmentation tasks. Our innovation work is focused on two blocks: (1) the preprocessed and resampled data produce multiple layers of information embedded in the input of the original network to further enhance the characteristics of the input; (2) a side output block is added, and the multi-loss function is calculated for each corresponding layer. Both blocks provide an easy performance boost to the original network.

Data Augmentation

To ensure that different sampling data can be obtained during multi-layer embedding, the original aneurysm point cloud data were enhanced [34]. Random scaling, jitter, translation, rotation, and other methods were tried to enhance the data. Through experiments, it was found that data enhanced by random scaling and translation had better adaptability to the network proposed here. The performance on data enhancement and non-data enhancement tasks will be given in Section 5.3.

PointNet + + network

PointNet + + is a very common 3D point data processing model and was taken as the backbone of the new network, as shown in Fig. 2; the blocks in black boxes are the original network blocks. The original classification work mainly involved taking two down samples to extract global features, whereas the new network has taken four down samples, as shown in Fig. 2(a). Figure 2(b) shows that the deformed PointNet + + segmentation network looks like a "U"-shaped network with a five-layer structure, but the original PointNet + + segmentation model had only three layers. This representation makes its hierarchical encoder-decoder structure clearer. In the encoding process, each down-sampling is a set abstraction (SA) process. During the decoding process, feature propagation (FP) is used for each up-sampled linear interpolation.

Multi-layer embedding

Due to the lack of sufficient training samples for the aneurysm data, good sample features could not be obtained [35]. It was then proposed to use multi-layer input embedded learning to obtain more local geometric features for monitoring global information transfer, enriching feature representation, and greatly improving classification and segmentation of the whole network. Therefore, multi-layer embedded

inputs were added to the five-layer network. As shown in Fig. 2, multi-layer embedded inputs were created for layers L_1, L_2 and L_3 respectively.

1.Given the input point set $X = \{x_1, x_2 \ldots, x_n\}$ as the input to the Remove and MLP (RM) block (as shown in Fig. 2), a multi-layer centroid $\widehat{X_l} = \{x_1, x_2 \ldots, x_m\}$ was generated after hierarchical down sampling, where $l = 1,2,3$ indicates the number of multi-layer embedded input layers. The set of centroids is a subset of the set of input points. When a layer L_1, L_2, or L_3 undergoes down sampling through the set abstraction layer, the embedded data remove the centroid points sampled at the farthest point of this layer. The duplicate points are filtered, and the multi-layer embedded input becomes:

$$\check{X}_l = X - \widehat{X_l} = \left\{ x_i | x_i \in X, and x_i \notin \widehat{X_l} \right\}, i = 1,2,3,\ldots n.$$

1

$\check{X}_l$ then goes through an MLP [36] and expands from 6 to 32 dimensions.

2. The embedding method adopts point feature embedding (FE). The number of point features of layers L_1, L_2 and L_3 was raised from $Ml \times (d + C)$ to $Ml \times (d + C + E)$, where Ml is the input and output point set size of set abstraction level l, d denotes the dimension of the coordinates, C and E are point features. The idea was initially to perform feature propagation at the decoding layer for set segmentation of PointNet++. The original FP involved interpolating the eigenvalue f of point Ml onto the coordinates of point Ml-1, as shown in Eq. (2). This study takes the opposite approach and uses the procedure for multi-layer embedding of encoding layers. First, the size of the sampled points in each layer can be maintained. In addition to the local information (e.g., maximum value) obtained by sampling + grouping + PointNet for the centroids of each layer during the original encoding, the feature information (coordinates + normal vectors) of several of the points closest to the centroids can be obtained. To embed the information of the nearest neighbor of a centroid, the weighted k-nearest neighbor approach was used. The weight $w_i(x)$ in Eq. (2) was generated by the cosine similarity coefficient [37], which means that the shorter the distance, the greater is the weight. In addition, the values of k for layers L_1, L_2 and L_3 were respectively 3, 6, and 8.

$$f^{(j)}(x) = \frac{\sum_{i=1}^{k} w_i(x) f_i^{(j)}}{\sum_{i=1}^{k} w_i(x)}, j = 1, \ldots, C \ (2)$$

where,

$$w_i(x) = \frac{x \bullet x_i}{\| x \| \| x_i \|} = \frac{\sum_{i=1}^{n} (x \times x_i)}{\sqrt{\sum_{i=1}^{n} (x)^2} \times \sqrt{\sum_{i=1}^{n} (x_i)^2}}$$

3

.

3. Unlike the procedure in PointNet++, FP is used to copy the linear features of the upper sampling points in the decoding process during segmentation, whereas the multi-layer embedding block is used to obtain as much local feature information as possible during the encoding process [35]. Multi-layer inputs can be embedded in one layer or in multiple layers, and their performance in experimental results can also vary. This will be described in detail in Section 5.3.

Multi-loss function

In the N-PointNet classification task, side output layers were added after L_1, L_2, and L_3 respectively, with each side output layer being given a local output score [38] [39] [40]. Together with the final layer (L_4) in the original network, a total of S side output layers were created in the proposed network. Let $Y = (Y^{(1)};...;Y^{(S)})$ represent the gold standard for all side output layers, and let the corresponding parameter $y = (y^{(1)};...;y^{(S)})$ represent the local output score of each side output layer. The objective function of the side output layer is then:

$$L_s(Y, y) = \sum_{n=1}^{S} \mu_n l_s^{(n)}(Y^{(n)}, y^{(n)})$$

4

where $S = 4$, parameter μ_n is the loss function fusion weight of each side output layer ($\mu_1, \mu_2, \mu_3, \mu_4$ =0.8,0.18,0.018,0.02), and $l_s^{(n)}(\bullet)$ denotes the cross-entropy loss function for each side output layer. To use the side output prediction results, the original final score from L_4 was linearly overlaid with the local output score of each side output layer as the final prediction result. There are many benefits to adding a side output layer to the original task. First of all, the side output layer can alleviate the gradient loss problem. Back-propagating the side output loss and the final layer loss to the early layer helps in training the early layer. Second, embedding multi-layer inputs can improve the performance of the whole network, and the side output layer can monitor the output score of each level, so that more accurate results can be obtained.

Experimental Configuration

Datasets

The IntrA [6] dataset is available for point cloud data classification and segmentation. IntrA collected whole-brain 3D models of 103 patients with the vascular system. A 3D model of the whole cerebrovascular system was formed by reconstructing 2D MRA images scanned from patients. IntrA consisted of 1694 healthy vessel parts, 215 aneurysm parts, and 116 aneurysm parts that were manually annotated by experts. All 2025 samples were represented as a 3D point cloud. These 2025 vessel parts and aneurysm parts were used as the dataset for the classification in this study. The 116 annotated

aneurysm parts were used as the data set for segmentation. Moreover, the files in the *fileSplit* folder were used to divide the dataset into five subsets to facilitate five-fold cross validation.

The ModelNet40 [10] dataset is a point cloud dataset used for classification, and all the images in it are hand-drawn CAD point cloud images. There are 9843 point cloud data points in the training set and 2468 point cloud data points in the validation set, for a total of 12311 points. The dataset contains 40 categories (e.g., aircraft, table, plant, etc.), and all point cloud data in each category are stored in folders named by category. Each point is a six-dimensional vector that includes its coordinates and its normal vector. This dataset was used to compare the method proposed in this study with other point-based methods.

Experimental settings

The experiments in this paper focused on three tasks: classification of IntrA, classification of ModelNet-40, and segmentation of IntrA. The training parameters were set differently for different tasks, as shown in Table 1, wherehere *Step size* denotes the learning rate at every 20 epochs multiplied by *Lr* decay. The test was repeated three times for each subsample in the same environment, and the average of the three results was taken as the final result.

Training Environment: All the experiments were implemented in Python 3.7, Torch-GPU 1.7.1, and Torchvision 0.8.2 on Linux 64 Red Hat 8.3.1-5 with 2*Intel(R) Xeon(R) Gold 5218R CPU @ 2.10 GHz, 512G RAM, 4*NVIDIA GeForce GTX 3090 24 GB, CUDA 11.0, and CuDNN v8.0.5.

Training Details: Taking classification of IntrA as an example, 1000 epochs were selected because the accuracy stabilized after 800−1000 epochs, with no further improvement.

Table 1
Training parameters of the three tasks.

Training parameter	Task		
	Classification of IntrA	Classification of ModelNet-40	Segmentation of IntrA
Epoch	1000	1000	1000
Batch size	32	16	32
Learning rate	2×10^{-4}	10^{-4}	10^{-3}
Optimizer	Adam	Adam	Adam
Weight decay	10^{-4}	10^{-4}	10^{-4}
Step size	20 (epochs)	20 (epochs)	20 (epochs)
Lr decay	0.4	0.5	0.7

Results

The experiment focused on three main tasks: the medical 3D point (IntrA) experiment, the non-medical 3D point experiment, and the ablation experiment. First, the experimental results of aneurysm (IntrA) classification and segmentation will be presented. Second, the results of experiments on a commonly used public data set (ModelNet40) to verify the generalization performance of the model will be summarized. Finally, the specific procedure for the ablation experiment will be given.

Table 2

Classification results for each network. #Points header in the second column indicates the number of input sample points. The accuracy of the healthy vessel part (V.) and the aneurysm part (A.) as well as the F1 score were calculated in each fold, and their mean values were obtained.

Method	#Points	V.(%)	A.(%)	F1
PointNet [24]	512	94.45	67.66	0.691
	1024	94.98	64.96	0.684
	2048	93.74	69.50	0.692
PointNet++ [8]	512	98.52	86.69	0.893
	1024	98.52	88.51	0.903
	2048	98.76	87.31	0.902
PointCNN [29]	512	98.38	78.25	0.849
	1024	98.79	81.28	0.875
	2048	98.95	85.81	0.904
PointConv [28]	512	**99.21**	91.96	0.915
	1024	98.89	83.57	0.883
	2048	98.61	90.47	0.883
SO-Net [26]	512	98.76	84.24	0.884
	1024	98.88	81.21	0.868
	2048	98.88	83.94	0.885
SpiderCNN [31]	512	98.05	84.58	0.869
	1024	97.28	87.90	0.872
	2048	97.28	84.89	0.866
DGCNN [32]	512	95.22	60.73	0.658
	1024	95.34	72.21	0.738
	2048	97.93	83.40	0.859
N-PointNet	512	98.53	**92.15**	**0.926**
	1024	98.48	92.02	0.921
	2048	98.53	92.13	0.925

Medical 3D point experiments

Classification of IntrA. Table 2 gives a comparison of the algorithms commonly used in point-cloud deep learning with the five-fold cross-validation results of N-PointNet. The #Points header in the second column indicates the number of input sample points. The accuracies of the healthy vessel part (V.) and the aneurysm part (A.) as well as the F1 score were calculated in each fold, and their mean values were obtained. The corresponding results in the table are the average of all folds. As for the proposed N-PointNet with PointNet++ backbone, 512 sample points performed the best in terms of the F1 score and of aneurysm accuracy, which surpassed the original work by 3.3% and 5.46% respectively. Compared with other numbers of sample points, the proposed method achieved the highest aneurysm accuracy of 92.15% and obtained an F1 score of 0.926 for 512 sample points. This accuracy was also 1% higher than the best of the other methods (PointConv), showing the superiority of the method proposed here for 3D medical point cloud analysis.

Table 3

Segmentation results for each network. The #Points header in the second column indicates the number of input sample points. The 95% confidence interval of IOU or DSC is indicated by CI95. V. and A. denote the healthy vessel part and the aneurysm part respectively.

Method	#Points	IOU(CI95)(%)		DSC(CI95)(%)	
		V.	A.	V.	A.
PointNet[24]	512	73.99(67.43 ~ 80.56)	37.30(26.17 ~ 48.44)	84.05(79.31 ~ 88.79)	48.96(36.53 ~ 61.38)
	1024	75.23(69.10 ~ 81.36)	37.07(25.66 ~ 48.48)	85.00(80.69 ~ 89.31)	48.38(35.63 ~ 61.13)
	2048	74.22(67.85 ~ 80.60)	37.75(26.85 ~ 48.64)	84.17(79.56 ~ 88.78)	49.59(37.48 ~ 61.70)
PointNet++[8]	512	93.42(90.91 ~ 95.92)	76.22(66.70 ~ 85.73)	96.48(95.04 ~ 97.92)	83.92(75.46 ~ 92.38)
	1024	93.35(91.10 ~ 95.60)	76.38(97.96 ~ 84.80)	96.47(95.20 ~ 97.74)	84.62(77.45 ~ 91.78)
	2048	93.24(90.93 ~ 95.56)	76.21(69.99 ~ 84.43)	96.40(95.08 ~ 97.72)	84.64(77.71 ~ 91.57)
PointCNN[29]	512	92.49(89.77 ~ 95.22)	70.65(58.89 ~ 84.42)	95.97(94.41 ~ 97.54)	78.55(67.37 ~ 89.73)
	1024	93.47(91.11 ~ 95.84)	74.11(63.54 ~ 84.68)	96.53(95.20 ~ 97.86)	81.74(71.88 ~ 91.59)
	2048	93.59(91.45 ~ 95.73)	73.58(62.81 ~ 84.35)	96.62(95.43 ~ 97.80)	81.36(71.39 ~ 91.33)
SO-Net [26]	512	94.22(92.12 ~ 96.32)	80.14(73.71 ~ 86.57)	96.95(95.79 ~ 98.12)	87.90(83.14 ~ 92.66)
	1024	94.42(92.39 ~ 96.45)	80.99(74.71 ~ 87.27)	97.06(95.93 ~ 98.19)	88.41(83.65 ~ 93.17)
	2048	94.46(92.51 ~ 96.41)	81.40(75.37 ~ 87.43)	97.09(96.02 ~ 98.16)	88.76(84.38 ~ 93.15)
SpiderCNN[31]	512	90.16(86.34 ~ 93.98)	67.25(55.21 ~ 79.29)	94.53(92.17 ~ 96.88)	75.82(64.07 ~ 87.56)
	1024	87.95(83.65 ~ 92.24)	61.60(48.97 ~ 74.23)	93.24(90.56 ~ 95.93)	71.08(58.50 ~ 83.67)
	2048	87.02(82.57 ~ 91.47)	58.32(45.21 ~ 71.44)	92.17(89.94 ~ 95.47)	67.74(54.5 ~ 80.98)
PointConv [28]	512	94.16(91.76 ~ 96.56)	79.09(70.26 ~ 87.92)	96.89(95.55 ~ 98.24)	86.01(78.34 ~ 93.69)
	1024	94.59(92.53 ~ 96.66)	79.42(70.55 ~ 88.29)	97.15(96.00 ~ 98.30)	86.29(78.33 ~ 94.25)

	#Points				
	2048	94.65(92.64 ~ 96.67)	79.53(70.96 ~ 88.10)	97.18(96.06 ~ 98.30)	86.52(78.95 ~ 94.09)
N-PointNet	512	95.53(94.58 ~ 96.06	**82.35**(78.45 ~ 85.26)	97.66(97.20 ~ 98.06)	**89.02**(85.32 ~ 91.43)
	1024	**95.54**(94.69 ~ 96.06)	82.15(78.43 ~ 85.07)	**97.69**(97.22 ~ 97.97)	88.88(85.31 ~ 91.30)
	2048	95.44(94.67 ~ 96.07)	82.28(78.49 ~ 85.19)	97.62(97.21 ~ 97.98)	88.96(85.48 ~ 91.41)

Segmentation of IntrA. Table 3 gives a comparison of the current algorithms commonly used in point-cloud deep learning with the five-fold cross-validation results of N-PointNet. The #Points header in the second column indicates the number of input sample points. The 95% confidence interval of IOU or DSC is indicated by CI95. V. and A. denote the healthy vessel part and the aneurysm part respectively. The corresponding results in the table are the average of all folds. Intersection over union (IOU) and the Sørensen-Dice coefficient (DSC) show that N-PointNet has surprisingly good performance in intracranial aneurysm segmentation. As for the proposed N-PointNet with the basic PointNet + + network, 512 sample points performed the best in terms of the aneurysm part, and 1024 sample points performed the best in terms of the healthy vessel part. The IOU and DSC coefficients of the 512-sample-point aneurysm segmentation results were 6.13% and 5.1% higher than those of the original task respectively. Moreover, the IOU and DSC coefficients of the vascular segmentation results at 1024 sampling points were 2.19% and 1.22% higher than those of the original task respectively. Table 3 shows that in the vessel segmentation results, the highest IOU and DSC values achieved by segmentation in the proposed model occurred in the 1024 sampling point case, which were 95.54% and 97.69% respectively. In the intracranial aneurysm segmentation results, the highest IOU and DSC values achieved by segmentation in the proposed model were 82.35% and 88.99% respectively for the 512 sampling point case. The method proposed in this study showed a substantial performance improvement over SO-Net and PointConv, with a 1.36% improvement in IOU and a 0.61% improvement in DSC for the index of intracranial aneurysm segmentation and a 0.89% improvement in IOU and a 0.54% improvement in DSC for the index of vessel segmentation. It can be seen that the proposed model can not only improve the intracranial aneurysm segmentation effect, but also the accuracy of vascular segmentation.

The data from the five-fold cross-validation experiments were then visualized and analyzed. Figure 3 shows a comparison of the segmentation results between N-PointNet and its original model at 1024 sampling points. Figure 3(a) shows the IOU results, and Fig. 3(b) shows the DSC results. The healthy vessel part for each fold of networks is found on the left side, and the aneurysm part for each fold of networks is found on the right side. These methods were then compared with their best performance. The numbers 0–4 denoted the five subsets, which were divided by the files (*annSplit_0.txt ~ annSplit_4.txt*) in the *fileSplit* folder of the IntrA dataset. In addition, Fig. 4 shows the segmentation results of the three models for some cases with 1024 sampling points. The first row shows the PointNet segmentation results, the second row the PointNet + + segmentation results, the third row the N-PointNet segmentation

results, and the last row the Ground truth. The visual analysis results reveal that the proposed model can improve the results of the original task in both the healthy vessel part and the aneurysm part and that the segmentation results of the deep network based on multilayer feature embedding are better. The results were closer to Ground truth.

Non-medical 3D point experiments

Classification of ModelNet40. Table 4 describes in detail the performance of the various methods in point cloud deep learning for the ModelNet40 dataset on the classification task. Only 1024 samples were used, but the data were then augmented with jitter. Comparison with the accuracy of other methods reveals that the model proposed in this paper is more capable than the original and can reach the average level. N-PointNet was found to have excellent performance, not only on medical data, but also on a general dataset, providing good classification results. This shows the strong generalization ability of N-PointNet.

Table 4
Classification results of different methods on ModelNet40

Method	Input	Accuracy (%)
PointNet[24]	xyz	89.2
SO-Net[26]	xyz + norm	90.9
PointNet++[8]	xyz + norm	91.9
SpiderCNN[31]	xyz + norm	92.4
PointCNN[29]	xyz	92.5
PointConv[28]	xyz + norm	92.5
DGCNN[32]	xyz	**92.9**
N-PointNet	xyz + norm	92.8

Ablation studies

This section describes an investigation of the effectiveness of data enhancement and also discusses the effect of multi-layer embedding on the experimental results under different conditions. In this section of the ablation experiments, the first fold was used as the test set and was calculated with 1024 sampling points if not otherwise specified.

Augmentation. Random scaling, jitter, translation, rotation, and other methods were considered for data augmentation. The objective was to find the most suitable data augmentation method for the proposed model. When using random scaling and translation, the combination of preprocessing methods and the proposed model gave the best results. Table 5 shows the ablation study on data augmentation for the segmentation task.

Table 5
Ablation study on data augmentation

Augmentation	IOU(%)		DSC(%)	
	V.	A.	V.	A.
Rotation	93.22	79.36	97.08	86.22
Jitter	95.36	82.02	97.71	88.58
Random scaling	94.95	80.13	97.32	87.43
Random scaling and translation	95.26	82.05	97.52	88.61
Random scaling, translation, and jitter	**95.54**	**82.15**	**97.69**	**88.88**

Multi-Layer Embedding. To study the effect of multi-layer embedding, different models were created in the segmentation task for comparison experiments. The segmentation results for the IOU and DSC indices are shown in Table 6. When the k-nearest neighbor features were embedded in the L_1 layer of the encoding process, the model segmentation effect was greatly improved. When continuing to embed local features in the L_2 and L_3 layers, the effect was only slightly improved over the original basis. The reason for this was that the L_2 and L_3 layers were obtained by down-sampling from the L_1 layer, so that the k-nearest neighbor information of the L_1 layer could be passed to the L_2 and L_3 layers. When continuing to embed k-nearest neighbor features in the L_2 layer, the number of k-nearest neighbor points could only be increased to slightly more than the original value ($k = 3$ in the L_1 layer, $k = 6$ in the L_2 layer, and $k = 8$ in the L_3 layer). The closer the sampled points are to each other, the greater the size of the point cloud segmentation task. Therefore, increasing the number of embedding layers can improve performance by a small margin, but only three-layer embedding is used in this paper.

Table 6
Ablation study on multi-layer embedding

Method	IOU(%)		DSC(%)	
	V.	A.	V.	A.
PointNet++	93.35	76.38	96.47	84.62
PointNet++ (+ 1 layer)	95.10	81.32	97.40	87.04
PointNet++ (+ 2 layers)	95.38	81.98	97.51	88.04
N-PointNet	**95.54**	**82.15**	**97.69**	**88.88**

Multi-Loss Function. To study the effectiveness of the multi-loss function, different models were built and subjected to experiments; the classification accuracy results and the F1 index are given in Table 7. The experimental results verify that the multi-loss function mentioned in Section 3.4 improves the performance of the hierarchical deep network model.

Table 7
Ablation study on multi-loss function

Method	V⁤(%)	A⁤(%)	F1
N-PointNet (without multi-loss function)	98.40	91.83	0.914
N-PointNet	**98.48**	**92.02**	**0.921**

Discussion

This article provides an improved encoder-decoder depth network structure (N-PointNet), which is used for aneurysm classification and segmentation tasks. The preprocessed and resampled data produce multiple layers of information embedded in the input of the original network to further enhance the characteristics of the input. The side output block that calculates the multi-loss function for each corresponding layer monitors the output score of each level, so that more accurate results can be obtained. Both blocks provide an easy performance boost to the original network.

The experimental results show that this method has great advantages in medical and non-medical point cloud datasets. Among them, the classification tasks in the medical dataset (IntrA) obtained the best score on aneurysm accuracy (92.15%) with F1 score (0.926), as shown in Table 2. The segmentation task has the highest score on the IOU index of vessels (95.54%) and aneurysms (82.35%), as well as the DSC index of vessels (97.69%) and aneurysms (89.02%), as shown in Table 3. In addition, it has relatively good performance for the non-medical dataset (ModelNet40). The adaptability of the proposed algorithm to other types of data is thereby proved.

Meanwhile, it is interesting to find that compared to the performance of other methods with 512 samples, the effect is improved simultaneously with 1024 and 2048 samples, and even slightly more than with more samples. This is due to the fact that multi-layer embedding can obtain more different information when resampling than the original sample points, especially if the original data themselves contain more sample points.

Although the algorithm proposed in this paper performs well on both medical and non-medical point cloud datasets, the algorithm has much room for improvement. For example, the next consideration is to optimize the model to reduce its complexity and improve its computational speed. In addition, how to improve the proposed algorithm to recognize tiny aneurysms is not obvious. Further research will be conducted to solve the problem of tiny aneurysm recognition.

Conclusion

An improved encoder-decoder depth network structure (N-PointNet) is proposed for IA classification and segmentation. It consists of a basic PointNet + + network, a multi-layer embedding block, and a side

output block. This study has demonstrated that the performance of PointNet + + can be easily improved beyond the current state of the art by introducing multi-layer embedding and side output blocks. More specifically, inserting a multi-layer embedding block into the input of the original network can further improve the input features, and the multi-loss function in the side output block facilitates fast convergence of the model and improves its performance. Experiments showed that N-PointNet not only performs well on medical 3D point cloud datasets, but also has good generalization ability on non-medical 3D point cloud datasets. Therefore, the model proposed in this paper still has limitations. In particular, the improvement of the recognition effect for tiny aneurysms is not significant. Further studies will be conducted to address such issues.

Abbreviations

IA: Intracranial Aneurysm

CT: computed tomography

MRA: magnetic resonance angiography

DL: deep learning

CAD: computer-aided diagnosis

MLP: Multilayer Perceptron

SA: set abstraction

FP: feature propagation

RM: Remove and MLP

FE: feature embedding

Declarations

Acknowledgements

Author Contributions

Jiaqi Wang performed the methodology and wrote the manuscript; Juntong Liu performed the data analysis and validation; Zhengyuan Xu performed the formal analysis; Pengzhan Yin performed the

conceptualization; Jinlong Yuan performed the supervision; Yunfeng Zhou performed the review and supervision; Mingquan Ye performed the review, editing, and supervision.

Data Availability Statements

The dataset analyzed during the current study are available in the IntrA and ModelNet40 dataset. These data were derived from the following resources available in the public domain:

1. https://github.com/intra3d2019/IntrA

2. https://modelnet.cs.princeton.edu/

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no conflicts of interest.

Author details

[1]Medical Information School of Wannan Medical College,Wuhu,241002,China

[2]Medical Imaging School of Wannan Medical College,Wuhu,241002,China

[3]Department of Radiology, the First Affiliated Hospital of Wannan Medical College, Wuhu, 241001, China

[4]Department of Neurosurgery, the First Affiliated Hospital of Wannan Medical College, Wuhu, 241001, China

[5]Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, 230001, China

References

1. He S, Wei ML, Xie F, Richard SA. A fenestrated persistent primitive hypoglossal artery harboring a ruptured aneurysm: A case report. Medicine. 2021;100(32):26904–9.

2. Hiraoka D, Manabe S, Hirooka K, Hirayama D, Onuki M. Surgical treatment for Takayasu arteritis complicated with thoracic aneurysm. Japanese J Cardiovasc Surg. 2018;47(6):289–92.

3. Gmeiner M, Dirnberger J, Fenz W, Gollwitzer M, Wurm G, Trenkler J, et al. Virtual cerebral aneurysm clipping with real-time haptic force feedback in neurosurgical education. World Neurosurg. 2018;112:313–23.

4. Wang J, Yuan Z, Qian G, Bao W, Jin G. 3D printing of intracranial aneurysm based on intracranial digital subtraction angiography and its clinical application. Medicine. 2018;97(24):11103–9.

5. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. (2016) 3D U-Net: Learning dense volumetric segmentation from sparse annotation. International Conference on Medical Image Computing and Computer-Assisted Intervention, pp 424–432.

6. Yang X, Xia D, Kin T, Igarashi T. (2020) IntrA: 3D intracranial aneurysm dataset for deep learning. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 2656–2666.

7. Schneider L, Niemann A, Beuing O, Preim B, Saalfeld S. MedMeshCNN - Enabling MeshCNN for medical surface models. Comput Methods Programs Biomed. 2021;210:106372.

8. Qi CR, Yi L, Su H, Guibas LJ. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. Adv Neural Inf Process Syst. 2017;30:5099–108.

9. Balatsoukas-Stimming A, Studer C. (2019) Deep unfolding for communications systems: A survey and some new directions. 2019 IEEE International Workshop on Signal Processing Systems (SiPS), pp 266–271.

10. Wu Z, Song S, Khosla A, Yu F, Zhang L, Tang X et al. (2015) 3D shapenets: A deep representation for volumetric shapes. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 1912–1920.

11. Fujita H. AI-based computer-aided diagnosis (AI-CAD): The latest review to read first. Radiol Phys Technol. 2020;13(1):6–19.

12. Shi Z, Hu B, Schoepf U, Savage R, Dargis D, Pan C, et al. Artificial intelligence in the management of intracranial aneurysms: Current status and future perspectives. Am J Neuroradiol. 2020;41(3):373–9.

13. Stember JN, Chang P, Stember DM, Liu M, Grinband J, Filippi CG, et al. Convolutional neural networks for the detection and measurement of cerebral aneurysms on magnetic resonance angiography. J

Digit Imaging. 2019;32(5):808–15.

14. Ueda D, Yamamoto A, Nishimori M, Shimono T, Doishita S, Shimazaki A, et al. Deep learning for MR angiography: Automated detection of cerebral aneurysms. Radiology. 2019;290(1):187–94.

15. Sichtermann T, Faron A, Sijben R, Teichert N, Freiherr J, Wiesmann M. Deep learning-based detection of intracranial aneurysms in 3D TOF-MRA. Am J Neuroradiol. 2019;40(1):25–32.

16. Serafin J, Grisetti G. (2015) NICP: Dense normal based point cloud registration. 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp 742–749.

17. Zeng J, Wang D, Chen P. A survey on transformers for point cloud processing: An updated overview. IEEE Access. 2022;10:86510–27.

18. Guo M, Cai J, Liu Z, Mu T, Martin RR, Hu S. PCT: Point cloud transformer. Comput Visual Media. 2021;7(2):187–99.

19. Qi CR, Su H, Nießner M, Dai A, Yan M, Guibas LJ. (2016) Volumetric and multi-view cnns for object classification on 3D data. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 5648–5656.

20. Maturana D, Scherer S. (2015) Voxnet: A 3D convolutional neural network for real-time object recognition. 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp 922–928.

21. Xie J, Dai G, Zhu F, Wong EK, Fang Y. Deepshape: Deep-learned shape descriptor for 3D shape retrieval. IEEE Trans Pattern Anal Mach Intell. 2016;39(7):1335–45.

22. Wu J, Zhang C, Xue T, Freeman WT, Tenenbaum JB. (2016) Learning a probabilistic latent space of object shapes via 3D generative-adversarial modeling. Proceedings of the 30th International Conference on Neural Information Processing Systems, pp 82–90.

23. Su H, Maji S, Kalogerakis E, Learned-Miller E. (2015) Multi-view convolutional neural networks for 3D shape recognition. Proceedings of the IEEE International Conference on Computer Vision, pp 945–953.

24. Qi CR, Su H, Mo K, Guibas LJ. (2017) PointNet: Deep learning on point sets for 3D classification and segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 652–660.

25. Atzmon M, Maron H, Lipman Y. Point convolutional neural networks by extension operators. ACM Trans Graphics (TOG). 2018;37(4):71:1–71.

26. Li J, Chen BM, Lee GH. (2018) So-Net: Self-organizing network for point cloud analysis. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 9397–9406.

27. Hermosilla P, Ritschel T, Vázquez PP, Vinacua À, Ropinski T. Monte Carlo convolution for learning on non-uniformly sampled point clouds. ACM Trans Graphics (TOG). 2018;37(6):1–12.

28. Wu W, Qi Z, Fuxin L. (2019) PointConv: Deep convolutional networks on 3D point clouds. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp 9621–9630.

29. Li Y, Bu R, Sun M, Wu W, Di X, Chen B. (2018) PointCNN: Convolution on x-transformed points.Advances in Neural Information Processing Systems, pp828–838.

30. Tatarchenko M, Park J, Koltun V, Zhou Q. (2018) Tangent convolutions for dense prediction in 3D. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 3887–3896.

31. Xu Y, Fan T, Xu M, Zeng L, Qiao Y. (2018) SpiderCNN: Deep learning on point sets with parameterized convolutional filters. Proceedings of the European Conference on Computer Vision (ECCV), pp 87–102.

32. Wang Y, Sun Y, Liu Z, Sarma SE, Bronstein MM, Solomon JM. Dynamic graph cnn for learning on point clouds. ACM Trans Graphics (TOG). 2019;38(5):1–12.

33. Kohonen T. (1990) The self-organizing map. Proceedings of the IEEE.78(9):1464–1480.

34. Shao D, Lu X, Liu X. (2022) 3D intracranial aneurysm classification and segmentation via unsupervised dual-branch learning.IEEE Journal of Biomedical and Health Informatics:1–10.

35. Fu H, Cheng J, Xu Y, Wong DWK, Liu J, Cao X. Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. IEEE Trans Med Imaging. 2018;37(7):1597–605.

36. Yu T, Li X, Cai Y, Sun M, Li P. (2022) S2-MLP: Spatial-shift MLP architecture for vision. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp 297–306.

37. Luo C, Zhan J, Xue X, Wang L, Ren R, Yang Q. (2018) Cosine normalization: Using cosine similarity instead of dot product in neural networks. International Conference on Artificial Neural Networks, pp 382–391.

38. Li G, Yu Y. Visual saliency detection based on multiscale deep CNN features. IEEE Trans Image Process. 2016;25(11):5012–24.

39. Liu Y, Cheng M, Hu X, Wang K, Bai X. (2017) Richer convolutional features for edge detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 3000–3009.

40. Huang L, Xia W, Zhang B, Qiu B, Gao X. MSFCN-multiple supervised fully convolutional networks for the osteosarcoma segmentation of CT images. Comput Methods Programs Biomed. 2017;143:67–74.
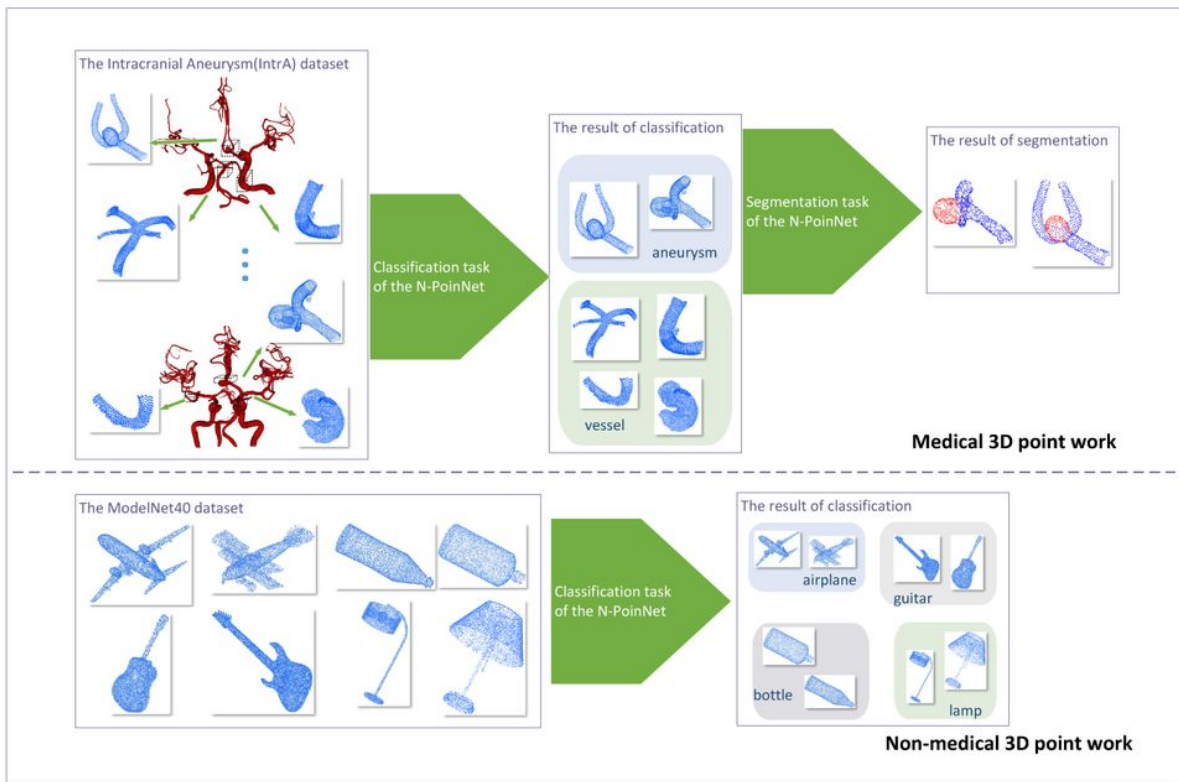
Figures

Figure 1

Workflow chart of the developed N-PointNet classification and segmentation system. Top part: classification and segmentation of the intracranial aneurysm (IntrA) dataset with N-PointNet; Bottom part: classification of the Modelnet40 dataset with N-PointNet.
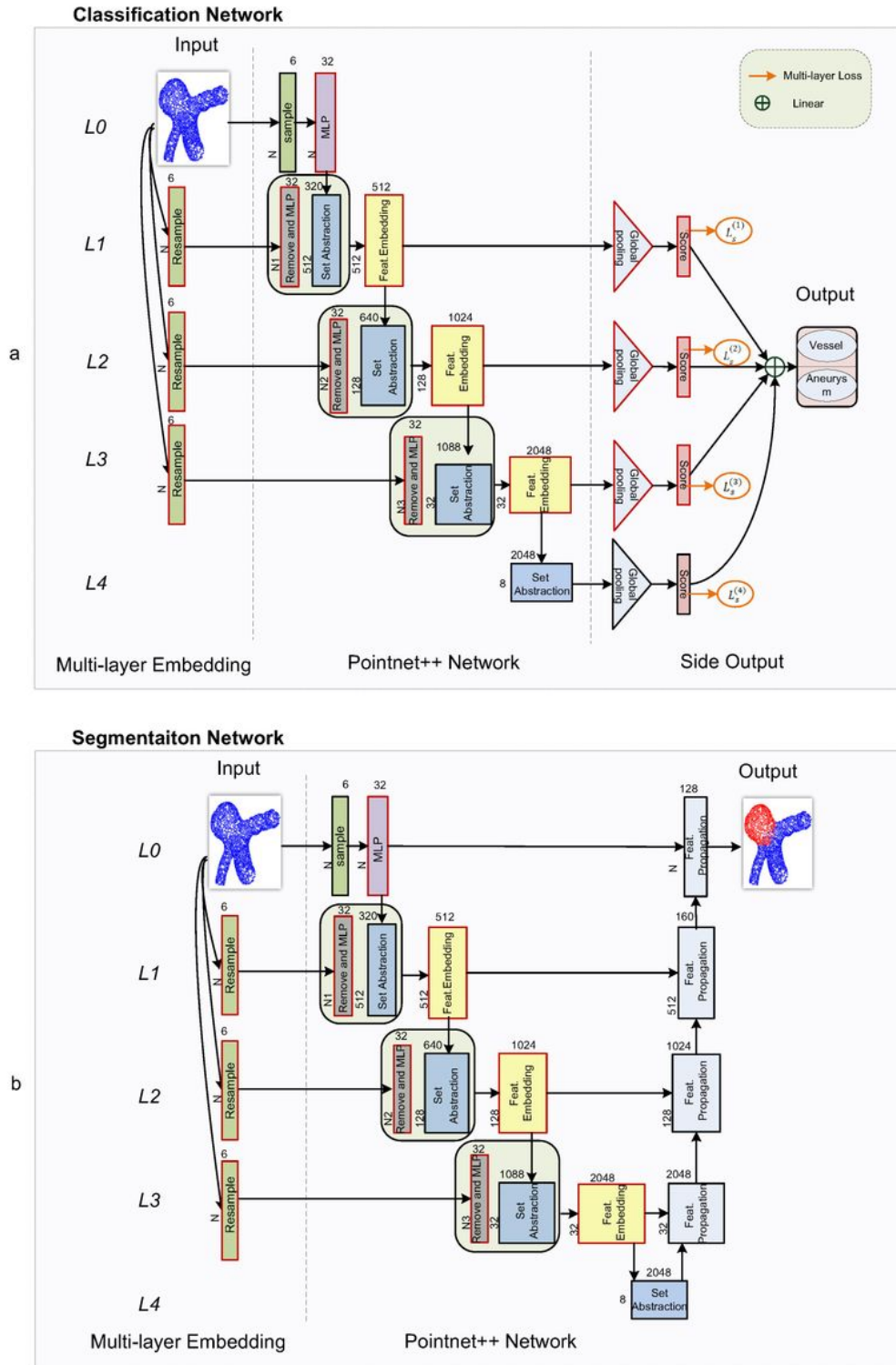
Figure 2

N-PointNet network. PointNet++ was taken as the backbone of the new network. The blocks in black boxes are the original network blocks, and the blocks in red boxes are newly added blocks: **a** the classification network; **b** the segmentation network.
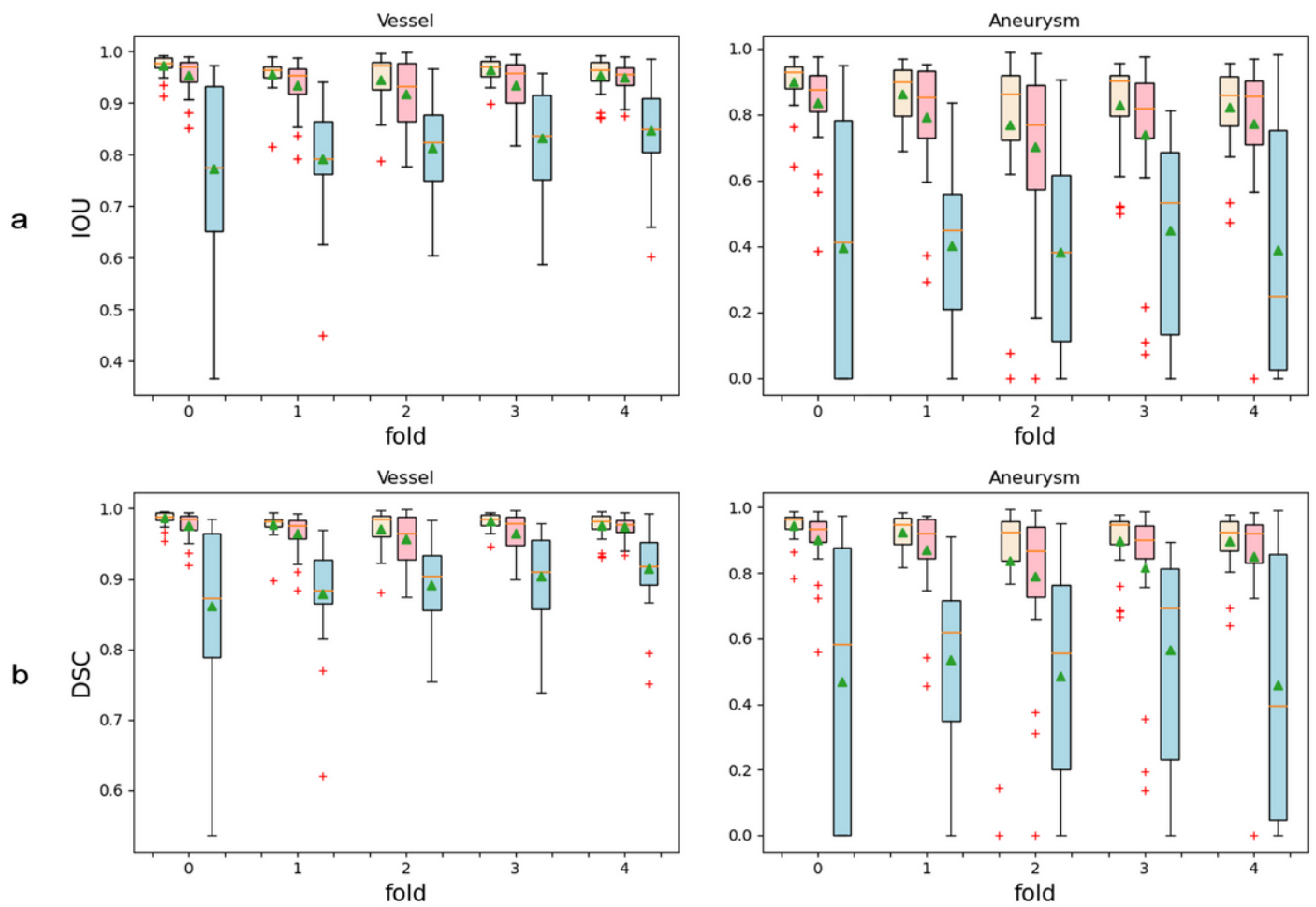
Figure 3

Box line diagram of the three networks. These networks are the result of segmentation over 1024 sampling points. The healthy vessel part is found on the left side, and the aneurysm part is found on the right side: **a** IOU results for each fold of networks; **b** DSC results for each fold of networks.
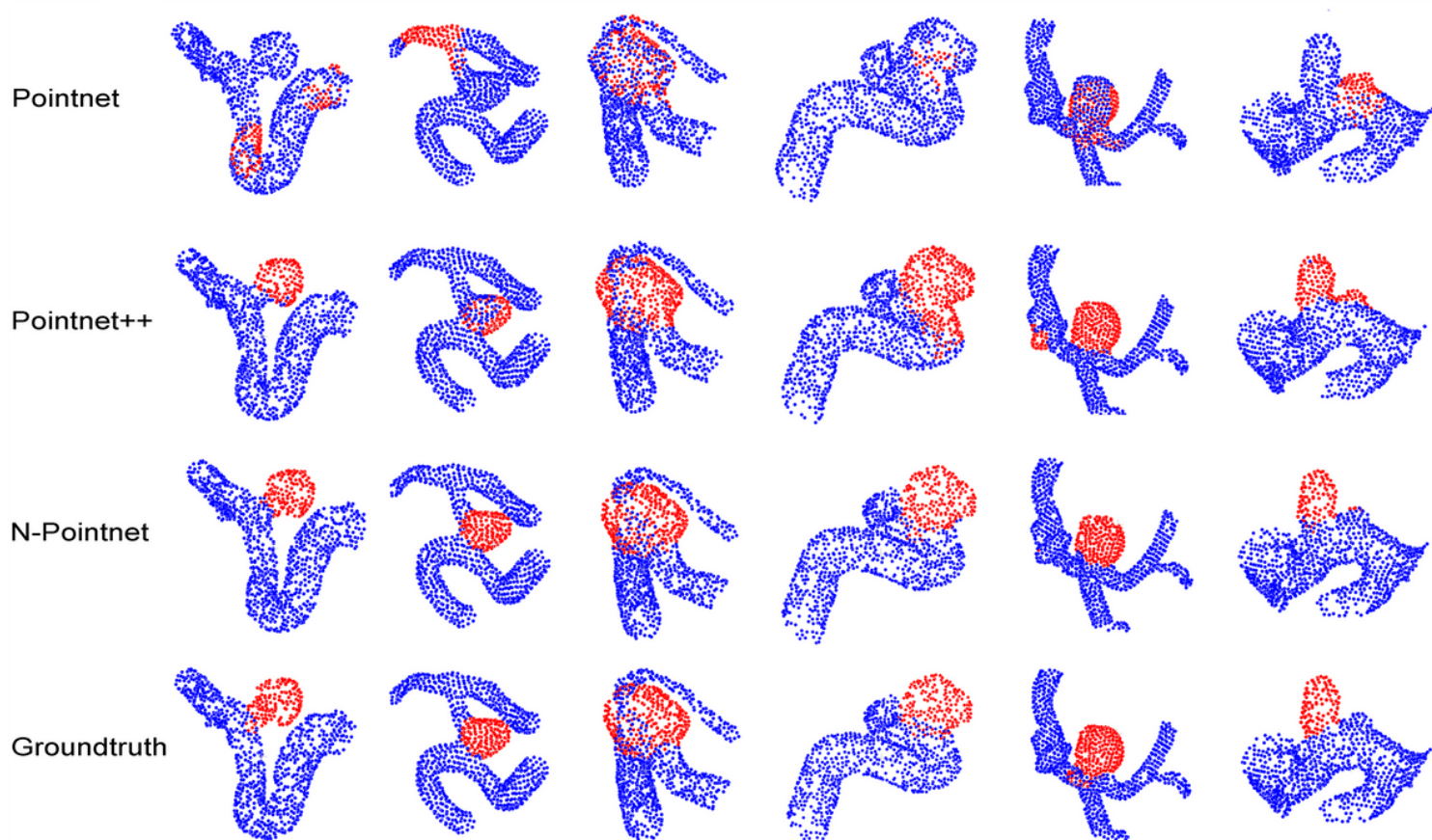
Figure 4

Results comparison for the three networks. These networks are the result of point cloud segmentation on 1024 sampling points. The first row shows the PointNet segmentation results; the second row the PointNet++ segmentation results; the third row the N-PointNet segmentation results; and the last row the Ground truth.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- GA.pdf