ORIGINAL RESEARCH ARTICLE

CA-UNet Segmentation Makes a Good Ischemic Stroke Risk Prediction
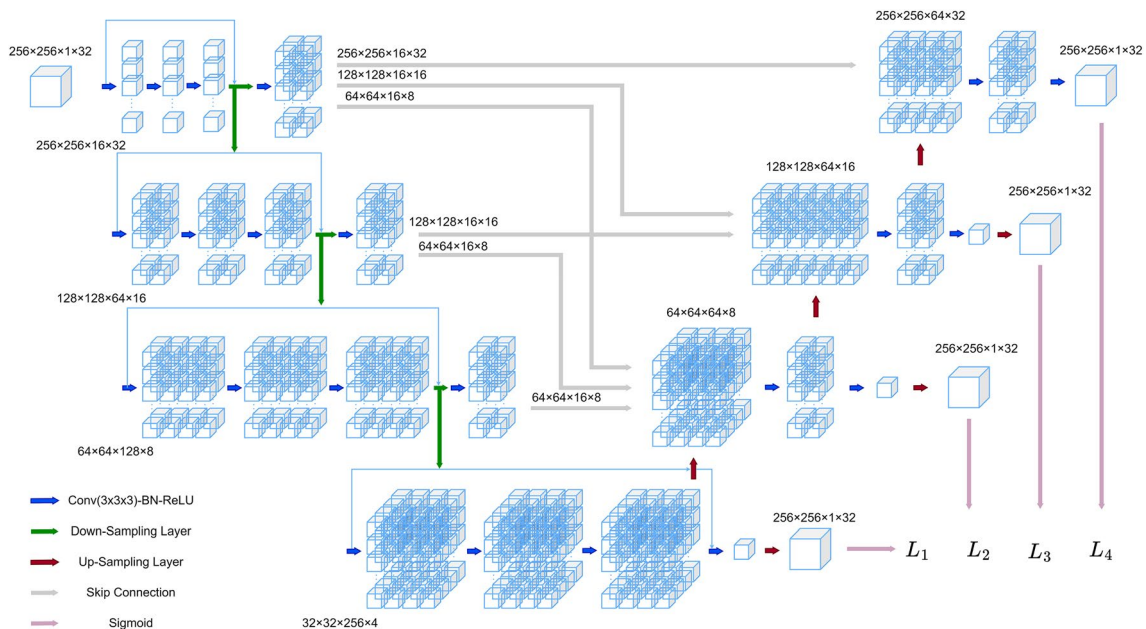
Yuqi Zhang[1,2] · Mengbo Yu[1,2] · Chao Tong[1,2] · Yanqing Zhao[3] · Jintao Han[3]

Abstract

Stroke is still the World's second major factor of death, as well as the third major factor of death and disability. Ischemic stroke is a type of stroke, in which early detection and treatment are the keys to preventing ischemic strokes. However, due to the limitation of privacy protection and labeling difficulties, there are only a few studies on the intelligent automatic diagnosis of stroke or ischemic stroke, and the results are unsatisfactory. Therefore, we collect some data and propose a 3D carotid Computed Tomography Angiography (CTA) image segmentation model called CA-UNet for fully automated extraction of carotid arteries. We explore the number of down-sampling times applicable to carotid segmentation and design a multi-scale loss function to resolve the loss of detailed features during the process of down-sampling. Moreover, based on CA-Unet, we propose an ischemic stroke risk prediction model to predict the risk in patients using their 3D CTA images, electronic medical records, and medical history. We have validated the efficacy of our segmentation model and prediction model through comparison tests. Our method can provide reliable diagnoses and results that benefit patients and medical professionals.

Graphical Abstract

Extended author information available on the last page of the article

1 Introduction

Based on the Global Stroke Fact Sheet 2022 published by the World Stroke Organization [1], stroke is still the World's second major factor of death, as well as the third major factor of death and disability. It is estimated that more than US$891 billion is spent globally for the treatment and prevention of stroke, which is 1.12% of the global GDP [2]. Stroke has become one of the major threats to global public health. The lifetime risk of stroke has risen by 50 percent in the last 20 years% [3]. In many countries, stroke disease has shown explosive growth with the accelerated urbanization, aging of society, and the prevalence of unhealthy lifestyles. Therefore, the prevention and treatment of stroke are facing great challenges.

Stroke is associated with high mortality, recurrence and disability rates [4], and can be classified into two categories based on etiology: ischemic and hemorrhagic strokes. Ischemic stroke patients account for over 80% of all stroke patients [5]. Common ischemic stroke causes are atherosclerotic plaques and carotid stenosis in the carotid region. Hence, the early detection of atherosclerosis and carotid stenosis is a fundamental approach to preventing ischemic stroke and assessing the risk of development.

Image segmentation is an important method to assist in the diagnosis of stroke. Over the past few decades, lots of segmentation methods for vessels have emerged [6–11]. In traditional vessel segmentation studies, researchers have proposed many methods [12–16], such as the hierarchical region growth algorithm [17], semi-automatic carotid segmentation algorithm based on a self-adaptive segmentation algorithm [18], the decision mechanism for venous bone separation points [19], the random wandering algorithm [20], the shortest path faster algorithm [21] and the method based on Hession matrix [22]. However, most traditional image segmentation methods require manual initialization or interactive operations, which cannot be fully automated.

In recent years, deep learning-based segmentation methods have been used with good results in various medical tasks [23–28]. For example, CarotidNet for 3D CTA carotid segmentation [29], 3D-UNet for coronary CCTA image segmentation [30], the SC2Net for detecting the detection of COVID-19 in X-rays [31], the atlas-based organ segmentation network MTL-ABS3Net [32], the CT image segmentation algorithm for liver [33], the transformers for 3D Medical Image Segmentation [34], combo loss-based spatio-temporal feature fusion network for coronary artery segmentation [35], MSRF-Net for medical image segmentation [36] and SONNET for cell nucleus segmentation [37]. Nevertheless, the application of deep learning-based methods for the carotid segmentation of 3D CTA images is still relatively few. It fails to achieve better results due to the limitation of privacy protection and labeling difficulties.

With the development of computer computing power [38], research on stroke risk prediction based on deep learning and machine learning has received more attention. Khosla et al. [39] proposed a new automatic feature selection algorithm for stroke risk prediction. Dritsas et al. [40] used machine learning techniques to predict stroke risk. Teoh et al. [41] predicted stroke risk from electronic health records. Arslan et al. [42] used data mining methods to predict stroke. However, these methods do not fully take advantage of clinical text and image data. They are limited by the difficulty of data acquisition and cannot obtain higher accuracy.

Given the above problems, we propose a 3D carotid CTA image segmentation model called CA-UNet and an ischemic stroke risk prediction model. CA-UNet is based on the encoder-decoder structure and improves on the down-sampling scheme, which could decrease the model parameters and effectively accelerate convergence speed. Skip connections enable the network to have information at each scale when decoding features at different scales. And we proposed a new fusion loss function for the characteristics of the task and introduced multi-scale training to balance the model's learning direction. Besides, our ischemic stroke risk prediction model is a fusion prediction network model that uses multiple data to predict jointly. It consists of a 3D image feature extraction network that uses carotid CTA images for prediction and a machine learning model that uses electronic medical records and medical history for prediction. The model predicts the risk of morbidity by the fusion of weights, fully uses clinical information and achieves good results. We validated the effectiveness of our models through comparative tests on our dataset. Our three contributions are listed below:

1. We propose a model which is more applicable for 3D CTA carotid segmentation.
2. We propose a multi-scale loss function for joint training which could solve the problem that features of image details would be lost in the process of down-sampling.
3. The proposed model for predicting the risk of ischemic stroke can effectively predict the risk of ischemic stroke in patients. The model could make a significant contribution to public health security.

2 Materials and Methods

2.1 Dataset

2.1.1 Private Dataset

The image data in the private dataset contain CTA images of 42 patients for the segmentation task and CTA images of 390 patients for the ischemic stroke risk prediction task. These data were provided by the partner hospitals and were desensitized for use in the study. Approximately 31,000 CTA images of 42 patients used for the segmentation task were annotated by two or more radiologists. We randomly selected 25 sets of 3D CTA images to be used for training and the remaining 17 sets of 3D CTA images to be used for testing. The CTA images of 390 patients for the ischemic stroke risk prediction task comprised approximately 290,000 CTA images. We randomly chose 80% of them as training samples and the others as test samples. Table 1 illustrates the sample distribution.

The text data in the private dataset contains electronic medical records and the medical history of 390 patients used for the ischemic stroke risk prediction task. The text data categories are age, gender, blood glucose level, body mass index (BMI), smoking status, type of residence, type of work, marital status, history of heart disease and history of hypertension.

2.1.2 Public Dataset

The public dataset, which contains only text data, is from the Stroke Prediction Dataset in Kaggle [43], with 4908 patient data. The public dataset assists in machine learning model training as the sample size of text data in the private dataset is relatively small. We randomly selected 80% of the samples as training samples and the rest as test samples, and the sample distribution is shown in Table 2. The data categories of the public dataset used are the same as the private dataset.

2.2 Data Preprocessing

2.2.1 Carotid Segmentation Task

In the carotid segmentation task, the preprocessing operation performed on the dataset is divided into two steps. First, the

Table 1 Sample distribution of private datasets for risk prediction

	Training set	Test set	Total
Positive samples	154	42	196
Negative samples	158	36	194
Total	312	78	390

**Table 2** Data distribution of public dataset

	Training set	Test set	Total
Positive samples	156	53	209
Negative samples	3770	929	4699
Total	3926	982	4908

image pixel values are intercepted according to the target of interest. The pixel values in the CTA images are called Hounsfield Units(HU). CTA imaging has a wide range, and the HU of the human body ranges from -1000 to $+1000$, for a total of 2000 values. Humans cannot distinguish such minor grayscale differences. Hence, the radiologists adjust the Window Width and Window Center of the CTA image according to the actual condition to see the target better. Based on the above idea, we eliminated the interference of irrelevant parts by limiting the range of HU values of CTA images, i.e., setting two HU thresholds: the minimum HU and the maximum HU. If the HU value of a pixel in the image is smaller than the minimum HU or larger than the maximum HU, the HU value is truncated to the threshold. As shown in Fig. 1, by this processing, the influences of most other tissues outside the range of carotid HU values are excluded, which can effectively reduce mis-segmentation and facilitate network training.

Next, the CTA image data were resampled and down-sampled in the cross-section. Since the CTA images used in this paper came from multiple scanning devices, the pixel z-axis spacing of the data is not uniform. When using 3D CTA images for deep learning model training, different slice thicknesses can impact the scale of extracted features, leading to degradation of model performance. Therefore, the input CTA images were resampled by trilinear interpolation. During resampling, the center of the input image is kept constant. The images were remapped to a spatial coordinate system with a pixel pitch of (1, 1, 1) mm, ensuring that all samples had the same pixel pitch. Since the 3D images used in this paper occupy hundreds of times more space than the 2D images, even if each batch is trained with only one set of 3D images for training, the GPU processor memory limitation causes the deep learning model to be untrained. Therefore, the input CTA images were down-sampled in the cross-section, and the cross-section size was adjusted from 512×512 to 256×25. The down-sampling was carried out by trilinear interpolation. Figure 1 shows the cross-sectional comparison of CTA images before and after image pre-processing.

2.2.2 Ischemic Stroke Risk Prediction Task

The two preprocessing steps described above in the ischemic stroke risk prediction task were also used to preprocess CTA

Fig. 1 Data preprocessing in carotid segmentation task

(a) Original CTA image

(b) After HU value interception

(c) After resampling and down-sampling
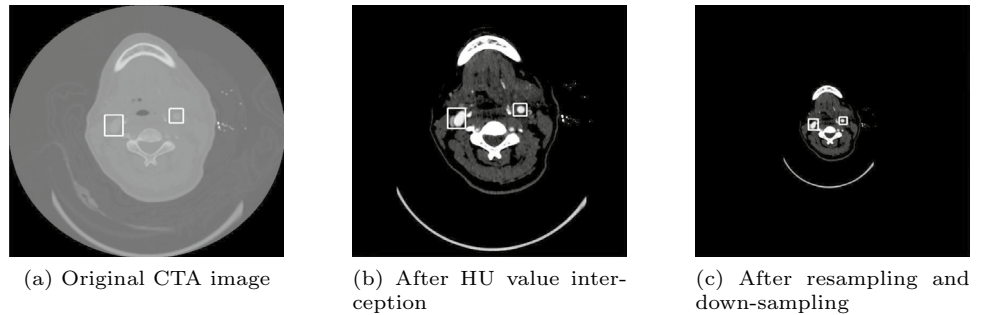
images. The difference was that we needed to intercept the network model's carotid artery region as the input image. The input to the image feature extraction network in the ischemic stroke risk prediction model is the segmented carotid region, which could be regarded as a subsequent task to the carotid segmentation task. The input was a 3D CTA image of the carotid region rather than the whole set of complete CTA images. Therefore, we selected 2 cm upward and 3 cm downward as the target area for interception according to the professional doctor's recommendation, i.e., we took 20 images upward and 30 images downward centered on the carotid bifurcation in the section direction, totaling 50 section images. Then, we intercepted the carotid artery region according to the image segmentation results. To facilitate network training, we adjusted the size of the intercepted carotid artery region to $40 \times 40 \times 50$.

Next, the public dataset was preprocessed. The public dataset contains some missing items, and the gaps need to be processed. The treatment of missing values was attempted, including deleting these records and filling the gaps with average values. After selection, we used a decision tree to predict the missing values.

2.3 CA-UNet Model

2.3.1 Main Structure

Compared with 2D segmentation tasks, the segmented target carotid region has prominent overall structural characteristics. The 3D structure of the central region of the carotid artery is shown in Fig. 2. The carotid artery can be roughly divided into the common carotid artery below the bifurcation, the internal carotid artery above the bifurcation, and the external carotid artery. The size of the common carotid artery is larger, while the size of the internal and external carotid arteries becomes smaller as the blood vessel extends upward. To adequately parse the information in the CTA images, the CA-UNet model adopts an encoder-decoder structure, and the CA-UNet is shown in Fig. 3. The left side is a contracting path and the right side is an expanding path. The primary function of the contracting path is to extract the feature from 3D CTA images. The expanding path fuses
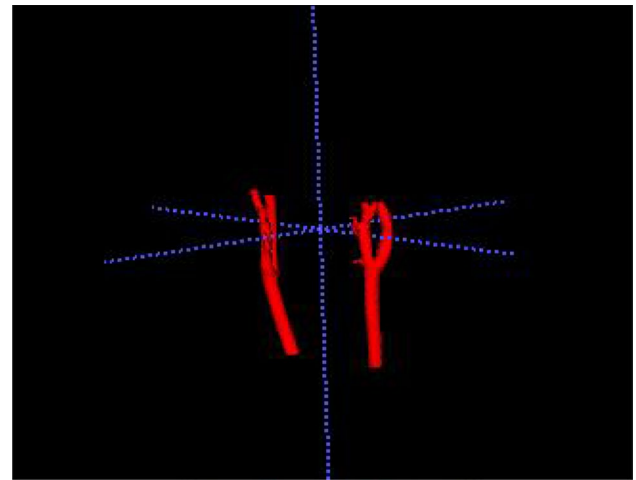
Fig. 2 3D structure of the main region of the carotid artery

multi-scale features and gradually restores the feature map to the identical size as the input image.

In the past work, image features of various scales in the model were not related. Hence, we use skip connections, which enable the network to get information at each scale when decoding features at different scales. We fuse features from different scale feature maps by skip connections which enables the network to decode the features of each layer with information of each scale and improves the network performance. We take the feature map of the third layer as an example. Its input data have two sources, one is the jump connections of the first three layers of the feature map in the contracting path, and the other is the up-sampling of the feature map in the layer above it, i.e., the feature map in the fourth layer of the extension path. In order to fuse the above four groups of feature maps, two problems need to be solved. The first problem is the different sizes of feature maps at different scales. To solve this problem, we use max pooling. For example, the first layer of the contracting path has an output feature map size of $256 \times 256 \times 32$, which is pooled using a 3D max pooling layer with the $4 \times 4 \times 4$ window size and 4 step size. The second problem is that the number of channels of feature maps at different scales. The number of channels of deep feature maps can be tens of
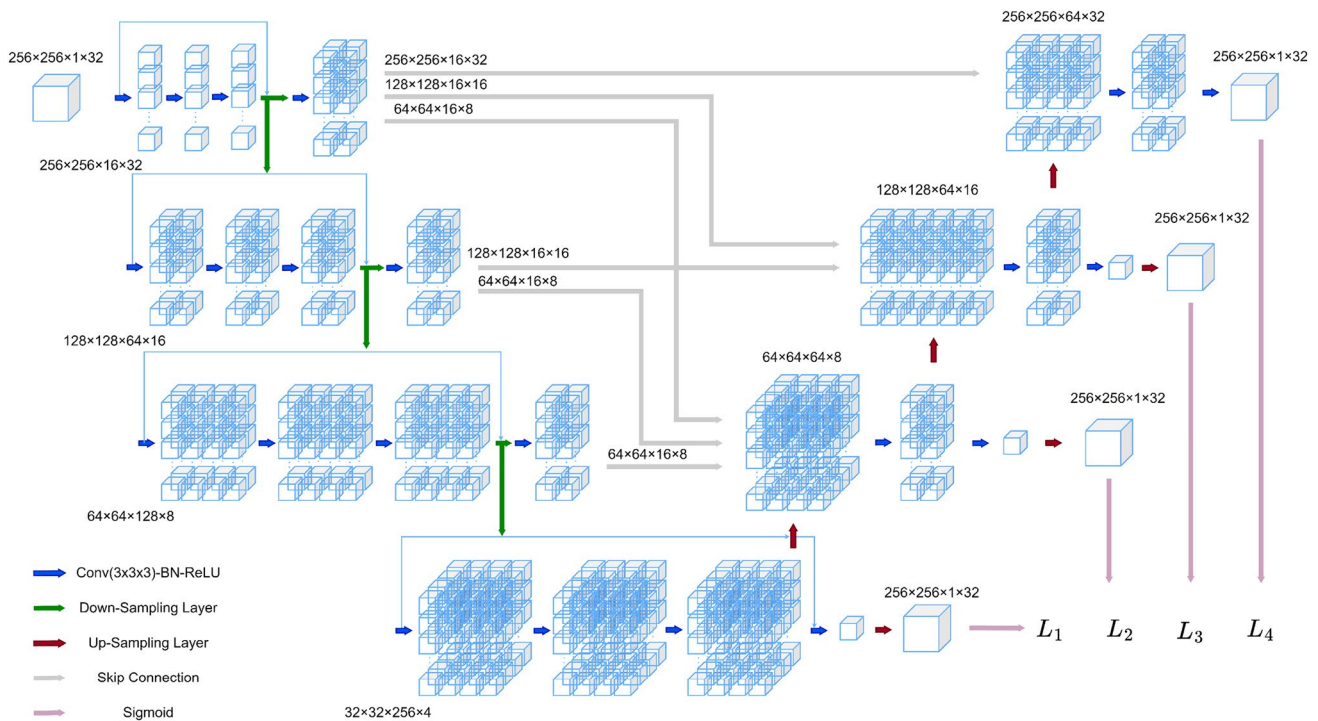
Fig. 3 CA-UNet. The input 3D CTA image size is $256 \times 256 \times 32$ ($H \times W \times D$). The model will first convolve the input image using a convolution kernel of size $3 \times 3 \times 3$. In the feature extraction process, the number of channels increases sequentially in the order of 16, 64, 128, and 256 as the size of the feature map decreases. In contrast, in the extended channel, the number of chan-

nels gradually decreases and the feature map becomes progressively larger. Note that the first result of each layer of the expanding path has a channel number of 64, since each has four sources. (For example, the second layer of the expanding path also has a jump connection from the fourth layer, which is not drawn in the figure for space constraints.)

times that of shallow feature maps, and direct splicing will result in a tiny percentage of shallow features in the final fused features. Here we use the standard convolution module of 3D convolution operation + batch normalization + ReLU activation function to process the three sets of pooled feature maps separately. After that, the above four sets of 3D feature maps are combined at the channel level by the concatenation

operation. Finally, the combined features are fed into the convolution kernel so that each input feature map is feature fused to obtain the output feature map for that scale layer of the expansion channel. The output feature map will be directly involved in the loss value calculation as one of the input predictions in the joint training scheme of the

Fig. 4 Comparison of different size blood vessels in carotid CTA images
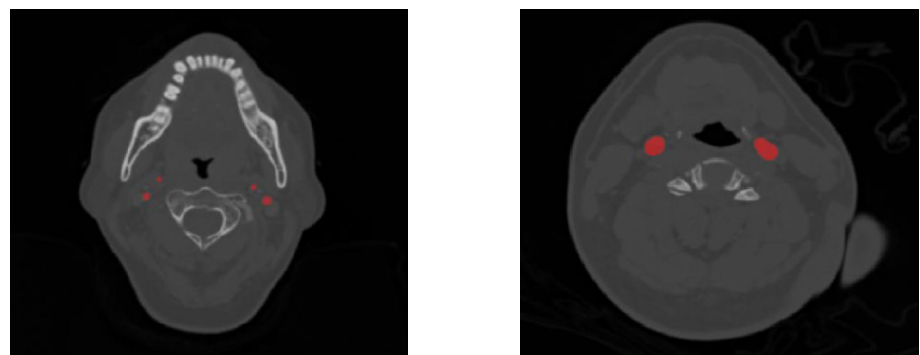
(a) Fine blood vessel area

(b) Coarse vascular area

multi-scale loss function, in addition to continuing to up-sampling as input to the subsequent network layers.

Furthermore, in the carotid 3D CTA image segmentation task, the carotid artery region accounts for a small proportion of the whole set of 3D images, and the blood vessels vary in thickness. As shown in Fig. 4, the difference in blood vessel size between the internal carotid artery, external carotid artery, and common carotid artery region is evident. Therefore, the original U-Net [44] four times down-sampling scheme in this task has the problem of excessive down-sampling. By removing the last layer of down-sampling, we not only increase the number of shallow convolutional layers and channels and reduce the model parameters, effectively accelerating the training, but also have no impact on performance.

2.3.2 Loss Function

Since the carotid region accounts for a small proportion of the whole 3D CTA images, there is a severe problem of positive and negative sample imbalance in the segmentation task. The statistics of the samples in the dataset reveal that the ratio of positive and negative samples is about 0.003, which is a severe imbalance. Thus, based on the Dice distance, we innovatively design an improved loss function. It effectively solves problems caused by inconsistent positive and negative samples and makes the network perform better on challenging classification samples. The single-scale loss function can be represented by Eq. (1).

$$L_{\text{Hybrid}} = \left(-\log\left(\frac{2 \cdot \sum_{i=0}^{N} (y_i \cdot \bar{y}_i)}{L_{H\_D}} \right) \right)^{\gamma}$$

$$L_{H\_D} = \sum_{i=0}^{N} (y_i \cdot \bar{y}_i) + \alpha \sum_{i=0}^{N} ((1 - y_i) \cdot \bar{y}_i) \qquad (1)$$

$$+ \beta \sum_{i=0}^{N} (y_i \cdot (1 - \bar{y}_i))$$

where α controls the weight of false positive samples in the loss value calculation, and β controls the weight of false negative samples in the loss value calculation. We can weigh the model prediction bias by adjusting these two parameters. In this paper, the α and β are taken as 0.4 and 0.6 to make the model balance and performance reach a better state. And γ is taken as 0.3 to improve the function's nonlinear performance.

In addition, in order to solve the problems of unstable training and difficult convergence of the original Dice loss function and help the model jump out of local extrema, we add a binary Cross Entropy loss function, which can effectively smooth the gradient. The final loss function used in this paper at the single-scale can be represented by Eq. (2).

$$L_{\text{Hybrid}} = \left(-\log\left(\frac{2 \cdot \sum_{i=0}^{N} (y_i \cdot \bar{y}_i)}{L_{H\_D}} \right) \right)^{\gamma} + \lambda L_{BCE}$$

$$L_{H\_D} = \sum_{i=0}^{N} (y_i \cdot \bar{y}_i) + \alpha \sum_{i=0}^{N} ((1 - y_i) \cdot \bar{y}_i) \qquad (2)$$

$$+ \beta \sum_{i=0}^{N} (y_i \cdot (1 - \bar{y}_i))$$

where the parameter γ is used to balance the participation weight of the binary cross-entropy loss function. Setting the parameter γ to 1 at the beginning of model training can help the training to be more stably and accelerate the model convergence. In the later stages of training, we gradually decrease the value of the parameter γ so that we can improve the model performance according to the Dice distance.

To better supervise the fusion of features at each scale of the CA-UNet, we propose a multi-scale multi-loss function joint training scheme, which can make good utilization of the image features extracted at each scale layer in training. In the calculation, an additional convolutional layer is connected after the expanding paths of all four scale layers. This convolutional layer uses a convolutional kernel of size $3 \times 3 \times 3$ with one channel to convolve the output of this scale layer. Then, We use trilinear interpolation to uniformly recover different feature maps to the input image size. We compute the corresponding loss values using the Sigmoid function and the above single-scale loss function. Finally, we assign different weights to the losses calculated by each scale layer to accumulate the final loss function. The multi-scale loss function can be represented by Eq. (3).

$$\text{Loss} = (L_{\text{Hybrid1}} + L_{\text{Hybrid2}} + L_{\text{Hybrid3}}) \cdot \omega + L_{\text{Hybrid4}} \quad (3)$$

where L_{Hybrid1}, L_{Hybrid2}, L_{Hybrid3} and L_{Hybrid4} denote the loss values computed from four different scale feature maps in the network model from deep to shallow. L_{Hybrid4} is the loss value calculated from the output image in the network, the specific calculation of these four loss functions is specifically calculated for the single-scale loss function design described above. ω iis the weight for the first three loss values. During the training process, this weight value is gradually decreased by a factor after each certain round of iterations. The proportion of L_{Hybrid4} in the overall loss value is gradually enlarged so that the model output is closer to the target effect in the later stages of training.

2.4 Ischemic Stroke Risk Prediction Model

2.4.1 3D Image Feature Extraction Network

In this paper, to extract the 3D features of carotid CTA images, we extend each convolutional layer of the

conventional DenseNet from 2D to 3D convolution. Our 3D DenseNet uses dense connectivity to ensure that each network layer is connected to all the networks in the previous layer. First, we perform the initial feature extraction work on the input image using a convolutional kernel of size $7 \times 7 \times 7$. Secondly, the model is followed by several 3D densely connected modules and transition modules. The transition modules consist of 3D convolution and 3D pooling, and the 3D densely connected modules are the core of image feature extraction in this paper. The structure of the 3D densely connected modules is presented as Fig. 5. Each feature map is directly connected to subsequent layers in the 3D densely connected module by skip connections before being fed to the next convolutional layer. The features from each layer are fused using the addition operation at the end of the connection. For the ith network layer within the module, the output x can be shown as Eq. (4).

$$x_i = H_i\Big(\big[x_0, x_1, \ldots, x_{i-1}\big]\Big) \tag{4}$$

where $[x_0, x_1, \ldots, x_{i-1}]$ represents the dense concatenation of the first few layers of the input feature maps, $H(\cdot)$ is the nonlinear transformation function, which is a set of composite functions performing 3DConv+ReLu+BN operations. 3DConv is a 3D convolution operation using a convolution kernel of size $3 \times 3 \times 3$. ReLU is the commonly used activation function. BN is the batch normalization operation. There is a direct connection between any two layers in the 3D DenseNet, which could obtain a larger range of receptive fields and preserve the features of the lower layers. Morever, using the bottleneck layers, the 3D DenseNet has fewer parameters than the 3D convolutional neural network. Fewer parameters make it easier to train the network when the 3D model is limited by GPU memory.

Besides, to extract the image features in the carotid region more effectively, we propose to introduce deformable convolution into the 3D image feature extraction network. The 3D deformable convolution adds an offset to the convolution kernel for learning, which enables the shape and size of the convolution window to be adjusted autonomously according to the characteristics of the carotid artery region. By this method, we make the convolution window focus on the carotid and take full advantage of the spatial structure of the data. The training of offsets and weights of the 3D deformable convolution can be represented by Eq. (5).

$$y(p_0) = \sum_{p_n \in R} w(p_n) x(p_0 + p_n + \Delta p_n) \tag{5}$$

where p_0 represents the position of the pixel point in the output feature map, $y(p_0)$ represents the feature value of the convolution layer at that position, and p_n represents the nth value in the convolution receptive field R. When using a 3D convolution kernel of size $3 \times 3 \times 3$, the receptive field $R = \{(-1, -1, -1), (-1, -1, 0), \ldots, (1, 1, 0), (1, 1, 1)\}$. $w(p_n)$ represents the weight of the corresponding position of the convolution kernel. Δp_n represents the offset corresponding to the nth value in the deformable convolutional receptive field R, and the exact position is obtained by the bilinear difference. The improved 3D deformable convolution structure is shown in Fig. 6.

2.4.2 Fusion Prediction Network

The overall of fusion prediction model is presented in Fig. 7. The fusion prediction model mainly consists of the image feature extraction sub-model and the machine learning sub-model. Since data such as electronic medical records and medical history in the private dataset are insufficient to train a machine learning model with good results, we introduce a large amount of data from the public dataset to assist in training the machine learning sub-model. Along this line, we perform migration learning on a 3D image feature extraction network trained on a CTA image dataset and a machine learning model trained on a public dataset containing electronic medical records and medical history data. Through parameter migration, we migrate the weight parameters into the image feature extraction sub-model and the machine learning sub-model in the fusion prediction network model. Finally, we derive the joint risk assessment results by weight

Fig. 5 3D dense connection module structure

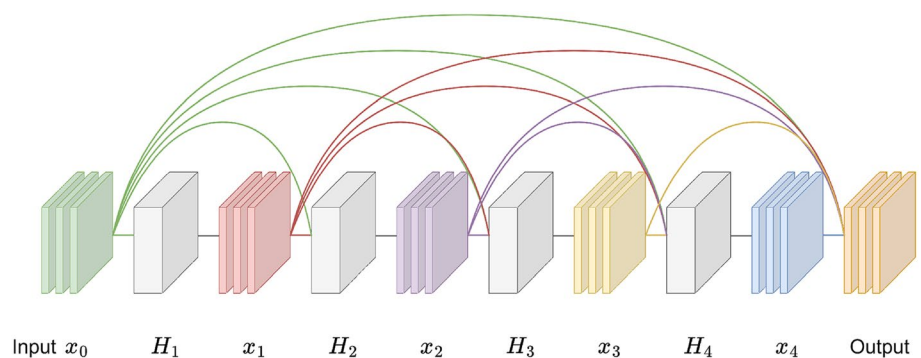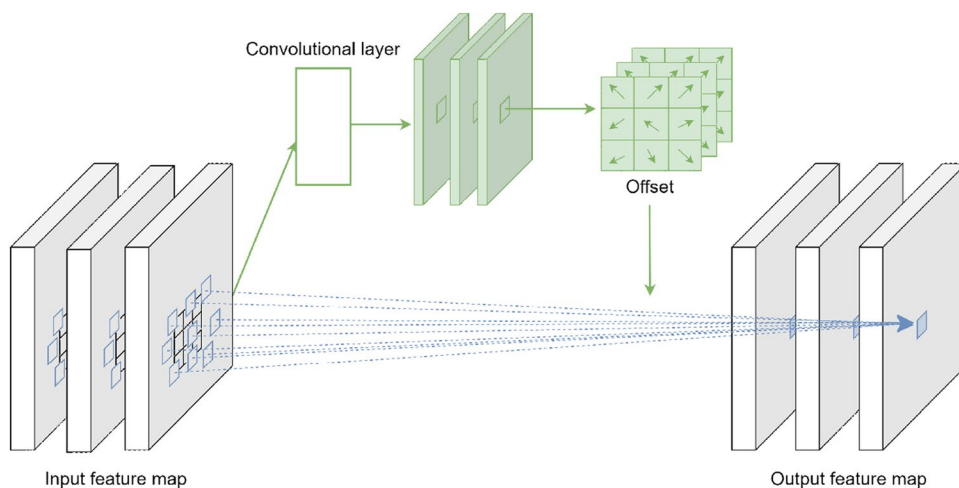Input x_0 H_1 x_1 H_2 x_2 H_3 x_3 H_4 x_4 Output

Fig. 6 3D deformable convolutional structure

fusion. The optimal weight values of the two sub-models in the fusion prediction network are derived by grid search.

In the fusion prediction network model, the outputs of the image feature extraction sub-model and the machine learning sub-model are fused according to the scale factors λ_1 and λ_2. In this way, we can obtain the final output prediction probability value of the fusion prediction network model, which can be calculated by Eq. (6).

$$p_{out} = p(\hat{y} = 1 \mid x) = \lambda_1 \cdot p(\hat{y}_1 = 1 \mid x) + \lambda_2 \cdot p(\hat{y}_2 = 1 \mid x) \tag{6}$$

where x indicates the input value, $\hat{y}_1$ indicates the evaluation result of the image feature extraction sub-model, $\hat{y}_2$ represents the evaluation result of the machine learning sub-model, λ_1 and λ_2 represent the weight values of both in the fusion prediction network model. We migrate the trained weight parameters of the sub-models to form a fusion prediction network model. The model can obtain joint prediction results that combine various types of information while the data are trained separately, making full use of the information in the data.

3 Results and Discussion

3.1 Evaluation Indicators

Dice coefficient, Jaccard Index, False Negative Rate (FNR), and False Positive Rate (FPR) are utilized evaluation metrics in image segmentation. The below is the formula representation for the above evaluation metrics, where R_{gt} represents the ground truth of the segmentation result and R_{seg} represents the segmentation result predicted by the network.

Dice coefficient indicates the ratio of the area of the intersection of two set regions to the total area and is usually used to represent the degree of overlap of two sets. A higher value

of the Dice coefficient indicates a better segmentation result. The calculation method is represented by Eq. (7).

$$\text{Dice} = \frac{2 * (R_{seg} \cap R_{gt})}{R_{seg} + R_{gt}} \tag{7}$$

Jaccard Index is expressed as the ratio of the area where two regions intersect to the area where they merge, which is compared to the similarity and difference of the two regions. The larger the value of the Jaccard coefficient, the more similar the two sets are, and the calculation is denoted by Eq. (8).

$$\text{Jaccard} = \frac{(R_{seg} \cap R_{gt})}{R_{seg} \cup R_{gt}} \tag{8}$$

False Negative Rate denotes the proportion of foreground pixels misclassified as background pixels to all pixels in the whole. A higher value of False Negative Rate indicates that more parts of the target object are not segmented completely and is calculated by Eq. (9).

$$FNR = \frac{FN}{R_{seg} \cup R_{gt}} \tag{9}$$

False Positive Rate denotes the proportion of background pixels misclassified as foreground pixels to all pixels in the whole. If the FPR value is higher, the more redundant parts of the result that do not belong to the target object. The calculation is represented by Eq. (10).

$$FPR = \frac{FP}{R_{seg} \cup R_{gt}} \tag{10}$$

Accuracy (Acc), sensitivity (Sen), and specificity (Spe) have commonly used evaluation metrics in medical image prediction classification tasks. The above evaluation metrics are calculated from a two-dimensional confusion matrix.
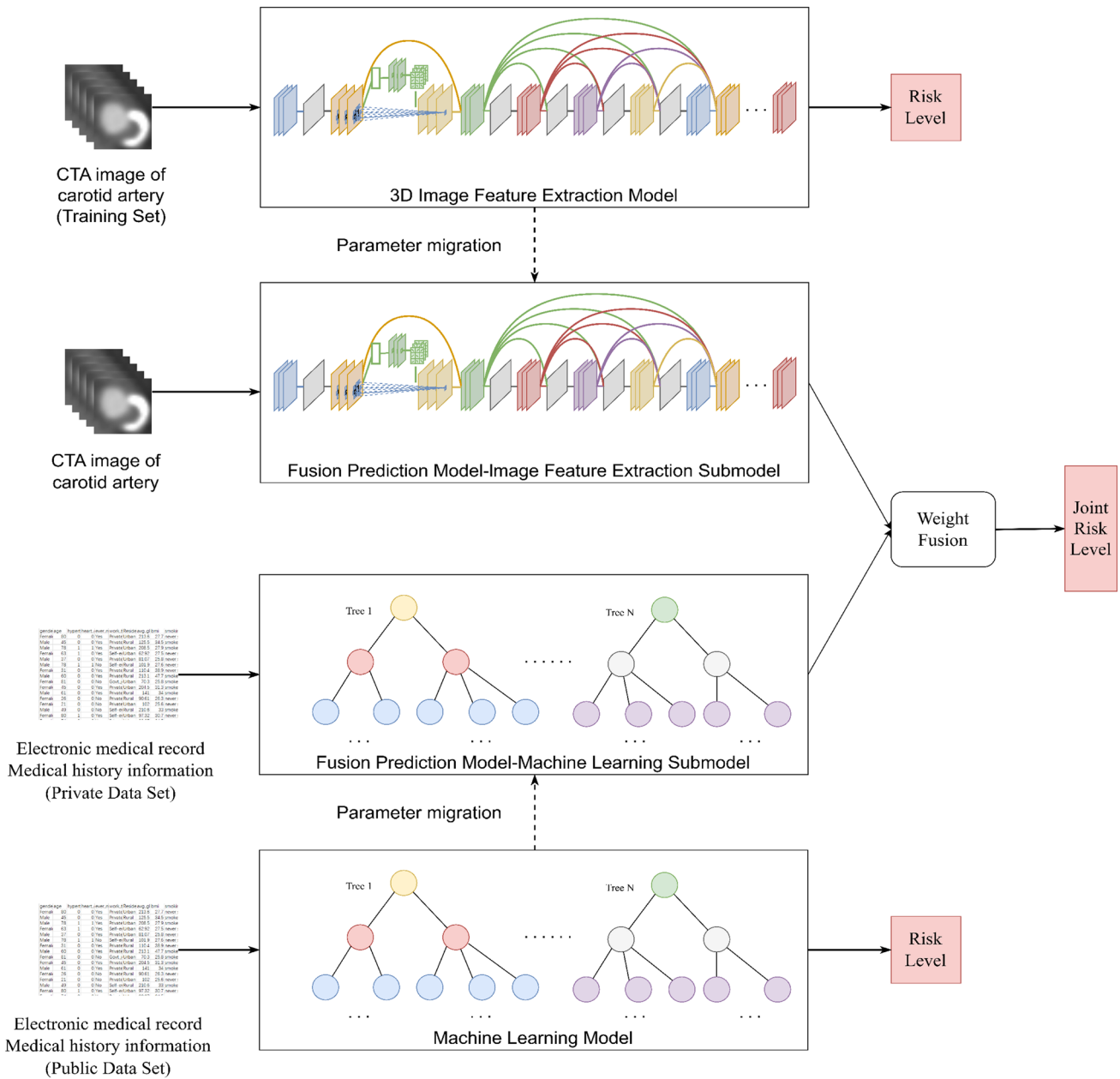
Fig. 7 Fusion prediction network. The input of the image feature extraction sub-model is the segmented 3D CTA image of the carotid region, which can reduce the influence of redundant image regions. The input of the machine learning sub-model is the electronic medical record. Ultimately, we obtain joint risk assessment results

Accuracy represents the proportion of all samples with correct predictions, which could be obtained by the equation 11.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{11}$$

Sensitivity represents the probability that an algorithm can correctly determine a positive sample and is calculated by Eq. (12).

$$Sensitivity = \frac{TP}{TP + FN} \tag{12}$$

Specificity represents the probability that an algorithm can correctly determine a negative sample and is calculated as shown in Eq. (13).

$$Specificity = \frac{TN}{FP + TN} \tag{13}$$

3.2 Results and Discussion of the Two Tasks

3.2.1 Carotid Segmentation Task

First, we investigated the appropriate number of down-sampling for CA-UNet by testing whether the four down-sampling layers used in the conventional encoder-decoder partition network have excessive down-sampling problems. After constructing a network containing four down-sampling layers, the down-sampling modules and network layers were removed layer by layer, starting from the bottom layer. The number of channels at the bottom layer was kept constant. Table 3 shows the results of comparative experiments with different numbers of down-sampled layers.

The experimental results show that when training CA-UNet, removing the lowest layer of down-sampling had little effect on the segmentation performance. The network performance showed a decline when the remaining down-sampling layer was removed. Removing the useless down-sampling layer could increase the number of shallow convolutional layers and channels and decrease the model parameters, effectively accelerating the training.

Different ways of calculating loss values have significant effects on the direction of learning in the training. Table 4 provides the test results of CA-UNet model using the fusion loss function compared with the Dice loss function.

The results show that compared with the Dice loss function, using the fusion loss function has significantly improved the results, and has better performance in all indicators. The most apparent decrease in the False Negative Rate indicates that by adjusting the variable parameters in the fusion loss function, the purpose of balancing the learning direction of the model is achieved.

To test the performance of CA-UNet and the fusion loss function, the same training set was used to train the models of 3D U-Net, V-Net, Zhou [29] and Zhu [35], and we tested on the same settings. The test results are presented at Table 5.

Compared with 3D U-Net, V-Net, [29], and [35], our CA-UNet model combined with fusion loss function gets the

Table 3 Comparison results of different down-sampling times

	Dice	Jaccard	FNR	FPR
CA-UNet + once down-sampling	83.53%	72.34%	**7.44%**	20.22%
CA-UNet + twice down-sampling	87.10%	77.64%	9.52%	12.84%
CA-UNet + three times down-sampling	**90.49%**	**82.90%**	9.96%	**7.14%**
CA-UNet+four times down-sampling	90.37%	82.62%	8.04%	9.34%

Bold indicates that the best results were achieved in the current indicator

Table 4 Comparison results of different loss functions (Bold indicates that the best results were achieved in the current indicator)

	Dice	Jaccard	FNR	FPR
CA-UNet + Dice loss function	88.77%	80.32%	12.02%	7.67%
CA-UNet + fusion loss function	**90.49%**	**82.90%**	**9.96%**	**7.14%**

Bold indicates that the best results were achieved in the current indicator

Table 5 Experimental results of different carotid artery segmentation models

	Dice	Jaccard	FNR	FPR
3D U-Net [45]	81.12%	69.22%	23.26%	7.52%
V-Net [46]	86.22%	76.45%	17.33%	6.22%
Zhou et al. [29]	86.61%	76.88%	14.13%	8.99%
Zhu et al. [35]	88.14%	79.76%	15.77%	**4.48%**
CA-UNet (Ours)	**90.49%**	**82.90%**	**9.96%**	7.14%

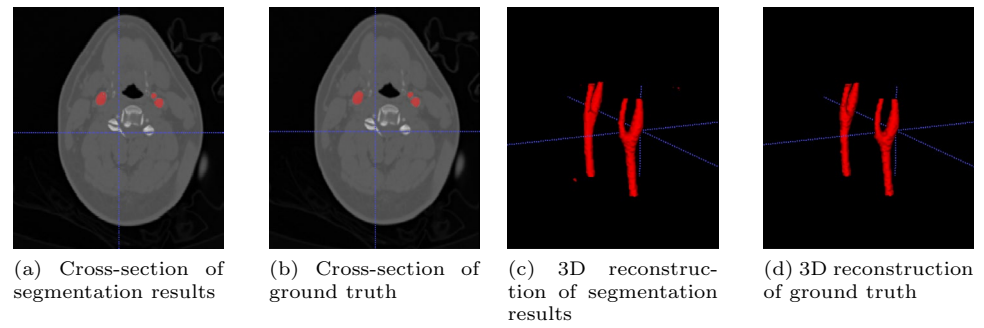Bold indicates that the best results were achieved in the current indicator

Table 6 Segmentation results for each volume in the test set

	Dice	Jaccard	FNR	FPR
V-1	91.96%	85.11%	10.44%	4.45%
V-2	93.47%	87.74%	8.81%	3.44%
V-3	94.95%	90.39%	4.28%	5.33%
V-4	93.07%	87.05%	4.94%	8.01%
V-5	93.69%	88.13%	7.39%	4.48%
V-6	92.44%	85.95%	5.78%	8.27%
V-7	95.24%	90.90%	**3.41%**	5.69%
V-8	90.69%	82.97%	13.79%	3.24%
V-9	92.43%	85.92%	3.54%	10.54%
V-10	**96.44%**	**93.13%**	5.53%	**1.35%**
V-11	94.87%	90.24%	7.16%	2.60%
V-12	91.42%	84.20%	11.44%	4.36%
V-13	92.11%	85.37%	6.09%	8.54%
V-14	93.52%	87.83%	6.13%	6.04%
V-15	91.76%	84.77%	4.14%	11.09%
V-16	93.50%	87.79%	6.92%	5.29%
V-17	94.10%	88.86%	7.00%	4.14%

Bold indicates that the best results were achieved in the current indicator

best evaluation performance with Dice coefficient, Jaccard Index, and False Negative Rate of 90.49, 82.90 and 9.96%, which is better than all other methods, and False Positive Rate of 7.14%, which is better than the other two methods.

Fig. 8 Comparison of CA-UNet segmentation results with ground truth

(a) Cross-section of segmentation results

(b) Cross-section of ground truth

(c) 3D reconstruction of segmentation results

(d) 3D reconstruction of ground truth

In addition, due to the optimization of the model structure and the number of down-sampling, the CA-UNet model has only 1/2 of the parameters of Zhu et al.

Table 6 shows the performance of each group of CTA images in the test set, applying the CA-UNet model. The best results were obtained for the sample numbered V-10, with Dice coefficient, Jaccard Index, False Negative Rate, and False Positive Rate of 96.44, 93.13, 5.53 and 1.35%. Figure 8 shows the segmentation results of the CA-UNet model proposed in this paper compared with the ground truth.

3.2.2 Ischemic Stroke Risk Prediction Task

The ischemic stroke risk prediction model proposed in this paper consists of three parts. The first part is a 3D image

feature extraction network for carotid CTA images. The second part is a machine learning model for predicting stroke risk using electronic medical records and medical history. The third part is the fusion network. We design comparison experiments for the machine learning model, the 3D image feature extraction model, and the fusion network model on their respective datasets.

First, we conduct comparative experiments with machine learning models. In this paper, SMOTE, random under sampling (RUS), and instance hardness threshold (IHT) were selected as methods to solve the sample size imbalance problem. The results of the comparative experiments are presented in Table 7.

In the comparison experiment of resampling methods, the accuracy and specificity of the XGBoost model using SMOTE are 90.63 and 95.16%, which are the best

Table 7 Comparison between machine learning model and data resampling method

Model	Resampling method	Accuracy	Sensitivity	Specificity
Decision Tree	SMOTE	85.23%	24.53%	88.70%
Decision Tree	RUS	67.52%	73.58%	67.17%
Decision Tree	IHT	75.25%	75.47%	75.24%
Random Forest	SMOTE	87.47%	18.87%	91.39%
Random Forest	RUS	69.25%	**84.91**%	68.35%
Random Forest	IHT	78.72%	71.70%	79.12%
SVM (Linear Kernel)	SMOTE	76.88%	62.26%	77.72%
SVM (Linear Kernel)	RUS	74.85%	83.02%	74.38%
SVM (Linear Kernel)	IHT	79.74%	77.36%	79.87%
SVM (RBF Kernel)	SMOTE	75.97%	79.25%	75.78%
SVM (RBF Kernel)	RUS	72.00%	**84.91**%	71.26%
SVM (RBF Kernel)	IHT	81.98%	60.38%	83.21%
XGBoost	SMOTE	**90.63**%	11.32%	**95.16**%
XGBoost	RUS	70.06%	**84.91**%	69.21%
XGBoost	IHT	77.70%	79.25%	77.61%
Logistic Regression	SMOTE	76.37%	58.49%	77.40%
Logistic Regression	RUS	73.42%	83.02%	72.87%
Logistic Regression	IHT	79.53%	79.25%	79.55%
MLP	SMOTE	82.38%	43.40%	84.61%
MLP	RUS	70.47%	83.02%	69.75%
MLP	IHT	85.13%	56.60%	86.76%

Bold indicates that the best results were achieved in the current indicator

Table 8 Comparison of fusion prediction models, where the 3D image feature extraction network is abbreviated as 3D CTA network

Model	Resampling method	Machine learning model weight	Accuracy	Sensitivity	Specificity
Only 3D CTA network	–	0	83.33%	91.67%	76.19%
3D CTA network + Random Forest	SMOTE	0.6	87.18%	75.00%	**97.62**%
3D CTA network + Random Forest	RUS	0.6	87.18%	91.67%	83.33%
3D CTA network + Random Forest	IHT	0.6	**89.74**%	91.67%	88.10%
3D CTA network + SVM (Linear Kernel)	RUS	0.8	88.46%	91.67%	85.71%
3D CTA network + SVM (Linear Kernel)	IHT	0.5	88.46%	**94.44**%	83.33%
3D CTA network + SVM (RBF Kernel)	SMOTE	0.6	88.46%	88.89%	88.10%
3D CTA network + SVM (RBF Kernel)	RUS	0.6	85.90%	88.89%	83.33%
3D CTA network + XGBoost	SMOTE	0.5	87.18%	83.33%	90.48%
3D CTA network + XGBoost	IHT	0.5	**89.74**%	**94.44**%	85.71%
3D CTA network + Logistic Regression	RUS	0.6	88.46%	91.67%	85.71%
3D CTA network + Logistic Regression	IHT	0.5	85.90%	88.89%	83.33%
3D CTA network + MLP	RUS	0.6	85.90%	88.89%	83.33%

Bold indicates that the best results were achieved in the current indicator

performance, but the sensitivity is only 11.32%. Because the SMOTE method oversamples a small number of classes for training, which leads to the misconception that the model can classify well during the training phase. However, it is still poor at classifying samples in a small number of classes in fact. Compared to the other two data resampling methods, Instance Hardness Threshold performs better on each machine learning model due to the removal of those data that are often misclassified in training. In the comparative experiments of machine learning models, using Instance Hardness Threshold as the resampling method, the Logistic Regression model performed the best overall, with accuracy, specificity, and sensitivity metrics of 79.53, 79.25 and 79.55%.

Next, we use the above machine learning model and 3D image feature extraction network to construct a fusion prediction network and conduct comparison experiments. The results are presented as Table 8.

The results indicate that the fusion prediction model with two sub-models weight ratios of 0.5 and 0.5 performs best. The accuracy, specificity, and sensitivity of the test set are 89.74, 94.44, and 85.71%. Its sensitivity was the highest, indicating that the model could correctly determine positive samples in the ischemic stroke risk prediction task.

In order to validate the performance of the proposed 3D image feature extraction network and the fusion prediction network, we train and test 3DResNet, 3D-CNN and 3D-DenseNet using the same settings. The test results are presented as Table 9.

The results indicate that 3D-ResNet achieved the best results in the specificity for the ischemic stroke risk prediction task. The proposed model achieved the best overall results with accuracy and sensitivity of 83.33 and 91.67%. When using machine learning models such as XGBoost

Table 9 Comparative results of predictive models for ischemic stroke risk

Model	Accuracy	Sensitivity	Specificity
XGBoost [47]	83.33%	83.33%	83.33%
3D-CNN [48]	75.64%	86.11%	66.67%
3D-ResNet [49]	79.49%	77.78%	80.95%
3D-DenseNet [50]	82.05%	86.11%	78.57%
3D Image Feature Extraction Network (Ours)	83.33%	91.67%	76.19%
Fusion Prediction Network (Ours)	**89.74**%	**94.44**%	**85.71**%

Bold indicates that the best results were achieved in the current indicator

alone for prediction, the results are relatively poor because the rich information in CTA images is not utilized. In contrast, the fusion prediction model, with accuracy, specificity, and sensitivity of 89.74, 94.44, and 85.71% on the test set, achieved the best results in all three metrics. Therefore, the fusion prediction network model proposed in this paper has significant advantages for the ischemic stroke risk prediction task.

4 Conclusion

We use the CA-UNet model to segment the carotid region and the fusion model to predict the risk of ischemic stroke for patients. According to the characteristics of the carotid segmentation task, we proposed to reduce the down-sampling layer and use skip connectionss which reduce the cost

of model training. And we apply a multi-scale loss function for joint training which could solve the problem that features of image details would be lost in the process of down-sampling. These novel designs resulted in a significant improvement of the assessment metrics compared to the work of others. In addition, based on CA-Unet, we propose to use a fusion prediction network to predict the risk of ischemic stroke in patients., with Acc, Sen and Spe of 89.74, 94.44 and 85.71%. Although we do not currently collect as much data as other vision tasks, our models can provide reliable diagnoses and outcomes, benefiting patients and healthcare professionals. In future research, we hope to expand more valuable data, enhance results, and investigate new ways to use more medical information, such as blood test information.

Declarations

Conflict of interest All authors declare that there are no potential conflicts of interest in this article.

Ethical approval Our study was approved by the Beihang University Biological and Medical Ethics Committee on September 28, 2022. And our study "CA-Unet segmentation makes a good ischemic stroke risk prediction" was approved for submission. (Protocol number: BM20220195)

References

1. Feigin VL, Brainin M, Norrving B, Martins S, Sacco RL, Hacke W, Fisher M, Pandian J, Lindsay P (2022) World stroke organization (WSO): global stroke fact sheet 2022. Int J Stroke 17(1):18–29. https://doi.org/10.1177/17474930211065917

2. Owolabi MO, Thrift AG, Mahal A, Ishida M, Martins S, Johnson WD, Pandian J, Abd-Allah F, Yaria J, Phan HT et al (2021) Primary stroke prevention worldwide: translating evidence into action. Lancet Public Health. https://doi.org/10.1016/S2468-2667(21)00230-9

3. Stroke Collaborators GLR (2018) Global, regional, and country-specific lifetime risks of stroke, 1990 and 2016. New Engl J Med 379(25):2429–2437. https://doi.org/10.1056/NEJMoa1804492

4. Doyle KP, Simon RP, Stenzel-Poore MP (2008) Mechanisms of ischemic brain damage. Neuropharmacology 55(3):310–318. https://doi.org/10.1016/j.neuropharm.2008.01.005

5. Members WC, Brott TG, Halperin JL, Abbara S, Bacharach JM, Barr JD, Bush RL, Cates CU, Creager MA, Fowler SB et al (2011) 2011 ASA/ACCF/AHA/AANN/AANS/ACR/ASNR/CNS/SAIP/SCAI/SIR/SNIS/SVM/SVS guideline on the management of patients with extracranial carotid and vertebral artery disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines, and the American Stroke Association, American Association of Neuroscience Nurses, American Association of Neurological Surgeons, American College of Radiology, American Society of Neuroradiology, Congress of Neurological Surgeons, Society of Atherosclerosis Imaging and Prevention, Society for Cardiovascular Angiography and Interventions, Society of Interventional Radiology, Society of Neurointerventional Surgery, Society for Vascular Medicine, and Society for Vascular Surgery. Stroke 42(8):464–540. https://doi.org/10.1161/STR.0b013e3182112cc2

6. Ma Y, Hao H, Xie J, Fu H, Zhang J, Yang J, Wang Z, Liu J, Zheng Y, Zhao Y (2020) Rose: a retinal oct-angiography vessel segmentation dataset and new model. IEEE Trans Med Imaging 40(3):928–939. https://doi.org/10.1109/TMI.2020.3042802

7. Wang D, Haytham A, Pottenburgh J, Saeedi O, Tao Y (2020) Hard attention net for automatic retinal vessel segmentation. IEEE J Biomed Health Inform 24(12):3384–3396. https://doi.org/10.1109/JBHI.2020.3002985

8. Wu H, Wang W, Zhong J, Lei B, Wen Z, Qin J (2021) SCS-Net: a scale and context sensitive network for retinal vessel segmentation. Med Image Anal 70:102025. https://doi.org/10.1016/j.media.2021.102025

9. Wu Y, Xia Y, Song Y, Zhang Y, Cai W (2020) NFN+: a novel network followed network for retinal vessel segmentation. Neural Netw 126:153–162. https://doi.org/10.1016/j.neunet.2020.02.018

10. Guo C, Szemenyei M, Yi Y, Wang W, Chen B, Fan C (2021) SA-UNet: spatial attention U-Net for retinal vessel segmentation. In: 2020 25th international conference on pattern recognition (ICPR). IEEE, pp 1236–1242. https://doi.org/10.1109/ICPR48806.2021.9413346

11. Samuel PM, Veeramalai T (2021) VSSC Net: vessel specific skip chain convolutional network for blood vessel segmentation. Comput Methods Progr Biomed 198:105769. https://doi.org/10.1016/j.cmpb.2020.105769

12. Manniesing R, Schaap M, Rozie S, Hameeteman R, Vukadinovic D, Lugt A, Niessen W (2010) Robust CTA lumen segmentation of the atherosclerotic carotid artery bifurcation in a large patient population. Med Image Anal 14(6):759–769. https://doi.org/10.1016/j.media.2010.05.001

13. Vukadinovic D, Walsum T, Manniesing R, Rozie S, Lugt A, Niessen WJ (2011) Region based level set segmentation of the outer wall of the carotid bifurcation in CTA. In: Medical imaging 2011: image processing, vol 7962. SPIE, pp 1176–1183. https://doi.org/10.1117/12.878114

14. Tang H, Walsum T, Hameeteman R, Shahzad R, Vliet LJ, Niessen WJ (2013) Lumen segmentation and stenosis quantification of atherosclerotic carotid arteries in CTA utilizing a centerline intensity prior. Med Phys 40(5):051721. https://doi.org/10.1118/1.4802751

15. Beare R, Chong W, Ren M, Das G, Srikanth V, Phan T (2010) Segmentation of carotid arteries in CTA images. In: 2010 international conference on digital image computing: techniques and applications. IEEE, pp 69–74. https://doi.org/10.1109/DICTA.2010.21

16. Freiman M, Joskowicz L, Broide N, Natanzon M, Nammer E, Shilon O, Weizman L, Sosna J (2012) Carotid vasculature modeling from patient CT angiography studies for interventional procedures simulation. Int J Comput Assist Radiol Surg 7:799–812. https://doi.org/10.1007/s11548-012-0673-x

17. Turani Z, Zoroofi RA, Shirani S (2013) 3D automatic segmentation of coronary artery based on hierarchical region growing algorithm (3D HRG) in CTA data-sets. In: 2013 20th Iranian conference on biomedical engineering (ICBME). IEEE, pp 275–279. https://doi.org/10.1109/ICBME.2013.6782234

18. Santos FLC, Joutsen A, Terada M, Salenius J, Eskola H (2014) A semi-automatic segmentation method for the structural analysis of carotid atherosclerotic plaques by computed tomography angiography. J Atheroscler Thromb 21(9):930–940. https://doi.org/10.5551/jat.21279

19. Bozkurt F, Köse C, Sari A (2017) Segmentation of carotid arteries in CTA images using region-based active contours and classification. In: 2017 international artificial intelligence and data processing symposium (IDAP), pp 1–8. IEEE. https://doi.org/10.1109/IDAP.2017.8090261

20. Bozkurt F, Köse C, Sarı A (2018) An inverse approach for automatic segmentation of carotid and vertebral arteries in CTA. Expert Syst Appl 93:358–375. https://doi.org/10.1016/j.eswa.2017.10.041

21. Hemmati H, Kamli-Asl A, Talebpour A, Shirani S (2015) Semi-automatic 3D segmentation of carotid lumen in contrast-enhanced computed tomography angiography images. Physica Med 31(8):1098–1104. https://doi.org/10.1016/j.ejmp.2015.08.002

22. Tenekecı ME, Pehlıvan H, Gümüşçü A, Karadağ K (2018) Using angio image sequence for coronary vessel segmentation. In: 2018 26th signal processing and communications applications conference (SIU) (2018). IEEE, pp 1–4. https://doi.org/10.1109/SIU.2018.8404754

23. Ma J, Chen J, Ng M, Huang R, Li Y, Li C, Yang X, Martel AL (2021) Loss odyssey in medical image segmentation. Med Image Anal 71:102035. https://doi.org/10.1016/j.media.2021.102035

24. Ramesh K, Kumar GK, Swapna K, Datta D, Rajest SS (2021) A review of medical image segmentation algorithms. EAI Endorsed Trans Pervas Health Technol 7(27):6–6. https://doi.org/10.4108/eai.12-4-2021.169184

25. Chen J, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, Lu L, Yuille AL, Zhou Y (2021) TransUNet: transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306. https://doi.org/10.48550/arXiv.2102.04306

26. Antonelli M, Reinke A, Bakas S, Farahani K, Kopp-Schneider A, Landman BA, Litjens G, Menze B, Ronneberger O, Summers RM et al (2022) The medical segmentation decathlon. Nat Commun 13(1):4128. https://doi.org/10.1038/s41467-022-30695-9

27. Jha D, Riegler MA, Johansen D, Halvorsen P, Johansen HD (2020) DoubleU-Net: a deep convolutional neural network for medical image segmentation. In: 2020 IEEE 33rd international symposium on computer-based medical systems (CBMS). IEEE, pp 558–564. https://doi.org/10.1109/CBMS49503.2020.00111

28. Stringer C, Wang T, Michaelos M, Pachitariu M (2021) Cellpose: a generalist algorithm for cellular segmentation. Nat Methods 18(1):100–106. https://doi.org/10.1038/s41592-020-01018-x

29. Zhou T, Tan T, Pan X, Tang H, Li J (2021) Fully automatic deep learning trained on limited data for carotid artery segmentation from large image volumes. Quant Imaging Med Surg 11(1):67 https://doi.org/10.21037/qims-20-286

30. Song A, Xu L, Wang L, Yang X, Xu B, Wang B, Yang B, Greenwald S (2022) Automatic coronary artery segmentation of CCTA images with an efficient feature-fusion-and-rectification 3D-UNet. IEEE J Biomed Health Inform. https://doi.org/10.1109/JBHI.2022.3169425

31. Zhao H, Fang Z, Ren J, MacLellan C, Xia Y, Li S, Sun M, Ren K (2022) SC2Net: a novel segmentation-based classification network for detection of Covid-19 in chest X-ray images. IEEE J Biomed Health Inform 26(8):4032–4043. https://doi.org/10.1109/JBHI.2022.3177854

32. Huang H, Chen Q, Lin L, Cai M, Zhang Q, Iwamoto Y, Han X, Furukawa A, Kanasaki S, Chen Y-W et al (2022) MTL-ABS 3 Net: Atlas-based semi-supervised organ segmentation network with multi-task learning for medical images. IEEE J Biomed Health Inform 26(8):3988–3998. https://doi.org/10.1109/JBHI.2022.3153406

33. Han K, Liu L, Song Y, Liu Y, Qiu C, Tang Y, Teng Q, Liu Z (2022) An effective semi-supervised approach for liver CT image segmentation. IEEE J Biomed Health Inform. https://doi.org/10.1109/JBHI.2022.3167384

34. Hatamizadeh A, Tang Y, Nath V, Yang D, Myronenko A, Landman B, Roth HR, Xu D (2022) UNetr: transformers for 3D medical image segmentation. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision, pp 574–584. https://doi.org/10.1109/WACV51458.2022.00181

35. Zhu H, Song S, Xu L, Song A, Yang B (2022) Segmentation of coronary arteries images using spatio-temporal feature fusion network with combo loss. Cardiovasc Eng Technol 13(3):407–418. https://doi.org/10.1007/s13239-021-00588-x

36. Srivastava A, Jha D, Chanda S, Pal U, Johansen HD, Johansen D, Riegler MA, Ali S, Halvorsen P (2021) MSRF-Net: a multi-scale residual fusion network for biomedical image segmentation. IEEE J Biomed Health Inform 26(5):2252–2263. https://doi.org/10.1109/JBHI.2021.3138024

37. Doan TN, Song B, Vuong TT, Kim K, Kwak JT (2022) SON-NET: a self-guided ordinal regression neural network for segmentation and classification of nuclei in large-scale multi-tissue histology images. IEEE J Biomed Health Inform 26(7):3218–3228. https://doi.org/10.1109/JBHI.2022.3149936

38. Tong C, Yin X, Li J, Zhu T, Lv R, Sun L, Rodrigues JJ (2018) A shilling attack detector based on convolutional neural network for collaborative recommender system in social aware network. Comput J 61(7):949–958. https://doi.org/10.1093/comjnl/bxy008

39. Khosla A, Cao Y, Lin CC-Y, Chiu H-K, Hu J, Lee H (2010) An integrated machine learning approach to stroke prediction. In: Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining, pp 183–192. https://doi.org/10.1145/1835804.1835830

40. Dritsas E, Trigka M (2022) Stroke risk prediction with machine learning techniques. Sensors 22(13):4670. https://doi.org/10.3390/s22134670

41. Teoh D (2018) Towards stroke prediction using electronic health records. BMC Med Inform Decis Mak 18(1):1–11. https://doi.org/10.1186/s12911-018-0702-y

42. Arslan AK, Colak C, Sarihan ME (2016) Different medical data mining approaches based prediction of ischemic stroke. Comput Methods Progr Biomed 130:87–92. https://doi.org/10.1016/j.cmpb.2016.03.022

43. Stroke prediction dataset. https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset

44. Ronneberger O, Fischer P, Brox T (2015) U-Net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention, pp 234–241. Springer. https://doi.org/10.1007/978-3-319-24574-4_28

45. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O (2016) 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: International conference on medical image computing and computer-assisted intervention, pp 424–432. Springer. https://doi.org/10.1007/978-3-319-46723-8_49

46. Milletari F, Navab N, Ahmadi S-A (2016) V-Net: fully convolutional neural networks for volumetric medical image segmentation. In: 2016 fourth international conference on 3D vision (3DV). IEEE, pp 565–571. https://doi.org/10.1109/3DV.2016.79

47. Chen T, Guestrin C (2016) XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp 785–794. https://doi.org/10.1145/2939672.293978

48. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324. https://doi.org/10.1109/5.726791

49. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778. https://doi.org/10.1109/CVPR.2016.90

50. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4700–4708. https://doi.org/10.1109/CVPR.2017.243

Authors and Affiliations

Yuqi Zhang[1,2] · Mengbo Yu[1,2] · Chao Tong[1,2] · Yanqing Zhao[3] · Jintao Han[3]

✉ Chao Tong
tongchao@buaa.edu.cn

1   School of Computer Science and Engineering, Beihang University, Beijing, China

2   State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China

3   Department of Interventional Radiology and Vascular Surgery, Peking University Third Hospital, Beijing, China