American Journal of Information Science and Computer Engineering
Vol. 7, No. 2, 2021, pp. 31-35
http://www.aiscience.org/journal/ajisce
ISSN: 2381-7488 (Print); ISSN: 2381-7496 (Online)

Classification of Cerebral Hemorrhage Based on CT Image Segmentation

Lei Peng[1, *], Yanli Xiao[2], Yufei Zhang[1]

[1]College of Medical Information Engineering, Shandong First Medical University & Shandong Academy of Medical Sciences, Taian, China
[2]College of Marxism, Shandong First Medical University & Shandong Academy of Medical Sciences, Taian, China

Abstract

Cerebral hemorrhage is a common cerebrovascular disease. It has the characteristics of high incidence rate, fast onset and high mortality. Therefore, it is of great significance to strengthen the timely and effective treatment of cerebral hemorrhage. Because CT technology is safe, fast, low cost and high efficiency, CT image has always been an important means for doctors to diagnose cerebral hemorrhage. However, it is difficult for doctors to calculate the amount of bleeding, because of the characteristics of brain CT images, such as large noise, uneven gray distribution and fuzzy boundary of hematoma area. Therefore, an automatic cerebral hemorrhage diagnosis method combining active contour segmentation model and random forest classification model is proposed to replace the traditional manual hematoma segmentation method to calculate the hemorrhage volume. Firstly, the hematoma area in CT image is accurately segmented by CV active contour model. Secondly, according to the number and thickness of CT image sequence, the volume of hemorrhage was calculated. Finally, the random forest algorithm is used for classification to assist diagnosis and treatment. The experimental results on the test set show that the proposed algorithm can achieve 95.33% accuracy. It has positive significance in the further classification and follow-up targeted treatment of cerebral hemorrhage.

Keywords

Cerebral Hemorrhage, Image Segmentation, Active Contour Model, Random Forest Algorithm

Received: May 28, 2021 / Accepted: July 22, 2021 / Published online: July 26, 2021

1. Introduction

Acute cerebrovascular disease, cancer and heart disease are the three major diseases leading to human death. As a common acute cerebrovascular disease, cerebral hemorrhage is also known as cerebral hemorrhage or hemorrhagic stroke. It usually refers to the rupture of cerebral vessels, leading to blood flow into the ventricle, brain parenchyma and other structures, further forming blood clots or aggregation areas. There are many inducing factors of cerebral hemorrhage [1, 2], such as hypertension, arteriosclerosis, heart disease, emotional stimulation, excessive force, etc. The most common sites of cerebral hemorrhage are basal ganglia hemorrhage, brainstem hemorrhage, cerebellar hemorrhage and lobar hemorrhage. Cerebral hemorrhage has the characteristics of high incidence rate, rapid progress of disease and high mortality. Without timely and effective treatment, it will lead to hemiplegia, disturbance of consciousness and other complications, and even endanger the life safety of patients [3]. Therefore, it is of great significance to strengthen the timely and accurate diagnosis and treatment of cerebral hemorrhage.

At present, the use of X-ray, computed tomography, magnetic resonance imaging, ultrasound imaging and other imaging technology combined with other clinical diagnostic data is a common method for the diagnosis of cerebrovascular

* Corresponding author
E-mail address: serpentsd@126.com (Lei Peng)

diseases [4]. Among them, CT image has become an important means for the diagnosis of cerebral hemorrhage due to its advantages of fast, noninvasive, safe, rapid, low cost and high efficiency. The area and volume of hematoma can be estimated by CT images of cerebral hemorrhage, which can provide evidence for doctors to diagnose the disease. The traditional method to measure the volume of hematoma is to manually segment the lesion area and hematoma area of each image. Then the size of hematoma was estimated according to the average length and width of hematoma and the number of layers of hematoma. This clinical measurement method is mainly based on the experience of doctors, with strong subjectivity, long time-consuming, extremely difficult, poor accuracy and repeatability.

In recent years, many segmentation algorithms based on brain CT images have been proposed. For example, threshold segmentation method [5], region growing and watershed algorithm [6], fuzzy C-means method [7] and various active contour model segmentation algorithms based on level set theory [8-11]. Although the above algorithms achieve the segmentation of brain CT image to a certain extent, they are still sensitive to the initial position of the segmentation curve and noise, and the segmentation quality is general, and the segmentation efficiency is poor. In addition, the CT image of cerebral hemorrhage has the characteristics of high noise, uneven gray level and fuzzy boundary, which brings great challenges to the segmentation and extraction of hematoma region in cerebral hemorrhage image. Based on the above problems, this paper proposes an automatic diagnosis method of cerebral hemorrhage combining active contour segmentation algorithm and random forest classification algorithm. First, the cerebral hemorrhage focus area in the CT image is segmented quickly and accurately by CV active contour segmentation model. Then the volume of hematoma was calculated. It should be pointed out that the number of three samples in the case data: large amount of bleeding (craniotomy is required), general amount of bleeding (puncture and blood drawing therapy can be used), and small amount of bleeding (drug treatment can be used) is not equal, which leads to the imbalance of the data set, and then affects the subsequent classification and diagnosis effect. Therefore, this paper uses SMOTE (Synthetic Minority Oversampling Technique) technology [12, 13] to balance the small sample data. Finally, the diagnosis model of intracerebral hemorrhage disease is constructed by random forest algorithm [14, 15]. The experimental results show that the algorithm proposed in this paper can help doctors in searching the focus area of cerebral hemorrhage and calculating the amount of bleeding, and help to shorten the diagnosis time of pathologists. The accuracy of diagnosis can reach 95.33%, which lays the foundation for further grading and follow-up targeted treatment of cerebral hemorrhage.

2. Method

According to the characteristics of uneven gray distribution, low contrast and fuzzy edge of brain CT image, firstly, the CV active contour model is used to segment and extract the bleeding area of each brain CT image. Secondly, the bleeding area is calculated according to the number of pixels occupied by the region. The volume of hematoma was calculated according to the thickness of CT slice and the bleeding area of each image. Thirdly, in order to ensure the accuracy of diagnosis, SMOTE algorithm was used to balance the above pathological data set. Finally, the classification model of cerebral hemorrhage is constructed according to the random forest classifier, and the classification results are further analyzed.

2.1. Active Contour Segmentation Model

Active contour model is a hot topic in the field of medical image segmentation. Compared with other image segmentation algorithms, active contour model has better segmentation accuracy when segmenting fuzzy edge, heterogeneous and noisy images. Among many active contour models, the more typical one is CV model.

CV model was proposed by Chan and Vese [8]. Its basic idea is to guide the evolution curve to converge to the target edge by using the mean value of inner and outer gray levels of contour curve. The energy function of CV model is as follows.

$$F(c_1,c_2,C)=\mu\cdot Length(C)+\nu\cdot Area(\Omega_{in})+\lambda_1\int_{\Omega_{\mathrm{in}}}\left|I(x,y)-c_1\right|^2dxdy+\lambda_2\int_{\Omega_{\mathrm{out}}}\left|I(x,y)-c_2\right|^2dxdy \tag{1}$$

Where Ω Is the image region. $I(x,y)$ is the gray value of a pixel whose coordinate is (x,y). C is the evolution curve. The region inside the curve C is represented as Ω_{in}, and the region outside the curve C is represented as Ω_{out}. $Length(C)$ is the length of the curve C. $Area(\Omega_{in})$ is the area of the region contained in the evolution curve C. c_1 and c_2 are the fitting centres, which are represented by the gray mean values of the inner and outer regions of the evolution curve C. $\mu\geq 0$, $\nu\geq 0$, $\lambda_1>0$ and $\lambda_2>0$ are fixed parameters.

The CV model is based on the global information of the image. This model is not sensitive to the initial position of the evolution curve, and also has a certain anti-noise ability.

2.2. Hematoma Volume Measurement

For the CT image sequence of cerebral hemorrhage, after the hematoma region segmentation of each image layer, the hematoma area S_i of the i th slice was calculated according to the number of pixels occupied by the area. Then, according to the slice thickness Δh_i of the corresponding cerebral hemorrhage slice and the number of CT image sequences m of each patient containing the hematoma area. The volume of hematoma was estimated by the following formula.

$$V = (m\text{-}1)\sum_{i=1}^{m} S_i \Delta h_i \quad (2)$$

2.3. Unbalanced Data Processing

The proportion of medical data in the three samples with large blood loss, general blood loss and small blood loss is unbalanced. In order to make accurate medical decision, we not only need to choose the appropriate classification algorithm, but also need to do unbalanced data processing on the data obtained. Sampling technology is a common method to deal with unbalanced data. It includes oversampling, undersampling, synthetic sampling and so on. In oversampling, the SMOTE method proposed by Chawla is most commonly used. The basic principle of SMOTE algorithm is to insert new values into the nearest neighbor minority samples to synthesize new minority samples. Specifically, Firstly, we find K nearest neighbor samples from each minority sample x. Secondly, a sample x' is randomly selected from the K neighbors. Then, a new sample x^{new} is synthesized from sample x and sample x' according to the following calculation method.

$$x^{\text{new}} = x + rand(0,1)\cdot(x' - x) \quad (3)$$

Where $rand(0,1)$ is a random number between intervals (0, 1).

After SMOTE algorithm processing, we can get a relatively balanced sample proportion of cerebral hemorrhage data set.

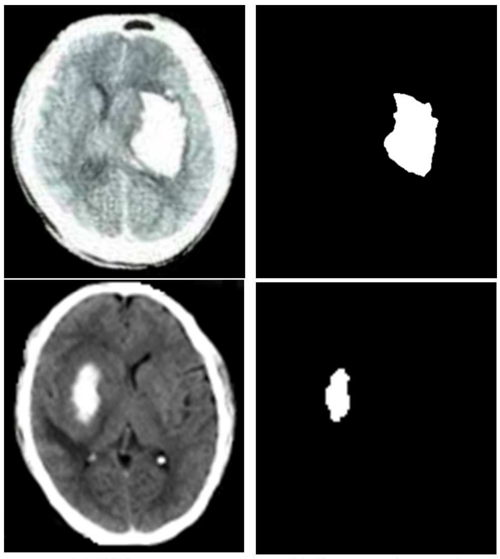
2.4. Random Forest Classification Model

Random forest (RF) is an ensemble learning algorithm based on bagging strategy. Firstly, RF algorithm randomly selects several samples from the original data set to generate a new training set. Then, through the node random splitting technique, multiple decision trees are generated for each sample to form a random forest. When there is a new test sample input, multiple decision trees will give their own prediction results. RF uses the voting result with more votes as the final decision result of the algorithm. Compared with classification models such as k-nearest neighbor (KNN), decision tree, support vector machine, naive Bayes and so on, RF can achieve better classification effect even when there are outliers or even missing values in the data set. Therefore, this paper selects RF as the classifier of cerebral hemorrhage diagnosis model.

3. Result

In this paper, the experimental data is to select 300 cases of cerebral hemorrhage patients with brain CT images. The original image was scaled to 256 × 256 pixel size.

Firstly, the CV model is used to segment the hematoma region of brain CT image sequence. As shown in Figure 1, there are some examples of hematoma region segmentation using this method. The algorithm proposed in this paper has a good anti-noise ability, and still shows good segmentation performance in the face of brain CT gray uneven and fuzzy boundary.



(a) CT images (b) Segmentation results

Figure 1. CT image segmentation results based on CV model.

For the CT image sequence of cerebral hemorrhage which has been segmented, we can count the number of pixels in each image one by one. The area of hematoma region in each image can be quantified by the number of pixels in the segmentation region. Then the area of hematoma in each image is multiplied by the thickness of each image to get the bleeding volume of this image. The volume of the whole hematoma area was estimated by further superposition.

According to the above method, the amount of cerebral hemorrhage in patients with cerebral hemorrhage can provide great help for doctors' diagnosis. For example, when the volume of hematoma is less than 20 ml, under normal circumstances, the patient is more conscious and can be treated with drugs without surgery. When the volume of hematoma is between 20 ml and 50 ml, the patient's consciousness is vague, so the injection of urinary hormone enzyme and puncture and blood drawing can be used. When the amount of bleeding is more than 50 ml, patients often appear semi coma or coma, and need immediate craniotomy. Therefore, the amount of cerebral hemorrhage can be used as an important basis for the diagnosis of cerebral hemorrhage, and it is also an important classification feature in the following random forest classification model.

There are 300 samples in the above data set. After manual labeling, there were 52 patients (marked as 0) with cerebral hemorrhage volume more than 50 ml who need immediate craniotomy. There were 68 patients (marked as 1) with cerebral hemorrhage volume between 20 ml and 50 ml who could choose puncture and blood drawing therapy. And there were 180 patients (marked as 2) with cerebral hemorrhage volume less than 20 ml who could be treated with drugs without surgery. The proportion of the three types of samples is not balanced. For this reason, SMOTE algorithm is used to sample two kinds of data with bleeding volume more than 50 ml and between 20 ml and 50 ml. After oversampling two kinds of small samples with appropriate sampling rate, there were 154 samples with cerebral hemorrhage more than 50 ml, and 166 samples with cerebral hemorrhage volume between 20 ml and 50 ml. The number of samples with cerebral hemorrhage less than 20 ml without operation remained unchanged. The total number of balanced samples is 500. The proportion of three types of samples is about 1: 1.08: 1.17, which is basically balanced.

Random forest algorithm is used to sample the balanced data. 70% was used for training (350 samples), and the remaining 30% was used for testing (150 samples). The classification results and confusion matrices on the test set are shown in Tables 1 and 2. The results show that the classification accuracy of the model designed in this paper is (47 + 42 + 54) / 150 = 95.33%.

Specifically, category 0 with bleeding volume more than 50 ml had the best classification effect. The recall rate is 98% and the accuracy rate is 96%. Only one sample was mistakenly classified into category 1 (i.e. the group with vague consciousness and needing puncture). In this case, doctors will still make secondary screening and judgment on CT images and other diagnostic data of patients with cerebral hemorrhage before puncture and blood drawing operation. It is very likely that the patient will be re classified into the craniotomy group.

In category 2, there were 58 samples with small bleeding without surgery. Among them, 54 cases were diagnosed correctly, 1 case was misdiagnosed as severe bleeding and needed craniotomy, and 3 cases were misdiagnosed as puncture and blood drawing group. Both misdiagnoses require surgery. Before surgery, doctors will still further review and screen all kinds of diagnostic data of patients, and there is still a great chance that they will be re classified into category 2 without surgery.

For category 1, i.e. the bleeding volume between 20 ml and 50 ml, the recall rate was 93% and the accuracy rate was 98%. Two samples were wrongly divided. One is to misdiagnose the patients who need puncture and blood drawing as the patients who need craniotomy, the other is to divide them into the group with small amount of bleeding and only need drug treatment without surgery. In the former case, the doctor will check the case data before operation to re judge and screen, so the probability of error correction is very high, but in the latter case, there will be great hidden danger.

In fact, the misjudged samples are just in the boundary area, that is, the bleeding volume is about 20 ml or 50 ml. How to improve the classification accuracy of boundary area samples is one of the follow-up research works.

Table 1. Classification results based on RF.

	precision	recall	f1-score	support
0	0.96	0.98	0.97	48
1	0.91	0.95	0.93	44
2	0.98	0.93	0.95	58
avg/total	0.95	0.95	0.95	150

Table 2. Confusion matrix based on RF classification.

	0	1	2
0	47	1	0
1	1	42	1
2	1	3	54

4. Conclusion

According to the characteristics of uneven gray distribution, low contrast and fuzzy edge of cerebral hemorrhage CT image, this paper proposes a diagnosis method of cerebral hemorrhage based on active contour segmentation model and random forest classification algorithm. Firstly, CV active contour model is used to segment the local hematoma area with fuzzy boundary accurately on the premise of high robustness to noise and initial contour curve. Then, the volume of hematoma area was calculated. In order to improve the accuracy of classification, SMOTE algorithm is used to balance the samples. Finally, the random forest model is used to do the classification experiment. The results show that the diagnosis method of cerebral hemorrhage designed in this paper has a high accuracy, and has a positive significance in assisting doctors to do regional segmentation of cerebral hemorrhage lesions, hematoma volume calculation, diagnosis of cerebral hemorrhage degree and follow-up treatment

Foundation Items

Shandong Province Medicine and Health Science and Technology Development Plan Project, China (No. 2019WS394).

Taian City Science and Technology Development Program, China (No. 2019GX017).

References

[1] Tiago, Gregório, Sara, et al. Original Intracerebral Hemorrhage Score for the Prediction of Short-Term Mortality in Cerebral Hemorrhage: Systematic Review and Meta-Analysis. [J]. Critical Care Medicine, 2019, 47 (6): 857-864.

[2] Jieqiong AN, Hao LI, Sun M, et al. Clinical Analysis of Acupuncture Combined with Nimodipine in the Treatment of Acute Cerebral Hemorrhage [J]. World Chinese Medicine, 2020, 15 (06): 892-895+900.

[3] Vaidyanathan S, Soni B M, Hughes P L, et al. Fatal Cerebral Hemorrhage in a Tetraplegic Patient Due to Autonomic Dysreflexia Triggered by Delay in Emptying Urinary Bladder after Unsuccessful Intermittent Catheterization by Carer: Lessons Learned [J]. International Medical Case Reports Journal, 2018, 11: 53-58.

[4] Li J, Cheng S Y, Xiong X Y, et al. Clinical Significances and Features of Prompt Brain CT Scan after Intracranial Artery Stenting: Analysis of 501 Cases [J]. Oncotarget, 2017, 8 (69): 114259-114267.

[5] Li P, Shi T, Y Zhao, et al. Design of Threshold Segmentation Method for Quantum Image [J]. International Journal of Theoretical Physics, 2020, 59 (2): 514-538.

[6] Mousavi M, Shariaty F, Orooji M, et al. The Performance of Active-Contour and Region Growing Methods Against Noises in the Segmentation of Computed-Tomography Scans [C]// International Youth Conference on Electronics, Telecommunications and Information Technologies (YETI-2020). 2020.

[7] Ahmed, M, Anter, et al. CT Liver Tumor Segmentation Hybrid Approach using Neutrosophic Sets, Fast Fuzzy C-means and Adaptive Watershed Algorithm – ScienceDirect [J]. Artificial Intelligence in Medicine, 2019, 97 (C): 105-117.

[8] Chan T F, Vese L A. Active Contours Without Edges [J]. IEEE Transactions on Image Processing, 2001, 10 (2): 266-277.

[9] Chen H, Yu X S, Wu C D, et al. Fast Image Segmentation Algorithm Based on Parametric Level Set Active Contour Model [J]. Dongbei Daxue Xuebao/Journal of Northeastern University, 2019, 40 (1): 6-10.

[10] Ma D, Liao Q, Chen Z, et al. Adaptive Local-fitting-based Active Contour Model for Medical Image Segmentation [J]. Signal Processing: Image Communication, 2019, 76: 201-213.

[11] Hsu W Y, Lu C C, Hsu Y Y. Improving Segmentation Accuracy of CT Kidney Cancer Images using Adaptive Active Contour Model [J]. Medicine, 2020, 99 (47): e23083.

[12] Chawla N V, Bowyer K W, Hall L O, et al. SMOTE: Synthetic Minority Over-sampling Technique [J]. Journal of Artificial Intelligence Research, 2002, 16 (1): 321-357.

[13] Xu Z, Shen D, Nie T, et al. An Oversampling Algorithm Combining SMOTE and K-means for Imbalanced Medical Data [J]. Information Sciences, 2021, (572): 574-589.

[14] Svetnik V, Liaw A, Tong C, et al. Random Forest: a Classification and Regression Tool for Compound Classification and QSAR Modeling [J]. Journal of Chemical Information and Computer Sciences, 2003, 43 (6): 1947-1958.

[15] Zaaboub N, Douik A. Early Diagnosis of Diabetic Retinopathy using Random Forest Algorithm [C]// 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP). 2020.