Scholar Publishing Group
Machine Learning Theory and Practice
https://doi.org/10.38007/ML.2022.030204
ISSN 2790-0983 Vol. 3, Issue 2: 32-39

The Application of Support Vector Machines in Medical Image Segmentation

Yunbo Li*

Department of Information Engineering, Heilongjiang International University, Harbin 150025, China

liyunbo@hiu.edu.cn

**corresponding author*

Keywords: Support Vector Machine, Medical Images, Image Segmentation, Gaussian Radial Basis Kernel Function

Abstract: Medical image segmentation is an interdisciplinary research area and an important application of graphics computing and image processing in biomedical engineering. Medical image segmentation has important applications in diagnostic medicine, surgical planning and anatomy teaching. The aim of this paper is to investigate support vector machine based medical image segmentation. The main algorithms of image segmentation, the problems of Mr image segmentation, the development trend of medical image segmentation and the current status of support vector machine research are described. Different kernel functions are selected to generate different support vector machine classifiers for segmentation of MR brain images using Mr images as the research object, and the experimental results show that the cross-validation of Gaussian radial basis kernel functions has a high correct rate.

1. Introduction

Medical image segmentation technology is a key technology for medical image processing and analysis. The development of image segmentation techniques has not only influenced the development of other related techniques in medical image processing, but medical image segmentation techniques have made significant progress due to the application of some emerging disciplines in medical image processing [1-2]. Along with the emergence and popularity of CT, MRI and PET image modalities, single segmentation techniques are no longer suitable for segmenting complex medical images generated by new image modalities [3].

In medical image segmentation research, Usman Ahmed proposed an ant colony algorithm based medical image segmentation method for selecting the optimal parameters for density peak clustering (DPC). The cut-off distance d_c is given for some shortcomings in the DPC algorithm, while the subjective randomness of the cluster centres is chosen by hand. The number of cluster centres is then quantified using a variable quantization method, allowing adaptive segmentation of the DPC algorithm and the interactive selection of the ideal d_c cut-off distance and cluster centres. Simulation results show that the algorithm is effective and practical [4].Jiwoong Jason Jeong proposed an automatic data enhancement algorithm for medical image segmentation. In order to improve the scale and diversity of medical images, they proposed a differential automatic data enhancement algorithm based on proximal update by finding the optimal enhancement strategy. Specifically, on the one hand, a dedicated search space is designed for the medical image segmentation task. On the other hand, they introduced a proximal micro-gradient descentable strategy to update the data augmentation strategy, which will improve the search efficiency [5]. Yoav Goldstein introduced a semi-automatic segmentation method for ivoct images based on image feature extraction and support vector machines (SVM). The image features used include optical attenuation factors and textures of the image based on a grey-scale co-occurrence matrix. Different groups of hyperparameters and image characteristics were tested. The method achieved an accuracy of 83% in the test images. Single-level accuracy was 89% for fibrous tissue, 79.3% for calcification and 86.5% for lipid tissue. The results suggest that the method can be used for semi-automatic segmentation of atherosclerotic plaque components in clinical IVOCT images [6]. Therefore, it is of practical significance to study the application of support vector machines in medical image segmentation.

This paper briefly introduces the background of medical image segmentation; secondly, it mainly introduces the image segmentation algorithm of support vector machine, describes the status of SVM algorithm and the theoretical basis of support vector machine, and analyses the specific steps of support vector machine application in MR brain images. The application of SVM in image segmentation is implemented, and the effectiveness of the method in simulating real data is verified.

2. A Study on the Application of Support Vector Machines in Medical Image Segmentation

2.1. Support Vector Machines

SVM methods successfully solve the small sample, local extremum problem. Traditional statistics apply empirical risk minimization criteria to optimize learning machine parameters. Due to limited training samples, learning machines based on the principle of minimising experimental risk often suffer from a lack of generalisability in practical applications [7-8]. Support vector machine methods based on statistical learning theory use standard optimisation parameters to minimise structural risk and use larger interval factors to control the training process of the learning machine [9-10].

2.2. Support Vector Machine Image Segmentation Principles

Support vector machines can be used in image segmentation as a recognition classifier. We analyse the correspondence of its individual variables in an image. Suppose that the feature for each pixel in the sample takes the values (x1, x2, ... , xn), corresponding to the ith pixel feature x. The final output of the support vector machine yi $\in$ {-1, l}. The output of yi can be set to be 1 when pixel i is the target pixel and -1 if pixel i is a non-target pixel. in this way the output of yi can be used to determine whether the ith element is the target pixel or not. A support vector machine trained with a dataset can perform target pixel recognition on an image [11-12]. In general, the more features included, the more complex the segmentation requirements and the higher the accuracy. Different features are selected for the requirements of image segmentation in different domains [13-14]. If several regions with different features need to be classified simultaneously, a multi-class support vector machine classifier needs to be constructed. The key aspect of the image segmentation method using support vector machines is how to construct the feature vectors and select the parameters of the support vector machine according to the actual situation [15].

2.3. Support Vector Machine in MR Brain Image Application

The specific steps of the support vector machine for segmentation of MR brain images in this paper are as follows:

(1) Read in the MR brain image to be segmented.

(2) Image pre-processing, i.e. removing non-brain tissue [16].

(3) Selection of training and test samples. By segmenting the reference image and determining the sample category, a certain number of samples are automatically selected randomly from various types of tissues of the image as training sample points, or manually, by left-clicking with the mouse to select a certain number of samples from various types of tissues of the original image as training samples. What is actually saved is the category to which each sample point belongs and the coordinates of the sample point [17].

(4) Feature extraction of the samples.

(5) Select the best kernel function type and its parameters and train the set of training samples obtained in (4) to obtain an SVM classifier, i.e. a classification model.

(6) The classifier obtained in (5) is used to classify the test set in (4), and different classes are represented by different colours according to the classification results; thus completing the segmentation of MR brain tissue images [18-19].

3. Investigation and Research on the Application of Support Vector Machines in Medical Image Segmentation

3.1. Experimental Data

The 120th typical image was selected, and all samples of the five tissue categories (cerebrospinal fluid, cerebral grey matter, cerebral white matter, bone dense matter, and background) were selected more precisely using a combination of k-mean clustering and manual classification, and were annotated one by one. In the training phase, 1000 samples are randomly selected from each tissue in each of the annotated images, making a total of 5000 samples to form the training sample set; the test sample set is the remaining pixels of the same slice and the pixels of other slices.

On the MR images, the SVM method was used to determine the classes of white matter, grey matter, cerebrospinal fluid, bone density and background to which each pixel belonged, so that the classification of brain tissue was a five-class classification problem for SVM. The SVM method was used to construct five classifiers, using cerebrospinal fluid as an example, with samples of cerebrospinal fluid labelled as +1 and samples of each of the remaining tissues labelled as -1. The support vector machine was trained.

3.2. Extraction of Image Texture Features

Smooth regions in an image contain grey pixels that are close to each other, while rough regions have widely varying grey levels, so the statistical time in the region histogram can be used as a metric to describe the image texture. If a region of an image has k levels of grey, the grey scale average is μ and the nth order moment of the histogram average is set to :

$$m_n = \frac{1}{N}\sum_{k=0}^{K-1}(k-\mu)^n h(k) \tag{1}$$

Second order moment's m2, also known as variance, is the more commonly used texture measure.

3.3. Feature Normalisation

Before using the SVM classification method, the extracted features must be normalised. The main advantage is that it prevents flooding smaller dynamic range resources, thus producing the same effect. Another advantage is that it is difficult to avoid extensive computation when calculating the internal gains of a resource vector, and larger attribute values may lead to computational overflow. Therefore, features usually need to be normalised. Normalisation is carried out using the following formula:

$$\text{Normalised features} = \frac{\text{Eigenvector - Eigenvector Minimum}}{\text{Eigenvector maximum - Eigenvector minimum}} \tag{2}$$

After normalisation with equation (2), the range of features is limited to [0, 1]. When normalising the features, it is necessary to normalise the features of the training set and test set samples in the same way.

4. Analysis and Research on the Application of Support Vector Machines in Medical Image Segmentation

4.1. Kernel Functions and Kernel Parameters

Different kernel functions and parameters generate different classifiers, so the first task is to compare the performance of different classifiers with various kernel functions and kernel parameters. The feature extraction method is the texture feature extraction introduced earlier, which includes 24 texture features and 3 grey scale features, and the number of training samples are all 1500.

Table 1, Table 2 and Figure 1 show the comparative results of the cross-validation correct rates obtained using the 5-fold cross-validation method after generating different classifiers under different kernel functions and their parameters.

From the data in the three tables, it can be seen that the Gaussian radial basis kernel function has the largest cross-validation correct rate under different models built on the basis of different kernel functions. The linear kernel function and polynomial kernel function were smaller, as shown in Figure 2. This is mainly because the three tissues in the MR brain tissue images are less different in terms of features and it is difficult for the low VC dimensional classifier to separate them well. Further analysis of the variation of the parameters γ and the penalty factor C of the Gaussian radial kernel function shows that when γ is constant and C changes, the cross-validation correct rate increases with the increase of C. As C increases, the number of training misclassification samples decreases, i.e. the empirical risk decreases, so the cross-validation correct rate increases; while when C is constant and γ changes, the cross-validation correct rate basically changes from small to large and then to small as γ decreases. This phenomenon is due to the fact that when γ is large, the classifier VC dimension is large and over-learning occurs, making the cross-validation correct rate smaller. For both linear and polynomial kernel functions, under-learning occurs due to the large number of MR brain image feature vectors and the overlap between brain tissues, resulting in poor classification performance. From the above analysis, it can be concluded that the kernel function and its parameters have an impact on the segmentation performance of the classifier. Experimentally, it is proved that the kernel function has a high correct rate of cross-validation and the best generalization performance when the Gaussian radial basis kernel function is selected.

Table 1. Cross validation accuracy under linear kernel function

Penalty factor C	Accuracy
1	70
10	72
100	75
1000	77
5427	79

Table 2. Cross validation accuracy under polynomial kernel function

Parameter d	C=10	C=100	C=1000	C=5427
5	76.78	77.13	78.24	78.21
10	75.55	78.71	78.44	75.56
15	78.67	78.46	76.54	77.43

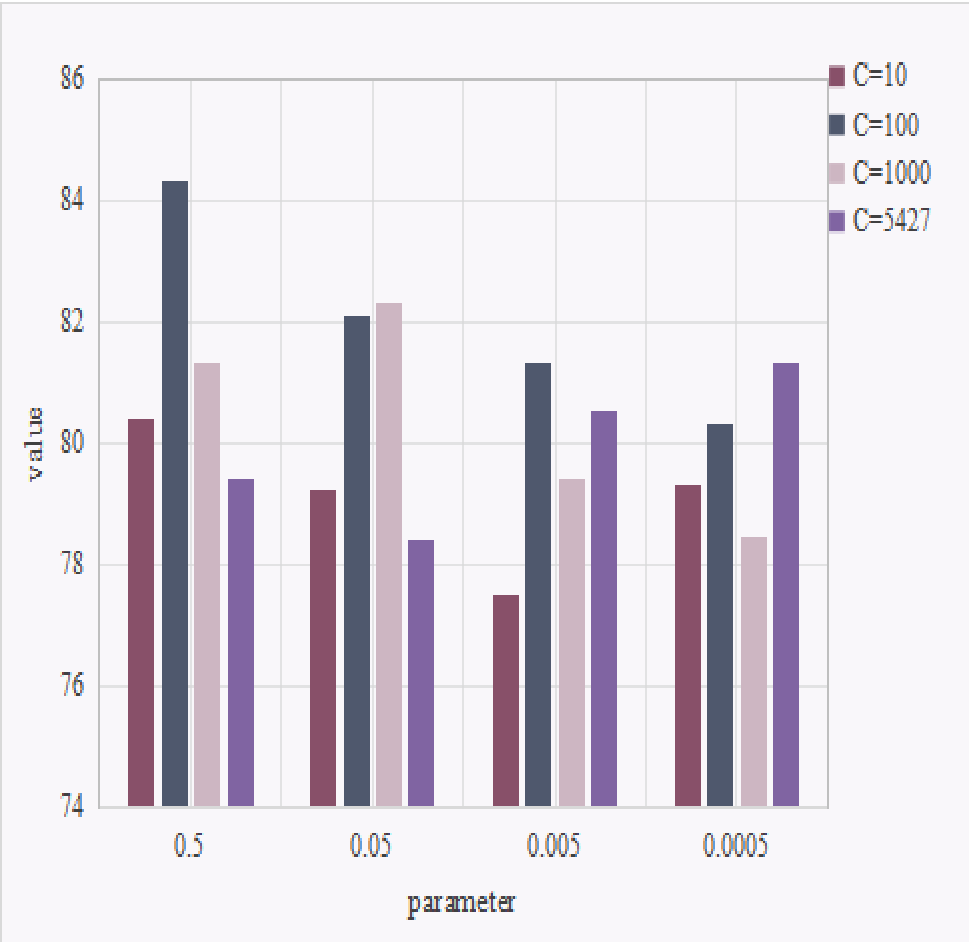

Figure 1. Cross validation accuracy under RBF kernel function

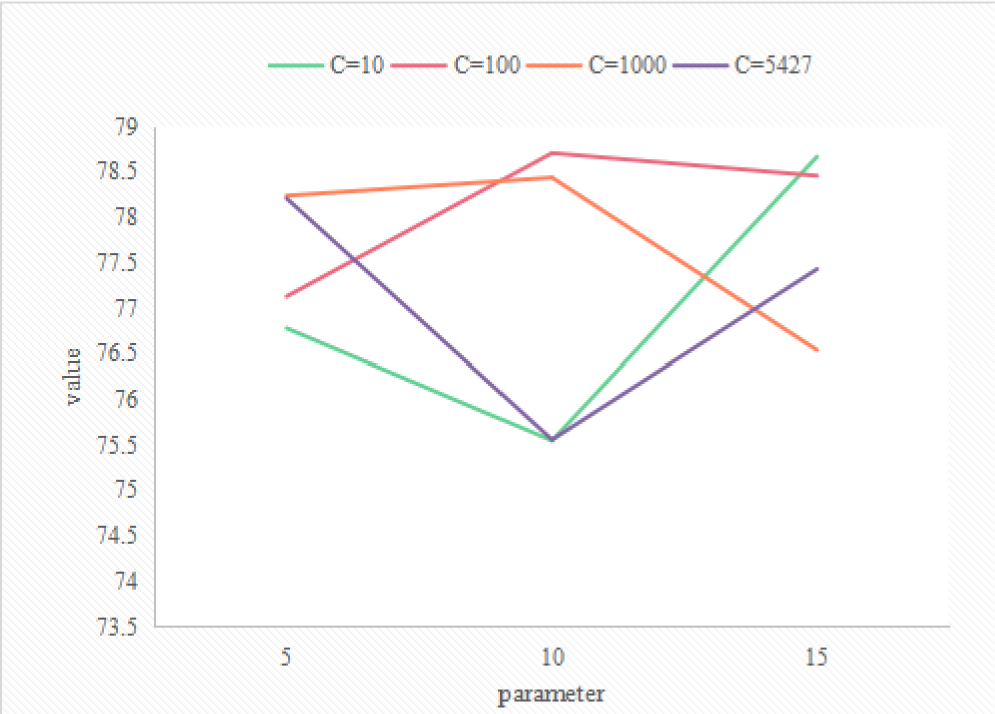

Figure 2. Polynomial kernel verification results

4.2. MR Image Segmentation with Different Noisy Images

The impact of noise on medical images is not estimable, and the accuracy of MR brain image segmentation under different noises is of important research significance. Therefore, this paper uses the Gaussian radial basis kernel function as the kernel function, in which the number of training samples are all taken as 1000, and the SVM classifier generated by taking $\gamma=0.0005$ and $C=5427$ to segment MR brain images with 0%, 3%, 7% and 9% noise. The accuracy of their segmentation was compared and the comparison results are shown in Table 3.

Table 3. SVM segmentation error rates under MR brain maps with different noises

Different noise	SVM segmentation error rate
0%	1.67%
3%	2.48%
7%	3.25%
9%	3.78%

As can be seen from Table 3, the segmentation error rate of SVM is low when MR brain images are not disturbed by noise, but as the noise of MR brain images increases, the segmentation error rate of SVM comes increasingly large, indicating that SVM is not effective in segmenting MR brain images with noise.

5. Conclusion

The purpose of medical image segmentation is to segment raw 2D or 3D images into regions with different attributes (e.g. grey scale, texture, etc.) to provide a reliable basis for clinical and pathological studies. This paper analyses the image segmentation of magnetic resonance medical imaging, which is a relatively new science today, and proposes the application of the currently popular support vector machine classification algorithm to process functional magnetic resonance medical imaging data. Due to the limited time available to carry out this paper, and the exploratory and experimental nature of the algorithms and results used in the paper, there is room for further refinement in the content of the research. Future research includes the following: the experiments in this paper are only a preliminary exploration of this, and a deeper understanding of support vector machine theory and consideration of the nature of the specific application problem may introduce improvements in the combination approach, thus improving the performance of the corresponding multi-class classification; the ambiguity of the medical images is also to be verified.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Julia Andresen, Timo Kepp, Jan Ehrhardt, Claus von der Burchard, Johann Roider, Heinz Handels: Deep learning-based simultaneous registration and unsupervised non-correspondence segmentation of medical images with pathologies. Int. J. Comput. Assist. Radiol. Surg. 17(4): 699-710 (2022) https://doi.org/10.1007/s11548-022-02577-4

[2] Risheng Wang, Tao Lei, Ruixia Cui, Bingtao Zhang, Hongying Meng, Asoke K. Nandi: Medical image segmentation using deep learning: A survey. IET Image Process. 16(5): 1243-1267 (2022) https://doi.org/10.1049/ipr2.12419

[3] Mourtada Benazzouz, Mohammed Lamine Benomar, Youcef Moualek: Modified U-Net for cytological medical image segmentation. Int. J. Imaging Syst. Technol. 32(5): 1761-1773 (2022) https://doi.org/10.1002/ima.22732

[4] Usman Ahmed, Jerry Chun-Wei Lin, Gautam Srivastava: Ensemble-based deep meta learning for medical image segmentation. J. Intell. Fuzzy Syst. 42(5): 4307-4313 (2022) https://doi.org/10.3233/JIFS-219221

[5] Jiwoong Jason Jeong, Amara Tariq, Tobiloba Adejumo, Hari Trivedi, Judy W. Gichoya, Imon Banerjee: Systematic Review of Generative Adversarial Networks (GANs) for Medical Image Classification and Segmentation. J. Digit. Imaging 35(2): 137-152 (2022) https://doi.org/10.1007/s10278-021-00556-w

[6] Yoav Goldstein, Martin Schätz, Mireille Avigal: Chest area segmentation in 3D images of sleeping patients. Medical Biol. Eng. Comput. 60(8): 2159-2172 (2022) https://doi.org/10.1007/s11517-022-02577-1

[7] Tin Tin Khaing, Pakinee Aimmanee, Stanislav S. Makhanov, Hideaki Haneishi: Vessel-based hybrid optic disk segmentation applied to mobile phone camera retinal images. Medical Biol. Eng. Comput. 60(2): 421-437 (2022) https://doi.org/10.1007/s11517-021-02484-x

[8] Michaela Kulasekara, Vu Quang Dinh, Maria Fernandez-del-Valle, Jon D. Klingensmith: Comparison of two-dimensional and three-dimensional U-Net architectures for segmentation

of adipose tissue in cardiac magnetic resonance images. Medical Biol. Eng. Comput. 60(8): 2291-2306 (2022) https://doi.org/10.1007/s11517-022-02612-1

[9] Rangu Srikanth, Kalagadda Bikshalu: Chaotic multi verse improved Harris hawks optimization (CMV-IHHO) facilitated multiple level set model with an ideal energy active contour for an effective medical image segmentation. Multim. Tools Appl. 81(15): 20963-20992 (2022) https://doi.org/10.1007/s11042-022-12344-x

[10] R. Premalatha, P. Dhanalakshmi: Enhancement and segmentation of medical images through pythagorean fuzzy sets-An innovative approach. Neural Comput. Appl. 34(14): 11553-11569 (2022) https://doi.org/10.1007/s00521-022-07043-5

[11] Boris Shirokikh, Alexandra Dalechina, Alexey Shevtsov, Egor Krivov, Valery Kostjuchenko, Amayak Durgaryan, Mikhail Galkin, Andrey Golanov, Mikhail Belyaev: Systematic Clinical Evaluation of a Deep Learning Method for Medical Image Segmentation: Radiosurgery Application. IEEE J. Biomed. Health Informatics 26(7): 3037-3046 (2022) https://doi.org/10.1109/JBHI.2022.3153394

[12] Soufiane Belharbi, Jérôme Rony, Jose Dolz, Ismail Ben Ayed, Luke McCaffrey, Eric Granger: Deep Interpretable Classification and Weakly-Supervised Segmentation of Histology Images via Max-Min Uncertainty. IEEE Trans. Medical Imaging 41(3): 702-714 (2022) https://doi.org/10.1109/TMI.2021.3123461

[13] Suraj Mishra, Yizhe Zhang, Danny Z. Chen, X. Sharon Hu: Data-Driven Deep Supervision for Medical Image Segmentation. IEEE Trans. Medical Imaging 41(6): 1560-1574 (2022) https://doi.org/10.1109/TMI.2022.3143371

[14] Rangayya, Virupakshappa, Nagabhushan Patil: An enhanced segmentation technique and improved support vector machine classifier for facial image recognition. Int. J. Intell. Comput. Cybern. 15(2): 302-317 (2022) https://doi.org/10.1108/IJICC-08-2021-0172

[15] Sana Ullah Khan, Naveed Islam, Zahoor Jan, Khalid Haseeb, Syed Inayat Ali Shah, Muhammad Hanif: A machine learning-based approach for the segmentation and classification of malignant cells in breast cytology images using gray level co-occurrence matrix (GLCM) and support vector machine (SVM). Neural Comput. Appl. 34(11): 8365-8372 (2022) https://doi.org/10.1007/s00521-021-05697-1

[16] K. Sathish, Y. V. Narayana, Mahammad Shareef Mekala, Rizwan Patan, Suresh Kallam: Efficient tumor volume measurement and segmentation approach for CT image based on twin support vector machines. Neural Comput. Appl. 34(9): 7199-7207 (2022) https://doi.org/10.1007/s00521-021-06769-y

[17] Mohamed Djerioui, Youcef Brik, Mohamed Ladjal, Bilal Attallah: Neighborhood Component Analysis and Support Vector Machines for Heart Disease Prediction. Ingénierie des Systèmes d Inf. 24(6): 591-595 (2019) https://doi.org/10.18280/isi.240605

[18] Syed Muhammad Saqlain Shah, Muhammad Sher, Faiz Ali Shah, Imran Khan, Muhammad Usman Ashraf, Muhammad Awais, Anwar Ghani: Fisher score and Matthews correlation coefficient-based feature subset selection for heart disease diagnosis using support vector machines. Knowl. Inf. Syst. 58(1): 139-167 (2019) https://doi.org/10.1007/s10115-018-1185-y

[19] Carl Leake, Hunter Johnston, Lidia Smith, Daniele Mortari: Analytically Embedding Differential Equation Constraints into Least Squares Support Vector Machines Using the Theory of Functional Connections. Mach. Learn. Knowl. Extr. 1(4): 1058-1083 (2019) https://doi.org/10.3390/make1040060