Zhao Di, Tang Yi

THE ADAPTIVE BOOSTING ALGORITHM IN BIOMEDICAL IMAGE SEGMENTATION

Adaptive Boosting is a powerful machine learning algorithm that has been widely used in biomedical image segmentation due to its ability to handle high-dimensional feature spaces and improve classification accuracy. In this paper, we will introduce the principles and applications of AdaBoost and discuss its advantages and disadvantages.

I. Introduction

The Adaptive boosting (AdaBoost) algorithm has been successfully applied to various biomedical image segmentations, including magnetic resonance imaging (MRI). In MRI, AdaBoost can be used to segment different anatomical structures and pathologies, such as brain tumours, white matter lesions and prostate cancer. AdaBoost-based segmentation in MRI is comparable to or better than other state-of-the-art segmentation methods.

II. The Advantages of AdaBoost algorithm

The AdaBoost algorithm can perform feature selection during training and reducing the effect of noisy or redundant features; AdaBoost is flexible and adaptable to different imaging modalities and applications and can be combined with other techniques to further improve segmentation accuracy and reduce false positives and misses.

III. The description of Adaptive boosting algorithm

The basic idea of AdaBoost is to iteratively train a series of weak classifiers on the data the end result is a strong classifier that combines the outputs of the weak classifiers[1]. we suppose the weak classifier is $G_i(x)$ and his weight in the strong classifier is α_i ,then the strong classifier f(x) can be obtained as Equation (1):

$$f(x) = \sum_{i=1}^{n} \alpha_i G_i(x) \tag{1}$$

In fact, in a strong classifier consisting of i weak classifiers, if the weak classifier is good at classifying, then it will take up more weight, and if the weak classifier is average, then it should take up less weight, so we need to decide the weight of this weak classifier according to its classification error rate of the sample, and the classification error rate of the sample is calculated in Equation (2):

$$\alpha_i = \frac{1}{2} log \frac{1-e_i}{e_i} \tag{2}$$

where e_i represents the error rate of the i_{th} weak classifier. The AdaBoost algorithm is implemented by changing the data distribution of the samples. We can use D_{k+1} to represent the set of weights of the samples in the $k+1_{th}$ training round, where $w_{k+1}, 1$ represents the weight of the first sample in the $k+1_{st}$ training round and w_{k+1}, N represents the weight of the Nth sample in the $k+1_{st}$ training round, and thus expressed in Equation (3):

$$D = (W_{k+1}, 1, W_{k+1}, 2, ..., W_{k+1}, N) \tag{3}$$

The sample weight in the $k+1_{th}$ round is determined according to the weight of the sample in the k_{th} round and the accuracy of the k_{th} classifier, which is Equation (4):

$$W_{k+1,1} = \frac{W_{i,j}}{Z_k} exp(-\alpha_k y_i G_k(x_i)), i = 1, 2, ..., N \tag{4}$$

The AdaBoost algorithm reflects the relative importance of each instance, and there is a dependency between these k weak classifiers during training, so that when the k_{th} weak classifier is introduced, it is actually optimising the previous k-1 classfiers.

IV. Conclusion

This paper describes how the Adaptive boosting algorithm works and summarises its advantages.In short, AdaBoost is a powerful machine learning algorithm that can be used for biomedical image segmentation.

1. Andersen, M. R., Winther, O., et al. Bayesian inference for structured spike and slab priors. In Advances in Neural Information Processing Systems,2014, 1745–1753.

Zhao Di, student of department of Information Technologies in Automated Systems Department, BSUIR, 3189124246@qq.com.

Tang Yi, student of department of Information Technologies in Automated Systems Department, BSUIR, tangyijcb@163.com.

Supervisor, Gourinovitch Alevtina, Associate Professor, PhD in Physics and Mathematics, the Belarusian State University of Informatics and Radioelectronics gurinovich@bsuir.by.