

# Paper understanding

2022/1/11 Lu Ziqi

By simply reading this paper, we explore a new way to deal with the recognition of tumor cells. In the past few decades, cancer identification and manual screening through X-ray and CT have some instability. However, recently, machine learning and deep learning have been used to reduce errors, accurately classify and accurately diagnose. By distinguishing and calculating the proportion of cancer cells in tumors and analyzing their DNA, we can screen their cancer. The article mentions the training of analysis and recognition through slices. First, it is found that the slices have instability and contingency. Therefore, using two slices can achieve an effect of reducing food and contingency, Secondly, the difference between the upper slice and the stratum slice also confirms that the experimental results have the feeling of doubt and unreliability. Therefore, the effect of improving the experimental accuracy is achieved by combining the two different slices of the upper and lower layers. The features are extracted by analyzing the slices and put into the bag model for hierarchical display. By training normal sections and tumor sections, we can learn their characteristics and distinguish the different probabilities of false positive and true positive, so as to achieve the effect of final screening and classification. It is mentioned in the article that the standby label for predicting tumor purity is completed by cutting the patch to form a bag, and a new MIL model is also developed. The feature module is extracted and by neural network, and the parametric learning process is improved. The pool filter is a relatively novel approach. Due to the standard bagged filter, the feature extraction module is given to a certain patch to extract the feature vector for each patch in the package. According to its advantages, the allocation pool filter obtains a powerful generation, which is represented by estimating the edge of the extracted feature. The training process also simulates the real clinical workflow.

Personally, it is very important to find the appropriate ground truth in the process of similar feature extraction. As the data of training set, it is necessary to achieve good labeling effect to distinguish the characteristics of normal and tumor parts. Through the combined training of multi-layer slices, it can improve the recognition accuracy. It is the core of any recognition algorithm, segmentation algorithm and classification algorithm. In addition, the grading of the bag indicates that the classification effect is also achieved in the training process, which is the same as the classification number in the code task. Generally speaking, the accurate labeling of data sets, the feature extractor and classifier of training network constitute the two core keys of this recognition task. In addition, the paper also mentioned that medical experts participate in the design to adjust the cardiogenic learning rate and other parameters, which can better improve the lack of knowledge of computer professionals in a specific field, the data processing and network writing structure adjustment of specific tasks, and is suitable for simple recognition of gray image data, such as recognition of numbers, cat and dog recognition, which are different from these obvious features. How to enlarge and reduce features, how to sample and how to set prefabricated parameters are the key to the in-depth learning of neural network. Finally, the paper also mentioned the limitations under certain circumstances. Therefore, we can consider and explore multi-dimensional from the aspects of data processing, data use and the adjustment of neural network parameters and even structure. Combined with my own research experience, a relatively perfect neural network applied to medical image recognition, such as UNET and its deformation UNET ++, 3D UNET, I think it should also be applicable to the research of tumor characteristics.

In addition, the limitations mentioned in his paper are also caused by factors such as data preservation. Therefore, data protection other than non computer problems and theoretical knowledge problems is also a factor limiting the accuracy of the project.