Pathological Diagnosis of Gastric Cancer Using Advanced Mifnet Algorithm

Basith Abdul Razack M[1], Ms.P.N.JeiprathaM.E(Ph.D)[2]
[1]Dept of Computer Science and Engineering
[2]Assistant Professor, Dept of Computer Science and Engineering
[1, 2]St.Joseph's College of Engineering, Chennai, Tamil Nadu

Abstract- *The development of a targeted and rapid method that can be used to diagnose early cancer of the stomach has an important clinical application value. Stomach cancer is still one of the deadliest diseases in the world. The worldwide shortage of pathologists provides a unique opportunity to use artificial intelligence systems to reduce labor and increase diagnostic accuracy. Most gastric cancer (GC) shows genetic instability, be it microsatellite instability or chromosomal instability, which is considered to be the first event in gastric carcinogenesis. New classification of gastric cancer based on histologic factors, genotypes and molecular phenotypes helps to better understand the characteristics of each subtype, and to improve early diagnosis, prevention and treatment. This project develops a method that uses in-depth study algorithms to help diagnose stomach cancer as it involves multiple tests to reach a conclusion. An advanced algorithm such as MIFNET is used to diagnose the presence of cancer more accurately. MIFNET is a combination of three different algorithms such as Multi task Net, Fusion Net and Global Net, a combination that provides accurate predictors of stomach cancer without further diagnosis. Therefore, this project helps in the effective diagnosis of stomach cancer with higher accuracy than existing models.*

Keywords- Pathological, Diagnosis, Gastric Cancer, Mifnet, Algorithm

I. INTRODUCTION

Stomach cancer (GC) is one of the most common infections from epithelial cells in the gastric mucosa. According to global data, GC is the fourth most common disease in the world and the second leading cause of cancer-related deaths. The occurrence and development of the GC is complex and is influenced by many factors such as nature and genetics, and the impact of these factors on the emergence of the GC has not yet been fully elucidated. The five-year survival rate of improved GC is less than 30% even after complete surgical treatment, chemotherapy, and radiotherapy, while the five-year survival rate after initial GC treatment may be more than 90%, even reaching therapeutic effect . Therefore, early diagnosis of GC is very important.

The implementation and development of the GC is a complex multi-stage, multistep, and multi-process process. There is a series of intermediate stages (including cancerous condition). Currently, the well-known pattern of human GC is proposed by Correa: "normal gastric mucosa - non-atrophic gastritis - atrophic gastritis - intestinal metaplasia - dysplasia - stomach cancer . " Atrophic gastritis (AG) and intestinal metaplasia (IM) are considered pre-cancerous lesions most closely associated with GC.AG and IM have a high risk of developing into GC if not treated in time. Their early detection and timely treatment have important implications for the prevention and treatment of GC.

Currently, there are a variety of early GC diagnostic methods, including endoscopic diagnoses [normal endoscopy, endoscopic ultrasonography, endoscopy augmentation, chromoendoscopy (CE), etc., diagnostic histopathological, imaging diagnosis (imaging diagnosis) X-ray examination, computed tomography examination, nuclear magnetic resonance,etc.), and tumor marker diagnosis (pepsinogen ,gastrin,GC symptoms, etc.). However, these diagnostic methods still have the following shortcomings: the endoscopic diagnostic method is dependent and easy to miss, histopathological diagnoses require invasive testing and time-consuming diagnostic procedures that require specialized knowledge and training, and the imaging diagnostic method cannot diagnose early. ulcers. Tumor markers are widely used to evaluate the therapeutic effect of GC, but there is still no effective tumor marker for GC. Therefore, a targeted, fast, accurate method of GC diagnosis should be developed.

Fluorescence hyperspectral imaging (FHSI) technology can be used to obtain local image information and spectral information of samples, making it possible to perform spectral analysis in all pixel regions of space, therefore, it can detect inaccessible activities directly. with traditional optical imaging and spectral method. This method can detect certain physiological and pathological changes of biological tissue through its spectral features. N. Bedard et al. found that when stimulated by blue light, normal tissues produce a pale blue autofluorescence while dysplastic and cancerous regions with reduced autofluorescence appear dark brown.Alternatively, this approach could establish a more comprehensive model for the early diagnosis of some diseases by using its “local + spectral” features.It have proven that spectral-spatial segregation method can achieve better differentiation effects than conventional spectral segregation methods and spatial segregation methods .

Because spectral separation uses only the spectral features of a single pixel, the close relative pixels are ignored. Spatial segregation is limited to image location information for a specific spectral band and ignores the structural information provided by the spectrum. In order to improve the ability to interpret and classify hyperspectral data, it is necessary to integrate spatial and spectral information into a low-dimensional environment using a spectral-spatial separation method.

In addition, hyperspectral images contain rich data with high magnitude. Due to the limitations of extracting features using standard machine learning algorithms (decision tree, random forest, and supporting vector machine), it is not possible to read presentations that represent features from this high-resolution data. A solution to the problem of gradient disappearance in deep network training was proposed by Hitonet al.2006: "Unsupervised weight training training + well-organized supervision training" which provides an effective in-depth learning solution. AlexNet, an in-depth learning network developed by the Hiton team in 2012, made significant strides in the field of image recognition, and made in-depth study one of the current areas of research.Hierarchical feature presentations can be automatically read to existing data using an in-depth reading method, which can overcome the limitations of the extracted features.

At present, in-depth studies have been widely reported in the field of medical image analysis.Inhyperspectral imaging combined with convolutional neural networks (CNNs) to diagnose normal cancer tissue and head and neck, and achieve 91.36% accuracy. Levin published a paper in Lancet Respiratory Medicine in which CT images of the lungs were combined with deep convolutional neural networks. He found that this could accurately diagnose pulmonary fibrosis with 95% accuracy. Byrne et al images of a small band affiliated with CNN were used to identify small colon polyps (adenoma and hyperplastic polyps) and found that the accuracy was 94%. Wang et al. used a two-dimensional shear-wave elastography integrated with CNN dataset.

II. DESIGN CCONCEPT

This section describes the background and motivation for preparing the database and performs a related function on the website of related research papers. Next, the methods for database preparation and database testing are described in detail, which includes details of each item.database.

Database preparation

GasHisSDB contains less than 245,196 images in two categories, including 97,076 abstract images and 148,120 standard images. Details of the database used are presented as follows.

Gastric Histopathology Small Size Photo Gallery:

Data source:

Stage 1: Four pathologists at Longhua Hospital at Shang-hai University of Traditional Chinese Medicine provide 600 photographs of stomach cancer pictures

size 2048 × 2048 and provide tissue level labels with a tightly monitored reading process [19, 23, 28].

Stage 2: Based on phase 1, five biomedical researchers

from the University of the Northeast prepared 245196 images of small size gastric pathology for poorly controlled studies, and two experienced pathologists from Liaoning Cancer Hospital and the Institute conducted the measurement.

2. Rules of preparation: By the way

by creating a database [30, 31, 32], the configuration rules for this site are as follows:

Stage 1: Three sizes (160 × 160, 120 × 120, 80 × 80 pixels) of normal pathological sections are cut directly. Also, it is necessary to select a cancerous region as an interesting region when dealing with rare pathological stages.

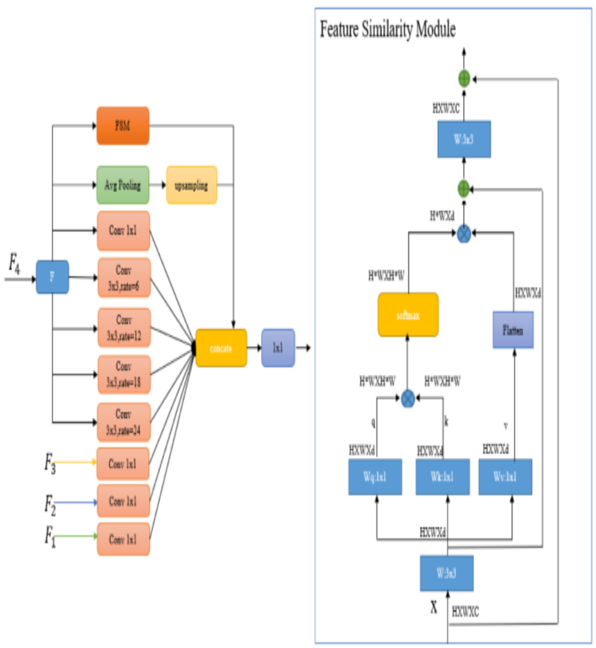
Stage 2: Region of pathological interest sec-
the basic truth is summarized simultaneously. This function can be used to filter images with very few cancer sites.

Stage 3: To minimize the correlation between small-sized images from the same original images, each is randomly rotated, and the image layout of all the details is slippery.

Dataset scale of GasHisSDB.

Sub-databasename	Croppingsize	Abnormal	Normal
Sub-databaseA	160×160pixels	13,124	20,160
Sub-databaseB	120×120pixels	24,801	40,460
Sub-databaseC	80×80pixels	59,151	87,500
Total		97,076	148,120

All standard images do not contain a cancerous region. Each cell does not have even the small atypia .Furthermore, the nuclei of cells in the image are virtually free of mi- tosis and are regularly organized into a single layer. Therefore, when viewed under an optic microscope, if there is no cancellation of any visible cells and tissues, and the characteristics of the normal image are satisfied, it can be determined that this is a normal image. In creating a standard image data set, we directly crop the entire image due to normal image attributes.



III. IMPLEMENTATION

Methods of feature extraction

Extract the various visual features of the site to prepare for classification using a machine learning separator. In this paper, we use three methods to extract the physical characteristics of a website, including the color histogram, the lo-cal binary pattern (LBP), and the gray level co-occurrence matrix (GLCM).

1. Color in the histogram

Color histogram is the most commonly used way to describe the color features of images. It simply represents the global distribution of colors in an image, that is, the number of different colors in the whole picture. Ideal for describing images that are difficult to distinguish automatically and images that do not need to be set aside the location of objects . Its advantage is that it is not affected by image rotation, shift changes, and additional familiarity with image scale changes. The downside is that it cannot define the local distribution of colors in an image, the location of each color, and specific objects.

2.Texture features

The texture is a visual element that reflects the similarity in the image. Shows the layout and layout of more buildings with slow or intermediate changes in the object area. It is not based on pixel properties but needs to perform arithmetic in an area that contains a lot of pixels. The texture is characterized by a gray distribution of pixels and its local neighbor, local information. In addition, knowledge of global texture is a level of repetition of local knowledge. This paper uses two methods to describe the composition of the GasHisSDB structure, namely Local Binary Patterns (LBP) and the Gray-level Co-occurrence Matrix (GLCM).

LBP is an operator to define functional texture that means-confirms and extracts local text information from an image that has important advantages such as gray level consistency and rotation, and the feature is easy to calculate. The LBP value of each pixel is eight values among its eight neighbors compared to its pixel. If the pixel value in eight neighbors is greater or equal to the median pixel value, the area is marked 1, and 0 if not, clockwise alignment after eight comparisons reveals the binary number of eight lengths, which gives LBP. values .

GLCM is defined as the density of the combined opportunities of pixels in two places. It not only shows the light distribution but also shows the distribution of space between pixels with the same or near light. It is a mathematical component of the second order in terms of light change of image. GLCM is the basis for defining a set of texture features .To accurately describe the texture of the symbiotic matrix, some parameters that indicate the state of the matrix are derived from the symbiotic matrix, usually the following:

1.Brightness: Demonstrates image sharpness and depth of texture.

2.Relation: Measure the similarity level of grayscale matrix features of consecutive gray grays or column directions.

3.Energy: The sum of the squares of gray symbiosis matrix elements, hence the so-called power, show similarities in gray distribution and intensity of image formation.
4.Homogeneity: Returns the ratio of the diagonal distribution of elements to GLCM.

Methods of separation

1) After the feature removal steps, two classic machine learning methods are used to separate the GasHisSDB, which includes Random Forest (RF) and Linear Support Vector Ma-chine (SVM line). In addition, three classic or novel
2) In-depth learning methods are used to classify GasHisSDB, including VGG16, ResNet50 and ViT.

1.Classic machine learning methods

The machine learning method of class translators is whether a picture is familiar or not with its visual features. RF is a learning curve for compatible collection and extended variation of Bagging. RF is based on the decision tree reader, who adds random selection of responsibility to the decision tree training process. SVMs are categorized and linear according to kernel functions. Examples of SVM line map training to space points to widen the gap between two phases. Then, the new models are mapped in the same area and predicted to belong to a category based on which side of the gap they fall into.

2. In-depth learning methods

The concept of deep learning comes from the study of Artificial Neural Network, In which a multi-layered perceptron with multiple hidden layers is a deep learning structure. In-depth reading creates a high-level attribute category or abstract feature by combining low-level features to obtain widely distributed data representations.In 2014, Visual Geometry Group and Google Deep-The brain has developed a new convolutional neural network: VGG.VGG is a Convolutional Neural Network (CNN) developed by AlexNet. A few types of VGG models are being released, and the most widely used is the VGG16 single image segment. In VGG16, three 3 × 3 convolution characters are used instead of 7x7 convolution characters, and two 3x3 convolution characters are used instead of 5x5 convolution kernels. The main purpose of this structure is to ensure uniform visual field conditions, to improve network depth, and to some extent improve the effect of neural network.

IV. CONCLUSION AND FUTURE WORK

In this paper, the gastric histopathological image sub-size imagery site, GasHisSDB, has been developed. GasHisSDB has three sub-frames, 160 × 160 pixels Minor web, 120 ×120 pixels Sub-site and 80 × 80 pixels Minor web. Each sub-site contains two standard image folders and unusual images. Each folder contains a cropped imagery that has been renamed and rendered. GasHisSDB is responsible for discriminating against the performance of class dividers. This paper is divided into the next two test sections. With classical machine learning methods, this paper presents five different features. Then evaluate the performance of the categories for the seven different categories of distinction on the three sub-bases of knowledge, and then analyze the differences in the accuracy of each category. With in-depth study methods, this paper examines three proven repetitive methods of CNN and ViT, which have recently been used in the field of image classification. This paper focuses on the analysis of four models from accuracy level, model size, training time and other indicators. In addition, this paper performs additional training time tests to determine the optimal performance of the ViT separation in GasHisSDB. GasHisSDB demonstrates that working on image classification research in this paper has the ability to explore existing image classification methods.

Database creation means more image categories-classification methods can be used in this database. We will try new ways to separate images in GasHisSDB to compare and analyze image classification methods and find the most effective ways to make medical progress.

REFERENCES

[1] Hu H., Gong L., Dong D. Identifying early Gastric cancer under magnifying narrow-band images with deep learning: a multicenter study. GastrointestEndosc. 2021;93(6):1333–1341. e3. [PubMed] [Google Scholar]
[2] Dong D., Fang M.J., Tang L. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced Gastric cancer: an international multicenter study. Ann Oncol. 2020;31(7):912–920. [PubMed] [Google Scholar]
[3] Dong D., Tang L., Li Z.Y. Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced Gastric cancer. Ann Oncol. 2019;30(3):431–438. [PMC free article] [PubMed] [Google Scholar]
[4] Zhang L., Dong D., Zhong L. Multi-focus network to decode imaging phenotype for overall survival prediction of Gastric cancer patients. IEEE J Biomed Health Inf. 2021;25(10):3933–3942. [PubMed] [Google Scholar]

[5] Huang B., Tian S., Zhan N. Accurate diagnosis and prognosis prediction of Gastric cancer using deep learning on digital pathological images: a retrospective multicentre study. EBioMedicine. 2021;71doi: 10.1016/j.ebiom.2021.103631. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

[6] Hu Y., Su F., Dong K. Deep learning system for lymph node quantification and metastatic cancer identification from whole-slide pathology images. Gastric Cancer. 2021;24:868–877. [PubMed] [Google Scholar]

[7] Song Z., Zou S., Zhou W. Clinically applicable histopathological diagnosis system for Gastric cancer detection using deep learning. Nat Commun. 2020;11(1) [PMC free article] [PubMed] [Google Scholar]

[8] Muti H.S., Heij L.R., Keller G. Development and validation of deep learning classifiers to detect epstein-barr virus and microsatellite instability status in Gastric cancer:a retrospective multicentre cohort study. Lancet Dig Health. 2021;3(10):e654–e664. [PMC free article] [PubMed] [Google Scholar]

[9] Wang X., Chen Y., Gao Y. Predicting Gastric canceroutcome from resected lymph node histopathology images using deep learning. Nat Commun. 2021;12(1) [PMC freearticle][PubMed] [Google Scholar]

[10] Wang S., Zhu Y., Yu L. RMDL: recalibrated multi-instance deep learning for whole slide Gastric image classification. Med Image Anal. 2019;58 [PubMed] [GoogleScholar