IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June Month 2021

DOI 10.17148/IJARCCE.2021.106120

Classification of Colorectal Cancer based on Multidimensional Features and CNN Model

Gayatri Ramesh Laware[1], Ranjit M. Gawande[2]

PG Scholar, Department of Computer Engineering, MCERC Nashik, India[1]

Asst. Prof., Department of Computer Engineering MCERC Nashik, India[2]

Abstract: According to incidence statistics, colon cancer is one of the most common types of cancer in the world. The correct diagnosis of this lesion will provide cancer patients with the most appropriate treatment. Diagnosis is made through visual analysis of tissue samples by a pathologist. This analysis is affected by intra-pathological and inter-pathological variation, and is also a complex and time-consuming task. In order to solve these problems, imaging techniques have been developed to be applied to histological images obtained by digitizing tissue samples. To this end, characterization and classification techniques are being explored to help pathologists and achieve faster and more objective diagnostic determinations. Therefore, this paper proposes a method that combines multi-dimensional fractal technology, curvature transformation and Haralik descriptors for research and captains. Detect colon cancer that has not been studied in the literature. The proposed method considers feature selection methods and various classification methods.

Keywords: colorectal cancer, feature associations, multiresolution features, fractal techniques, curvelet transforms, haralick descriptors.

I. INTRODUCTION

Colorectal cancer (CRC) is one of the most common cancers in the world and one of the leading causes of cancer deaths. According to the latest epidemiological data, this type of cancer constitutes a serious burden in most European countries and is still associated with a very high mortality rate. Therefore, early diagnosis and tumour differentiation are critical to the survival and well-being of a large number of patients. Traditionally, pathologists diagnose CRC by visually examining excised tissue samples that have been fixed and stained with haematoxylin and eosin under a microscope. Evaluate the existence and degree of malignant tumours by observing the tissue changes in their tissues.

Various features are studied to identify benign and malignant case models of colon cancer that show the best correlation between features and classifiers. The study showed an effective correlation between the features of the best classification results and the histological features seen in the HE image, as well as evidence of correlation not related to the classification of colon cancer.

Colon cancer is cancer that occurs in the lining of the large intestine (colon) or rectum. A study submitted by the International Agency for Research on Cancer (IARC) shows that colon cancer is the third most common cancer in men worldwide (746,000 cases) and the second most common cancer in women (614,000 cases): 7 Almost 55% of cases are carried out in more developed areas After testing, 8 the death toll was 694,000, and the highest death rate was 9 52% of the death toll in the underdeveloped regions of the world (International Agency for Research on Cancer, 2012).In Brazil, the annual estimates for the 2018-2019 biennium are 36,360 new cases and 16,696 deaths per year; by 2018, there are 140,000 new infections and 50,000 deaths in the United States.

II. RELATED WORK

Colon cancer can be diagnosed by sigmoidoscopy or colonoscopy and tissue biopsy. The result is a tissue sample stained with haematoxylin, which is visually analysed by a pathologist. This is a complex task, and experts need time to minimize these problems. Computational methods have been proposed to assist pathologists in pattern classification and recognition of HE-stained colorectal tissue (Masood and Rajpoot, 2009; Kalkan et al., 2012; Rathore et al.2013; Naiyar et al., 2015; Jorgensen et al., 2017).

After the digitization of tissue samples, those photos may be analysed via way of means of making use of computational structures. These structures permit the improvement and research of latest photograph processing techniques. Moreover,

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June Month 2021

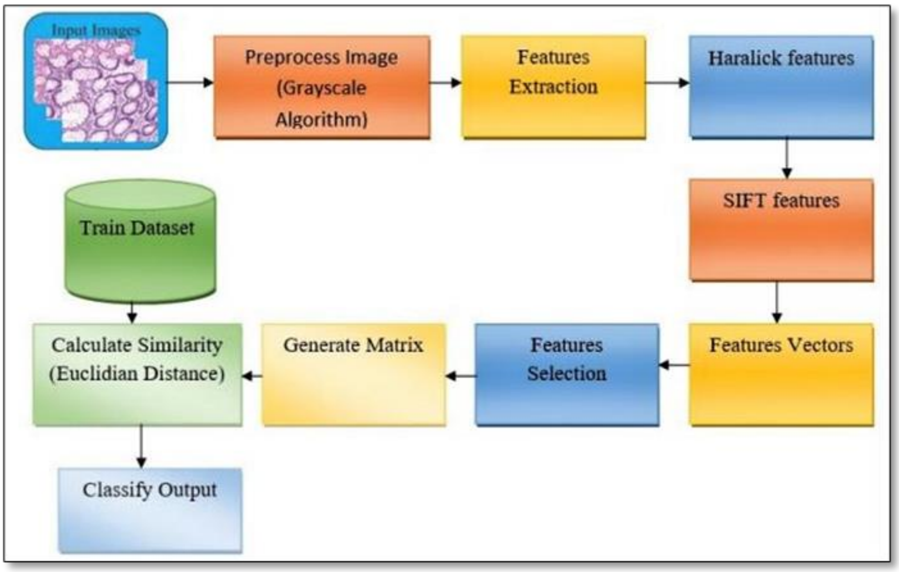
DOI 10.17148/IJARCCE.2021.106120

those structures can aid pathologists of their prognosis and diagnosis definitions which, consequently, cause the maximum good enough remedies for the patients. The steps of characteristic extraction and type compose a applicable a part of those computational strategies. By them, it's far viable to discover the intrinsic photograph statistics described via way of means of the computational strategies and relate them to the 31 investigated kinds of cancer. Besides, those steps additionally make contributions to a quicker and greater correct prognosis.

In this context, the extraction of features can be achieved through the use of various approaches, such as the techniques examined in this proposal. For example, the properties of Haralick (Haralick et al., 1973; Haralick, 1982) on studies of liver diseases (Suganya and Rajaram, 2013), the detection of epileptic activity (Boubchir et al., 2014), the diagnosis of glaucoma (Samanta et al., 2014) and oral cancer detection (Chakraborty et al., 2016). Fractal techniques are useful for quantifying self-similarity traits and the most important approaches are the fractal dimension (FD), lacunity (Lac) and, more recently, percolation (Perc). Irregularity of the image. a fractal (Ivanovici and Richard, 2011). Lac indicates how pixels are distributed and organized in an image. Perc consists of the appearance of a path connecting the upper partfrom a fractal to its bottom (Gould et al., 2005).

Pattern recognition systems that incorporate fractal techniques achieved significant success in psoriasis lesions (Ivanovici et al., 2009), glaucoma detection (Lamani et al., 2014), diagnosis of breast disease (Dobrescu et al., 2014), study of behavior of prostate cancer (Neves et al., 2014) and the quantification of non-Hodgkin lymphomas (Roberto et al., 2017) as examples. Another method of feature extraction is curvature transformation, which takes scales and angles from observations into account to analyze singularities of curves. The curvilinear transformation has been successful in the study of ulcers in endoscopic images (Eid et al., 2013), the classification of various degrees of prostate cancer (Lin et al., 2015), breast cancer diagnosis (Saraswathiet al., 2016) and. applied detection of cerebrovascular and neoplastic diseases (Nayak et al., 2017).

III.PROPOSED WORK



IV. METHODOLOGY

Modules:

- Image Pre-processing
- Feature Extraction
- Haralick Features
- SIFT Features
- Features Vectors
- Feature Selection

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June Month 2021

DOI 10.17148/IJARCCE.2021.106120

- Generate matrix
- Calculate Similarity
- Classification

Grayscale Algorithm:

- How do you convert a color image to grayscale? If each color pixel is described by a triple (R, G, B) of intensities for red, green, and blue, how do you map that to a single number giving a grayscale value?
- The lightness method averages the most prominent and least prominent colors: (max(R, G, B) + min(R, G, B)) / 2.
- The average method simply averages the values: (R + G + B) / 3.
- The luminosity method is a more sophisticated version of the average method. It also averages the values, but it forms a weighted average to account for human perception.

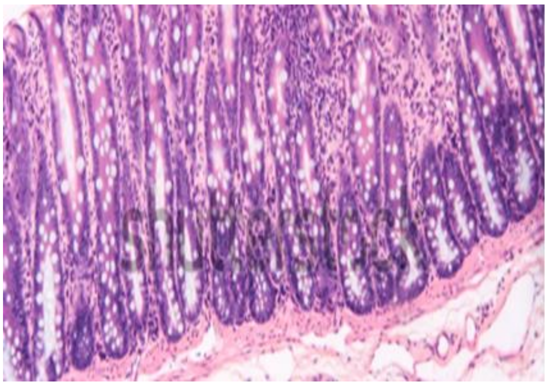


Fig-Original Image

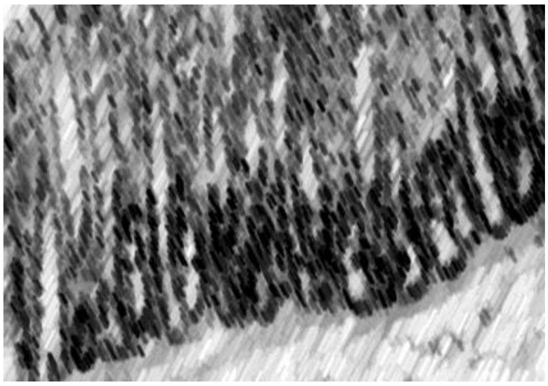


Fig-Greyscale Image

SIFT Algorithm

- SIFT is quite an involved algorithm. It has a lot going on and can become confusing, So I've split up the entire algorithm into multiple parts. Here's an outline of what happens in SIFT.
- Constructing a scale space This is the initial preparation. You create internal representations of the original image to ensure scale invariance. This is done by generating a "scale space".
- LoG Approximation the Laplacian of Gaussian is great for finding interesting points (or key points) in an image. But it's computationally expensive. So we cheat and approximate it using the representation created earlier.
- Finding keypoints with the super-fast approximation, we now try to find key points. These are maxima and minima in the Difference of Gaussian image we calculate in step 2.
- Get rid of bad key points Edges and low contrast regions are bad keypoints. Eliminating these makes the algorithm efficient and robust. A technique similar to the Harris Corner Detector issued here.
- Assigning an orientation to the keypoints An orientation is calculated for each key point. Any further calculations are done relative to this orientation. This effectively cancels out the effect of orientation, making it rotation invariant.
- Generate SIFT features Finally, with scale and rotation invariance in place, one more representation is generated.

Euclidean Distance
Euclidean distance is computed using the following formula:

$$d(p,q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}.$$

- The library contains both procedures and functions to calculate similarity between sets of data.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June Month 2021

DOI 10.17148/IJARCCE.2021.106120

- The function is best used when calculating the similarity between small numbers of sets.
- The procedures parallelize the computation and are therefore more appropriate for computing similarities on bigger datasets.

V. RESULT AND DISCUSSION

In this section, present the acquired results of the performed ML experiment. The CNN model created to classify the colorectal cancer images. We used 70% of the images (randomly chosen) to train this supervised learning model and the remaining 30% image to test it. Since we are working with a balanced dataset (i.e., each class has the same number of samples), the model will be less prone to bias towards a particular class while making decisions. We will present the model's performance on both subsets and evaluate it based on the widely used evaluation parameters, including accuracy, precision, recall, F-measure, and confusion matrix of the classification. The results of the different variants of the proposed framework are presented to show the superior performance of all the variants over simple patch-based methods.

The proposed network is well-suited for the grading task which relies on recognizing abnormalities in glandular structures. These clinically significant structures vary in size and shape that cannot be captured efficiently with standard patch classifiers due to computational and memory constraints. Nevertheless, based on the discussions provided earlier in this section, it can be concluded that the proposed methods can fulfil the task of colon cancer tissue classification with convenient accuracy and high reliability.

Fig- Input Image

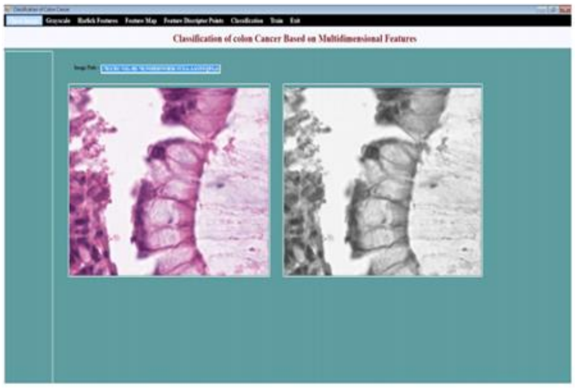


Fig- Greyscale (Pre-processing)

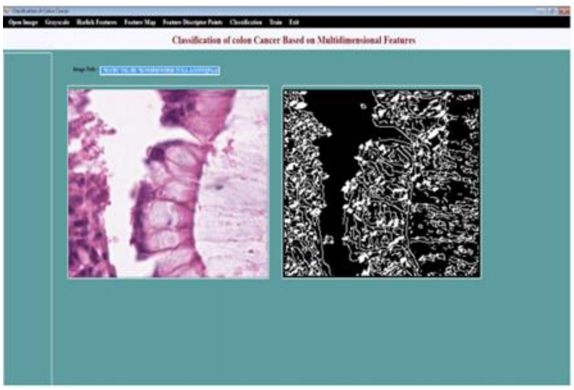


Fig- Harlick Features

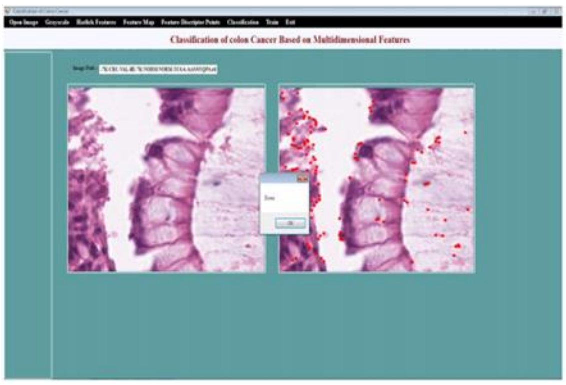


Fig-SIFT Feature

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 6, June Month 2021
DOI 10.17148/IJARCCE.2021.106120

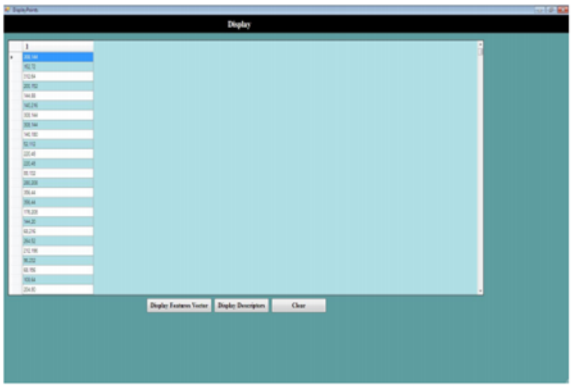


Fig- Feature Vector

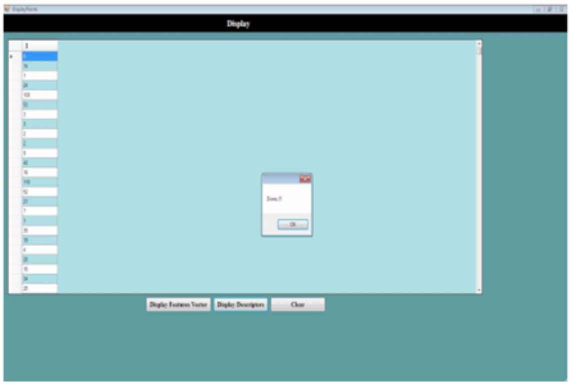


Fig- Feature Descriptor

Fig- Classification

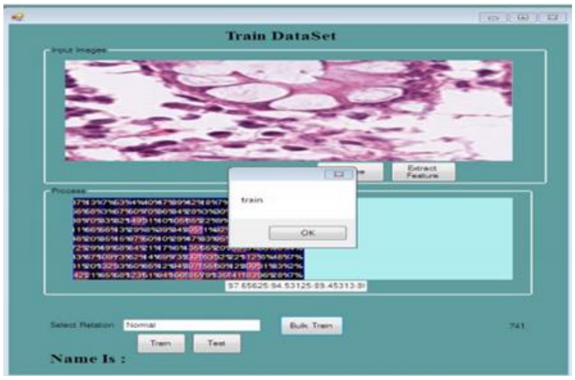


Fig- Training Model

VI. CONCLUSION

It was proposed an investigation of different feature extraction techniques and classifiers to identify colorectal cancer images as benign and malignant cases. Computer-aided diagnosis systems improves the detection of tumours for colorectal cancer can be easily adapted for detecting many other types of diseases.

This work gives promising results with some acceptable errors. However current analysis techniques can be further improved by using combination of multiple techniques in neural network.

REFERENCES

[1] Muhammad Shaban, Ruqayya Awan, Muhammad Moazam Fraz, Senior Member, IEEE, Ayesha Azam, Yee-Wah Tsang, David Snead, and Nasir M. Rajpoot, Senior Member, IEEE, Context-Aware Convolutional Neural Network for Grading of Colorectal Cancer Histology Images, IEEE Transactions on Medical Imaging, Vol. 39, No. 7, July 2020.
[2] Di Cataldo, S. and Ficarra, E. (2017). Mining textural knowledge in biological images: Applications, methods and trends. Computational and Structural Biotechnology Journal, 15:56 – 67.
[3] Hastie, T., Tibshirani, R., and Friedman, J. (2009). Overview of supervised learning. In The elements of statistical learning. Springer.
[4] Janowczyk, A. and Madabhushi, A. (2016). Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. Journal of Pathology Informatics, 7(1):29.
[5] Masood, K., Rajpoot, N., 2009. Texture based classification of hyperspectral colon biopsy samples using clbp. In: Biomedical Imaging: From Nano to Macro,2009. ISBI'09. IEEE International Symposium on. IEEE, pp. 1011–1014.
[6] Charisis, V. S., Hadjileontiadis, L. J., 2016. Potential of hybrid adaptive filtering in inflammatory lesion detection from capsule endoscopy images. World Journal of Gastroenterology 22 (39), 8641.
[7] Saraswathi, D., Dharani, D., Srinivasan, E., 2016. An efficient feature extraction technique for breast cancer diagnosis using curvelet transform and swarm intelligence. In: Wireless Communications, Signal Processing and Networking (WiSPNET), International Conference on. IEEE, pp. 441–445.
[8] Korbar, B., Olofson, A. M., Miraflor, A. P., Nicka, C. M., Suriawinata, M. A., Torresani, L., Suriawinata, A. A., and Hassanpour, S. (2017). Deep learning for classification of colorectal polyps on whole-slide images. Journal of Pathology Informatics, 8:30.
[9] Rabidas, R., Midya, A., Chakraborty, J., Sadhu, A., Arif, W., 2018. Multi resolution analysis using integrated microscopic configuration with local patterns for benign-malignant mass classification.
[10] Bianconi, F., Bello-Cerezo, R., Napoletano, P., 2017. Improved opponent color local binary patterns: an effective local image descriptor for color texture classification. Journal of Electronic Imaging 27 (1), 011002.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June Month 2021

DOI 10.17148/IJARCCE.2021.106120

[11] Breiman, L., 2001. Random forests. Machine learning 45 (1), 5–32. 811 Candes, E., Demanet, L., Donoho, D., Ying, L., 2006. Fast discrete curvelet transforms. Multiscale Modeling Simulation 5 (3), 861–899.
[12] Chakraborty, M., Mukhopadhyay, S., Dasgupta, A., Patsa, S., Anjum, N., Ray, J., 2016. A new approach of oral cancer detection using bilateral texture features in digital infrared thermal images. In: Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the. IEEE, pp. 1377–1380.
[13] Gopi, V. P., Babu, V. S., 2015. Colon cancer detection based on structural and statistical pattern recognition. In: Electronics and Commu786 nication Systems (ICECS), 2015 2nd International Conference on. IEEE, pp. 1735–1739.
[14] A., Hassan, M., Alquhayz, H., 2016. A colon cancer grade prediction model using texture and statistical features, smote and mrmr. In: Multi-Topic Conference (INMIC), 2016 19th International. IEEE, pp. 1–7.
[15] Kumar R, Srivastava R, Srivastava S. Detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features. J Med Eng 2015; 2015:457906.
[16] L. Pantanowitz, N. Farahani, and A. Parwani, “Whole slide imaging in pathology: Advantages, limitations, and emerging perspectives,” Pathol. Lab Med. Int., vol. 7, p. 23, Jun. 2015.
[17] A. Bentaieb, H. Li-Chang, D. Huntsman, and G. Hamarneh, “A structured latent model for ovarian carcinoma subtyping from histopathology slides,” Med. Image Anal., vol. 39, pp. 194–205, Jul. 2017.