International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 07 | July 2022 www.irjet.net p-ISSN: 2395-0072

Shap Analysis Based Gastric Cancer Detection

Varanasi L V S K B Kasyap[1], Athuluri Sai Kushal[1], Sunkara Vinisha[1]

[1]School of Computer Science and Engineering, VIT-AP University, Inavolu, India.

---***---

Abstract - Gastric Cancer is one of the most widely reported problem in the world causing high modality rates in the recent times. Gastroscopy is one of the efficient methods that is widely used to analyze the gastric problems. The advent technology of deep learning helps the doctor to assist in detecting the gastric cancer in the early stages. The performance of the existing methods in detecting gastric cancer from the images are not so high and consuming longer times. This paper proposes a novel deep learning frame work that can be used to detect gastric cancer from the gastric slice images. The proposed methos is based on patch-based analysis of the given input image. Specifically, the model selects and extracts the features from the images in training phase and evaluates the true risk of the patients. This is one of the novel contributions of the proposed work. The Bag-of-features techniques will be applied on the extracted features in the proposed network for the selected patches for better analysis. Experimental results prove that the proposed framework is able to detect the gastric cancer from the images effectively and efficiently. The model is robust enough to detect the minute lesions that can cause the gastric tumor in the further stages. The dataset used in this analysis is publicly available and the results achieved by this model are higher than the other models that use the same dataset. The proposed framework is also compared with existing frameworks, giving the accuracy scores higher for the proposed model.

Key Words: Gastric Cancer, Stomach neoplasms, Endoscopy, Convolutional Neural Network, Deep Learning

1.INTRODUCTION

This Stomach cancer, often known as Gastric Cancer(GC), is a type of cancer. When cells in the stomach's lining grow out of control, they develop into tumours that can infiltrate healthy tissues and spread to other regions of the body. Global data show that GC is the second most common cause of cancer-related fatalities and the fourth most prevalent malignancy worldwide [1]. Environmental and genetic factors, among others, play a complex role in the onset and development of GC, and their effects on these processes have not yet been fully understood. Even after receiving a full course of treatment that includes surgery, chemotherapy, and radiotherapy, the five-year survival percentage for advanced GC is still less than 30% [2], whereas the five-year survival rate for early GC can be over 90%, sometimes even having a curative impact [3]. The incidence and development of GC is a complicated process involving numerous mechanisms, steps, and stages. There are several transitional phases that includes the precancerous state namely "Normal gastric mucosa - chronic non-atrophic gastritis - atrophic gastritis - intestinal metaplasia - dysplasia - gastric cancer," as per the observations of Correa's current more widely accepted pattern of human GC [4]. Atrophic gastritis (AG) and intestinal metaplasia (IM) are two conditions that are thought to be precancerous lesions that are strongly linked to GC [5]. If not treated in a timely manner, AG and IM have a higher chance of turning into GC. For the prevention and treatment of GC, their early detection and prompt treatment have significant practical implications.

Examination of GC could be done with the help of various sources namely imaging tests, pathological images and endoscopy. To initiate with, stomach cancer has to be detected successfully via endoscopy. The surface structure can be precisely analysed by image-enhanced endoscopic techniques including narrow-band imaging [6] and linked colour imaging [7]. According to the studies, the precision of gastrointestinal tumour diagnosis [8] could be augmented by the deployment of endoscopic techniques. However, there is a research which has stated that even endoscopy examinations lead to still missed 10% of upper gastrointestinal malignancies [9]. There would be missed diagnoses in an endoscopic unit even if two experts participated [10]. The cause was that accurate gastroscopy image diagnosis requires years of practise to develop. Second, the gold standard for tumour diagnosis is histological image recognition. Diagnostic mistakes and a heavy workload for pathologists have been brought on by the dearth of pathologists [11]. Last but not least, imaging tests are crucial in assessing the lymph node metastases of stomach cancer. The primary focus of an imaging evaluation is on the morphological characteristics of the lesions. For instance, the perigastric adipose tissue is so dense that it resembles lymph nodes. Doctors may make errors in diagnosis due to inexperience and missing diagnoses. The accuracy of the diagnosis will eventually decline, particularly when there are several cases [12].

Artificial intelligence (AI) is exploding in the field of medicine due to the growing demand for detection, categorization, and segmentation or delineation of margins that is more accurate. After various findings in this recent scenarios, the universal ground truth is that the AI makes the machines to think as humans. One of the most crucial components of AI is machine learning. Deep learning is more accurate and flexible than standard machine learning techniques like support vector machines and Bayesian networks, and it is also easier to adapt to other fields and applications. Although AI-based technologies have shown impressive outcomes in the medical field, they have not been widely used in clinics. The main reasons are due to the unique feature of the black-box technique, as well as other factors including high computing costs. It results from the inability to clearly represent the knowledge for a particular task that is been carried out by a deep learning model, despite the existence of the underlying statistical principles. Simpler AI techniques, such as linear regression and decision trees, are self-explanatory since the model parameters allow one to visualise the classification decision border in a few dimensions. However, they do not possess the complexity needed for activities like classifying 3D and the majority of 2D medical images. Trust could be built among the patients only when the medical diagnosis done by the doctor is found to be open, clear, and explicable. It should ideally be able to fully explain to all parties concerned the reasoning behind a certain choice. Deploying deep learning models in the healthcare sectors is a challenging task as the black box models needs more interpretations. A model in AI needs to act as an aid for the medical professionals and in addition it also should permit the human expert to review the choices and exercise judgement. It has been understood from various articles that AI is used in various applications. This has drastically changed over the past ten years as a result of advancements in machine learning (ML) and the broad industrial adoption of ML, which were made possible by more potent machines, better learning algorithms, and easier access to enormous amounts of data [13]. Deep Learning techniques [14] began to rule accuracy metrics around 2012, through which better results are obtained within the stipulated time. As a result, many real-world issues are now being solved using machine learning models in a variety of industries, from fashion, education and finance [15] to medicine and healthcare. Explainability is essential for the safe and trustworthy use of artificial intelligence (AI) and a vital facilitator for its practical application. By dispelling misconceptions about artificial intelligence (AI), end users can develop trust by seeing what a model considered while making a choice. For users who do not use deep learning, such as the majority of medical professionals, it is even more crucial to display the domain-specific attributes used in the conclusion. Machine learning algorithms' output and outcomes can now be understood and trusted by human users when those are obtained through a set of procedures and techniques known as explainable artificial intelligence (XAI). An AI model, its anticipated effects, and potential biases are all described in terms of explainable AI. It contributes to defining model correctness, fairness, transparency, and outcomes in decision-making supported by AI. When putting AI models into production, a business must first establish trust and confidence. A model to be established could adopt a suitable approach to the development of AI by the deployment of AI explainability.

2. SHAP ANALYSIS

The (SHapley Additive exPlanation) SHAP analysis framework is adopted in the proposed framework because of its diversified properties. In this framework, the prediction variability is distributed among covariates that are available; the contribution of explanatory variable prediction at each point is assessed as the underlying model. The SHAP analysis results the Shapley values demonstrating the model predictions as the binary variable linear combination that describes the presence of the covariate in the proposed model or not. The SHAP algorithm estimates the prediction p(x) with the p(x'), linear function of binary variables where z belongs to $\{0,1\}^N$ and the quantities belong to real number, defined in (1)

$$p(N') = \varphi_0 + \sum_{i=1}^{N} \varphi_i N'_i \tag{1}$$

Here N is the count of explanatory variables.

(2) shows that the properties of the local accuracy, consistency and missingness obtained at each variable,

$$\varphi_i(p,x) = \sum_{z'x'} \frac{|N'||(M - |x'| - 1||'}{N!} |p(x') - (z')| \tag{2}$$

The function f in the proposed model, x is the available variable and x' will be the selected variable. The Shapley variables differ in the mean at the i^{th} variable.

3. CONCLUSIONS

In this study, a novel deep learning framework is presented for detection of Gastric Cancer. In this framework different architectures were adopted at both shallow and deep layers i.e., MSN-module and In-Net module. The proposed framework is evaluated on BOT Gastric dataset and the results obtained shows that the model is robust and effective. The model also outperforms well existed frameworks with fewer number of the layers. The Average-Classification Accuracy of the model at pixel level is 99.82%. This work can be improved further by integrating White Light Endoscopic images-based prediction and the H&E-stained images for finer and early predictions at the root levels of the tumor.

REFERENCES

1) D. F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," Ca-Cancer J. Clin. **68**(6), 394–424 (2018). [CrossRef]

2) K. D. Miller, R. L. Siegel, C. C. Lin, A. B. Mariotto, J. L. Kramer, J. H. Rowland, K. D. Stein, R. Alteri, and A. Jemal, "Cancer Treatment and Survivorship Statistics, 2016," Ca-Cancer J. Clin. **66**(4), 271–289 (2016). [CrossRef]

3) H. Katai, T. Ishikawa, K. Akazawa, Y. Isobe, I. Miyashiro, I. Oda, S. Tsujitani, H. Ono, S. Tanabe, T. Fukagawa, S. Nunobe, Y. Kakeji, and A. Nashimoto, "Five-year survival analysis of surgically resected gastric cancer cases in Japan: a retrospective analysis of more than 100,000 patients from the nationwide registry of the Japanese Gastric Cancer Association(2001-2007)," Gastric Cancer **21**(1), 144–154 (2018). [CrossRef].

4) P. Correa and M. B. Piazuelo, "Natural history of helicobacter pylori infection," Dig. Liver Dis. **40**(7), 490–496 (2008). [CrossRef]

5) Y. H. Park and N. Kim, "Review of atrophic gastritis and intestinal metaplasia as a premalignant lesion of gastric cancer," J. Cancer Prev. **20**(1), 25–40 (2015). [CrossRef]

6) Sumiyama K (2017) Past and current trends in endoscopic diagnosis for early stage gastric cancer in Japan. Gastric Cancer 20(Suppl 1):20–27. https://doi.org/10.1007/s10120-016-0659-4

7) Shinozaki S, Osawa H, Hayashi Y, Lefor AK, Yamamoto H (2019) Linked color imaging for the detection of early gastrointestinal neoplasms. 12:1756284819885246. https://doi.org/10.1177/1756284819885246

8) Dohi O, Majima A, Naito Y, Yoshida T, Ishida T, Azuma Y, Kitae H, Matsumura S, Mizuno N, Yoshida N (2020) Can image-enhanced endoscopy improve the diagnosis of Kyoto classification of gastritis in the clinical setting? 32(2):191–203. https://doi.org/10.1111/den.13540244(5):512–524. https://doi.org/10.1002/path.5028.

9) Cooper LA, Demicco EG (2018) PanCancer insights from The Cancer Genome Atlas: the pathologist's perspective. 244(5):512–524. https://doi.org/10.1002/path.5028.

10) Toyoizumi H, Kaise M, Arakawa H, Yonezawa J, Yoshida Y, Kato M, Yoshimura N, Goda K, Tajiri H (2009) Ultrathin endoscopy versus high-resolution endoscopy for diagnosing superficial gastric neoplasia. Gastrointest Endosc 70(2):240–245. https://doi.org/10.1016/j.gie.2008.10.064.

11) Xu Y, Jia Z, Wang L, F Z YA (2017) Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. BMC Bioinform 18(1):281. https://doi.org/10.1186/s12859-017-1685-x.

12) Gao Y, Zhang ZD, Li S, Guo YT, Wu QY, Liu SH, Yang SJ, Ding L, Zhao BC, Li S, Lu Y (2019) Deep neural network-assisted computed tomography diagnosis of metastatic lymph nodes from gastric cancer. Chin Med J 132(23):2804–2811. https://doi.org/10.1097/cm9.0000000000000532.

13) Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. Science 2015, 349, 255–260. [CrossRef]

14) LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436–444. [CrossRef] [PubMed]

15) Khandani, A.E.; Kim, A.J.; Lo, A.W. Consumer credit-risk models via machine-learning algorithms. J. Bank. Financ. 2010, 34, 2767–2787. [CrossRef].

16) Meyes, R.; de Puiseau, C.W.; Posada-Moreno, A.; Meisen, T. Under the Hood of Neural Networks: Characterizing Learned Representations by Functional Neuron Populations and Network Ablations. arXiv 2020, arXiv:2004.01254.

17) Arrieta, A.B.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; García, S.; Gil-López, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Inf. Fusion 2020, 58, 82–115. [CrossRef].

18) Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. Chemom. Intell. Lab. Syst. 1987, 2, 37–52.[CrossRef]

19) Maaten, L.V.D.; Hinton, G. Visualizing data using t-SNE. J. Mach. Learn. Res. 2008, 9, 2579–2605.

20) Scott M. Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 4768–4777.

21) Molnar, C. (2020). Interpretable machine learning. Lulu.com.

22) Bhatt, U., Xiang, A., Sharma, S., Weller, A., Taly, A., Jia, Y., Ghosh, J., Puri, R., Moura, J. M., and Eckersley, P. (2020). Explainable machine learning in deployment. In ACM FAT, pages 648–657.

23) Bracke, P., Datta, A., Jung, C., and Sen, S. (2019). Machine learning explainability in finance: an application to default risk analysis.

24) Mokhtari, K. E., Higdon, B. P., and Bas¸ar, A. (2019). Interpreting financial time series with shap values. In Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering, pages 166–172

25) Oikawa K, Saito A, Kiyuna T, Graf HP, Cosatto E, Kuroda M. Pathological Diagnosis of Gastric Cancers with a Novel Computerized Analysis System. J Pathol Inform. 2017 Feb 28;8:5. doi: 10.4103/2153-3539.201114. PMID: 28400994; PMCID: PMC5359998.