

Exploring ChatGPT's Potential for Consultation, Recommendations and Report Diagnosis: Gastric Cancer and Gastroscopy Reports' Case

Jiaming Zhou¹, Tengyue Li¹, Simon James Fong^{1,2*}, Nilanjan Dey³, Rubén González Crespo⁴

¹ Department of Computer and Information Science, University of Macau, SAR, Macau (China)

² Chongqing Key Laboratory of Intelligent Perception and Blockchain Technology, Department of Artificial Intelligence, Chongqing Technology and Business University, Chongqing (China)

³ Department of Computer Science and Engineering, Techno International New Town, Kolkata (India)

⁴ Computer Science and Technology Department, Universidad Internacional de La Rioja, Logroño (Spain)

Received 20 March 2023 | Accepted 22 April 2023 | Published 25 April 2023



ABSTRACT

Artificial intelligence (AI) has shown its effectiveness in helping clinical users meet evolving challenges. Recently, ChatGPT, a newly launched AI chatbot with exceptional text comprehension capabilities, has triggered a global wave of AI popularization and application in seeking answers through human-machine dialogues. Gastric cancer, as a globally prevalent disease, has a five-year survival rate of up to 90% when detected early and treated promptly. This research aims to explore ChatGPT's potential in disseminating gastric cancer knowledge, providing consultation recommendations, and interpreting endoscopy reports. Through experimentation, the GPT-4 model of ChatGPT achieved an appropriateness of 91.3% and a consistency of 95.7% in a gastric cancer knowledge test. Furthermore, GPT-4 has demonstrated considerable potential in consultation recommendations and endoscopy report analysis.

KEYWORDS

Artificial Intelligence, ChatGPT, E-assessment, Gastric Cancer, Medicine.

DOI: 10.9781/ijimai.2023.04.007

I. INTRODUCTION

ARTIFICIAL Intelligence (AI) applications in the medical field are rapidly expanding, with computer vision and natural language processing technologies playing a crucial role in achieving more efficient and accurate diagnosis and treatment solutions than ever. Computer vision technology, particularly deep learning algorithms, has made profound achievements in medical image analysis. For instance, convolutional neural networks (CNNs) are widely used to detect and identify various diseases, such as cancer [1], diabetic retinopathy [2], pneumonia [3], oral anomalies [4] and brain tumor [5]. These advanced technologies assist physicians in analysing imaging data more accurately, thus enhancing diagnostic accuracy and reducing misdiagnoses.

In the field of natural language processing (NLP), AI technologies have made breakthroughs in medical record text mining, disease risk prediction, and clinical decision support [6]. For example, NLP technologies can analyse patients' electronic medical records, helping doctors quickly access patients' medical history, past diagnoses, and treatment plans, thereby improving the efficiency of diagnosis and treatment. As AI technology continues to advance, the application prospects of computer vision and natural language processing technologies in the medical field will become increasingly extensive, contributing to various domains of healthcare applications.

* Corresponding author.

E-mail address: ccfong@umac.mo

A. Gastric Cancer

Gastric cancer is a globally prevalent disease. According to data from the World Health Organization, gastric cancer is the fifth most common type of cancer globally and the fourth leading cause of cancer-related deaths. In 2020, there were over 1 million new cases of gastric cancer worldwide, and approximately 769,000 people died from the disease [7]. As the symptoms of gastric cancer are often not apparent in the early stages, many patients are only diagnosed at late stages with poor treatment outcomes and prognosis. This makes early diagnosis and screening crucial for survival. Furthermore, the causes of gastric cancer are complex and include genetic factors, environmental factors, lifestyle factors such as poor dietary habits and smoking, and Helicobacter pylori (H. pylori) infection [8]-[12].

B. ChatGPT

ChatGPT is a large-scale natural language processing model developed by OpenAI. As a general-purpose deep learning language model, it has extensive applications in all walks of life. By learning from a vast amount of text data, ChatGPT can understand, generate, and respond to a wide range of natural language questions, providing assistance and suggestions to users [13]-[18]. It demonstrates outstanding performance in many aspects, including question-responding systems, article generation, summarization, translation, and other natural language processing tasks. During the development of GPT-4, Open-AI continuously optimized and improved the model's performance to meet the needs of more users. From GPT-3.5 to GPT-4, significant progress has been made, including higher appropriateness, consistency, and error-correction capabilities.

C. Purpose of the Study

If gastric cancer is detected and treated early, the patient's five-year survival rate could be as good as 90% [19]. Unfortunately, gastric cancer is often diagnosed at late stages. This lateness of diagnosis is due to the similarity between early-stage gastric cancer symptoms and those of common gastric diseases, as well as the fact that most people are not well informed about the risk factors for gastric cancer. Therefore, they rarely proactively undergo proper examinations when experiencing gastric discomfort. The advent of ChatGPT has transformed many people's work and lifestyles, promoting a global wave of AI popularization and application. This article reports insights into exploring ChatGPT's potential in disseminating gastric cancer knowledge, providing consultation recommendations, and endoscopy report analysis, ultimately benefiting both patients and doctors.

II. MATERIALS AND METHODS

A. Materials

For the medical knowledge test on gastric cancer, we tested twenty-three questions on five aspects, including concept, prevention, screening, treatment, and complications. For the consultation recommendations and endoscopy report analysis, our case materials were obtained from the Chinese Medical Case Repository, Journal of Medical Case Reports, and F1000 Research [20]-[26].

B. Methods

The tests were executed on GPT-3.5 and GPT-4, generating three responses per question. The responses were evaluated in two aspects:

- **Appropriateness:** If one of the responses is wrong, including obvious false statements, inaccurate diagnoses, or inappropriate suggestions, the response will be judged as inappropriate.
- **Consistency:** The consistency between the three responses is judged at an approximate level; there are two situations that are considered consistent in the experiments: being all correct for the same idea or making the same mistake for the same idea; if they are inconsistent, they will be considered unreliable.

There are four situations shown in Table I and Fig.1. Thumbs-up means the response is right and thumbs-down means the response is wrong. The same color means that all three responses are for the same idea and are consistent.

TABLE I. EVALUATION OF APPROPRIATENESS AND CONSISTENCY

| | Appropriate | Inappropriate |
|--------------|-----------------------|-----------------------|
| Consistent | All thumbs-up | All thumbs-down |
| | All in same color | All in same color |
| Inconsistent | All thumbs-up | Not all thumbs-up |
| | Have different colors | Have different colors |

If inappropriate responses are generated, we tested the error-correction capability of ChatGPT. We classify its error-correction capabilities into two categories: automatic error-correction and guided error-correction.

- **Automatic error correction** refers to ChatGPT's ability to detect and correct errors upon its first attempt after receiving a prompt.
- **Guided error correction** refers to ChatGPT's capacity to identify errors after a user narrows down the paragraph range, even if it did not initially detect the error.

In consultation recommendations and endoscopy report analysis tests, with the support of real pathological diagnostic results, we focus more on whether ChatGPT can detect abnormalities and the consistency of the three responses.

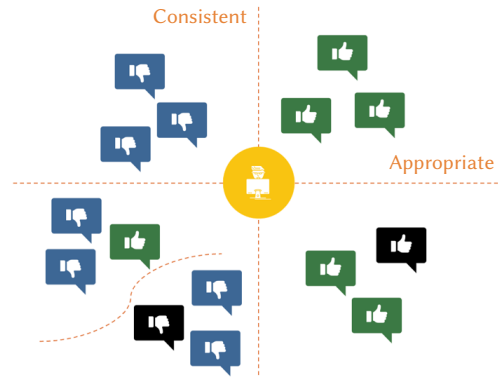


Fig. 1. Evaluation of appropriateness and consistency.

III. EXPERIMENT

In this section, we will sequentially introduce the experiments, including medical knowledge, consultation recommendations, and endoscopy report analysis tests. The general flow of the experiments is shown in Fig. 2.

A. Medical Knowledge Test

In this knowledge test about gastric cancer, we tested each question separately with GPT-3.5 and GPT-4. Three responses were generated for each question, and we evaluated the appropriateness and consistency. Table II shows the 23 fundamental questions about gastric cancer that we used for testing.

TABLE II. EVALUATION OF CHATGPT FOR QUESTIONS ABOUT GASTRIC CANCER

| Question | GPT-4 ^a | GPT-3.5 ^a |
|--|--------------------|----------------------|
| 1. What is gastric(stomach) cancer? | C/A | C/A |
| 2. What are the stages of gastric(stomach) cancer? | C/NA | NC/NA |
| 3. According to TNM stage, which stage does early gastric(stomach) cancer belong to? | C/A | NC/NA |
| 4. How to prevent gastric(stomach) cancer? | C/A | C/A |
| 5. How to prevent H.pylori infection? | C/A | NC/NA |
| 6. What are the symptoms of H. pylori infection? | C/A | C/A |
| 7. Who should seek diagnosis and treatment of an H. pylori infection? | C/A | C/A |
| 8. Is my risk for gastric(stomach) cancer higher if my family member has it? | C/A | C/A |
| 9. Can gastric ulcers develop into gastric(stomach) cancer? | C/A | C/A |
| 10. What is gastric(stomach) cancer screening? | C/A | C/A |
| 11. Are there risks associated with gastric(stomach) cancer screening? | C/A | C/A |
| 12. How do I know if I need to be screened for gastric(stomach) cancer? | C/A | C/A |
| 13. Is there anything I should prepare for before the gastric(stomach) cancer screening? | C/A | NC/NA |
| 14. What are the symptoms of gastric(stomach) cancer? | C/A | C/A |
| 15. What is cagA-positive H. pylori and how does it affect the risk of gastric(stomach) cancers? | C/A | C/A |
| 16. Does gastroscopy require anaesthesia? | C/A | C/A |
| 17. Is it possible for children to get gastric(stomach) cancer? | C/A | C/A |
| 18. How to treat gastric(stomach) cancer? | NC/NA | C/NA |
| 19. How to treat stage 0 gastric(stomach) cancer? | C/A | C/A |
| 20. How to treat early gastric(stomach) cancer? | C/A | C/NA |
| 21. How to treat advanced gastric(stomach) cancer? | C/A | C/A |
| 22. What is HER2 testing? | C/A | C/A |
| 23. Can a patient with gastric(stomach) cancer get pregnant? | C/A | C/A |

^a where C for consistent, A for appropriate, NC for inconsistent, and NA for inappropriate

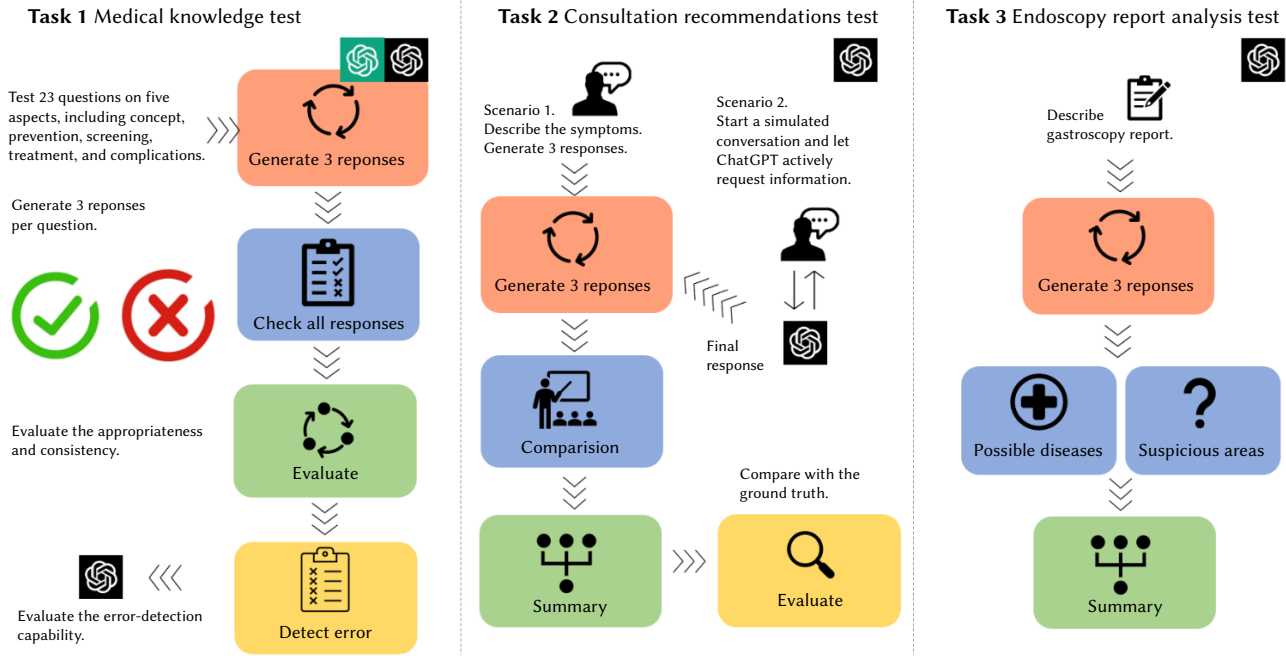


Fig. 2. General flow of the experiments.

Human: Preventing *Helicobacter pylori* (*H. pylori*) infection can be challenging, as it is often spread through close contact with an infected person, and can also be acquired through contaminated food or water. However, there are several steps you can take to reduce your risk of *H. pylori* infection: Practice good hygiene: Wash your hands frequently with soap and water, especially before eating or preparing food, and after using the bathroom or changing a diaper. Avoid contaminated food and water: Drink clean, safe water, and avoid eating food that may be contaminated, such as undercooked meat or shellfish, unpasteurized milk or dairy products, or food that has been improperly stored or handled. Avoid sharing personal items: *H. pylori* can be spread through saliva, so avoid sharing utensils, cups, or other personal items with someone who may be infected. Get vaccinated: Although there is no vaccine specifically for *H. pylori*, vaccination against other diseases such as hepatitis A and B, and human papillomavirus (HPV) can help reduce your risk of infection and may indirectly reduce the risk of *H. pylori* infection. Avoid smoking: Smoking has been linked to an increased risk of *H. pylori* infection, so quitting smoking or avoiding exposure to secondhand smoke may help reduce your risk. If you suspect you may have been exposed to *H. pylori* or are experiencing symptoms such as abdominal pain, nausea, or indigestion, talk to your doctor about testing and treatment options. Early detection and treatment of *H. pylori* infection can help prevent complications and reduce the risk of gastric cancer. Is it correct?

GPT-4: Yes, the information provided in the paragraph is correct. It outlines various steps individuals can take to reduce their risk of *H. pylori* infection, including practicing good hygiene, avoiding contaminated food and water, not sharing personal items, getting vaccinated for related diseases, and avoiding smoking. Additionally, it encourages people to consult a doctor if they suspect exposure to *H. pylori* or experience related symptoms

Human: Get vaccinated: Although there is no vaccine specifically for *H. pylori*, vaccination against other diseases such as hepatitis A and B, and human papillomavirus (HPV) can help reduce your risk of infection and may indirectly reduce the risk of *H. pylori* infection. Is it correct?

GPT-4: I apologize for the confusion in my previous response. The statement about vaccination is not correct. There is currently no vaccine for *H. pylori*, and vaccination against other diseases like hepatitis A and B, and human papillomavirus (HPV) will not directly or indirectly reduce the risk of *H. pylori* infection. These vaccinations are important for preventing the diseases they target, but they have no effect on *H. pylori* infection risk.

Human: The stages of gastric (stomach) cancer are determined based on the size and extent of the tumor, as well as the spread of cancer to nearby lymph nodes or distant organs. The stages are as follows:
 Stage 0: Also known as carcinoma in situ, the cancer cells are only in the inner lining of the stomach and have not spread to other tissues.
 Stage I: The cancer has grown into the deeper layers of the stomach wall but has not spread to the lymph nodes or other organs.
 Stage II: The cancer has grown through the full thickness of the stomach wall and may have spread to nearby lymph nodes.
 Stage III: The cancer has spread to nearby lymph nodes and may have invaded nearby organs such as the spleen or pancreas.
 Stage IV: The cancer has spread to distant organs such as the liver, lungs, or bones.
 Doctors use a variety of tests to determine the stage of gastric cancer, including imaging tests like CT scans, MRI scans, and PET scans, as well as endoscopic biopsies and blood tests. Treatment options for gastric cancer depend on the stage and location of the cancer, but may include surgery, chemotherapy, radiation therapy, targeted therapy, and immunotherapy.
 Is it correct?

GPT-4: Yes, the information provided in the paragraph is correct. It accurately describes the different stages of gastric (stomach) cancer, ranging from Stage 0 (carcinoma in situ) to Stage IV (cancer that has spread to distant organs). The paragraph also lists various tests that doctors use to determine the stage of gastric cancer and mentions possible treatment options based on the stage and location of the cancer.

Human: Stage I: The cancer has grown into the deeper layers of the stomach wall but has not spread to the lymph nodes or other organs. Is it correct?

GPT-4: The statement is not entirely accurate. Stage I gastric cancer is further divided into two sub-stages, which consider the involvement of lymph nodes:
 Stage IA: The cancer has grown into the deeper layers of the stomach wall (either the submucosa or the muscle layer) but has not spread to the lymph nodes or other organs.
 Stage IB: The cancer has either grown into the muscle layer of the stomach wall and spread to 1 or 2 lymph nodes, or it has grown through the muscle layer but has not spread to any lymph nodes or other organs.
 Thus, the correct description of Stage I gastric cancer should account for these sub-stages and their relation to lymph node involvement.

Fig. 3. Example of guided error correction.

TABLE III. SUMMARY OF CONSULTATION RECOMMENDATIONS

| Case ID | Pathological confirmation | Symptom Keywords | Inference results/Possible diseases ^a |
|---------|--|---|--|
| 1 | Gastric adenocarcinoma & Pernicious anemia | 1. 61-year-old female 2. intermittent social alcohol consumption 3. epigastric pain 4. heartburn 5. decreased appetite 6. weakness 7. unintentional weight loss | 1. Gastritis (3) 2. Peptic ulcer disease (3) 3. Gastroesophageal reflux disease (GERD) (2) 4. Gastrointestinal malignancy (3) |
| 2 | Gastric adenocarcinoma of the fundic gland type, GA-FG | 1. 73-year-old male 2. discomfort in the upper and middle abdomen 3. more than 1 month 4. H. pylori history | 1. Recurrent H. pylori infection (2) 2. Gastritis (3) 3. Peptic ulcer disease (3) 4. Functional dyspepsia (3) 5. Gastroesophageal reflux disease (GERD) (3) |
| 3 | Gastric adenocarcinoma of the fundic gland type, GA-FG | 1. 52-year-old female 2. vague stomach pains 3. pains worse when hungry 4. pains be relieved after eating | 1. Peptic ulcer disease (3) 2. Gastritis (3) 3. Functional dyspepsia (3) 4. Gastroesophageal reflux disease (GERD) (3) |
| 4 | Gastric cancer combined with duodenal cancer | 1. 81-year-old male 2. abdominal pain 3. black stools, unformed 4. persistent epigastric pain 5. abdominal distension unrelated to food intake 6. occasional nausea and vomiting 7. weight loss | 1. Upper gastrointestinal bleeding (3) 2. Gastrointestinal bleeding (3) 3. Gastrointestinal malignancy (3) 4. Peptic ulcer disease (3) 5. Gastritis (2) 6. Gastric outlet obstruction (2) |

^a where the number in (*) after each possible disease indicates the number of occurrences in the three responses generated by ChatGPT.

After the experiment, the results are recorded and compared to the ground truth from the golden standards. For GPT-4, 91.3% of the responses are appropriate with a consistency of 95.7%; for GPT-3.5, the accuracy is not good, and only 73.9% of the responses are appropriate with a consistency of 82.6%. GPT-4 has a more powerful reasoning ability, but that comes with a paid subscription fee, it does not respond as fast as GPT-3.5. Prior to this, we also found that GPT-4 has higher completeness and expansiveness in its responses.

Next, we will focus only on GPT-4 because the tolerance for errors is relatively low in medical scenarios. Although GPT-3.5 is fast in response, it is unsuitable for application in medical scenarios. Based on our experimental results, a more rigorous statement would be that GPT-3.5 may be less suitable for this specific scenario of gastric cancer consultation.

In the responses obtained from the experiment, some were considered inappropriate. We used these data to examine GPT-4's error correction ability (shown in Fig. 3). The prompt we used in the error correction experiment is:

[Paragraphs to be checked]

Is it correct?

Unfortunately, under this prompt, GPT-4 does not perform well in error correction. When the prompt was first sent, GPT-4 did not immediately detect the paragraph error, but once the scope was artificially narrowed, GPT-4 still had the ability to correct the error (Fig. 3). Despite this, it shows that ChatGPT has the potential for error correction.

B. Consultation and Diagnosis Recommendations

In this experiment, we tested the feasibility of GPT-4 acting as if it is an "internet doctor." Again, we had ChatGPT generating three responses for each question. Table III contains some key points derived from the patients' chief complaints and ChatGPT's responses. The prompt we used is:

[Paragraph]

You are a specialist gastroenterologist,
please make a diagnosis of this case description.

The table presents four cases of patients with gastric cancer; however, only cases 1 and 4 were suggested as potentially having gastrointestinal malignancy. All three responses alluded to this information, emphasizing its significance. These two cases provided comprehensive information, while the other two cases did not indicate a risk of gastric cancer for two reasons: the information provided was insufficient, and there was a lack of critical information. It is important to note that many gastric diseases share similar symptoms, and complications may coexist, making diagnosis difficult with limited information. Consequently, GPT-4 listed several possible diseases for each case.

To address the issue of insufficient information, there are two possible scenarios. The first is human-driven, wherein users actively provide adequate information. This approach is crucial, as ChatGPT may not necessarily seek additional information but directly guess a diagnosis, which is beyond the user's control and may leave patients uninformed due to a lack of medical knowledge. The second scenario involves having ChatGPT proactively seeking information, allowing it to determine the necessary details (Fig. 4). It is crucial to emphasize the importance of providing sufficient information for accurate diagnosis, to consider the potential benefits of ChatGPT, and to actively seek additional information to improve diagnostic accuracy. For example:

Let's have a conversation. Suppose you are a professional gastroenterologist, and I am a patient who comes to consult you, [What specifically do you want to consult] please try to ask for more information from me to analyze my disease and give me suggestion.

In the realm of gastric cancer and other gastric diseases, GPT-4 consistently emphasizes the necessity for patients to undergo a comprehensive evaluation by a healthcare professional. This evaluation should include appropriate laboratory tests, imaging studies, and potentially an endoscopic examination to reach a definitive diagnosis and initiate suitable treatment. It is important to recognize that GPT-4 can only indicate potential risks, and a definitive diagnosis can only be made following a thorough examination. Additionally, GPT-4 can offer guidance on diagnostic procedures.

| | |
|---|--|
| <p>Clinician: <i>Let's have a conversation. Suppose you are a professional gastroenterologist, and I am a patient who comes to consult you, please try to ask for more information from me to analyze my disease and give me suggestions.</i></p> <p>GPT-4: <i>Hello, I'm Dr. ChatGPT, a gastroenterologist. I'm here to help you with your gastrointestinal concerns. To better understand your situation, I will need to ask you a few questions.</i></p> <ol style="list-style-type: none"> 1. Can you please describe your main symptoms, including their severity, frequency, and duration? 2. Have you noticed any specific triggers for these symptoms, such as certain foods or activities? 3. Have you experienced any recent changes in your bowel habits (e.g., constipation, diarrhea, or alternating between the two)? 4. Are you experiencing any pain or discomfort in your abdomen? If so, can you please describe the location, intensity, and any factors that seem to relieve or worsen the pain? 5. Have you noticed any blood in your stool, or black, tarry stools? 6. Do you have a history of gastrointestinal issues, either personally or in your family? 7. Are you currently taking any medications or supplements? 8. Have you recently lost or gained weight unintentionally? <p><i>Please provide as much detail as possible, as this will help me better understand your situation and provide appropriate suggestions.</i></p> | <p>Clinician: <i>Let's have a conversation. Suppose you are a professional gastroenterologist, and I am a patient who comes to consult you. I want to consult if I have gastric cancer, please try to ask for more information from me to analyze my disease and give me suggestions.</i></p> <p>GPT-4: <i>Hello, I'm glad you've come to consult with me. As a gastroenterologist, I'm here to help you with any concerns you may have about your gastrointestinal health. To better understand your situation and give you the most accurate advice possible, I'll need to ask you some questions. Please remember that I am an AI language model and not a real doctor, but I will try my best to provide helpful information based on your responses.</i></p> <ol style="list-style-type: none"> 1. Can you please tell me your age, gender, and any relevant medical history, including any family history of gastric cancer or other gastrointestinal diseases? 2. What specific symptoms have you been experiencing and for how long? 3. Have you noticed any changes in your appetite, weight, or bowel habits recently? 4. Do you have any history of gastrointestinal problems, such as ulcers, GERD, or gastritis? 5. Are you currently taking any medications, or have you taken any medications recently, particularly nonsteroidal anti-inflammatory drugs (NSAIDs) or antacids? 6. Do you smoke or consume alcohol? If so, how much and how often? <p><i>Please provide as much information as you can so I can better understand your situation and provide appropriate suggestions.</i></p> |
|---|--|

Fig. 4. Example of ChatGPT requesting information.

C. Endoscopy Report Analysis

The primary objective of this experiment was to assess GPT-4's supplementary role in endoscopic examination reports. By describing the observations made during endoscopy, GPT-4 can assist in analysing suspicious regions and reconfirming biopsy sites. It is crucial to note that the focus of this experiment is not on the accuracy of GPT-4's analysis of medical mechanisms but rather on its ability to detect abnormalities. As in previous experiments, three responses were generated for each query, and if the content of the responses was not entirely consistent, they could all be used as reference points. The prompt used in this experiment is as follows:

Gastroscopy report: [Description]

With this description, please give me some diagnostic hints.

Fig. 5 summarizes the results of GPT-4's endoscopy report analysis. Through experiments, we find that GPT-4 does have the ability to analyse endoscopy reports.

IV. DISCUSSION

The launch of ChatGPT has changed the way many people work and live. As a general deep learning language model, it has applications in many areas. In this experiment, we primarily focus on the medical field, using gastric cancer as a specific scenario to test ChatGPT's application. Our tests revolve around gastric cancer medical knowledge test, consultation recommendations, and endoscopy report analysis.

First, in the medical knowledge test, we compared the performance of GPT-4 and GPT-3.5 by testing both models with the same questions, generating three responses per question. GPT-4's appropriateness reached 91.3%, with a consistency of 95.7%. As a general model, it is already quite powerful without having been specifically trained on a particular type of data. However, it is important to note that we should not overly rely on ChatGPT for clinical diagnosis; as it has stated itself, it is not a doctor who can provide sound advice. According to the experiment, ChatGPT is prone to making mistakes in certain details or expressions, such as the specific staging of cancer, and its responses are not very satisfying in regard to detailed classifications. In the error correction test, GPT-4 also overlooked many details and did not detect errors as expected, only considering them after narrowing down the scope. Of course, this could also be an issue with the prompt. After ChatGPT's popularity, prompt engineering has become an important technology. One thing to note in the medical knowledge test session is that GPT-4 has tried to provide more comprehensive responses, including expanding on related but unasked content. However, by generating three responses to the same question, it sometimes

still misses some key points. Therefore, one suggestion when using ChatGPT is to generate multiple responses to consider.

In the consultation recommendations test, with rich case information, we observed GPT-4's excellent diagnostic assistance capabilities. However, due to the complex gastric environment in our body, numerous complications, and similar symptoms across various gastric diseases, GPT-4's diagnostic suggestions are within a relatively broad range. This could serve as preliminary information before seeking medical attention, but a hospital examination is still necessary and very advisable for a definite diagnosis.

In the endoscopy report analysis test, GPT-4 also demonstrated outstanding reasoning abilities, identifying possible diseases based on the text description of the report, extracting information on suspicious areas, and suggesting biopsy and pathological examinations. However, there are limitations to this, as in many cases, biopsies are directly taken during endoscopic examinations before the report is available, making the diagnostic assistance a secondary confirmation. Nevertheless, it can still provide diagnostic suggestions in ordinary gastric disease endoscopic diagnoses.

In summary, through the experiment, we discovered more things that ChatGPT can do than expected. We should learn to utilize it without relying on it and always remember that it is a chatbot, not a person. Although ChatGPT has a wide range of application potential, it is important to note that as an AI model, it is not flawless [27]. In some cases, it may not provide accurate responses or suggestions. Therefore, users should exercise caution when using ChatGPT and seek professional advice from certified experts when necessary.

V. FUTURE PERSPECTIVES

At present, most of the application of artificial intelligence in gastric cancer is primarily concentrated on the field of computer vision, such as autorecognition of gastric lesions under digital endoscopy. This task of object detection aims to identify suspicious targets from endoscopic images and classify gastric lesion types. Although the experiments in this study focus on the field of NLP, we can gain valuable insights from our endoscopic report analysis experiment. With the advancement of multimodal research, we believe that natural language processing and computer vision can be more deeply and effectively integrated in the future, such as through some automatic co-analysis and generation of unanimous reports when suspicious areas are detected during endoscopy inspection. The results of object detection and language reasoning can also be mutually verified, which would collaboratively produce a significant contribution to medical diagnosis.

Case 1 Description:

Gastroscopy report: (1) The gastric mucosa is smooth and free of erosions and ulcers. Local patchy redness is observed in the gastric antrum. The mucosa from the gastric antrum to the gastric angle is thinned, with visible blood vessels, and atrophy extends beyond the cardia. A yellowish nodular lesion is observed near the cardia at the gastric fundus. The comprehensive evaluation suggests a presentation after *Helicobacter pylori* (HP) eradication. (2) A discolored, slightly yellowish lesion is observed on the anterior wall of the upper gastric body near the greater curvature, measuring approximately 4 mm x 5 mm. Upon closer inspection with a regular endoscope, the lesion appears slightly elevated with a yellowish hue, with visible dilated blood vessels and black dot-like changes on the surface. Under narrow-band imaging (NBI), there is a color difference boundary between the lesion and the surrounding normal mucosa. Under narrow-band imaging-magnifying endoscopy (NBI-ME), there is no distinct boundary between the lesion and the surrounding normal mucosa. The microsurface structure of the lesion is regular/y arranged, with slightly widened crypt intervals. No significant irregularities are observed in the microvascular structure. After acetic acid staining, the whitening time of the lesion is shortened. The lesion becomes more prominent after indigo carmine staining.

Inference results (Possible diseases):

1. Atrophic gastritis (3)
2. Macular tumor near the cardia -> suggest biopsy and histopathological examination (3)
3. Discolored and slightly yellowish lesion -> suggest biopsy and histopathological examination (3)

Case 2 Description:

Gastroscopy report: A 1.0 cm diameter region with slightly red coloration is observed near the cardia on the lesser curvature of the stomach, with a slightly depressed center and a small amount of thin white coating on the surface. The lesion boundary is clear under narrow-band imaging magnification endoscopy (NBI-ME), with thickened, tortuous microvessels forming a mesh-like pattern, and partial disappearance of microstructures. In the gastric angle, there is rough mucosa with nodular changes, measuring approximately 4.0 cm x 5.0 cm. NBI-ME reveals multiple foci of dilated, thickened, and irregular microvessels in the gastric angle mucosa, with partial fusion of microstructures. On the anterior wall of the gastric sinus near the greater curvature, there is a 0.8 cm diameter area of rough mucosa. The lesion boundary is clear under NBI-ME, with fused glandular ducts, partial disappearance, and malformed microvessels. On the greater curvature side of the gastric sinus and body junction, there is a 1.0 cm diameter area with a reddish center. NBI-ME reveals a brownish coloration of the lesion, with clear boundaries, partial fusion, and even absence of glandular ducts, as well as thickened, malformed microvessels. A 1.0 cm diameter area of slightly raised mucosa is observed on the anterior wall of the gastric sinus. NBI-ME reveals a small amount of fused glandular ducts and partially linear, dilated, and deformed microvessels.

Inference results (Possible diseases):

1. Gastric lesion with slight depression -> suggest biopsy and histopathological examination (3)
2. Rough and nodular mucosa of the gastric body -> suggest biopsy and histopathological examination (3)
3. Gastric sinus lesions -> suggest biopsy and histopathological examination (3)

Case 3 Description:**Gastroscopy report:**

Esophagus: The mucosal surface is smooth and moist, with normal vascular patterns.

Cardia: Good closure, with smooth mucosa and no folds.

Gastric fundus: Mucus pool with bile-like fluid present. Smooth mucosa and clear, visible blood vessels. Large folds resembling gyri.

Gastric body: Smooth mucosa, with submucosal blood vessels visible. Regular arrangement of large folds along the greater curvature: normal peristaltic waves can be seen.

Gastric antrum: Intact morphology, arch-shaped, with smooth and orderly angular notches.

Pylorus: Alternating opening and closing; when open, it appears round with smooth and even edges.

Duodenal bulb: Spherical, with smooth mucosa without folds, presenting a velvety appearance.

Descending duodenum: Circular folds are clearly visible, and the papilla is normal.

Inference results (Possible diseases):

1. The presence of bile-like fluid in the mucus lake may suggest some reflux of bile from the duodenum into the stomach. (3)

Fig. 5. Summary of gastroscopy report analysis.

ACKNOWLEDGEMENT

The authors are thankful for the financial support from research grants by Guangzhou Development Zone Science and Technology (2021GH10, 2020GH10), research grant by the University of Macau (MYRG2022-00271-FST) and research grant by the Science and Technology Development Fund of Macau (0032/2022/A).

REFERENCES

- [1] Y. Horie et al., "Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks," *Gastrointestinal Endoscopy*, vol. 89, no. 1, pp. 25-32, 2019, doi: 10.1016/j.gie.2018.07.037.
- [2] A. S. Kumarasuvamy and R. S. Rajendran, "Design an Early Detection and Classification for Diabetic Retinopathy by Deep Feature Extraction based Convolution Neural Network," *Journal of Trends in Computer Science and Smart Technology*, vol. 3, no. 2, pp. 81-94, 2021, doi: 10.36548/jtcsst.2021.2.002.
- [3] T. Rahman et al., "Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-ray," *Applied Sciences*, vol. 10, no. 9, Art. no. 9, 2020, doi: 10.3390/app10093233.
- [4] A. Laishram and K. Thongam, "Automatic Classification of Oral Pathologies Using Orthopantomogram Radiography Images Based on Convolutional Neural Network," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 7, no. 4, pp. 69-77, 2022, doi: 10.9781/ijimai.2021.10.009.
- [5] M. A. Khemchandani, S. M. Jadhav, and B. R. Iyer, "Brain Tumor Segmentation and Identification Using Particle Imperialist Deep Convolutional Neural Network in MRI Images," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 7, no. 7, pp. 38-47, 2022, doi: 10.9781/ijimai.2022.10.006.
- [6] B. G. Patra et al., "Extracting social determinants of health from electronic health records using natural language processing: a systematic review," *Journal of the American Medical Informatics Association*, vol. 28, no. 12, pp. 2716-2727, 2021, doi: 10.1093/jamia/ocab170.
- [7] H. Sung et al., "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209-249, 2021, doi: 10.3322/caac.21660.
- [8] E. C. Smyth, M. Nilsson, H. I. Grabsch, N. C. van Grieken, and F. Lordick, "Gastric cancer," *The Lancet*, vol. 396, no. 10251, pp. 635-648, 2020, doi: 10.1016/S0140-6736(20)31288-5.
- [9] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, "Cancer statistics, 2022," *CA A Cancer J Clinicians*, vol. 72, no. 1, pp. 7-33, 2022, doi: 10.3322/caac.21708.
- [10] K. Yashima, M. Shabana, H. Kurumi, K. Kawaguchi, and H. Isomoto, "Gastric Cancer Screening in Japan: A Narrative Review," *Journal of Clinical Medicine*, vol. 11, no. 15, Art. no. 15, 2022, doi: 10.3390/jcm11154337.
- [11] A. Jarenlapnopparat, K. Bhatia, and S. Coban, "Inflammation and Gastric Cancer," *Diseases*, vol. 10, no. 3, Art. no. 3, 2022, doi: 10.3390/diseases10030035.
- [12] S. Kikuchi, Y. Obata, T. Sasakabe, S. Kawai, C. Wang, and Y. Lin, "Relative risk of gastric cancer between those with and without *Helicobacter pylori* infection history in Japan," *JGH Open*, vol. 6, no. 3, pp. 166-170, 2022, doi: 10.1002/jgh3.12714.

- [13] H. L. Haver, E. B. Ambinder, M. Bahl, E. T. Oluyemi, J. Jeudy, and P. H. Yi, "Appropriateness of Breast Cancer Prevention and Screening Recommendations Provided by ChatGPT," *Radiology*, p. 230424, 2023, doi: 10.1148/radiol.230424.
- [14] T. H. Kung et al., "Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models," *PLOS Digital Health*, vol. 2, no. 2, p. e0000198, 2023, doi: 10.1371/journal.pdig.0000198.
- [15] R. Vaishya, A. Misra, and A. Vaish, "ChatGPT: Is this version good for healthcare and research?," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 17, no. 4, p. 102744, 2023, doi: 10.1016/j.dsx.2023.102744.
- [16] S. Biswas, "ChatGPT and the Future of Medical Writing," *Radiology*, vol. 307, no. 2, p. e223312, 2023, doi: 10.1148/radiol.223312.
- [17] M. Sallam, "ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns," *Healthcare*, vol. 11, no. 6, Art. no. 6, 2023, doi: 10.3390/healthcare11060887.
- [18] E. Kasneci et al., "ChatGPT for good? On opportunities and challenges of large language models for education," *Learning and Individual Differences*, vol. 103, p. 102274, 2023, doi: 10.1016/j.lindif.2023.102274.
- [19] P.-H. Niu, L.-L. Zhao, H.-L. Wu, D.-B. Zhao, and Y.-T. Chen, "Artificial intelligence in gastric cancer: Application and future perspectives," *World J Gastroenterol*, vol. 26, no. 36, pp. 5408–5419, 2020, doi: 10.3748/wjg.v26.i36.5408.
- [20] Z. X. Dong et al., "Endoscopic diagnosis and treatment for gastric adenocarcinoma of fundic gland type: report of 3 cases," *Chin J Dig Endosc*, vol. 39, no. 11, pp.931-934, 2022, doi: 10.3760/cma.j.cn321463-20210325-00206
- [21] H. Y. Dong, X. Y. Jia, Q. P. Pang, G. D. Li, "A case of simultaneous five-origin early gastric cancer treated by gastroscopy," *Chin J Dig Endosc*, vol. 34, no. 12, pp. 913-914, 2017, doi: 10.3760/cma.j.issn.1007-5232.2017.12.018
- [22] H. C. Huang et al., "[A case report of gastric adenocarcinoma of fundic gland type as suspected gastric neuroendocrine tumor]," *Zhonghua Nei Ke Za Zhi*, vol. 61, no. 6, pp. 685–687, 2022, doi: 10.3760/cma.j.cn112138-20210714-00482.
- [23] S. Kamran, M. K. Dilling, N. A. Parker, J. Alderson, N. D. Tofteland, and Q. V. Truong, "Case Report: Simultaneously, diagnosed gastric adenocarcinoma and pernicious anemia – a classic association." *F1000Research*, 2020. doi: 10.12688/f1000research.24353.2.
- [24] M. Moriya, A. Uehara, T. Okumura, M. Miyamoto, and Y. Kohgo, "Stress-induced hemorrhagic gastric ulcer after successful Helicobacter pylori eradication: two case reports," *Journal of Medical Case Reports*, vol. 5, no. 1, p. 252, 2011, doi: 10.1186/1752-1947-5-252.
- [25] K. Barrett, M. W. Hii, and R. J. Cade, "Benign gastro-colic fistula in a woman presenting with weight loss and intermittent vomiting: a case report," *Journal of Medical Case Reports*, vol. 5, no. 1, p. 313, 2011, doi: 10.1186/1752-1947-5-313.
- [26] I. B. Ismail, H. Zenaidi, R. Jouini, S. Rebi, and A. Zoghalmi, "Case Report: Primary pure clear cell gastric carcinoma." *F1000Research*, 2020. doi: 10.12688/f1000research.25810.1.
- [27] F. C. Kitamura, "ChatGPT Is Shaping the Future of Medical Writing However, Still Requires Human Judgment," *Radiology*, vol. 307, no. 2, p. e230171, 2023, doi: 10.1148/radiol.230171.

Jiaming Zhou



She received a B.Ec. degree in Economic Statistics from Guangdong University of Foreign Studies. She is currently pursuing an M.Sc. degree with the University of Macau, Macau China. She is a Research Assistant in Data Analytics and Collaborative Computing Laboratory (DACC). She has experience in data mining and data analysis. Her research interests are now focused on deep learning, computer

vision algorithms, medical image processing, and other artificial intelligence applications in medicine.

Tengyue Li



Gloria Tengyue Li is currently a PhD student at the University of Macau. She is also the Head of Data Analytics and Collaborative Computing Laboratory, Zhuhai Institute of Advanced Technology, Chinese Academy of Science, Zhuhai, China. Ms Li is leading and managing the laboratory in R&D as well as technological transfer and incubation. She is an entrepreneur with experience in innovative I.T.

contest, with her award-winning team in the Bank of China Million Dollar Cup competition. Her latest winning work includes the first unmanned supermarket in Macau enabled by the latest sensing technologies, face recognition and e-payment systems. She is also the founder of several Online2Offline dot.com companies in trading and retailing both online and offline. Ms Li is also an active researcher, manager and chief knowledge officer in the DACC laboratory at the Faculty of Science and Technology, University of Macau.

Simon James Fong



He graduated from La Trobe University, Australia, with a 1st Class Honours BEng. Computer Systems degree and a PhD. Computer Science degrees in 1993 and 1998, respectively. Simon is now working as an Associate Professor at the Computer and Information Science Department of the University of Macau, as an Adjunct Professor at Faculty of Informatics, Durban University of Technology, South

Africa. He is a co-founder of the Data Analytics and Collaborative Computing Research Group in the Faculty of Science and Technology. Prior to his academic career, Simon took up various managerial and technical posts, such as systems engineer, IT consultant and e-commerce director in Australia and Asia. He published more than 200 papers in SCI journals. He serves on the editorial boards of the Journal of Network and Computer Applications of Elsevier, IEEE IT Professional Magazine, and various special issues of SCIE-indexed journals. Simon is also an active researcher with leading positions such as Vice-chair of IEEE Computational Intelligence Society (CIS) Task Force on "Business Intelligence and Knowledge Management", TC Chair of IEEE ComSoc e-Health SIG and Vice-director of International Consortium for Optimization and Modelling in Science and Industry (iCOMSI).

Nilanjan Dey



Nilanjan Dey is an Associate Professor in the Department of Computer Science and Engineering at Techno International New Town (Formerly known as Techno India College of Technology), Kolkata, India. He is a visiting fellow of the University of Reading, UK. He was an honorary Visiting Scientist at Global Biomedical Technologies Inc., CA, USA (2012-2015). He was awarded his PhD. from

Jadavpur University in 2015. He is the Editor-in-Chief of International Journal of Ambient Computing and Intelligence, IGI Global. He is the Series Co-Editor of Springer Tracts in Nature-Inspired Computing, Springer Nature, Series Co-Editor of Advances in Ubiquitous Sensing Applications for Healthcare, Elsevier, Series Editor of Computational Intelligence in Engineering Problem Solving and Intelligent Signal processing and data analysis, CRC. His main research interests include Medical Imaging, Machine learning, Computer Aided Diagnosis, Data Mining, etc. He is a senior member of IEEE and fellow of IETE.

Rubén González Crespo



Dr. Rubén González Crespo has a PhD in Computer Science Engineering. Currently he is Vice Chancellor of Academic Affairs and Faculty from UNIR and Global Director of Engineering Schools from PROEDUCA Group. He is an advisory board member for the Ministry of Education at Colombia and an evaluator from the National Agency for Quality Evaluation and Accreditation of Spain (ANECA).

He is a member from different committees at the ISO Organization. Finally, he has published more than 200 papers in indexed journals and congresses.