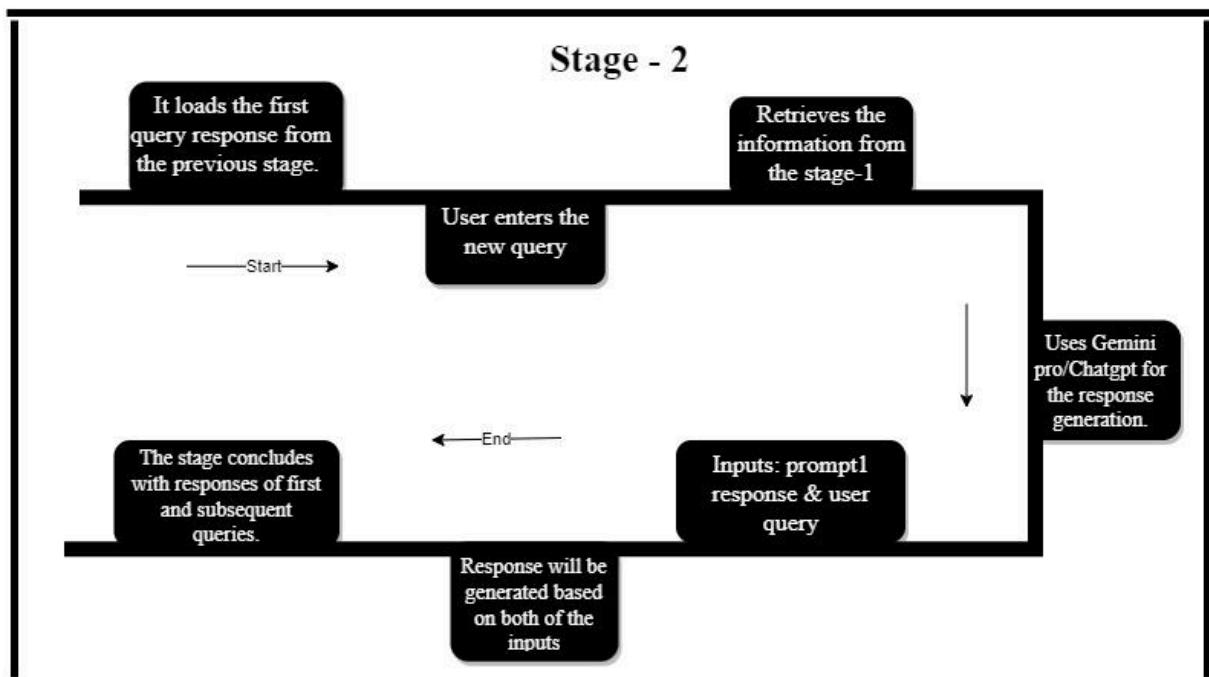
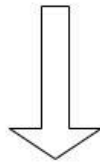
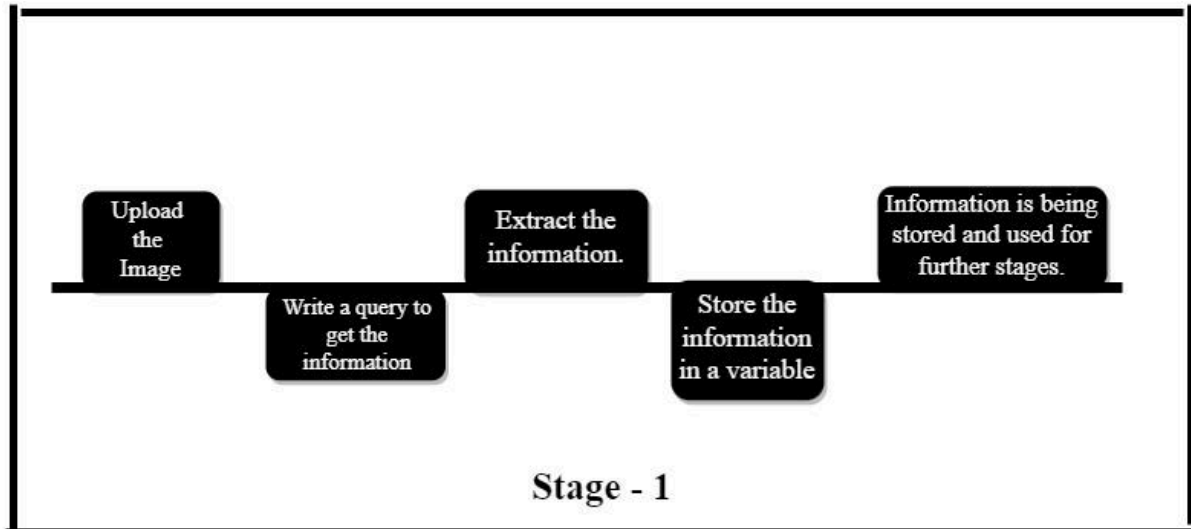


IMAGE ANALYSIS AND CLASSIFICATION

We have implemented the image analysis in the two stages :



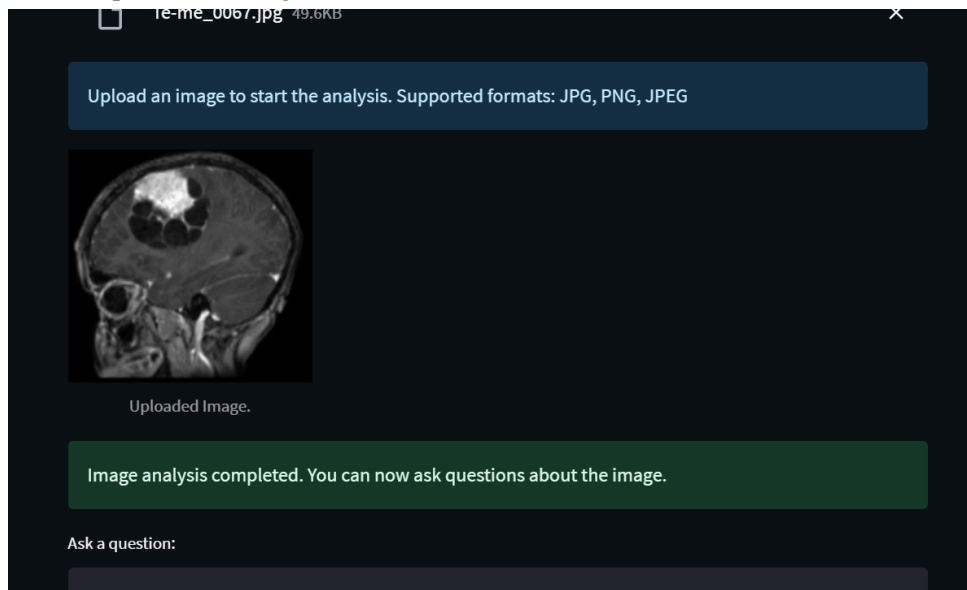
1. Image analysis using Gemini api in both the stages:

In this version we used the Gemini pro vision in the first stage to analyze the image and giving the response for this prompt: "Please analyze this image thoroughly. Identify and describe the organ or object presented in the image. Provide insights into the type of medical or scientific imaging technique that might have been used to capture this image. Additionally, assess the features or anomalies that could indicate whether the depicted subject is in a healthy or unhealthy state. Look for any specific signs or characteristics that contribute to your analysis."

And then saving the generated response in the `{img_analysis}` variable and then in the second stage we are using the Gemini-pro for the text generation, where we are giving the input as the response from the first-stage and the user question.

The generated response from the stage 2 is given to user as the answer to the user question:

When I upload this image:



The image analysis is done in the backend and saves the response.

This are the some of the user questions and the Gemini pro generated answers for that:

Chat History:

User: which organ the image belongs?

AI: Brain

User: is the condition is normal or abnormal? why?

AI: The condition is abnormal. The presence of a mass or tumor in the brain is not normal and can indicate several health conditions. While some tumors are benign and do not pose a significant health risk, others are malignant and can spread to other parts of the body. The type of tumor can be determined through further medical tests, such as a biopsy. In this case, the tumor appears to be located in a critical area of the brain, which could affect the person's cognitive and motor functions. The presence of the tumor may also cause other symptoms, such as headaches, seizures, and nausea. The overall health of the person will depend on the type of tumor, its size, and its location in the brain.

User: which doctor is to be consulted?

AI: Neurologist

User: who is the pm of india?

AI: The provided document does not contain any information about the current Prime Minister of India, so I cannot answer this question from the provided text.


2. Image Analysis using Gemini api for stage-1 and OpenAi Api for stage - 2:

For this approach we get all the details about the image using the prompt:


"Analyze the provided image in detail, identifying and describing the depicted image, whether it's an organ, object, or any other relevant entity. Look for insights into the type of medical or scientific imaging technique that might have been utilized to capture this image. What is the image related to? Additionally, assess features or anomalies indicative of the subject's health status. Try looking for any specific signs or findings that can be gathered for further analysis."

So the above prompt is provided first followed by the user's questions in the textbox after the image is uploaded in the streamlit. Then the response generated will be stored in a variable "prompt_response".

Upload an image:


 Drag and drop file here
Limit 200MB per file • JPG, PNG, JPEG, TIF, TIFF

Browse files

 kidney.jpg 9.1KB ×

Enter your questions (separate each question with a new line):
to? Additionally, assess features or anomalies indicative of the subject's health status. Try looking for any specific signs or findings that can be gathered for further analysis.
What are the symptoms of the disease?
How could this disease happen?

Image uploaded.



Uploaded Image.

Response to the prompt: The provided image is a medical imaging scan of a patient's abdomen. The imaging technique used is most likely a computed tomography (CT) scan, which utilizes X-rays and computer processing to create detailed cross-sectional images of the body.

In the image, we can observe the patient's right kidney, which appears enlarged and contains a large mass. The mass is likely a tumor, as it is significantly different in density from the surrounding healthy tissue. The presence of such a mass could indicate kidney cancer, and further imaging or biopsy may be necessary for a definitive diagnosis.

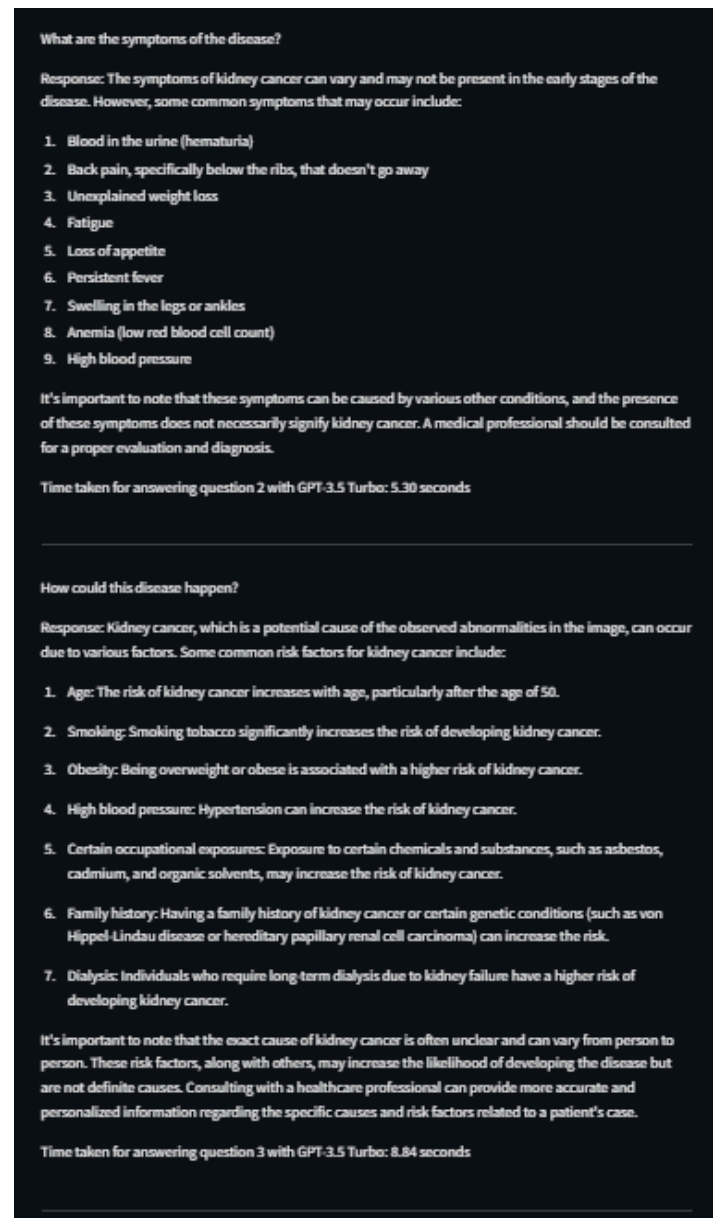
Additionally, there is a presence of contrast dye within the collecting system of the kidney, which is enhancing the visibility of the structures. This suggests that the patient may have been given an intravenous contrast agent prior to the CT scan, in order to improve the visibility of the urinary tract and enhance the diagnostic capabilities of the scan.

Overall, while this image provides valuable information about the patient's condition, a thorough assessment by a medical professional is necessary to determine the exact nature of the mass and the appropriate course of treatment.

Time taken for image reading with gemini-pro-vision: 8.71 seconds

Now the users will be answered subsequently by using "gpt-3.5-turbo" of OpenAi which takes the "prompt_response" and the user's question both as the input.

These are some of the example responses for the user's question which were answered using OpenAI.



3. Image Classification and labeling using Gemini Pro Vision:

This version is of classifying the images based on the labels given by the gemini pro vision for the image. When the user uploads the multiple images at once ,For each image, the script sends the image to the Gemini pro vision model with a label prompt.

The prompt: 'Label this image and return the top three labels with confidence scores. Give the response just the label and its confidence score, just like "label""confidence score"

After labeling for every image with their confidence scores, all the responses with their image names are saved in the json.

This are some sample images tried this is the response saved in the json:

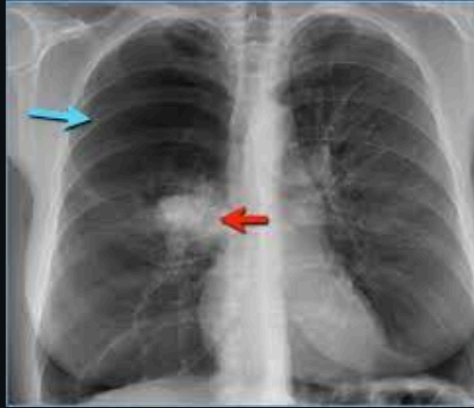
```
"6.jpg": {
  "Image Name": "6.jpg",
  "Response": "arlo\0.9999999\n\nsecurity camera\0.9999999\n\noutdoor camera\0.9999999\n\n",
},
"7.jpg": {
  "Image Name": "7.jpg",
  "Response": "\nsurveillance camera\99.9999998\n\n\ncctv camera\99.999996\n\n\nsecurity camera\99.9999999\n\n",
},
"8.jpg": {
  "Image Name": "8.jpg",
  "Response": "\ngrassland\0.9999999\n\n\nhill\0.9999999\n\n\nmountain\0.9999999\n\n",
},
"9.jpg": {
  "Image Name": "9.jpg",
  "Response": "\nmountain\0.9999999\n\n\nhill\0.9999999\n\n\nfell\0.9999998\n\n",
},
"10.jpg": {
  "Image Name": "10.jpg",
  "Response": "\nmountain\0.87\n\n\nhill\0.86\n\n\nnature\0.84\n\n",
},
"Te-me_0064.jpg": {
  "Image Name": "Te-me_0064.jpg",
  "Response": "\nastrocytoma\0.99999999\n\n\n glioma\0.99999999\n\n\n brain tumor\0.99999999\n\n",
},
"Te-me_0065.jpg": {
  "Image Name": "Te-me_0065.jpg",
  "Response": "astrocytoma\0.9999999\n\n\n\n glioma\0.9999999\n\n\n\n brain tumor\0.9999999\n\n\n"
```

Pro's:

Getting the most correct labels when compared to OpenAI and Langchain.

4. Image classification using Langchain and Transformers:

This approach to classify the multiple images based on the common objects found is done using Langchain and also with transformers specifically CLIPModel and CLIPProcessor of OpenAI. In this path of the images are placed is given in streamlit, so using



Original Image

Detected Objects:

- Class: toothbrush, Confidence: 0.99, Box: [8, 53, 43, 66]



Original Image

Detected Objects:

- Class: airplane, Confidence: 0.91, Box: [0, 0, 154, 154]

Method - 2 using CLIPModel and CLIPProcessor

This method is grouping every image under a group based on the common keywords from the caption generated. So after executing the method all the images are grouped under the same category which is not appropriate as they are different types of images.