CS492 Final Project Proposal - Group 10 - Visual Search Engine

Team members
Omar Iltaf
Kenon Kahoano
Soo-Min Yoo

App Description

The general idea of our app is to allow users to analyse pictures using the Microsoft Computer Vision API and then generate a small list of possible search queries. Once clicked, these search queries will generate a URL and implicitly open said URL in an internet browser. Additionally, all past searches will available to the user as the app will store them on-device in a database. The data returned by the Computer Vision API can include the image's description, tags, colors, and more. This means a diverse search query selection will be available, however only the most accurate and relevant will be shown to the user.

API Description

The Microsoft Vision API has a feature that can return information about visual content found in an image. Tags, descriptions, and domain-specific models can be used to identify and label content. It can also Identify image types and color schemes in pictures.
This API uses optical character recognition (OCR) to detect text in an image and extract the recognized words into a machine-readable character stream. This allows for analyzing images to detect embedded text, generate character streams, and enable searching.
Handwritten OCR can even detect and extract handwritten text from notes, letters, essays, whiteboards, forms, etc. It works with different surfaces and backgrounds, such as white paper, yellow sticky notes, and whiteboards. Handwritten text recognition saves time and effort and can increase productivity by allowing users to take photos of their text, instead of having to transcribe it. It makes it possible to digitize notes, which then allows users to implement quick and easy search.

Specific API Methods

Our app should be able to analyse images, then perform an internet search based on the image data.

- We will use HTTP POST to request data from the Computer Vision API.
- This is the general form of the request URL:
  https://[location].api.cognitive.microsoft.com/vision/v1.0/analyze[?visualFeatures][&details][&language]
- This API response will return JSON from which the necessary data about the image will be extracted to allow the user to perform a search.
- Images will be attached to the API request with the setEntity method called on the HTTP POST object.
- The response will be received by calling execute(request) on the HTTPClient object.
- From the response getEntity will be used to obtain the JSON data about our image.

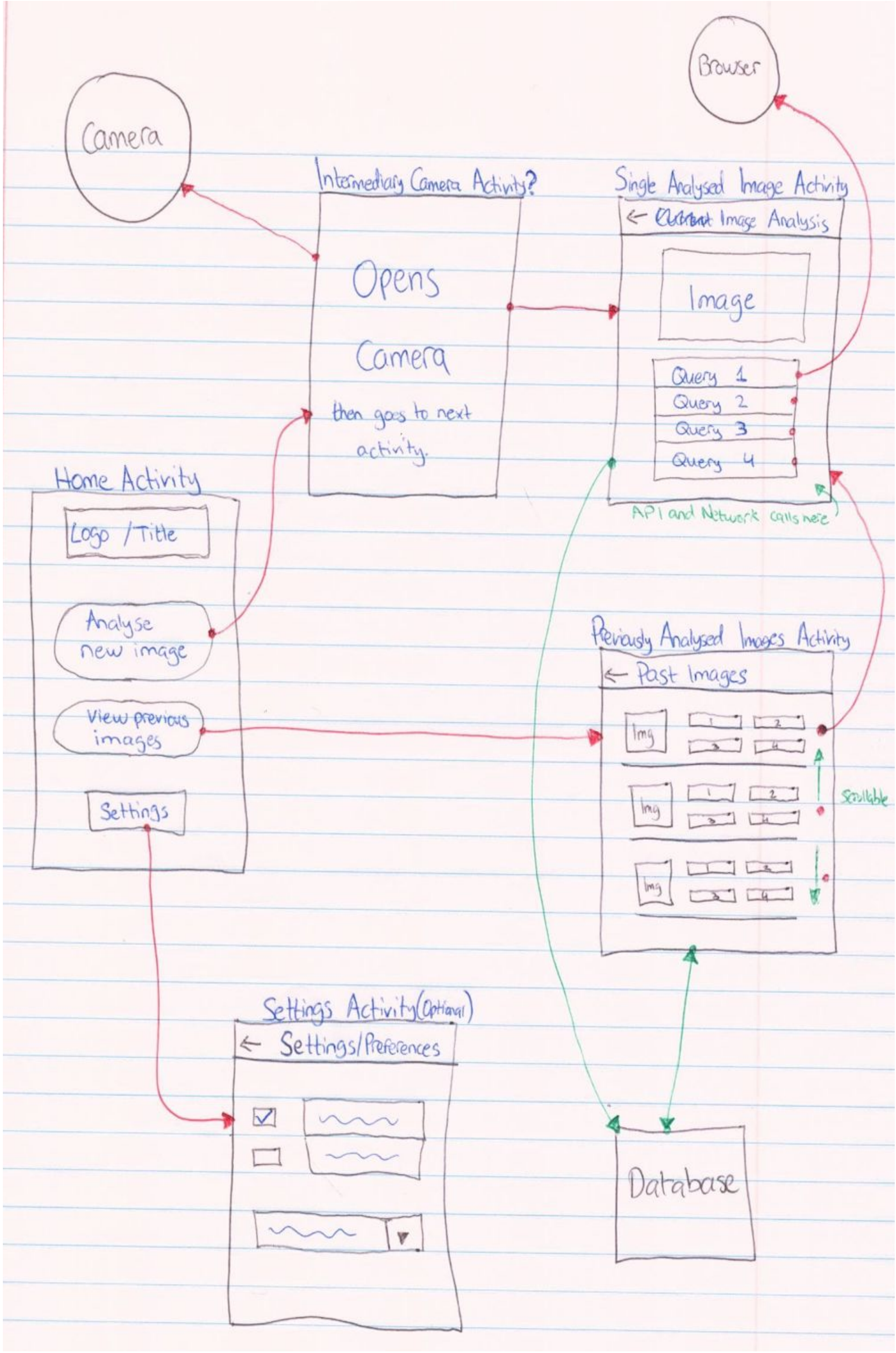
This is the general procedure of how the app will interact with the API.

User Interface Description

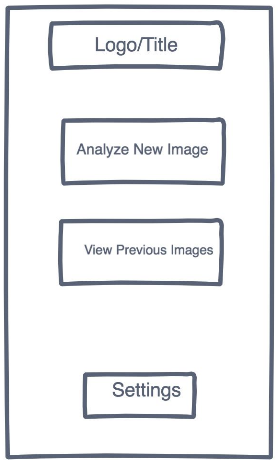
Activities:

- Home Activity
  - Title/logo at top.
  - Would contain a button to launch the Intermediary Camera Activity, then once a picture has been taken, the Single Analysed Image Activity would be launched.
  - Another button to launch the Previously Analysed Image Activity.
  - Another final button to launch the Settings Activity.
  - Buttons organised in a list.
- Intermediary Camera Activity
  - This may or may not be an actual activity. However, the camera will need to be implicitly launched from the Home Activity, at which point the Single Analysed Image Activity will be launched.
- Single Analysed Image Activity
  - This activity is where the API and network calls will take place.
  - It will contain information specific to a certain image.
  - This includes the actual image displayed at the top of the page, with the generated search queries in a list underneath (maybe upto 5 queries).
  - Clicking on a search query will open up a browser and perform a Google search.
  - When the image data has been received it, along with the image will need to be stored in the database, so it can be accessed from the Previously Analysed Images Activity.
- Previously Analysed Images Activity
  - This activity will contain a list of all previously analysed images and their generated search queries.
  - A RecyclerView may be best to implement this activity.
  - The data to fill the RecyclerView will obtained from an on-device database which will hold all data for past analysed images.
  - Clicking on an individual item will launch a Single Analysed Image Activity with the appropriate data.
- Settings Activity (Optional)
  - This activity would contain various user preferences.

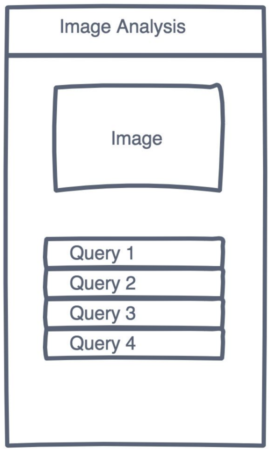
Mocks


Camera
Browser
Intermediary Camera Activity?
Opens
Camera
then goes to next activity.
Single Analysed Image Activity
← ~~Current~~ Image Analysis
Image
Query 1
Query 2
Query 3
Query 4
API and Network calls here
Home Activity
Logo / Title
Analyse new image
View previous images
Settings
Previously Analysed Images Activity
← Past Images
Img
1
2
3
4
Img
1
2
3
4
Img
1
2
3
4
Scrollable
Settings Activity (Optional)
← Settings/Preferences
☑
□
Database

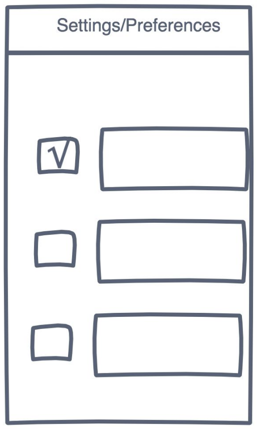
Home Activity



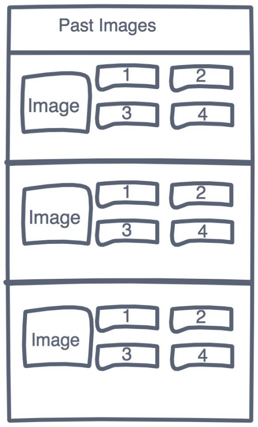
Single Analyzed Image Activity



Settings (optional)



Previously Analyzed Images Activity


Past Images
Image
1
2
3
4
Image
1
2
3
4
Image
1
2
3
4