Distribution of automatic image segmentation and tracking using Python, React, and the Segment Anything Model

Anil Kamat[1], Vincent Lau[1], Kevin Liu[1]
1. Rensselaer Polytechnic Institute, Troy, NY

Introduction

- Image segmentation is increasingly being used in a variety of applications, especially with object detection and tracking.
- However, automatic image segmentation is typically unavailable to those who are not programming-inclined.
- We use the Segment Anything Model (SAM) created by Facebook in order to automatically segment images.
- This project aims to create a publicly accessible application that is easy to set up for users, assuming limited or no experience with programming.

Tech Stack

- Back End
  - Flask, matplotlib, numpy, cv2, and SAM
- Front End
  - ReactJS

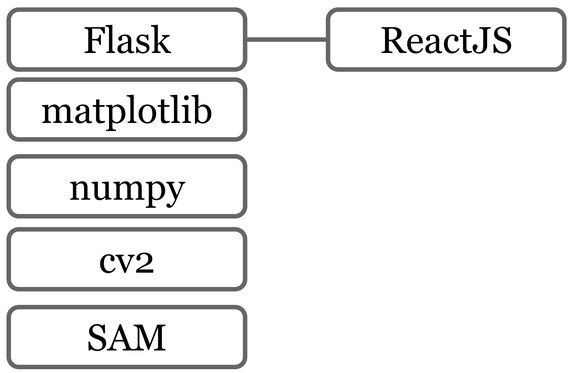


Figure 1: Tech stack visualized

Project Details

Project Flow

- User moves to the main page
- User chooses their input method, and uploads their chosen media
- Media is sent to the Flask server, where it is then processed and fed to the SAM
- Resulting masks are overlaid with the input image, saved, and sent back
- The main page parses image data and displays it

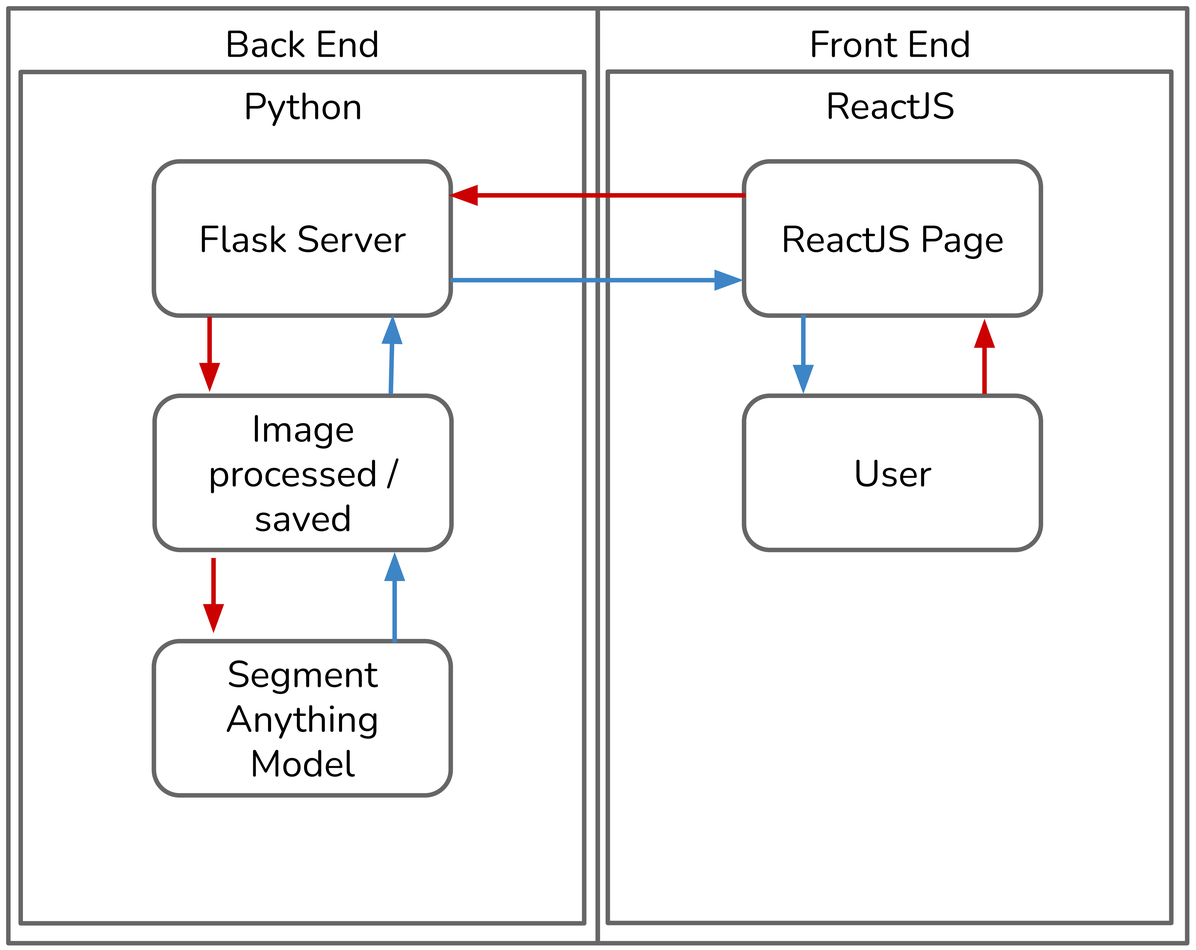
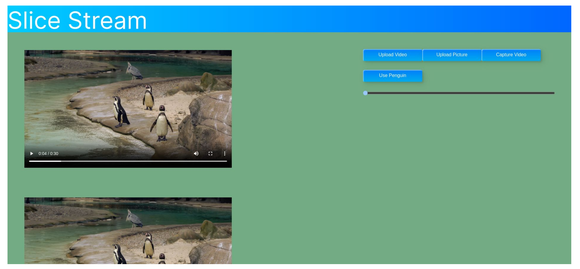


Figure 2: Diagram of project flow. User input is indicated in red, and application output is indicated in blue.

Front End

We used basic react to form all front end components and mostly implemented JSX (JavaScript XML)



We have a few features:

- ➢ Capture - takes the current frame in the video to be used to apply masks to the current video
- ➢ Upload Video - allows the user to upload any video from their local desktop and track those videos
- ➢ Upload Picture - allows the user to upload any image that from the video they want to apply masks to
- ➢ Slide Bar - moves the video frame by frame allowing the user to more easily capture the frame that they want.

Results

- While we were unable to fully create the entire website with object tracking, we were able to set the groundwork for the website to be completed.
- We have a working backend and a working front end and the only thing left to do is to clean up the front end look and connect the back end to the front end.
- While additional features would be nice, the general object tracking features are there.

Conclusion

- Ultimately we have yet to accomplish our goal in making a publicly accessible application. In order to accomplish this we need the following features
  - Host our project somewhere on the web
  - Connect front end data to backend calculations