



ELDIANS

# HACHATHON2023

## *VISIQNISTA:*

EMPOWERING SIGHT

ENABLING INDEPENDENCE

Team :

EL BAHÀ Ali

LOURARHI Yahya

ZAOUG Imad





# CONTENT

**01**

About Us

**02**

Challenges and  
Solution

**03**

Macro-planning

**04**

Version 1.0

**05**

Version 1.1

**06**

Limitation

**07**

Perspective



# ABOUT US



We are students at École Centrale Casablanca, and we seized the opportunity presented by the 2023 hackathon organized by MoroccoAI to introduce our innovative idea, VisonVista. Our project focuses on the creation of assistive glasses for individuals with visual impairments.



Recognizing that social needs are often overlooked by investors, particularly in areas where profit opportunities may not be immediately apparent, we embarked on the mission to address the unmet needs of the visually impaired community. Our vision with VisonVista is to not only create a groundbreaking solution but also to inspire increased investment in socially impactful ventures.





# CHALLENGES AND SOLUTION: INTRODUCING VISIONVISTA.

01

Blind individuals confront various challenges in their daily lives due to visual impairment. **Navigating unfamiliar environments** becomes challenging, and they face difficulties crossing streets and avoiding obstacles without visual cues.

02

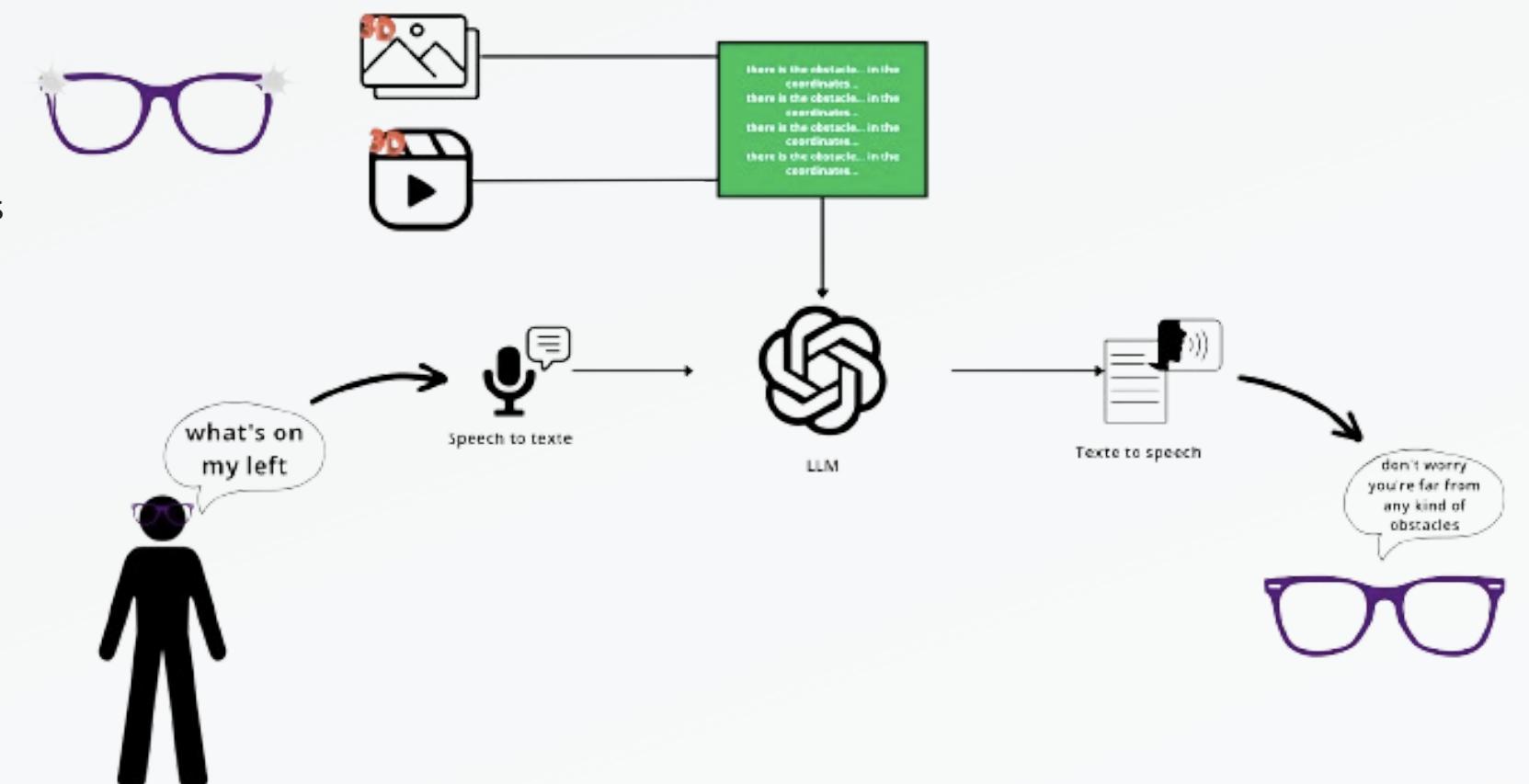
**Limited access** to printed materials and visual information in public spaces impacts their ability to obtain information. Communication hurdles arise in **non-verbal interactions**, including interpreting facial expressions and recognizing people without visual cues. **Social inclusion** may be compromised, leading to feelings of isolation and difficulties participating in visually reliant social activities.

03

**safety concerns** increase as blind individuals may struggle to identify potential hazards in their environment. Addressing these challenges is crucial for enhancing the overall well-being and inclusivity of individuals with visual impairments.

04

We introduce **VisionVista**, intelligent glasses designed to address this specific deficiency experienced by individuals with visual impairments.

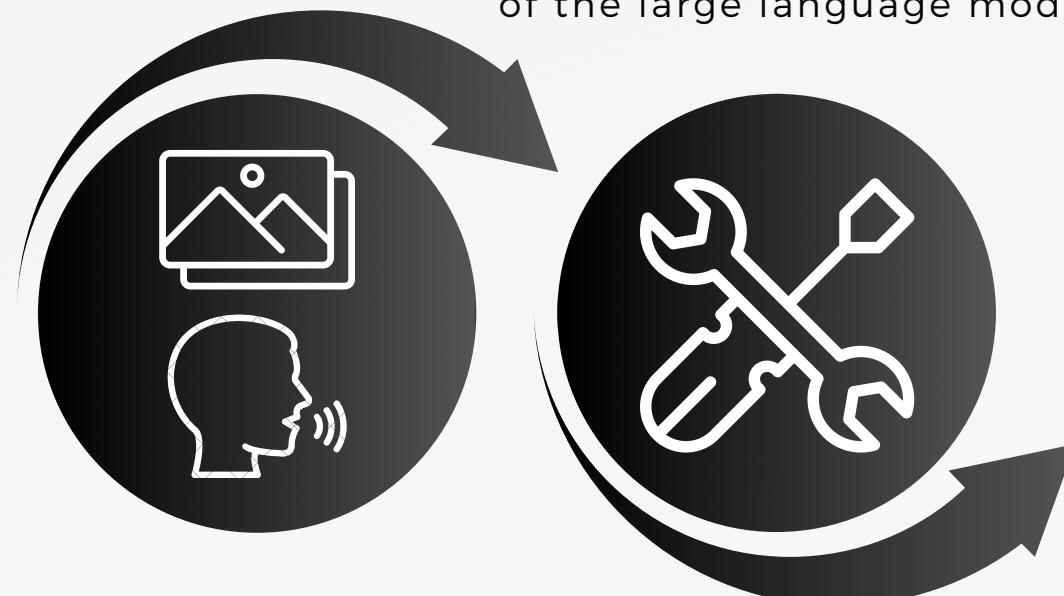




# MACRO-PLANNING

## January and February 2024

Focus on refining computer vision capabilities and optimizing the performance of the large language model.



## December 2023

Initiate the development of the first version enabling user interaction through images (V 1.0 and V 1.1).

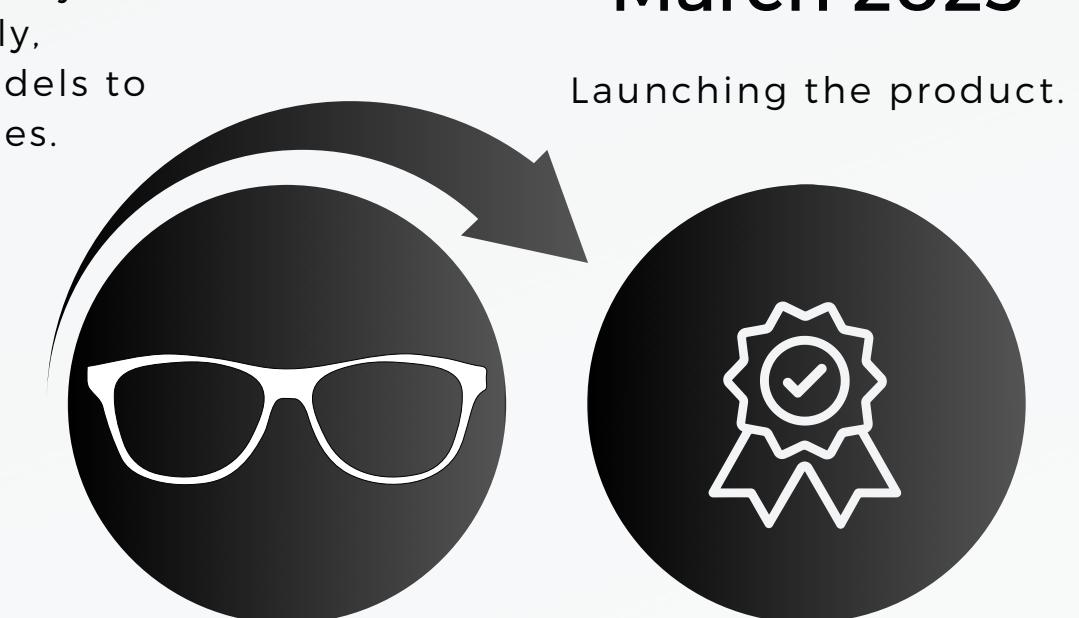
## April to August 2024

Continuously refine new features, improve the performance of the large language model, and address any identified issues. Additionally, incorporate pause estimation models to detect environmental activities.



## March 2024

Integrate video and OCR features, enhancing the system's ability to read text in images.



## March 2025

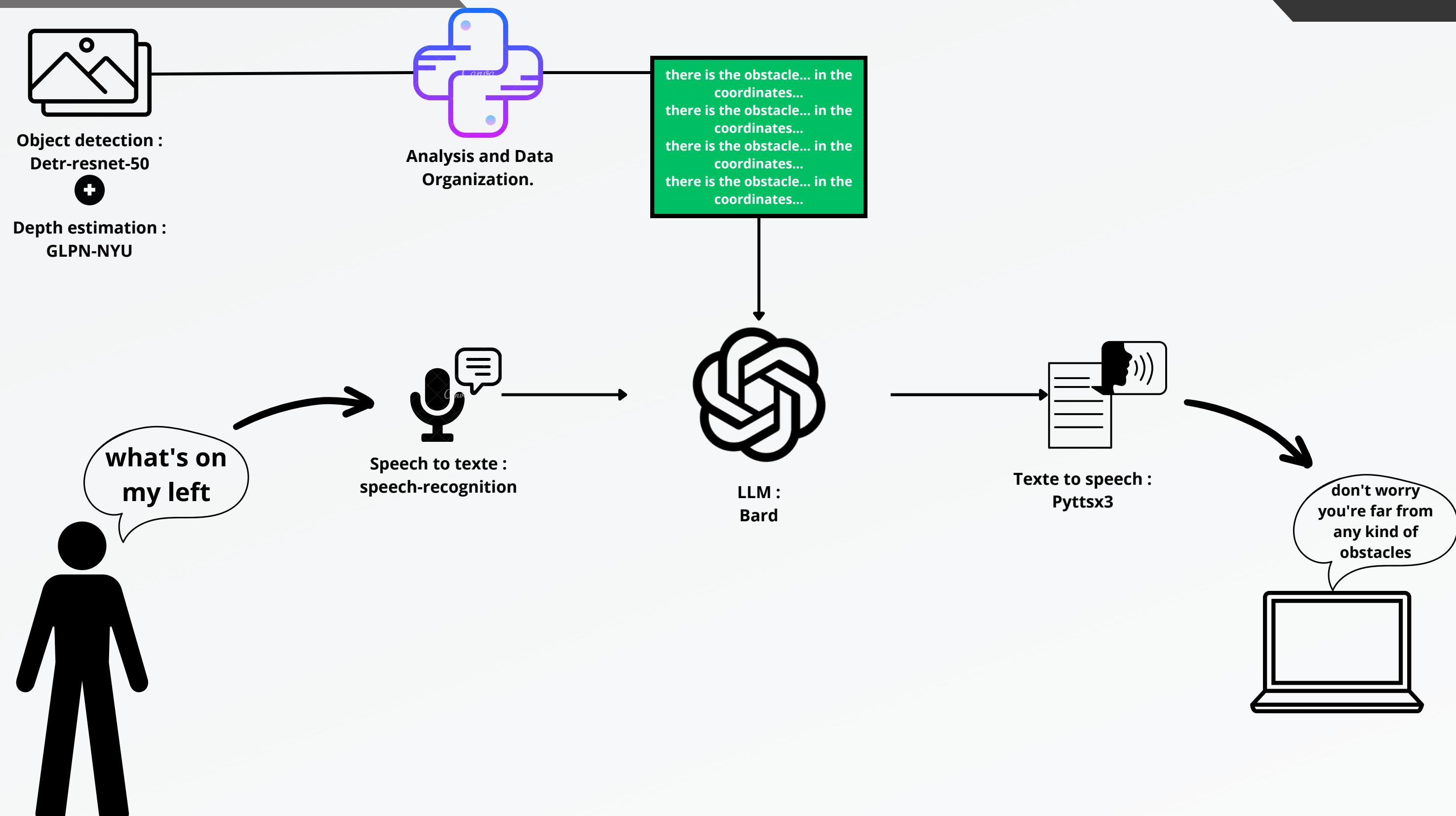
Launching the product.

## September to December 2024

Commence the creation of the initial VisionVista glasses while actively addressing and correcting any errors identified during the development process.

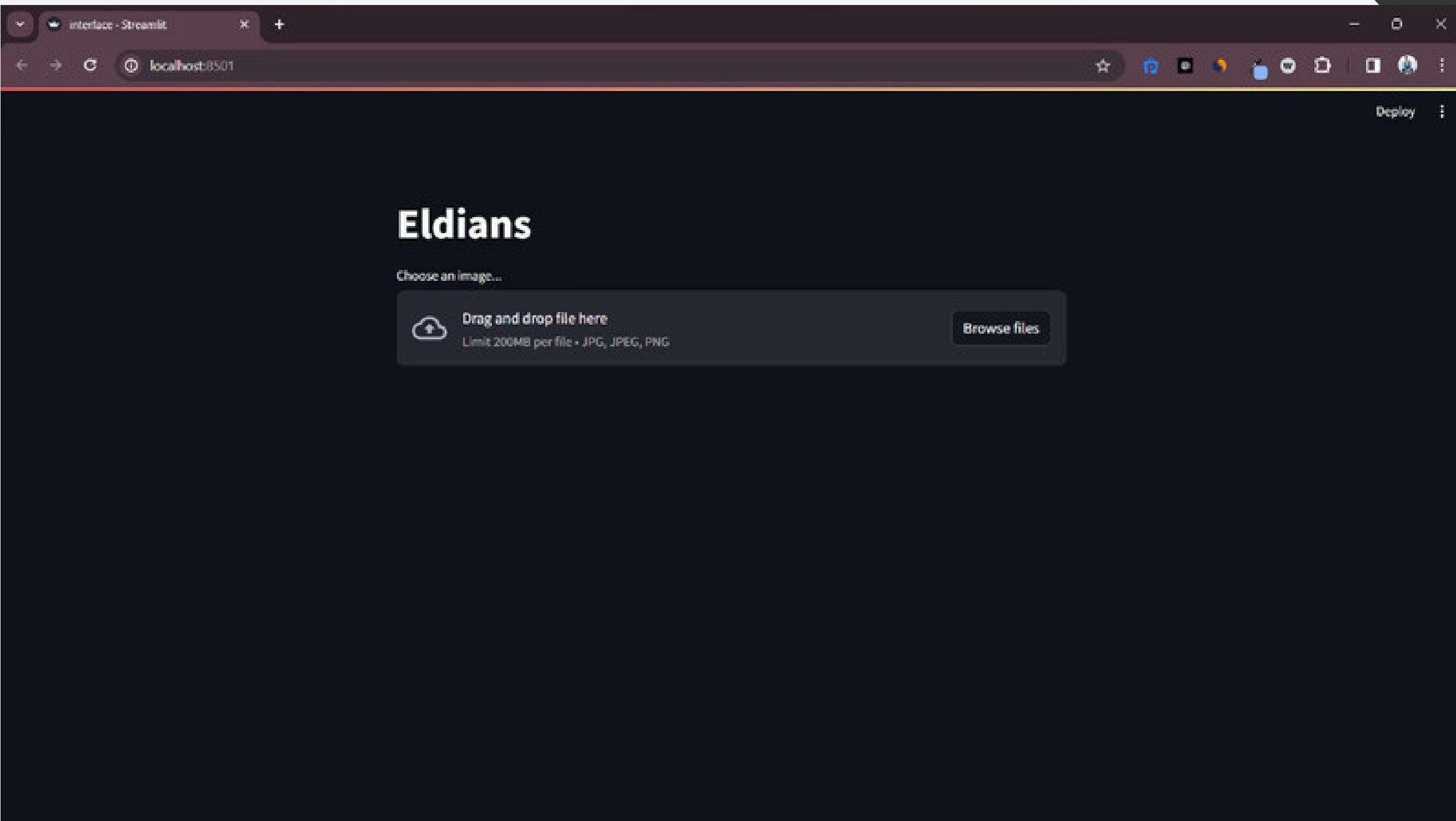


# VERSION 0.1 : OVERVIEW





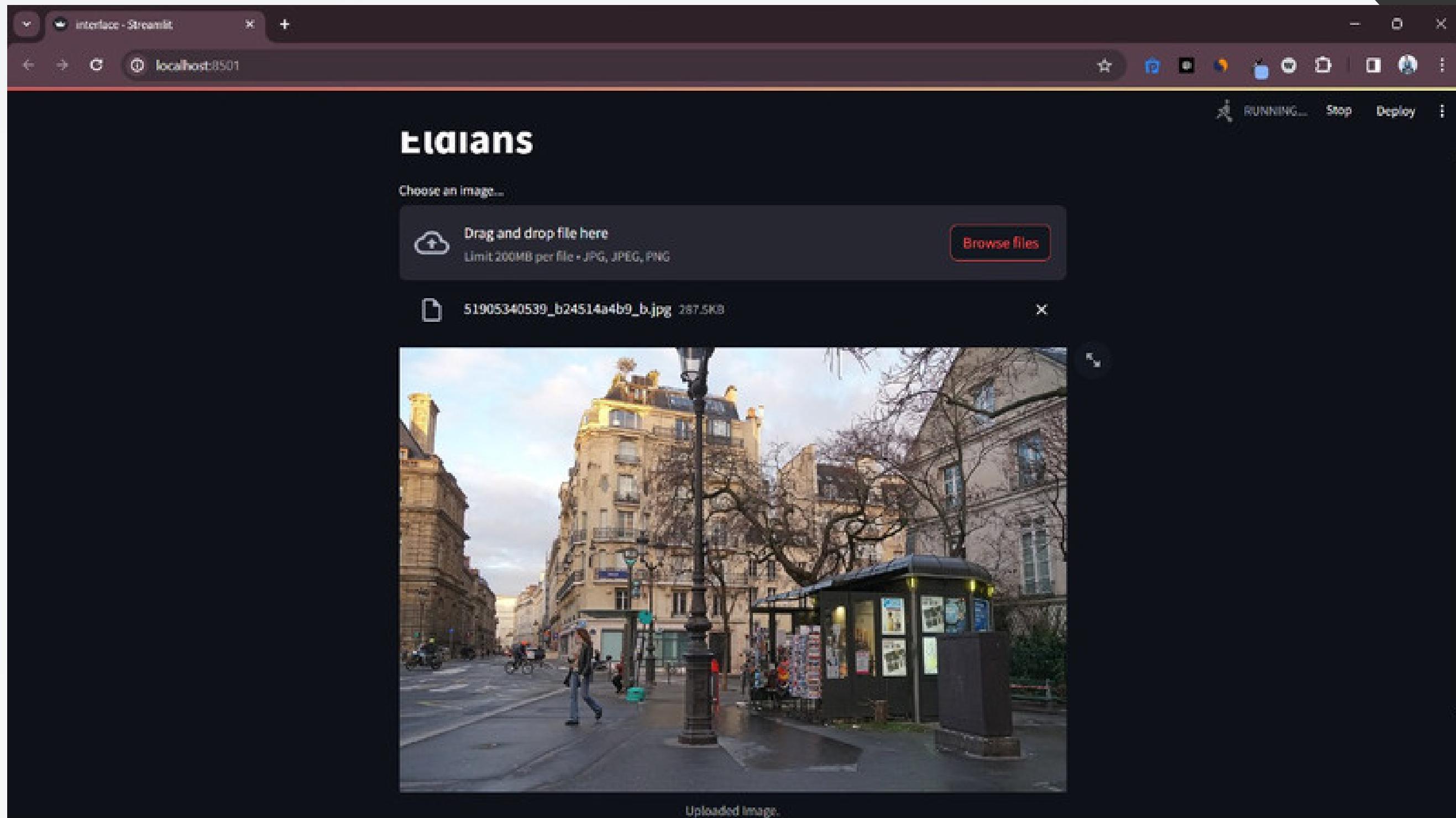
# VERSION 0.1 : TEST



Step 1



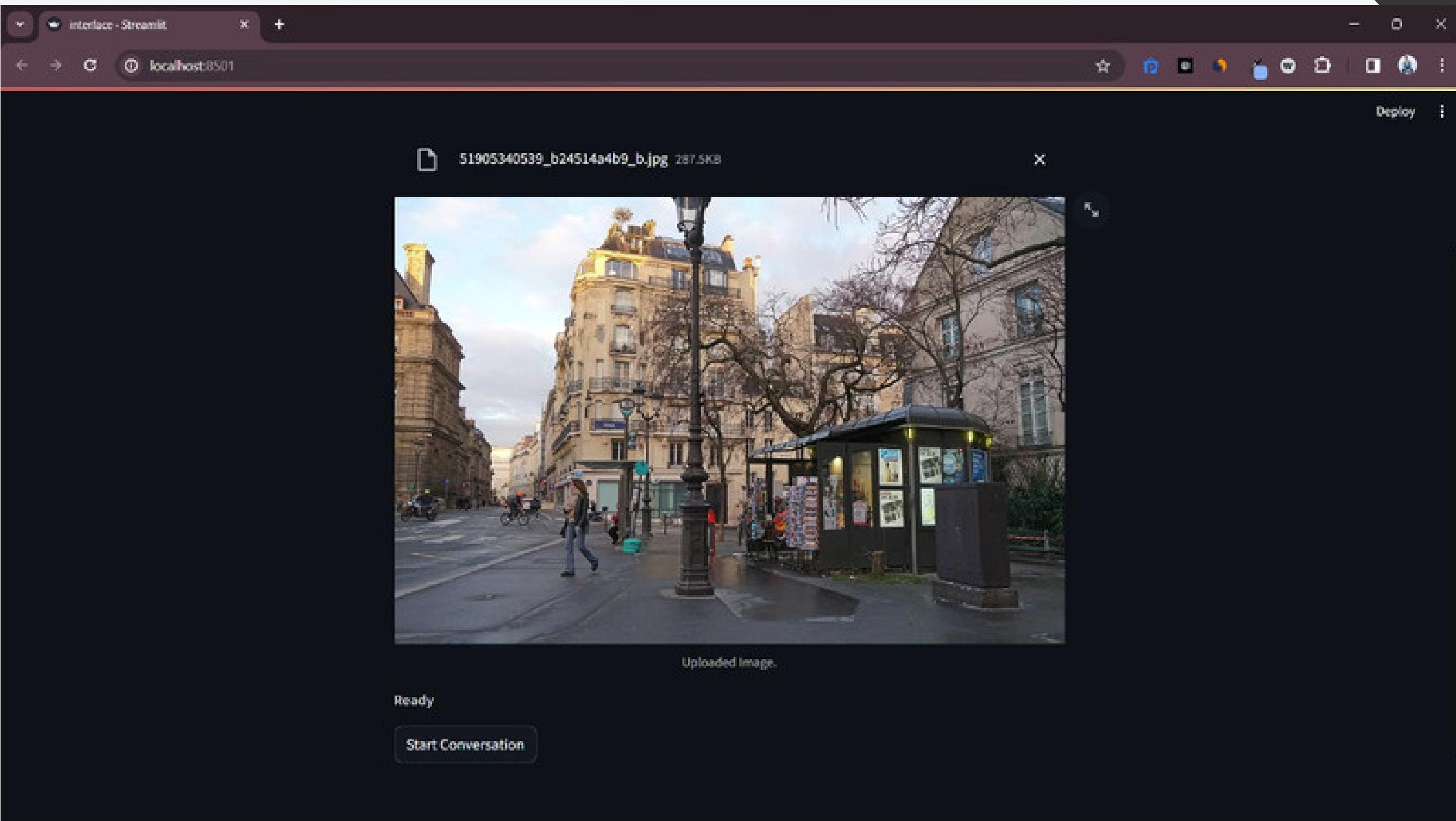
# VERSION 0.1 : TEST



Step 2



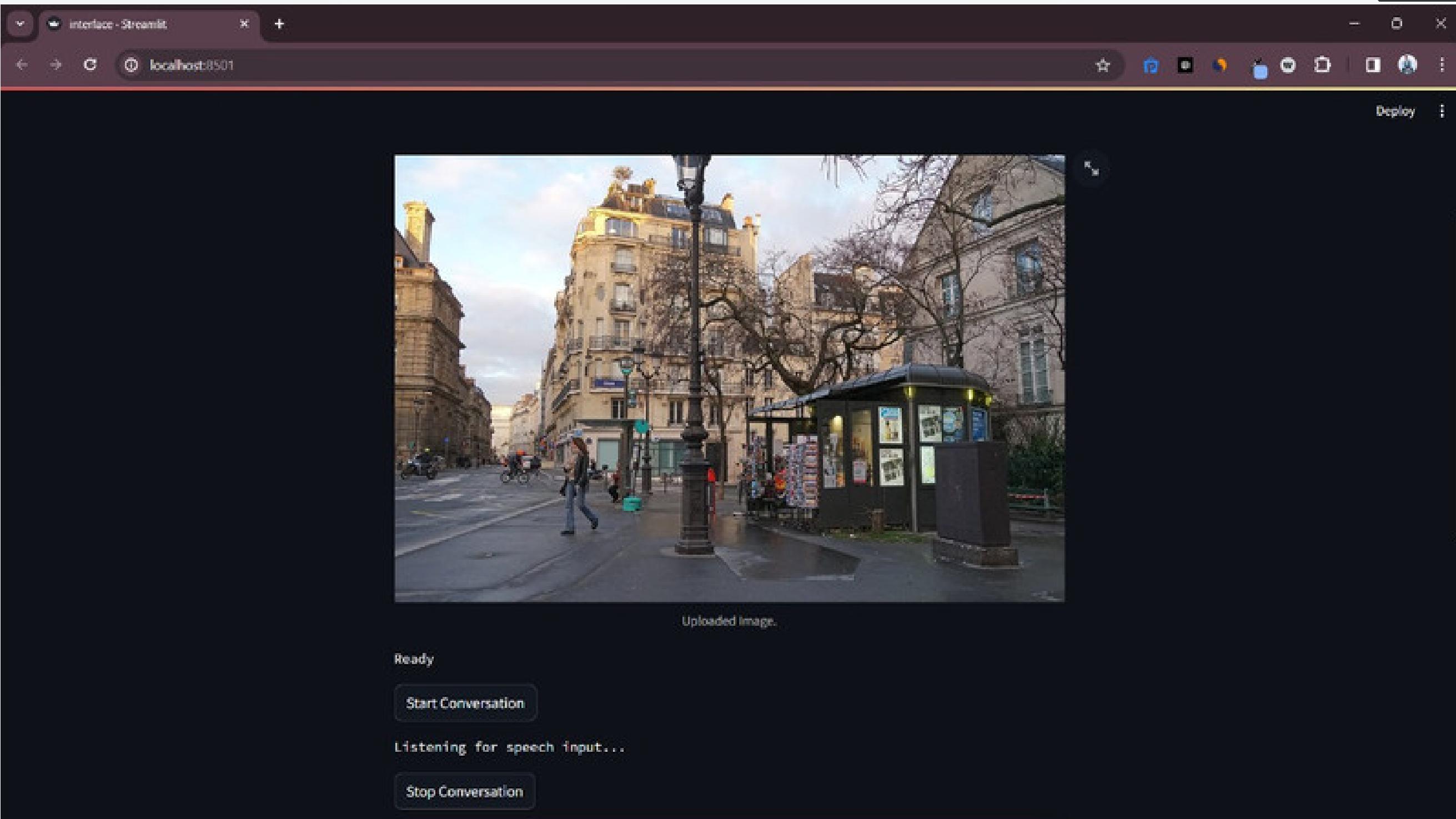
# VERSION 0.1 : TEST



Step 3



# VERSION 0.1 : TEST



Step 4



# VERSION 0.1 : TEST

Start Conversation

Listening for speech input...

Stop Conversation

Vocal Prompt: is there something near me

Response: Sure, here's the response based on the information you provided:

Response:

Yes, there are several people and objects around you. Here's a breakdown of what's nearby.

People:

There are two people with a high confidence level (0.996 and 0.975) standing in front of you.

There are two more people with a confidence level of 0.849 and 0.827 standing to your right.

In total, there are 9 people detected around you, with varying confidence levels.

Objects:

There are two motorcycles detected with confidence levels of 0.981 and 0.786 at a depth of 0.5m and 0.6m respectively.

There are three handbags detected with confidence levels of 0.899, 0.875, and 0.842 at a depth of 0.4m, 0.5m, and 0.6m respectively.

There is also a bicycle detected with a confidence level of 0.543 at a depth of 0.3m.

It's important to be cautious of your surroundings, especially with people and objects.

I hope this information is helpful! Let me know if you have any other questions.





# VERSION 0.1 : LIMITATIONS

- **Time Lag:**

The duration between analysis and the initial response falls within a range of 1 minute to 1 minute and 20 seconds.

- **Improvement of Analysis and Data Organization:**

Enhance the quality of analyses before incorporating them into the prompt.

- **Response Length:**

Responses are extensive and not concise, offering detailed engagement rather than brevity.

- **Memory Issues:**

The model tends to forget previous questions, resulting in a lack of continuity in the conversation flow.

- **Image Processing Quality:**

There is room for improvement in the detection quality of the image processing models.





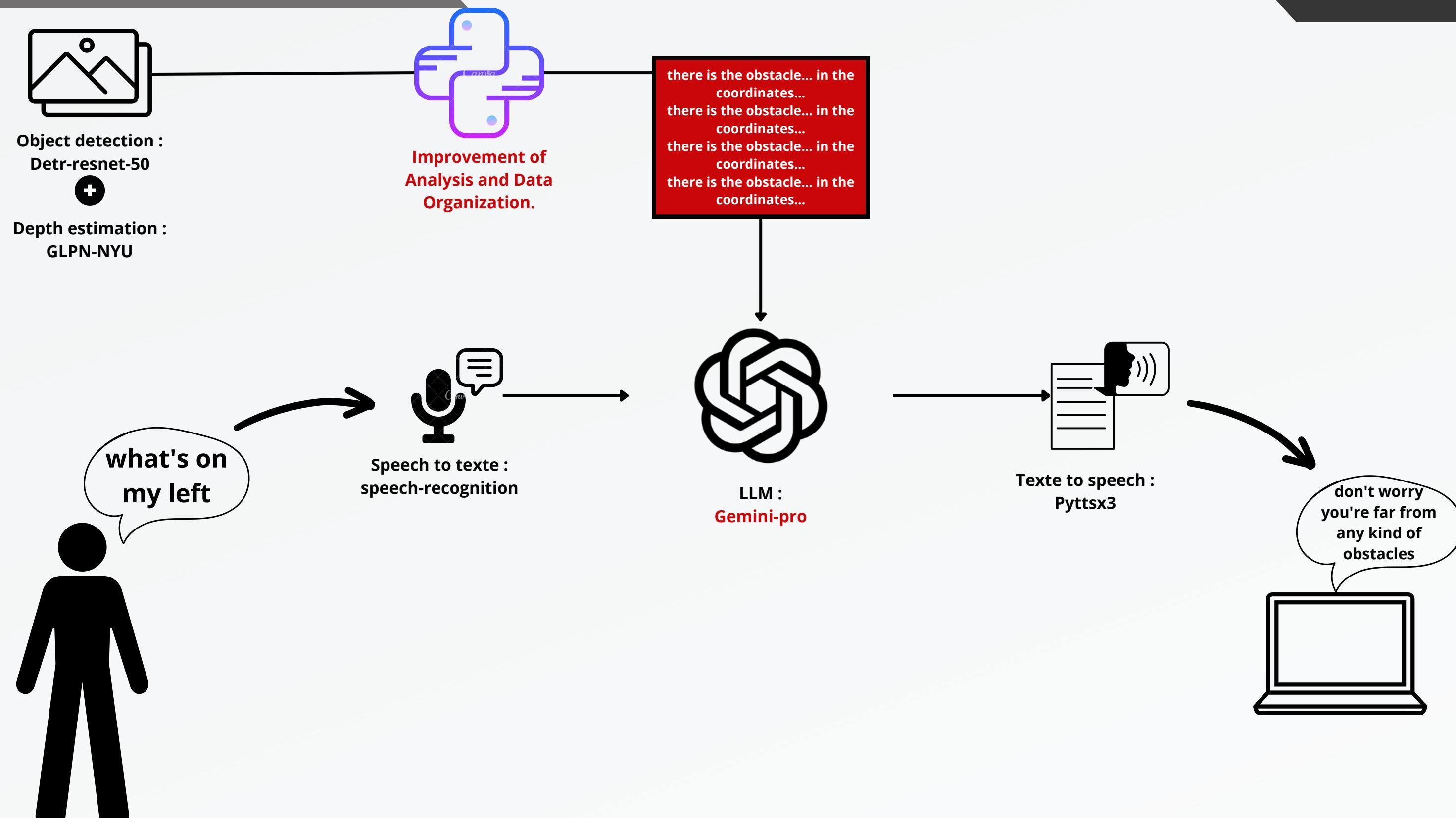
## VERSION 0.2

- ✓ • **Time Lag:**  
The duration between analysis and the initial response falls within a range of 1 minute to 1 minute and 20 seconds.
- ✓ • **Improvement of Analysis and Data Organization:**  
Enhance the quality of analyses before incorporating them into the prompt.
- ✓ • **Response Length:**  
Responses are extensive and not concise, offering detailed engagement rather than brevity.
- ✓ • **Memory Issues:**  
The model tends to forget previous questions, resulting in a lack of continuity in the conversation flow.
- ✗ • **Image Processing Quality:**  
There is room for improvement in the detection quality of the image processing models.





# VERSION 0.2 : OVERVIEW





## VERSION 0.2 : LIMITATIONS

- **Time Lag:**

The duration between analysis and the initial response falls within a range of 20 seconds to 30 seconds.

- **Improvement of Analysis and Data Organization:**

Enhance the quality of analyses before incorporating them into the prompt.

- **Prompt Quality:**

While we have improved the prompt, further efforts are required to achieve better results.

- **Image Processing Quality:**

There is room for improvement in the detection quality of the image processing models. And we should start thinking about the integration of the video.





# CONCLUSION AND PERSPECTIVES

## Initial Promise:

- Version 0.2 shows promise with smooth and contextual responses.



## Adaptation Required:

- In-depth work needed to align the model with the project.
- Crucial understanding of interactions of visually impaired individuals.
- Fine-tuning for continuous improvement.
- Exploration of new classes or transfer learning.

## Future Perspectives:

- Integration of real-time functionalities.
- Improvement in interpreting text within images.
- OCR for texts in images and videos



# OUR TEAM



## ZAOUG Imad

3rd year engineering student

+212 691-790850

<https://www.linkedin.com/in/imad-zaoug/>



## EL BAHÀ Ali

2nd year engineering student

+212 644-119780

<https://www.linkedin.com/in/ali-elbaha/>



## LOURARHI Yahya

2nd year engineering student

+212 618-346613

<https://www.linkedin.com/in/yahya-lourarhi/>





**ELDIANS**

