Design Review III - AR Object Detection and Text Recognition for Language Learning

Team LangLens

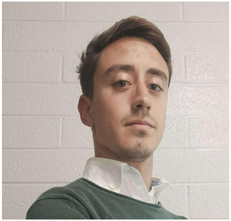
Team LangLens

Stefan Mihailovic
Team Lead

Brian Ruiz
Recorder

Sami Tanquary
Architect / Web Developer

Kyle Young
Customer Communicator



Daniel Navarrete
Release Manager

Italo Santos
Faculty Mentor

Our Client

Dr. Okim Kang

- Professor in the Department of English at Northern Arizona University
- Specialized knowledge in Applied Linguistics and Teaching English as a Second Language (TESL)
- Received many awards and honors for her contributions to language pedagogy research
- Worked on developing systems / apps dedicated to linguistics in the past

NORTHERN ARIZONA UNIVERSITY

Problem Statement

1

Lack of **free**, **accessible** tools for language learners that utilize **both** object detection and text recognition

2

Current learning applications **do not focus** on the key elements of word learning: ***meaning***, ***usage***, *and* ***form***

3

Current tools are **challenging** for those who struggle with text or are inexperienced in navigating a foreign language

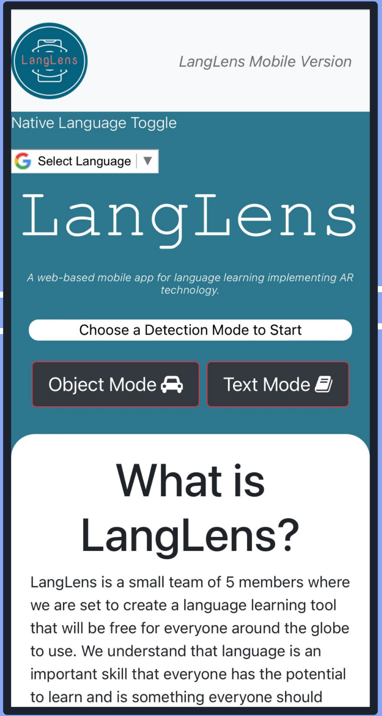
Our Solution

Free

Free-to-use and accessible to anyone

Mobile

Web-app optimized for mobile devices



LangLens Mobile Version

Native Language Toggle

Select Language ▼

LangLens

A web-based mobile app for language learning implementing AR technology.

Choose a Detection Mode to Start

Object Mode 🚘 Text Mode 📘

What is LangLens?

LangLens is a small team of 5 members where we are set to create a language learning tool that will be free for everyone around the globe to use. We understand that language is an important skill that everyone has the potential to learn and is something everyone should

Homepage GUI

Modern

Employs both object AND text recognition capabilities

Word Learning

Covers key aspects of word learning: *meaning, usage,* and *form*

Requirements Review

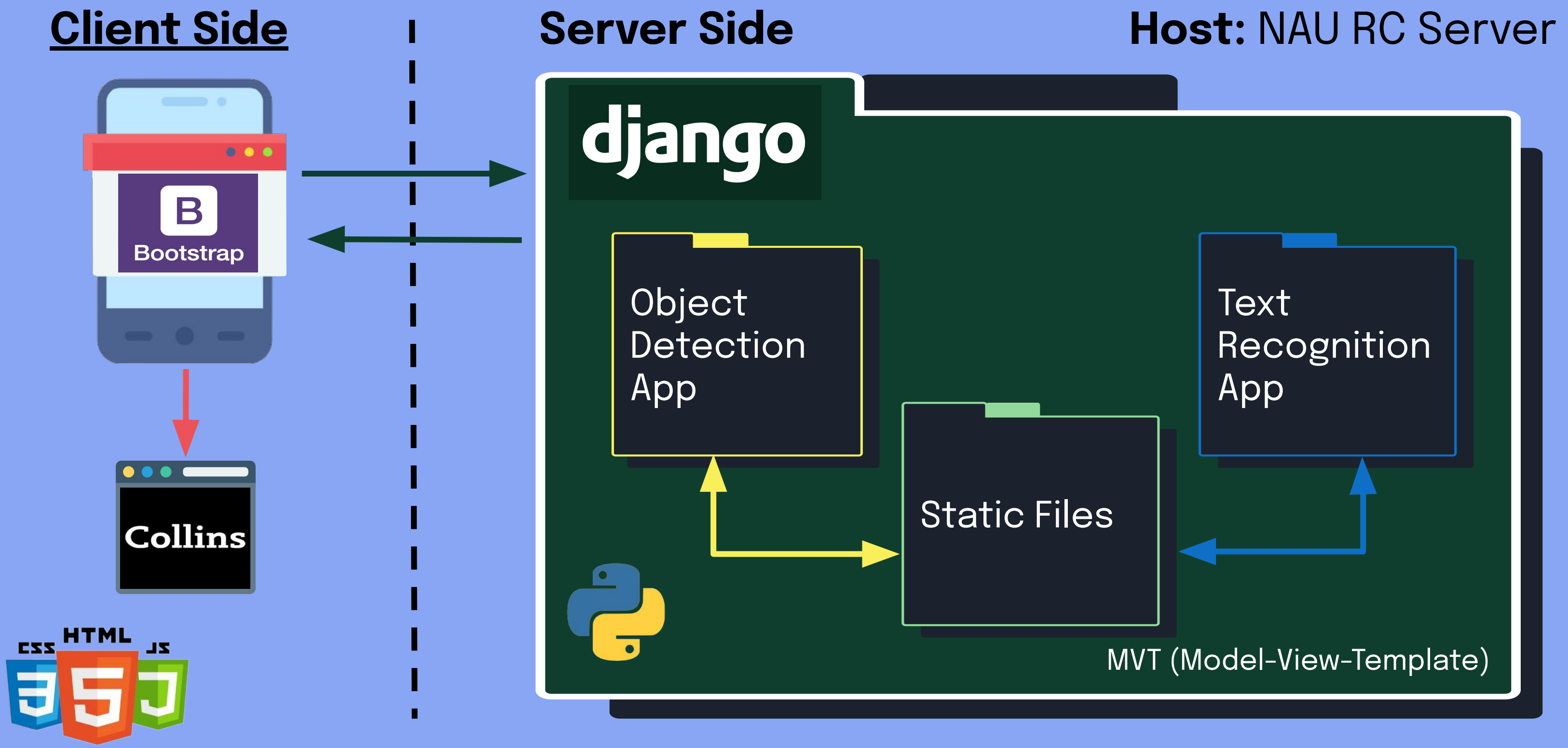
Key Requirements

- ❏ Toggle between object detection and text recognition modes
- ❏ Target language selector (MVP: Spanish, French, Korean)
- ❏ Scanning process can be restarted
- ❏ Scans environment in real time
- ❏ Displays a link to an external learning page after each scan

Other Requirements

- ❏ Web-app optimized for mobile devices
- ❏ Free software
- ❏ Ease of use
- ❏ Accurate and efficient

Architecture Overview



Implementation Overview

YOLOv5

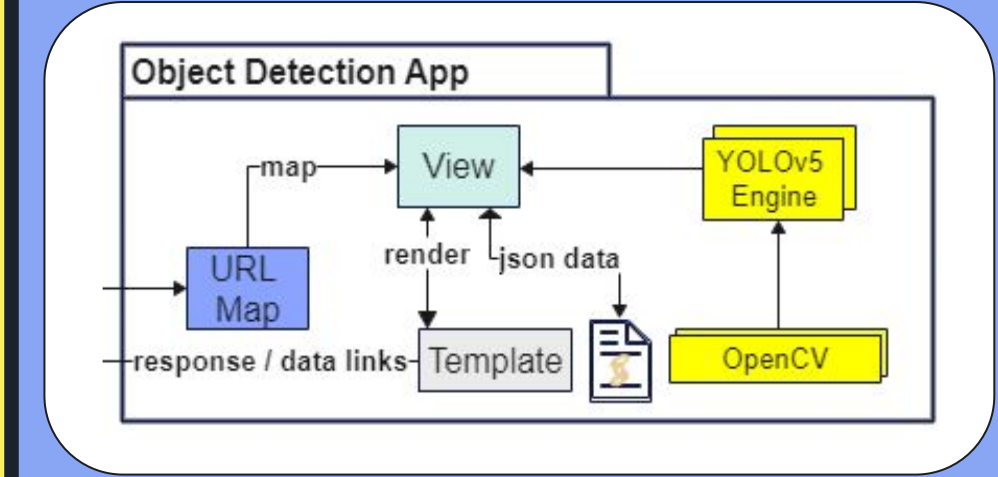
- ➔ Python-based object detection algorithms
- ➔ uses trainable object models
- ➔ returns object classes and confidences

OpenCV

- ➔ image preprocessing

Base64 Img → View → JSON Response

Object Detection Mode



Implemented with MVT design

Implementation Overview

Tesseract OCR

→ trained language data models

→ returns text array w/ coords

GoogleTrans API

→ translate text to target language

OpenCV

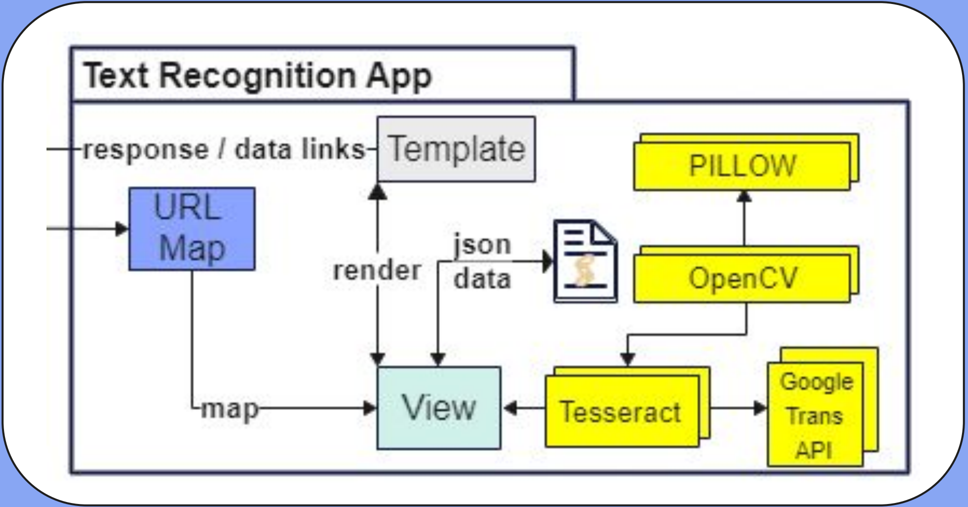
→ image preprocessing

PILLOW

→ write non-ascii text on image

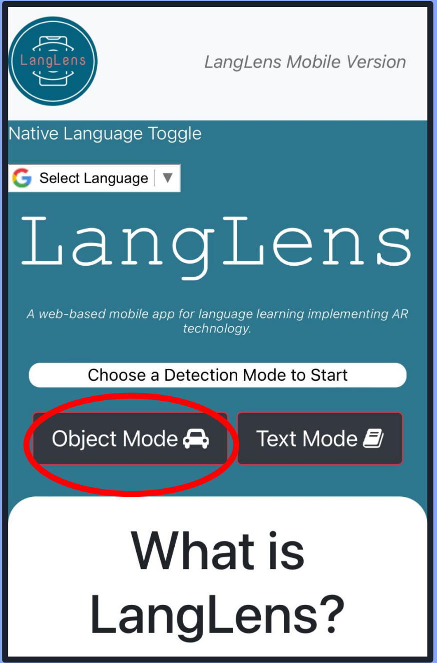
Base64 Img → View → JSON Response

Text Recognition Mode



Implemented with MVT design

Prototype Review



LangLens Mobile Version

Native Language Toggle

Select Language ▼

LangLens

A web-based mobile app for language learning implementing AR technology.

Choose a Detection Mode to Start

Object Mode

Text Mode

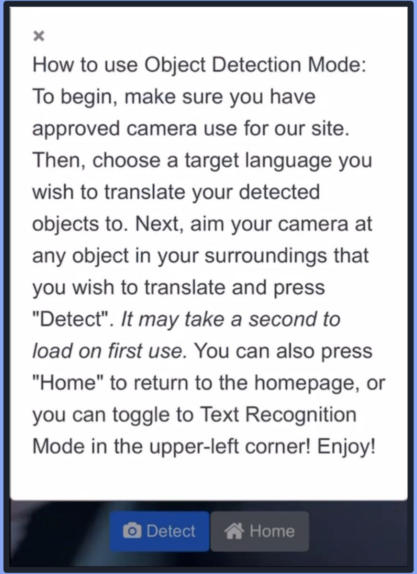
What is LangLens?

Homepage GUI

Click

Object Mode

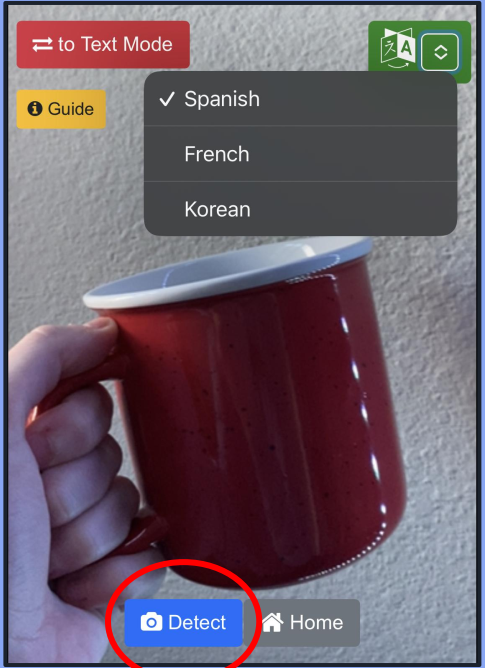
Guide



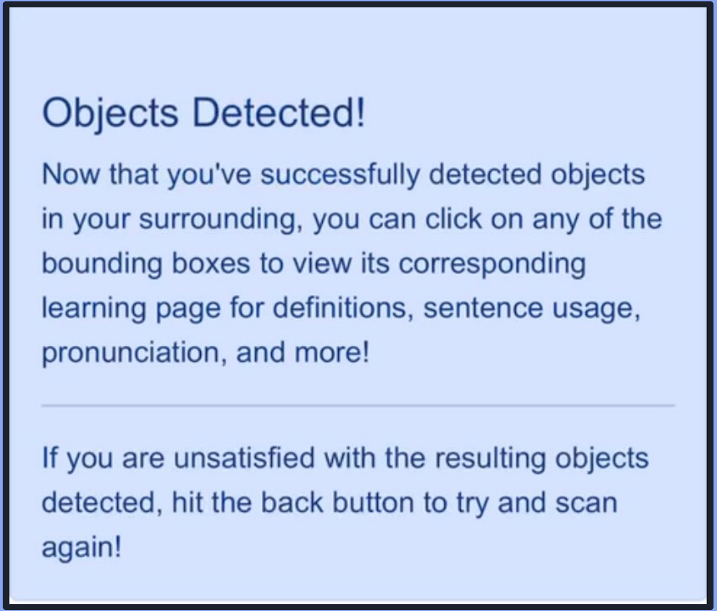
How to use Object Detection Mode: To begin, make sure you have approved camera use for our site. Then, choose a target language you wish to translate your detected objects to. Next, aim your camera at any object in your surroundings that you wish to translate and press "Detect". *It may take a second to load on first use.* You can also press "Home" to return to the homepage, or you can toggle to Text Recognition Mode in the upper-left corner! Enjoy!

Guide (Object Mode)

Prototype Review



GUI (Object Mode)



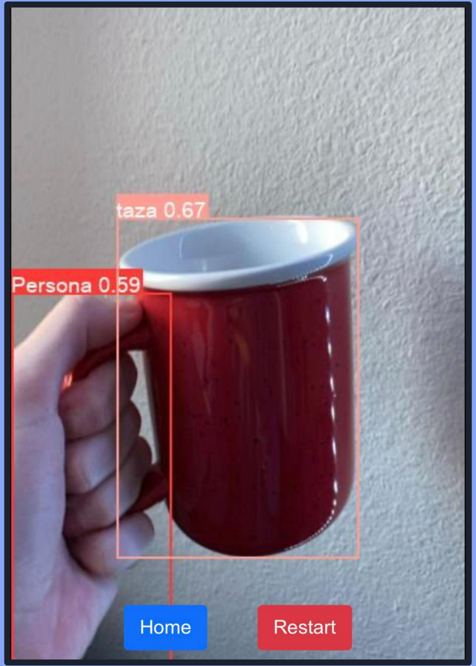
Objects Detected!

Now that you've successfully detected objects in your surrounding, you can click on any of the bounding boxes to view its corresponding learning page for definitions, sentence usage, pronunciation, and more!

If you are unsatisfied with the resulting objects detected, hit the back button to try and scan again!

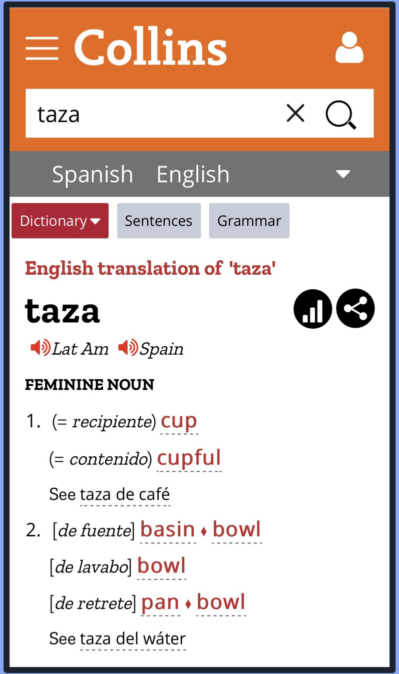
Successful Detection Popup (Objects)

Prototype Review



Result Image (Object Mode)

Click bounding box



English translation of 'taza'

taza

🔊*Lat Am* 🔊*Spain*

FEMININE NOUN

1. (= *recipiente*) cup

   (= *contenido*) cupful

   See taza de café

2. [*de fuente*] basin ♦ bowl

   [*de lavabo*] bowl

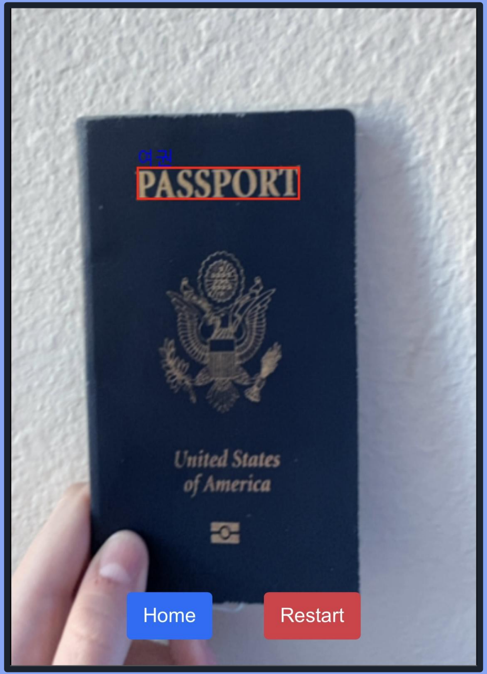
   [*de retrete*] pan ♦ bowl

   See taza del wáter

Redirect to Learning Page (Collins Dictionary)

Search Query parameter set to “ **taza** ” *when redirected*

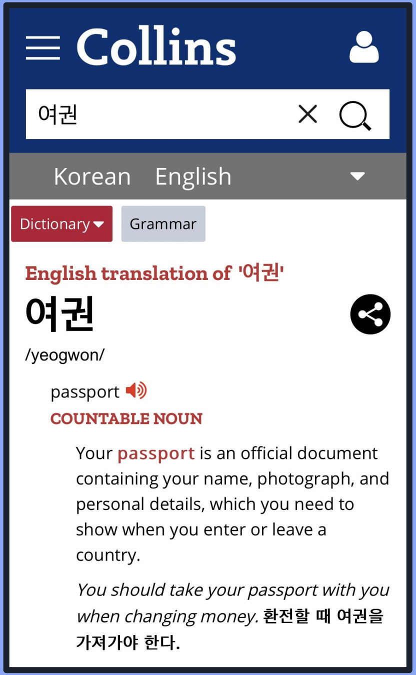
Prototype Review



Result Image (Text Mode)

Click

bounding box



여권

English translation of '여권'

여권

/yeogwon/

passport

COUNTABLE NOUN

Your passport is an official document containing your name, photograph, and personal details, which you need to show when you enter or leave a country.

You should take your passport with you when changing money. **환전할 때 여권을 가져가야 한다.**

Redirect to Learning Page (Collins Dictionary)

Search Query parameter set to " 여권 " when redirected

Challenges and Resolutions

Challenge	Resolution
Low Performance running YOLOv5 / OCR in real time	Instead of running it in a real time video, scan a picture taken of the user's current environment
OpenCV can't display non-ascii characters	Convert OpenCV frame to PIL image and use imported TrueType fonts for non-ascii supported languages
Way to send user to learning page without hyperlinks	Create clickable bounding boxes around each of the objects/words that have been scanned that redirect to the external learning page.

Testing Plan

Unit Testing

Focused on certain parts of the application or components of code

Integration Testing

See if the actions that we want the application to perform are running smoothly

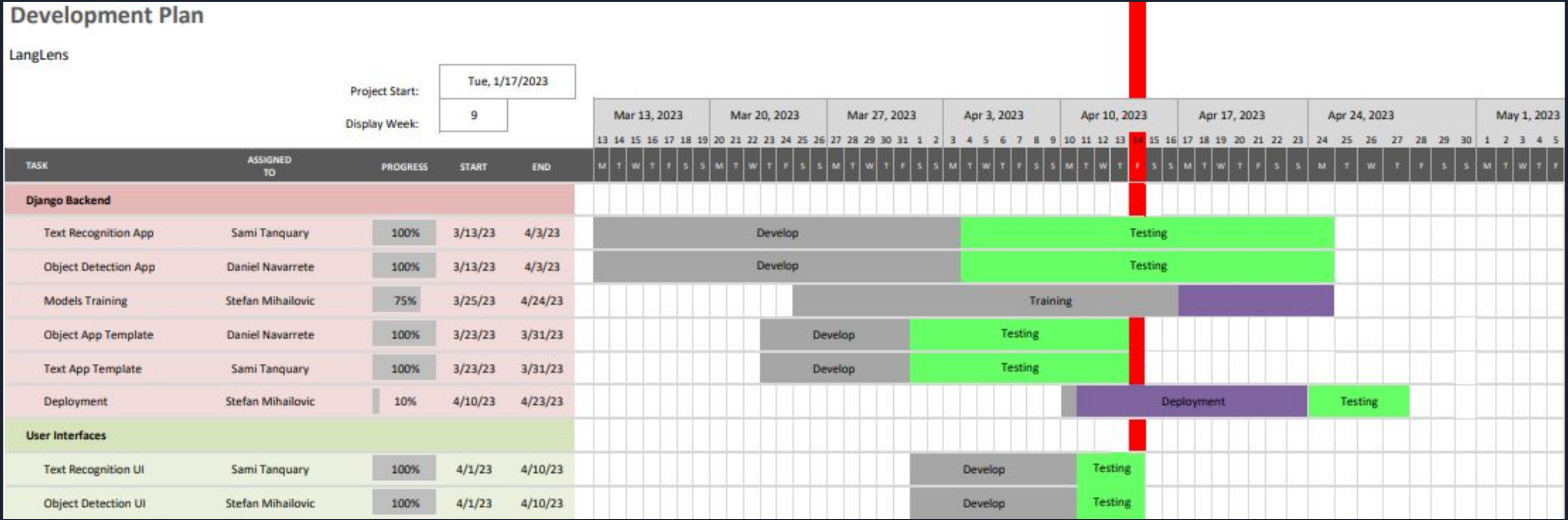
Usability Testing

Make sure that the application is intuitive and easy to use.

Testing Plan

Kind of tests	Approach
Unit Tests	Test Ajax requests for the main backend functions checking if the returned response is the expected one
Integration Tests	Test the expected and unexpected cases of use for the camera, language, object detection, text recognition and learning page integration
Usability Tests	Ensure that the software is intuitive and easy to use for the end-users. Participants will test the usability for the main menu, text recognition/object detection modes, and learning page redirection

Schedule



Development Plan

LangLens

Project Start:	Tue, 1/17/2023
Display Week:	9

TASK	ASSIGNED TO	PROGRESS	START	END
Django Backend				
Text Recognition App	Sami Tanquary	100%	3/13/23	4/3/23
Object Detection App	Daniel Navarrete	100%	3/13/23	4/3/23
Models Training	Stefan Mihailovic	75%	3/25/23	4/24/23
Object App Template	Daniel Navarrete	100%	3/23/23	3/31/23
Text App Template	Sami Tanquary	100%	3/23/23	3/31/23
Deployment	Stefan Mihailovic	10%	4/10/23	4/23/23
User Interfaces				
Text Recognition UI	Sami Tanquary	100%	4/1/23	4/10/23
Object Detection UI	Stefan Mihailovic	100%	4/1/23	4/10/23

🟪 = In Progress

⬜ = Complete

🟩 = Testing

Conclusion

Problem:

- lack of accessible AR language learning tools
- challenging for novice learners
- no focus on meaning, use, and form

Solution:

- free, easy to use, mobile, web app
- object and text recognition features
- focus on key elements of word learning

Future Plans:

- Improve learning models
- Deploy to server
- Finish testing

Current Status:

- On-track
- Dr. Okim is very happy with progress

Thank you!

Questions?