

# User Guide

## GazeDetectionApp

October 18, 2024

# Table of Contents

## Contents

<b>1 Purpose and Function of the GazeDetectionApp</b>	<b>2</b>
<b>2 General Instructions for Using the Application</b>	<b>2</b>
<b>3 Installing and Opening the Application</b>	<b>2</b>
3.1 Installing the Application . . . . .	2
3.2 Opening the Application . . . . .	2
<b>4 User Manual</b>	<b>3</b>
4.1 Collect Data . . . . .	4
4.2 Prepare Data . . . . .	11
4.3 Create Video . . . . .	15
4.4 Compare Video- and Model Annotation . . . . .	20

## 1 Purpose and Function of the GazeDetectionApp

Annotating drivers' gaze direction is often still done manually. This is a time-consuming process that requires accuracy. Annotators review video recordings (of drivers) and manually note the driver's gaze direction at specific moments. This method is error-prone and inefficient. At the same time, it is crucial to monitor drivers' gaze direction, as distraction and improper mirror checks are significant causes of traffic accidents. By analyzing gaze behavior, insights are gained that contribute to improving road safety. The GazeDetectionApp automates the annotation process, allowing researchers to collect data more quickly and accurately. This supports the development of AI models, testing with annotated test videos, and the automatic annotation of live driver recordings. The application is designed so that technical (IT) expertise is not required to use the application. This allows researchers to focus on the research itself. The application features an intuitive interface and simple operation, enabling users without a technical background to use it easily.

## 2 General Instructions for Using the Application

The application is a tool to facilitate video recording. It is an application used in the "field," in this case, the car. The application and accompanying settings, such as camera positioning and recording setup, are executed on location.

## 3 Installing and Opening the Application

### 3.1 Installing the Application

For the installation of the GazeDetectionApp, refer to the readme.txt file in the GitHub Repository GazeDetectionApp. This file contains the detailed steps for installing and starting the application.

### 3.2 Opening the Application

To open the application on the desktop (after installation), open a command prompt and navigate to the folder where the application is installed. Then, execute the following command:

```
python GazeDetection.py
```

See Figure 1 for an example:

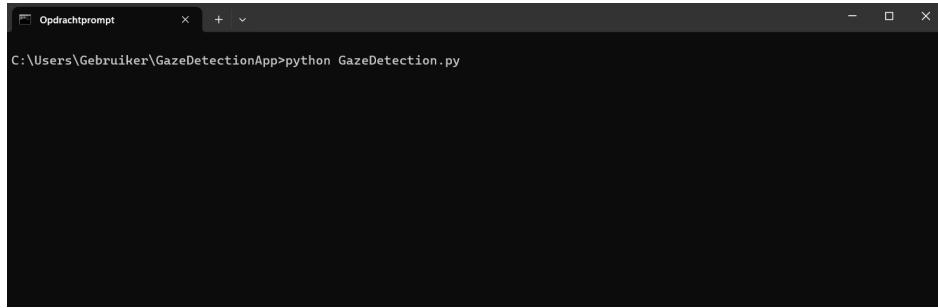


Figure 1: Start the application by running the command in the command prompt.

After entering the command 'python GazeDetection.py' and pressing <Enter>, the application's start screen will appear, as shown in Figure 2.

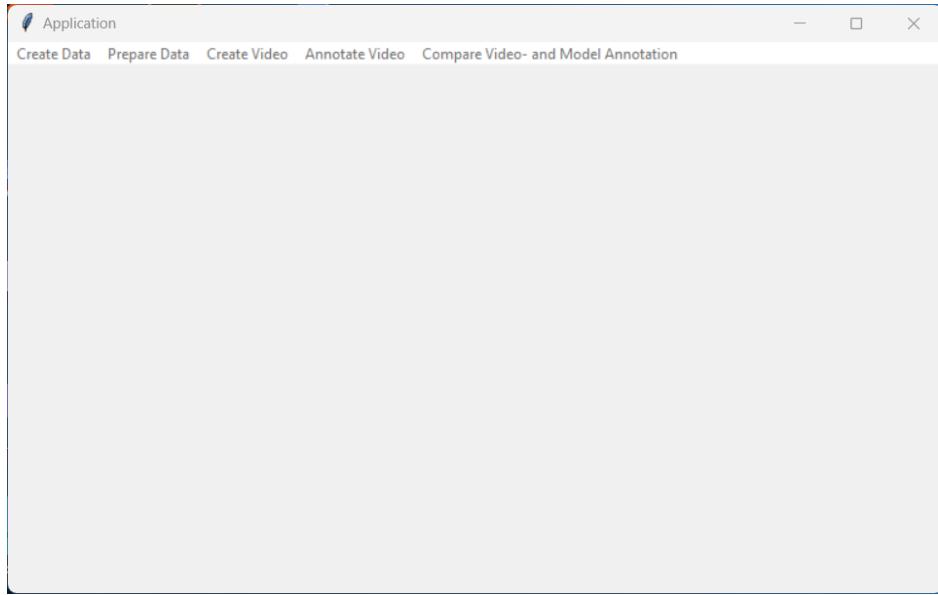


Figure 2: Application Start Screen

## 4 User Manual

The GazeDetectionApp includes five main functions, accessible via the menu bar at the top of the screen, as shown in Figure 2. A description of these functions follows:

- **Collect Data:** This function guides the process of recording video footage of a driver. These recordings are intended to train a model that will later be used to automatically annotate the driver's footage.
- **Prepare Data:** This function generates a photo collection based on adjustable variables from the video recordings. This collection is used to train, test, and validate the model.
- **Create Video:** This function allows the user to create an annotated (live) recording for comparison with the annotations generated by the model from the same recording.
- **Annotate Video:** This function annotates a video recording of a driver based on the trained model.
- **Compare Video and Model Annotation:** This function compares the annotation made by the model with the (live) video that has been annotated using user audio.

In this manual, all functions are explained step-by-step, so you know exactly how to use the GazeDetection-App.

## 4.1 Collect Data

Clicking on the "Create Data" menu option in the menu bar opens a submenu. See Figure 3.

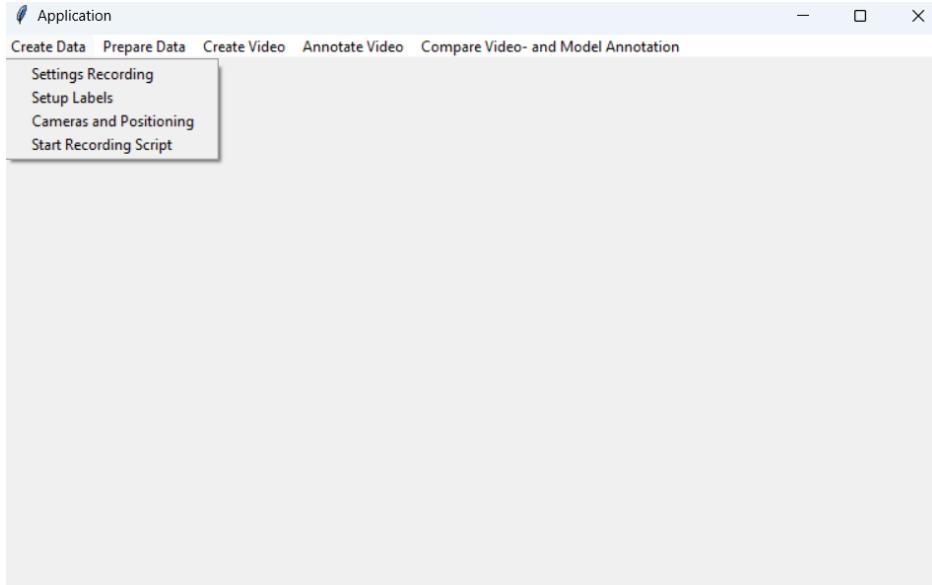


Figure 3: submenu of the "Create Data" menu option

The submenu options "Settings" and "Setup Labels" can be configured from a workstation. The options "Cameras and Positioning" and "Start Recording Script" must be executed in the car. For this, the laptop needs to be placed on the passenger seat, and the cameras - connected to the laptop via USB - need to be positioned correctly in the car. The first camera should be above the rearview mirror, and the second camera should be on the dashboard, aligned with the middle of the passenger seat. Both cameras should be set up so that they are maximally zoomed in on the face while keeping the entire face in the frame. A tool for adjusting the cameras - specifically for the Logitech C920 cameras used - can be the application provided with the webcams. This allows for detailed panning, as well as zooming in and out if needed.

**Settings** The first step is to configure some general settings necessary for collecting and storing video recordings. When selecting the submenu item "Settings Recording," the screen shown in Figure 4 will open.

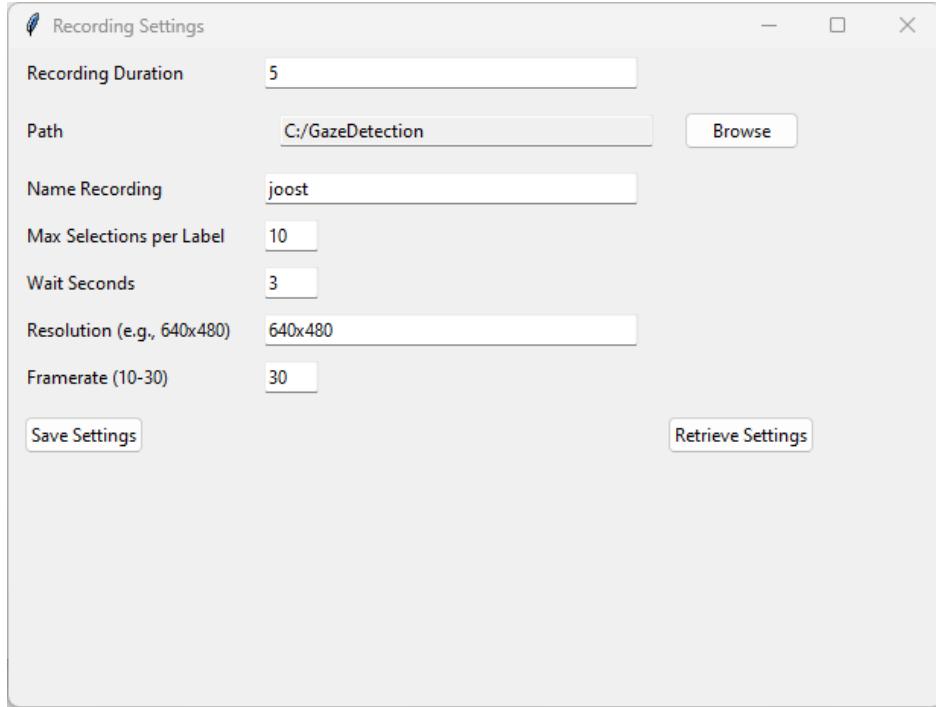


Figure 4: Submenu "Recording Settings"

The following variables can be configured here:

- **Recording Duration:** Enter the duration of each video recording (in seconds).
- **Path:** Select the folder where the recordings should be saved by clicking the "Browse" button.
- **Name Recording:** Provide a unique name for the recording session.
- **Max Selections per Label:** Enter the maximum number of recordings that should be made per label during this session.
- **Wait Seconds:** Enter the number of seconds the application should wait between consecutive recordings.
- **Resolution (e.g., 640x480):** Enter the resolution at which the recordings should be made.
- **Framerate (10-30):** Enter the number of frames per second for the recordings.

The following actions can be performed:

- **Save Settings:** Click this button to save the settings to a JSON file.
- **Retrieve Settings:** Click this button to load the most recently saved settings from the JSON file.

**Setup Labels** Before data can be collected, the labels (gaze directions) need to be set up. Selecting the submenu item "Setup Labels" opens the screen shown in Figure 5.

#	Label	Label (NL)	Label (EN)	Gaze Direction	Sentence	Dutch Translation	Use (Yes/No)
1	forward	Vooruit	Forward	Forward	Look forward.	Kijk vooruit.	<input checked="" type="checkbox"/>
2	left	Links	Left	Left	Look to the left.	Kijk naar links.	<input checked="" type="checkbox"/>
3	right	Rechts	Right	Right	Look to the right.	Kijk naar rechts.	<input checked="" type="checkbox"/>
4	mirror_interior	Binnenspiegel	Interior Mirror	Interior Mirror	Look at the interior mirror.	Kijk in de binnenspiegel.	<input checked="" type="checkbox"/>
5	mirror_right	Rechter Buitenspiegel	Right Side Mirror	Right Side Mirror	Look at the right side mirror.	Kijk in de rechter zijspiegel.	<input checked="" type="checkbox"/>
6	mirror_left	Linker Buitenspiegel	Left Side Mirror	Left Side Mirror	Look at the left side mirror.	Kijk in de linker zijspiegel.	<input checked="" type="checkbox"/>
7	shoulder_right	Rechterschouder	Right Shoulder	Right Shoulder	Look over your right shoulder.	Kijk over je rechter schouder.	<input checked="" type="checkbox"/>
8	shoulder_left	Linkerschouder	Left Shoulder	Left Shoulder	Look over your left shoulder.	Kijk over je linker schouder.	<input checked="" type="checkbox"/>
9	dashboard_straight_down	Dashboard	Dashboard	Dashboard Straight Down	Look straight down at the dashboard.	Kijk recht naar beneden.	<input checked="" type="checkbox"/>
10	dashboard_down_right	Middenconsole	Center Console	Dashboard Down Towards Center Console	Look down towards the center console.	Kijk naar beneden naar de middenconsole.	<input checked="" type="checkbox"/>
11	forward_right	Rechts vooruit	Forward Right	Forward Right	Look forward and to the right.	Kijk vooruit en naar rechts.	<input checked="" type="checkbox"/>
12	forward_left	Links vooruit	Forward Left	Forward Left	Look forward and to the left.	Kijk vooruit en naar links.	<input checked="" type="checkbox"/>

**Save Labels** **Retrieve Labels**

Figure 5: submenu "Setup Labels"

The following variables are configured for each gaze direction:

- **Label:** The name of the label.
- **Label (NL):** The Dutch translation of the label.
- **Label (EN):** The English translation of the label.
- **Gaze Direction:** The gaze direction associated with the label.
- **Sentence:** The instructional sentence that should be spoken.
- **Dutch Translation:** The Dutch translation of the instructional sentence.
- **Use (Yes/No):** Select whether this label should be used for this recording.

The following actions can be performed:

- **Save Labels:** Click this button to save the configured labels to a JSON file.
- **Retrieve Labels:** Click this button to load the most recently saved labels from the JSON file.

The relationship between the configured labels (gaze direction) and the actual direction in which the user looks must be established for consistency and reliability. Each defined gaze direction corresponds to a sticker in the car that indicates the center or midpoint of the gaze direction. Below is an example of how this can be set up. Figure 6 provides an overview of the gaze directions linked to a color.

Tabel: Labelcategorieën, Bijbehorende Kijkrichtingen en Instructiezinnes			
Label	Gaze Direction	Sentence	Color
forward	Forward	Look forward.	
left	Left	Look to the left.	
right	Right	Look to the right.	
mirror_interior	Interior Mirror	Look at the interior mirror.	
mirror_right	Right Side Mirror	Look at the right side mirror.	
mirror_left	Left Side Mirror	Look at the left side mirror.	
shoulder_right	Right Shoulder	Look over your right shoulder.	
shoulder_left	Left Shoulder	Look over your left shoulder.	
dashboard_straight_down	Dashboard Straight Down	Look straight down at the dashboard.	
dashboard_down_right	Dashboard Down Towards Center Console	Look down towards the center console.	
forward_right	Forward Right	Look forward and to the right.	
forward_left	Forward Left	Look forward and to the left.	

Figure 6: The labels (gaze directions) and the corresponding sticker

Figure 7 provides an impression of how the configuration of the gaze direction, combined with the sticker, indicates the gaze direction. In this case, the stickers for the right side mirror (Right Side Mirror) and right (Right) are visible.



Figure 7: The sticker indicating the gaze direction

**Cameras and Positioning** The next step is to detect and position the cameras. Detection involves the system identifying the connected cameras, while positioning involves correctly placing the cameras so that the driver is appropriately framed. Positioning is done on-site (i.e., in the car where the recordings will be made).

The laptop is placed on the passenger seat (possibly on an elevated surface), and the cameras are set up in the car. Place one camera above the rearview mirror, aimed at the driver's face, and the second camera on the dashboard at the level of the passenger seat, also aimed at the face. Ensure that both cameras are positioned so that the entire face remains visible, with sufficient detail to capture the gaze direction.

Selecting the submenu "Cameras and Positioning" opens the screen shown in Figure 8.

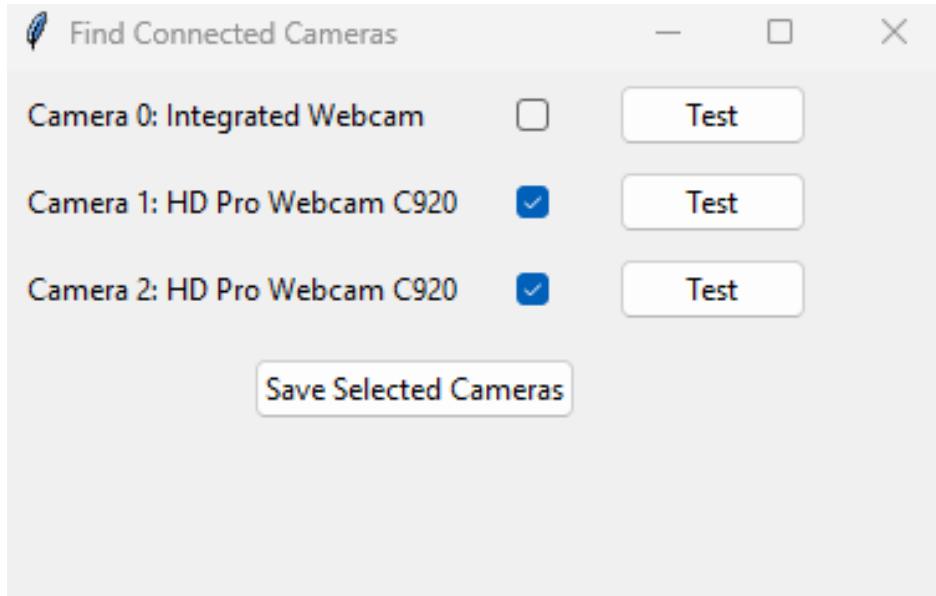


Figure 8: submenu "Cameras and Positioning"

The application automatically detects connected cameras and displays them in a list. The cameras are indexed from 0 to 10, where 0 is often the built-in webcam. In this application, various external cameras are used, connected to the laptop via USB ports.

- **Select Cameras:** Check the cameras you wish to use for the recordings.
- **Test:** Click the "Test" button next to each camera to view a live feed and position the camera correctly. When you click this button, the camera opens, and its output appears on the screen. You can adjust the camera's position by moving it or by using the standard configuration software for the webcams.<sup>1</sup> Note: It may take some time for the screen to become visible, due to the initialization of the available cameras.
- **Save Selected Cameras:** Click 'Save Selected Cameras' to save the current settings to a JSON file, so you can easily reload them later.

**Start Recording Script** After correctly setting the general settings, labels, and cameras, you can begin recording. Selecting the submenu "Start Recording Script" opens the screen shown in Figure 9.

<sup>1</sup>For the cameras used (Logitech C920 Pro), there is an app called "Logitech Camera Settings," which can be found in the Windows start menu. This app allows you to zoom in and out and adjust the position of the cameras.

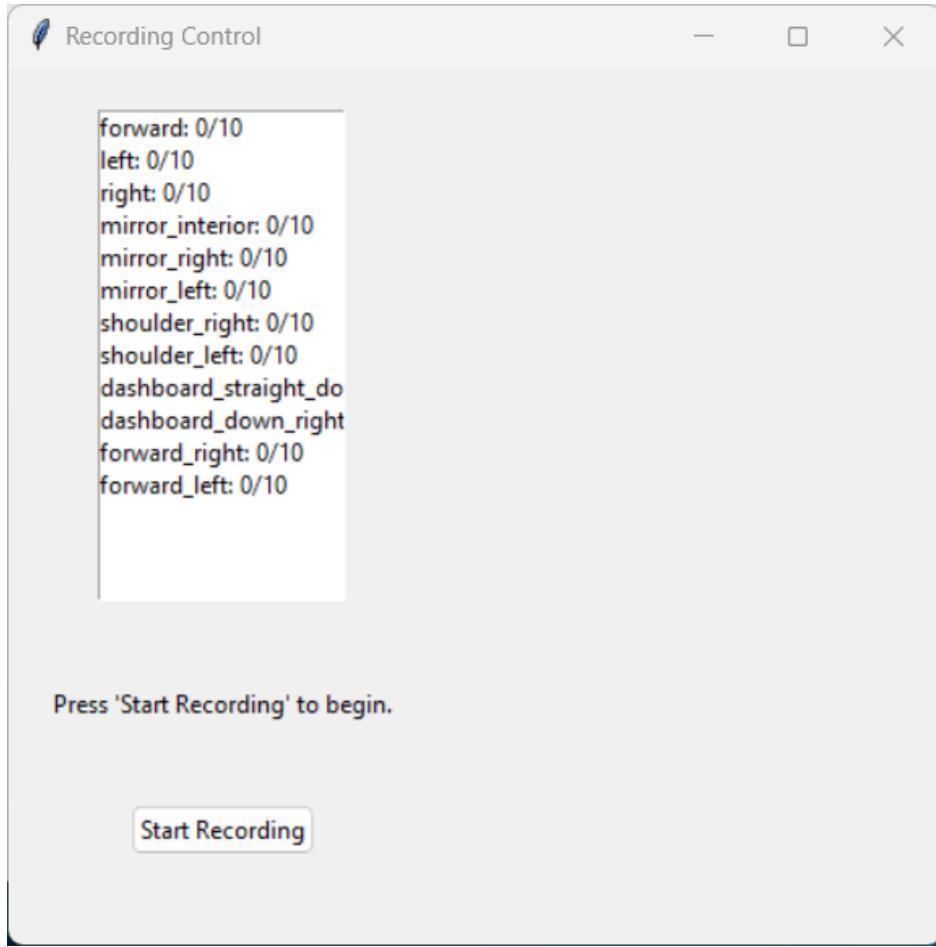


Figure 9: Recording Control

Note: It may take some time for the screen to load due to the initialization of the selected cameras. Until the loading is complete, the message "Loading, Please wait..." will be displayed, and the cursor will indicate that the system is processing.

Clicking the "Start Recording" button initiates the recording. The system will instruct you, via a spoken command, to look in a certain direction. For the label "forward," the phrase "Please look forward" will be used. You will have the configured number of seconds to look in that direction. The command "start" will then follow, and a recording will be made for the set number of seconds. During this time, you should continue looking in the specified direction. After each recording, a new direction will be instructed until all pre-configured directions have been recorded the required number of times. On the screen, you can see how many times a specific direction has been completed and how many times it is required. When the system has completed all directions - recording each direction the specified number of times - this will be indicated with a pop-up message and on the screen. You can now close the notification and the screen. The following diagram 10 outlines the directory structure used to organize the saved recordings. Each directory and file type is described in table 1 below.

```
root <Path / Name Recording>
|-- recording
|   |-- train
|   |   |-- label1
|   |   |   |-- 1
|   |   |   |   |-- camera_1.avi
|   |   |   |   |-- camera_2.avi
|   |   |   |-- 2
|   |   |   |   |-- camera_1.avi
|   |   |   |   |-- camera_2.avi
|   |   |   |-- ...
|   |   |-- label2
|   |   |   |-- 1
|   |   |   |   |-- camera_1.avi
|   |   |   |   |-- camera_2.avi
|   |   |   |-- 2
|   |   |   |   |-- camera_1.avi
|   |   |   |   |-- camera_2.avi
|   |   |   |-- ...
|   |-- test
|   |   |-- label1
|   |   |   |-- 1
|   |   |   |   |-- camera_1.avi
|   |   |   |   |-- camera_2.avi
|   |   |   |-- 2
|   |   |   |   |-- camera_1.avi
|   |   |   |   |-- camera_2.avi
|   |   |   |-- ...
|   |   |-- label2
|   |   |   |-- 1
|   |   |   |   |-- camera_1.avi
|   |   |   |   |-- camera_2.avi
|   |   |   |-- 2
|   |   |   |   |-- camera_1.avi
|   |   |   |   |-- camera_2.avi
|   |   |   |-- ...
|   |-- val
|   |   |-- label1
|   |   |   |-- 1
|   |   |   |   |-- camera_1.avi
|   |   |   |   |-- camera_2.avi
|   |   |   |-- 2
|   |   |   |   |-- camera_1.avi
|   |   |   |   |-- camera_2.avi
|   |   |   |-- ...
|   |   |-- label2
|   |   |   |-- 1
|   |   |   |   |-- camera_1.avi
|   |   |   |   |-- camera_2.avi
|   |   |   |-- 2
|   |   |   |   |-- camera_1.avi
|   |   |   |   |-- camera_2.avi
|   |   |   |-- ...
|   |-- readme.txt
```

Figure 10: Structure of the saved recordings

Directory	Description
Path	The root path where the recording files are saved.
Recording	Main directory for the recording
train	Directory for training data
val	Directory for validation data
test	Directory for test data
label<n>	Labels for categorization (e.g., label_forward, label_left)
sequence<n>	Number of times a label is repeated (e.g., 1, 2)
camera_x	Video files for each sequence with camera index (e.g., camera_1.avi)

Table 1: Directory structure

The laptop and cameras can now be removed from the car.

## 4.2 Prepare Data

This function can be executed from the workstation. When you click on the 'Prepare Data' menu option in the menu bar, a submenu appears with various options for preparing data for model training. See Figure 11.

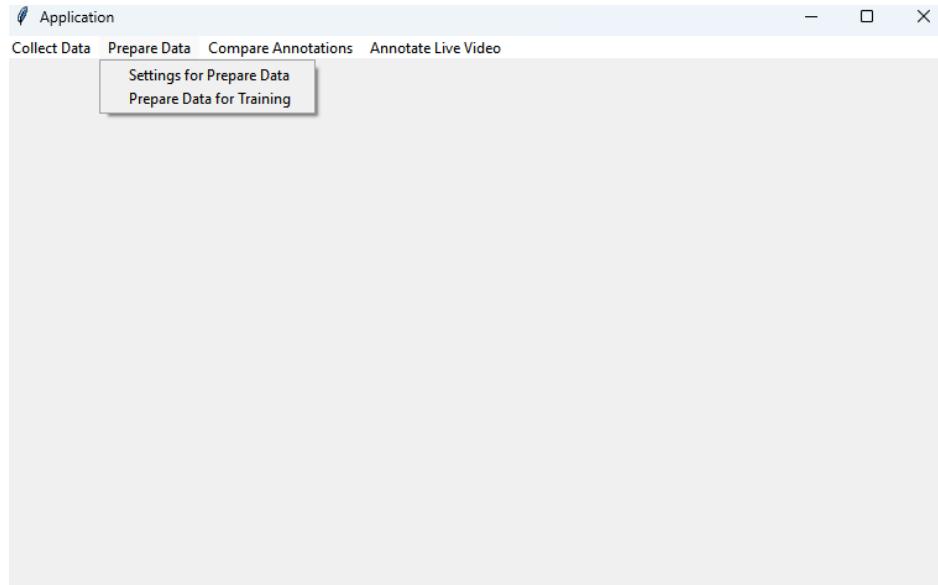


Figure 11: submenu of the "Prepare Data" menu option

**Settings for Prepare Data** Before you can use video recordings to train a model, datasets of photos need to be created (extracted) from the video recordings. Several variables need to be configured for this. Selecting the submenu item "Settings Prepare Data" opens the screen shown in Figure 12.

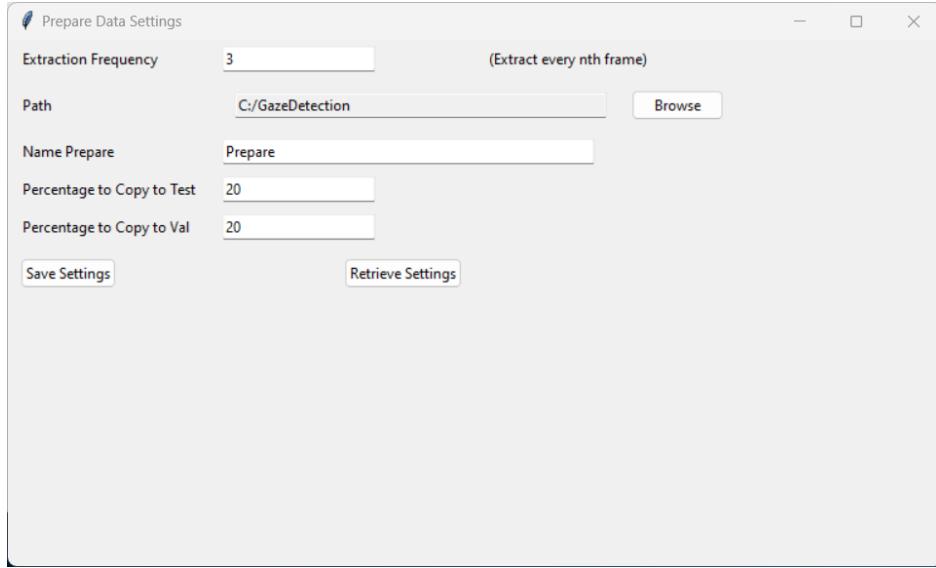


Figure 12: Prepare Data Settings

The following variables are configured on this screen:

- **Extraction Frequency:** How many frames to skip between each extracted frame from the video (Extract every nth frame).
- **Path:** The path to the video recordings.
- **Name Prepare:** The name of the video recordings.
- **Percentage to Copy to Test:** The percentage of the total extracted frames to be saved in the test directory.
- **Percentage to Copy to Val:** The percentage of the total extracted frames to be saved in the val (validation) directory.

The following actions can be performed:

- **Save Settings:** Click 'Save Settings' to save the current settings to a JSON file, so you can easily reload them later.
- **Retrieve Settings:** Click this button to load the most recently saved settings from the JSON file.

**Prepare Data for Training** Selecting the submenu item "Prepare Data for Training" opens the screen shown in Figure 13.

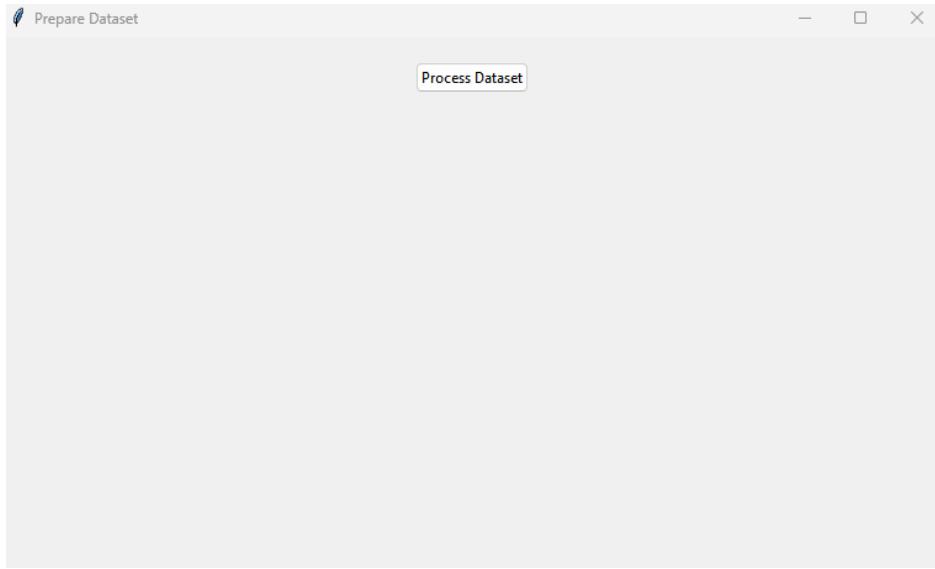


Figure 13: Prepare Data

Clicking the "Process Dataset" button will actually generate and save the dataset. The generated dataset follows a clear structure, with video files sorted based on their label and category (train, test, val). Each video file is named according to the 'cycle' (repetition), camera number, and frame number. The following diagram (14) outlines the directory structure used to organize the saved recordings. The diagram 14 outlines the directory structure used to organize the saved recordings. Each directory and file type is described in table 2 below.

```
root <Path / Name Recording>
|-- dataset
|   |-- train
|   |   |-- label_forward
|   |   |   |-- cycle_1_camera_1_frame001.avi
|   |   |   |-- cycle_1_camera_2_frame001.avi
|   |   |   |-- cycle_1_camera_1_frame002.avi
|   |   |   |-- cycle_1_camera_2_frame002.avi
|   |   |   |-- cycle_2_camera_1_frame001.avi
|   |   |   |-- cycle_2_camera_2_frame001.avi
|   |   |   |-- ...
|   |   |-- label_left
|   |   |   |-- cycle_1_camera_1_frame001.avi
|   |   |   |-- cycle_1_camera_2_frame001.avi
|   |   |   |-- cycle_2_camera_1_frame001.avi
|   |   |   |-- cycle_2_camera_2_frame001.avi
|   |   |   |-- ...
|   |-- test
|   |   |-- label_forward
|   |   |   |-- cycle_1_camera_1_frame001.avi
|   |   |   |-- cycle_1_camera_2_frame001.avi
|   |   |   |-- cycle_2_camera_1_frame001.avi
|   |   |   |-- cycle_2_camera_2_frame001.avi
|   |   |   |-- ...
|   |   |-- label_left
|   |   |   |-- cycle_1_camera_1_frame001.avi
|   |   |   |-- cycle_1_camera_2_frame001.avi
|   |   |   |-- cycle_2_camera_1_frame001.avi
|   |   |   |-- cycle_2_camera_2_frame001.avi
|   |   |   |-- ...
|   |-- val
|   |   |-- label_forward
|   |   |   |-- cycle_1_camera_1_frame001.avi
|   |   |   |-- cycle_1_camera_2_frame001.avi
|   |   |   |-- cycle_2_camera_1_frame001.avi
|   |   |   |-- cycle_2_camera_2_frame001.avi
|   |   |   |-- ...
|   |   |-- label_left
|   |   |   |-- cycle_1_camera_1_frame001.avi
|   |   |   |-- cycle_1_camera_2_frame001.avi
|   |   |   |-- cycle_2_camera_1_frame001.avi
|   |   |   |-- cycle_2_camera_2_frame001.avi
|   |   |   |-- ...
|   |-- readme.txt
```

Figure 14: Training dataset structure according to naming conventions

Directory	Description
Path	root <Path / Recording Name>
dataset	Main directory for the dataset
train	Directory for training data
test	Directory for test data
val	Directory for validation data
label_name	Labels for categorization (e.g., label_forward, label_left)
videoname.avi	Video files named as follows: cycle_<sequence>_camera_<camera number>_frame<frame number>.avi

Table 2: Dataset structure

### 4.3 Create Video

This feature allows for the recording of videos that are directly annotated by the driver. The purpose of these videos is to verify if the driver's manual annotation matches the model's automatically generated annotation, adding an extra validation step to the process.

During the recording, the system plays an audio prompt instructing the driver to look in a specific direction. The driver follows the prompts and speaks aloud the number visible at the center of the viewing direction. After a set number of seconds, the system gives the next prompt, repeating until all instructions are completed. The driver wears headphones (wired or wireless) to hear the instructions, and the spoken numbers are recorded and later translated by the application into gaze direction annotations.

This feature offers the following options: - Videos can be recorded with spoken instructions, where the driver speaks aloud the number they see in each prompted direction. This is recorded for later analysis of both video and audio. - Alternatively, videos can be recorded without prompts. As the driver navigates a quiet area (e.g., an industrial zone), they say aloud any number they see while looking around, capturing spontaneous annotations. - Finally, a (live) video can be recorded without any spoken instructions.

All videos can be subsequently annotated by the model. Videos with spoken numbers can then be compared to the model's annotations, allowing for analysis of the model's accuracy relative to actual driving conditions. This function must be executed from the car with the camera's in place. When you click on the 'Create Video' menu option in the menu bar, a submenu appears. See Figure 15.

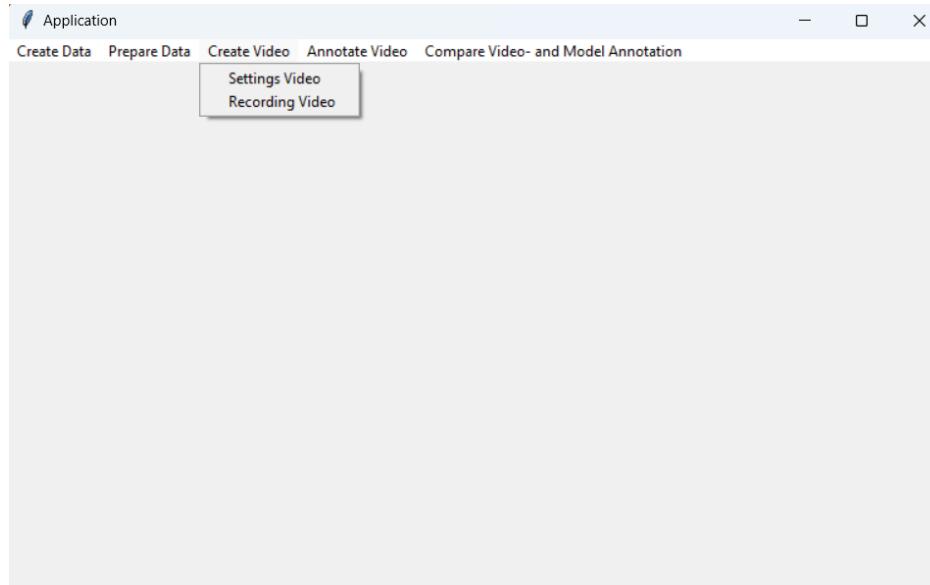


Figure 15: submenu of the "Create Video" menu option

**Settings Video** Before using video recordings for annotation, several settings need to be configured. Selecting the submenu item "Settings Video" opens the screen shown in Figure 16.

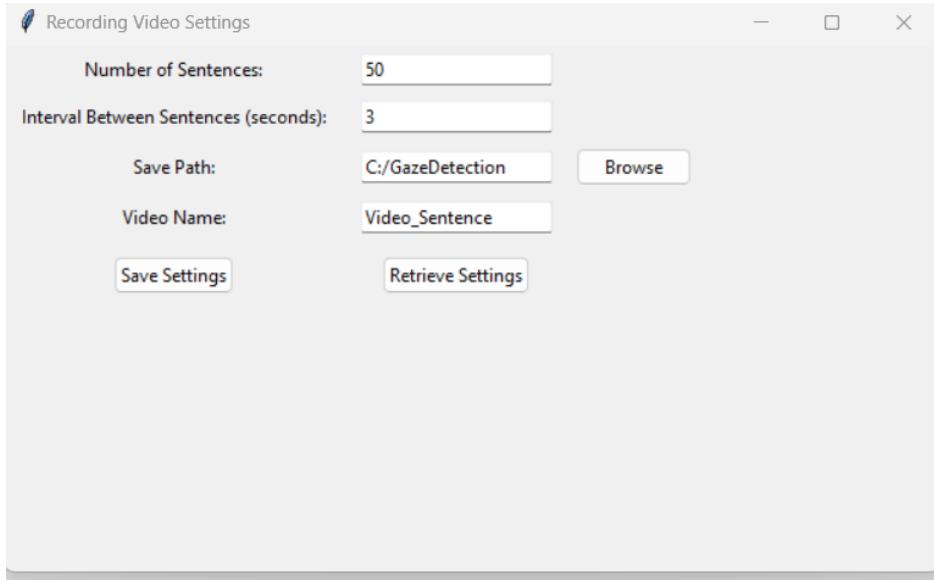


Figure 16: Recording Video Settings

The following settings can be configured:

- **Number of Sentences:** How many sentences should be spoken for this test video?
- **Path:** The path to the video recording.
- **Interval Between Sentences (seconds):** How often should a sentence be spoken (in seconds)?
- **Save Path:** The directory path where the video should be saved.
- **Video Name:** The name of the video.

The following actions can be performed:

- **Save Settings:** Click this button to save the configured settings to a JSON file.
- **Retrieve Settings:** Click this button to load the most recently saved settings from the JSON file.

Note: To record a video with instructions, you need to specify which labels will be used for the spoken instructions. This can be done using the "Setup Label" function in the Create Data menu. Camera setup can be completed using the "Cameras and Positioning" function, also located in the Create Data menu.

**Recording Video** Selecting the submenu item "Recording Video" opens the screen shown in Figure 17.



Figure 17: Start Recording Video

Clicking the "Start Recording With Sentence" button initiates the recording. The progress is displayed on the screen. Once the recording is complete (when the specified number of sentences has been recorded) the video is saved to the specified path (as entered in the settings) with the designated name.

Clicking the "Start Recording No Sentence" button begins the recording. The recording will continue until the "Stop Recording" button is pressed. The recording then ends, and the video is saved to the specified path (as entered in the settings) with the designated name.

The **Annotate Video** function allows the user to analyze and automatically annotate a video using a previously trained YOLOv8 model. Annotations are generated for each frame of the video, and the results are saved to a CSV-file. This ensures efficient and accurate detection of the gaze direction.

Additionally, the **Annotate Video** function can annotate a video where the driver speaks numbers aloud to indicate the direction they are looking. Each time a number is spoken, that moment is annotated—based on timestamps—with the corresponding label and stored in a CSV-file. This allows for model evaluation using videos recorded during actual driving conditions.

This function can be executed with previously recorded videos. When you click on the 'Create Video' menu option in the menu bar, a submenu appears. See Figure 18.

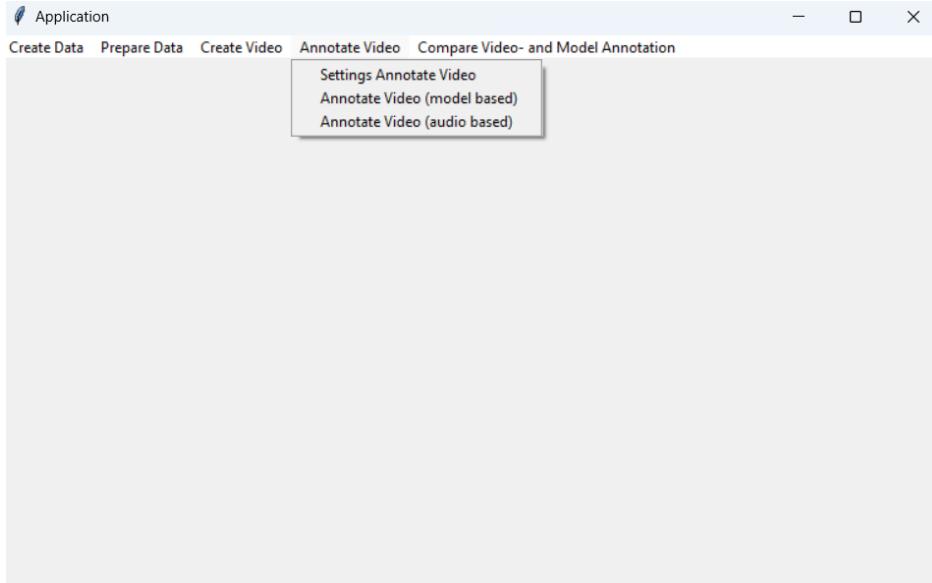


Figure 18: submenu of the "Annotate Video" menu option

**Video Settings** Before annotate a video, several settings need to be configured. Selecting the submenu item "Settings Annotate Video" opens the screen shown in Figure 19.

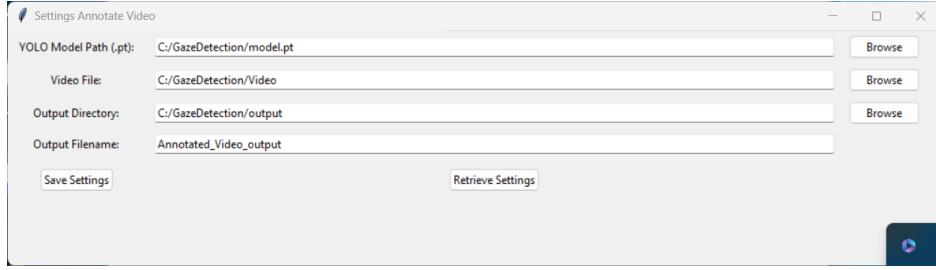


Figure 19: Video Recording Settings

The following settings can be configured:

- **YOLO Model Path (.pt)**: The name and the path to the trained YOLO model file.
- **Video File**: The path to the video file to be annotated.
- **Output Directory**: The directory where the annotation results will be saved.
- **Output Filename**: The name of the output CSV-file for storing annotation results.

The following actions can be performed:

- **Save Settings**: Click this button to save the configured settings to a JSON file.
- **Retrieve Settings**: Click this button to load the most recently saved settings from the JSON file.

**Annotate Video (Model Based)** When the user clicks on the submenu item "Annotate Video (Model Based)," the screen shown in Figure 20 is displayed.

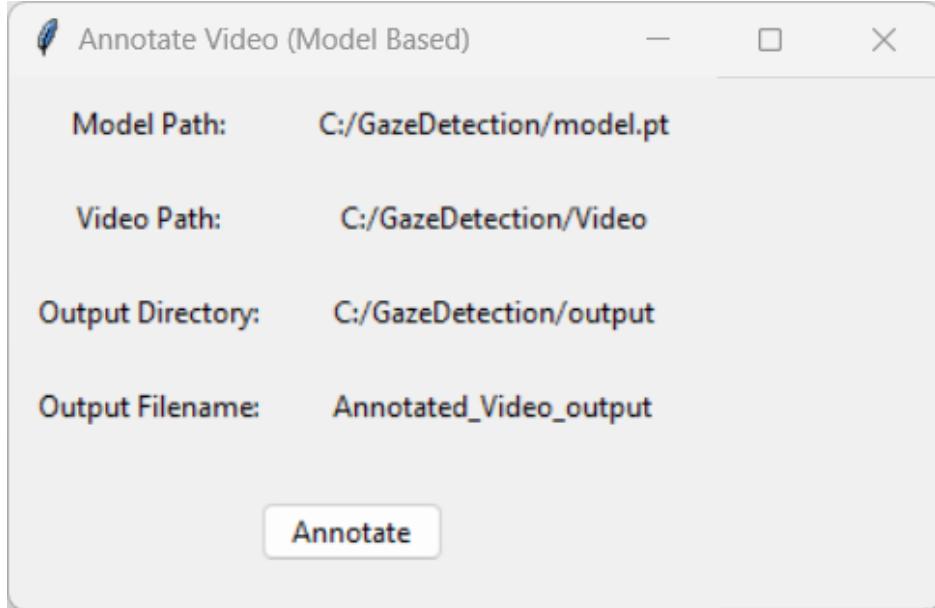


Figure 20: Annotate Video (Model Based)

The screen displays the settings as they were entered in the configuration. Clicking the **Annotate** button starts the system's analysis of the video. The YOLOv8 model annotates the gaze direction in each frame and saves the annotations to a CSV-file. This file contains frame numbers and the top 5 predictions with the highest probabilities. See Figure 21 for an example of the generated output.

A	B	C	D	E	F
Frame	Top1 (Label, Confidence)	Top2 (Label, Confidence)	Top3 (Label, Confidence)	Top4 (Label, Confidence)	Top5 (Label, Confidence)
0	dashboard_straight_down,53.17	mirror_left,29.30	shoulder_right,6.04	dashboard_down_right,3.68	forward_left,3.60
1	mirror_left,41.87	dashboard_straight_down,34.86	dashboard_down_right,9.47	forward_left,8.31	left,2.23
2	mirror_left,67.22	dashboard_straight_down,13.72	forward_left,7.10	dashboard_down_right,4.49	shoulder_right,4.01
3	mirror_left,69.47	dashboard_straight_down,18.84	forward_left,5.18	shoulder_right,2.42	dashboard_down_right,1.43
4	mirror_left,68.24	dashboard_straight_down,16.64	forward_left,6.00	dashboard_down_right,2.38	shoulder_right,2.31
5	mirror_left,35.84	dashboard_straight_down,27.80	forward_left,18.58	dashboard_down_right,5.67	shoulder_right,3.46
6	mirror_left,47.05	forward_left,19.35	dashboard_straight_down,16.32	shoulder_right,8.69	dashboard_down_right,2.68
7	mirror_left,40.58	dashboard_straight_down,26.06	forward_left,11.22	shoulder_right,10.85	dashboard_down_right,3.27
8	mirror_left,55.63	dashboard_straight_down,28.14	forward_left,8.74	shoulder_right,3.31	dashboard_down_right,1.45
9	mirror_left,59.95	forward_left,21.12	dashboard_straight_down,5.57	right,4.44	mirror_right,3.95
10	mirror_left,61.83	dashboard_straight_down,19.14	forward_left,12.14	shoulder_right,2.78	right,1.74
11	mirror_left,65.98	forward_left,19.02	dashboard_straight_down,8.53	shoulder_right,2.19	right,2.02
12	mirror_left,59.91	forward_left,16.68	dashboard_straight_down,12.98	shoulder_right,4.16	right,3.24
13	mirror_left,69.29	forward_left,12.95	dashboard_straight_down,10.31	shoulder_right,4.17	right,1.07
14	mirror_left,53.48	dashboard_straight_down,27.84	forward_left,8.49	shoulder_right,6.02	dashboard_down_right,1.32
15	mirror_left,58.41	dashboard_straight_down,31.43	forward_left,4.34	shoulder_right,3.32	dashboard_down_right,0.74
16	mirror_left,49.72	dashboard_straight_down,38.60	forward_left,4.80	shoulder_right,3.14	right,1.14
17	mirror_left,42.91	dashboard_straight_down,38.90	shoulder_right,7.16	forward_left,4.87	dashboard_down_right,2.10
18	mirror_left,45.60	dashboard_straight_down,41.15	shoulder_right,4.04	forward_left,3.70	dashboard_down_right,1.93
19	mirror_left,59.78	dashboard_straight_down,29.02	forward_left,5.53	dashboard_down_right,1.79	shoulder_right,1.05
20	mirror_left,62.10	forward_left,16.22	dashboard_straight_down,13.14	shoulder_right,3.66	dashboard_down_right,3.13
21	mirror_left,40.29	dashboard_straight_down,17.84	dashboard_down_right,17.73	forward_left,16.53	shoulder_right,4.42
22	dashboard_down_right,48.24	forward_left,16.29	mirror_left,14.48	shoulder_right,8.65	dashboard_straight_down,7.14
23	dashboard_down_right,39.52	dashboard_straight_down,27.33	mirror_left,12.93	forward_left,7.75	forward_right,7.35
24	dashboard down right,41.83	dashboard straight down,25.34	mirror_left,13.01	forward_left,7.40	forward_right,7.34

Figure 21: Example of an model based annotation \*.CSV output

Once the process is complete, a message appears indicating that the video has been successfully annotated, along with the location of the saved CSV-file.

**Annotate Video (Audio Based)** When the user clicks on the submenu item "Annotate Video (Audio Based)," the screen shown in Figure 22 is displayed.

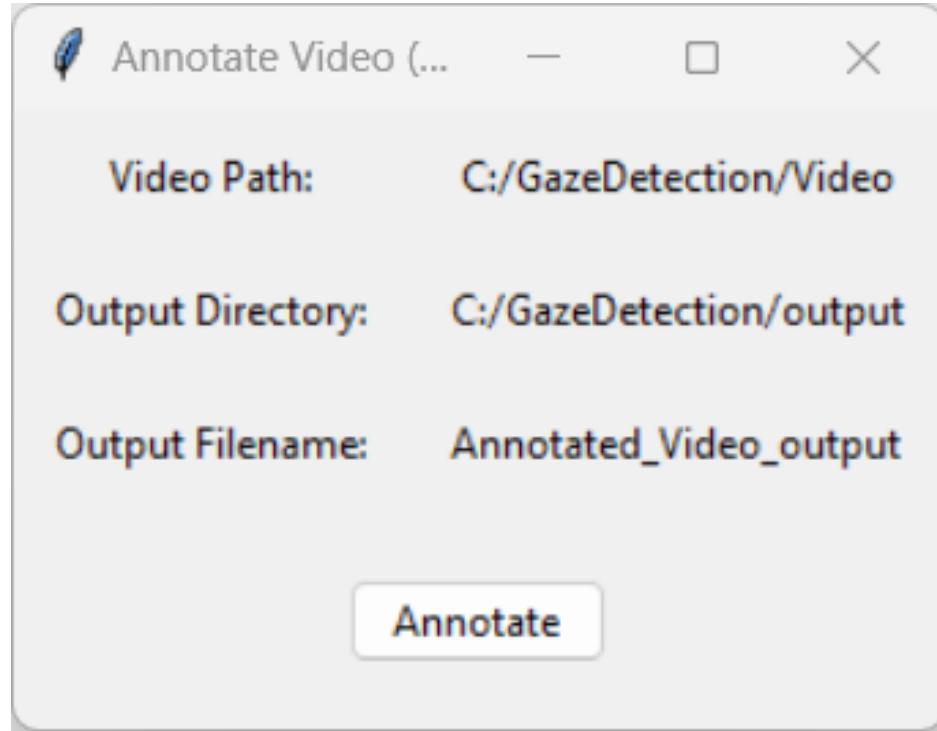


Figure 22: Annotate Video (Audio Based)

The screen displays the settings as entered in the configuration. Clicking the **Annotate** button starts the system's analysis of the video. The video will be analyzed for spoken numbers, which correspond to gaze directions. The annotations will be saved in a CSV-file containing timestamps and the corresponding gaze direction. See Figure 23 for an example of the generated output.

	A	B	C	D	E
1	Timestamp (s)	Number	Label	Start Frame	End Frame
2	0.0	8	shoulder_left	0	120
3	8.0	7	shoulder_right	120	180
4	16.0	3	right	180	240
5	21.0	9	dashboard_straight_down	240	300
6	26.0	3	right	300	360
7	31.0	2	left	360	420
8	37.0	2	left	420	480
9	42.0	3	right	480	540
10	48.0	9	dashboard_straight_down	540	600
11	55.0	7	shoulder_right	600	660
12	61.0	4	mirror_interior	660	720
13	66.0	1	forward	720	780
14	66.0	11	forward_right	720	780
15	72.0	8	shoulder_left	780	840
16	78.0	1	forward	840	900
17	78.0	10	dashboard_down_right	840	900
18	85.0	1	forward	900	960
19	85.0	11	forward_right	900	960
20	91.0	9	dashboard_straight_down	960	1020
21	97.0	1	forward	1020	1080
22	97.0	2	left	1020	1080
23	97.0	12	forward_left	1020	1080
	104.0			1080	1140

Figure 23: Example of an audio based annotation \*.CSV output

Once the process is complete, a message appears indicating that the video has been successfully annotated, along with the location of the saved CSV-file.

This function allows for accurate annotations of videos, which is essential for testing and validating models in various scenarios.

#### 4.4 Compare Video- and Model Annotation

Under construction