

Assignment 1

Intelligent Interactive Systems (1MD032)

Deadline: 2021/11/12 17:00 Stockholm Time

Overview

The objective of assignment 1 is for you to collect, clean, and annotate a small dataset. Later, we will aggregate these datasets into a large dataset which will become the base for your final project. To help you with the assignment, we have created a notebook to guide you through the individual tasks. You can find it on Studium.

The grade for this assignment is pass/fail, and to pass this assignment you must create and submit the following documents:

1. A filled-out information sheet and consent form (details below).
2. A set of 6 *anonymized* videos showing hand gestures (details below).
3. A CSV file containing the annotations for the 6 videos (details below).
4. A set of 6 *annotated* videos visualizing your annotations (details below).

Information Sheet and Consent Form

We want this course to be as relevant and cutting-edge as possible. As such, we want the practical part of the course to be similar in kind to the problems you might encounter in the wild. This requires that you collect and annotate enough data to build machine learning models, which can – frankly – get quite boring.

To minimize the data collection grind, we decided to go for a crowd-sourcing approach, where each of you collects and annotates a small amount of data, which is then shared with the rest of the course. As a result, *you must sign a form* that acknowledges that you have been informed about what we are doing here.

Additionally, you can choose to participate in our effort to extend this dataset beyond a single year and – over time – build a dataset that can be used to train general deep learning models. For this, we would like to add the result of your assignment to this extended dataset. Participation is completely voluntary, and if you are okay with it, you can tick the respective box while filling out the form.

Once you are done filling out the form, submit it on Studium before the deadline.

Task 1 - Anonymized Videos

Record 6 videos of your hand, each showing a different gesture. The current theme is American Sign Language (ASL) and this year's gestures are:

1. ASL_letter_A: The hand forms a fist with the thumb on the side.
2. ASL_letter_B: The hand forms a stop-sign with the thumb facing inwards.
3. ASL_letter_C: The hand forms a claw.
4. ASL_letter_L: The hand forms a fist with thumb and index finger extended.
5. ASL_letter_R: The hand forms a fist with the index and middle finger extended and with the middle finger crossed over the index finger.
6. ASL_letter_U: The hand forms a fist with index and middle finger extended and parallel.

An example picture of each gesture, and a drawing of the full ASL sign language alphabet, can be found in the appendix below. Additionally, you can find videos of demonstrations of the full ASL alphabet here:

https://www.youtube.com/watch?v=cGavOVNDj1s&ab_channel=LanguageofEarth

https://www.youtube.com/watch?v=tkMg8g8vVUo&ab_channel=ASLTHAT

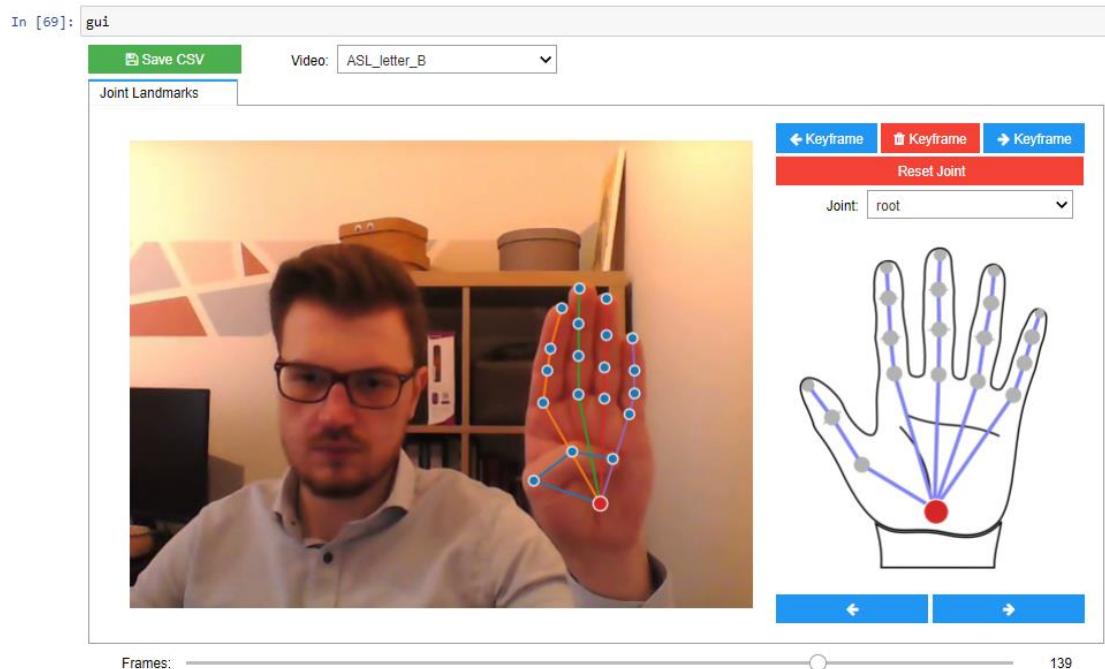
Each video you create must fulfil the following **requirements**:

- The name of the file must match the gesture shown in the video.
- The video must be in MP4 format.
- The gesture must be shown for at least 6 seconds.
- The hand showing the gesture must not be occluded by any objects.
- The hand showing the gesture must not occlude your face (if present).
- The video must be anonymized.

Submit the anonymized videos as a ZIP file on Studium before the deadline. The ZIP file must be named <last name>_<first name>_<student number>_videos.zip. For example, if your name is Sebastian Wallkötter and your student number is 1234 then the file name should be wallkötter_sebastian_1234_videos.zip (all lowercase ASCII characters).

Task 2 – Annotations

Create a CSV file containing the x and y position of the visible hand joints that can be seen in each frame of the videos you created during task 1. The joints correspond to the grey circles shown in the figure below. For this we provide an annotation tool. Upload the CSV file it creates on Studium before the deadline.



A picture of the GUI. On the right you can see a visualization of the skeleton and an indicator which joint is currently being annotated.

Task 3 - Annotated Videos

Write code that visualizes the annotations you created in task 2 on top of each video you recorded in task 1. For this, draw the hand's skeleton onto each frame according to the annotations found in the CSV file you created in task 2. Represent visible joints with a circle and connect visible adjacent joints with a line (representing the bone). Then, save the resulting frames as a new video called `<gesture>_annotated.mp4` replacing `<gesture>` with the respective name). The result should look similar to the drawing you saw while using the annotation tool.

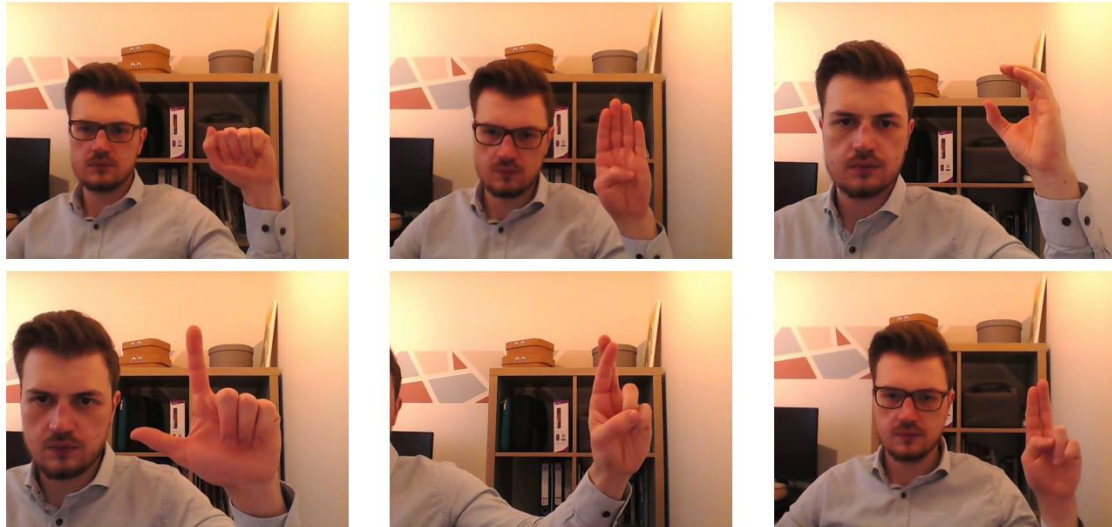
Submit the annotated videos as a ZIP file on Studium before the deadline. The file must be named `<last name>_<first name>_<student number>_annotated.zip`, replacing the tags as you did in task 1.

Hint: A rough outline of the program could look something like this:

```
for each video in gesture_videos:
    for each frame in video:
        place a label into the top left corner that reads "annotated"
        get the corresponding joint positions from the CSV file
        draw a circle at each (non-missing) joint position
        for each pair of visible joints connected by the skeleton:
            draw a line between the two joints using cv2.line
        create a new_video file called <gesture>_annotated from the annotated frames
```

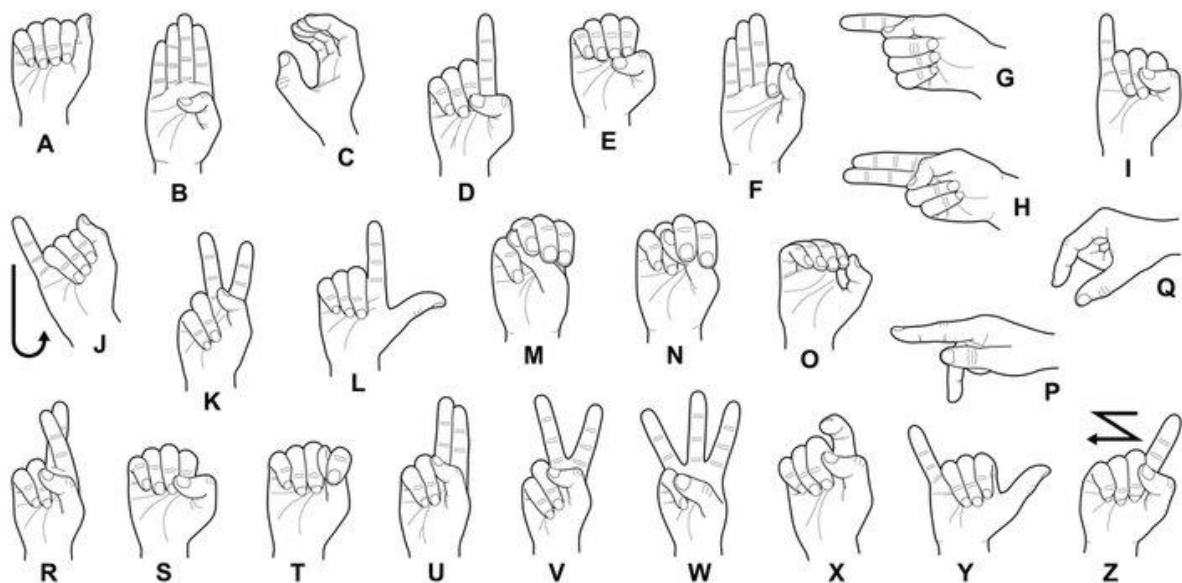
Appendix

Example Pictures of each hand gesture



From left to right, top to bottom:
ASL letter A, ASL letter B, ASL letter C
ASL letter L, ASL letter R, ASL letter U

The American Sign Language Alphabet



(Source: Adobe Stock)