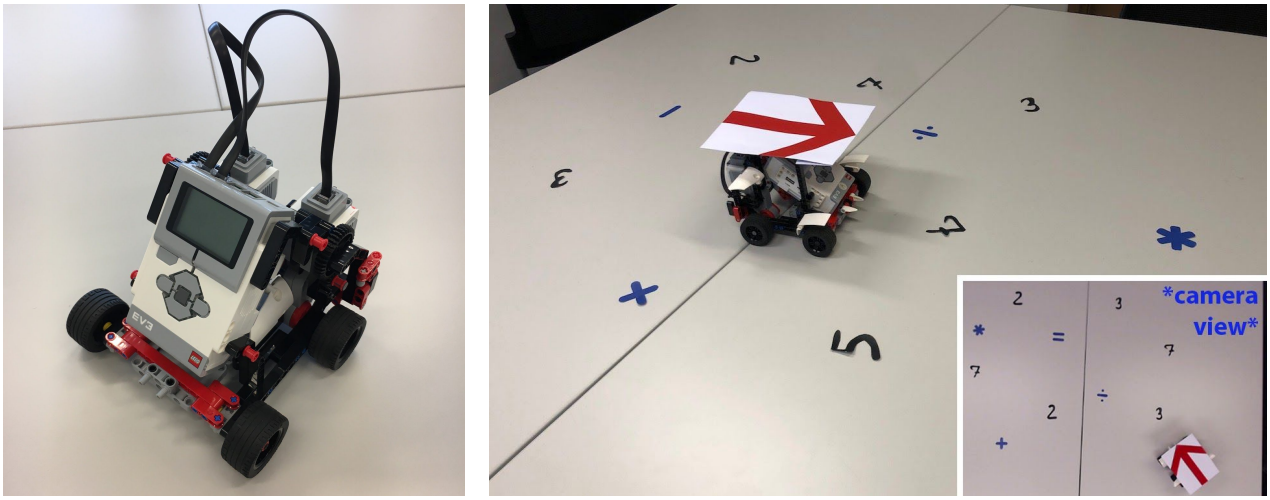
LTS5 - Image analysis and pattern recognition
lts5www.epfl.ch
EPFL, Lausanne, Switzerland

Special Project description

1st May 2020

1 Introduction

In this special project, your task is to use your image analysis skills to analyse the behaviour of a Lego© Mindstorm© robot (see Fig. 1.a) in a specific environment. The environment is a flat arena of approximately 3 meters by 3 meters containing visual elements such as different mathematical operators and handwritten digits with different colors (see fig. 1.b). The exact position and orientation of these elements may vary. In this project, we use recordings of the environment by a camera mounted above the arena and pointed directly at the arena, such that the plane of the arena is parallel to the image plane of the camera.



(a) (b)

Figure 1. (a) Lego© Mindstorm© EV3 robot (b) Project environment

In the following section, the detailed scenario of the project is presented. During the last session, Friday, May 29th, you will run your code on a pre-recorded test sequence and present your approach to the problem, explaining briefly how you implemented your solution.

2 Task: Find and solve the math problem

Your main task is to find the result of an equation based on a video sequence. The equation will be indicated by a moving robot. The detailed scenario is defined as follows:

- Several mathematical operators (multiplication, division, minus, plus, equal) are placed on the table. The color of operators is **blue**.
- Several handwritten digits (0 to 8) are placed on the table. Digits are always **black**.
- From an initial location somewhere on the table, the robot moves around the table. The **red arrow** on the top side of the robot will help you to localize it and detect it's orientation. Each time the robot passes above an operator or a digit, the symbol located below the robot is added to the equation. For example the sequence "2" ➔ "+" ➔ "3" ➔ "=" becomes "2+3=".
- The sequence always starts with a digit and ends with the operator "=".

The goal is, given a new video sequence, to retrieve the formula and its associated answer. During the final session, you will be presented with three test scenarios:

    - SC1: All operators/ digits have vertical orientations.
    - SC2: Both operators and digits have random orientations.
    - SC3 **(Bonus)**: Both operators and digits have black colors. Orientations are random.

The input of your algorithm is a **".avi"** video sequence, recorded at **2 FPS**. Your approach should output a video sequence with the same frame rate, duration and resolution as the input video. **Each frame** (e.g. frame at time **t**) of the output video should contain the following information, printed on the same frame:

- The current state of the formula at time **t**.
- The trajectory of the robot from start to time **t**.

Note 1: The visual aspects of the output video is up to you. No special requirement from our side.

Note 2: There are multiple operators and digits. Please note that their sizes and brightness can slightly change at the test time.

3 Organization

3.1 Sessions

Four sessions are dedicated to the special project, organized similarly to exercise sessions. During the last session on May 29th, you will present your work and test your implementation.

May 1	Intro+Q&A	Zoom
May 8	Q&A	Zoom
May 15	Q&A	Zoom
May 22	Q&A	Zoom
May 29	Presentations	Zoom

Table 1: Calendar and organization of the special project sessions

3.2 Teams

You will work in teams of three (preferably same teams that you had for the labs). You should have already registered your teams on the Moodle; if not, please do it as soon as possible.

3.3 Tools

The complete implementation of this special project will be done in Python. You are free to use any external libraries and pre-implemented functions from previous labs.

3.4 Data

We provide you an example ("robot_parcours_1.avi") in our github, taken in similar conditions to the final test. You are free to extract different shapes for this video, crop and use them to train classifiers, if needed. For digit classification, you are free to use the MNIST dataset that you have seen in Lab 3.

4 Deliverables

Your final deliverable should contain a "main.py" file (No Jupyter Notebook is accepted for the project), and getting the path of the input video, as well as the path of the desired output as an argument:

```
$ python main.py --input /path/to/input/video.avi --output
/path/to/your/result.avi
```

Your approach should output a ".avi" video sequence (see section 2).

During the last session, Friday, May 29th, each team will also present their approach in 10 minutes. You should focus on your solution and provide a clear overview of the choices you made and the steps you implemented. No report is needed. You will also run your code on a pre-recorded test sequence (SC1, SC2 and SC3) at the time of the presentation.

You should provide an archive containing all the python files + list of packages used in your code by **May 28th, 11:59 PM**. Please name the archive file with the number of your team, e.g. Team28.zip.