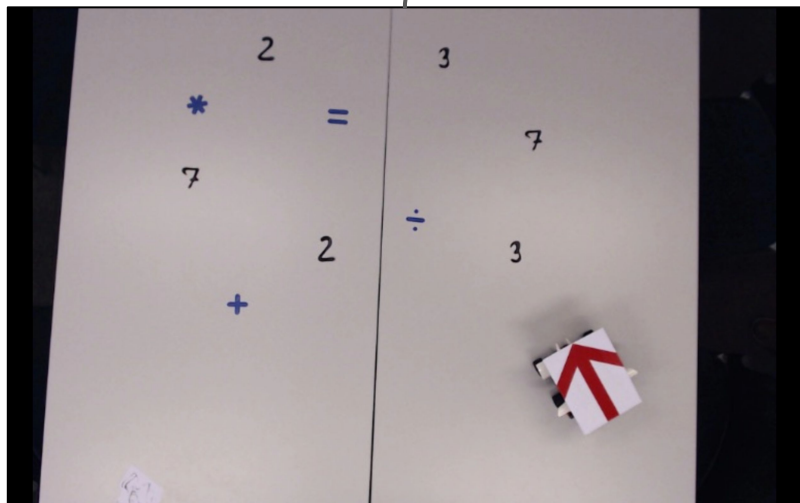


# Special project

Image analysis and pattern recognition

Oganes Manasian  
Maksim Kriukov  
Henry Declety

# General overview



## Robot tracking:

- Frame differencing (by default)
- Red channel tracking
  - Otsu thresholding

## Object detection:

- Canny edge detector (by default)
- Otsu thresholding
- Filtering based on size, shape and area

## Intersection detection:

- Robot center coordinates overlap with digit bounding boxes (by default)
- Robot bounding box overlaps with digits centers

## Operator prediction:

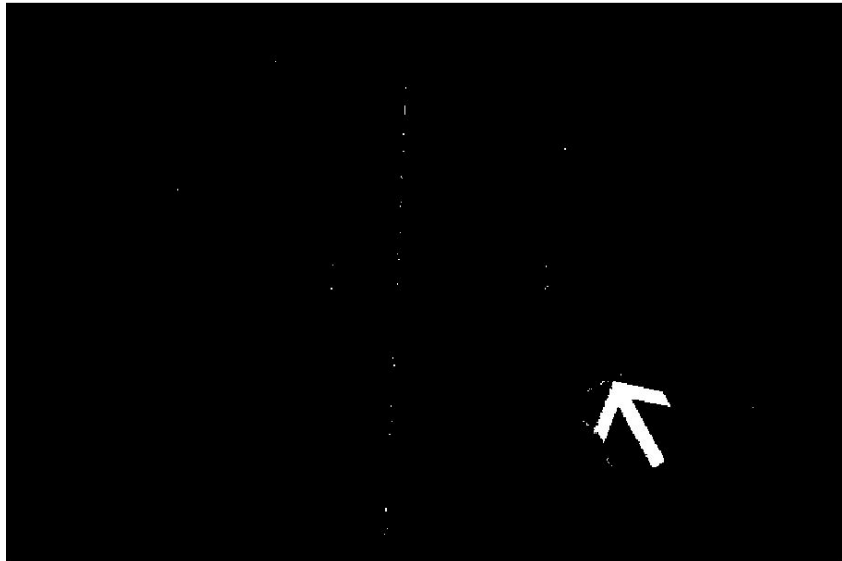
- Pretrained CNN on augmented operator images

## Digit prediction:

- Pretrained CNN on MNIST images

Video generation and equation calculation

# Robot tracking



The obtained mask after Otsu thresholding

## 1st method - by default

1. Blur the frames by Gaussian filter
2. Subtract adjacent image arrays to detect moving arrow

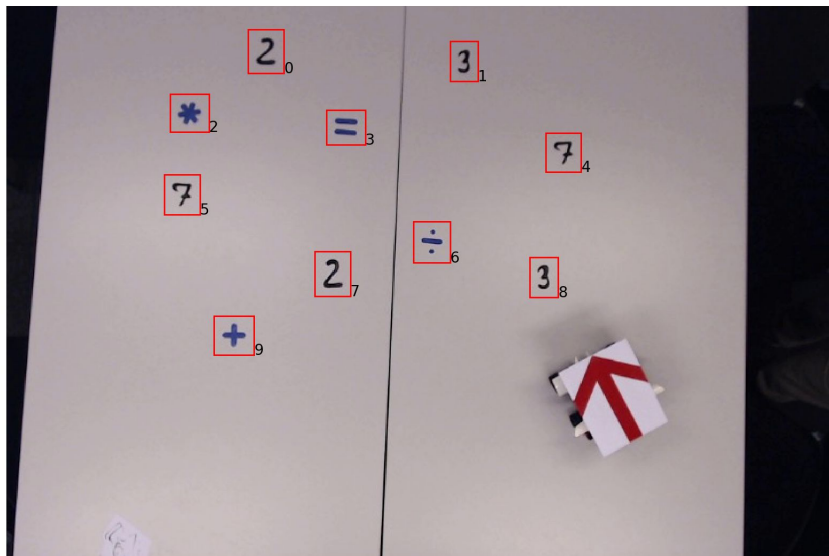
## 2nd method

1. Apply Otsu thresholding to only **red** channel → minimizes intra-class variance of bimodal distribution
2. Get the mask in each frames by:  
 $R > \text{otsu\_threshold}$   
 $G < \text{otsu\_threshold}$   
 $B < \text{otsu\_threshold}$
3. Binary closing + region growing
4. Take biggest region

## Automatic mode

Chooses method with the lowest difference in the steps' length

# Object detection

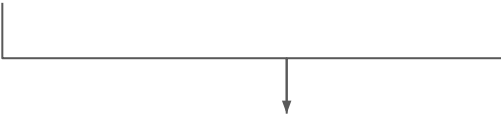


## 1st method - by default

1. Take only first frame and grayscale it
2. Apply Canny detector  
→ multi-step algorithm with blurring, gradient calculation and extraction of the "largest" edge

## 2st method

1. Take only first frame and grayscale it
2. Apply Otsu thresholding

- 
1. Binary closing
  2. Region growing
  3. Filtering
    - a. By pixel area
    - b. By shape and size of the bounding box

# Classification

## Classification of rotated symbols

by default

### Approach 1

1. Augment MNIST with rotations
2. Train CNN
3. Do predictions as usual
4. Loss: Cross entropy
5. Early stopping on rotation augmented video sample set



### Approach 2

1. Take binary MNIST
2. Train CNN
3. Do predictions of image at different rotations
4. Loss: Cross entropy + label smoothing
5. Early stopping on rotation augmented video sample set