Tracking of an Equation Reading Robot

EPFL - Image Analysis and Pattern Recognition

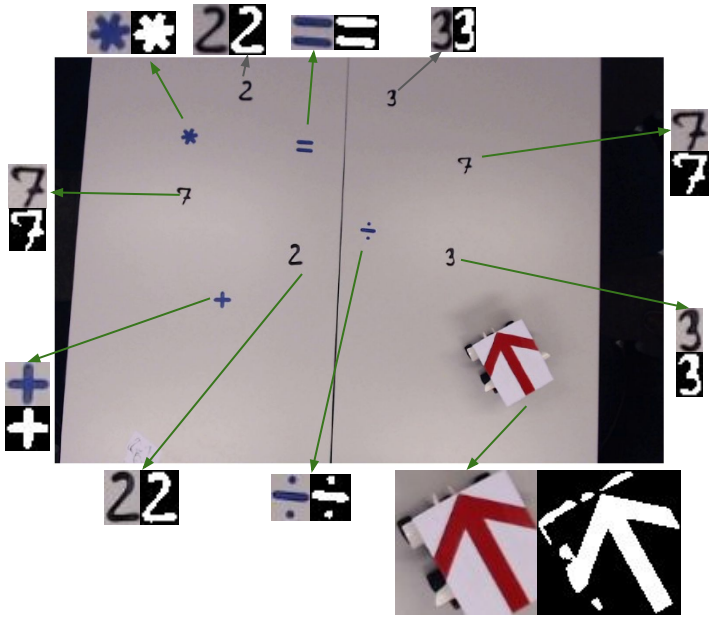
Team 35: Francesca Luongo, Antoine Spahr, Mathieu Mach

General Approach

- Detection of elements on first frame only
- Extract objects from mask
- Classification Algorithms
  - Type of objects → KMeans (on colors)
  - Operators → KNN (on Fourier descriptors)
  - Digits → MLP trained on rotated MNIST
- Tracking of Arrow
- Equation reading by overlap detection

Detection - General and Color-Kmeans

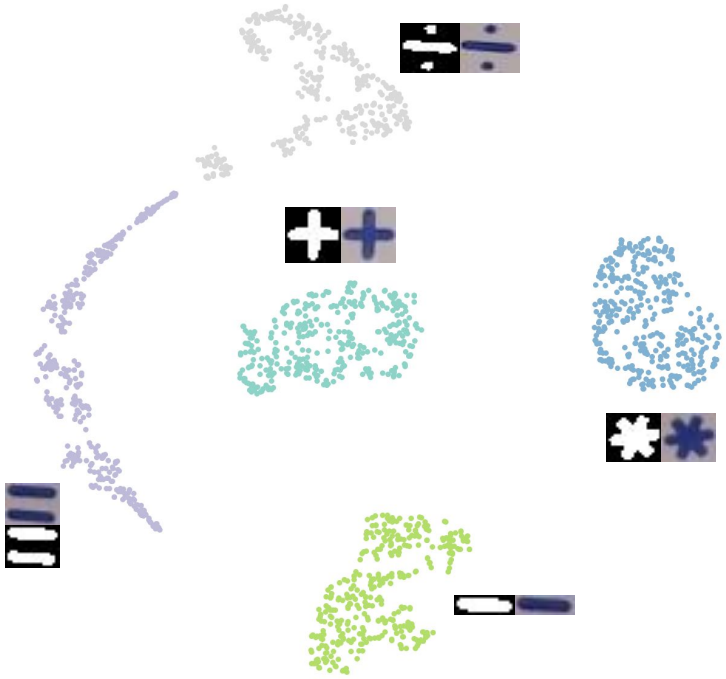
- Extraction of frames, masks and equation elements.
- Kmeans on average element pixels color
- Pixels as samples, RGB as features
- Manual class attribution to detected color cluster.



Detection - KNN

- Perform KNN on the 5 first Fourier Descriptors.
- 1800 images generated from 360 rotations of 1° on the 5 operators images.
- Train Accuracy : 100%
- Test Accuracy : 100%

t-SNE representation of 5 Fourier descriptors.



Detection - MLP

- Trained on rotated data (original + 2 rotated version)

  → 162’153 train images

  - Converged in 121 epochs with learning rate of 1e-4
  - Layers: (784,200) → (200,150) → (150,9)
  - ReLU activation function
  - Adam optimizer
- Train Accuracy : 99.58%
- Test Accuracy : 93.04%

Tracking

- Iterative approach of bounding box on every frame
- Compute all masks but only keep biggest one → the arrow