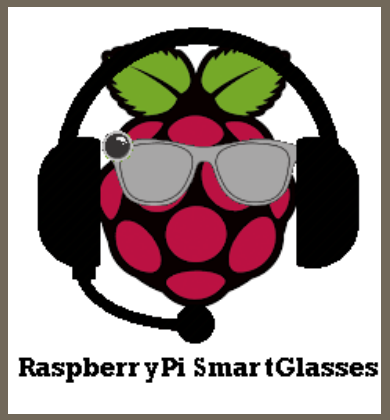


Final Project: Hand Gestures Interface for Low-Cost DIY Smart Glasses

Matan Argaman
Ronen Tamari
Advisor: Ofer Springer

Project Goal

- Develop fluent and natural interface for using hand gestures with smart glasses

The Problems

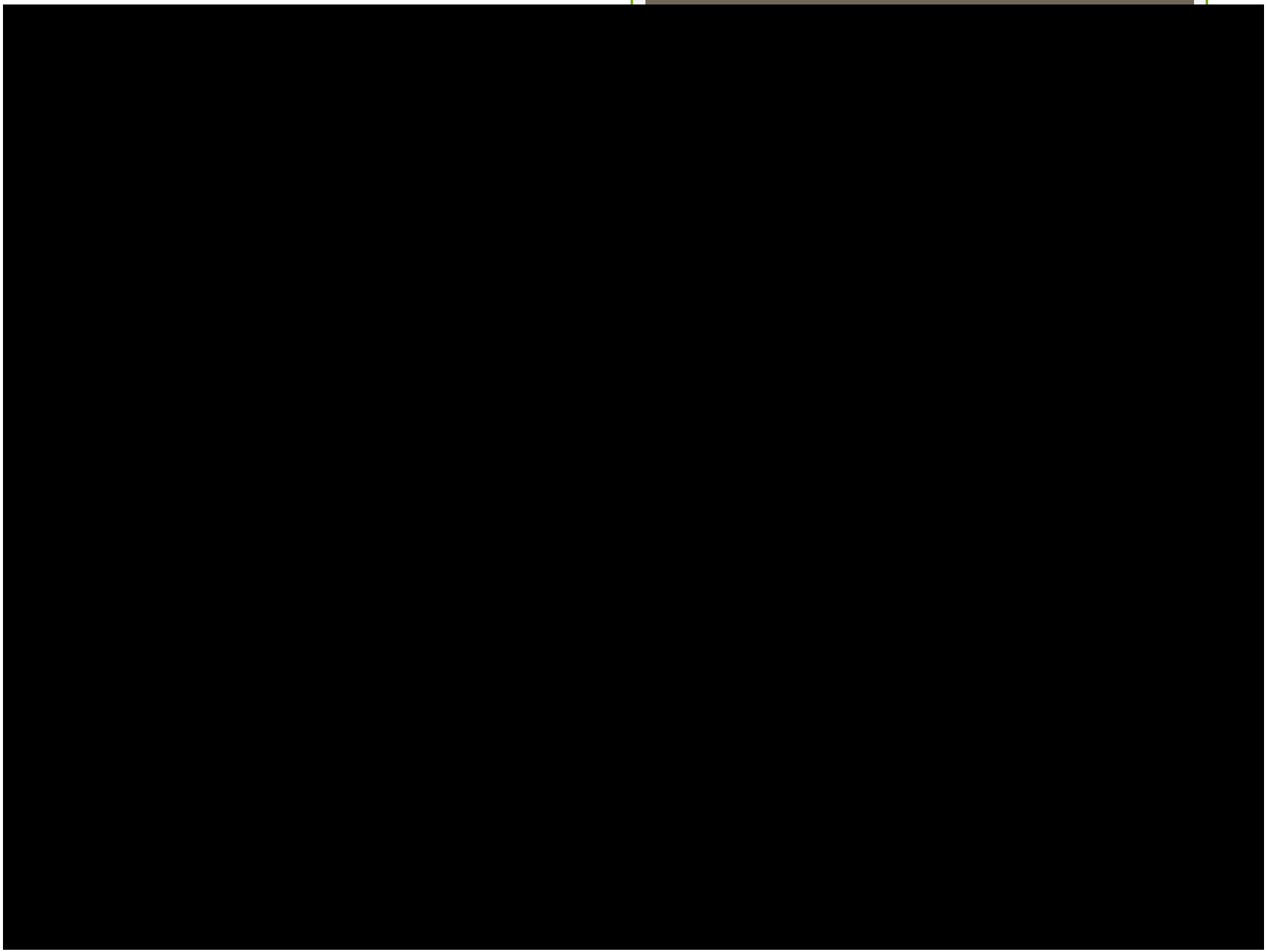
- Un-natural interfaces for smart glasses
  - Overuse of voice commands
  - Limited use of hand gesture recognition
  - **Limited interaction with the real world**

The Problems

- Current hand gesture recognition methods
  - Have trouble with noisy backgrounds/camera movement
  - Sensitive to lighting variation

Our Solution

- Use of hand gestures to make queries referencing the world seen by the user

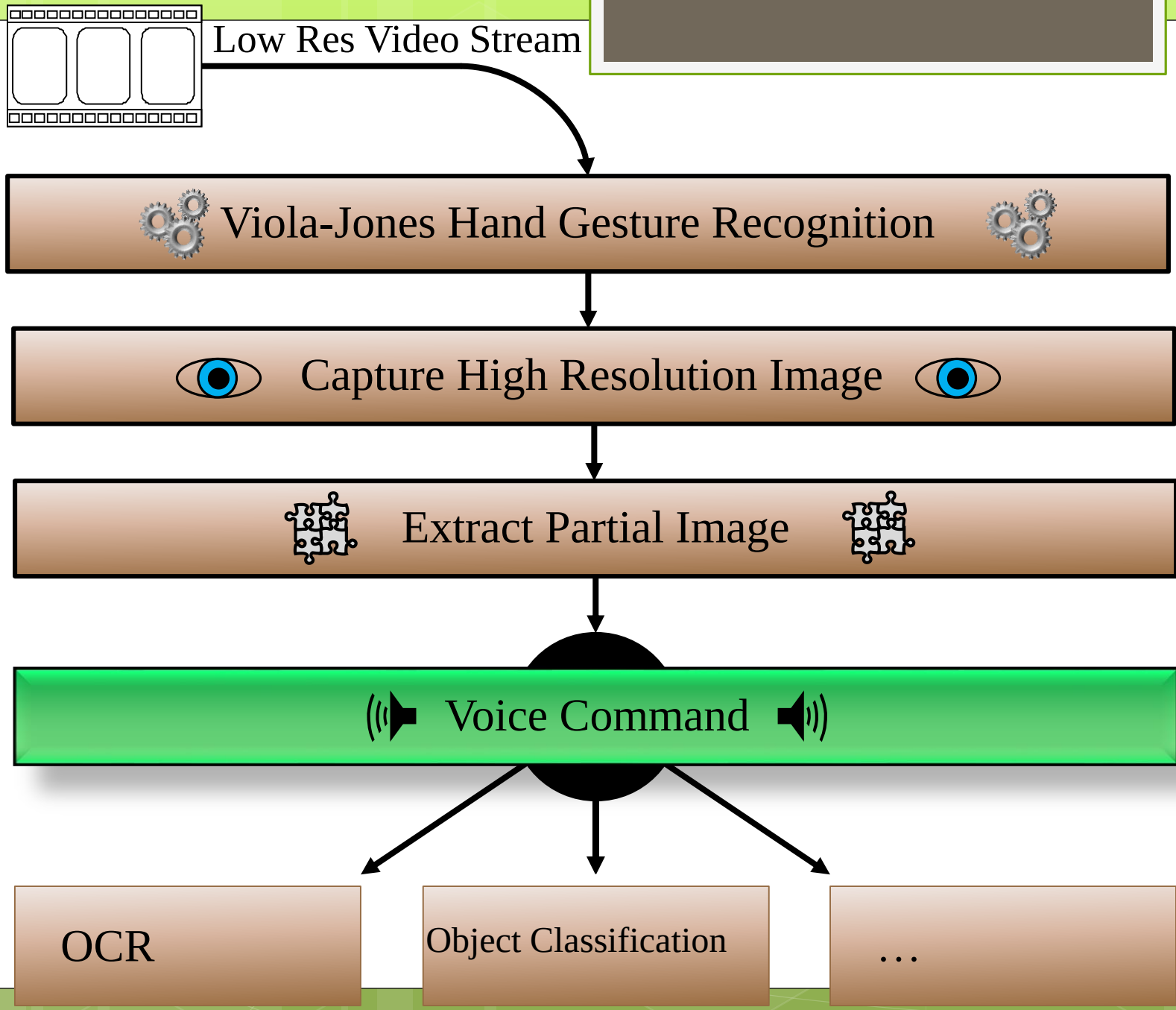


Hardware

- On paper:
- Vuzix M100 Smart-Glasses (1000$)

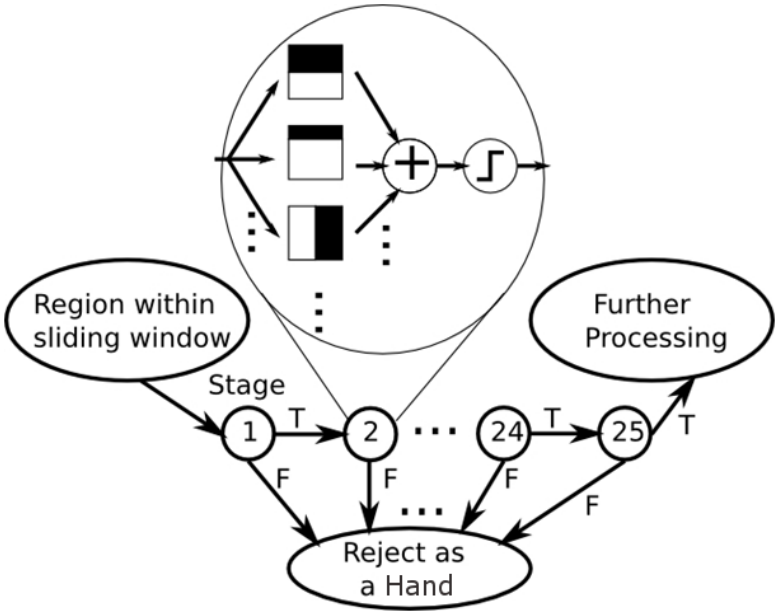
Hardware

- In reality….
- “Poor Man’s Smart Glasses”
- Solution based on Raspberry Pi with Wifi module and custom wide angle camera
- ~140$ price tag compared to 1k ++


Low Res Video Stream
Viola-Jones Hand Gesture Recognition
Capture High Resolution Image
Extract Partial Image
Voice Command
OCR
Object Classification
…

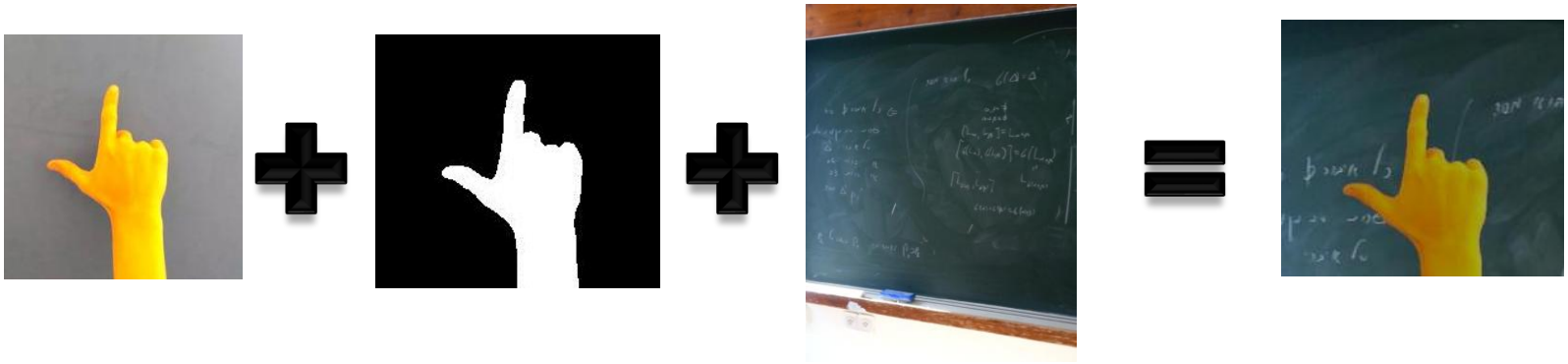
Hand Detection Algorithm

- Cascade classifier based on Viola Jones over Haar feature set



Challenges

- To handle high scene variance, classification task requires large dataset (thousands of samples)
- Sample synthesis using small core initial dataset created using automated hand segmentation
- Large selection of backgrounds from web based datasets and videos



Challenges

Training two classifiers for one gesture – double the work!

Avoid training left hand classifier by right hand classifier + symmetry

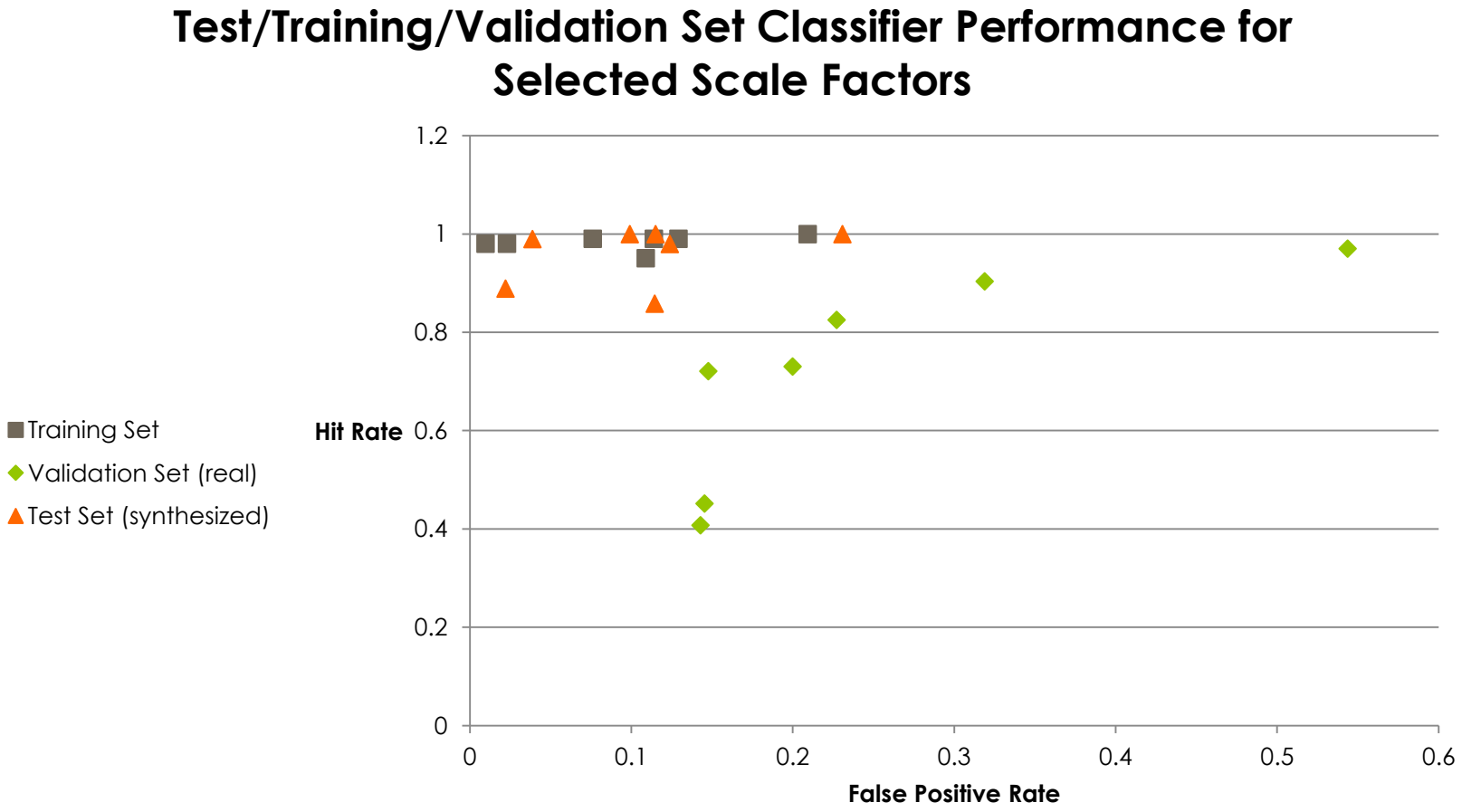
Flip horizontally

Rotate CW 90°

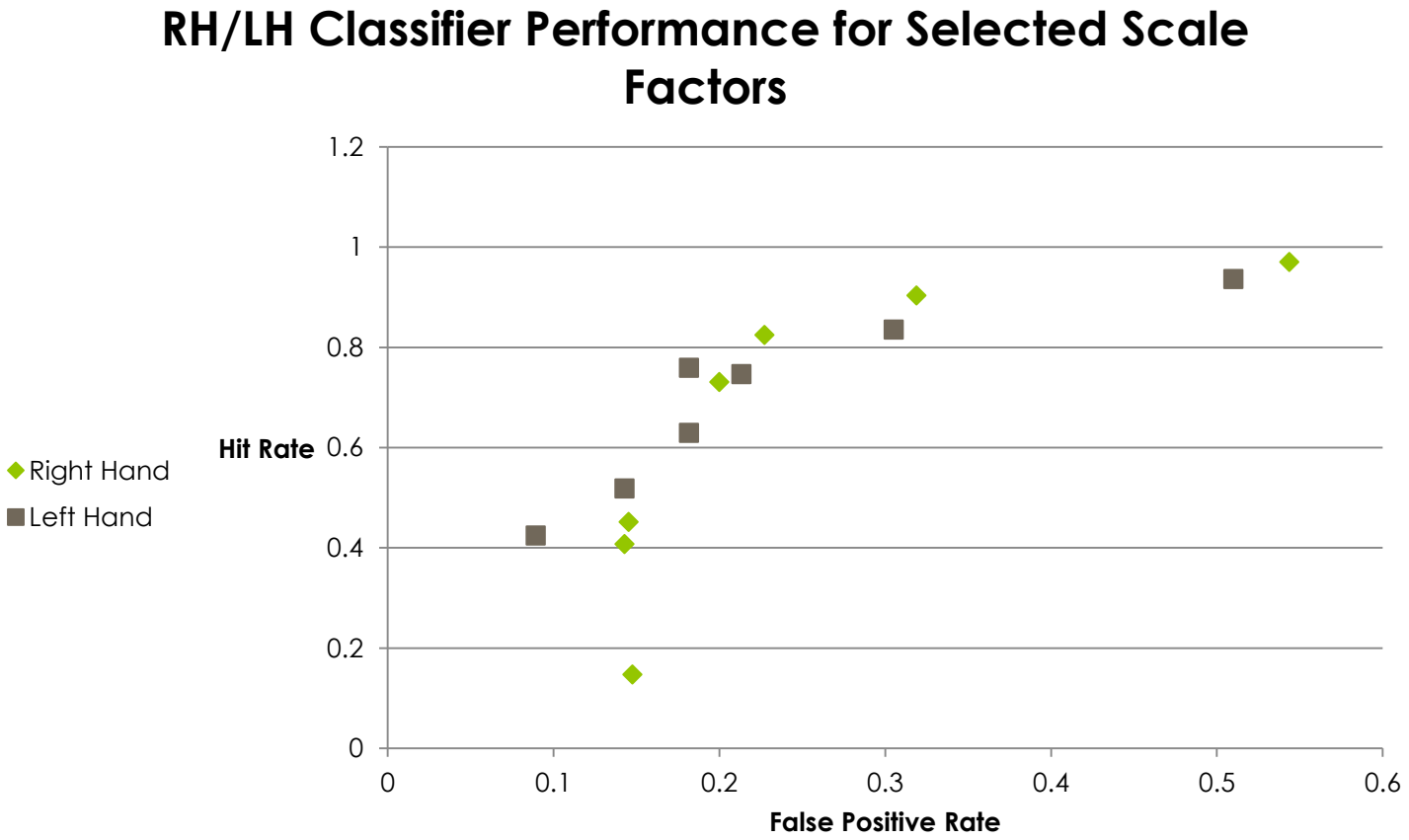
Challenges

- Achieve real-time, quality detection on Raspberry Pi (Single-core 700 Mhz CPU)
  - Detection on low resolution, single channel image. Extract high res color image only if gesture detected.
  - Detect left hand only after right detected (to avoid 2 detection tasks per frame)

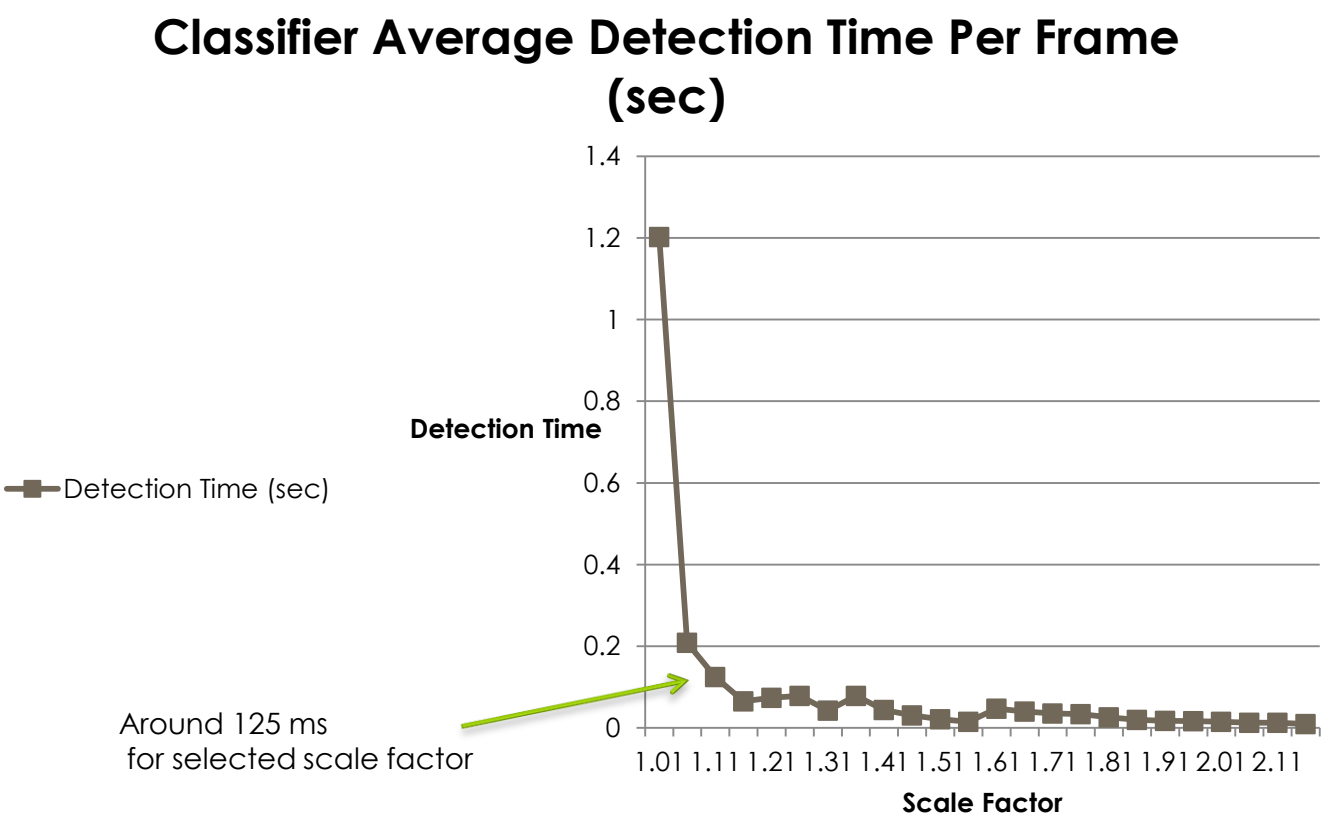
Results


Test/Training/Validation Set Classifier Performance for Selected Scale Factors
Training Set
Validation Set (real)
Test Set (synthesized)
Hit Rate
1.2
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
False Positive Rate

Results


RH/LH Classifier Performance for Selected Scale Factors
Right Hand
Left Hand
Hit Rate
1.2
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
False Positive Rate

Results


Classifier Average Detection Time Per Frame (sec)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Detection Time
Detection Time (sec)
1.01 1.11 1.21 1.31 1.41 1.51 1.61 1.71 1.81 1.91 2.01 2.11
Scale Factor
Around 125 ms
for selected scale factor

Overall Classifier Performance

- Total False Positive Rate:

$$FPR_{total} \approx FPR_{right} \cdot \frac{1}{4} FPR_{left} \approx 0.25 \cdot \frac{1}{4} 0.5 \approx 3\%$$

- Total False Negative Rate (1-TPR):

$$FNR_{total} \approx 1 - \underbrace{TPR_{right} \cdot TPR_{left}}_{TPR_{total}} \approx 1 - 0.85 \cdot 0.93 \approx 21\%$$

Conclusions

- Good classifier performance and generalization ability
- Significant reduction in data collection time, which is the largest bottleneck of training process.
- Scalability still relatively low (adding new gestures is costly)
- Versatile interface due to open source and web connectivity .

Future Directions

- Improvement of dataset and data collection automating
- Classifier sensitive to scale - train classifier over more scale invariant features?
- Create API and “unleash on web” for community to develop own apps.
- Hardware streamlining and improvement- 3D printing of glasses , tailor power solution.
- General hand gesture learner – using neural networks and simulated 3D hands to synthesize massive datasets?

Questions?