THE ANDREW & ERNA VITERBI FACULTY OF ELECTRICAL ENGINEERING

SIPL
Signal and Image Processing Lab

TECHNION Israel Institute of Technology

Eye and Head Tracking for Controlling A Reading Software for Children

Dor Arad and Tomer Keren, Supervised by Ori Bryt

Introduction

- Analyzing gazed points on a screen computer ,during task performing on that computer can be great tool for analyzing the cognitive state of the kid.
- Traditional gaze detection tool require special hardware or fixed head position
- Predicting the above can be hard using domestic simple equipment

Goals

- Creating efficient online gaze estimator with head pose consideration for educational software.
  - Domestic easy-to-use
  - Robust treatment for varying users and conditions
  - Computationally efficient for online applications

Challenges

- Various room conditions and computers manufacture could harm results
- Not using special hardware and yet obtain satisfactory results
- Dealing with a lot of data efficiently

Solution Over-View

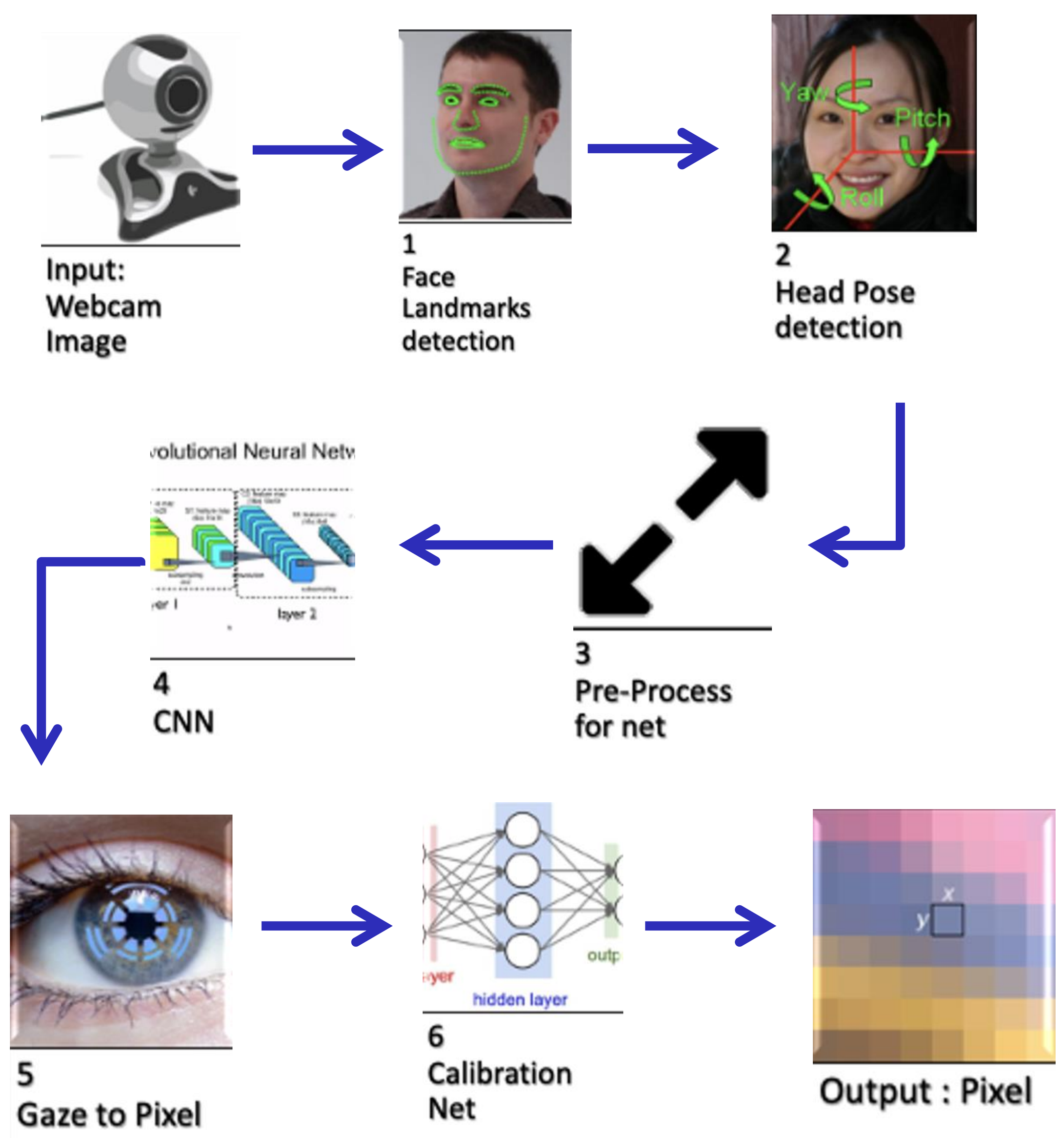
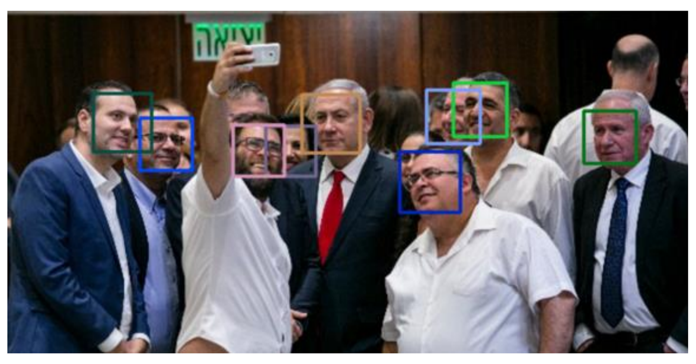
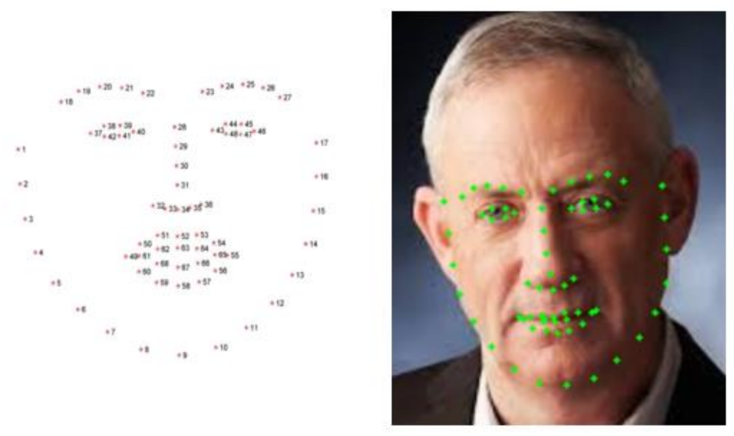


Image to CNN

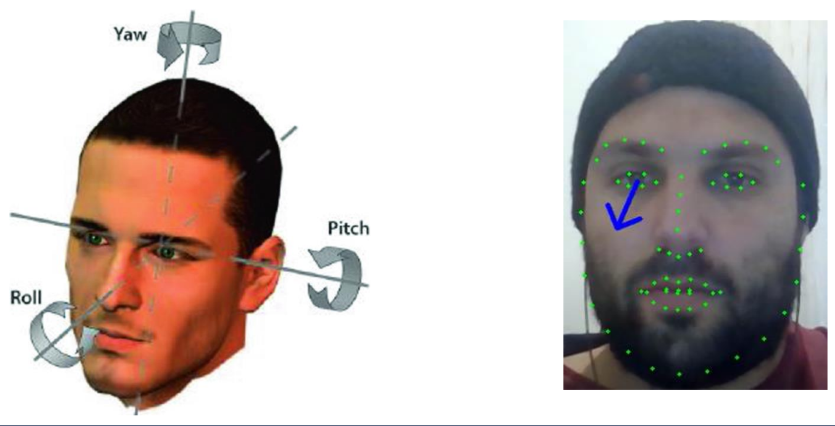
- **Face detection** using the Viola & Jones algorithm



- **Face Landmarks detection** using dlib library detector
  - Determine 68 key-points in the user's face

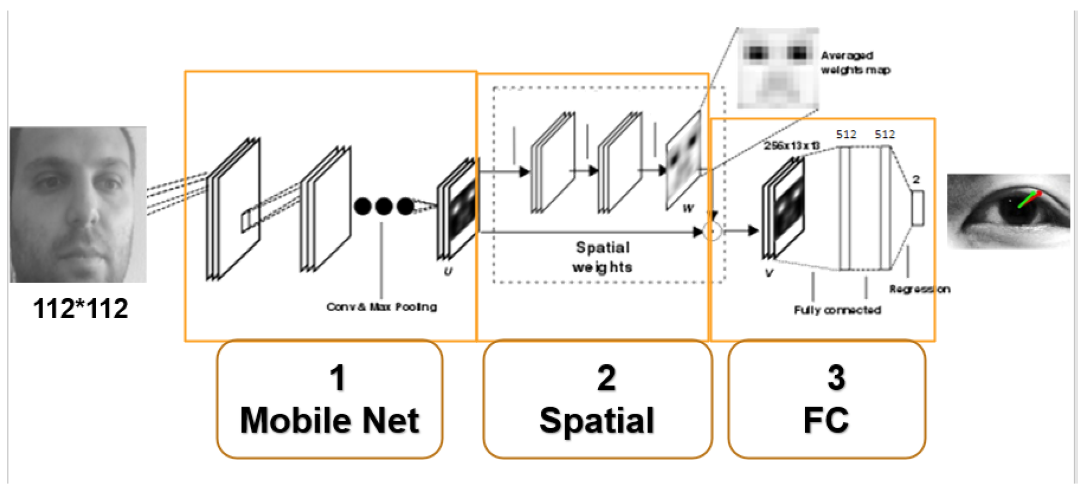


- **Head Pose calculations** using the Perspective-N-Points algorithm

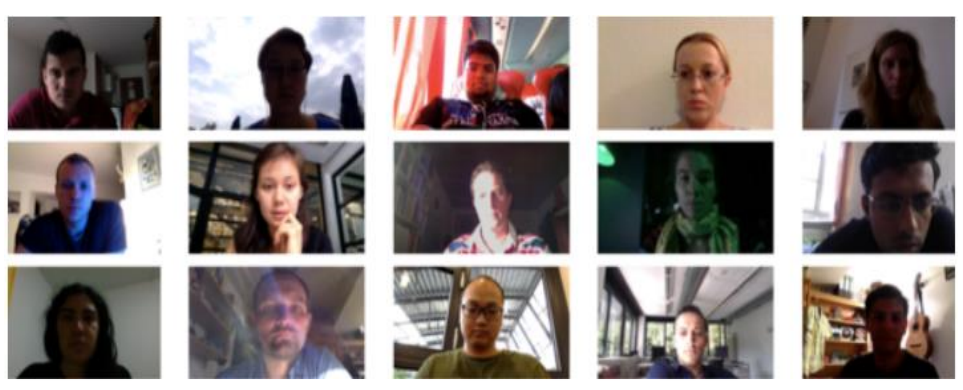


Convolutional neural network

- Machine learning system, broadly use to solve computer vision problems
  - Inputs are face image, fixed eye location and fixed head pose
  - Outputs are two angles that determine the Gaze direction, calculated from face center
  - Pre-Process for each face, normalizing the image using rotation and translation vectors

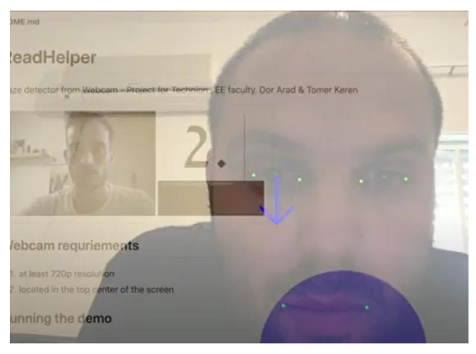


- CNN was trained over a large dataset
  - MPII Gaze Database contain 214,000 images over large variance of participants



Gaze to Pixel

- Sub-system used to convert the gaze angles, into a pixel on our screen, using Two mathematically-different approaches
  - Linear method, get relation between angle and pixel due to small angle approximation
  - Geometric method, solves 3D trigonometric problem between user and screen surface.

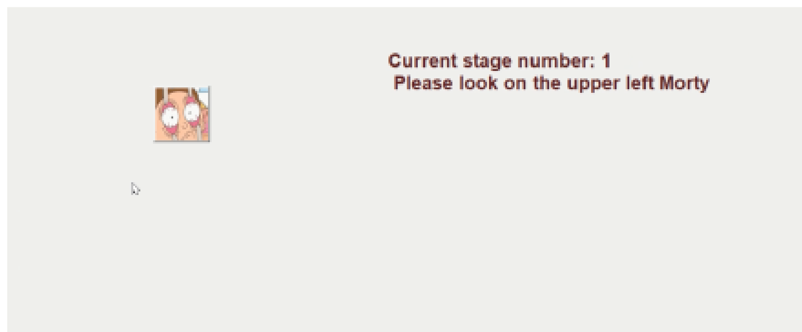


Calibration Net

- Using simple linear regression, which is trained at pre-process calibration stage by the user.

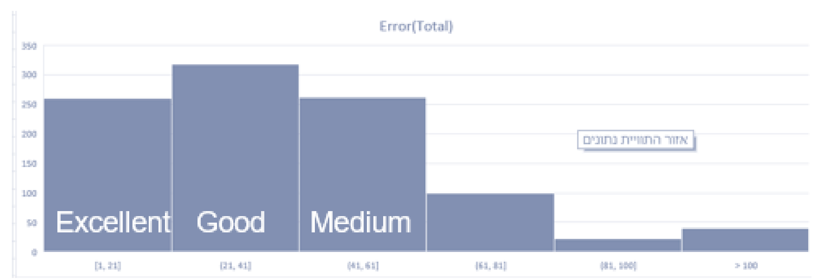
$$pixel_{fixed} = (x_{fixed}\ y_{fixed}) = (w_{xx}\ w_{xy}\ w_{yx}\ w_{yy}) * (x_{res}\ y_{res}) + (b_x\ b_y)$$

- Trying to approximate the errors as linear, handle over-fitting and determine threshold for outliers.

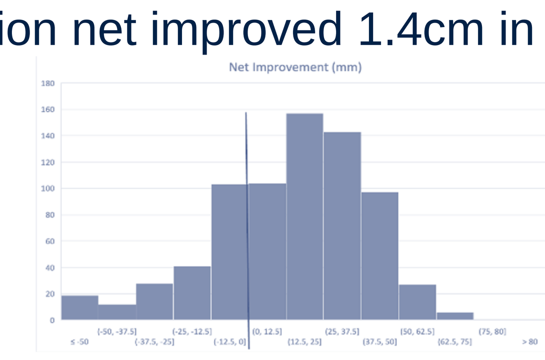


Results

- Defined usable and non-usable outputs, arranging usable to excellent, good and medium



- Mean Square Error for usable results is 3 cm
- Calibration net improved 1.4cm in average



- System can calculate 8-12 FPS

Conclusions

- Using a combination of computer vision and learning methods to create robust system.
- Errors can be minimized by Calibration fix-net
- The project yield a great easy to use tool to build gaze-based apps.

February 2021