Workshop on Biometrics

Hierarchical Network for Facial Palsy Detection

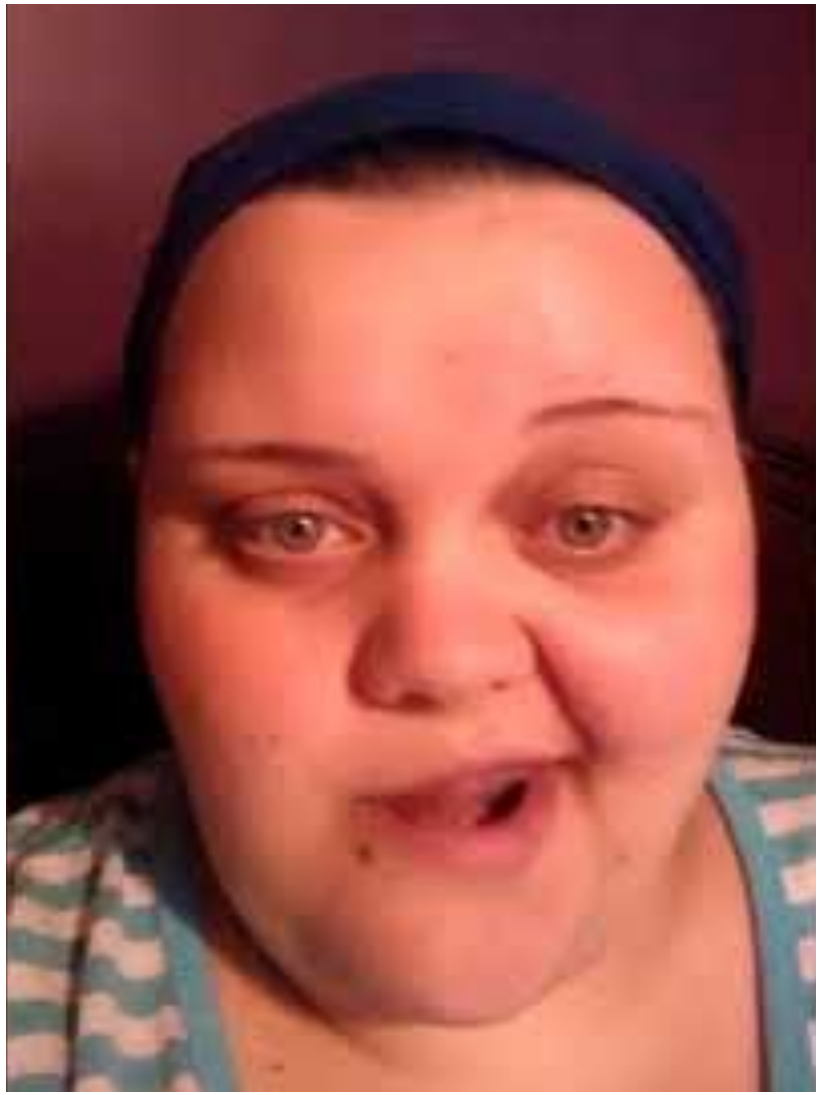
Gee-Sern Jison Hsu[1], Wen-Fong Huang[1], Jiunn-Horng Kang[2]

[1]National Taiwan University of Science and Technology, [2]Taipei Medical University

2018
CVPR
SALT LAKE CITY • JUNE 18-22

Contents

- What is Facial Palsy?
- Issues about Detection of Facial Palsy
- Proposed Solution
- Experimental Study
- Conclusion

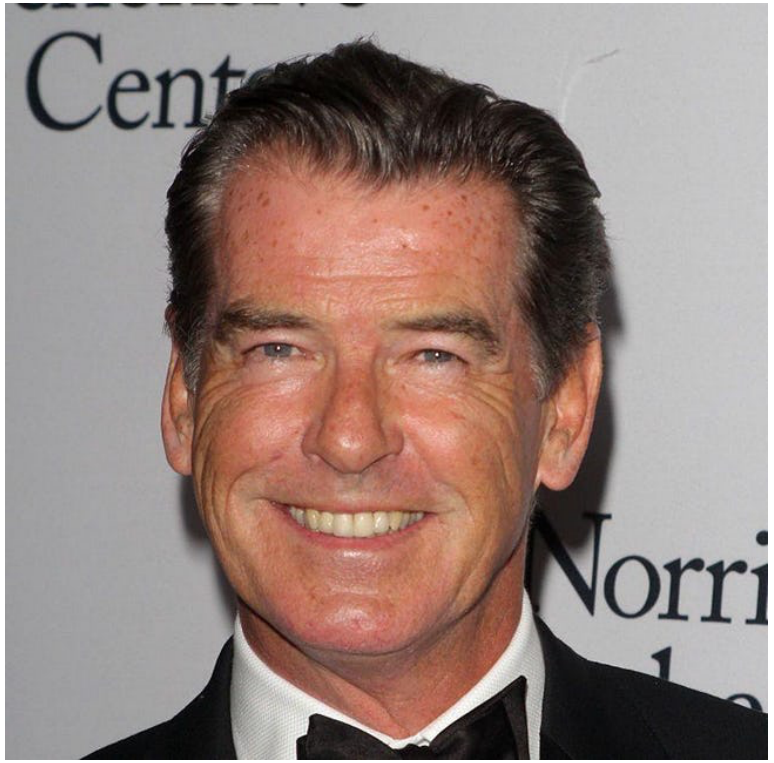


CLICK HERE

What is Facial Palsy?

- Also known as Bell's palsy, which is the most common type of facial paralysis, affecting 1/5,000 people a year.
- Temporary loss of muscle control in one side of the face, causing asymmetric facial deformation.
- Causes are not completely known, but herpes or syphilis could be responsible.
- People with diabetes and HIV are at higher risk of developing facial palsy.

Those that You Know



Pierce Brosnan

Sylvester Stallone

Angelina Jolie

Issues

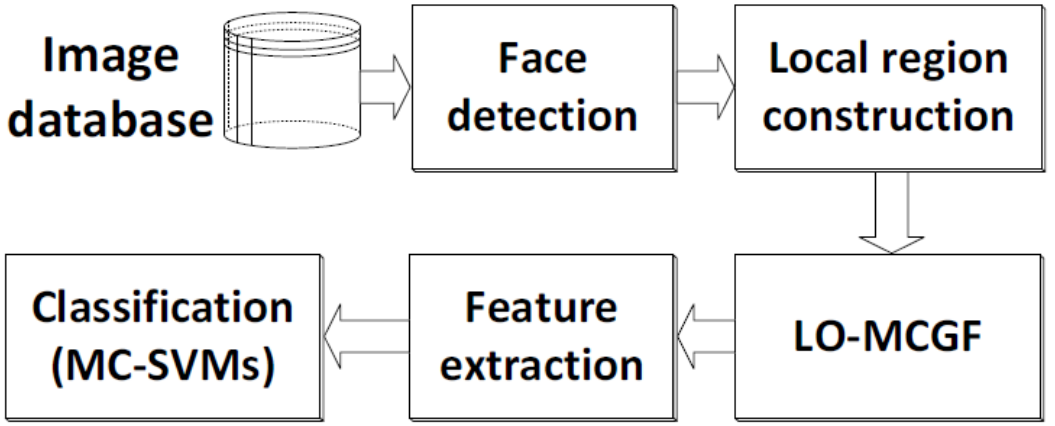
- Automatic detection/diagnosis is emerging recently; however, all use handcrafted features.
- Previous approaches were evaluated on proprietary databases, making performance comparison difficult.

Our Contributions

- First deep-learning approach for facial palsy detection, able to identify affected regions in a still image and the frequency of the syndrome in a video.
- Release of the first public database, labeled by clinicians, for facial palsy study.

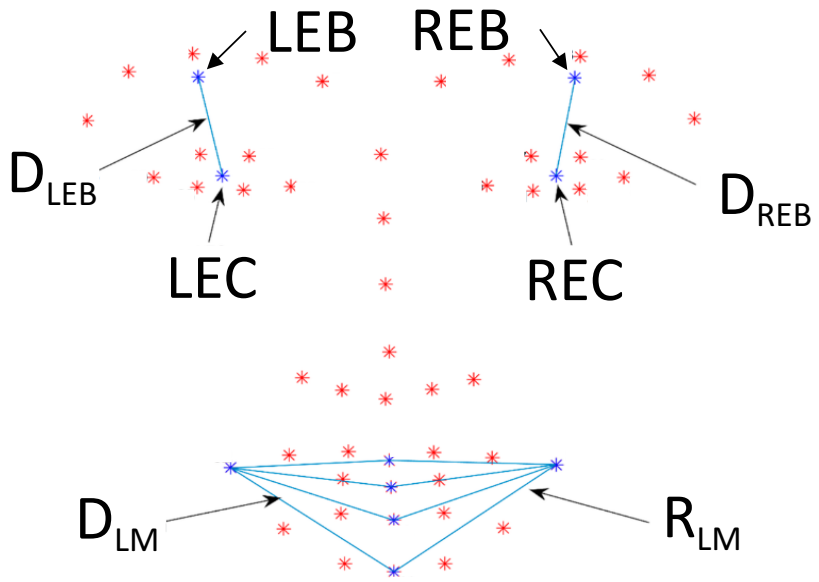
Just a Couple Years Ago

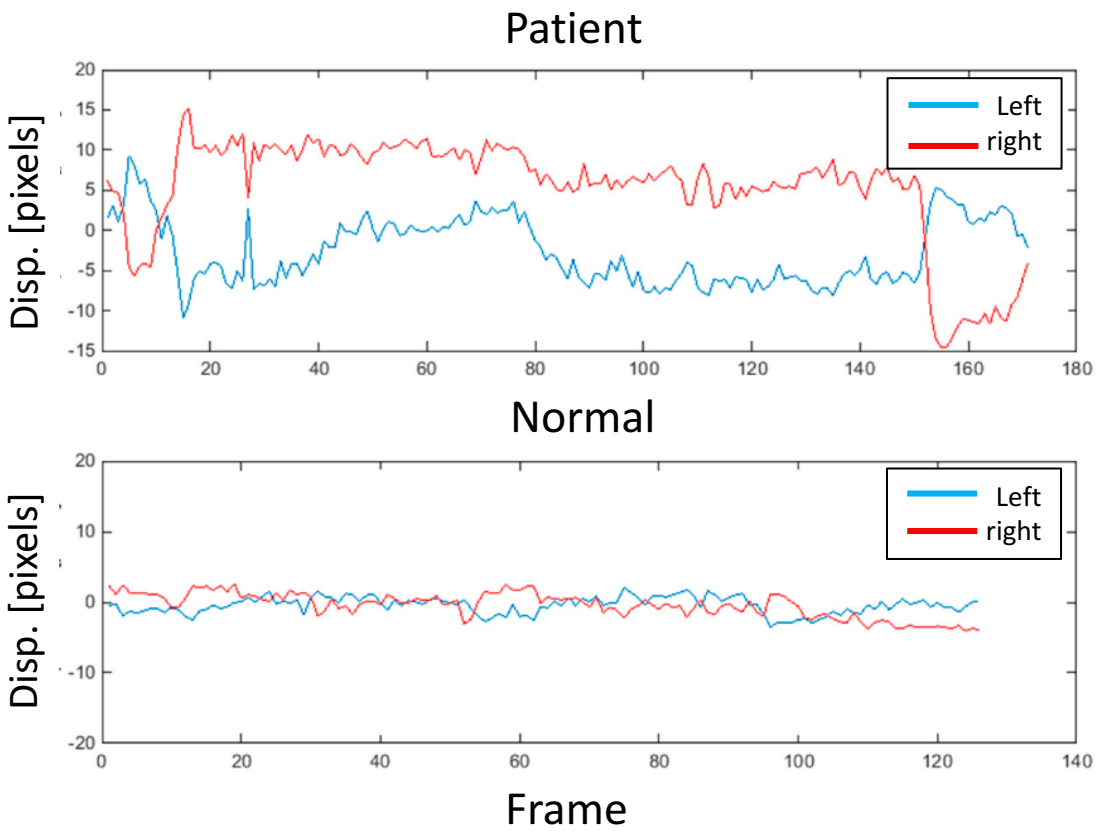
- The LO-MCGFs [2] use uniform pass-bands to enhance the desired spatial frequencies, and bounded filter support to specify the region of interest.



- The proprietary database provided by Osaka Police Hospital has 85 subjects (75 patients + 10 healthy volunteers).

[2] Ngo, Truc Hung, et al. "Quantitative analysis of facial paralysis based on limited-orientation modified circular Gabor filters." ICPR 2016.



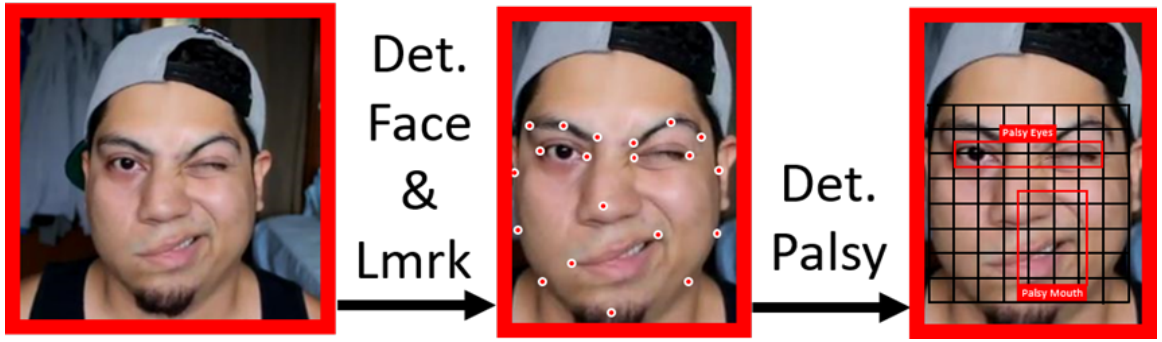


Given the facial landmarks, compute the asymmetric score using the landmark displacements at eye and mouth regions while the subjects change expression.

[3] Kim, Hyun Seok, et al. "A smartphone-based automatic diagnosis system for facial nerve palsy." Sensors 2015.

Proposed Network with 3 Component Nets

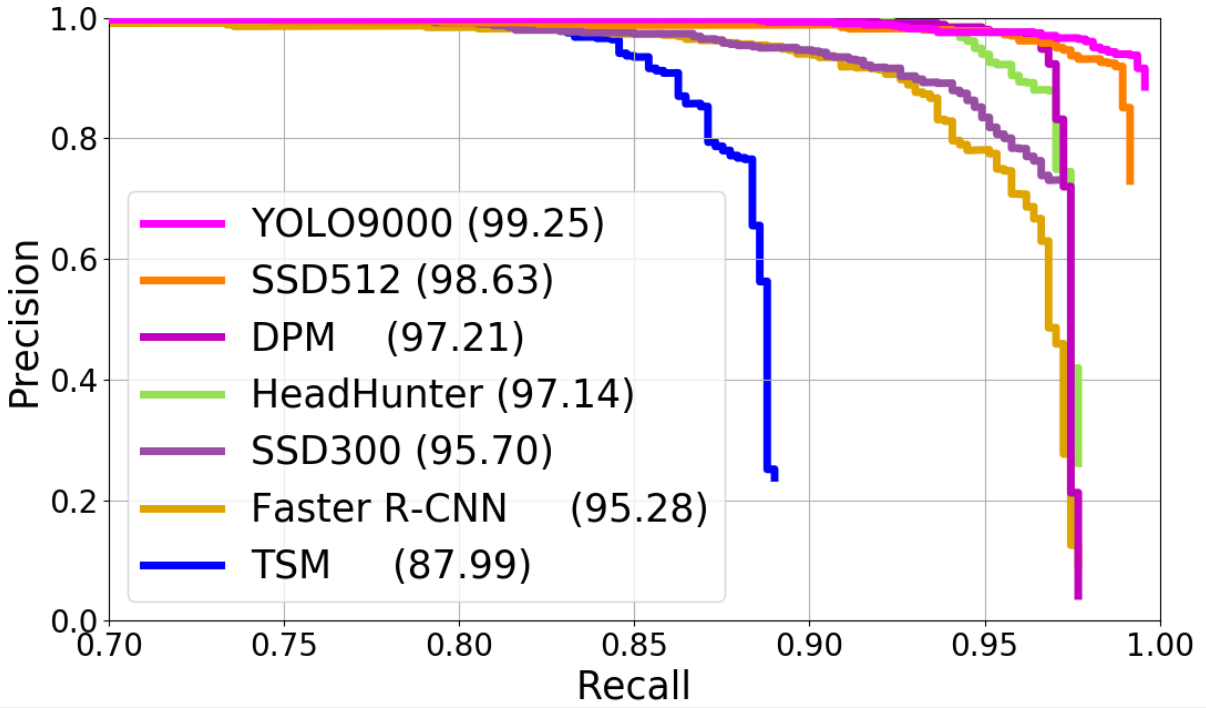
- FaceNet
  - YOLO-9000, retrained on WiderFace, yielding 99.25% AP on AFW.
- LandmarkNet
  - H.G. Face Alignment Network (ICCV 17) for locating facial landmarks.



- PalsyNet
  - Landmark-defined grid of 8×8 cells covering all possible palsy regions.

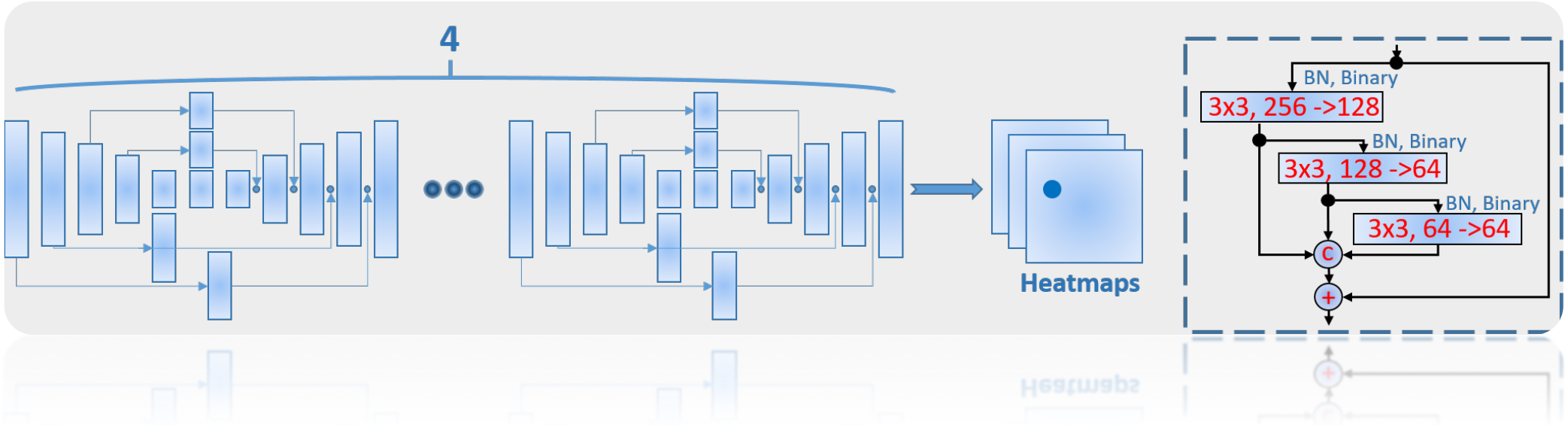
FaceNet Net_f

- Built on YOLO-9000 and retrained on the WiderFace database (393,703 labeled faces)
- AP 99.25% on AFW benchmark, better than DPM (97.2%), HeadHunter (97.1%), SSD-512 (98.6%) and Faster RCNN (95.3%).



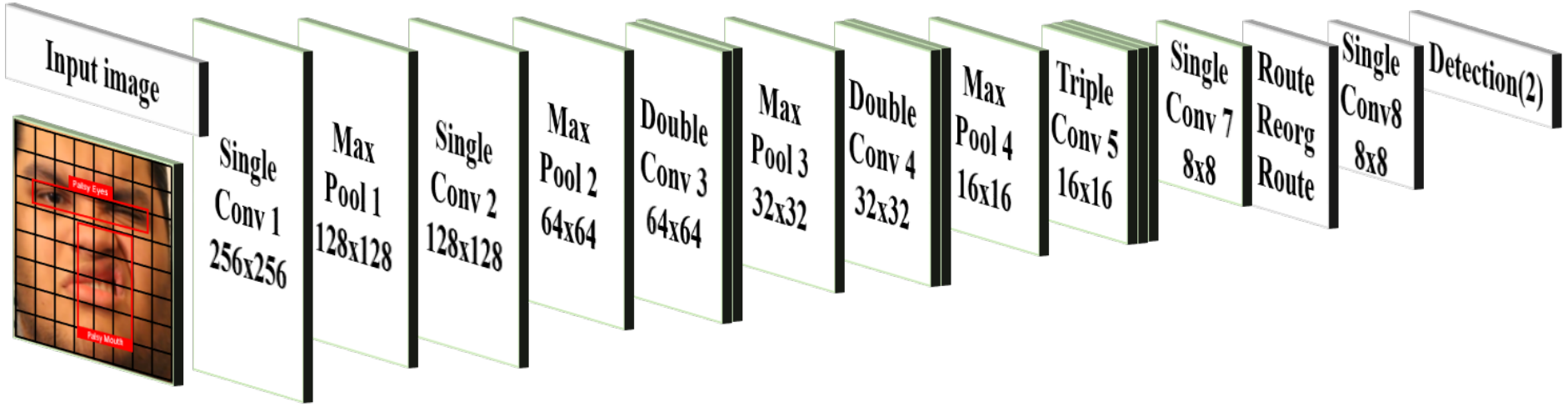
LandmarkNet Net_m

- LandmarkNet :
  - Face Alignment Network (FAN) built on Hour-Glass with bottleneck block replaced by residual block.
  - Hour-Glass consists of multi stacked hourglass modules allowing for repeated bottom-up, top-down inference.



PalsyNet Net_p

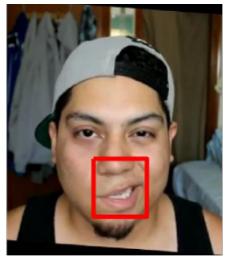
- Fewer layers: 11 convolution layers and 4 max-pooling layers (v.s. 19 convolution layers and 5 max-pooling in Darknet-19).
- Tailor-made anchor boxes for fast locating of palsy regions.

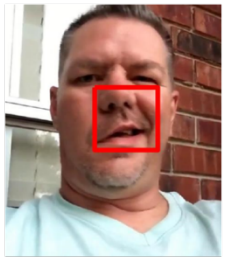


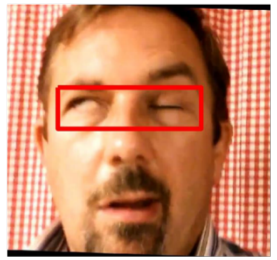
Facial Palsy Dataset

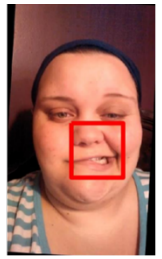
- 32 videos of 21 patients from YouTube;
- As the shortest facial palsy session lasts for a second, we convert each video into an image sequence with 6FPS;
- Manually labeled local palsy regions when the deformation intensity was considered sufficiently high by clinicians.

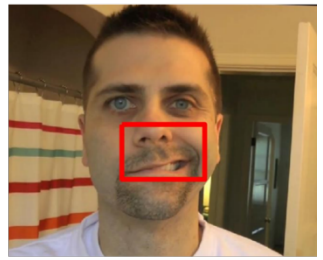
Facial Palsy Dataset

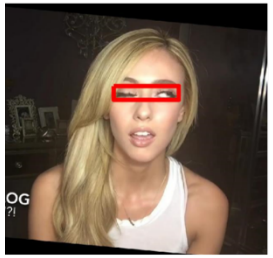

Person 1.

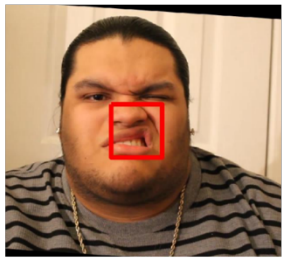

Person 2.

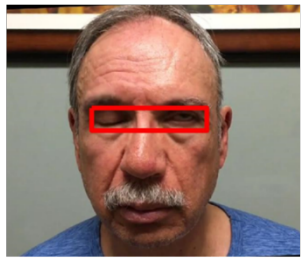

Person 3.

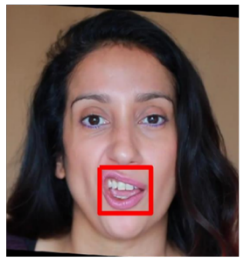

Person 4.

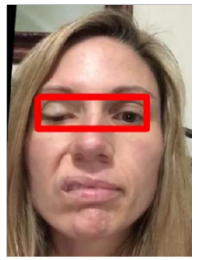

Person 5.

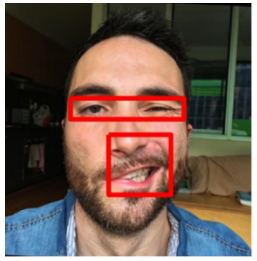

Person 6.

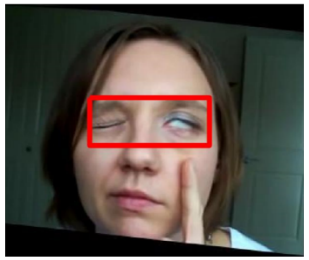

Person 7.


Person 8.

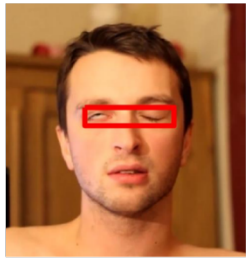

Person 9.

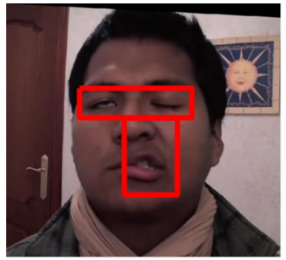

Person 10.

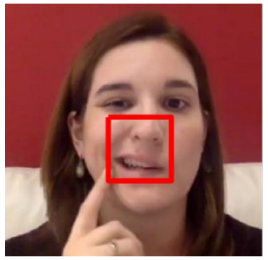

Person 11.

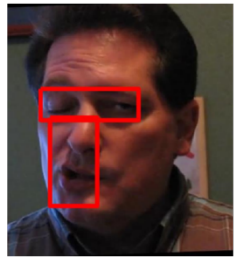

Person 12.

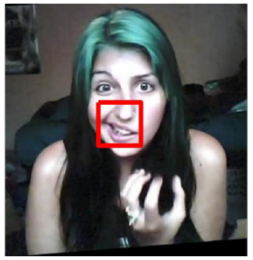
Person 13.

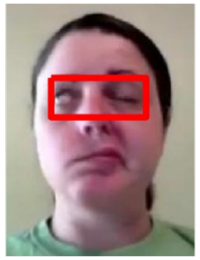

Person 14.

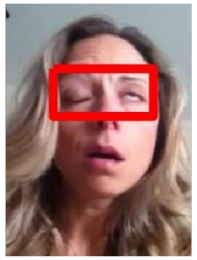

Person 15.

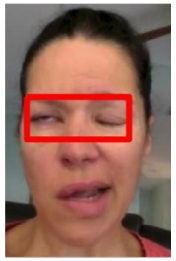

Person 16.


Person 17.


Person 18.


Person 19.


Person 20.


Person 21.

Data Labeling

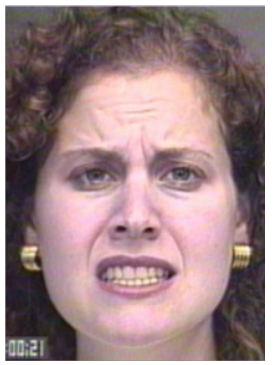
CK+ Included for Robustness

123 subjects, 593 image sequences with 7 expressions

0-Neutral

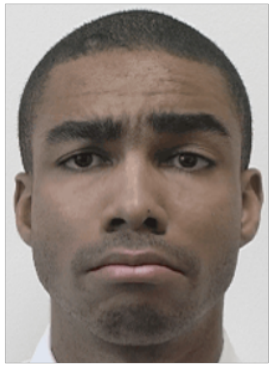
1-Angry

2-Contempt


3-Disgust

4-Fear

5-Happiness

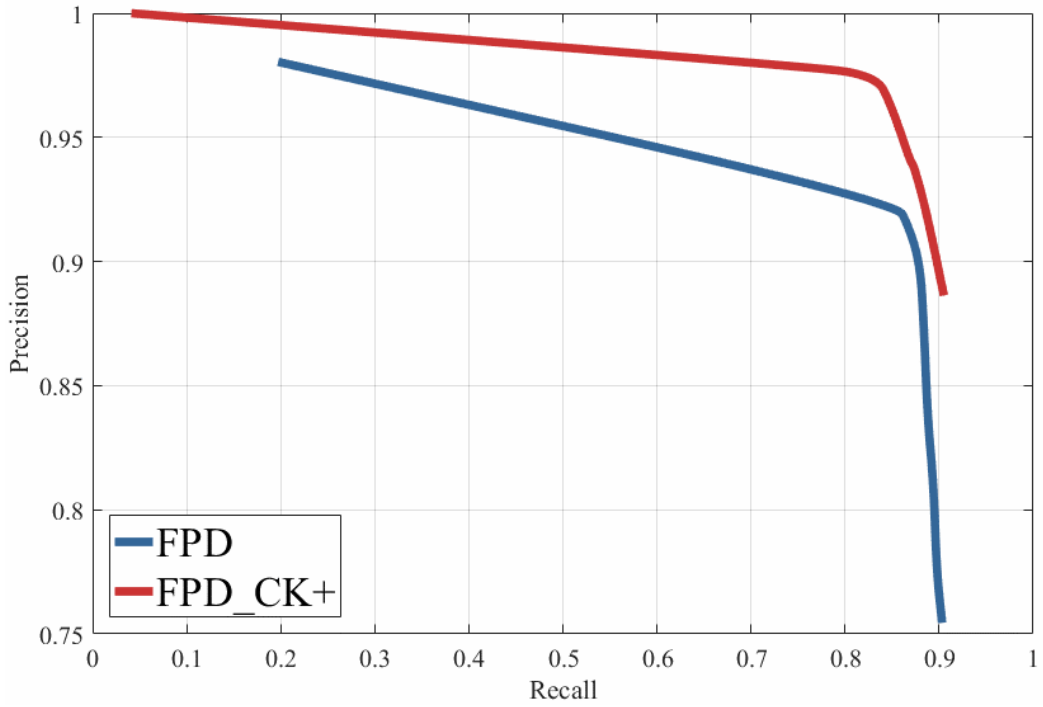

6-Sadness

7-Surprise

CLICK HERE

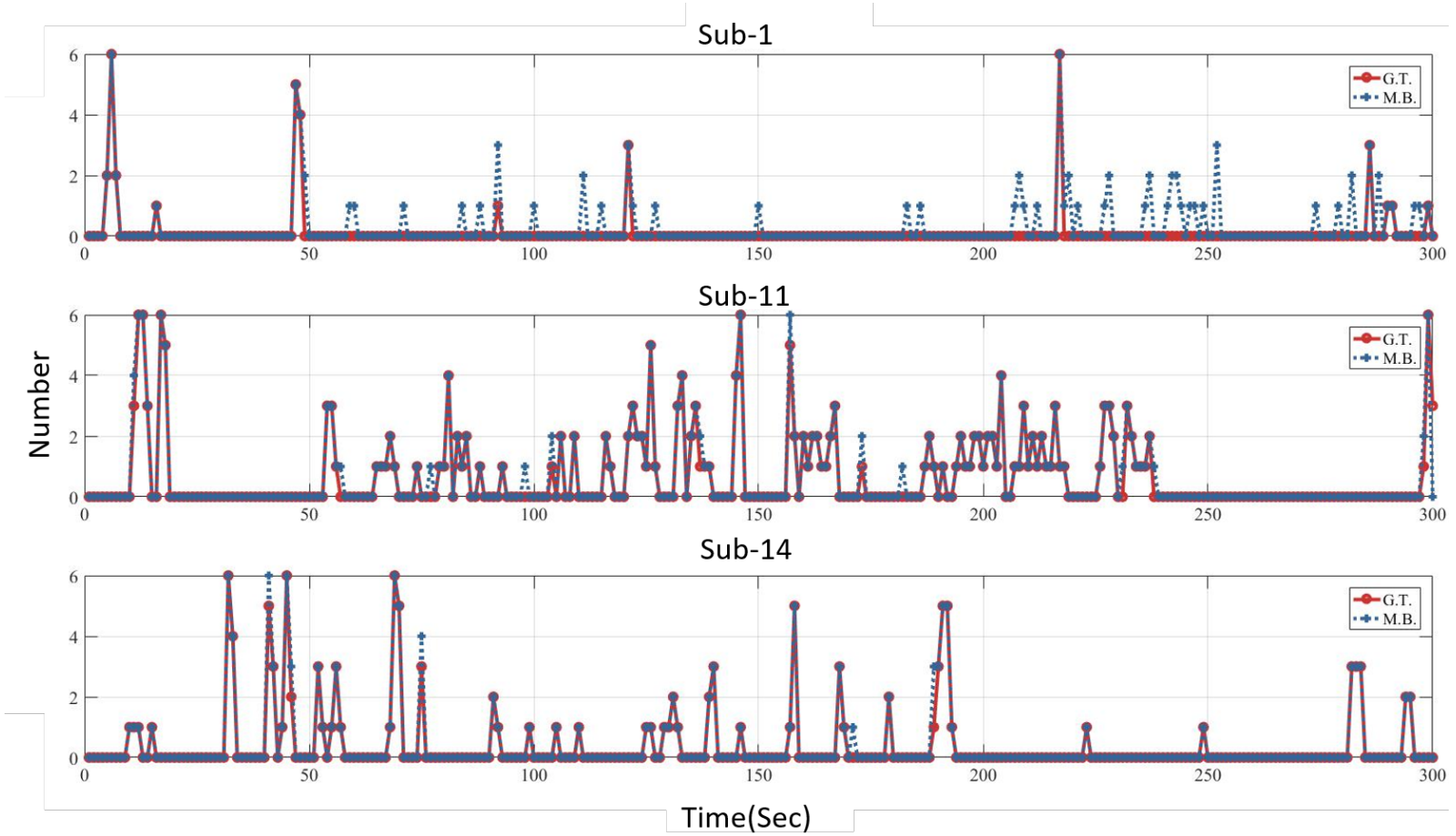
Performance Evaluation

- 21 leave-one-out on FPD and 5-fold cross validation on CK+
- Performance with and without CK+ expression dataset.



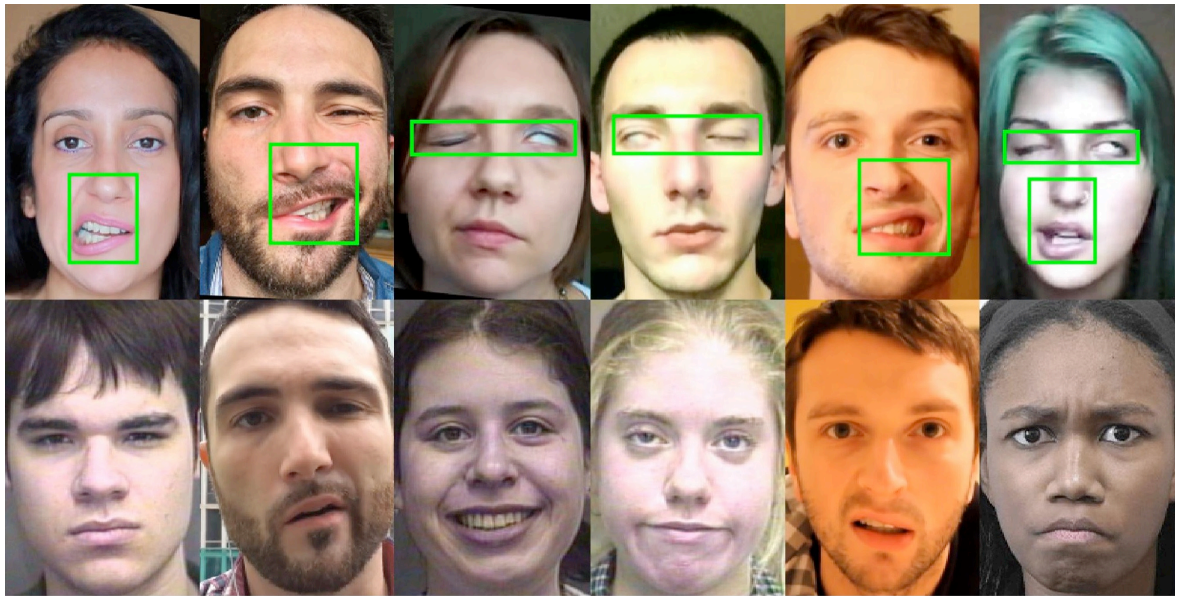
Frequency of Occurrence

- Detection of palsy-caused deformation over time when patients were talking.



Conclusion

- Identification of facial palsy formulated as detection of local regions, we propose a top-down hierarchical framework as a solution.
- A practical case study for the application/modification of deep learning framework for auto detection/diagnosis of (facial) medical disorders.



Our Team

Gee-Sern Jison Hsu
National Taiwan University of Science and Technology
jison@mail.ntust.edu.tw

Jiunn-Horng Kang
Taipei Medical University
jhk@tmu.edu.tw

Wen-Fong Huang
National Taiwan University of Science and Technology
m10503416@mail.ntust.edu.tw