Emotions and emotion detection

"Emotions can save our lives, but they can also cause terrible harm" Paul Ekman

dirk.Labudde@hs-Mittweida.de

☎+49 3727 58 1077

D. Labudde

Hochschule Mittweida, University of Applied Sciences
Technikumplatz 17, D-09648 Mittweida

A interactive Introduktion - Paul Ekman

1967-1968
Universal Facial Expressions
Dr. Ekman travels to Papua New Guinea to study the nonverbal behavior of the Fore people. He chose these people as they were an isolated, Stone Age culture located in the South East Highlands. Ekman's research provided the strongest evidence to date that facial expressions are universal.

https://www.youtube.com/watch?time_continue=1&v=zwNwaaP5hV0&feature=emb_logo

2018

More Information on: https://www.paulekman.com/about/paul-ekman/

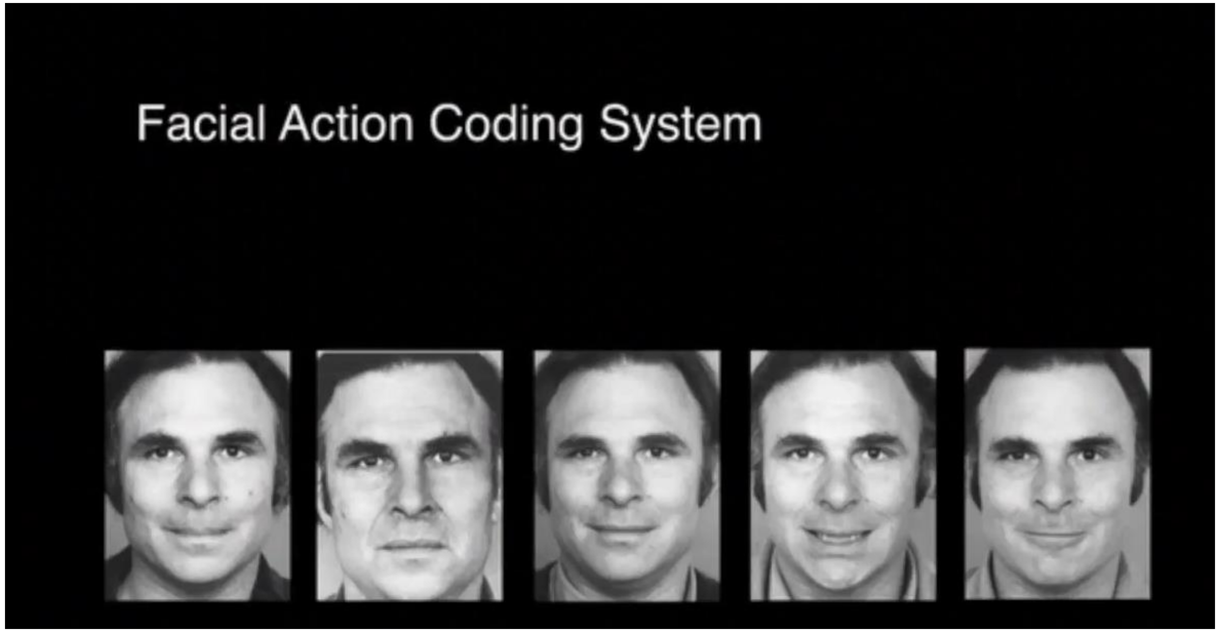
A interactive Introduktion - Paul Ekman

1972-1978

Coding the Face

Dr. Ekman's findings inspired the development of the Facial Action Coding System. FACS was the first and only comprehensive tool for objectively measuring facial movement. Ekman developed this tool along with W. Friesen in 1978 and later revised it in 2003 with J. Hagar as a third author.

https://www.youtube.com/watch?time_continue=24&v=6RzCWRxnc84&feature=emb_logo



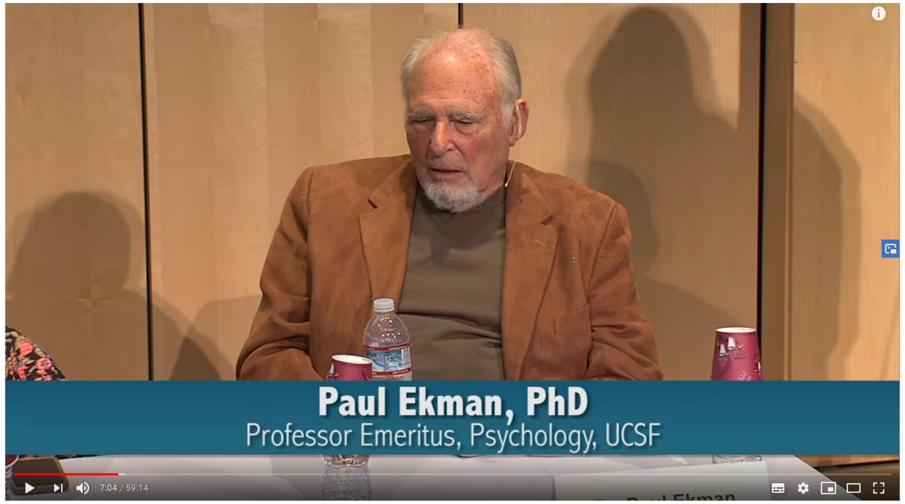
More Information on: https://www.paulekman.com/about/paul-ekman/

A interactive Introduktion - Paul Ekman

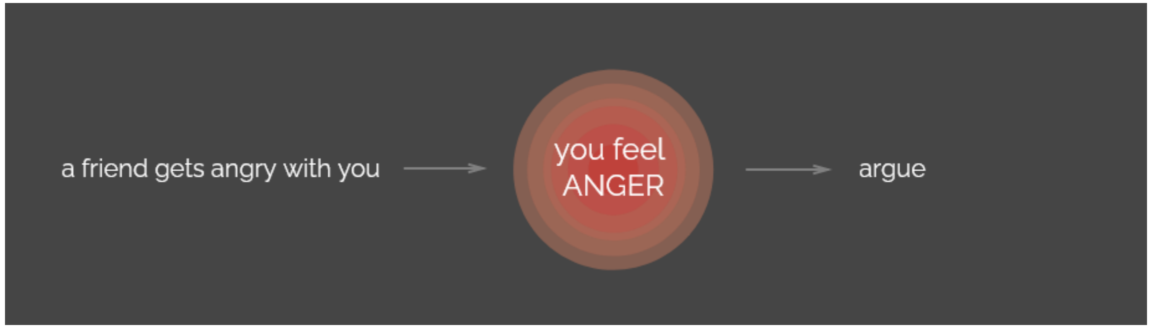
2016

Atlas of Emotions

The Dalai Lama imagined "a map of our emotions to develop a calm mind." He asked his longtime friend and renowned emotion scientist Dr. Paul Ekman to realize his idea. Ekman took on the creation of the Atlas alongside his daughter, Dr. Eve Ekman, a second-generation emotion researcher and trainer. The Atlas of Emotions represents what researchers have learned from the psychological study of emotion.

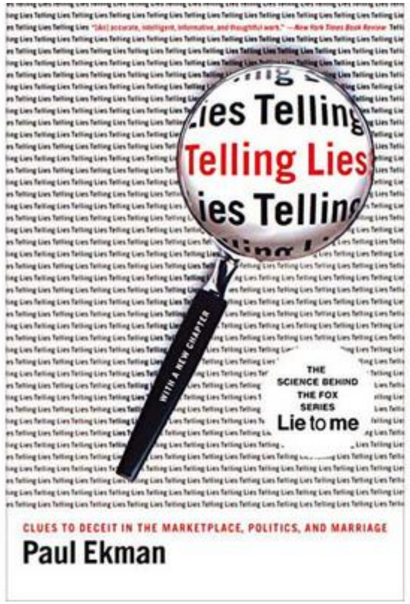


https://www.youtube.com/watch?v=AaDzUFL9CLE

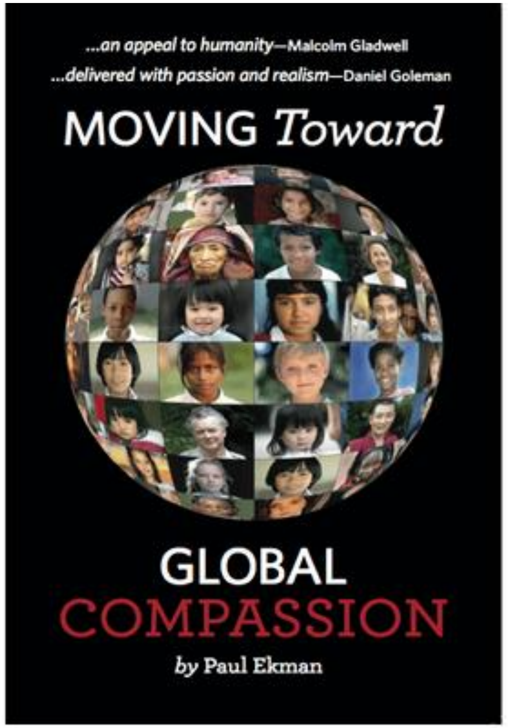


http://atlasofemotions.org

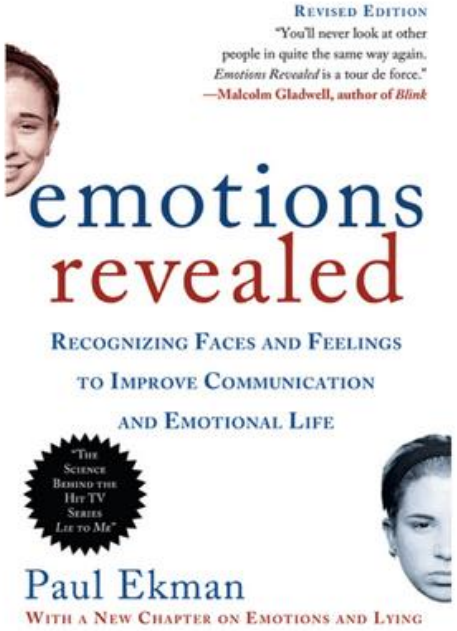
Books (most popular)



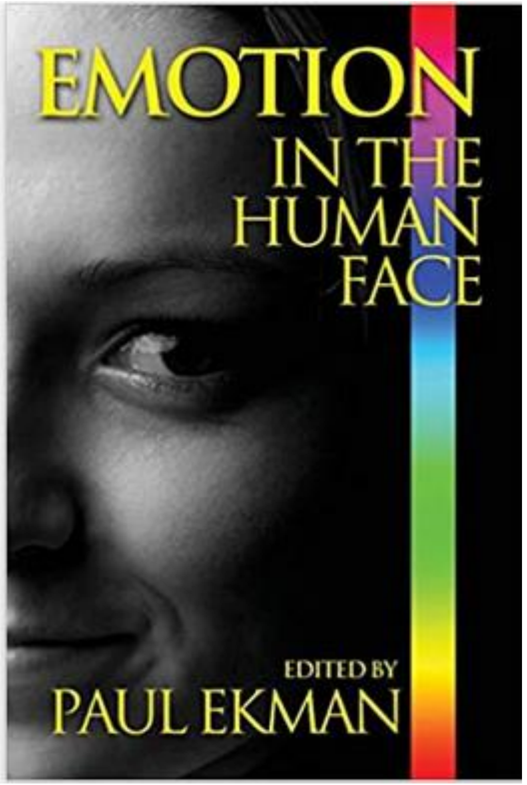
Telling Lies



Moving Toward Global Compassion



Emotions Revealed



Emotion in the Human Face, originally published in 1972, was the first volume to evaluate and integrate all research on facial expression of emotion since Darwin published

Emotions and emotion detection

Can we use the FACS together with AI today?

Facial Expression Recognition FER - *analog*

Cognitive analysis of a video:

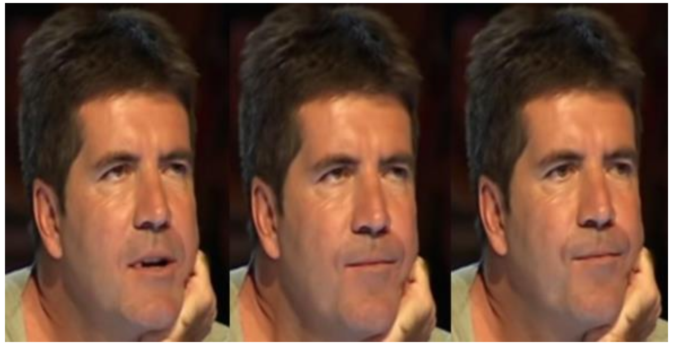
- watch the video
  - observe the main actor
  - observe the jury
  - watch the audience

- Pay attention to the interactions of all

https://www.youtube.com/watch?v=dnp-8GrHOIk

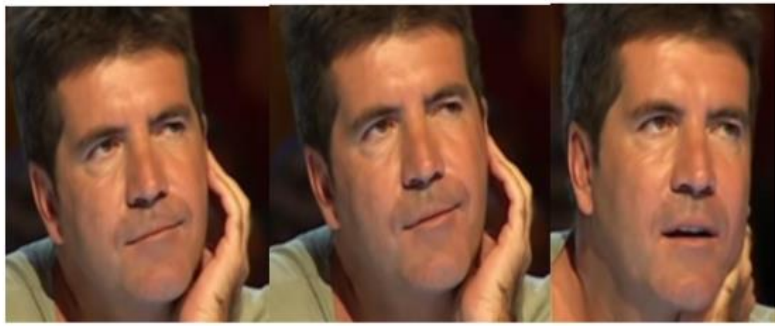
Facial Expression Recognition FER - *analog*

People follow this scene
Observer / Viewer



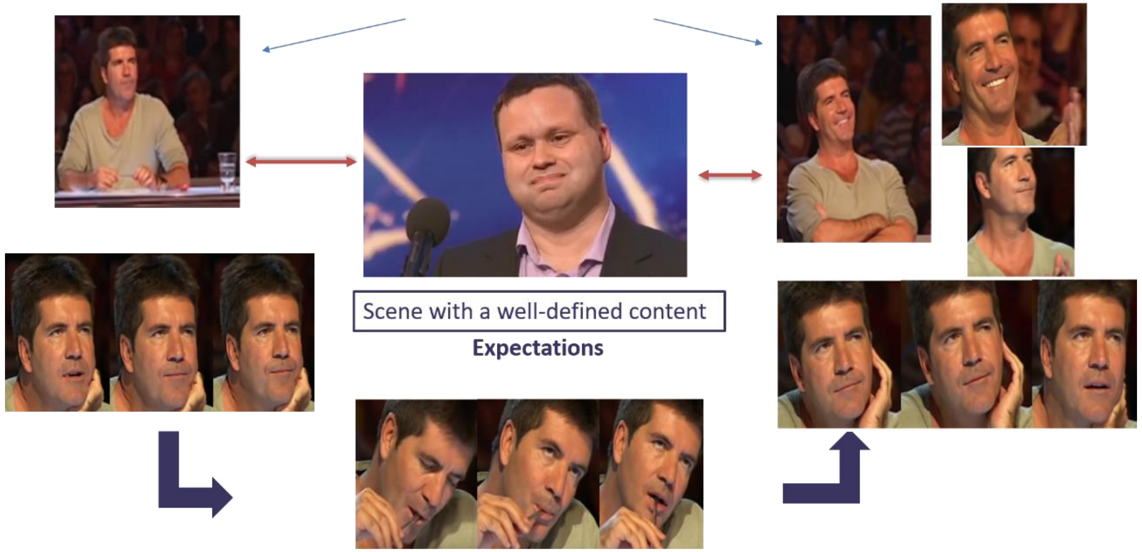
Scene with a well-defined content

Expectations



Facial Expression Recognition FER - *analog*

People follow this scene
Observer / Viewer



The individual actors interact (red arrows) and show different emotions (figures).
These emotions are evoked by other actors or the scene.

For details see the videos on: **Atlas of Emotions**

Facial coding

One of the strongest indicators for emotions is our face. As we laugh or cry we're putting our emotions on display, allowing others to glimpse into our minds as they "read" our face based on changes in key face features.

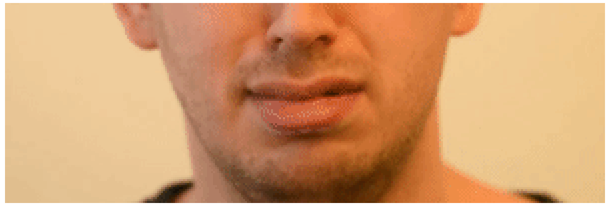
Computer-based facial expression analysis mimics our human coding skills quite impressively as it captures raw, unfiltered emotional responses towards any type of emotionally engaging content. These expressed emotional states are detected in real time using fully automated computer algorithms that record facial expressions via webcam.

Facial Expression Recognition FER - *analog*

facial movements - musculature (muscle groups) - facial expressions

AU 11

Pull the outer part of the upper lip diagonally upwards



zygomaticus minor

AU 25

Opening the lips

Depressor labii inferioris or relaxation of the mentalis or Orbiscularis oris

AU 43

Close the eyes by lowering the upper eyelid

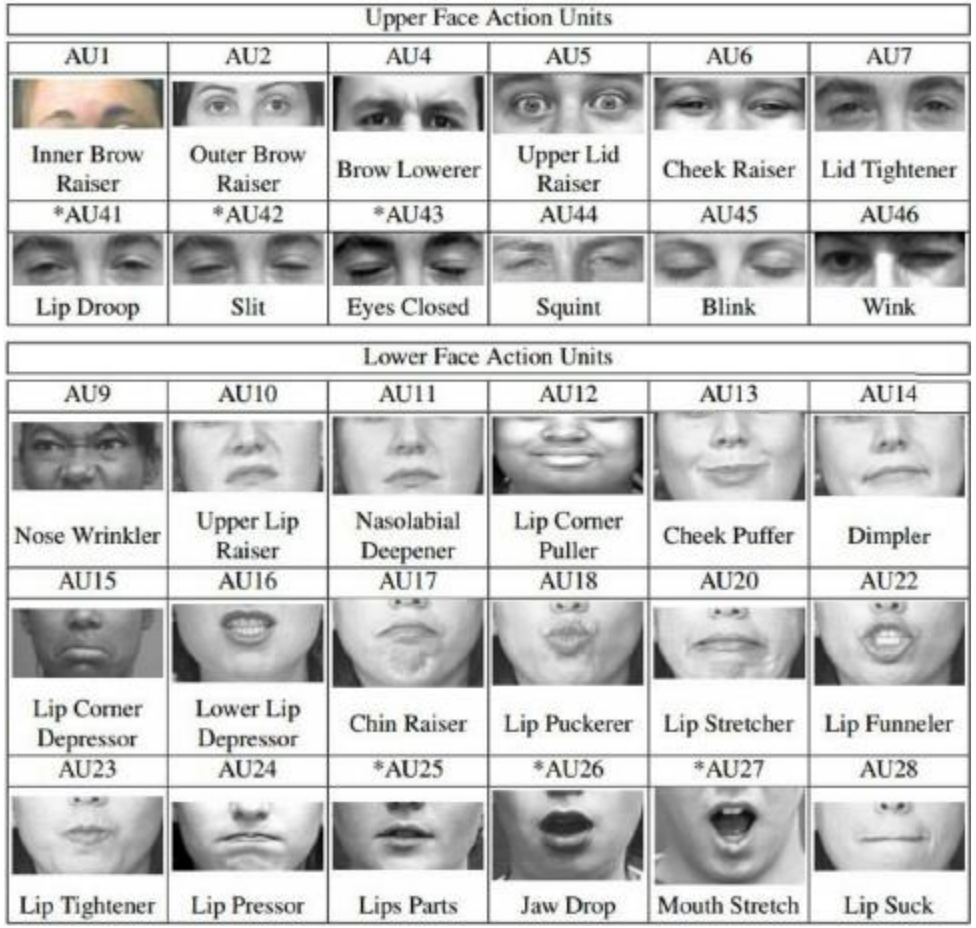
Relaxation of the levator palpebrae superioris; orbicularis oculi, pars palpebralis

***Facial Action Coding System* (FACS)**

For details see the videos on: **Coding the Face**

All visually recognizable face movements can be described with the 44 Action Units.

Facial Action Coding System (FACS)



Facial Action Units (AUs) of upper and lower face

Facial Expression Recognition FER - *analog*

"Emotions have subjective components that can be experienced and objective components that can be recorded, that accompany or promote goal-oriented behavior that enables the organism to adapt to its living conditions".

- Subjective component (feeling)
- Cognitive component (perception and interpretation)
- **Communicative component** (facial expressions, gestures, voice pitch)
- Physiological component (heart and respiratory rate, adaptation of blood vessels)
- Motivational component (energy, behavioral preparation)

facial movements - musculature (muscle groups) - facial expressions

↕

Emotions

Emotions are visible in the face and evoke new emotions in other people.

Facial Expression Recognition FER - *analog*

Facial expression as part of emotions

Basic emotions: Sadness, anger, fear, disgust, contempt, surprise and joy.

***Facial Action Coding System* (FACS)**

Composition (combination) of Action Units

7 basic emotions are universal and show the same characteristics for most people in the world

For details see the videos on **: Universal Facial Expressions**

Facial Expression Recognition FER - *analog*

Facial expression as part of emotions

Basic emotions: Sadness, anger, fear, disgust, contempt, surprise and joy.

1 - Lifting of the eyebrows inside
2 - Lifting of the eyebrows outside
5 - Lifting the upper eyelid
26 - Opening the mouth by relaxing the
Lower jaw musculature

- Anger: 4CDE + 5CDE + 7CDE + 17 + 23 + 24
- Disgust: 9 + [10 und/oder 16] + 19 + 26
- Fear: 1 + 2 + 4 + 5ABCDE + 7 + 20ABCDE + 26
- Joy: [6 und/oder 7] mit 12CDE
- Sadness : 1 + 4 + [6 und/oder 7] + 15ABC
- Surprise: 1CDE + 2CDE + 5AB + 26

AU 1,2,5

AU26

- *1CDE* und *2CDE : eyebrows raised with a medium to maximum intensity*
- *5AB: upper eyelids slightly raised*

***Facial Action Coding System* (FACS): Number = Action Unit and A-E = increasing intensity**

Video and content

People follow a scene
- **Observer**
- **Viewer**

People act in this scene
Acting persons

Scene with a well-defined content

- **Emotions**
- **Language (paraverbal)**
- **Screams**
- **Body posture (non-verbal)**
- **Facial expression/gestures**
- **Statements**

Witness Interview:
- by the court
- by the police

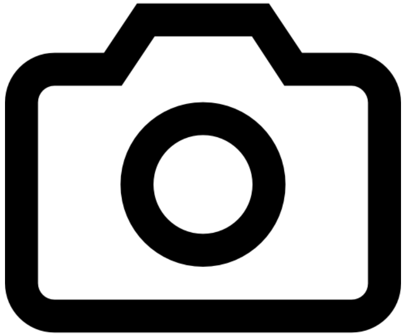
Communication Models

verbal, non-verbal and paraverbal consistency

Development over time

Analysis of videos and their content

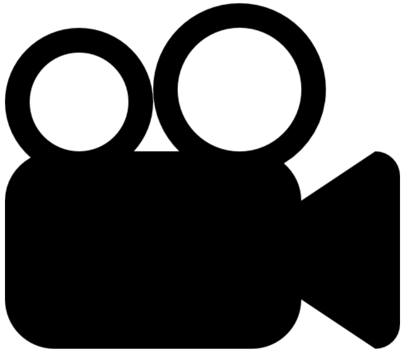
Conflicting information can occur between facial expressions, gestures, spoken words, voice, etc.!



Photos/Pictures

What is producible/ detectable?

State of the art



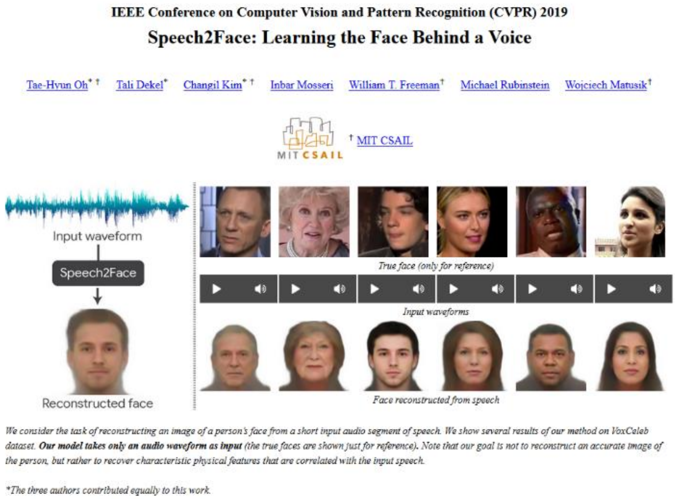
Videos/Films

video ' I see' --- videre ' see

Important components

- Body language and posture (gestures)
- Voice
- Scene
- Environment

Social Engineering


IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2019

Speech2Face: Learning the Face Behind a Voice

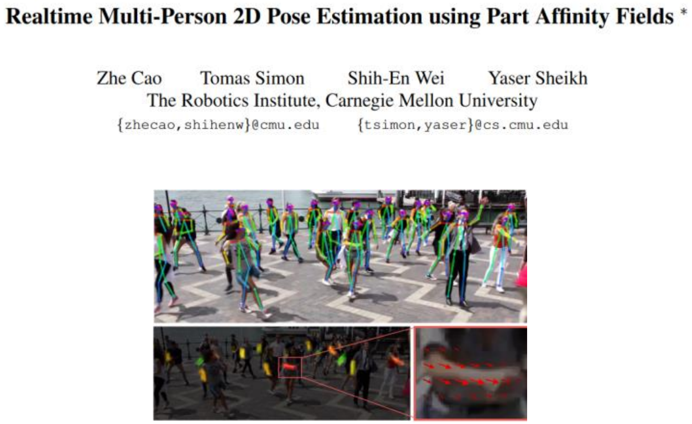
Tae-Hyun Oh[*†] Tali Dekel[*] Changil Kim[*†] Inbar Mosseri William T. Freeman[†] Michael Rubinstein Wojciech Matusik[†]

[†] MIT CSAIL

We consider the task of reconstructing an image of a person's face from a short input audio segment of speech. We show several results of our method on VoxCeleb dataset. **Our model takes only an audio waveform as input** (the true faces are shown just for reference). Note that our goal is not to reconstruct an accurate image of the person, but rather to recover characteristic physical features that are correlated with the input speech.

*The three authors contributed equally to this work.

Speech2Face

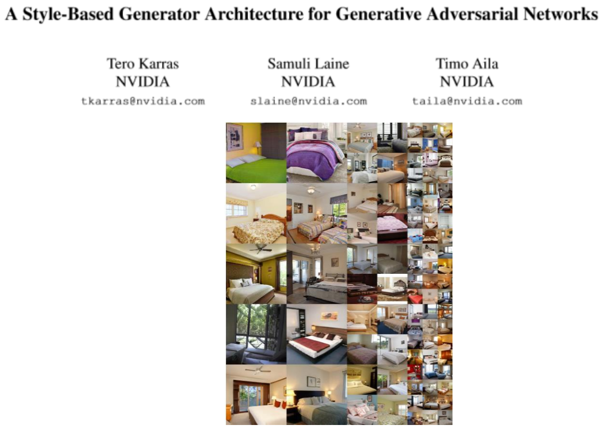

Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields *

Zhe Cao Tomas Simon Shih-En Wei Yaser Sheikh

The Robotics Institute, Carnegie Mellon University

{zhecao,shihenw}@cmu.edu {tsimon,yaser}@cs.cmu.edu

Pose Estimation


A Style-Based Generator Architecture for Generative Adversarial Networks

Tero Karras NVIDIA tkarras@nvidia.com

Samuli Laine NVIDIA slaine@nvidia.com

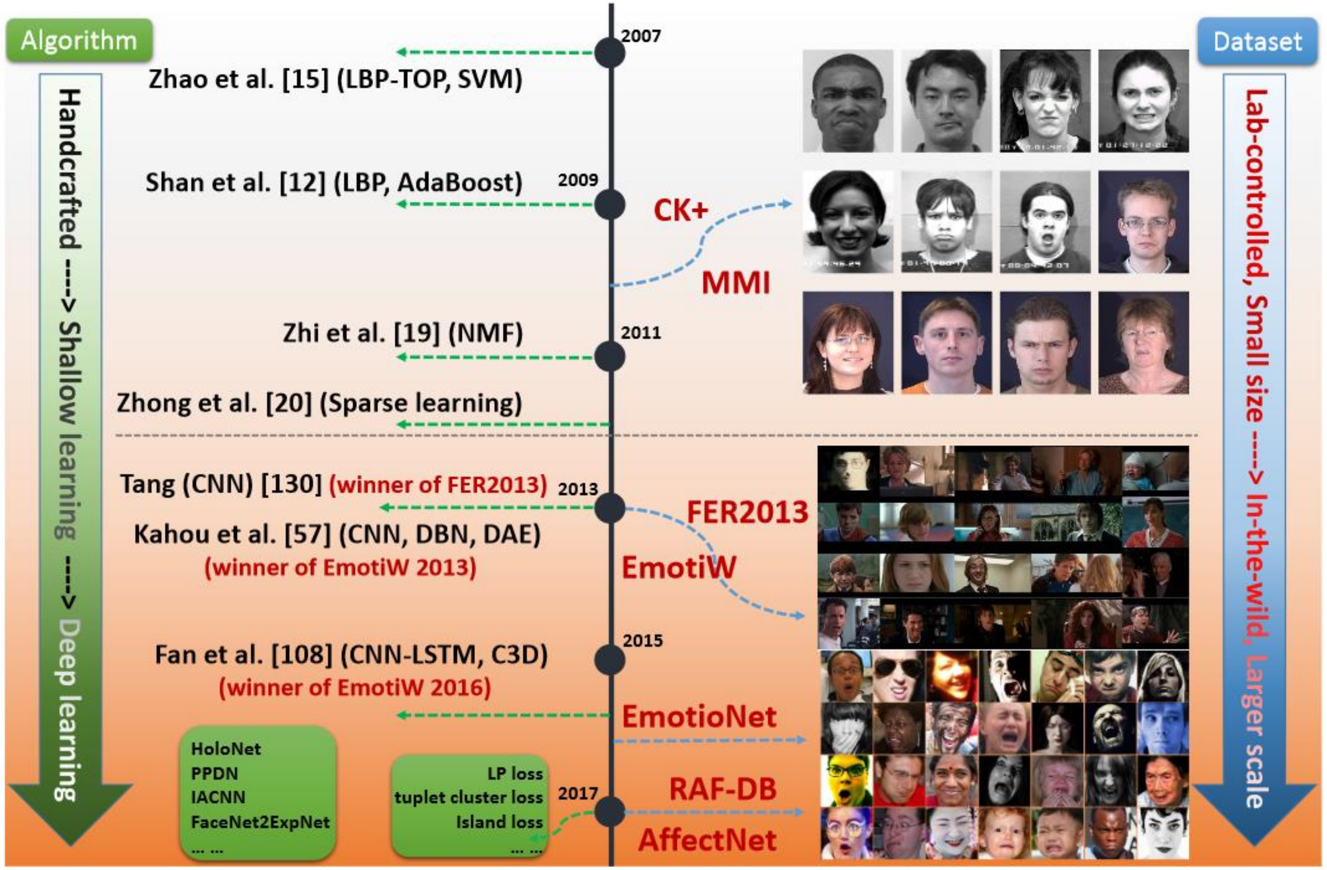
Timo Aila NVIDIA taila@nvidia.com

Style-Based Generator

https://www.youtube.com/watch?v=kSLJriaOumA

Emotions

Imagined by a GAN (generative adversarial network). StyleGAN (Dec 2018) — Karras et al. and Nvidia



2018

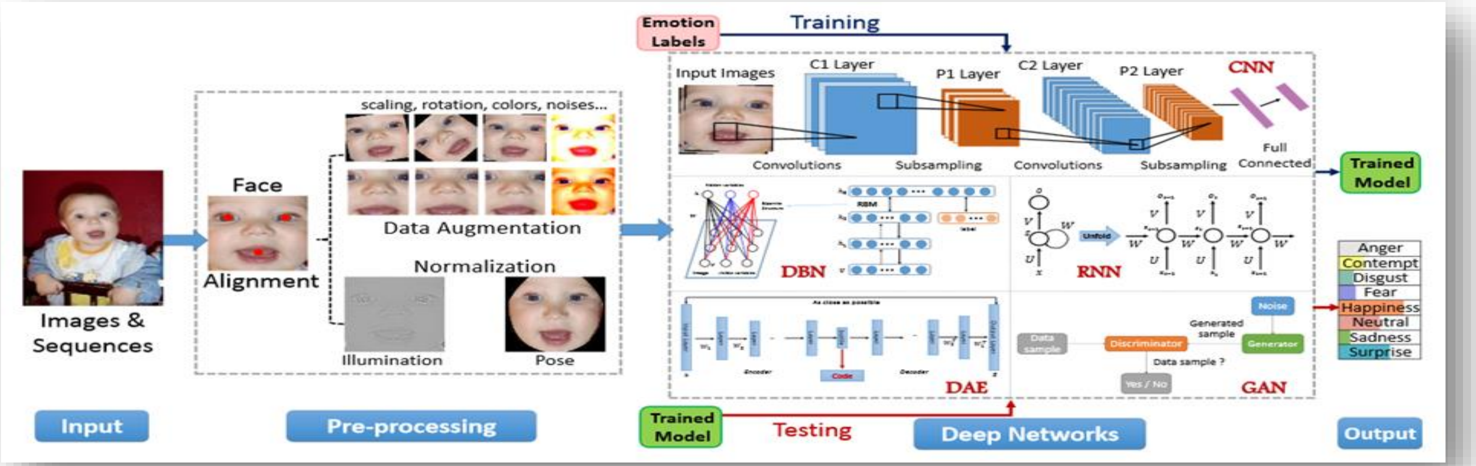
Deep Facial Expression Recognition: A Survey

Shan Li and Weihong Deng*, *Member, IEEE*

Deep Facial Expression Recognition (Deep FER)

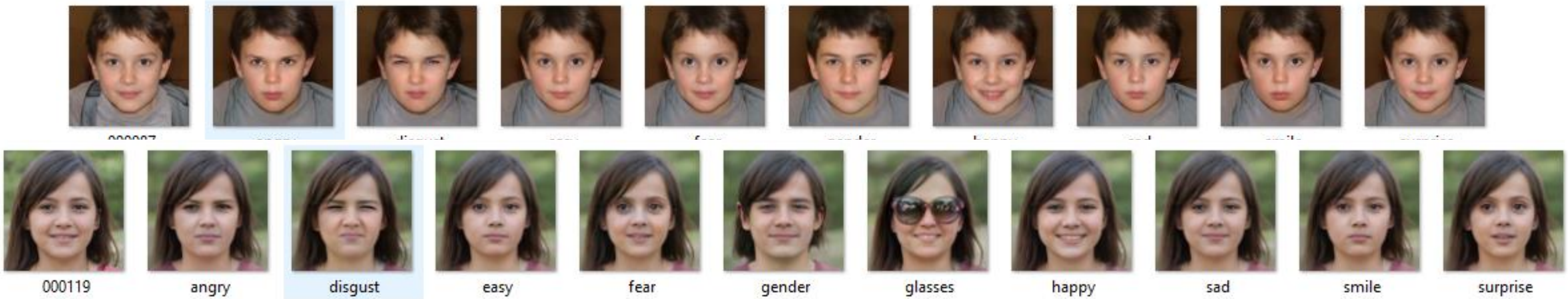
Deep FER: Neural Networks (NNs) learn facial expressions to emotions and can then recognize them on images or in real time

https://arxiv.org/pdf/1804.08348.pdf



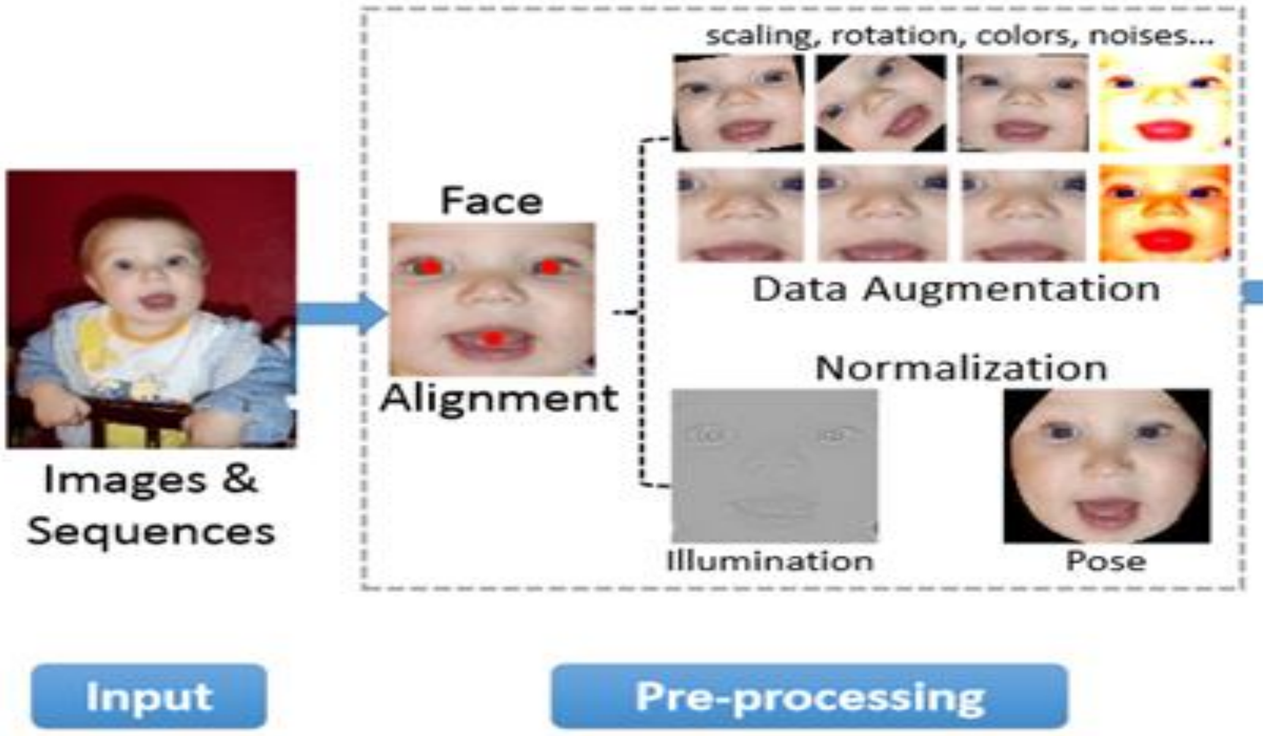
Emotions ....

Deep Facial Expression Recognition (Deep FER)



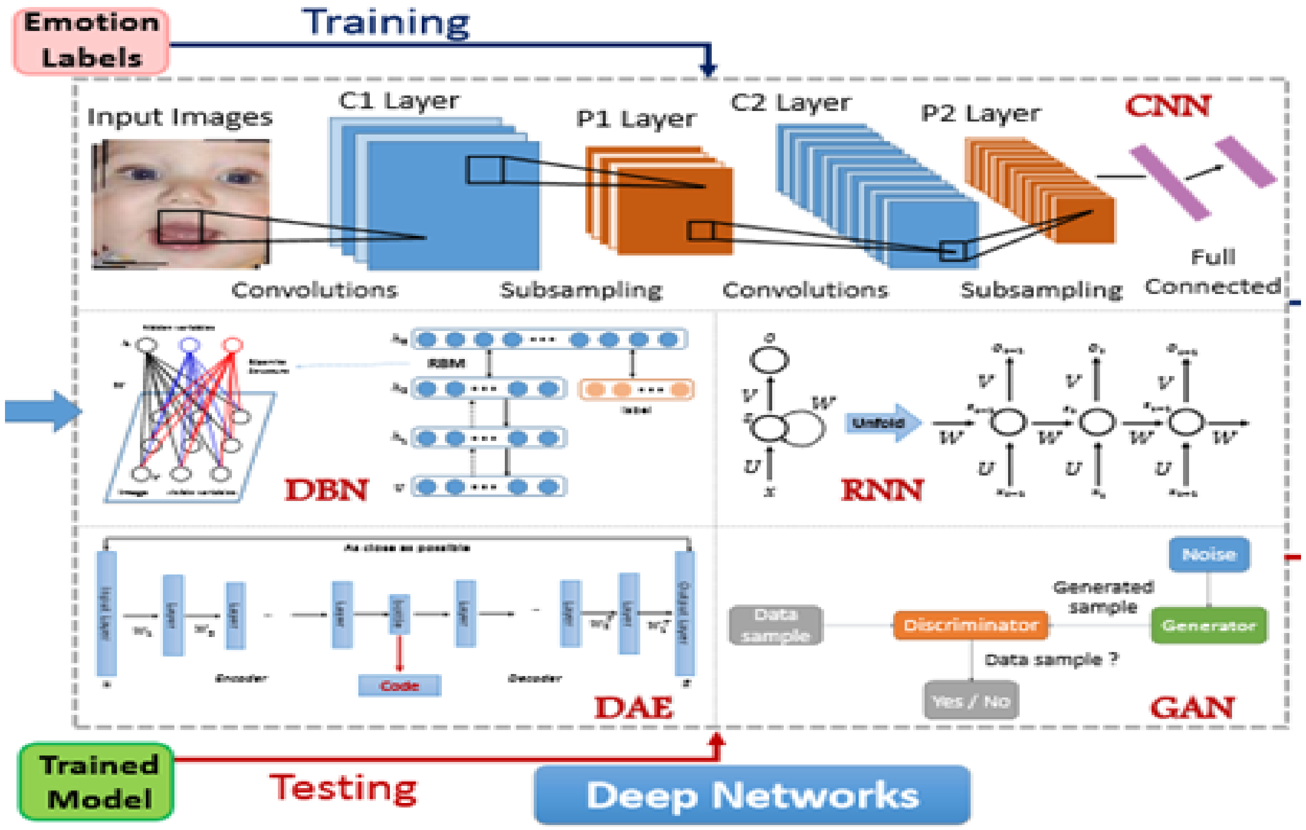
Step 1: Pre-processing

Face Alignment, Data Augmentation, Normalization



Face Alignment:
The face in the image is detected by means of certain aspects.
Data Augmentation:
Modification of the images by e.g. rotation, other scaling or changing the contrasts and colors. This allows the data set to be enlarged without the need for additional photos.
Illumination and Pose Normalization:
Solves the problem of different exposures and poses.

Step 2: Selection of the neural network



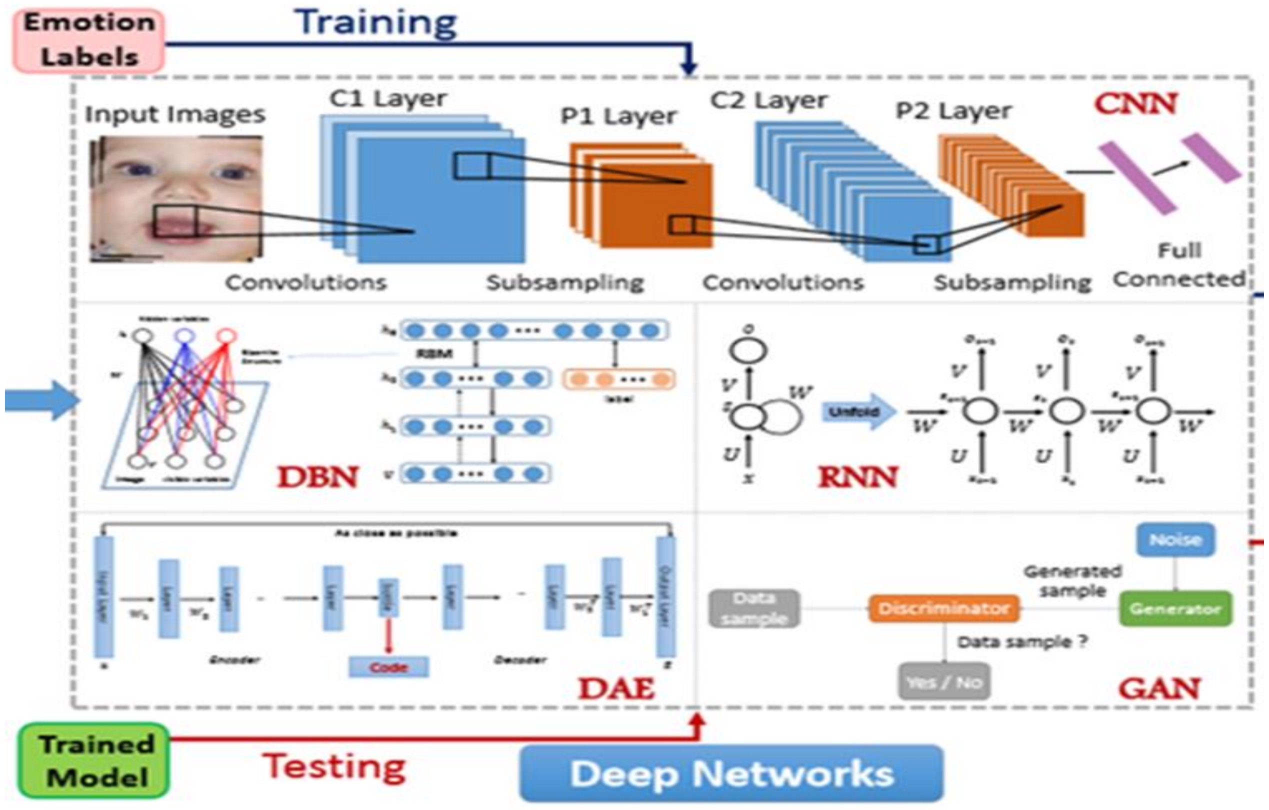
https://arxiv.org/pdf/1804.08348.pdf

- For Deep FER Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs) are preferred
- Combinations of different types of neural networks are also possible

Step 3: Training

The labeled and pre-processed images are transferred to the selected neural network.

- After the image has been handed over, the NN gives a prognosis which emotion will be shown there.
- The prognosis is compared with the true emotion on the image and the so called Loss Value is calculated.
- With the help of the Loss Value the parameters and weights of the NN are adjusted to improve the performance.
- This process is repeated until the best possible performance is achievedTranslated with www.DeepL.com/Translator (free version)



https://arxiv.org/pdf/1804.08348.pdf

What can be trained with?

It can be trained with different types of data:

- Static images (posed or real emotions are possible)
- sequences of pictures showing the development and progression of the emotion
- Image sequences showing different faces for comparison

fear, sadness, anger, joy, disgust, contempt, surprise

https://github.com/thoughtworksarts/EmoPy

Facial Expression Recognition FER – *analog/digital*

AU pattern
Variation of intensity

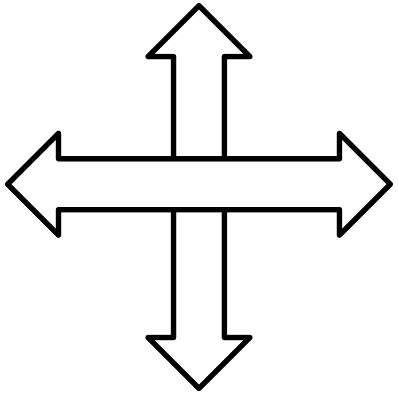
Temporal development of emotions as modulation of the AU by Paul Ekman

Benefits and process

Pictures and Videos

Fake Videos
Documents

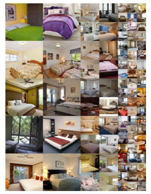
Recognition



Manipulation

Speech2Face

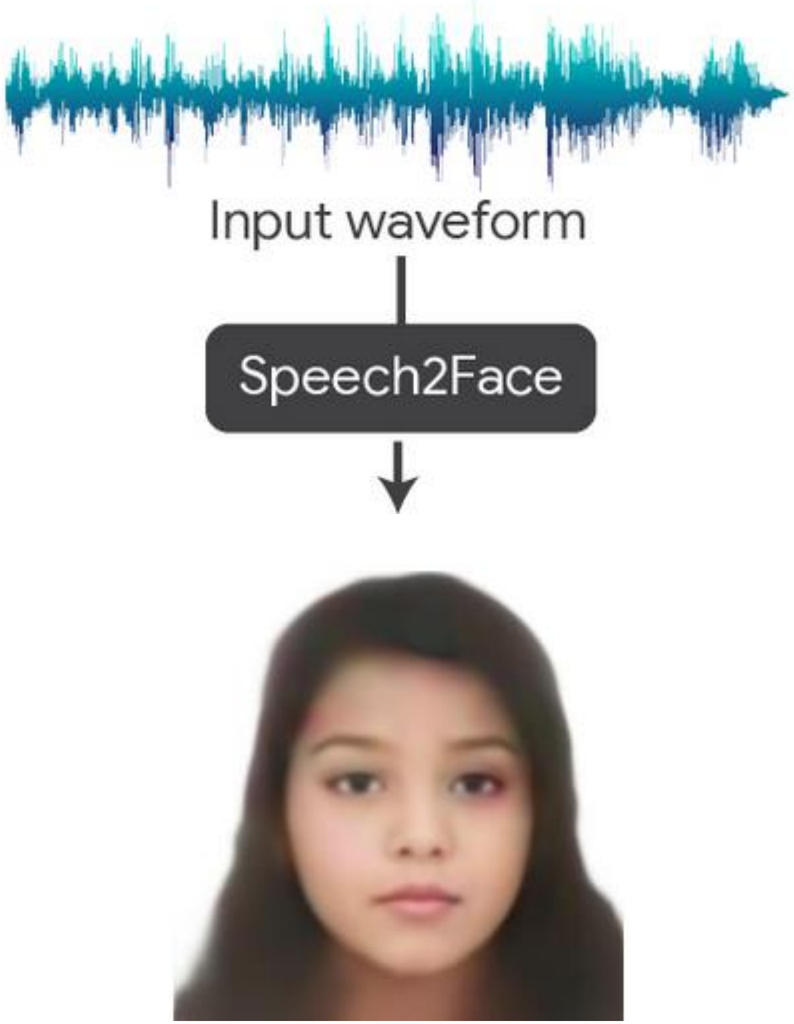
Pose Estimation



Style-Based Generator

Deep FER

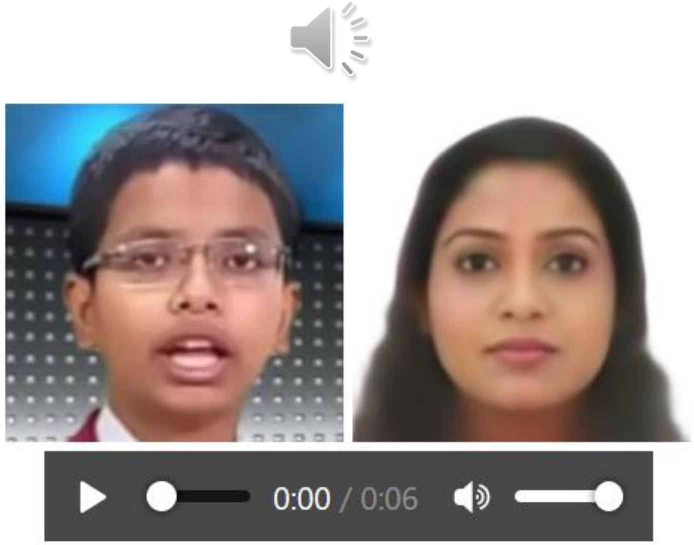
Problem area - "Language and appearance”



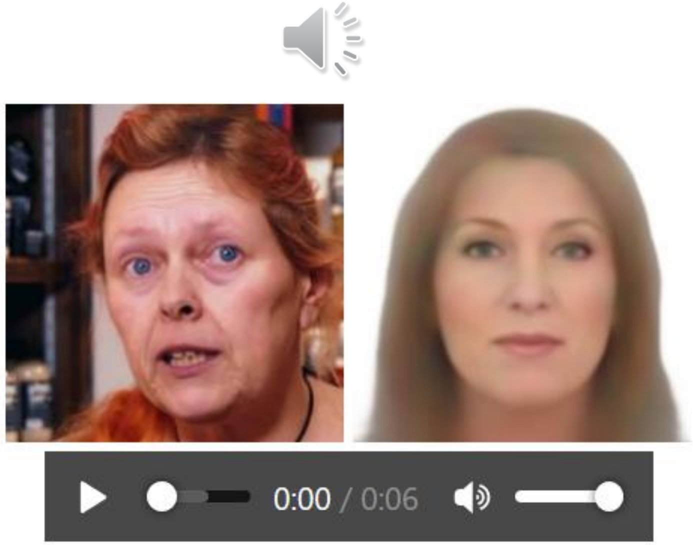
Original image (ref. frame) | Reconstruction from image | Reconstruction from audio

Speech2Face

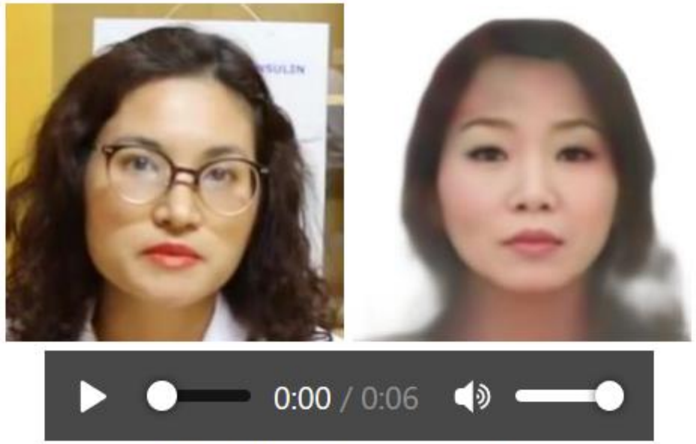
Problemfeld – „Sprache und Aussehen“



(a) Gender mismatch

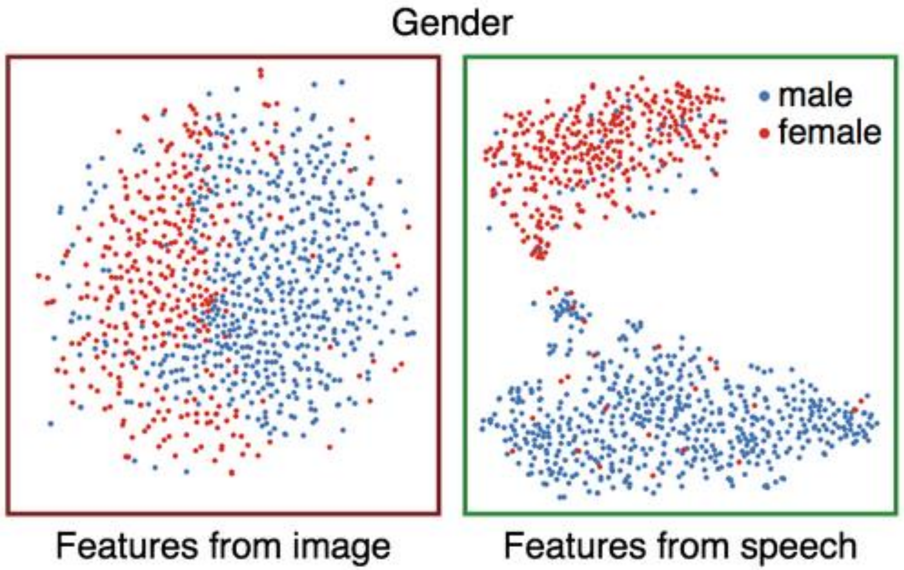


(c) Age mismatch (old to young)



(b) Ethnicity mismatch

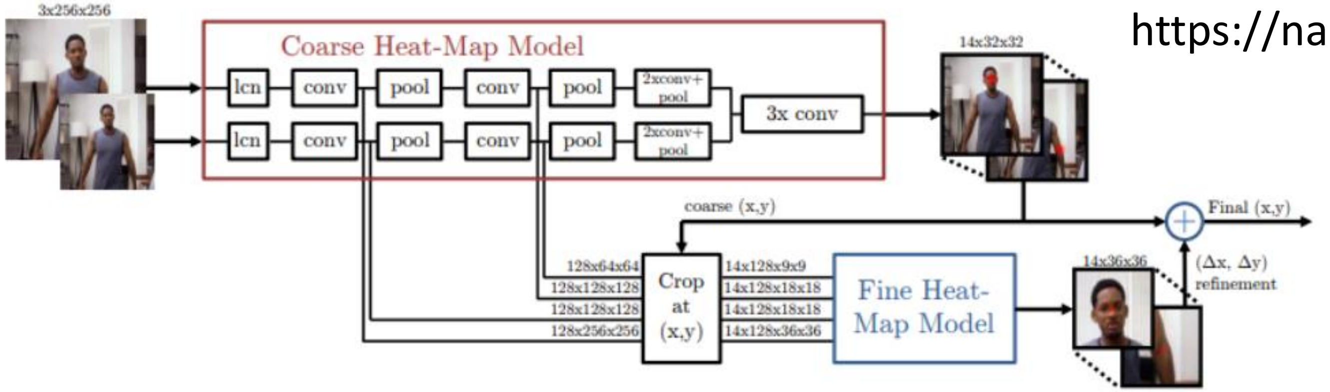
(d) Age mismatch (young to old)

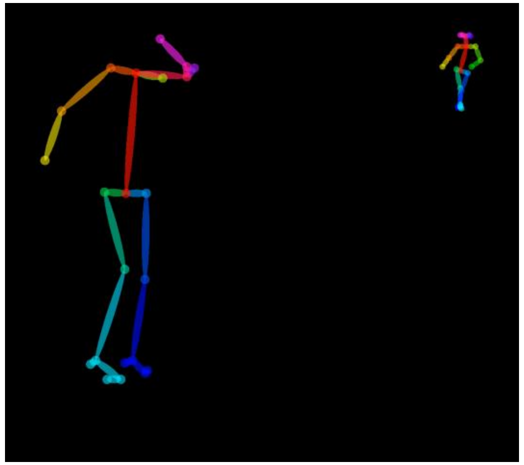


Problem area - movement

Cascaded Pyramid Network for Multi-Person Pose Estimation

https://nanonets.com/blog/human-pose-estimation-2d-guide/

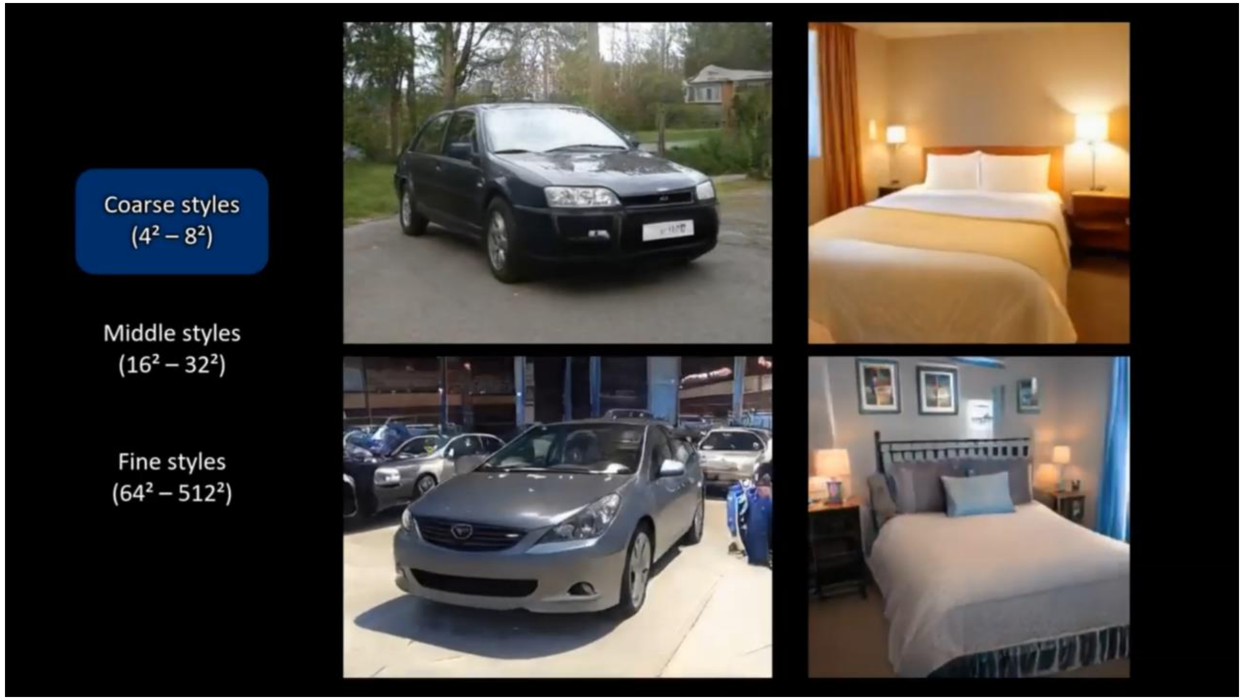




Pose Estimation

Detection and prediction of natural poses and movement

Problem area - scene/background



Style-Based Generator

Possibilities and overall picture

Training with the different algorithms

AI algorithms for various problems

Scene

Language

Movement

Face/Emotions

Analysis **Generation**

By linking different algorithms it is possible to search for contradictions in videos and images.

Use cases can be:

- Surveys
- Insurance videos
- accident videos ...

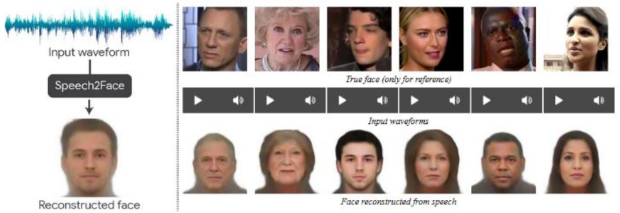
Supplements and literature

IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2019

Speech2Face: Learning the Face Behind a Voice

Tae-Hyun Oh[*,†] Tali Dekel[*] Changil Kim[*,†] Inbar Mosseri William T. Freeman[†] Michael Rubinstein Wojciech Matusik[†]

[†] MIT CSAIL



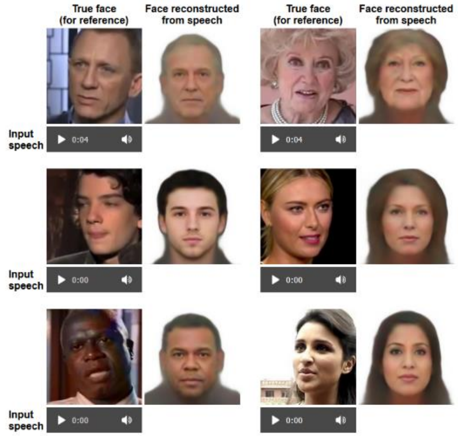
We consider the task of reconstructing an image of a person's face from a short input audio segment of speech. We show several results of our method on VoxCeleb dataset. ***Our model takes only an audio waveform as input*** *(the true faces are shown just for reference). Note that our goal is not to reconstruct an accurate image of the person, but rather to recover characteristic physical features that are correlated with the input speech.*

**The three authors contributed equally to this work.*

https://speech2face.github.io/

https://speech2face.github.io/supplemental/index.htm

Figure 1: Teaser



Real-time Convolutional Neural Networks for Emotion and Gender Classification

Octavio Arriaga
Hochschule Bonn-Rhein-Sieg
Sankt Augustin Germany
Email: octavio.arriaga@smail.inf.h-brs.de

Paul G. Plöger
Hochschule Bonn-Rhein-Sieg
Sankt Augustin Germany
Email: paul.ploeger@h-brs.de

Matias Valdenegro
Heriot-Watt University
Edinburgh, UK
Email: m.valdenegro@hw.ac.uk

https://arxiv.org/pdf/1710.07557.pdf

Fig. 5: Results of the provided real-time emotion classification provided in our public repository

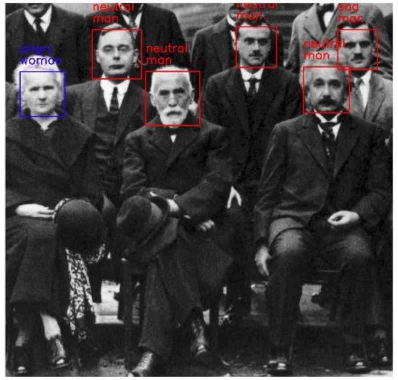


Fig. 6: Results of the provided combined gender and emotion inferences demo. The color blue represents the assigned class *woman* and red the class *man*

Supplements and literature

Facial Emotion Detection Using Convolutional Neural Networks and Representational Autoencoder Units

Prudhvi Raj Dachapally
School of Informatics and Computing
Indiana University

https://arxiv.org/ftp/arxiv/papers/1706/1706.01509.pdf

Fig. 3. Seven classes of emotions in the JAFFE dataset (taken from Dennis Hamester et al., 2015)

FSIL

Forensic Science Investigation Lab

dirk.Labudde@hs-Mittweida.de

☎+49 3727 58 1077