Università degli Studi di Modena e Reggio Emilia

LM Ingegneria Informatica
A.A. 2018/2019

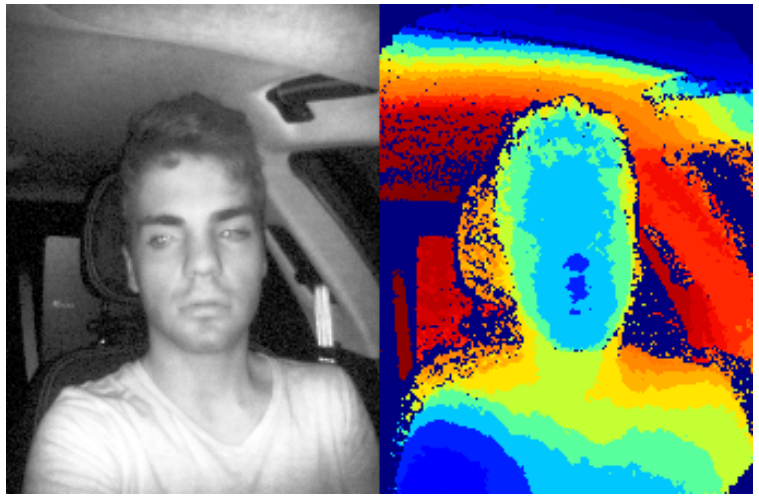
Vision and Cognitive Systems

Depth-based Vision for Driver Monitoring

Guido Borghi

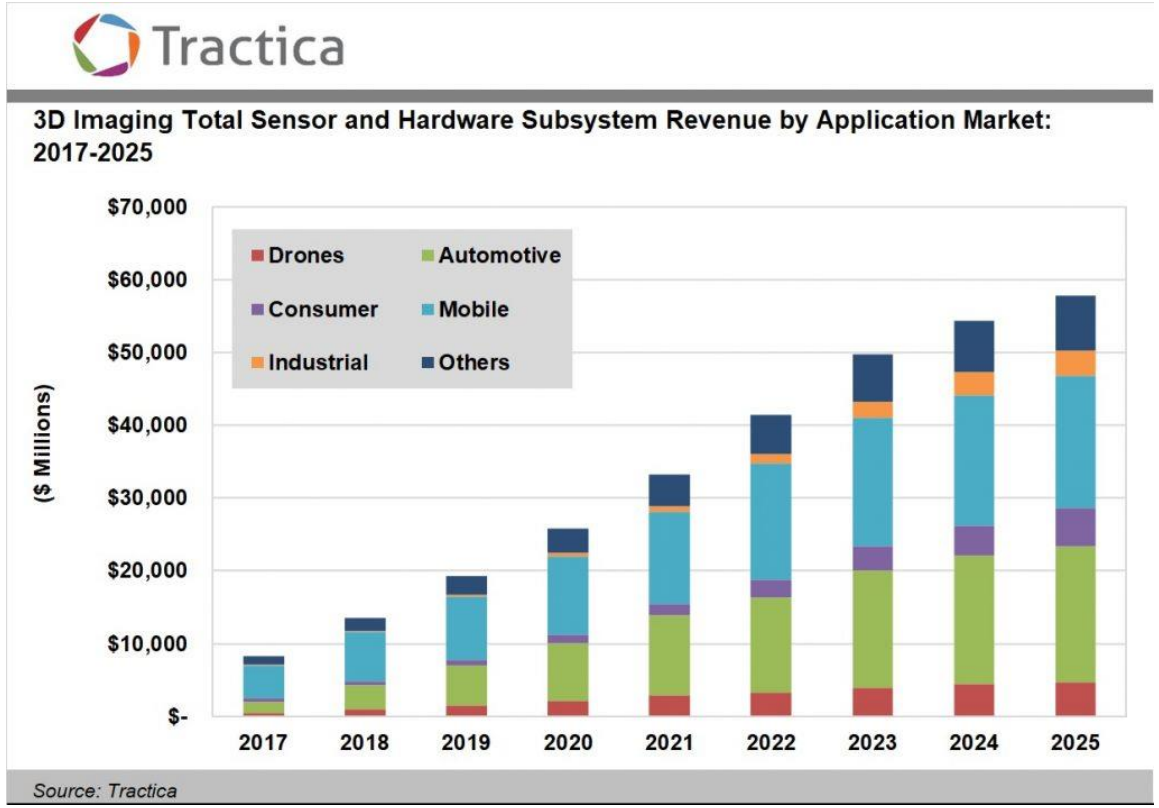
Depth Map

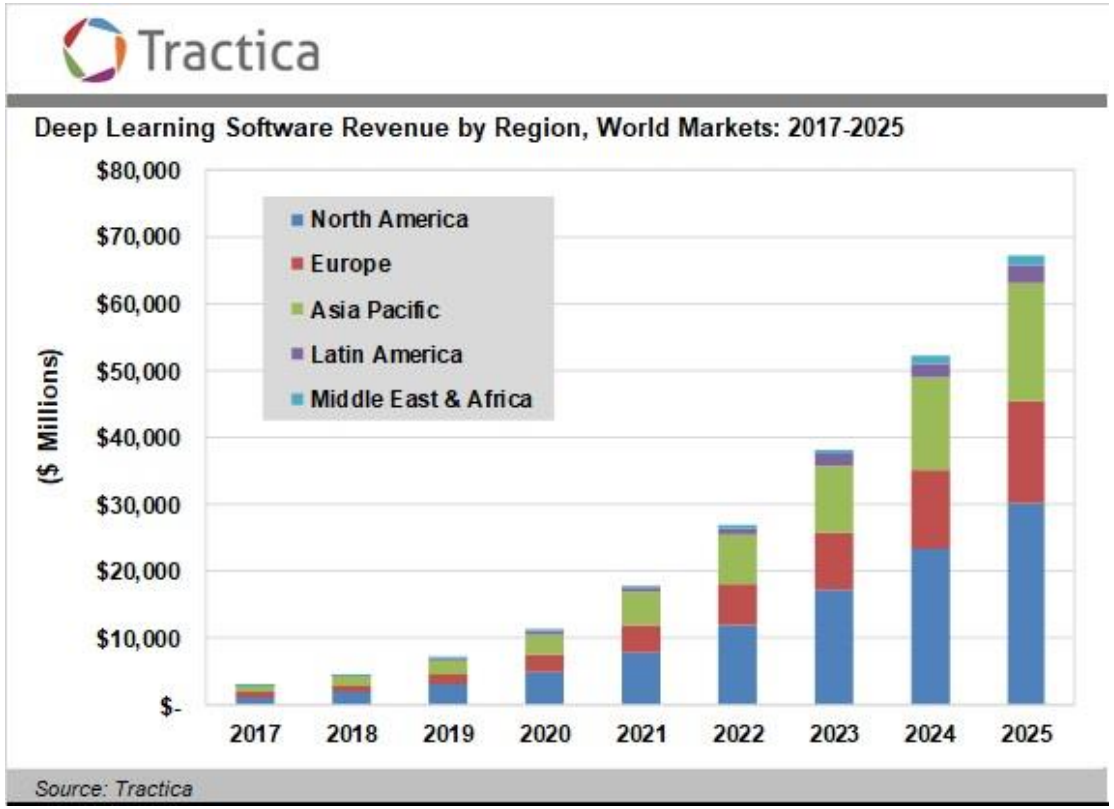
- A *depth map* is an image, or an image channel, that **contains information about the distance** between two objects, *e.g.* the acquisition device and a surface into the acquired scene, *i.e.* an object visible from the camera's point of view
- From a **2D** perspective, **depth maps are usually coded as a gray-level image**, *i.e.* a single image channel with a 0 - 255 range (8 bit or more)
- From a **3D** perspective, **depth map is a projection of a point cloud**, in which every point contains the 3D position in respect to the camera coordinate system
- In the literature, depth maps are also referred as *depth images*, *range images* and *2.5D images*



Depth Sensors

- **Depth** sensors: devices that are able to provide in output distances
- Recently, a lot of **new depth sensors have been introduced in the market**
- **A new trend is acquiring increasing importance: Deep Learning + Depth Maps**

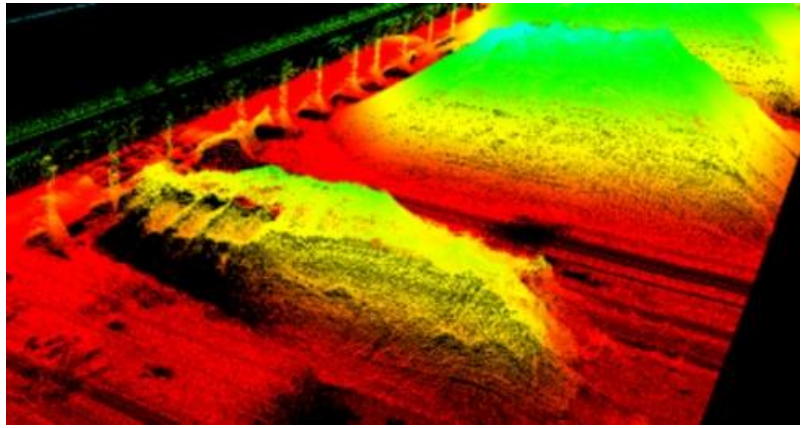




Depth Sensors

- **Goal**: perceiving the 3D world, as the humans do with their eyes

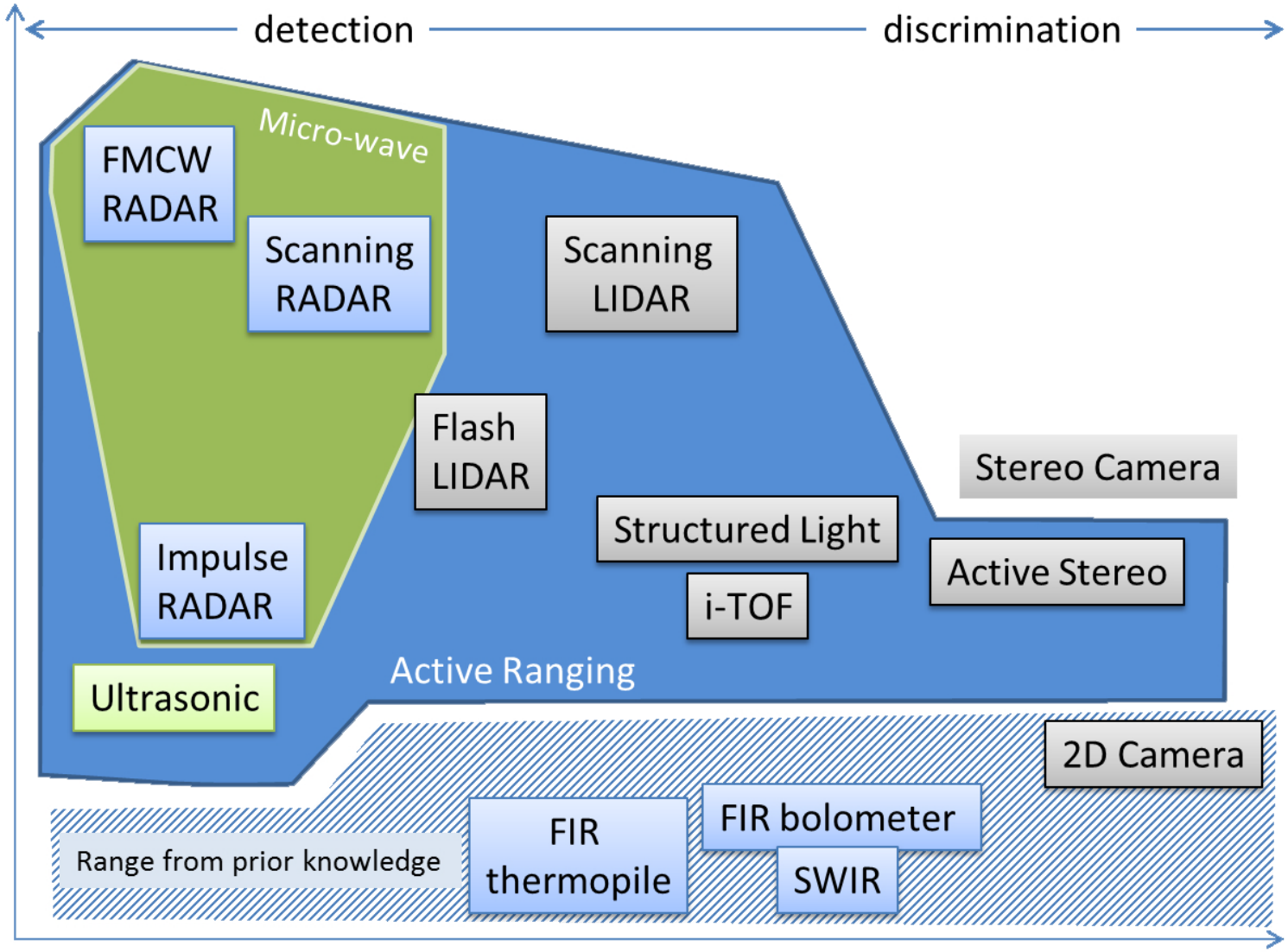
- **Applications**:
  - **Automotive** (ADAS, Driver Monitoring…)
  - **Robots** (*Human-Robot Interaction*)
  - Mobile (**Facial Recognition** on *IPhone X*, laptops…)
  - Industrial applications (**safety** app…)
  - **Drones**



Depth Sensors

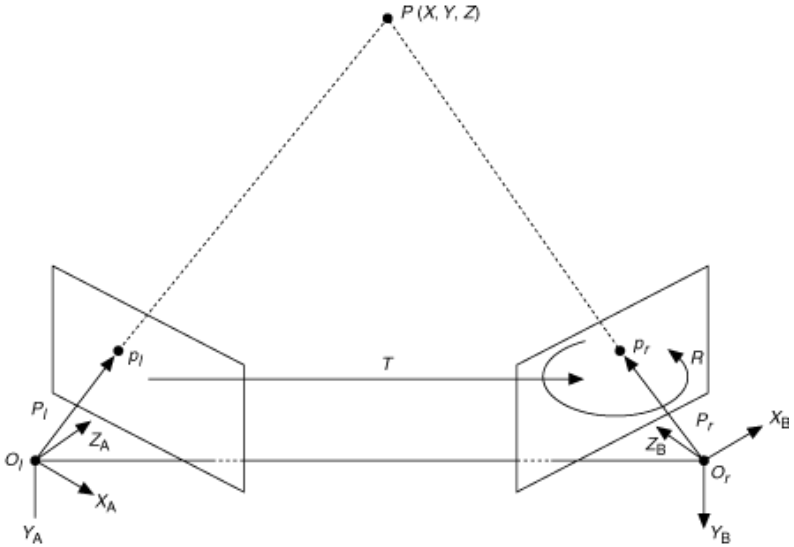
- **Issue with depth maps**: each depth sensor produces depth maps in its own convention
  - *Microsoft Kinect One* (second version): 16-bit depth maps, in which each pixel contains a distance expressed as millimeters
- There are three main technologies for depth devices:
  - **Stereo Cameras**
  - **Structured Light**
  - **Time-of-Flight (ToF)**
- We briefly analyze these types of sensors in order to better understand the following part of the presentation.

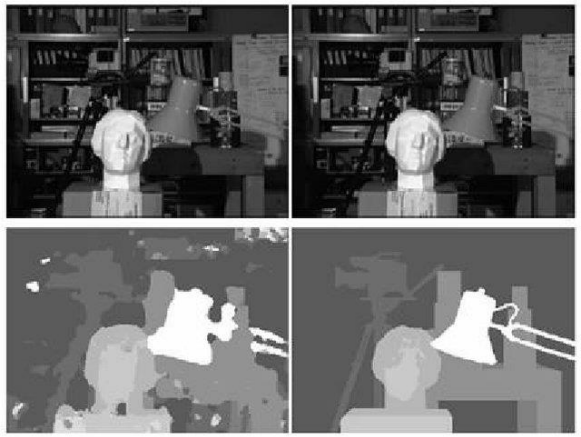
Depth Sensors



Stereo Cameras

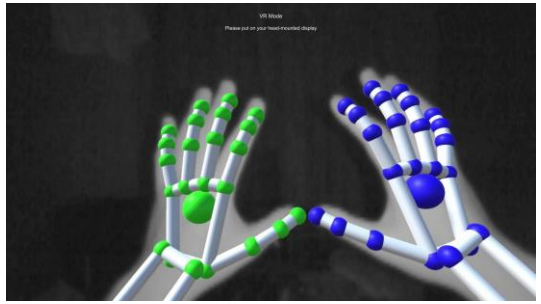
- Similar to what happens in the **human body with the eyes**
- Two similar cameras are placed in a **fixed distance** on the same plane
- The depth of the scene is computed by combining acquired images of these two different intensity (gray-level or RGB) cameras
- Given a pair of rectified images, it is possible to retrieve the distance of a point in the scene applying a **triangulation method** on a set of corresponding points that lying on epipolar lines.





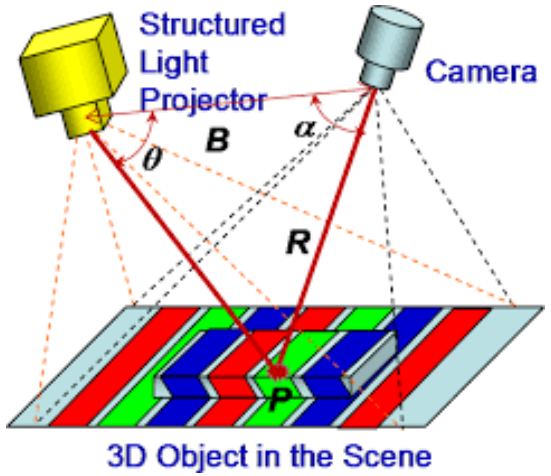
Example of Stereo Cameras

- **Leap Motion** device
  - Two **infrared** cameras with a spatial resolution of 640x240
  - Up to **200 frame per second**
  - Field of view: 135° (**fish-eye** lens)
  - Size: 7 x 1.2 x 3 mm and only 32g of weight
  - SDk for real time **hand tracking** (robust and accurate)



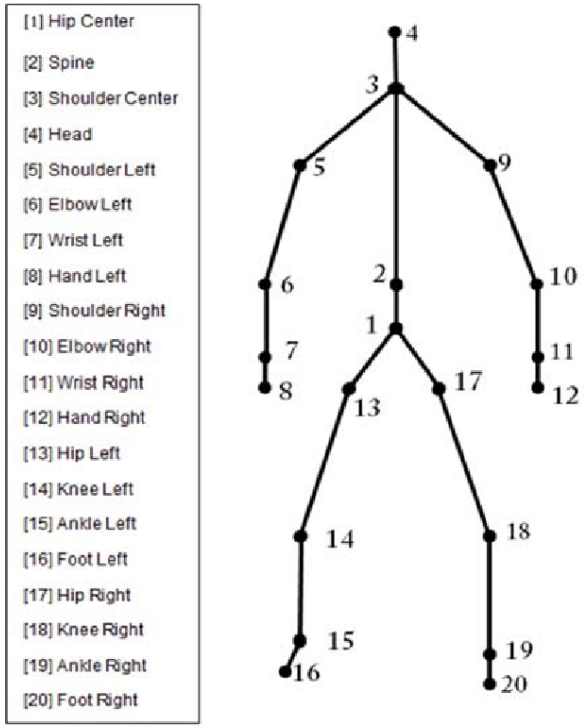
Structured Light

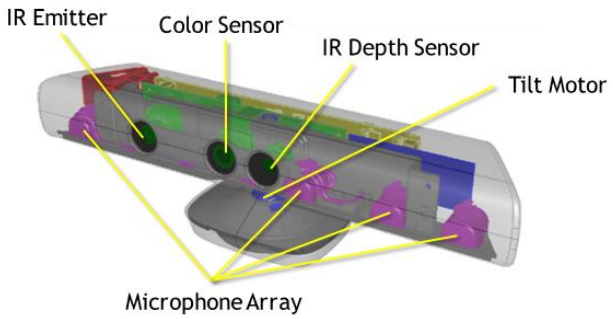
- These scanners project a **specific pattern** inside of the scene
- The **deformation** in the projected pattern introduced by the objects present inside of the scene allows, through appropriate geometric transformations, to return for every projected point its 3D position
- The hardware of these devices includes a **laser projector** and a **sensor** that is sensitive to the corresponding bandwidth



Example of Structured Light Cameras

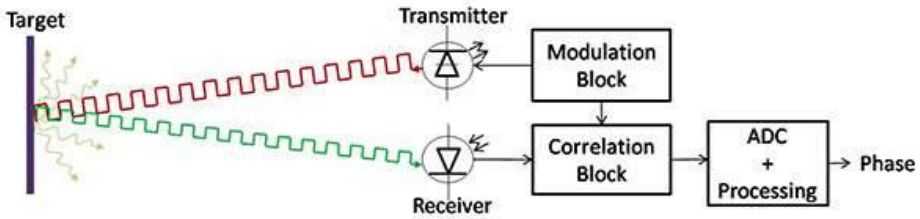
- **Microsoft Kinect** (*first* version)
  - RGB camera: 640x480 up to 30fps
  - CMOS depth sensor (320x240)
  - Range: 0.4 – 4.5 m
  - 2 full skeleton tracked (20 joints)
  - Power consumption: 2.5W
  - Field of view: 57° x 43°
  - Tilt motor



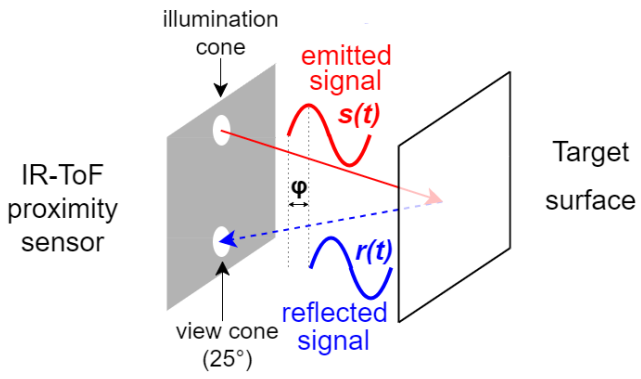


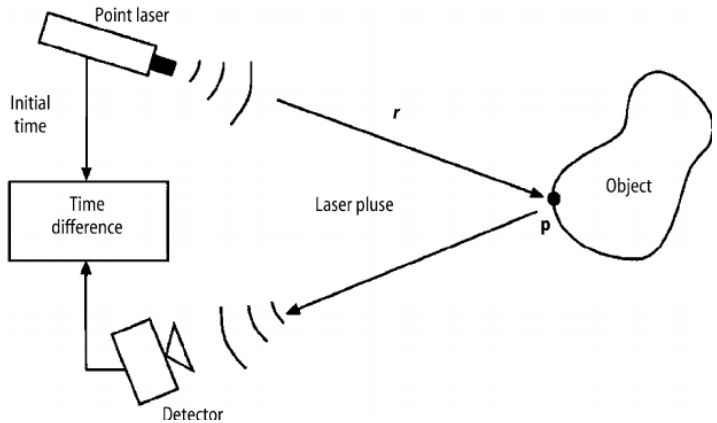
Time-of-Flight

- The distance is computed measuring the **time interval** (the **phase difference**) taken for infrared light to be reflected by the object in the scene



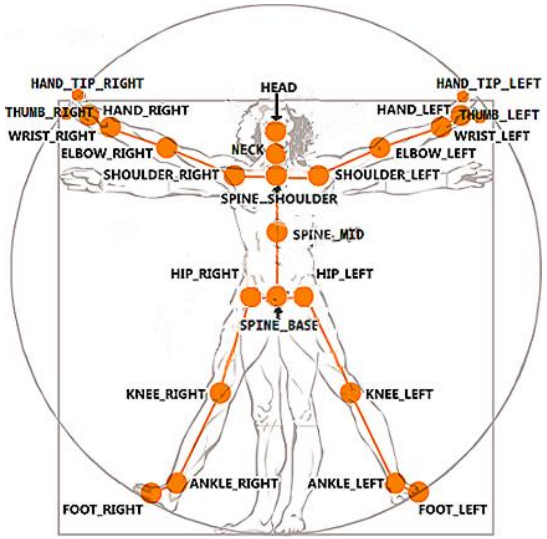
- Also in this case, it is necessary to have a **laser projector** and a **sensor** sensitive to the corresponding bandwidth





Example of ToF

- **Microsoft Kinect One** (*second* version)
  - RGB camera: 1920x1080 up to 30fps
  - CMOS depth sensor (512x424)
  - Range: 0.5 – 5 m
  - 6 full skeleton tracked (26 joints)
  - Power consumption: 2.5W
  - Field of view: 70° x 60°
  - No tilt motor



Example of ToF

- CamBoard Pico Flexx
  - Depth sensors with a spatial resolution of 224x171
  - Only 68 x 17 x 7.35 mm (8g)
  - Up to 45 fps
  - Range: 0.1 – 4m

- Pico Zense DCAM710
  - 69mm x 25mm x 21.5mm
  - Depth resolution: 640 * 480 @ 30FPS
  - RGB resolution: 1920 * 1080 @ 30FPS
  - Viewing angle: 69 ° (horizontal) 51 ° (vertical)

Depth Sensor Comparison

- **ToF vs Stereo Cameras**:
  - The main challenge with Stereo Cameras is solving the **correspondence problem**: this requires **computational intensive algorithms** for feature extraction and matching
  - Stereo Cameras cameras **rely on standard intensity images**, and so are prone to failure in case of low quality of images or an insufficient variation level of textures and colors into the scene
  - ToF devices are not affected by these limitations, since they do not depend on color or textures
  - With Stereo Cameras the **reconstruction error is a quadratic function of the distance**, while ToF sensors, that are based on reflected light, are usually able to increase the illumination energy when necessary
  - Benefits of stereo cameras are represented by the **relative simplicity of implementation**: no laser projectors or laser sensors are required, but just only two cameras

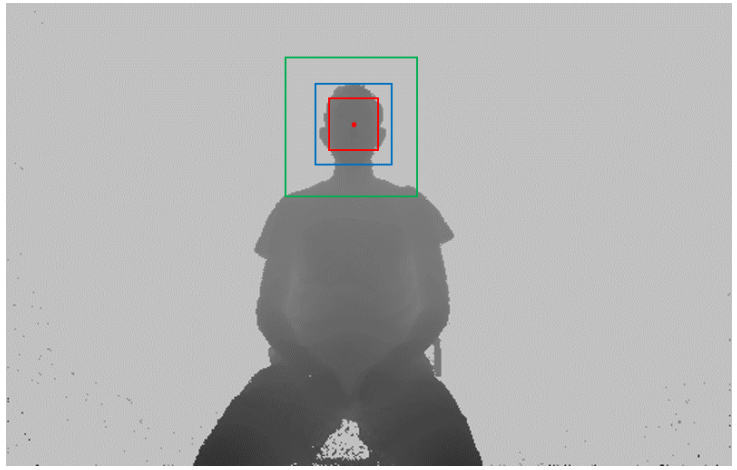
Depth Sensor Comparison

- **ToF vs Structured Light**:
  - ToF: **high spatial resolution** in conjunction with low hardware costs
  - **SL are influenced by external sources of light** (and near-infrared light, *e.g.* the sunlight)
  - The parallel use of several structured-light or ToF sensors is limited by interference problems
  - SL: **lower frame rate usually corresponds to undesired blur effects** if the subject does not remain relatively still during the projection sequence.
  - The majority of the aforementioned problems are only partially present in ToF devices, that are less influenced by external light sources, and **they can achieve a higher frame rate just increasing the size of the laser projector and the related sensor**

Sarbolandi, H., Lefloch, D., & Kolb, A. "Kinect range sensing: Structured-light versus Time-of-Flight Kinect, CVIU 2015

Depth Maps and Objects

- Depth maps are useful for many tasks of the Computer Vision, such as background suppression, scene segmentation, object tracking and 3D reconstruction
- The embedded distance of objects to the camera of depth maps can be exploited to crop object of our interest (for instance, the subject's face), including only a small portion of the background
- Without any additional information or constraint, a face can be located everywhere in the image with an unknown scale. As a result, a complete set of face candidates can be obtained with a sliding-window approach performed at different scales

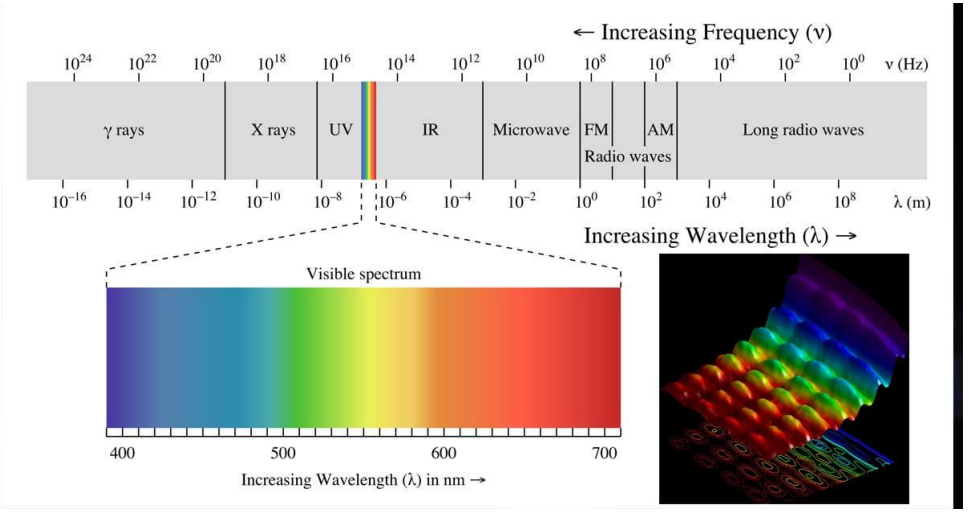


$$w, h = \frac{f_{x,y} \cdot R}{D}$$

- $f_{x,y}$ are the horizontal and vertical focal lengths
- R is the average value representing the width of a face (200mm)
- D is the distance between the acquisition device and the head in the scene

Infrared Images

- Infrared images are similar to depth maps, since they are based on infrared light
- Non-visible light (0.75 – 15 microm)
  - **NIR** (Near Infrared, 0.75 - 1)
  - **SWIR** (Short-Wave Infrared, 1 – 2.7)
  - **MWIR** (Medium-Wave Infrared, 3 - 5)
  - **LWIR** (Long-Wave Infrared, 8 – 14)
    - **Thermocameras**



Driver Distraction

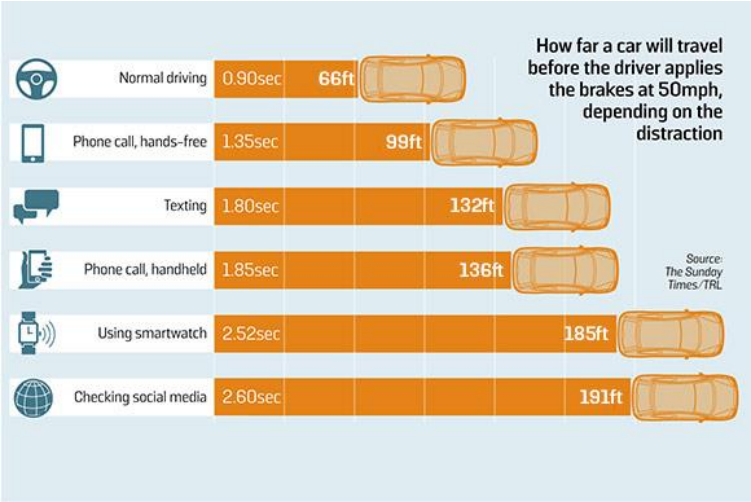
- **What is Driver Distraction**?
  - It is a form of **inattention** (*i.e. failure to pay attention*)

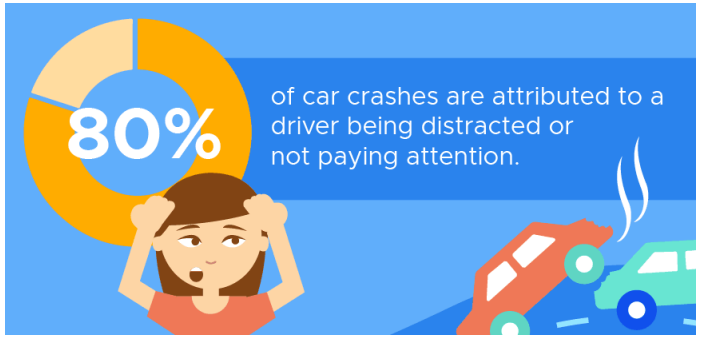
- **Definition**:
  - «*The diversion of attention away from activities critical* ***for safe driving toward a competing activity****, which may result in insufficient or no attention to activities critical for safe driving*»
    (Regan, Hallet & Gordon, 2011)

- Generally, it is **hard** to find a **complete** and accurate description
  - It is also difficult to investigate and find **general solutions** for the problem

Why study Driver Distraction?

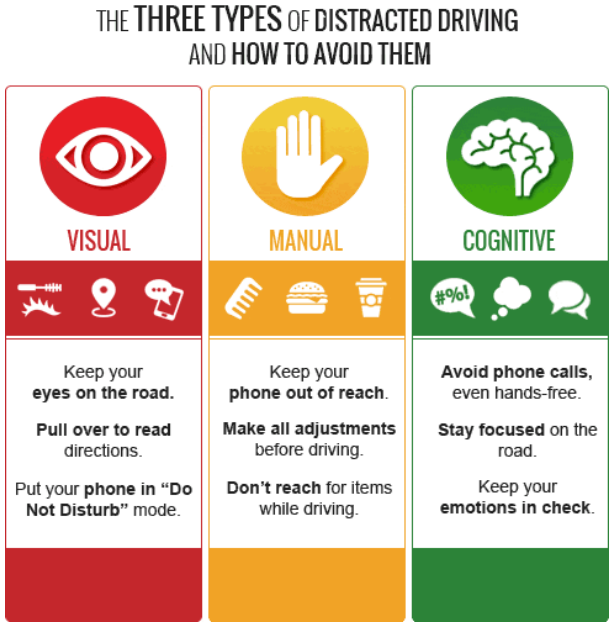
- About **5% to 25%** of **car crashes** have been attributed to driver distraction (and more…)
- Drivers spend about **25-30% of total driving time on distracting activities**
  - 50% concerns **conversation** with a passenger
  - 30% concerns distraction **outside** the vehicle
  - 20% is a **technology-related** type of distraction
- About **70%** (!) of truck crashes are related to driver distraction





Types of Driver Distraction

- Generally, there are **four types of distractions**:
  - <u>Visual Distraction</u>: driver is looking away from the roadway
  - **Biomechanical Distraction**: driver's hands are not on the steering wheel
  - **Auditory Distraction**: driver is responding to a ringing cell phone
  - **Cognitive Distraction**: driver is not focused on driving activity
- Activities that cause **visual distraction** (*e.g.* looking away from the road during texting) appear to be the **most dangerous**
- In this presentation, we will focus on visual distraction

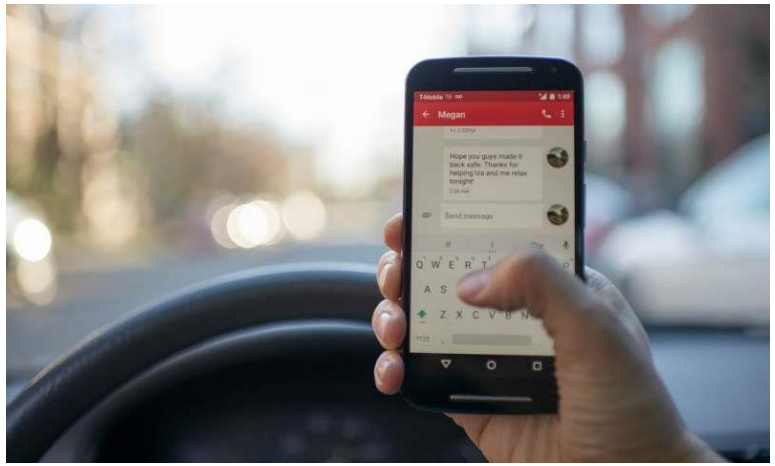


Sources of Driver Distraction

- The sources of driver distraction can reside:
  - **Inside** the vehicle (*e.g.* technology-related)
  - **Outside** the vehicle (*e.g.* traffic-related)
- Driver Distraction can be monitored through:
  - **Physiological signals**: *electroencephalography, electrocardiography and electromyography* (collected with sensors placed usually inside the steering wheel or the car seat)
  - **Vehicle signals**: parameters acquired from the car bus (*e.g.* steering wheel angles)
  - **Physical signals**: image acquisition and elaboration (our field)

The role of the mobile phone

- **Mobile phone** is a «recent» invention:
  - **1973** - *Martin Cooper* introduces the first mobile phone
  - **1997** - Nokia 6610: mobile phones have a small factor form, good battery and they are affordable
- It is one of the **most important cause of fatal driving distractions**
  - It involves **all four distraction categories**
  - It represents about **18% of fatal driver accidents** in North America (source: *National Highway Traffic Safety Administration*)
  - It is also a risk factor for **pedestrian** and **cyclists**!



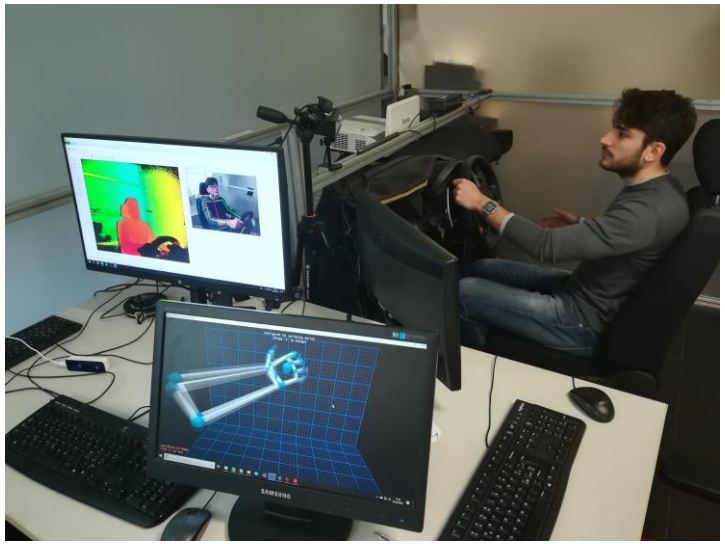
Types of Driver Fatigue

- Moreover, concerning driver distraction, **driver fatigue** plays a crucial role
- Four categories of driver fatigue can be considered:
  - **Local Physical Fatigue**: fatigue in a skeletal or ocular muscle
  - **General Physical Fatigue**: the consequence of a heavy manual labor
  - **Central Nervous Fatigue**: *drowsiness*
  - **Mental fatigue**: low concentration due to bad physical conditions

Countermeasures

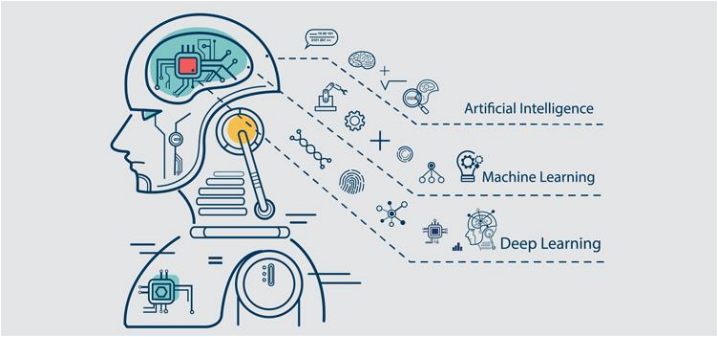
- There are 5 broad categories of **countemeasures** to address distraction:

  1. **Legislation** and enforcement
  2. Driver **training**
  3. Publicity campaigns
  4. **Technology-based** countermeasures
  5. Road **infrastructures**



They can be directed at drivers, transport companies, roads and vehicles.
The number 4 represents our application field.

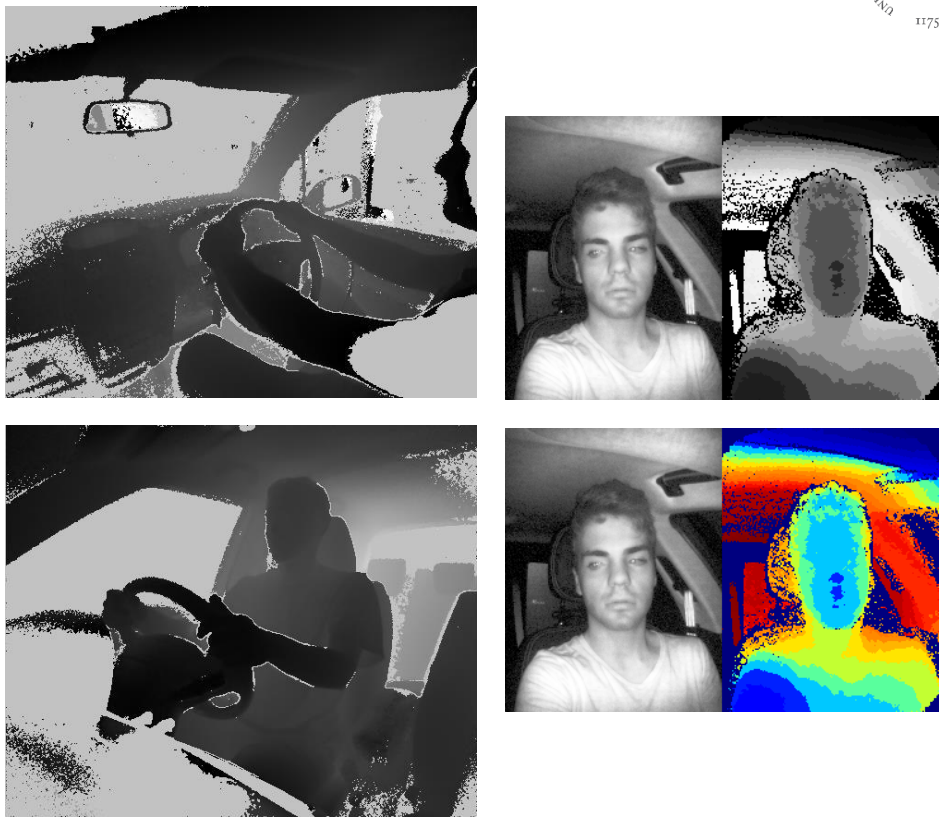
AI for Driver Monitoring - Requirements

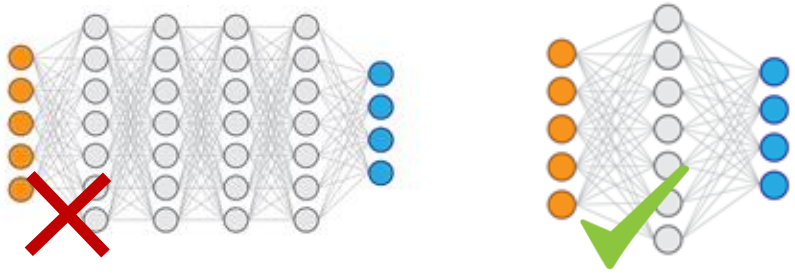


- We investigate how to use:
  - **Computer Vision** algorithms
  - **Artificial Intelligence** (*Machine Learning* and *Deep Learning* techniques)
- Some **requirements** are needed for the **Automotive and Computer Vision** contexts:
  - **Light Invariance**: vision-based systems have to be reliable even in presence of dramatic light changes (day, night, tunnels, bad weather conditions)
  - **Non-invasiveness**: driver's movements and gaze must not be impeded during the driving activity
  - **Real Time performance**: monitoring and interaction systems have to quickly detect anomalies and provide a fast feedback

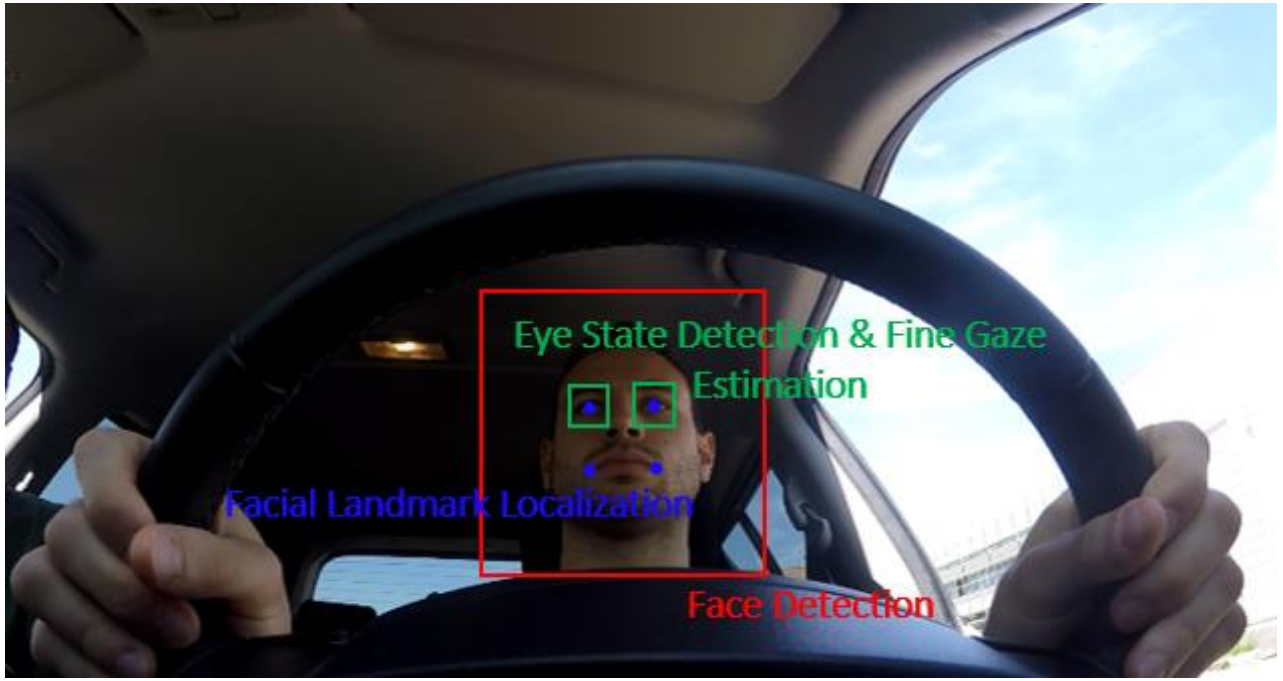
AI for Driver Monitoring - Solutions

- Possible solutions:
  - Light Invariance → **infrared** sensors:
    - Depth Sensors and Depth Maps
  - Non-invasiveness → **new depth** sensors
    - **Cheap** but **accurate**
    - **Small factor form**
  - Real Time → **design strategy**
    - **Shallow** Deep Neural networks
    - Graphics Processing Units





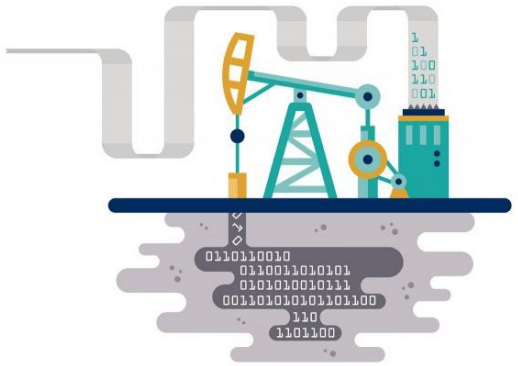
- Useful tasks for *Driver Attention Monitoring*:
  - Head Pose Estimation
  - Facial Landmark Localization
  - Head Detection
  - Face Recognition



- **Deep learning-based techniques**
  - First works in late ‘70, but limited by mathematical and computational problems
  - Today, it is a revolution in the Computer Vision (but not only) field
  - Powerful models, but there are some limitations related to:
    - **Availability of a huge amount of training data**
    - **Data must be annotated** (for supervised approaches)

      • «*Data is the new oil*» (Clive Humby)
    - **High computational power needed**

      • Nvidia GPU



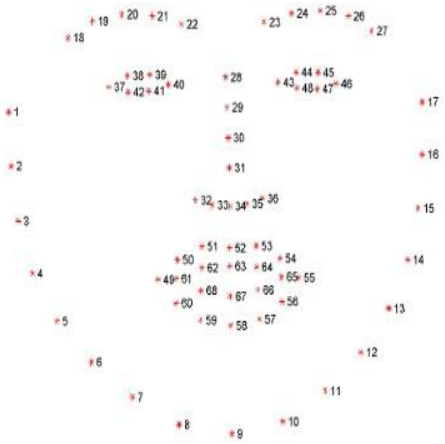
- Data collection and annotation is a key element in developing new AI algorithms

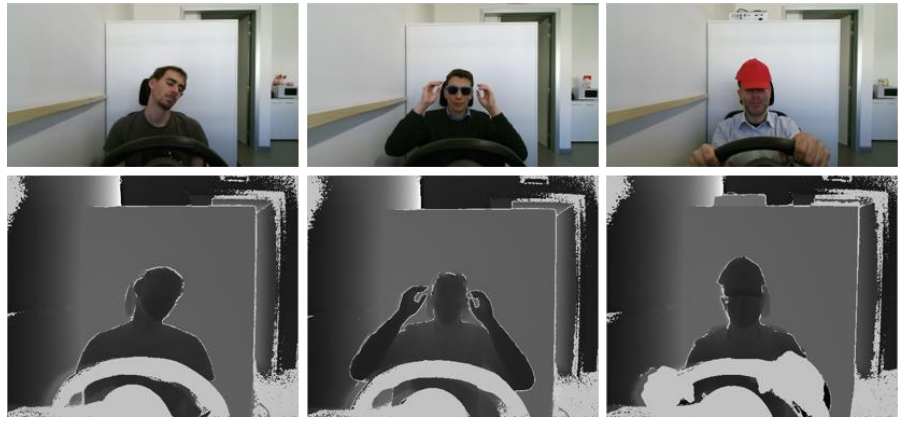
Pandora dataset

- Goal: *Head and Shoulder Pose Estimation* task
- **250k** images (**Full HD RGB** and **depth**) of the upper body
- **22 subjects**
- Annotations: **head** and **shoulder** angles (*yaw*, *pitch* and *roll*)
  - **Head**: ±70° roll, ±100° pitch and ±125° yaw
  - **Shoulder**: ±70° roll, ±60° pitch and ±60° yaw
- Challenging **camouflage** (glasses, scarves, caps…)

MotorMark dataset

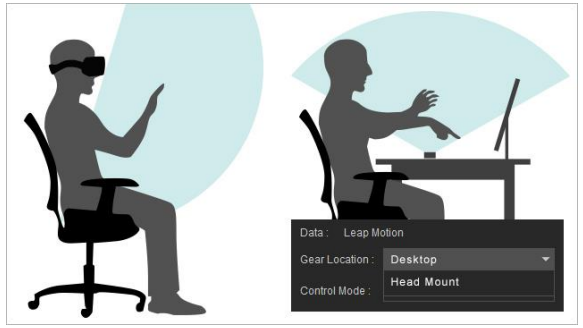
- Goal: *Facial Landmark Localization* task
- **30k** images (RGB and depth) of the upper body
- **35 subjects** with head garments
- Real automotive context
- Facial Landmark annotations following the *ISO MPEG-4* standard

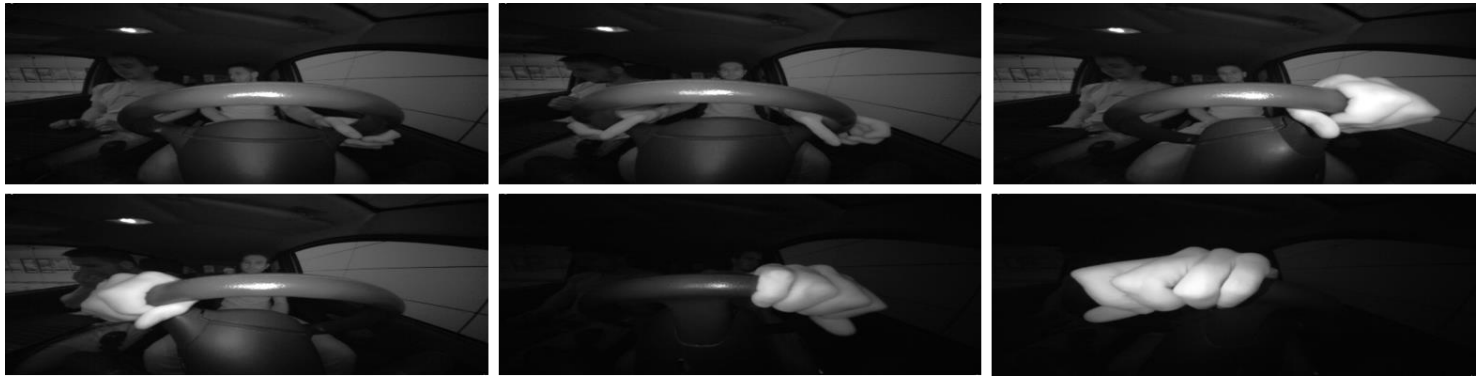




Turms dataset

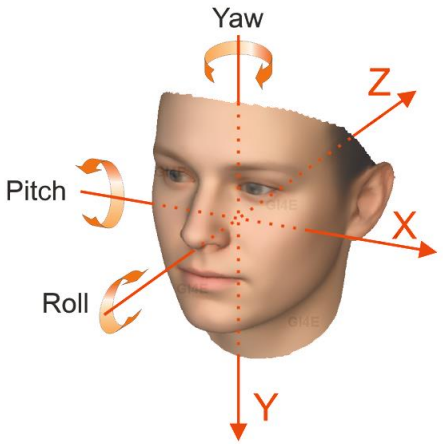
- Goal: *Driver's hands detection and tracking*
- **14k infrared** (*IR*) frames acquired
- *Leap Motion device* (**stereo camera**)
- **Annotations**: bounding boxes of right and left hands
- Original position: back to the steering wheel

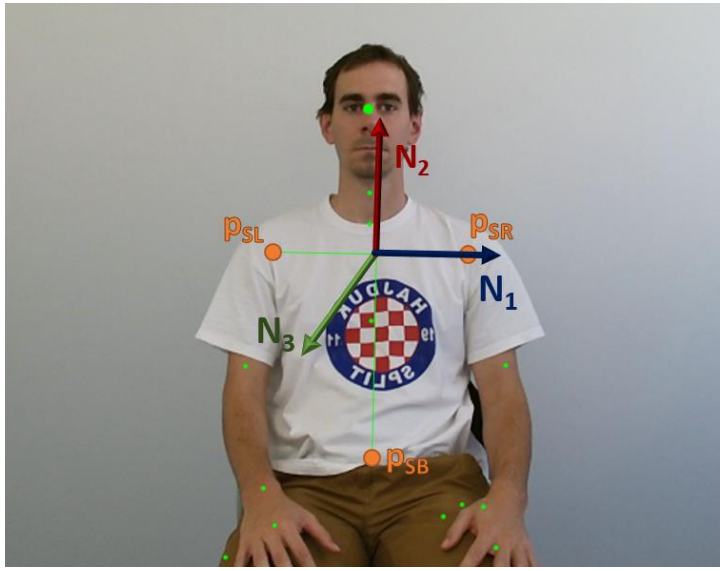




Head and Shoulder Pose Estimation

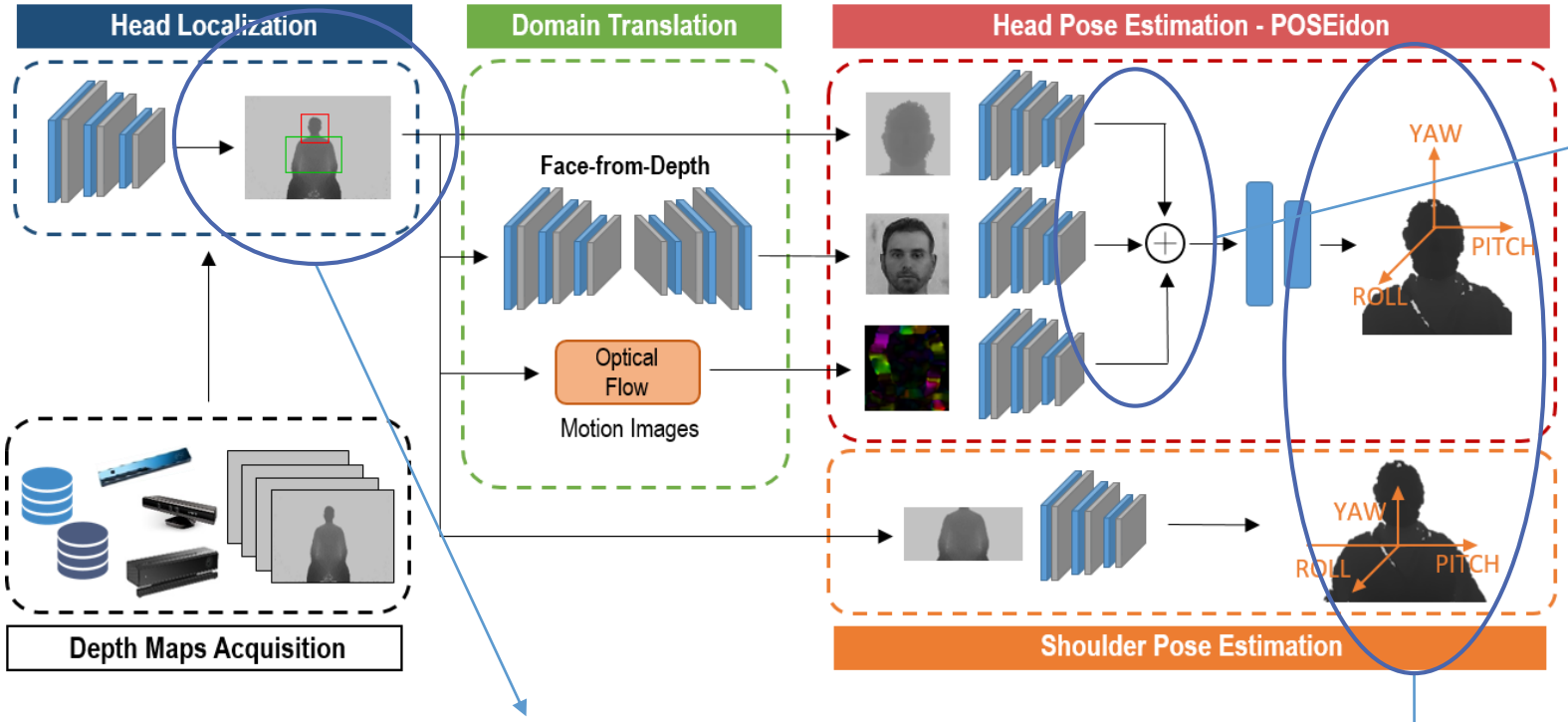
- It is the ability to infer the orientation of the person's head (shoulders) relative to the view of a camera
- Head (shoulder) pose can be expressed in the 3D space with three angles:
  - **Yaw**
  - **Pitch**
  - **Roll**





- Remember the requirements:
  - Infrared sensors (depth maps)
  - Real Time performance

Head and Shoulder Pose Estimation



Training Procedure

- Double-step procedure
  - 1st: each individual network is trained
  - 2nd: the last fc layer is removed, networks are then merged through a *conv* and *concat* operations:

$$y^{cat} = [x^a | x^b], \qquad d^y = d_a^x + d_b^x$$

$$y^{cnv} = y^{cat} * k + \beta, d^y = \frac{(d_a^x + d_b^x)}{2}$$

x^a , x^b: feature maps
d_a^x , d_b^x: feature channel

- Loss function:

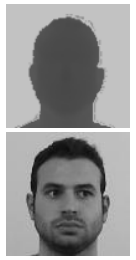
$$L = \sum_{i=1}^{3} |w_i \cdot (y_i - f(x_i))|_2^2$$

$$\boldsymbol{w} = [0{,}2, 0{,}35, 0{,}45]$$

Head crop formula

$$w, h = \frac{f_{x,y} \cdot R}{D}$$

- $f_{x,y}$ are the horizontal and vertical focal lengths
- R is the average value representing the width of a face (200mm)
- D is the distance between the acquisition device and the head in the scene.



Final Outputs

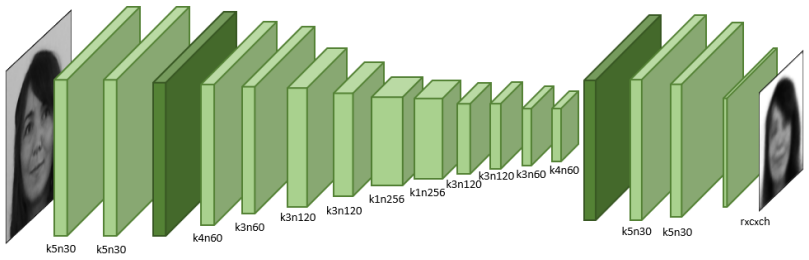
- 3D angles *yaw*, *pitch* and *roll* for head and shoulders

Head and Shoulder Pose Estimation

Face Translation: *Face-from-Depth*

- The main idea is to **add knowledge** (the *generated* faces) at training and testing time, **to improve the performance**
- We propose (two versions) of a **new** neural network called *Face-from-Depth*:

FfD v1: CNN



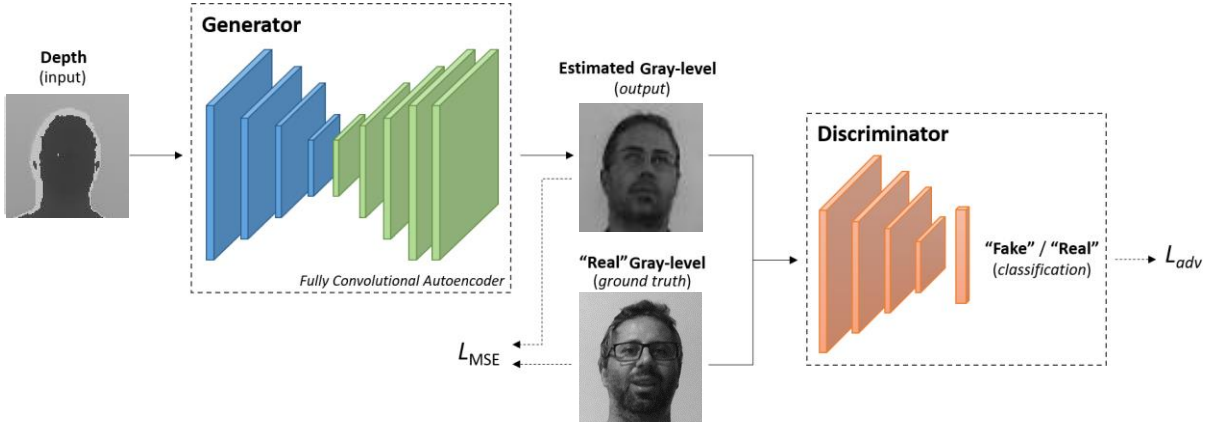
- Elements from **autoencoder** and **CNN**
- **Weighted Loss**:

$$L = \frac{1}{R\,C}\sum_{i}^{R}\sum_{j}^{C}\left(\left|y_{ij} - y'_{ij}\right|_2^2 \cdot w_{ij}^{\mathcal{N}}\right)$$

$$\mathcal{N}:\ \mu = \left[\frac{R}{2}, \frac{C}{2}\right]^T \quad \sum = \mathbb{I}\cdot\left[\left(\frac{R}{\alpha}\right)^2, \left(\frac{C}{\beta}\right)^2\right]^T \qquad \alpha = 3.5, \beta = 2.5$$

- *Bi-variate Gaussian* prior mask to highlight the central area

FfD ^{v2}: conditional GAN



- **Min-max game** (Generator - Discriminator):

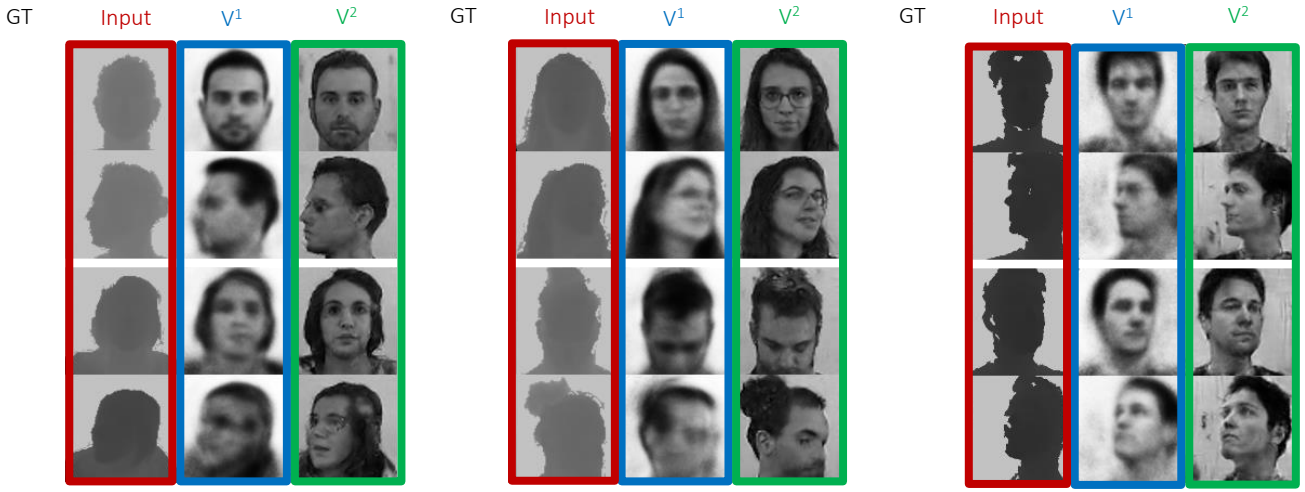
$$\min_{\theta_d}\max_{\theta_g} \mathbb{E}_{x\sim p_{gray}(x)}[\log(G(x))] + \mathbb{E}_{y\sim p_{dpt}(y)}[\log(1 - G(D(y)))]$$

- **2 loss functions**:

$$L_{MSE}(s^g, s^d) = \frac{1}{N}\sum_{i=1}^{N}\left|G(s_i^g) - s_i^d\right|_2^2$$

$$L_{adv}(y, t) = -\frac{1}{N}\sum_{i=1}^{N}\left[t_i \log y_i + (1 - t_i)\log(1 - y_i)\right]$$

Face Translation: *Face-from-Depth*

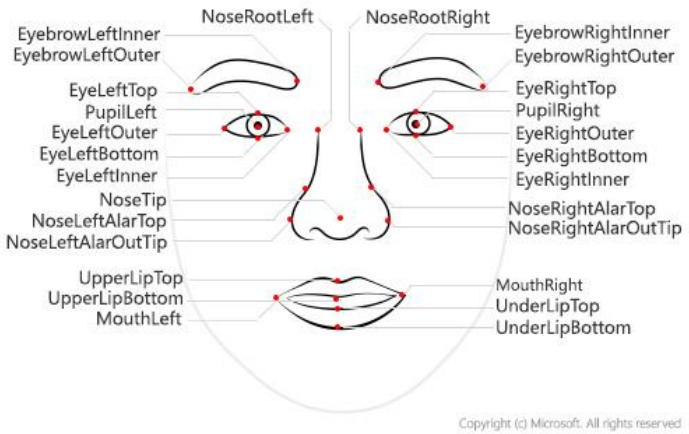


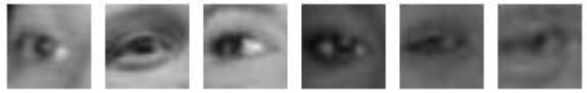
Pandora dataset

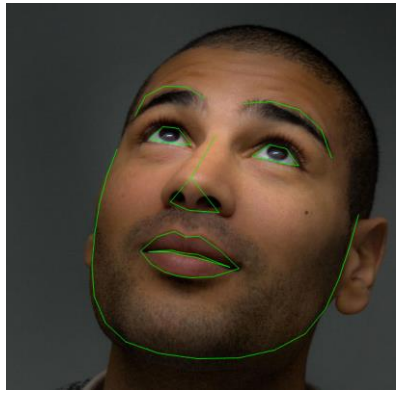
Biwi dataset

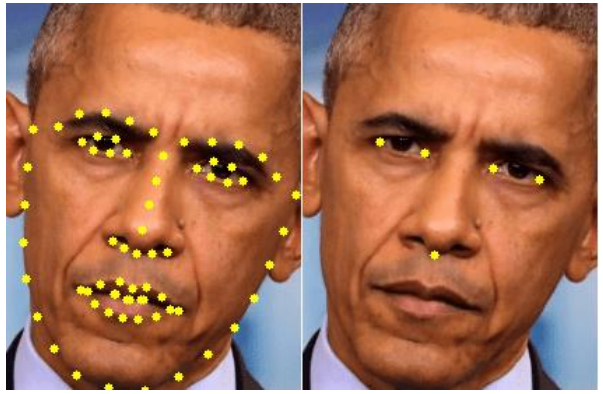
Facial Landmark Localization

- Task: detection and localization of the position of salient points of the human face
- In our case:
  - Eyes' pupils
  - Mouth corners
  - The nose tip

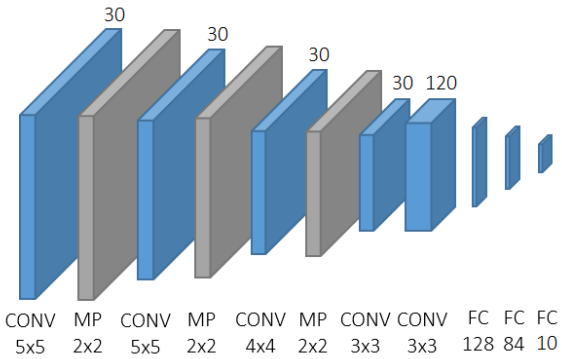








Facial Landmark Localization



Input pre-processing

- *Contrast Limited Adaptive Histogram Equalization* algorithm applied
- Values scaled:
  mean = 0, *variance* = 1
- A fixed window containing the head is cropped and all the cropped images are resized to 64x64 pixels

Deep Network

Shallow model (real time performance)

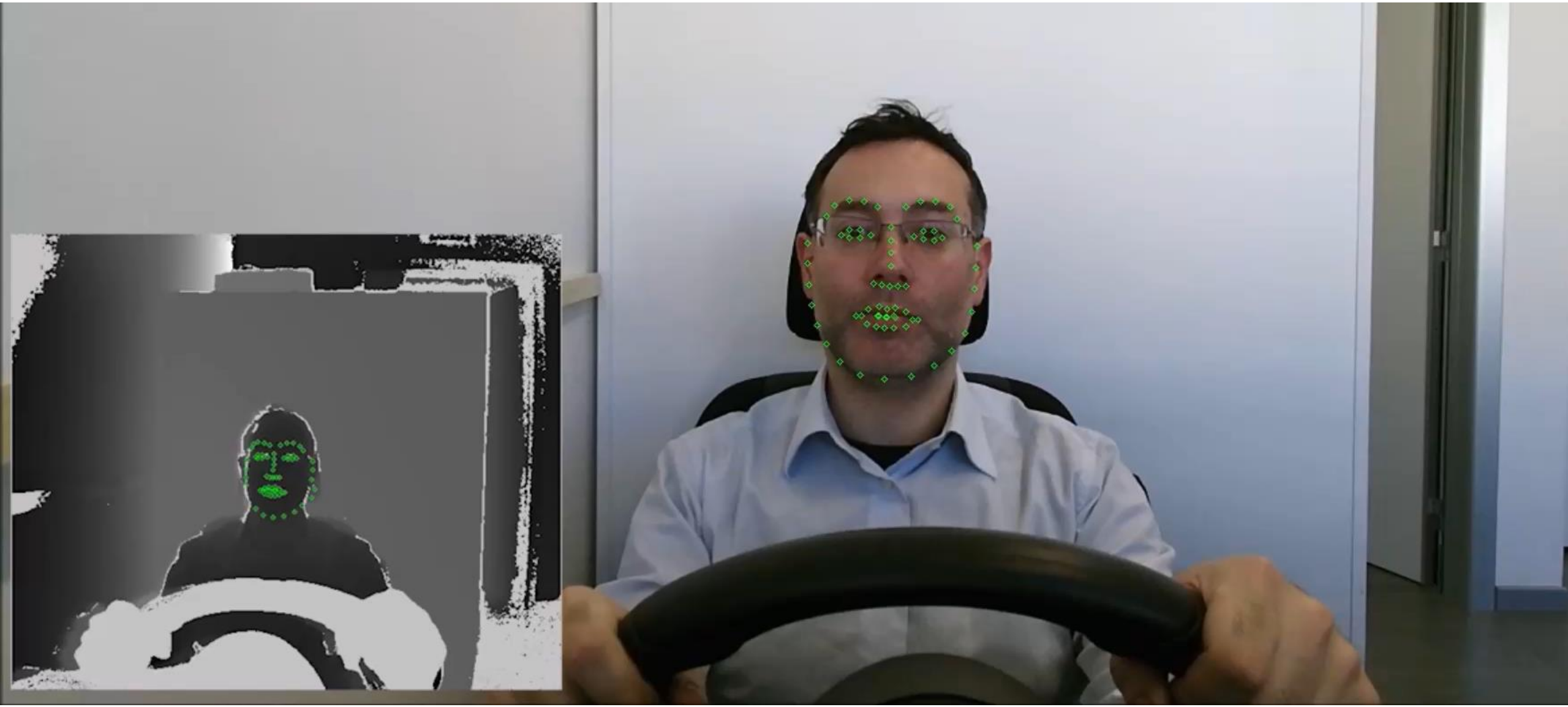
- 5 *convolutional* layers
- 3 *max-pooling* layers
- 3 *fully connected* layers
- Activation function: *tanh*
- L_2 loss:

$$L_2 = \sum_{i=1}^{3} ||(y_i - f(x_i))||_2^2$$

Output

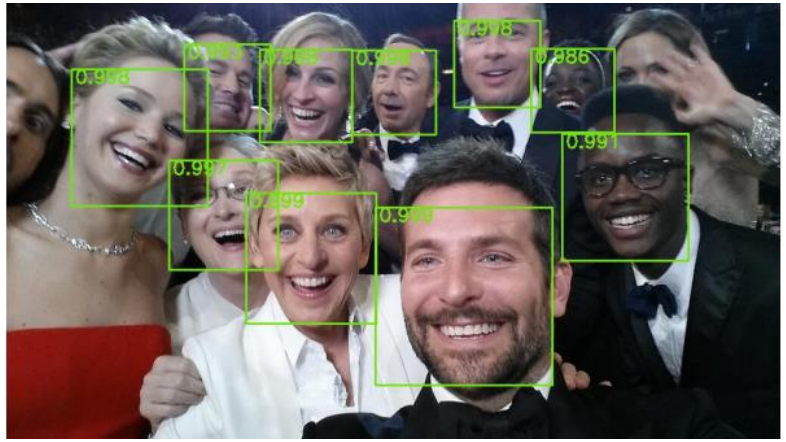
- **10 coordinates** (*x*,*y*) of 5 facial landmarks
- Ground truth coordinates are normalized in the range [-1, 1] accordingly to the specific activation function of the output network layer

Facial Landmark Localization

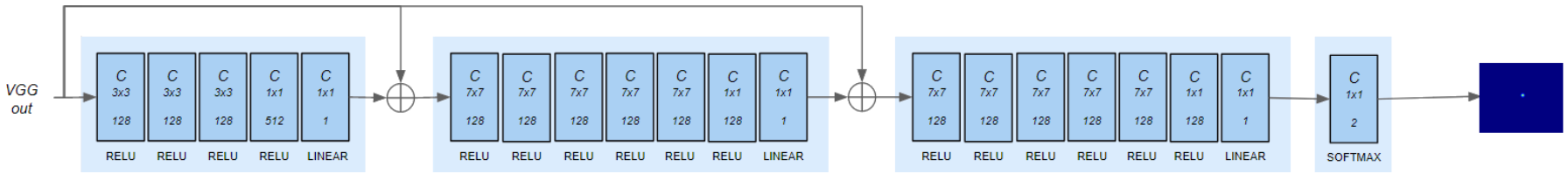


Head Detection

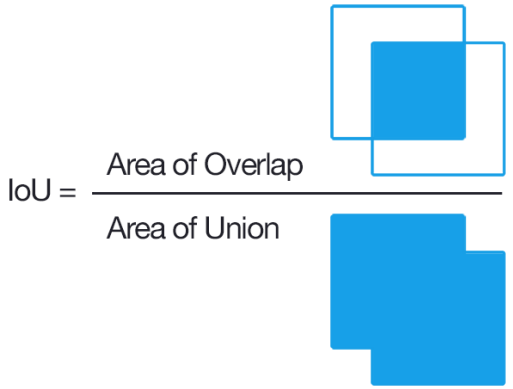
- It is the ability to detect and localize one or more heads/faces in an image.
- This is a traditional problem of the computer vision field, but only few works tackle this task on different types of images, like depth maps.



Head Detection



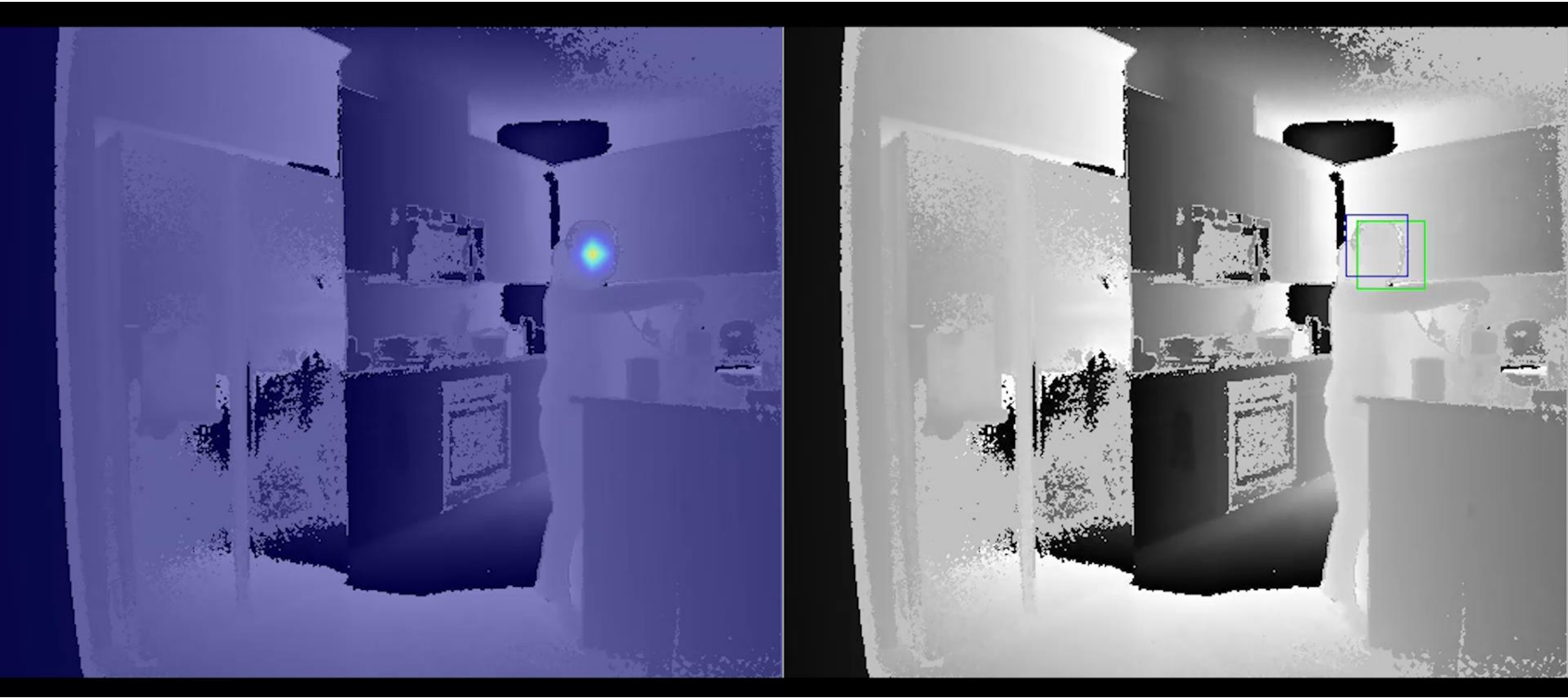
- **Input**: depth maps (512x424 16-bit images, from *Microsoft Kinect One*)
- **Output**: 64x53 probability map (as *bi-variate Gaussian* function)
- **Network Architecture**: Fully Convolutional Network (inspired by CPM [1])
  - We **modify** the original architecture and **reduce** the number of parameters to deal with:
    - Real time performance
    - Lack of training data
- **Network Details**: (*ReLU* + linear activation for each block) + *softmax*
- **Loss function**: categorical cross-entropy



The final accuracy is evaluated through the *Intersection over Union* metric:

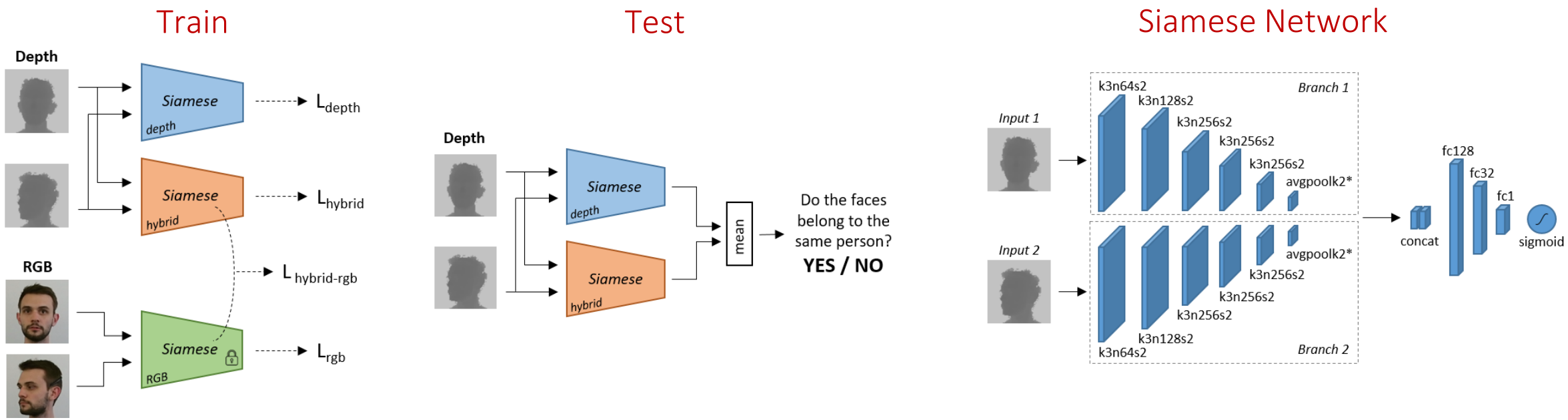
$$IoU(A,B) > 0.5 \qquad IoU(A,B) = \frac{Overlap\ Area}{Union\ Area} = \frac{|A \cap B|}{|A \cup B| - |A \cap B|}$$

Head Detection



Face Recognition

- **Privileged Information**: additional (privileged) knowledge available only during the training but it improves the performance of the system at testing time
- The **JanusNet** system: *Siamese* networks + 3 models (train) + 2 models (test)



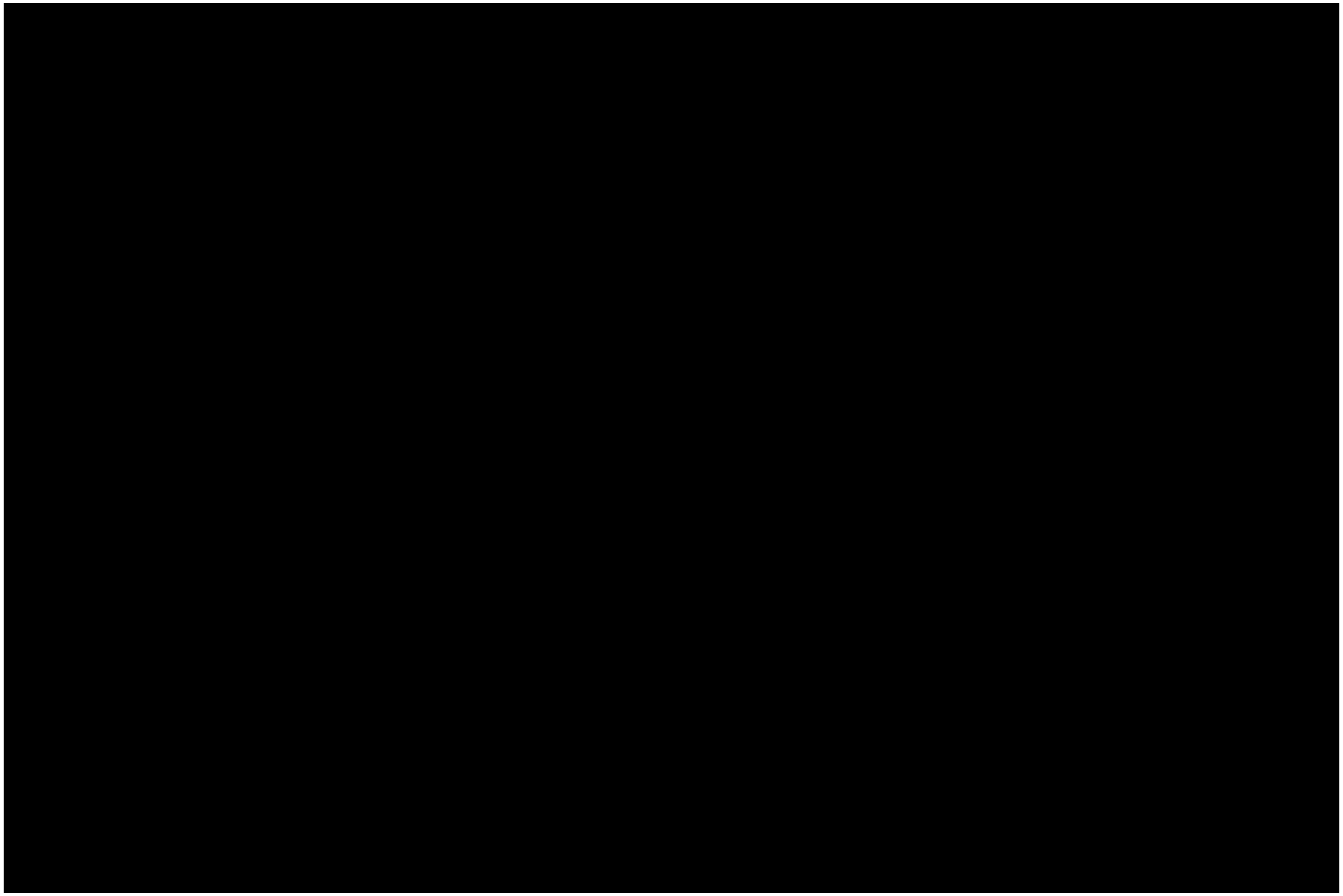
Privileged Information Loss:

$$L_{hybrid-rgb_{1,2}} = \frac{1}{N}\sum_n^N (y_n^{hybrid} - y_n^{rgb})^2$$

Final Loss:

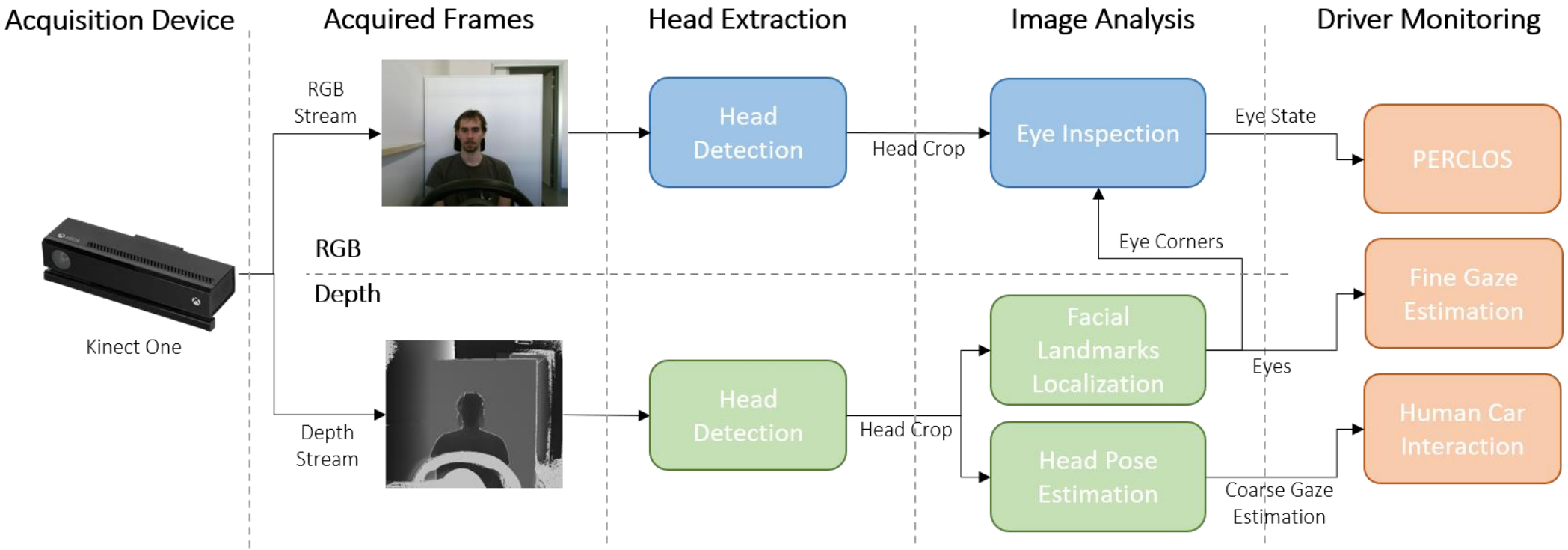
$$L = \alpha\left(L_{hybrid-rgb_1} + L_{hybrid-rgb_2}\right) + \beta\left(L_{depth} + L_{hybrid} + L_{rgb}\right)$$

Face Recognition

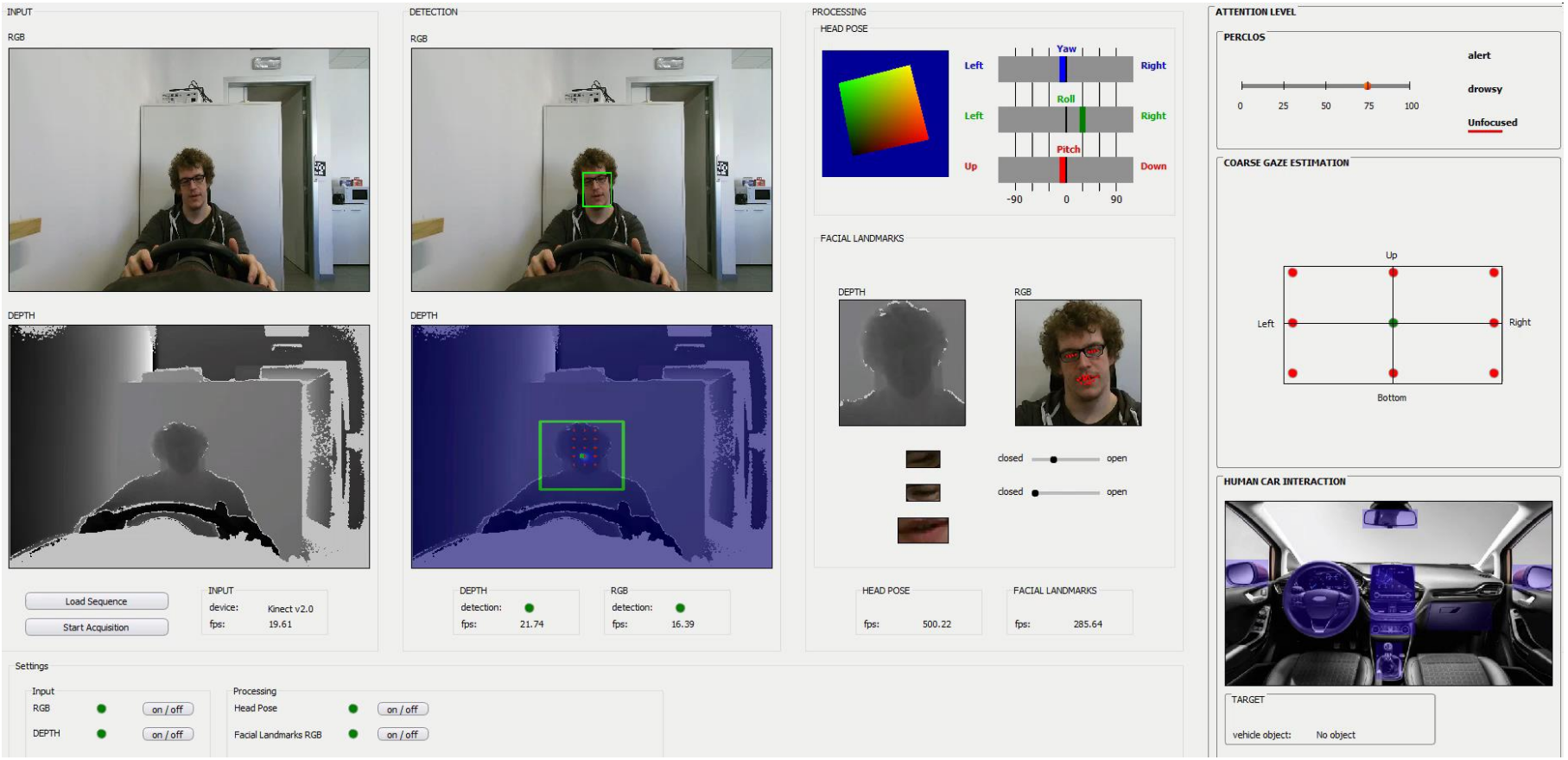


Mercury framework

- A framework for Driver Monitoring and Human Car Interaction
- Let’s merge all the studied algorithms!



Mercury framework



Perclos
Driver drowsiness level

Coarse Gaze Est.
Where is the driver looking at?

Driver-Car Interaction
What object is the driver looking at?

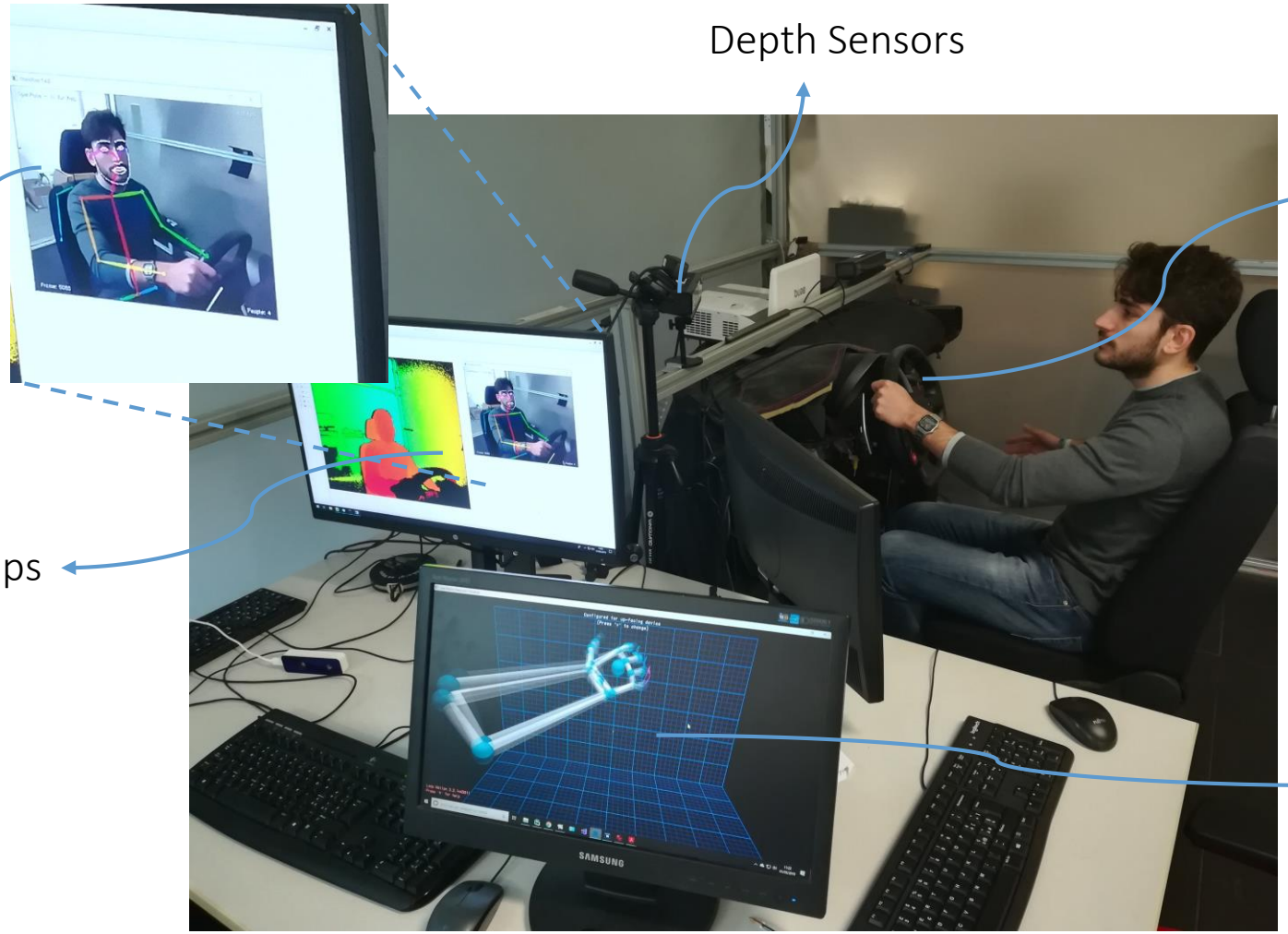
Microsoft Kinect
(RGB + Depth)

Head Detection
(on RGB and Depth)

Head Pose Estimation
Facial Landmarks

Driver's Attention
Eye State

Mercury setup



Depth Sensors

Body Joints
Facial Landmarks

Real Car
Dashboard
(donated by Ferrari SpA)

Depth Maps

Hand Joints for
3D Gesture
Segmentation and
Classification

Commercial Driver Monitorin Systems

- **Real systems proposed by different car companies**
- Vision-based (but not only)
- Changes in **eye movement**:
  - **Gaze Tracking**: cameras capture images of the driver's face and a number of cues including eye gaze direction that are used to infer driver states
  - **Blink Frequency**
  - **Eye closure ratio**

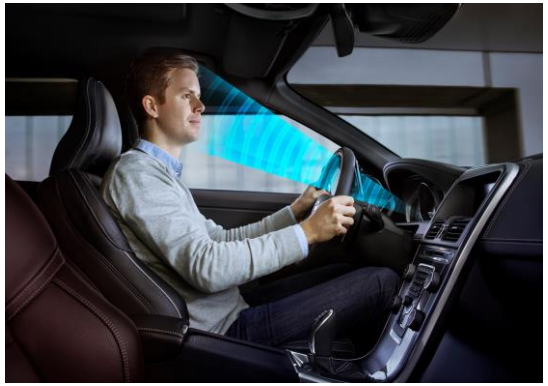
- Physical measures :
  - **Face Tracking / Head Detection**: infra-red light to locate pupils

Driver Monitoring System (Toyota - Lexus)

- First introduced by Toyota in 2006
- Steering wheel mounted infrared sensor which monitors movement of driver's head and eyes.
- Detects if the driver's head has turned to the side for a few seconds or if their eyes are closed.
- Warns the driver with *flashing light and sound* whenever the driver is not paying attention and an obstacle is detected. If no action is taken, the vehicle will turn to active intervention, *applying brakes*.

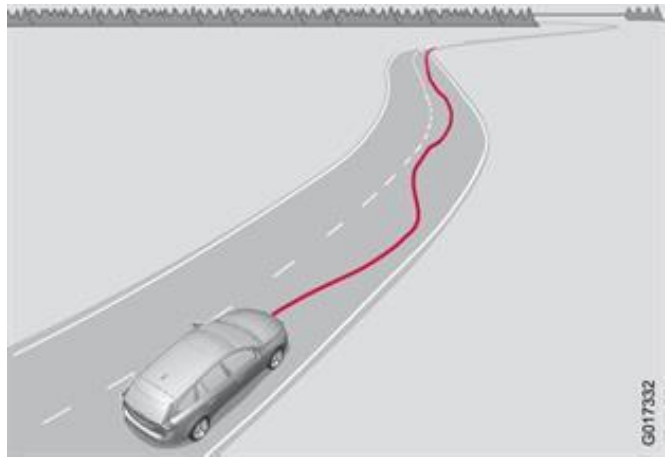
Driver State Estimation (Volvo)

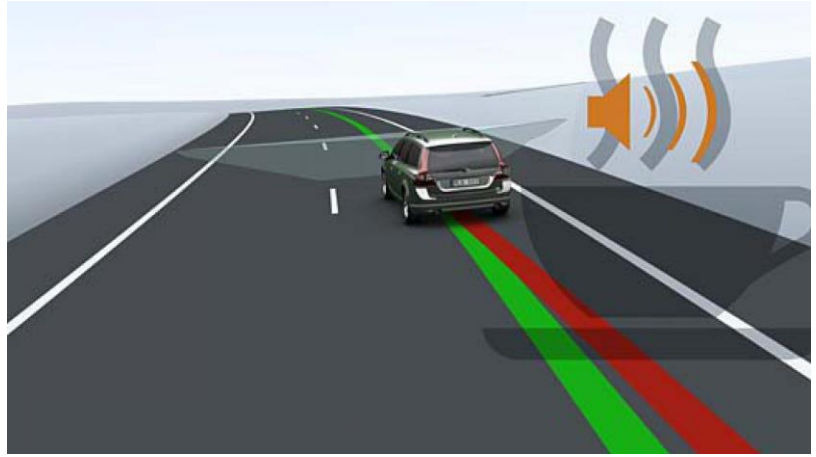
- Dashboard mounted infrared sensor
- Detects which direction the driver is looking, how open the eyes are, as well as head position and angle.
- Linked to other external system monitors that intervene if there is a risk of crashing e.g. lane keeping, collision warning with automatic braking, and adaptive cruise functions



Driver Alert Control (Volvo)

- 2007
- A **RGB camera** detects the side markings painted on the carriageway and compares the section of the road with the driver's steering wheel movements
- The driver is alerted if the vehicle does not follow the carriageway evenly



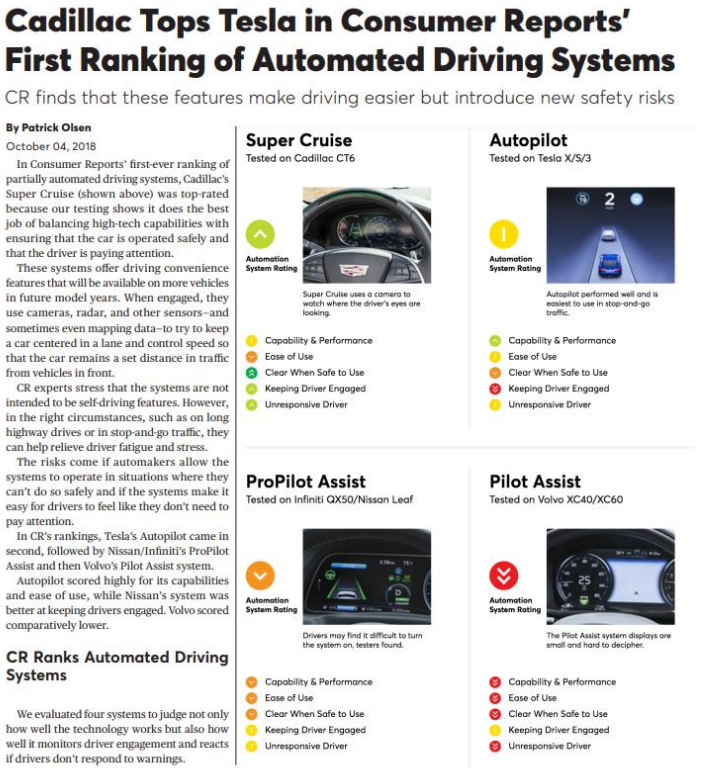


Driver Assistance Systems (BMW)

- In-cabin camera to ensure the driver is paying attention to the road.
- Optical camera, mounted in the instrument cluster, checks to see whether the driver's eyes are open and facing the road.

Super Cruise (Cadillac)

- It uses LIDAR, cameras, sensors, and mapping data to allow drivers taking off their hands from the steering wheel while driving on highways
- It has an infrared camera on the steering column to make sure eyes stay on the road



Cadillac Tops Tesla in Consumer Reports' First Ranking of Automated Driving Systems

CR finds that these features make driving easier but introduce new safety risks

By Patrick Olsen

October 04, 2018

In Consumer Reports' first-ever ranking of partially automated driving systems, Cadillac's Super Cruise (shown above) was top-rated because our testing shows it does the best job of balancing high-tech capabilities with ensuring that the car is operated safely and that the driver is paying attention.

These systems offer driving convenience features that will be available on more vehicles in future model years. When engaged, they use cameras, radar, and other sensors–and sometimes even mapping data–to try to keep a car centered in a lane and control speed so that the car remains a set distance in traffic from vehicles in front.

CR experts stress that the systems are not intended to be self-driving features. However, in the right circumstances, such as on long highway drives or in stop-and-go traffic, they can help relieve driver fatigue and stress.

The risks come if automakers allow the systems to operate in situations where they can't do so safely and if the systems make it easy for drivers to feel like they don't need to pay attention.

In CR's rankings, Tesla's Autopilot came in second, followed by Nissan/Infiniti's ProPilot Assist and then Volvo's Pilot Assist system.

Autopilot scored highly for its capabilities and ease of use, while Nissan's system was better at keeping drivers engaged. Volvo scored comparatively lower.

CR Ranks Automated Driving Systems

We evaluated four systems to judge not only how well the technology works but also how well it monitors driver engagement and reacts if drivers don't respond to warnings.

Super Cruise
Tested on Cadillac CT6

Automation System Rating

Super Cruise uses a camera to watch where the driver's eyes are looking.

- Capability & Performance
- Ease of Use
- Clear When Safe to Use
- Keeping Driver Engaged
- Unresponsive Driver

Autopilot
Tested on Tesla X/S/3

Automation System Rating

Autopilot performed well and is easiest to use in stop-and-go traffic.

- Capability & Performance
- Ease of Use
- Clear When Safe to Use
- Keeping Driver Engaged
- Unresponsive Driver

ProPilot Assist
Tested on Infiniti QX50/Nissan Leaf

Automation System Rating

Drivers may find it difficult to turn the system on, testers found.

- Capability & Performance
- Ease of Use
- Clear When Safe to Use
- Keeping Driver Engaged
- Unresponsive Driver

Pilot Assist
Tested on Volvo XC40/XC60

Automation System Rating

The Pilot Assist system displays are small and hard to decipher.

- Capability & Performance
- Ease of Use
- Clear When Safe to Use
- Keeping Driver Engaged
- Unresponsive Driver

Driver Monitoring System (Subaru)

- The system keeps a watchful eye over drivers to ensure their attention is focused on the road
- Using an infrared LED and a camera, the system monitors the driver for signs of inattention or sleepiness and if detected, warns them to refocus attention
- The system can also recognise up to five individual drivers, memorising their preset preferences and adjusting the cabin environment for both their safety and comfort. It provides:
  - Distraction warning
  - Dozing/Drowsiness Warning
  - Facial Recognition

- **Smart Eye** (http://smarteye.ai)
  Smart Eye develops artificial intelligence powered eye tracking technology that understands, assists and predicts human intentions and actions. By studying a person's eye, face and head movements, Smart Eye technology can draw conclusions about an individual's alertness, attention and focus as well as gain insights into a person's awareness and mental status.

- **Ambarella** (https://www.ambarella.com)
  Ambarella's products are used in a wide variety of human and computer vision applications, including surveillance, Advanced Driver Assistance Systems (ADAS), electronic mirror, drive recorder, driver/cabin monitoring, autonomous driving, and robotic applications. Ambarella's low-power and high-resolution video compression, image processing, and deep neural network processors and software enable cameras to become more intelligent by extracting valuable data from high-resolution video streams.

Driver Fatigue Detection (Škoda-Volkswagen)

- It closely monitors driver behaviour, noting any erratic steering wheel movements, pedal use and any lane deviations (working with Lane Assist), so that it can judge the moment when you start to feel sleepy

- Alerts the driver with a visual display on the dashboard and a warning sound.
  After 15 minutes, the system repeats the warning

Tiredness Recognition System (SEAT)

- The system detects if a driver is feeling tired by analysing steering wheel movements.
- Once detected drowsy behaviour, the system sends an audible and visual alert suggesting the driver takes a break

Driver Attention Alert (Nissan)

- Using steering wheel angle sensors, DAA monitors steering input patterns to establish a baseline or a “snapshot” of how you were driving.

- It continuously compares subsequent driving patterns to the most recent snapshot using a statistical analysis of steering corrections

Driver Attention Alert (Mazda)

- 2015
- After 20 minutes of driving over 65 km/h speed, the system learns the driver's behavior, in order to detect anomalies caused by drowsiness or inattention
- A visual and sound feedback is given to the driver

Attention Assist (Mercedes)

- The system creates an individual profile of the driver, recognizing behavior while the driver is fully alert. That profile is used as the basis for comparison during the rest of the drive.
- Uses a highly sensitive sensor that monitors and records steering movement and speed.

- **Driver Attention Alert** (Peugeot- Citroën):
  - It detects lane changing above a speed of 65 Km/h warning the driver
- **Driver Alert** (Ford):
  - It detects lane changing warning the driver.
- **Driving assistant** (BMW):
  - The Lane Departure Warning detects lane markings and alerts the driver to an unintentional lane change at speeds above approx. 70 km/h by means of vibrations in the steering wheel.

MANUFACTURERS	IN-CABIN VISION SYSTEMS	STEERINGWHEEL SYSTEMS	LANE DETECTION
LEXUS	X (IR camera)		
VOLVO	X (IR camera)		
BMW	X (RGB camera)		X
CADILLAC	X (IR camera)		
SUBARU	X (IR camera)		
ŠKODA-VOLKSWAGEN		X	X
SEAT		X	
NISSAN		X	
MERCEDES		X	
PEUGEOT- CITROËN			X
FORD			X