

# C.A.S.A. in C.V.

Context Aware Security Analytics in  
Computer Vision

**Paola Barra**



UNIVERSITÀ  
DEGLI STUDI  
DI SALERNO



# One to Many



"black and white  
dog jumps over  
bar."

*[karpathy.github.io](http://karpathy.github.io)*

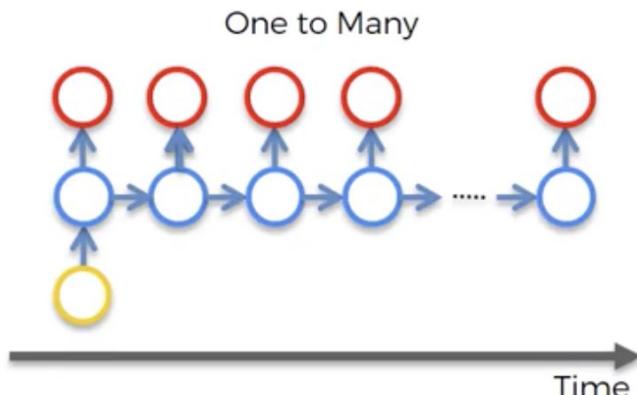
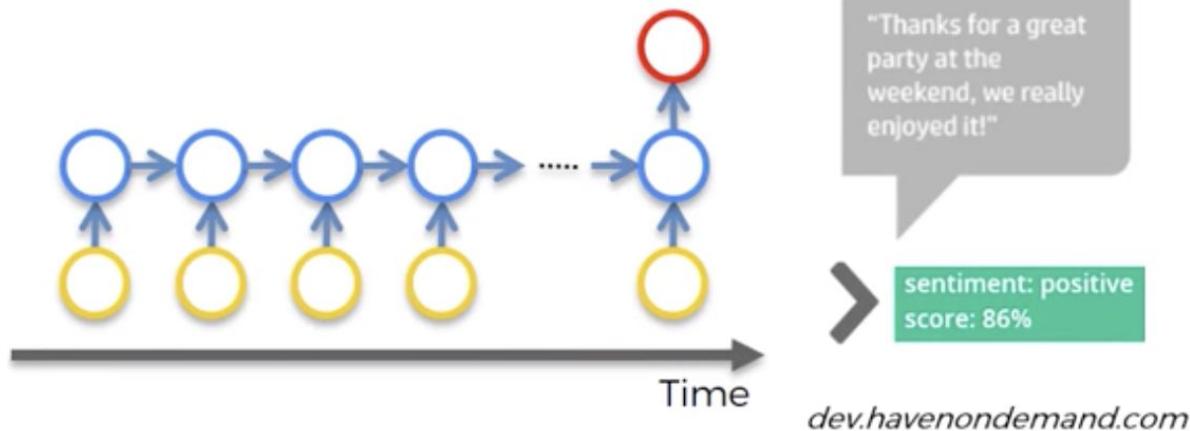


Image in input, a sentence in output = many words.

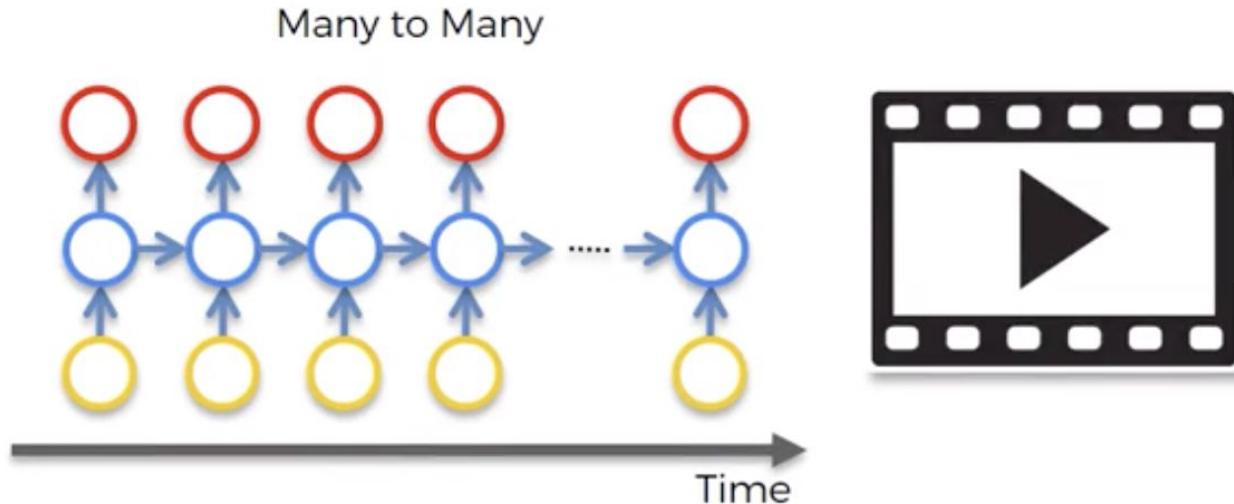
# Many to One

Many to One



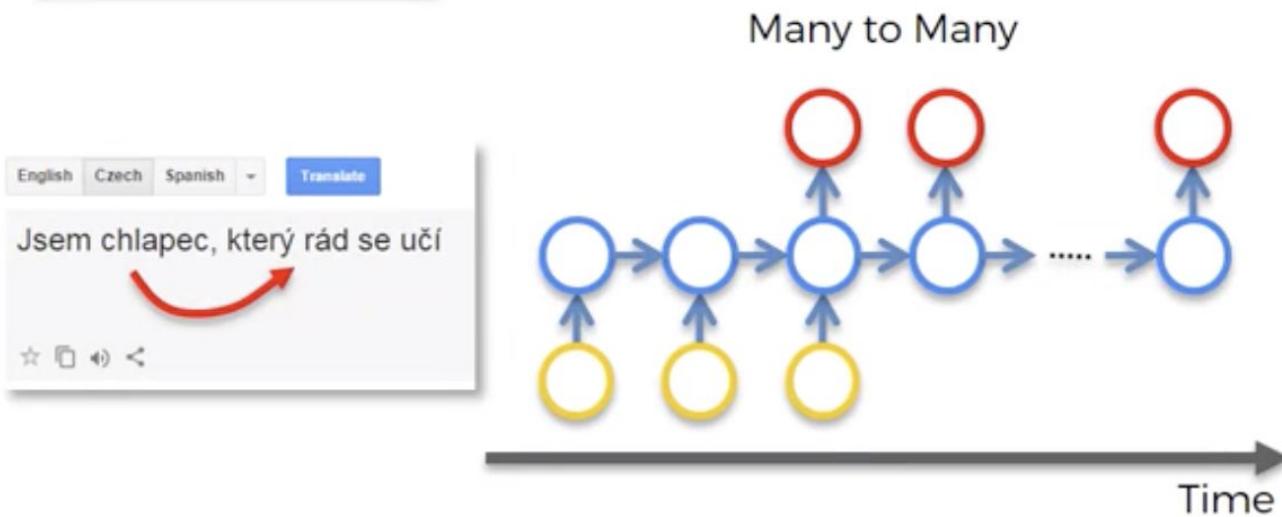
A sentence in input, a sentiment score in output.

# Many to Many



Subtitling a film: a video in input, a sentence in output.

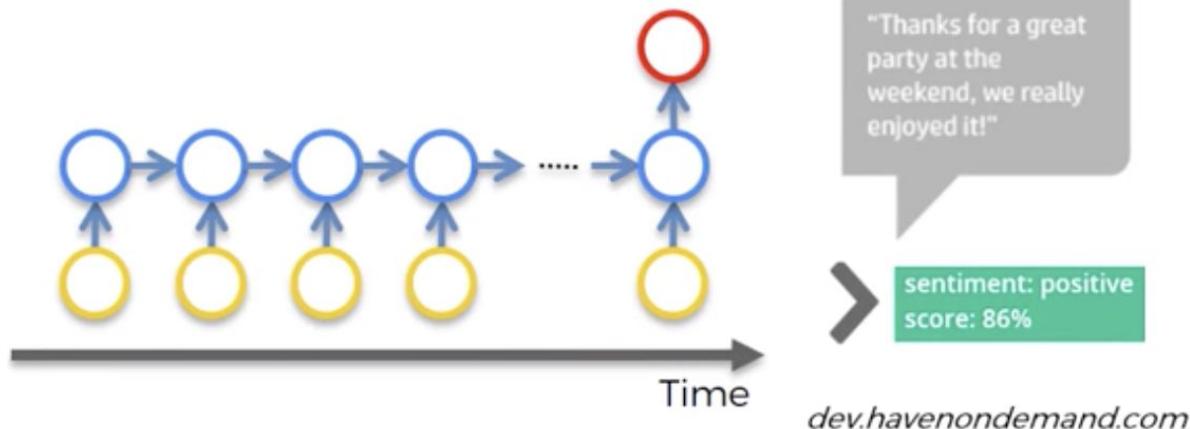
# Many to Many



Google translate: a sentence in input, a sentence in output

# RNN

Many to One





<https://arstechnica.com/gaming/2016/06/an-ai-wrote-this-movie-and-its-strangely-moving/>



UNIVERSITÀ  
DEGLI STUDI  
DI SALERNO



# Re-identification

*“Given an image/video of a person taken from one camera, person re-identification is the task of associating images of the same person taken from different cameras or from the same camera in different occasions.”*

In other words **assigning a stable ID to a person in multi-camera setting**. Usually the re-identification is constrained to a small time period and a small area covered by cameras.

**Humans are easily able to Re-ID** others by leveraging descriptors based on the person's face, height, and build, clothing, hair style, walking pattern, etc but this seemingly easy problem is extremely difficult for a machine to solve.

Camera A

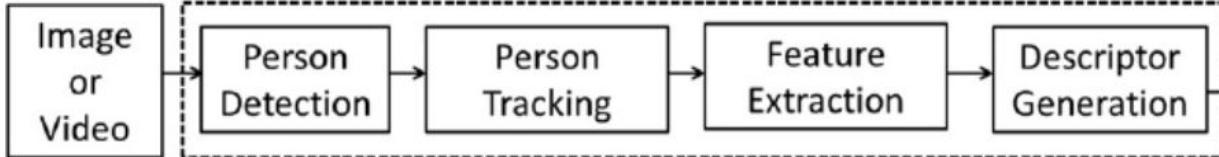
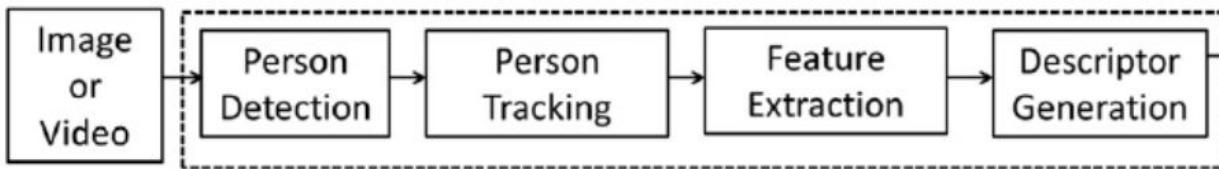


Image  
or  
Video

Camera B



Step 1:

Extraction of person descriptors from multiple cameras

Step 2:

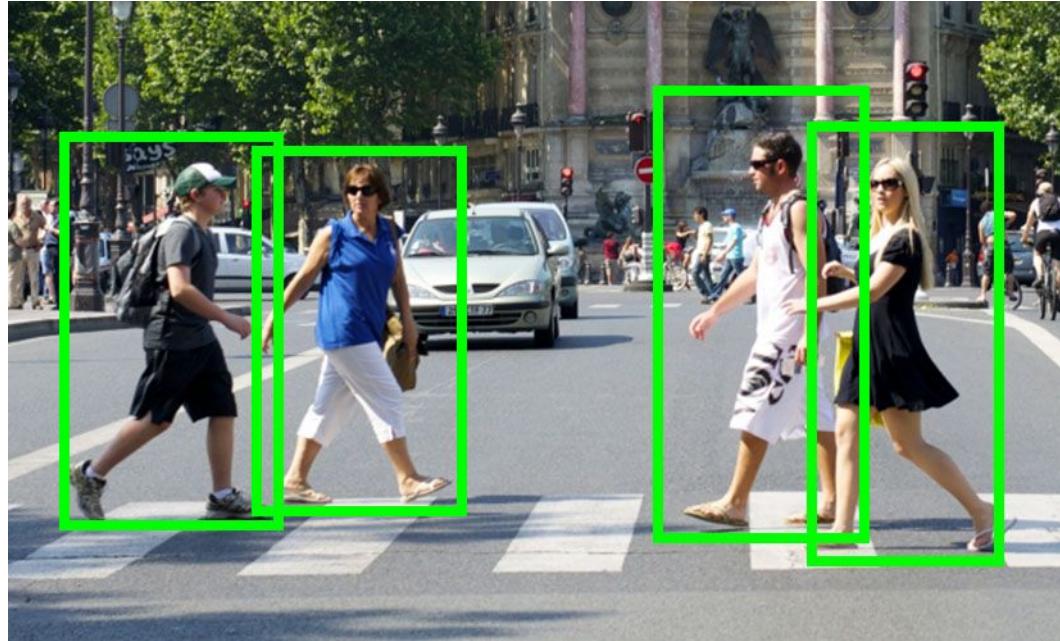
Establish correspondence

# Intra-class variation problem

The **main challenge comes from the intra-class variations**, i.e. the variations in appearance of the same person across different cameras. Main challenges are discussed below.



**1.**Prior to Re-ID the system has to **detect a person** and define the bounding box of the person in an image.



**2.** A Re-ID system may take an image (called single-shot) or a video (multi-shot) as input. In a video input we need to be able to **establish correspondence between detected subjects across frames**. This process is called **tracking**. Tracking multiple persons is also a challenging task.

**3. Illumination changes.** Intensity of daylight, shade, reflected light from colored surfaces, indoor lighting can cause the same subject to appear in different shades and colors across cameras

**4. Low resolution.** Many old CCTV systems are with cameras of low resolution. Due to the lack of information person Re-ID becomes even more difficult

**5. Occlusion.** In crowded environments partial or even complete occlusion of persons by others presents challenge in extracting features

- 6. Uniform clothing** at schools and even some workplaces will confuse Re-ID algorithms which extract information from clothing/appearance
- 7. Scalability.** Public areas are covered by thousands of cameras and current technologies are only beginning to address multi-camera surveillance problem

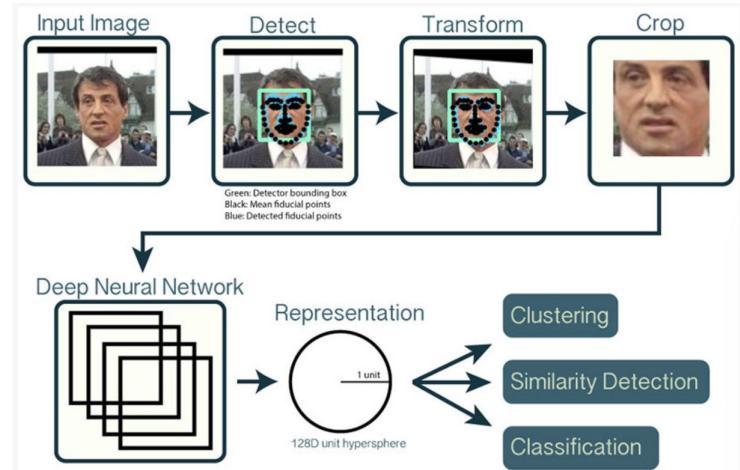
See tutorial for Re-identification here:

<https://medium.com/@niruhan/a-practical-guide-to-person-re-identification-using-alignedreid-7683222da644>

# OpenCV Face recognition

How we can apply deep learning and OpenCV together (with no other libraries other than scikit-learn)?:

1. Detect faces
2. Compute 128-d face embeddings to quantify a face
3. Train a Support Vector Machine (SVM) on top of the embeddings
4. Recognize faces in images and video streams



# Pose Detection

There are many state-of-the-art pose detection software. You are free to use what you prefer:

- Open Pose <https://github.com/CMU-Perceptual-Computing-Lab/openpose>
- wrnchAI <https://www.learnopencv.com/pose-detection-comparison-wrnchai-vs-openpose/>
- BIPLAB software (you'll find it in the folder)

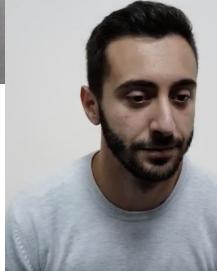
If you need you can use the PC of our laboratory to implement this.

# wrnchAI vs OpenPose



# Project 1 : Person re-identification

Build a NN to recognize a person only by a photo of the face.



Input images.

# Project 1 : Person re-identification

Steps:

- Download Gotcha dataset <https://gotchaproject.github.io/>
- extract the images of the face or the body of a subject (Indoor).
- follow the instructions here to implement a re-identification  
<https://medium.com/@niruhan/a-practical-guide-to-person-re-identification-using-alignedreid-7683222da644>
- perform the re-identification on outdoor video.

Useful code:

- face extraction: see Face detector file.

If you need help just contact me at [pbarra@unisa.it](mailto:pbarra@unisa.it)

# Project 2 : Action classification walk/stairs

Build a NN to classify the action “walking” or “climbing the stairs”.



stairs

stairs

stairs

stairs

walk

stairs

# Project 2 : Action classification walk/stairs

Steps:

- Download Gotcha dataset <https://gotchaproject.github.io/>
- extract the gait of the person walking
- extract the gait of the person climbing the stairs.
- perform a NN to classify if a person in a video is walking or is climbing the stairs

If you need help just contact me at [pbarra@unisa.it](mailto:pbarra@unisa.it)

# Project 3 : Action classification fight

Build a NN to classify the action “fighting” or “not fighting”.

Dataset here: <https://www.kaggle.com/mohamedmustafa/real-life-violence-situations-dataset/download>



**fight**



**non fight**