Computer vision methods timeline

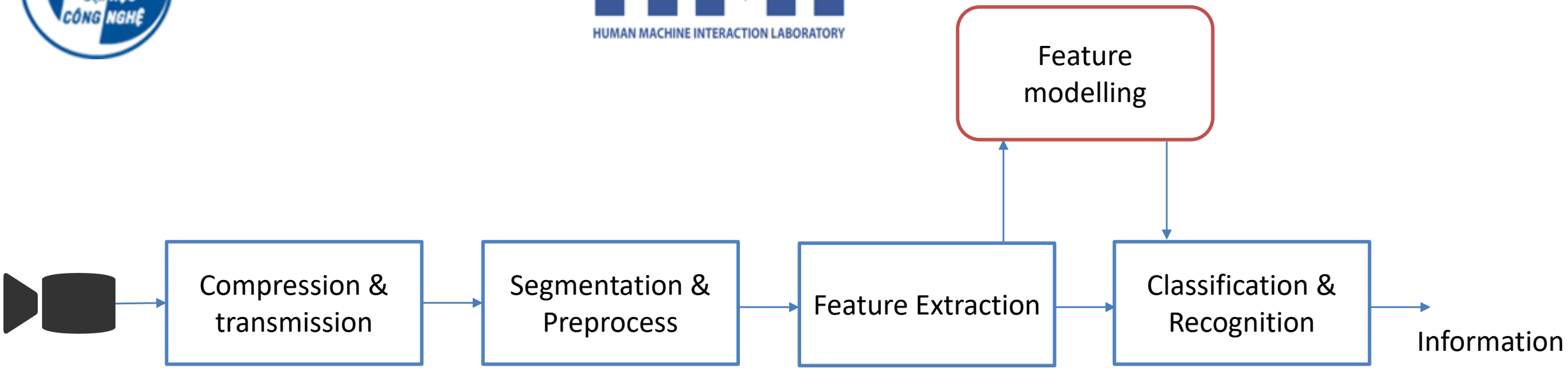
Le Thanh Ha, Ph.D
Assoc. Prof. at University of Engineering and Technology,
Vietnam National University
ltha@vnu.edu.vn; lthavnu@gmail.com; 0983 692 592

HMI
HUMAN MACHINE INTERACTION LABORATORY

Computer vision methods timeline

- Before 1995: Template matching
  - Match patches to find the correspondence between images.
- 1995 – 2000: Local feature descriptor
  - Scale Invariant Feature Transform (SIFT)
- 2000 – 2005: Machine learning for feature descriptors
  - Modern image dataset
  - Bins, Grids, and Visual Words
- 2005 – 2010: Object template
  - Histogram of Oriented Gradients (HOG)
  - Deformable Parts-based Model (DPM)
- 2010 – now: Big data, Convolutional network, Deep learning

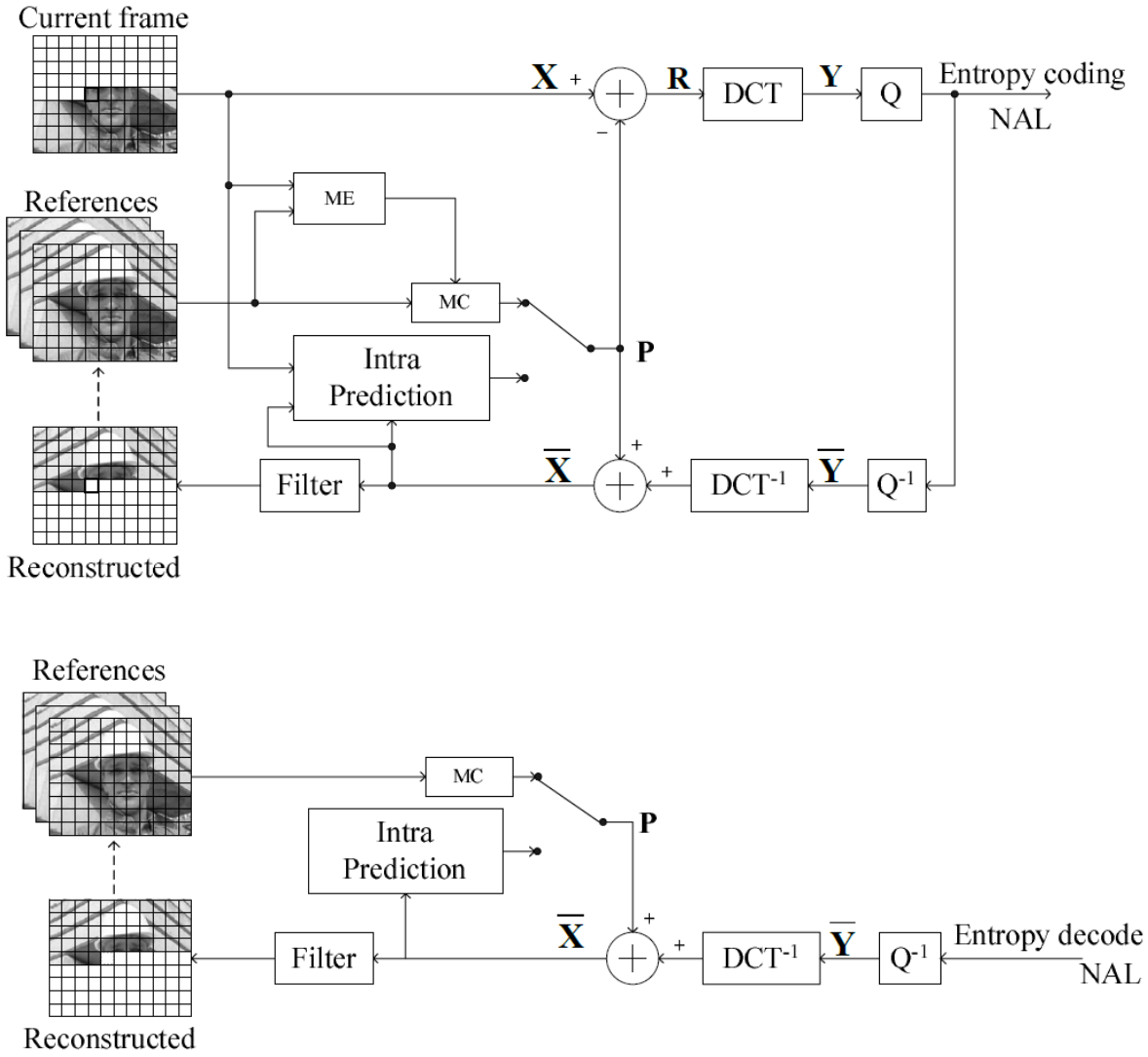
HMI
HUMAN MACHINE INTERACTION LABORATORY



TYPICAL IMAGE PROCESSING WORKFLOW

Compression & Transmission

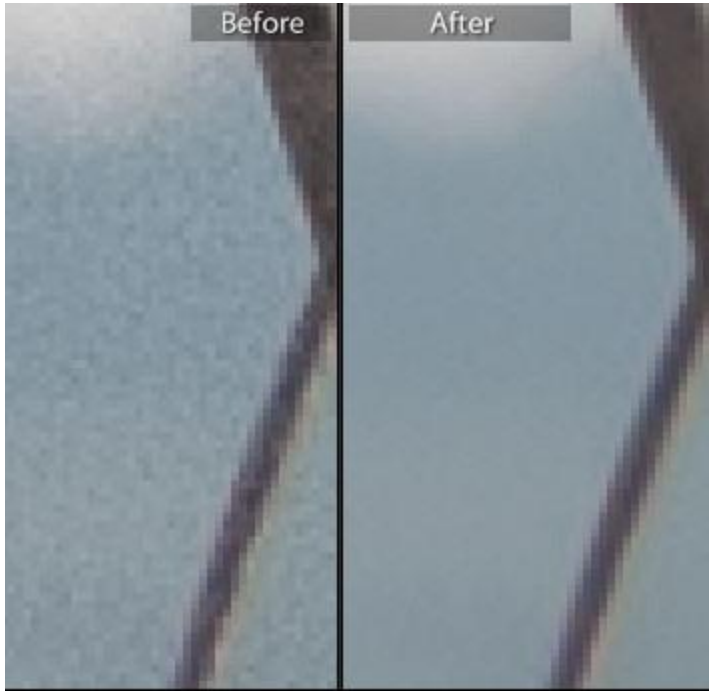
- Reduce the size of video presentation to be sent effectively over error-prone network or to be stored
- Smallest in size, best in quality
- Tolerate with network conditions
- H.264/AVC, H.265/HEVC, VP9, Raspro



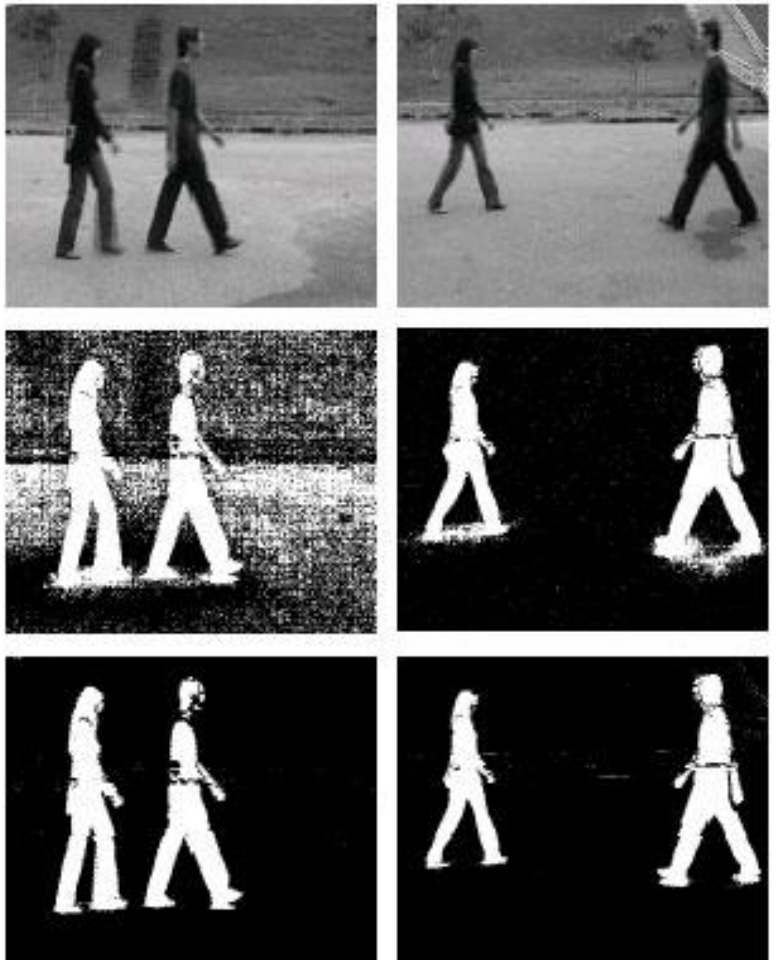
Segmentation & Preprocess

- Noise removal
- Image enhancement
- Temporal segmentation: scene cut, key frames,…
- Spatial segmentation: grabcut, background subtraction

Noise removal



Bankground subtraction



Feature Extraction

- Target object must be detected and described with small number of dimensions. It should be:
  - Robustness
  - Repeatable
  - Noise tolerable
  - Invariant to illumination, rotation, scaling, and view point.

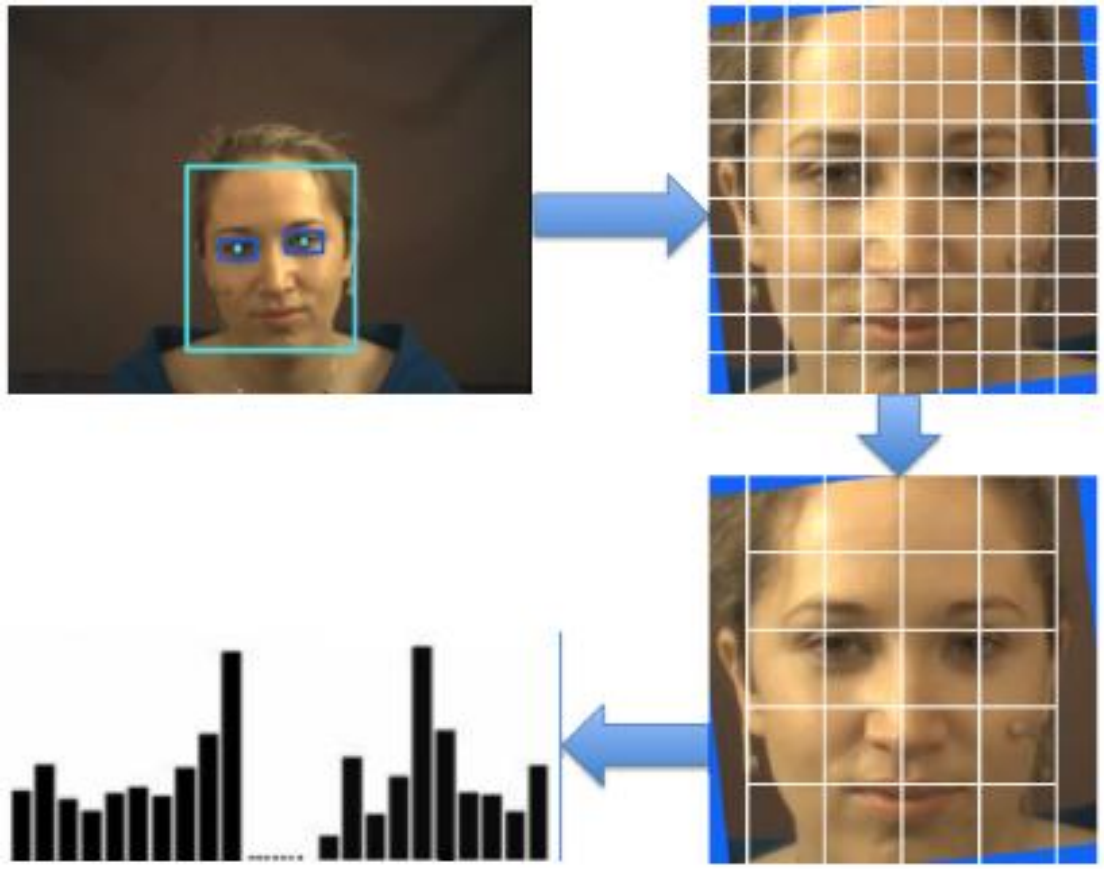
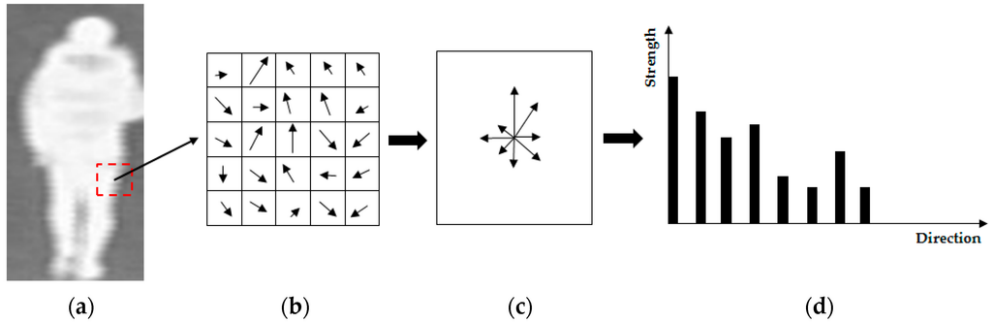
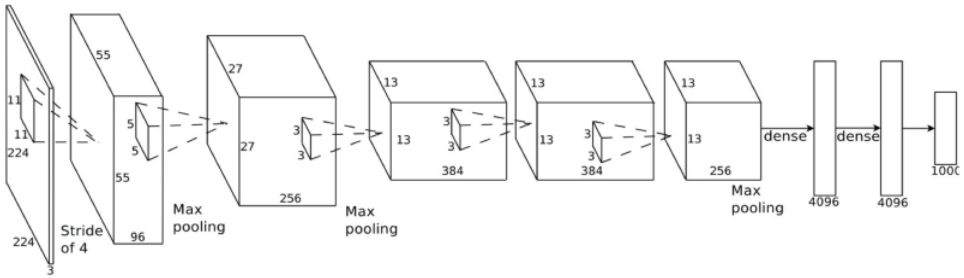


Image feature examples

- Histogram based: Histogram of oriented gradients (HOG)



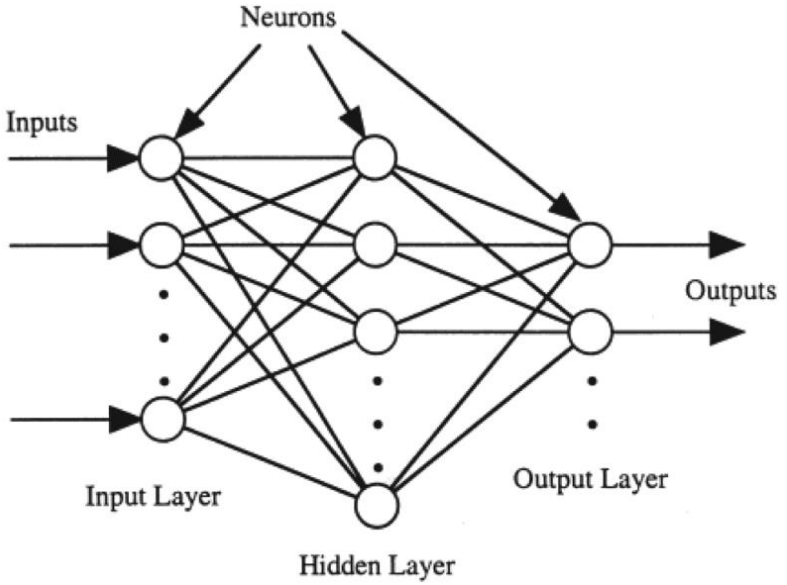
- Feature points:
  - Harris corners
  - Sussan corner
- Feature points description:
  - SIFT
  - SURF

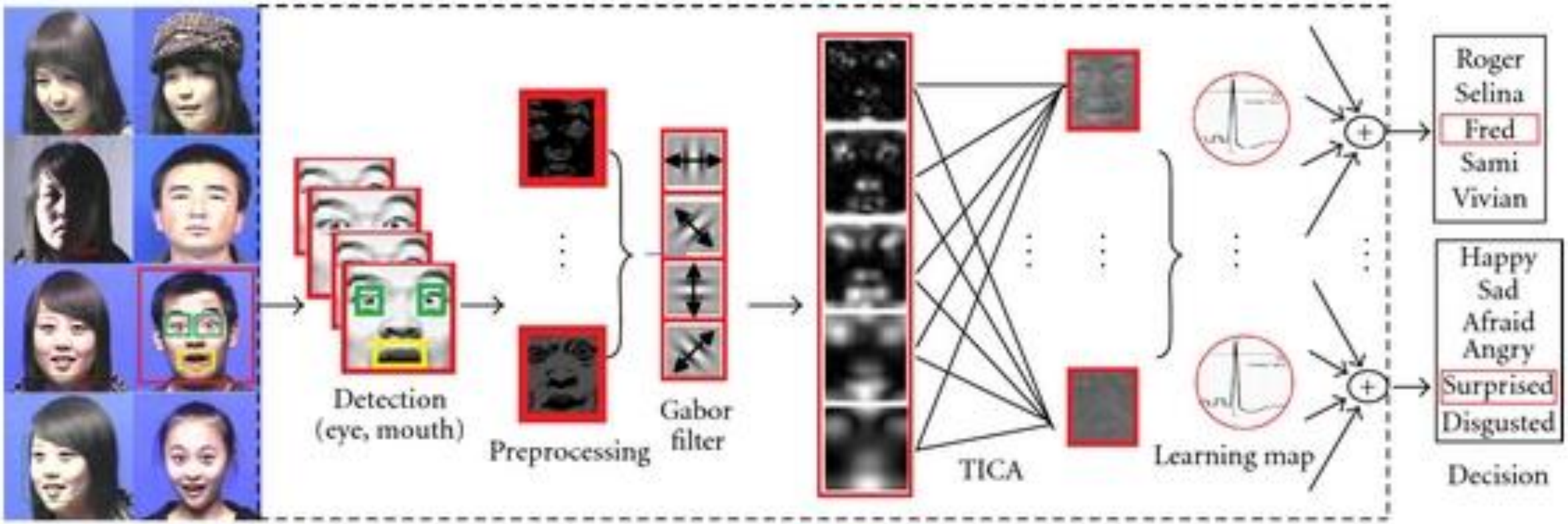


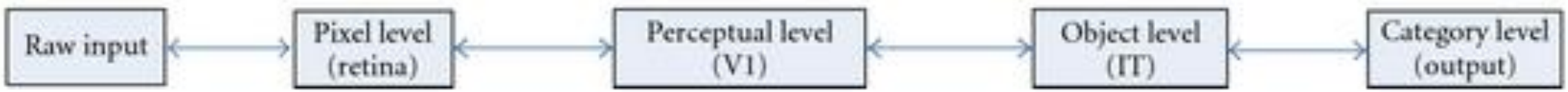
Deep learning

Feature modeling/learning

- A training dataset must be collected.
- A model must be used to generalize the input features and map it to desired output (class):
  - Support Vector Machines
  - Neuron networks
  - Logistics
  - Decision tree
  - Ada-boost
  - …
- After trained, the model can be used to classify the testing data.




Detection
(eye, mouth)
Preprocessing
Gabor
filter
TICA
Learning map
Roger
Selina
Fred
Sami
Vivian
Happy
Sad
Afraid
Angry
Surprised
Disgusted
Decision


Raw input
Pixel level
(retina)
Perceptual level
(V1)
Object level
(IT)
Category level
(output)