Object Detection

Objectives

- What is Object Detection?
- The techniques used in object detection.
- Learn about face detection
- Pedestrian detection
- Applications of face detection
- Object detection challenges

What is Object Detection?

Object detection???

What is Object Detection?

- Object detection is the task of detecting instances of objects of a certain class within an image.
- Object detection allows us to identify and locate objects in an image or video.
- Object detection deals with detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos.
- Object detection is merely to recognize the object with a bounding box in the image.

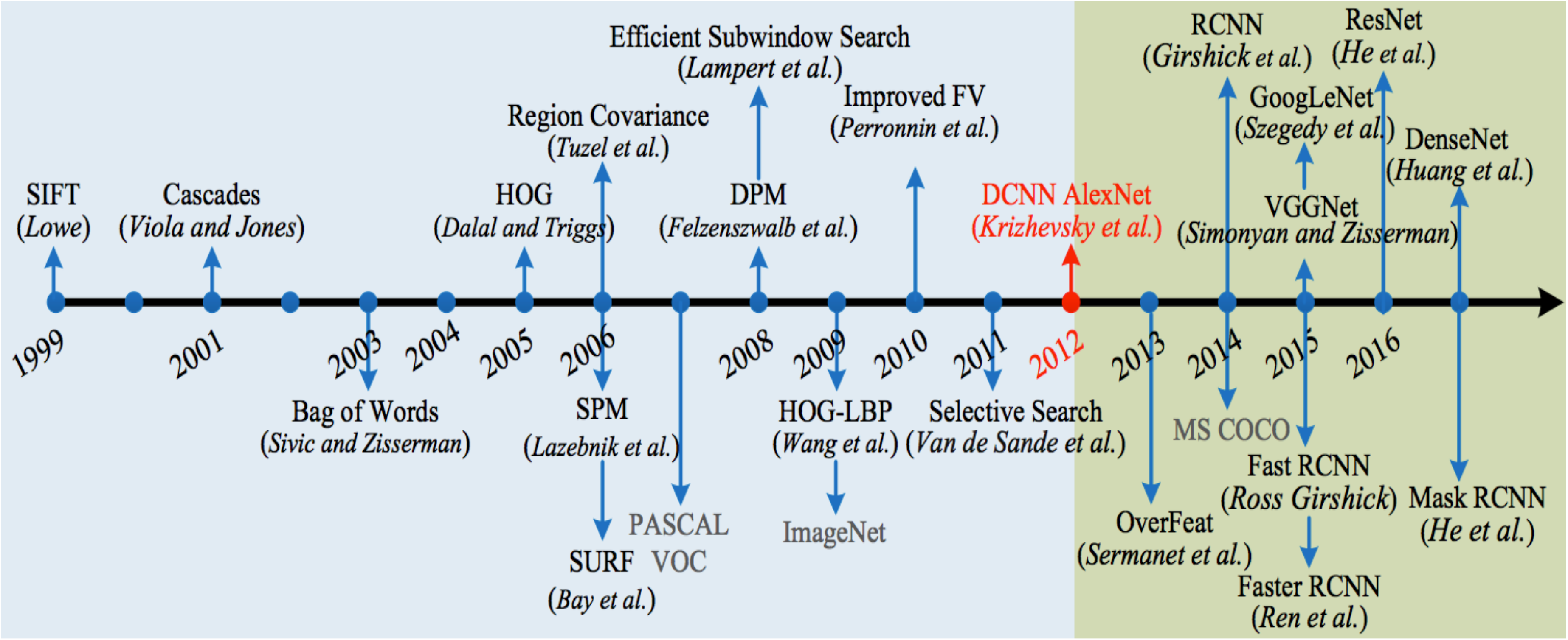
Object Detection Concept

- Every object class has its own special features that help in classifying the class – Object class detection uses object special features
- Examples:
  - Circles: objects that are at a particular distance from a center.
  - Squares, objects that are perpendicular at corners and have equal side lengths.
  - Face: eyes, nose, and lips can be found, and features like skin color and distance between eyes

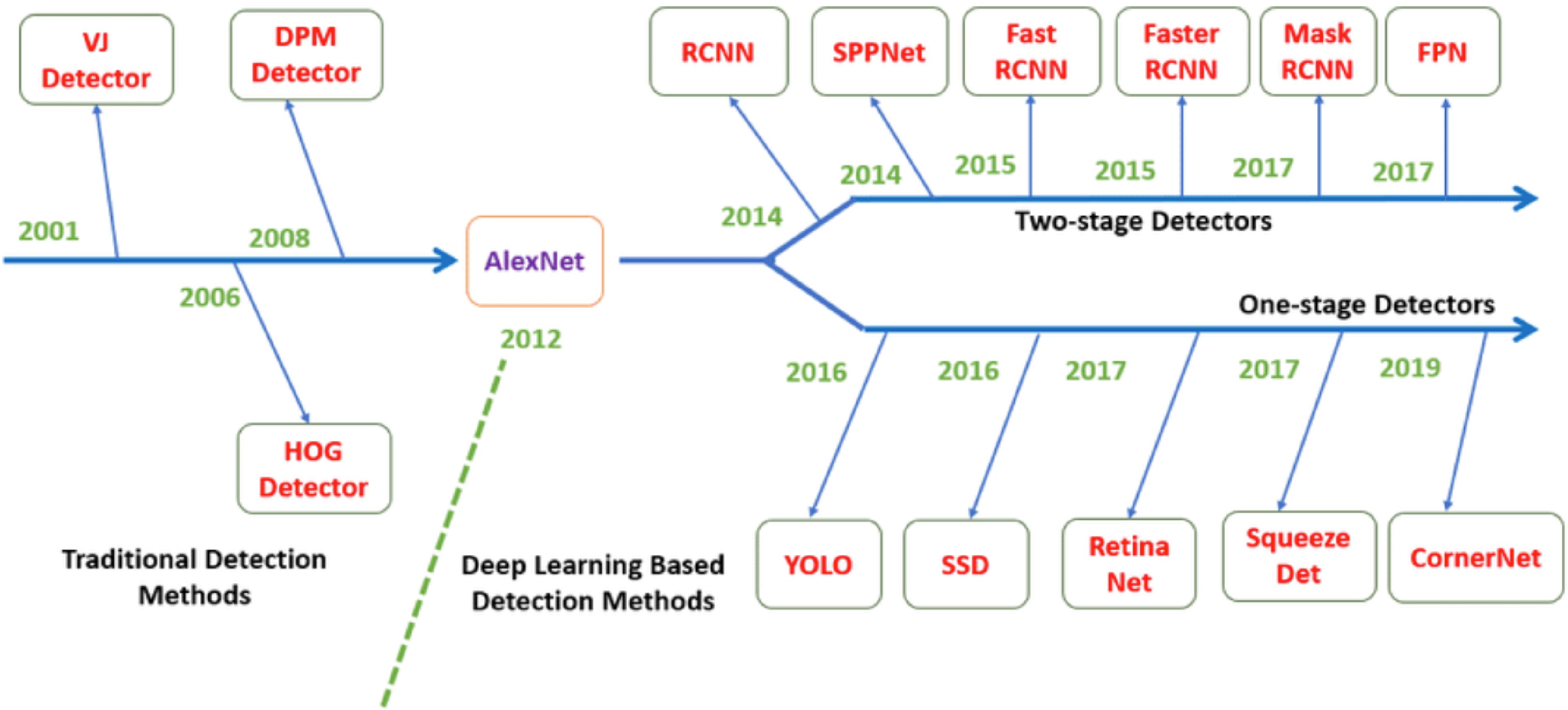
Object Detection methods

- Non-neural approaches:
  - Viola–Jones object detection framework based on Haar features
  - Scale-invariant feature transform (SIFT)
  - Histogram of oriented gradients (HOG) features
- Neural network approaches:
  - Region Proposals (R-CNN, Fast R-CNN, Faster R-CNN, cascade R-CNN)
  - Single Shot MultiBox Detector (SSD)
  - You Only Look Once (YOLO)
  - Single-Shot Refinement Neural Network for Object Detection (RefineDe)
  - Deformable convolutional networks

Object Detection methods

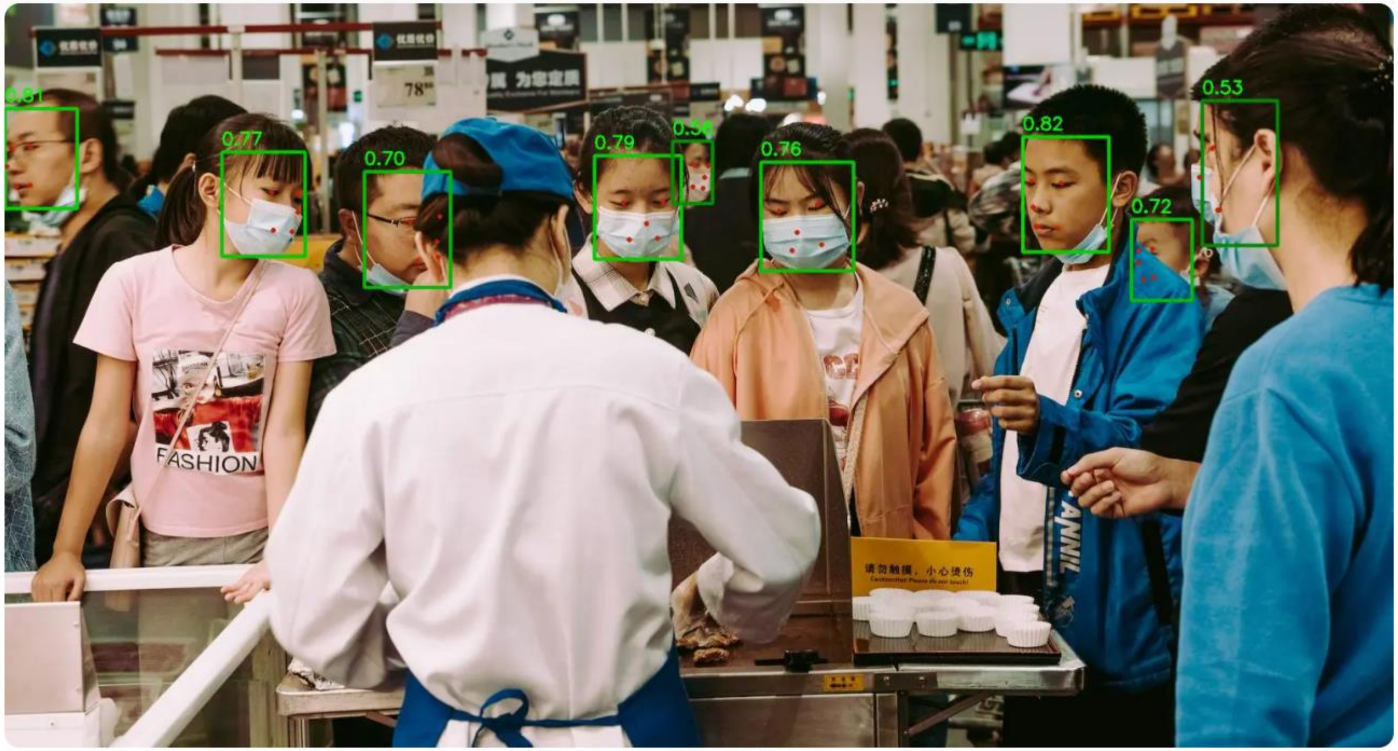


Object Detection methods



Face Detection

- The face detection system will confirm detection by overlaying a rectangle on each face in the scene displayed on the camera's LCD

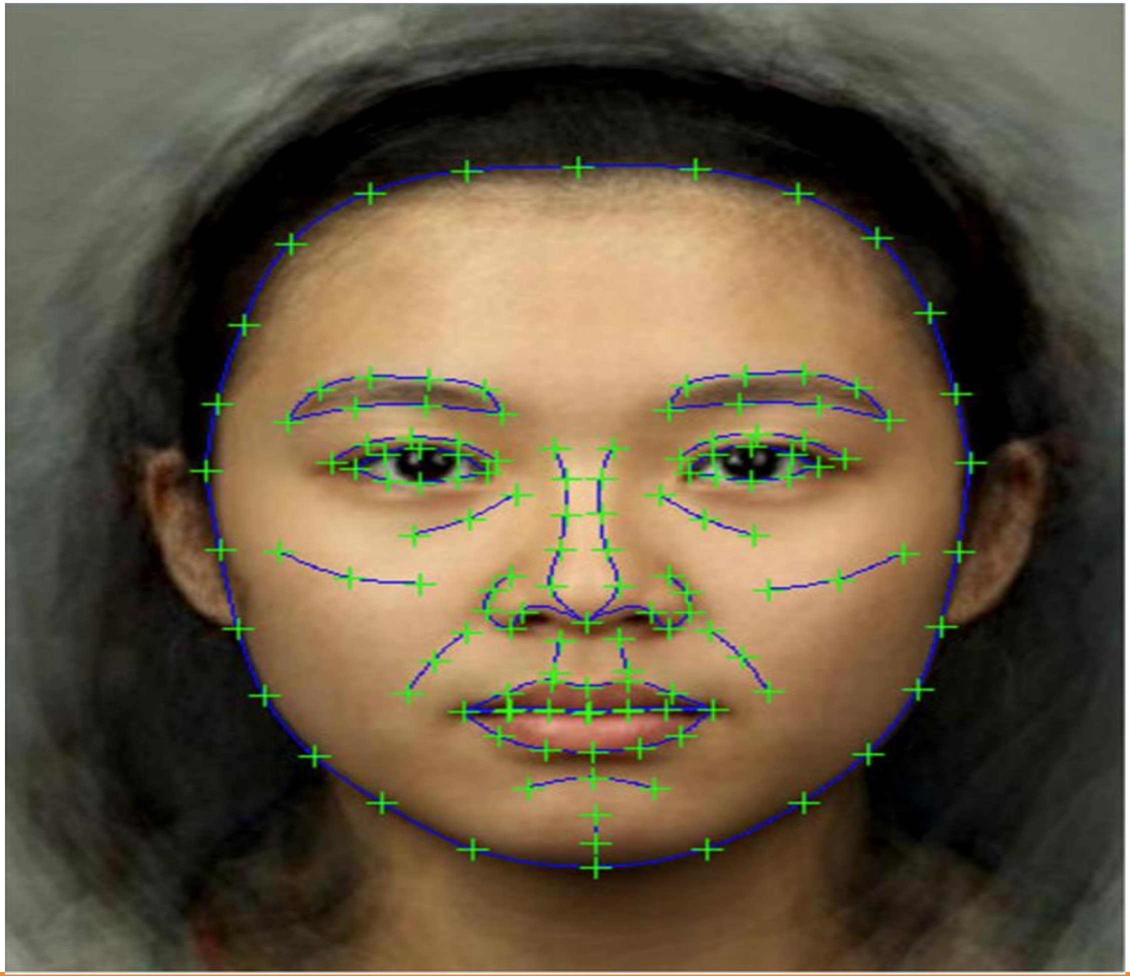


Face Detection

- Face detection is that identifies human faces in digital image. It refers to the psychological process by which humans locate and attend to faces in a visual scene.
- Face-detection algorithms focus on the detection of frontal human faces.
- Face Features:
  - Eyes
  - Eyebrows
  - Nose
  - Mouth
  - Jaw

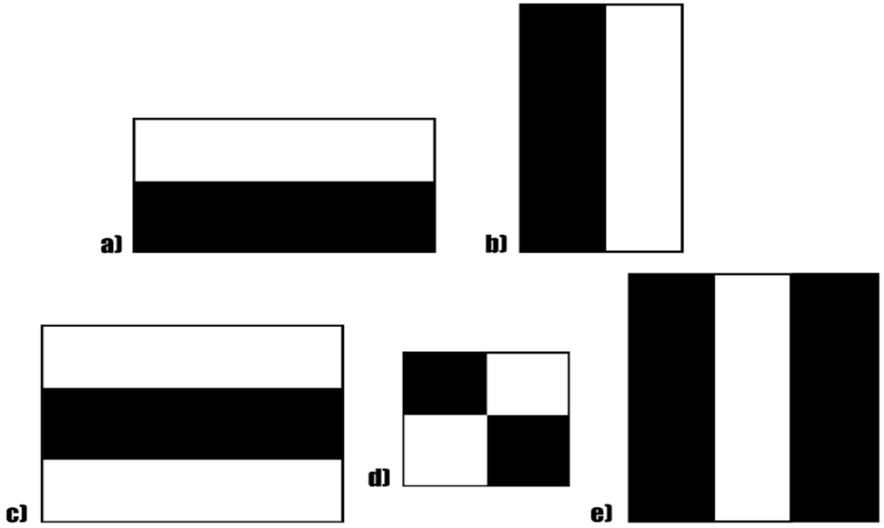
Face Detection

- Face Features:
  - Eyes
  - Eyebrows
  - Nose
  - Mouth
  - Jaw



Face Detection - Haar Cascade

- The algorithm uses edge or line detection features
- These features on the image makes it easy to find out the edges or the lines in the image, or to pick areas where there is a sudden change in the intensities of the pixels.

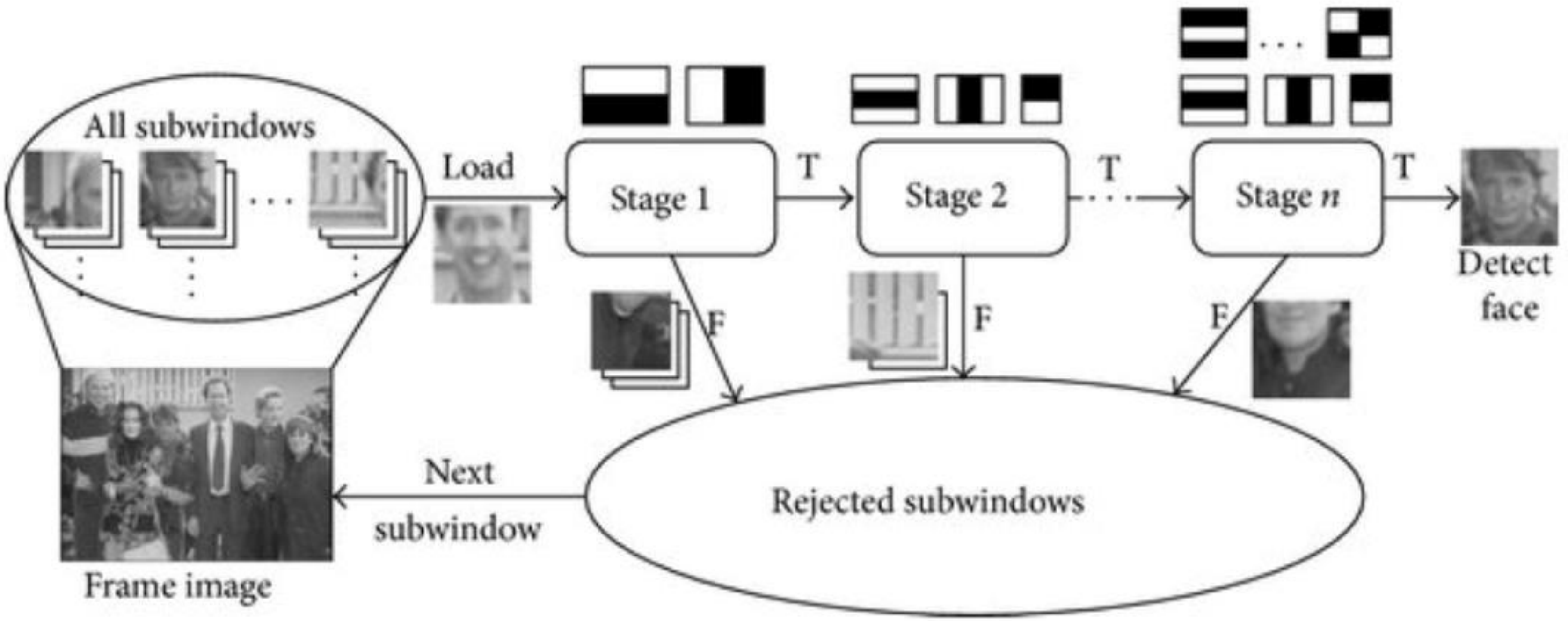


Face Detection - Haar Cascade

- The haar calculation is done by finding out the difference of the average of the pixel values at the darker region and the average of the pixel values at the lighter region. If the difference is close to 1, then there is an edge detected.

0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1

Face Detection - Haar Cascade

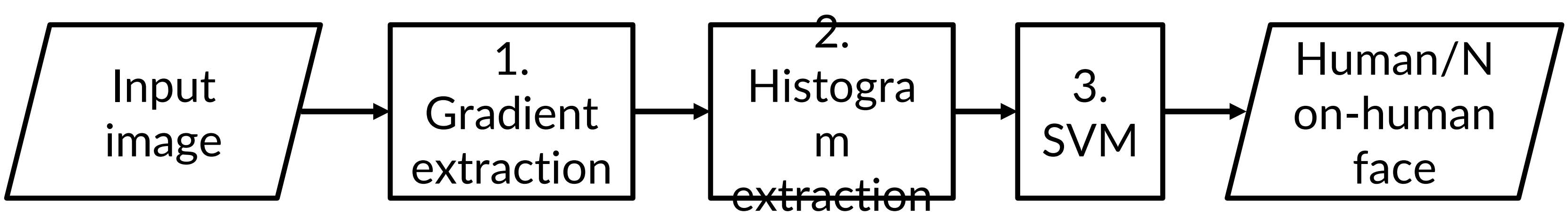


Face Detection - Haar Cascade

– Steps:

- Step 1: The image (that has been sent to the classifier) is divided into small parts (or subwindows as shown in the illustration)
- Step 2: We put N no of detectors in a cascading manner where each learns a combination of different types of features from images are passed through. Supposedly when the feature extraction is done each sub-part is assigned a confidence value.
- Step 3: Images (or sub-images) with the highest confidence are detected as face and are sent to the accumulator while the rest are rejected. Thus the cascade fetches the next frame/image if remaining and starts the process again.

Face Detection
- Histogram of oriented gradients

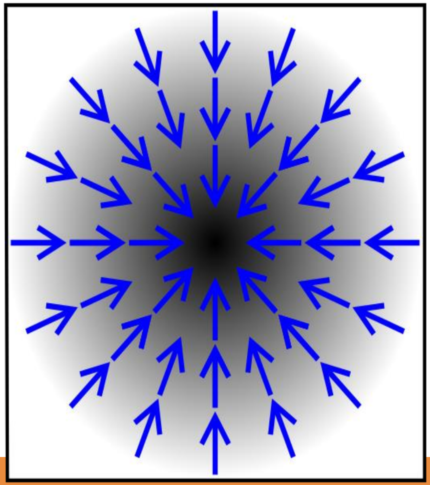


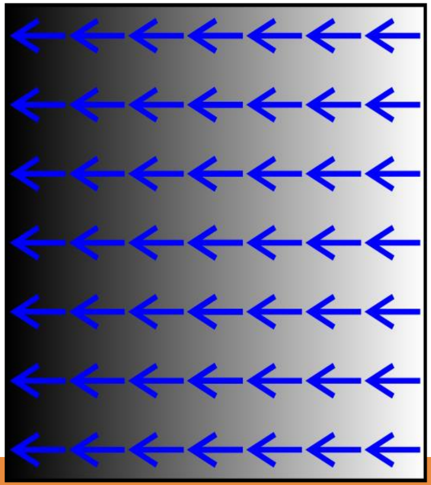
Face Detection
- Histogram of oriented gradients

FPT Education
FPT UNIVERSITY

- An image gradient is a directional change in the intensity or color in an image.
- Gradients can be used for extracting the useful information (structure, feature, and properties of objects) from images.
- Purpose: To calculate the magnitude and direction of gradient

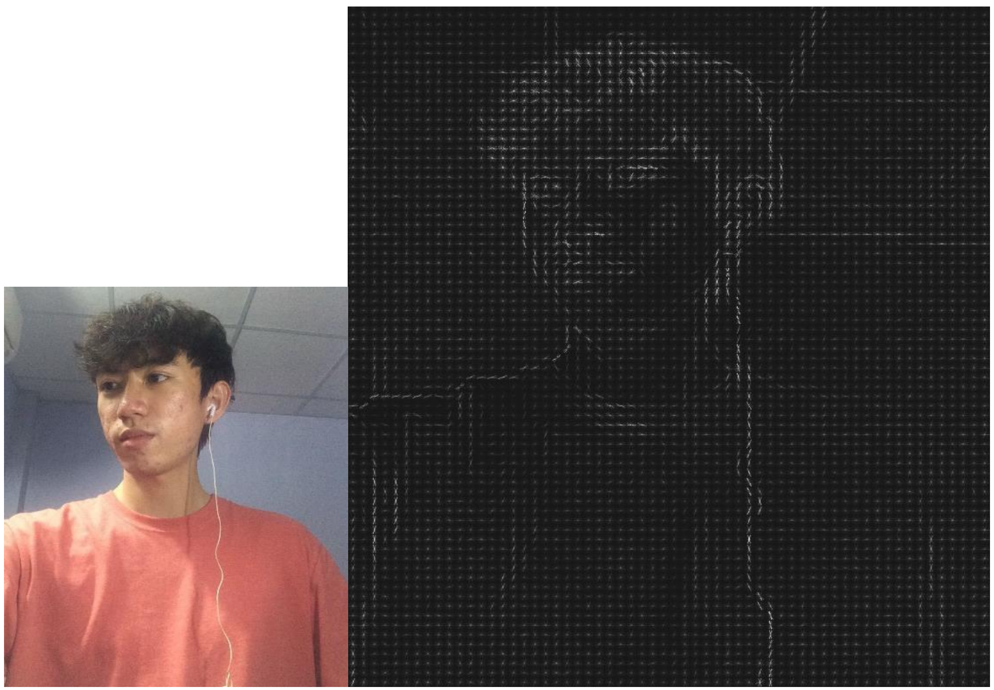
Two types of gradients, with **blue arrows** to **indicate** the **direction** of the gradient. **Dark areas** indicate **higher values**

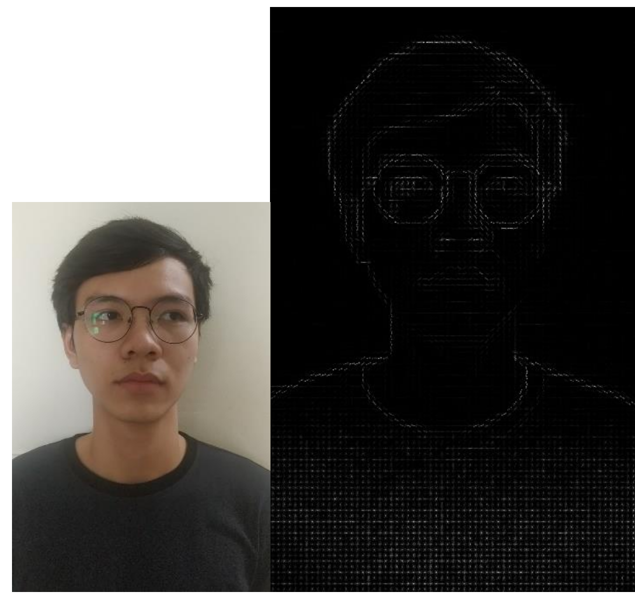




Face Detection
- Histogram of oriented gradients

- In face images, the important features of the size, shape, and face orientation are in points of the lines of eyes, nose, lips, and cheekbones

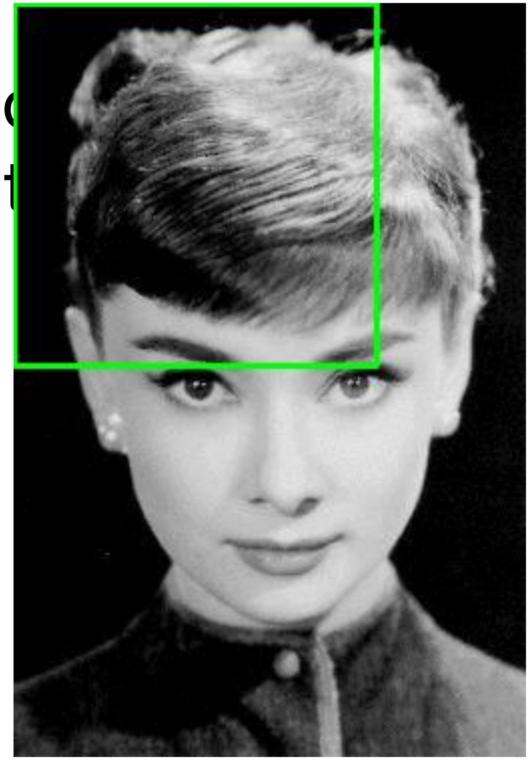




Images after calculating magnitude and direction

Face Detection
- Histogram of oriented gradients

- We use a window slide on frame from left to right and top to bottom to calculate magnitude and direction for every pixel
- To detect face from multiple scales and slide with 4 sizes of window concurrent
  - 108x108
  - 90x90
  - 74x74
  - 62x62

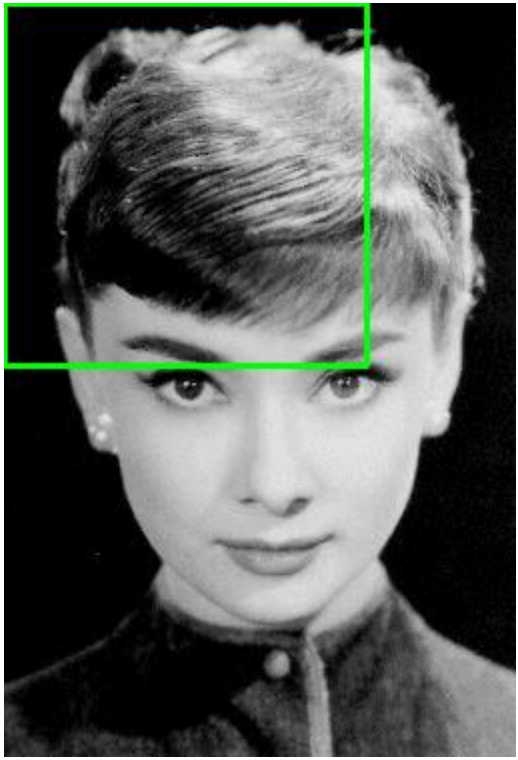


Face Detection
- Histogram of oriented gradients

- HG and VG is directional gradient with:
  - HG is Horizontal gradient
  - VG is Vertical gradient

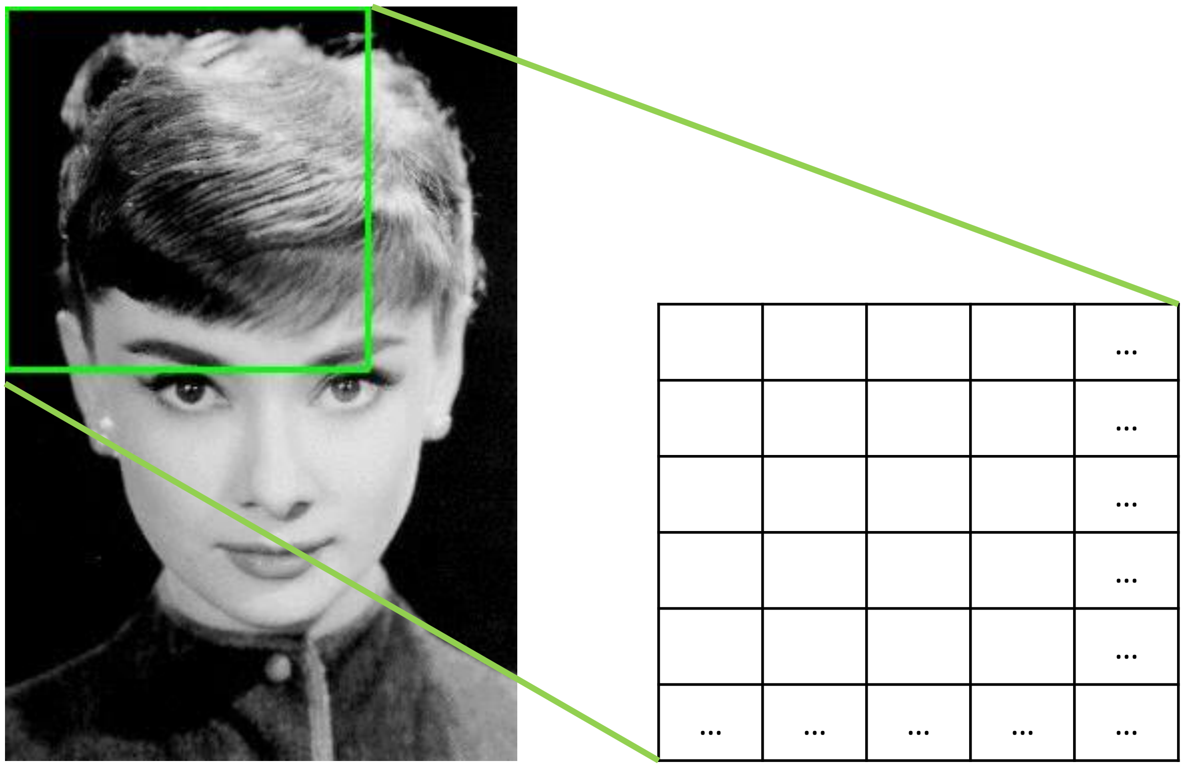
$$Gradient\ Magnitude = \sqrt{HG^2 + VG^2}$$

$$Gradient\ Direction = tan^{-1}\left(\frac{HG}{VG}\right)$$



Face Detection
- Histogram of oriented gradients

- An



108x108 kernel

-1	0	1

Centred

-1	1

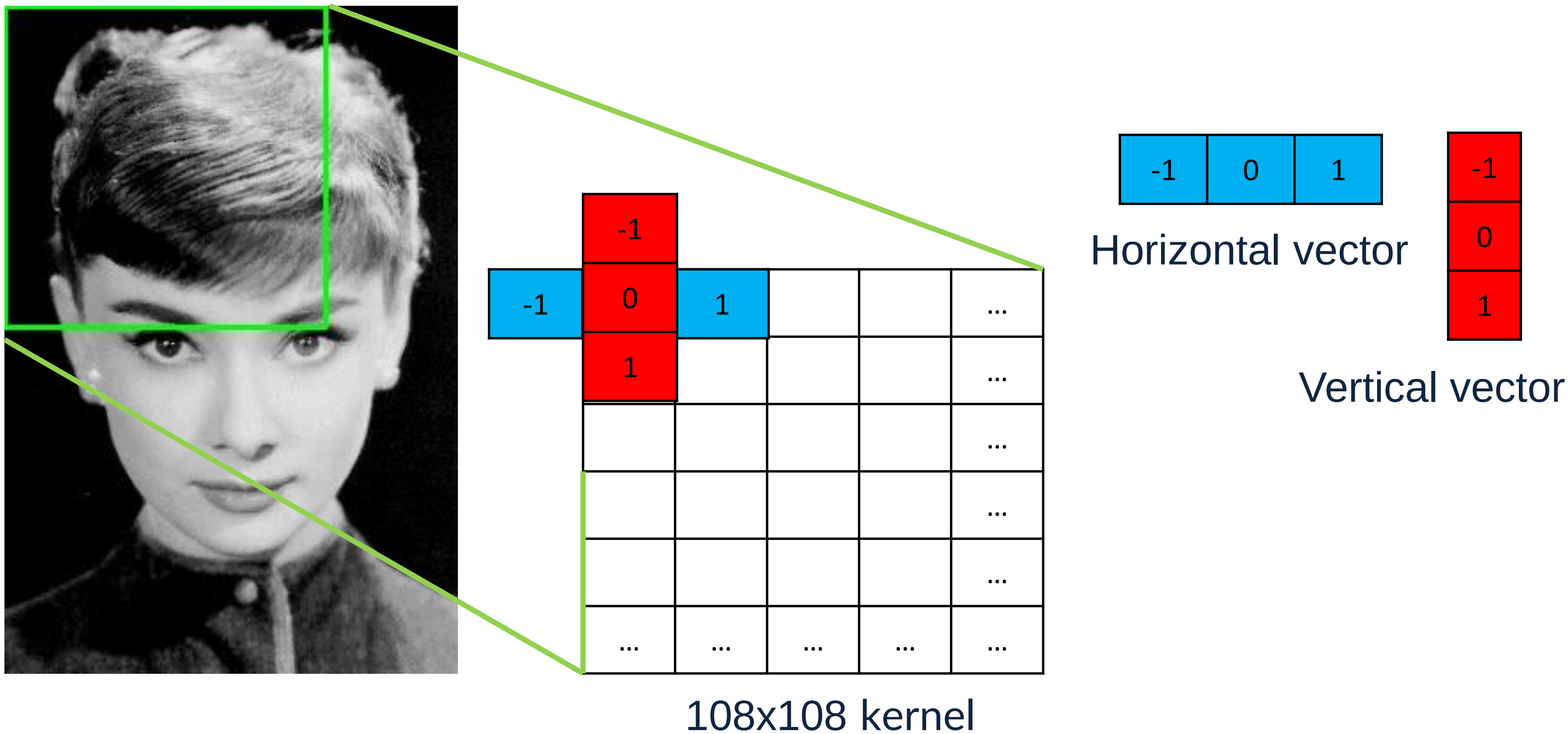
Un-centred

1	-8	0	8	-1

Cubic-corrected

Slide directional gradient on kernel

Face Detection - Histogram of oriented gradients



Slide directional gradient on kernel

Face Detection
- Histogram of oriented gradients

- Using **dot product** to **calculate directional gradient**
- The **dot product** of **two vectors** **a** = [a_1, a_2, …, a_n] and **b** = [b_1, b_2, …, b_n] is defined as:

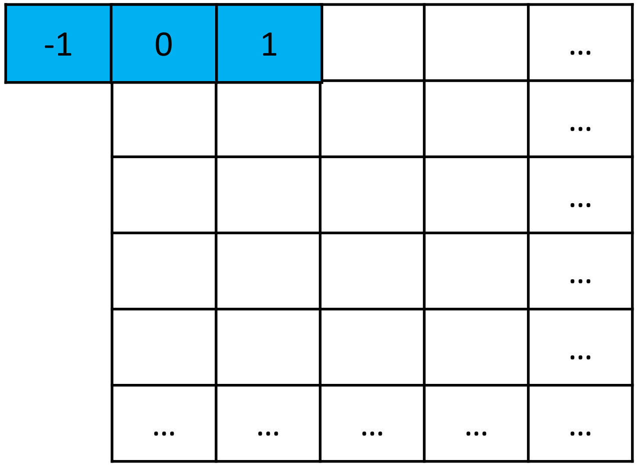
-1	0	1			…
					…
					…
					…
					…
	…	…	…	…	…

108x108 kernel

$$\boldsymbol{a} \cdot \boldsymbol{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

Face Detection
- Histogram of oriented gradients

- Using **dot product** to **calculate directional gradient**
- The **dot product** of **two vectors** **a** = [a_1, a_2, …, a_n] and **b** = [b_1, b_2, …, b_n] is defined as:



108x108 kernel

$$\boldsymbol{a \cdot b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

n = 3

a:

-1	0	1

Horizontal vector

b:

a	b	c

3x3 vector at slide position

$$\boldsymbol{a \cdot b} = \sum_{i=1}^{3} a_i b_i = (-1 \ * a) + (0 \ * b) + (1 * c)$$

Face Detection
- Histogram of oriented gradients

What is HOG?

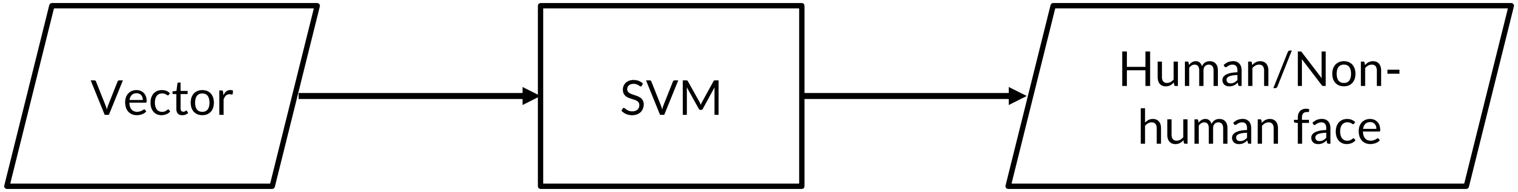
- **HOG** is a **representational vector** of an image that **simplifies** the image by extracting **useful information**
- **HOG** is **calculated** from **magnitude** and **direction** gradients
- The histogram **contains 9 elements** corresponding to **angles 0, 20, 40 … 160**
- **Each element** is selected based on **the direction**, and the **value** that goes into the bin is selected based on **the magnitude**

0	20	40	60	80	100	120	140	160

Histogram of Oriented Gradient (HOG)

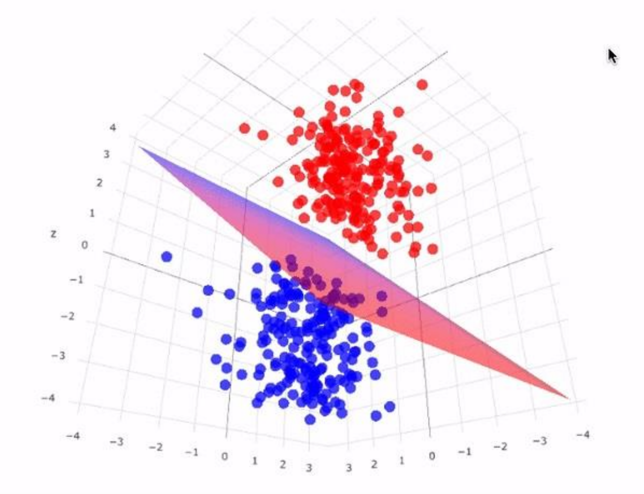
Face Detection
- Histogram of oriented gradients

- Support Vector Machine (SVM) is a supervised learning algorithm that we use to predict a vector is human or non-human face
- Input: a vector
- Output: Human or non-human face



Face Detection - Histogram of oriented gradients

- How SVM works?
  - In training phase, SVM finds a optimal hyperplane that divides the dataset into 2 parts (positive and negative)
  - In prediction phase, SVM uses the hyperplane to finds the part that vector belongs

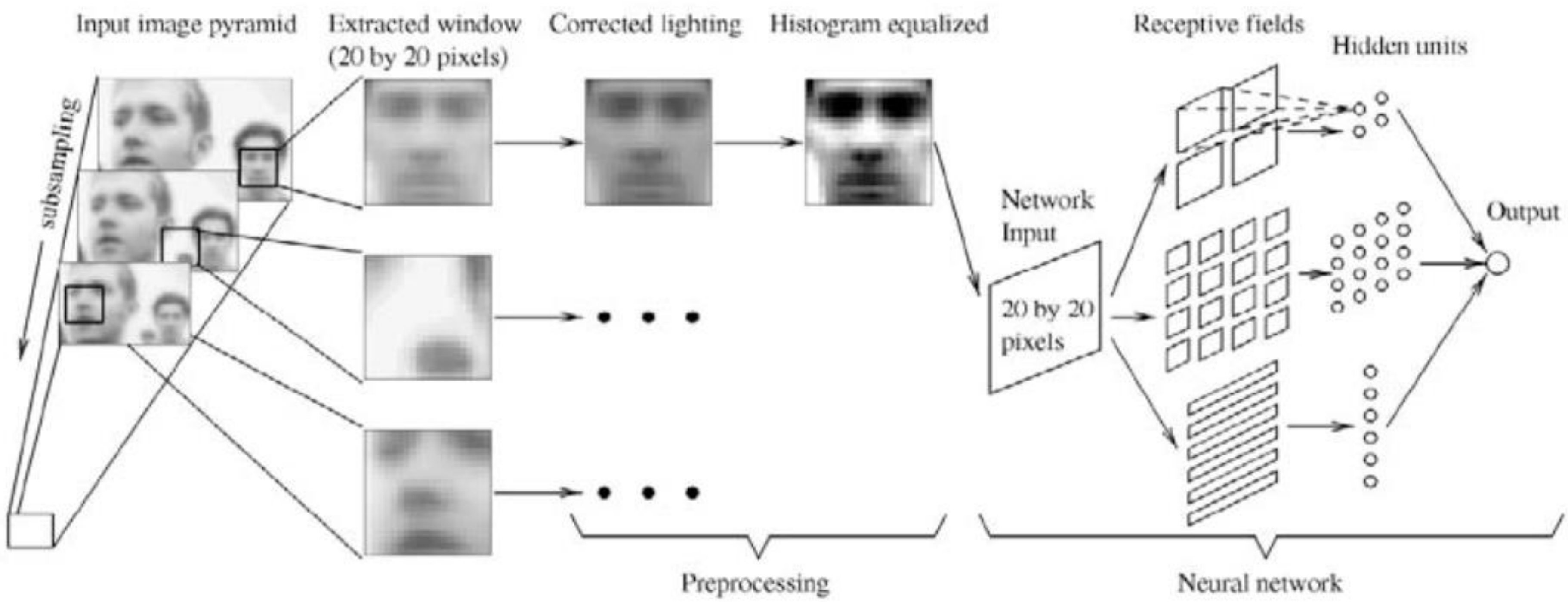


Face Detection - Histogram of oriented gradients

- Out put→ After finishing face detection, the output will be a frame with face location if the image contains human face

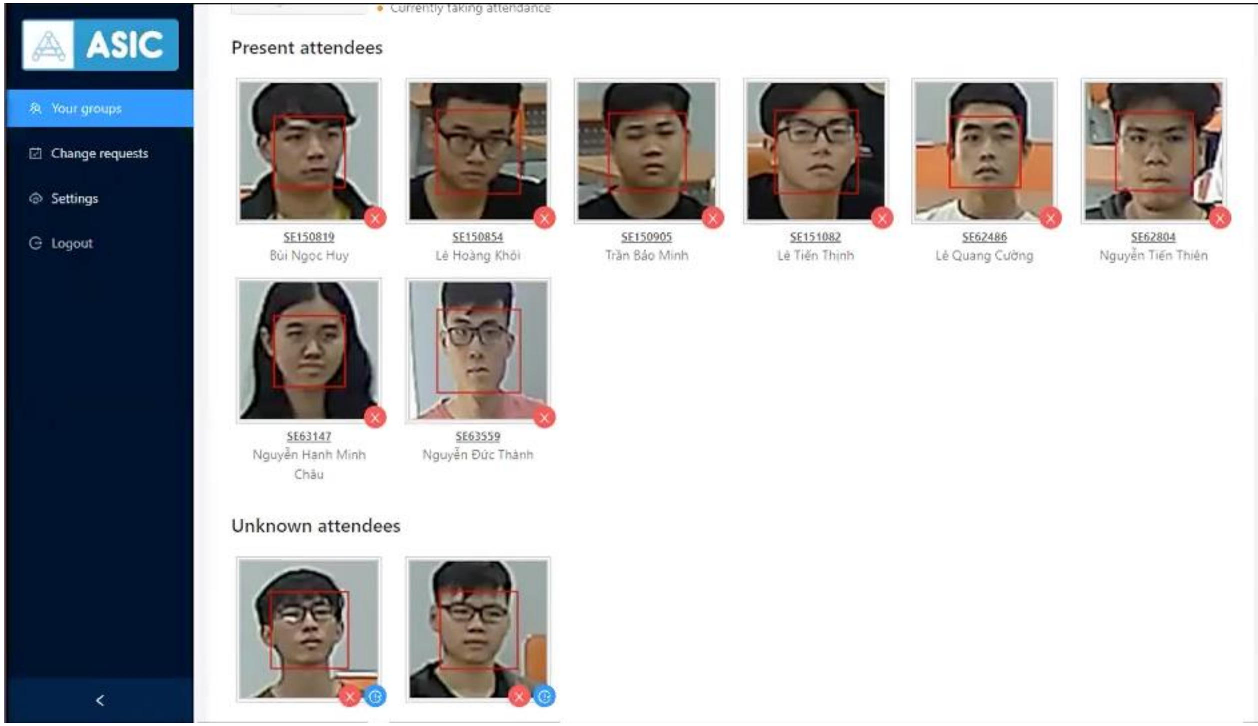
Face Detection - Neural network

- Face



Applications

- Facial motion capture
- Facial recognition
- Photography
- Marketing



Challenges

- Viewpoint variation
  - One of the biggest difficulties of object detection is that an object viewed from different angles may look completely different.
- Deformation
  - The subject of computer vision analysis is not only a solid object but also bodies that can be deformed and change their shapes, which provides additional complexity for object detection.
- Occlusion
  - Sometimes objects can be obscured by other things, which makes it difficult to read the signs and identify these objects.

Challenges

- Illumination conditions
  - Lighting has a very large influence on the definition of objects. The same objects will look different depending on the lighting conditions.
- Cluttered or textured background
  - Objects that need to be identified may blend into the background, making it difficult to identify them.
- Variety
  - The same object can have completely different shapes and sizes. Computer vision needs to do a lot of research to read an object and understand what it means.

Challenges

- Speed
  - When it comes to video, detectors need to be trained to perform analysis in an ever-changing environment. It means that object detection algorithms must not only accurately classify important objects but also be incredibly fast during prediction to be able to identify objects that are in motion.

Summary

- What is Object Detection?
- The techniques used in object detection.
- Learn about face detection
- Pedestrian detection
- Applications of face detection
- Object detection challenges