1605 The prototype model of SIH – Women Safety Analytics

Opencv – Frame Work(in-built RTSP)

Path to YOLO

v

**Websocket**

To send safety signals to the authorithy and also the person in danger

Path to Websocket

Path to CNN

**YOLO + CNN**

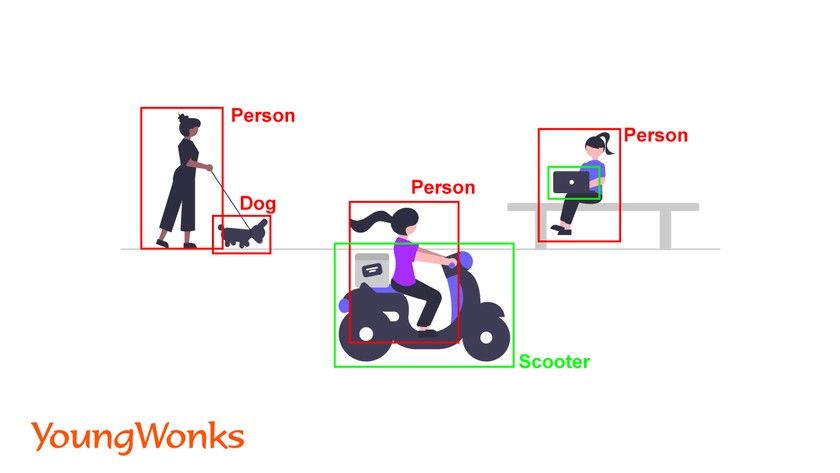
Facial detection Gender Distribution, Gender Classification

**CNN+STE**

Gesture recognition Expression recognition Voice recognition

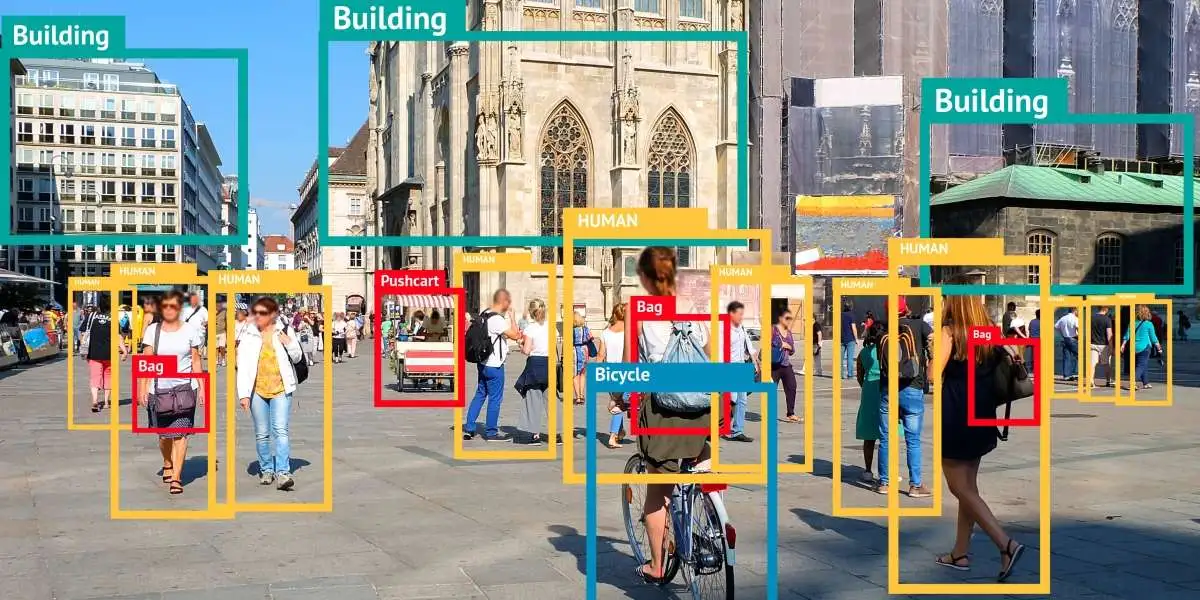
OpenCV:

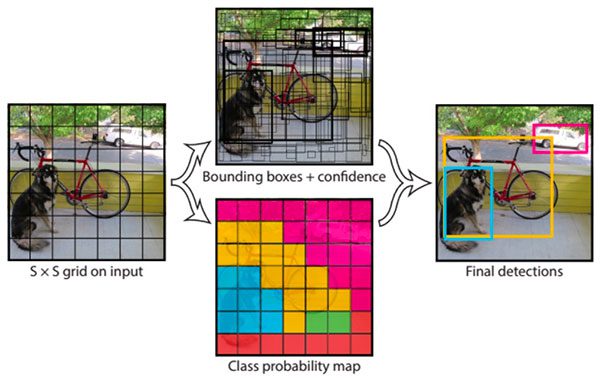
* For face detection and image preprocessing.
* To capture video frames from the CCTV camera ,can be done using RTSP (Real-Time Streaming Protocol) which is in-built in opencv



YOLO:

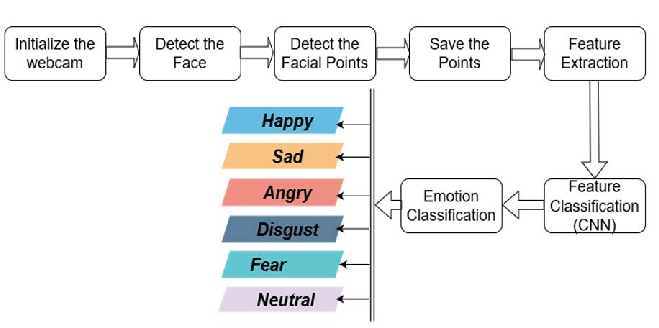
* Use YOLO to detect people in each frame. YOLO will return **bounding boxes**, confidence scores, and class labels for each detected object
* Extract regions of interest (ROIs) corresponding to the detected people.
* Yolo being integrated is able count the number of people present.
* Resize and pre-process these ROIs to feed them into the CNN model for gender classification.





CNN:

* Use the pre-trained CNN model to classify the gender of each detected person.
* Integrate results with the bounding boxes from YOLO for analysis.
* Being trained in both gesture and expression recognition aids in analysis the situation in better terms.



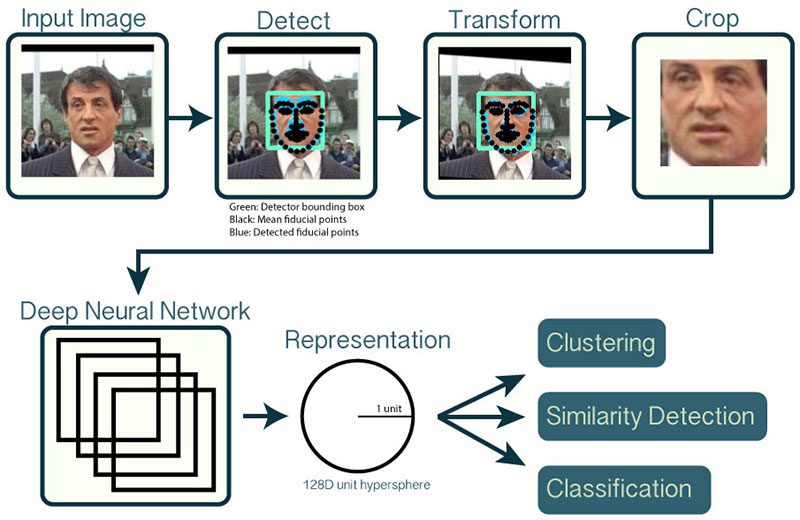
YOLO (You Only Look Once) + Gender Classification CNN:

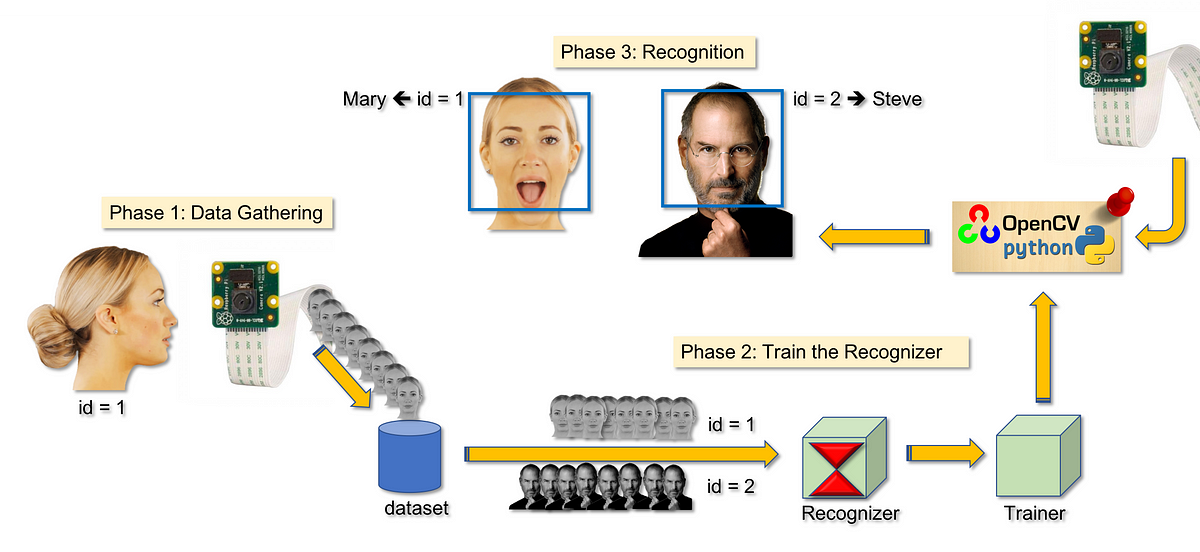
YOLO is a state-of-the-art object detection algorithm that detects people in real time with high speed and accuracy. After detecting people, a CNN (Convolutional Neural Network) can classify each detected person’s gender .

Both yolo and cnn can be integrated and used to predict both the class and location of objects in an image here location meaning place and class meaning the gender of the person.

Advantages: Fast, accurate, and suitable for real-time applications; works well with high-resolution CCTV footage.

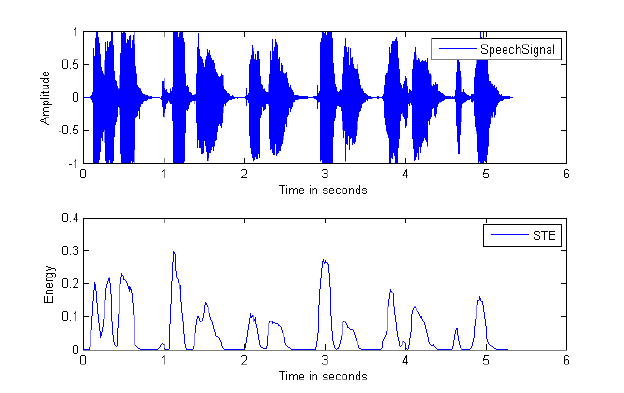
Implementation: Use YOLO for person detection, crop the detected regions, and pass them through a gender classification CNN.

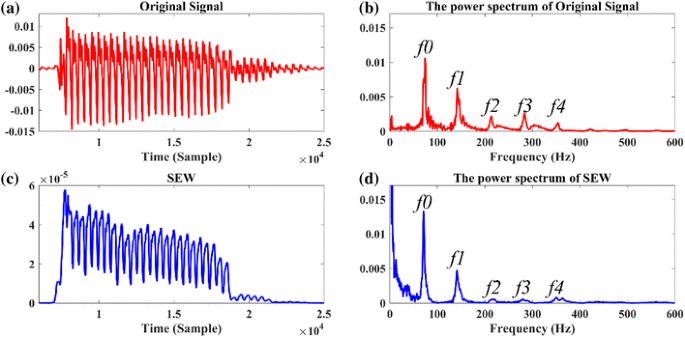




Short Time Energy (STE):

* Neural networks, particularly Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs), can be trained to recognize patterns in voice intensity across different contexts.
* **Short-Time Energy** is calculated by summing the squares of the amplitude values over short time frames (e.g., 20-30 milliseconds). It provides a measure of how intense or energetic the signal is within each frame.





Websocket:

* WebSockets provide a full-duplex communication channel over a single, long-lived TCP connection. This method allows instant communication between the detection system and authorities.
* The surveillance system maintains a persistent WebSocket connection with the control center. Upon detection of an event, an alert message is sent instantly over this channel.
* Advantages:
* Real-time, bidirectional communication.
* Low overhead compared to traditional HTTP requests.
* Ideal for continuously updating the status of a monitored area.
* 