Unmanned object detection

LincolnspartacusJames
Pavithran Ravichandiran

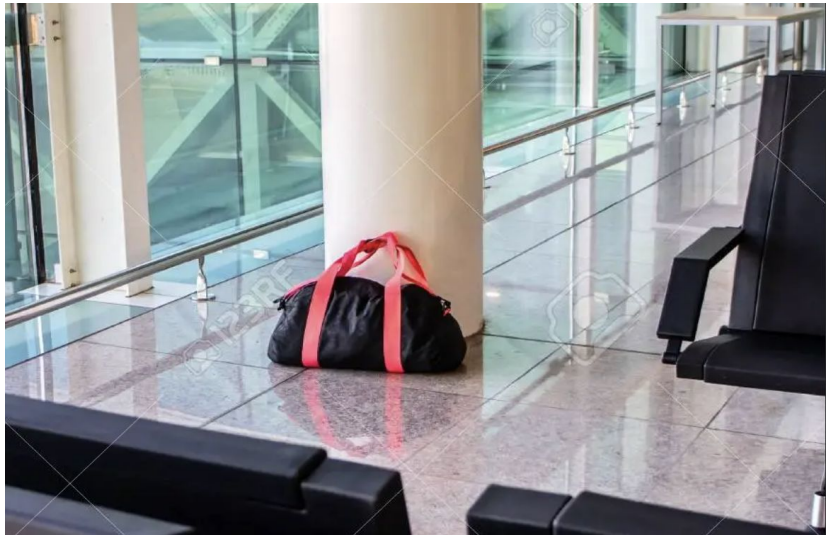
Problem statement

- Video surveillance system using computer vision
- Detect **suspicious objects left** by human to determine any **illegal** activities
- System can be used to aid existing video surveillance systems.

Motivation - Why is this important?

In public places like railway stations, Airports there may be scenarios where a person enters a scene with an object, **places the object that may be suspicious** and leaves the scene after placing the object.

Domain: Security, Video Surveillance.



Motivation - Why is this important?

Even there are incidents where a vehicle is **parked in a no parking zone** or in the middle of a busy interstate highway.

Domain: Road traffic & safety, Video Surveillance.

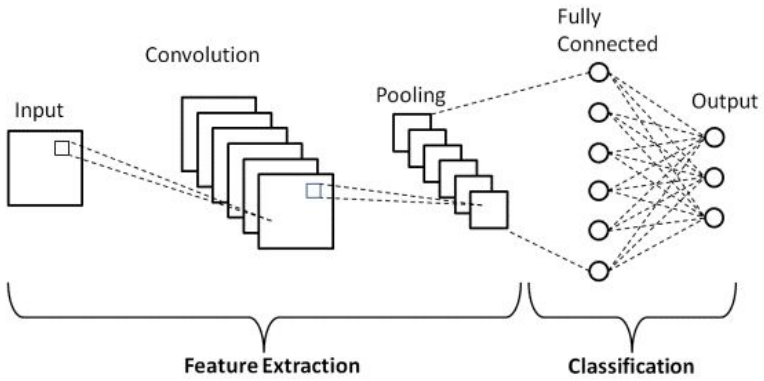
Related work

Morphological and thresholding techniques

- Identifying suspicious objects by morphological and different thresholding techniques.
- Example techniques:
  - Temporal average filter
  - Using frame differencing
  - Erosion/dilation
- Able to detect foreground and background objects
- Cons:
  - Suffers from performance issue and longer time to process as video quality diminishes
  - Faster movements may require higher thresholds. **Choosing the right threshold value is a hyperparameter problem**.

Deep Learning models

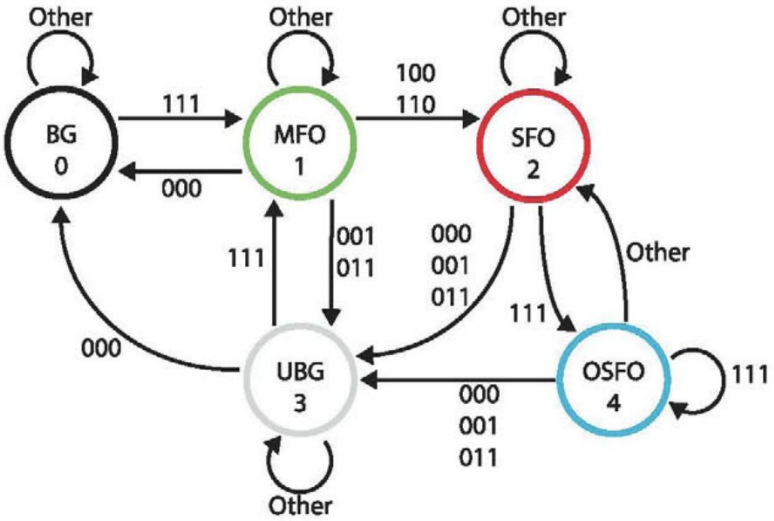
- Deep convolutional neural networks trained with different category of objects.
- Good detection speed and accuracy.
- **Cons:**
  - Might fail to distinguish and find the objects that belong to different categories outside the pre-trained objects
  - **Can't pretrain all scenarios** because suspicious objects in a video surveillance can be anything

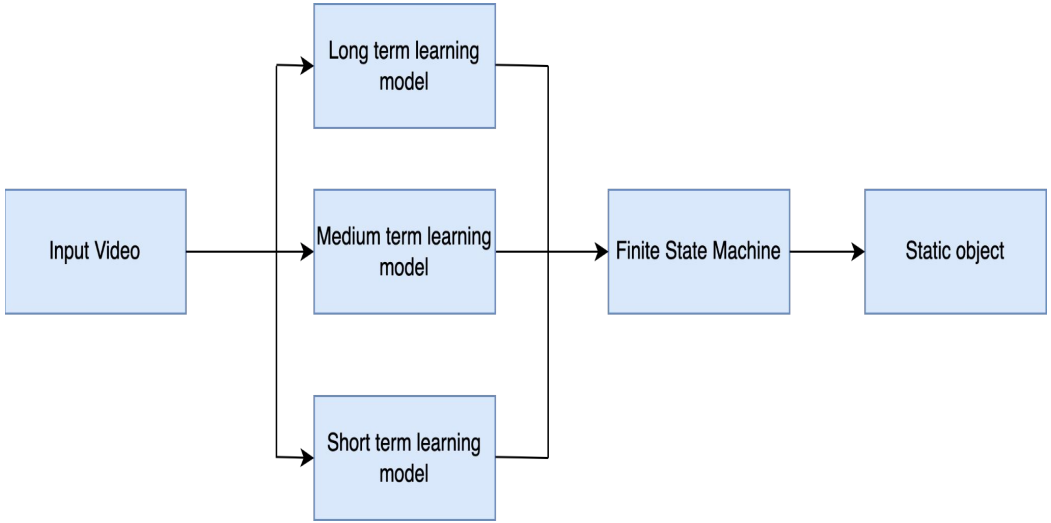


Finite State Machine approach

- Static foreground objects are detected using a state machine with three gaussian mixture models for background subtraction.
    - **Long, short and medium** term detectors. These models differ in learning rate.
    - Initially all pixels are assumed to be background pixels.
- Based on the state outputs from detectors, pixels can be classified as **foreground or background pixels.**
- Three detectors leads to **8 different states** consuming processing and time overhead

Finite State Machine approach





FSM approach - new proposed work

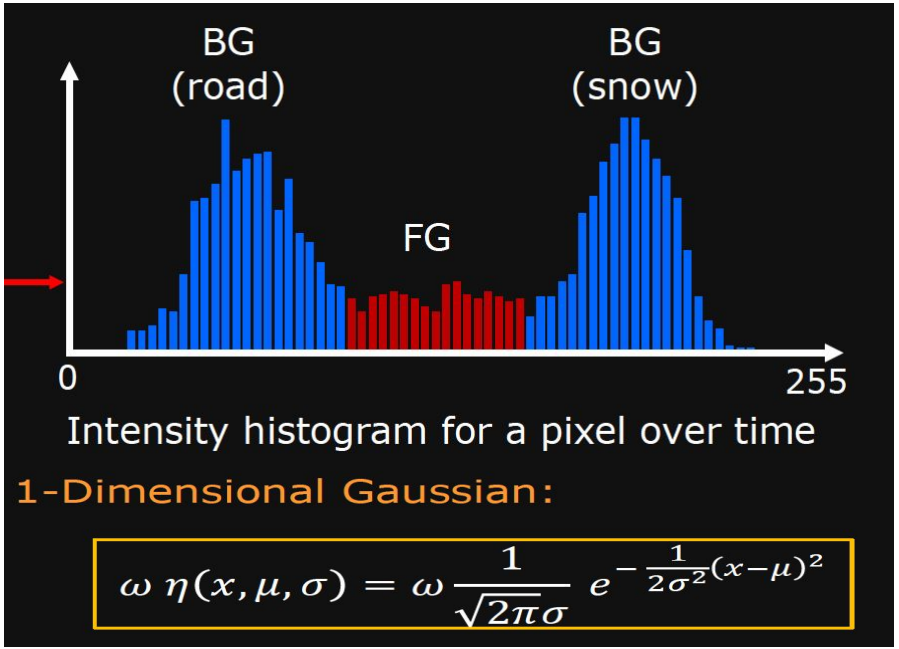
- **Objectives**:
  - Find the static foreground object that changes state from moving to static with less overhead of time
  - Track the owner of the candidate static object
- Proposed work eliminates the medium term detector and reduces the total detectors **only to long and short term detectors**.
  - Reduce the total number of state in the state machine from 8 to 4.
- Further, we've used an extended Gaussian mixture model to detect **moving** objects that come to a **static state**.
  - **Kalman filter** is combined with Gaussian Mixture Model to reduce false alarms and noises.

ROI selection

- Helps **remove unwanted space**
- **Reduces processing overhead**
- Input points from user
- **Example**:
    - Camera in parking area might cover nearby road

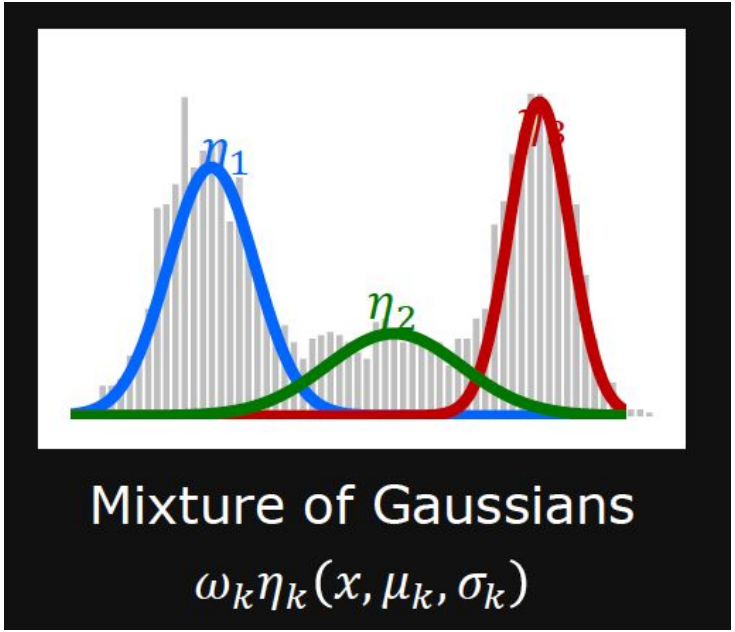
Detect foreground objects- Background subtraction

- Widely deployed method using gaussian mixture model
- Ability to detect various scenarios in a video
- Moving objects can be easily detected
- An extension using Kalman filter is added to find static objects

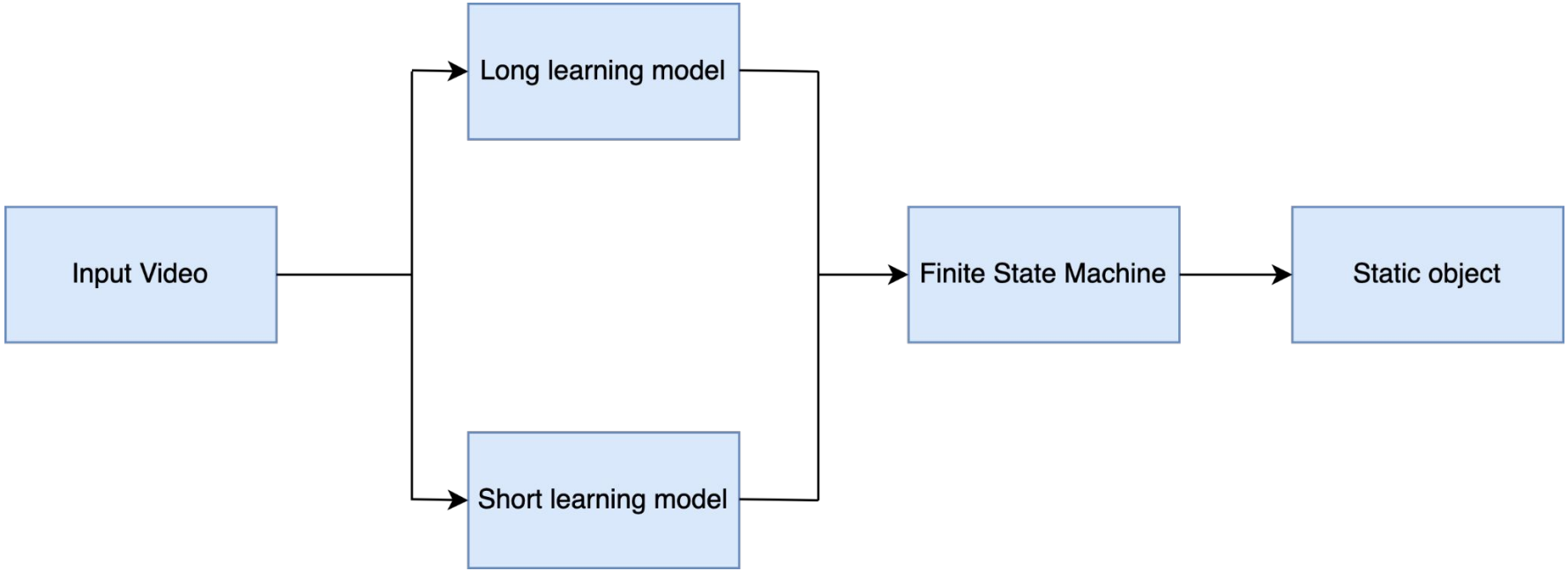


Background subtraction - Gaussian Mixture model

- Uses background model based on pixels learnt sequentially from the prior frames of input video
- Pixels are classified as background or foreground pixel
- Every pixel is made as a mixture of m Gaussian distributions



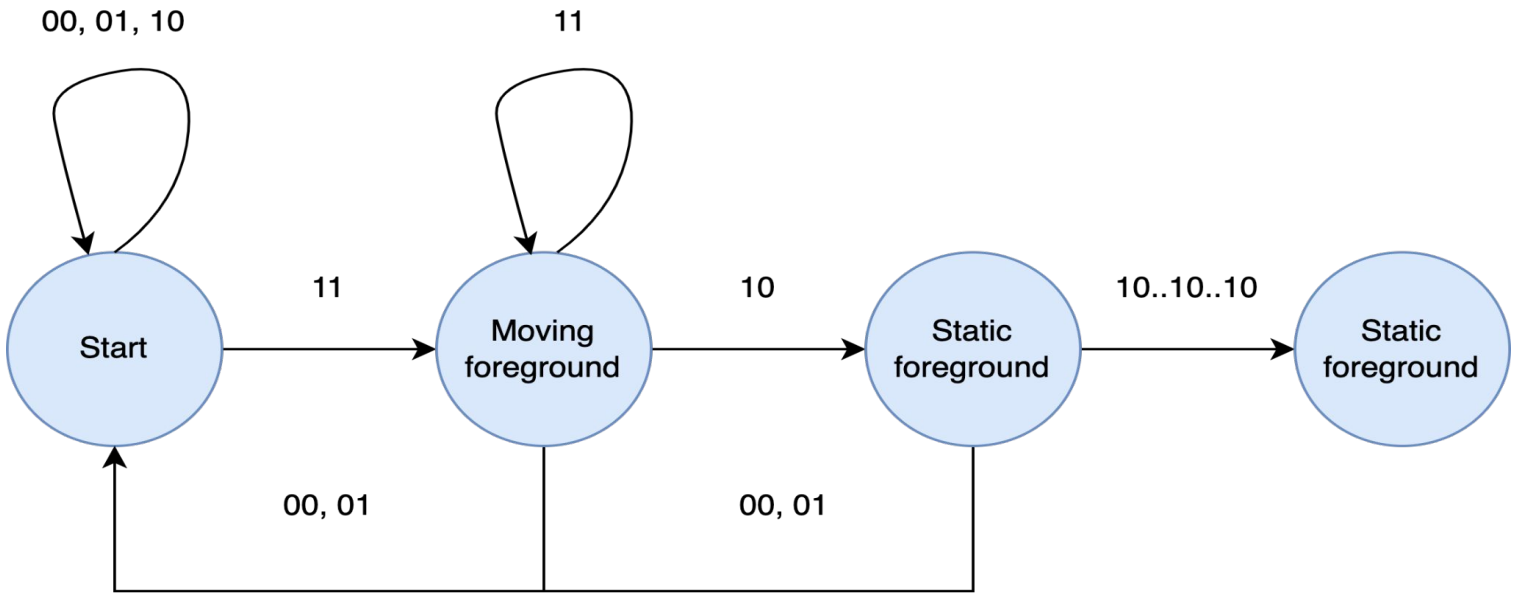
Architecture overview



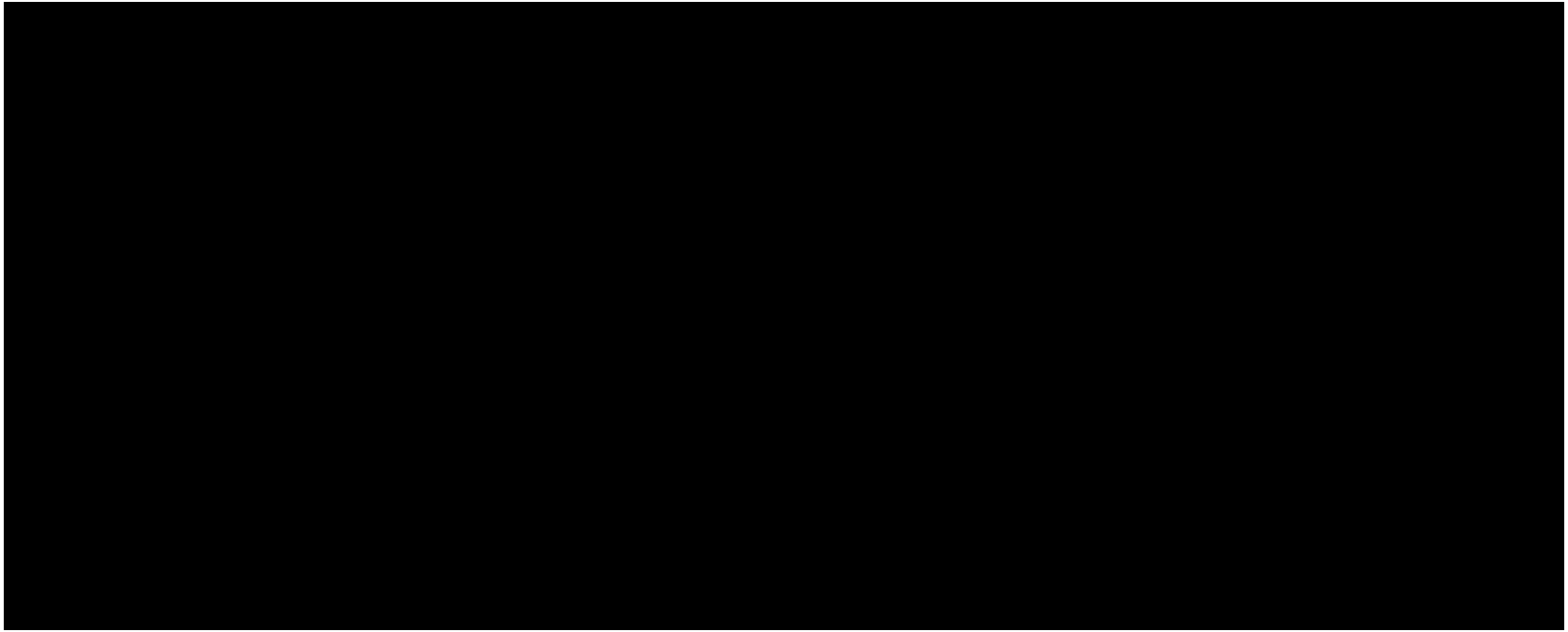
Pixel value and its types

Pixel values P (Long learning,Short Learning)	Pixel type
00	Background pixel
01	Occluded that is exposed in recent image
10	Likely to be static object
11	Moving object pixel

Finite State Machine



Demo



Accuracy

Video Scenarios	Ground Truth	GS	Three detector model	Proposed model
S1	T	F	T	T
S2	F	F	T	T
S3	F	T	T	T
S4	F	T	T	T
S5	F	F	T	T
S6	F	F	T	T
S7	F	F	T	T

Future scope - WIP

- **Owner tracking** - Differentiating between intentional and unintentional object placement.
  - Once the static object is detected, the focus is shifted to the previous frame (likely frame the object could have been kept on the ground).
  - A window near the detected the static object is created.
  - Any person near this window is tracked using the **Dalal-Triggs Human Detector** where a histogram of oriented gradients with a single filter can be used

Thank you. Any questions?

Learning rate

- Higher value of learning rate set for slower model and lower value for faster model
- <Equation for weight updates>