# Conference Submission

## Architecture Diagram



Figure 1: The proposed video processing framework architecture.

## System Evaluation

### Performance

### Limitations

## Use Case

The proposed framework provides core video processing functionality that enables extension and flexibility to meet a broad range of client use cases. Presented below is a conceptual use case showing the frameworks successful adoption to a specific domain, providing core insights and analytics into the video footage in real-time.

### Scenario

The Abbey Road in London, pictured on the front of the famous Beetles album “Abbey Road”, is a popular destination for tourists and locals attempting to recreate the album cover for themselves. Due to this, it is under live surveillance by a multitude of video cameras at all times. This presents an opportunity to gain real-time analytics of the activities occurring at the crossing. Deployed is the proposed framework configured to detect cars and people within the live video stream, extrapolate activities to track walking, standing and people attempting to recreate the album cover photo. It will analyse the activity log for anomalies, alerting the users upon discovery an anomalous event.



Figure 2: The famous Beetles cover “Abbey Road” (<https://www.google.co.uk/search?q=beatles+abbey+road&safe=off&rlz=1C5CHFA_enGB750GB750&source=lnms&tbm=isch&sa=X&ved=0ahUKEwi5rJ_ehO_ZAhUMCsAKHTLaA9IQ_AUICigB&biw=1280&bih=612#imgrc=sJtcKFJFhoc5cM:>)

In order to evaluate the proposed frameworks adoption success, we will focus on the performance and extensibility of the framework, showcasing its flexibility to support real-life client domains.

### Object Detection and Tracking

As the proposed architecture diagram shows (Figure 1), client-side processing is enabled in order to maintain real-time requirements and limit the stream of data from the device. This enables the tracking and detection of people and vehicles to occur at or near the device source (edge computing REFERENCE). When identifying cars and people within the video feed, the base framework allows the use of Haar Cascades (REFERENCE). This is a common, high-performant, detection methodology that provides an acceptable accuracy rate (Figure 3).

Between frames, tracking needs to occur in order to give people and vehicles a persistent identity, enabling the computation of further server-side analytics in identifying object movement patterns. Providing this is a Kernalized Correlation Filters (KCF) tracker (REFERENCE), which tracks an object by treating its location as a set of positions (‘bags’) that each could contain the objects location based on its previous location. It looks at overlapping positive regions identified to possibly contain the person or cars location to provide an accurate final decision on the objects location.



Figure 3: The deployed video processing framework identifying people and cars within the “Abbey Road” real-time video stream.

### Event Detection

### Anomaly Detection

### Performance

### Conclusion