Assignment 3.1P – A case study of AI Solution

SIT 788

Domain: Surveillance Systems.

Application: AI-enabled Traffic Surveillance.

Why is it important to have AI solutions in this domain? Provide the reason and justification.

- Increasing security risks and technological advancements have led to the development of smart cities. Reduced hardware costs have made hardware such as CCTV easily accessible.
- There is a growing demand for CCTV installations across public and private premises. One such example solution is AI-enabled Traffic Surveillance.
- The solutions are generally deployed across cities or even across states with several command-and-control centers set up to handle the data in real-time. This leads to a demand for an effective strategy to process the data. This data is in form of video recordings or real-time video streams from CCTV cameras.
- Optimization in knowledge discovery from a huge volume of data becomes essential as the response time for such violations is generally required to be as less as possible. Because of the large volume of data being collected, AI finds applications in various use cases for detecting traffic violations such as red-light violations, overspeeding violations, lane driving violations, etc.
- A major advantage of AI-enabled traffic surveillance is a reduction in human time and effort involved in processing videos on a scale of Giga/terabytes.

What is the difference between traditional applications and AI-based applications in this domain?

- Traditional methods of traffic surveillance include on-site monitoring on crossroads or patrol vehicles, these methods require tiresome manual work in extreme conditions.
- The introduction of CCTV cameras solves the problem of outdoor manual work to some extent, but traditional CCTV solutions rely on live human monitoring of video feed or storing large volumes of video recording on persistent storage such as a database or hard disk. This approach requires many human hours to monitor live video feeds or watch video recordings for possible traffic violations.
- Even traditional rule-based computer vision solutions designed to automate traffic surveillance are not generalized as they rely on hand-crafted rules. This makes them specific to certain use cases and cannot handle unseen behavior effectively.
- AI and ML techniques handle these pain points as AI/ML solutions are generalized and require minimum human intervention once trained effectively. This leads to a faster response time for traffic violations.

What are the challenges and Strengths of AI in that domain?

Strength:

- Traffic surveillance videos, whether live or recorded, can be processed in near real-time which reduces human effort drastically. AI-based computer vision solutions have boasted of very high accuracy which surpasses even human capabilities in recognition and detection scenarios, this makes the AI-enabled surveillance systems state of the art in this domain.
- Due to the generalisability of the architecture, a single ML/AI model can be trained for various business use cases such as license plate detection, pedestrian detection, etc. Multiple models can be deployed into a single system which can solve complex use cases such as red light or stop line violations, Speed violations, Lane driving violations, etc. This removes the need for having multiple systems for different scenarios.
- Flexibility to deploy these solutions on-premises as well as on-cloud makes AI-based surveillance systems highly scalable in deployment as they can handle a high number of CCTV camera sources based on the demand size. Moreover, AI solutions can reduce the space required to store these videos by intelligently removing redundant video frames.

Challenges:

- As the video contains information that can be linked to an individual, there is always a trade-off between security and privacy. This requires employing methods such as differential privacy to make sure an individual's privacy is not compromised.
- To achieve the state-of-the-art accuracy a large amount of data is required. Generally, raw video frames do not contain information that can be used to train an ML/Deep Learning model. Manual annotations are required to train a model, which can be a time-consuming task. Although once a model with considerable accuracy is trained, fine-tuning the model requires a relatively smaller volume of data.
- The stakeholders involved in deploying these solutions must bear the additional cost of training human operators.