

Over the last few years surveillance cameras have become ubiquitous. With so many cameras, analyzing the output manually has become very laborious and inefficient. In recent years, however, a lot of development has been focused on automatic video processing using artificial intelligence. There are many deep learning models for object detection offering basic low-level analysis. This thesis builds upon these models and creates an efficient video processing pipeline that serves as a base for further higher-level analyses. We aim to develop sufficiently fast video processing pipeline that will be able to process surveillance camera video streams in real-time while maintaining low CPU utilization. We move as much of the pipeline as possible to the GPU, with the data never leaving the GPU memory before the very end of the pipeline, and therefore leaving most of the CPU computational power for further data analysis. Our testing shows that our implementation achieves performance very close to real-time with 1080p video even on common consumer hardware.