



INTISARI

Seiring dengan meningkatnya jumlah kamera pengawas yang terpasang pada ruang publik pada 2017 diperkirakan 106 juta CCTV baru terpasang. Secara konvensional pengawasan dari video tersebut menggunakan tenaga manusia. Semakin banyaknya CCTV terpasang mendorong untuk dikembangkannya sistem pengawasan yang terotomasi.

Untuk menyelesaikan permasalahan tersebut maka dapat digunakan sistem berbasis *computer vision*. Semakin cepat dan akurat pengenalan terjadinya anomali tindak kriminal pada video maka semakin cepat juga respon dari penegak hukum terhadap anomali tersebut.

Pada tesis ini akan digunakan sistem *state-of-the-art* dalam *action recognition* untuk menyelesaikan masalah tersebut. Struktur dari penelitian ini terdiri dari 2 langkah. Langkah pertama adalah ekstraksi ciri dengan menggunakan I3D. Arsitektur ini dipilih karena didesain untuk menangkap ciri spatio dan temporal dari video. Kemudian langkah kedua adalah pelatihan pengklasifikasi dengan ciri yang didapat dari dataset *weakly supervised*. Dengan menggunakan dataset UCF-Crime pendekatan ini mampu mencapai hasil *state-of-the-art* sebesar 78% AUC dan waktu inferensi 99.1 fps.

Kata Kunci : Pengenalan Anomali Video, *UCF-Crime*, *deep learning*, *I3D*, *CCTV*.



ABSTRACT

Surveillance cameras number keep increasing from year to year, in 2017 alone estimated 106 million new CCTV installed. Surveillance cameras are installed to widen surveillance coverage in order to detect anomalies. Conventionally, detecting anomalies through surveillance cameras are using human observers. The increasing number of surveillance cameras installed is increasing the need for an automatic system to replace unreliable human observer.

Thus, an intelligent computer vision-based system to detect anomalies need to be developed. The goal of this system is to provide a warning to the first responder accurately and as fast as possible while the system is running all the time. The faster the response of the first responder, the better chance of an anomaly to be resolved.

In this thesis, we proposed usage of state-of-the-art deep-learning based action recognition system to resolve said problem. The structure of this research consisted of two main steps, the first is feature extraction using I3D. This architecture is designed to capture both spatio and temporal information from a video. The second step is training a classifier using extracted features from a weakly supervised dataset. Using UCF-Crime dataset, this approach could achieve state-of-the-art performance with 78% of AUC with inference time of 99.1 fps.

Keywords: action recognition, anomaly recognition, UCF-Crime, deep learning, I3D