

Deteksi Aktivitas Tidak Wajar pada CCTV ATM Menggunakan 3D Convolutional Neural Network = Detection of Suspicious Activities on CCTV at ATM Using 3D Convolutional Neural Network

Aldi Hilman Ramadhani, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20506499&lokasi=lokal>

Abstrak

Penelitian ini memiliki tujuan untuk mencari model machine learning yang dapat mengenali kegiatan yang dilakukan pengguna ATM, serta mencari algoritma terbaik untuk mengetahui kapan suatu kegiatan pengguna ATM dimulai dan selesai pada suatu video. Terdapat sembilan jenis aktivitas berbeda yang ingin dideteksi. Penelitian ini dapat dibagi dalam dua fase, yaitu fase mencari rentang waktu aktivitas pada video yang disebut fase deteksi aktivitas, dan fase mengenali aktivitas tersebut yang disebut fase pengenalan aktivitas. Pada fase pengenalan aktivitas, penulis mengajukan suatu rancangan arsitektur 3D CNN, serta melakukan eksperimen terhadap parameter pada arsitektur tersebut. Setelah melakukan beberapa eksperimen, didapatkan model terbaik dengan kernel berukuran $3 \times 3 \times 3$, menggunakan input video dengan piksel berukuran 20×20 per frame, dan menggunakan dua lapis layer ekstraksi fitur. Pada fase deteksi aktivitas, penulis mengajukan suatu rancangan fungsi deteksi aktivitas, yang mengikuti framework 'classification lalu post-processing' yang merupakan salah satu framework untuk deteksi aktivitas (Yao et al., 2018), serta melakukan eksperimen terhadap parameter pada fungsi tersebut. Setelah melakukan beberapa eksperimen, didapatkan performa terbaik dengan parameter θ sebesar 20, dan konstanta C sebesar 365. Pada kedua eksperimen tersebut, terdapat beberapa kesalahan yang dilakukan, sehingga diperlukan eksperimen lanjutan dimana kesalahan tersebut tidak dilakukan. Kesalahan tersebut adalah model kemungkinan besar masih underfit, dan terdapat permasalahan pada pemotongan video manual pada dataset. Setelah menyelesaikan kesalahan tersebut, model untuk fase pengenalan aktivitas mendapatkan akurasi sebesar 93.94%, presisi sebesar 96.36%, recall sebesar 93.94%, dan f-score sebesar 93.69%. Pada sisi lain, dalam fase deteksi aktivitas didapatkan akurasi sebesar 94.44%, presisi sebesar 96.30%, recall sebesar 96.30%, dan f-score sebesar 94.07%.

This research aims to find a machine learning model that can recognize the activities of ATM users, and find the best algorithm to find when each ATM user activity starts and finishes on a video. There are nine different types of activities that this study want to detect. This research can be divided into two phases, namely the phase of detecting for a time span of activity on a video that is called the activity detection phase, and the phase of recognizing that activity that is called the activity recognition phase. In the activity recognition phase, I propose a 3D CNN architecture design, and conduct experiments on the parameters of the architecture. After carrying out several experiments, the best model is obtained with a kernel with dimensions of $3 \times 3 \times 3$, using video input with pixels measuring 20×20 per frame, and using two layers of feature extraction layer. In the activity detection phase, I propose an activity detection function, which follows the 'classification then post-processing' framework, which is one of the frameworks for activity detection (Yao et al., 2018), and conducts experiments on the parameters of the function. After carrying out several experiments, the best performance was obtained with a θ parameter of 20, and a constant C of 365. In both experiments, there were some errors made, so that further experiments were needed to be done, where the errors were not carried out. The error is that the model is most likely still in underfit phase, and

there are problems with manual video clipping on the dataset. After resolving these errors, the model for the activity recognition phase gained an accuracy of 93.94%, a precision of 96.36%, a recall of 93.94%, and an f-score of 93.69%. On the other hand, in the activity detection phase an accuracy of 94.44% is obtained, a precision of 96.30%, a recall of 94.44%, and an f-score of 94.07%.