

INTISARI

Saat ini teknologi *deep learning* telah banyak digunakan dalam berbagai bidang mulai dari pengolahan citra, pengenalan suara, klasifikasi bahkan deteksi suatu objek. Telah banyak jaringan deep learning yang dikembangkan. Untuk deteksi objek terdapat YOLO (*You Only Look Once*), SSD, dan Fast-RCNN sedangkan untuk data yang sifatnya sekuensial waktu terdapat LSTM, RNN dan GRU.

Pada penelitian ini akan dibahas mengenai perancangan sebuah sistem agar dapat melakukan prediksi arus lalu lintas secara real-time menggunakan kamera CCTV jalan raya. Pendekatan yang dilakukan berbasis deep learning dengan menggunakan beberapa arsitektur yang telah ada. Secara garis besar, proses perancangan sistem akan dibagi menjadi 3 bagian yaitu perancangan objek detektor menggunakan arsitektur YOLOV3 untuk diimplementasikan di CCTV, perancangan LSTM untuk prediksi jumlah kendaraan serta perancangan *dashboard* untuk visualisasi sistem. Objek detektor akan bertugas untuk melakukan deteksi jenis kendaraan, *tracking* dan perhitungan jumlah kendaraan. LSTM bertugas untuk memprediksi arus lalu lintas berdasarkan data jumlah kendaraan yang telah dibuat oleh objek detektor. Untuk memberikan hasil yang optimal, kedua jaringan akan dilatih dan dievaluasi menggunakan metriknya masing-masing. Objek detektor akan dilatih menggunakan data sebanyak 1688 gambar dan diuji menggunakan 168 gambar pada *framework* Darknet, serta dievaluasi menggunakan nilai *precisions*, *recall*, *mAP* dan *F1-Score*. Sedangkan LSTM akan dilatih menggunakan 1170 data (9 hari) dengan memvariasikan parameter jumlah lapisan LSTM, jumlah neuron tiap lapisan, jumlah data input, serta jumlah data output LSTM dan akan dievaluasi menggunakan 260 data uji (2 hari) untuk mengetahui nilai *Root Mean Squared Error* (RMSE) dan *Mean Average Percentage Error* (MAPE). Data dari objek detektor dan LSTM akan ditampilkan jadi satu pada *dashboard* sehingga informasi dari kedua model dapat dipahami lebih mudah.

Berdasarkan hasil pelatihan dan evaluasi, model YOLOV3 yang digunakan memiliki nilai *precision* = 87.59%, *recall* = 88%, dan *f1-score* = 87.79%. Untuk LSTM, diperoleh nilai RMSE terbaik sebesar 25.933 dan MAPE = 10.362%

Kata kunci : *Deep Learning, You Only Look Once, Long-Short Term Memory, tracking, precision, recall, F1-score, Root Mean Squared Error, Mean Average Percentage Error*

ABSTRACT

Currently, deep learning technology has been widely used in various fields ranging from image processing, voice recognition, classification and even object detection. Many deep learning networks have been developed. For object detection there are YOLO (You Only Look Once), SSD, and Fast-RCNN while for time-series data there are LSTM, RNN and GRU.

This research will discuss the process of designing a system that can predict traffic flow using CCTV cameras in real time. The approach taken is based on deep learning by using several existing architectures. The system design process will be divided into 3 parts they are: designing an object detector using YOLOV3 architecture to be implemented on CCTV, designing a LSTM model for predict the number of vehicle, and designing a dashboard for system visualization. The object detector used to detect and classify type of vehicle, tracking and counting total number of vehicle every 5 minutes. LSTM is used to predict traffic flow based on the number of vehicles data that have been created by the object detector. To provide optimal results, both networks will be trained and evaluated using their respective metrics. The object detector will be trained using 1688 images and tested using 168 images in Darknet framework, and evaluated using the precisions, recall, mAP and F1-Score metric. Whereas LSTM will be trained using 1170 data (9 days) by varying the training's parameters like number of LSTM layers, number of neurons per layer, amount of input data, and the amount of LSTM output data. After training process, LSTM will be evaluated using 260 test data (2 days) to find out the values of Root Mean Squared Error (RMSE) and Mean Average Percentage Error (MAPE). Data from the object detector and LSTM will be displayed together on the dashboard so that information from both models can be understood more easily.

Based on the results of training and evaluation, the YOLOV3 model used has a value of precision = 87.59 %, recall = 88 %, and f1-score = 87.79 %. For LSTM, the best RMSE value is 25,933 and MAPE = 10,362 %.

Keywords : Deep Learning, You Only Look Once, Long-Short Term Memory, tracking, precision, recall , F1-score, Root Mean Squared Error, Mean Average Percentage Error