ABSTRAK

Deteksi aktivitas manusia adalah sebuah sistem yang dapat secara otomatis mengidentifikasi aktivitas dan telah banyak dikembangkan di berbagai bidang seperti kesehatan, keamanan dan industri. Terdapat dua pendekatan utama dalam memperoleh data aktivitas manusia, yaitu *vision-based* dan *sensor-based* yang masing-masing memiliki kelebihan dan kekurangan. Salah satu pendekatan yang dapat digunakan untuk mengatasi keterbatasan atau kelemahan tersebut adalah *skeleton-based* HAR *(Human Activity Recognition)* yang memanfaatkan koordinat titik sendi tubuh sebagai representasi aktivitas manusia. Namun, implementasi deteksi aktivitas manusia secara *real-time* masih menghadapi tantangan akibat kebutuhan komputasi yang tinggi. Oleh karena itu, penelitian ini bertujuan untuk mengimplementasikan model yang dapat melakukan deteksi aktivitas manusia berbasis data *skeleton* menggunakan EfficientNet-Lite guna menghasilkan model yang akurat dan efisien untuk skenario *real-time*.

Penelitian dilakukan menggunakan *dataset* NTU RGB+D dengan empat kelas aktivitas, yaitu berjalan, berdiri, duduk, dan terjatuh. Data *skeleton* akan direpresentasikan dalam bentuk *heatmap* dan disusun dengan melakukan *stacking* pada *frame* agar data mudah diproses oleh CNN sekaligus mempertahankan informasi temporal yang ada. Model EfficientNet-Lite juga dimodifikasi pada bagian *input layer*, MBConv *block*, dan *classification layer* untuk menyesuaikan karakteristik atau bentuk dari data serta mengurangi beban komputasi, dan meminimalkan risiko *overfitting*. Evaluasi dilakukan menggunakan akurasi, presisi, *recall,* jumlah parameter, FLOPs (*Floating Point Operations*), dan *inference time* serta FPS (*Frame Per Second*).

Hasil pengujian menunjukkan bahwa model dapat memperoleh akurasi sebesar 97% dengan nilai FLOPs sebesar 21.26 MFLOPs dan kecepatan inferensi sebesar 4.02 ms. Selain itu, pada skenario *real-time* menggunakan *pose estimator* MediaPipe diperoleh hasil rata-rata proses deteksi sebesar 26.8 FPS. Hasil tersebut menunjukkan bahwa EfficientNet-Lite dengan representasi *skeleton heatmap* memiliki potensi untuk menghasilkan model HAR yang efisien dan sesuai untuk implementasi *real-time*.

Kata Kunci: *Deep Learning*, Klasifikasi, Deteksi Aktivitas Manusia, EfficientNet-Lite, *Real-Time*

ABSTRACT

Human Activity Recognition (HAR) is a system capable of automatically identifying human activities and has been widely developed across various domains, including healthcare, security, and industry. There are two primary approaches to acquiring human activity data, vision-based and sensor-based methods, each with it's own advantages and disadvantages. Skeleton-based HAR offers an alternative representation by utilizing body joint coordinates to represent human activities. However, the implementation of real-time HAR still faces challenges related to high computational demands. Therefore, this research aims to implement a HAR model based on skeleton data using EfficientNet-Lite to produce an accurate and efficient model suitable for real-time scenarios.

This research utilizes the NTU RGB+D dataset with four activity classes: walking, standing, sitting, and falling. Skeleton data is represented as heatmaps and organized through frame stacking, enabling the CNN to process the data efficiently while preserving temporal information from the original input. The EfficientNet-Lite model is modified at the input layer, MBConv block configuration, and classification layer to accommodate the input data characteristics or shape, reduce computational load, and minimize the risk of overfitting. Model evaluation is conducted using accuracy, precision, recall, parameter count, FLOPs (Floating Point Operations), inference time, and FPS (Frames Per Second).

The experiment results show that the model achieves an accuracy of 97%, with FLOPs of 21.26 MFLOPs and an inference time of 4.02 ms per frame. Furthermore, in a real-time scenario using the MediaPipe pose estimator, the system achieves an average processing time of 26.8 FPS. These results demonstrate that EfficientNet-Lite with skeleton heatmap representation produces an efficient HAR model and has potential for real-time implementation.

Keywords: *Deep Learning, Classification, Human Activity Recognition, EfficientNet-Lite, Real-Time*