

## ABSTRACT

Human Pose Estimation (HPE) is a field in computer vision where models attempt to estimate the pose from an image or video. Research on HPE has been extensively conducted using Convolutional Neural Networks (CNN). In this study, a system model for estimating human activity based on vision is developed using Mediapipe, and the activities are classified using a random forest classifier. The results of this study are then compared with sensor-based activity estimation. The findings reveal that the random forest model can achieve an accuracy of 47.2% in recognizing fast walking activity. However, there are still many misclassifications in fast walking and running activities. In contrast, sensor-based human pose estimation can achieve an accuracy of 69% in recognizing running activities. The imbalance in the training data, with the largest distribution in the 'walking' class at 40%, leads to poor classification performance by the model. Other factors, such as variations in individual styles and data collection techniques, also affect the classification accuracy.

**Keywords:** internet of things, pose estimation, mediapipe, random forest