

ABSTRACT

Traffic accidents in Indonesia have increased by 34.37% over the past four years, with drowsy driving identified as a major contributing factor. To address this issue, this thesis presents a real-time drowsiness detection system designed for resource constraint device, utilizing lightweight deep learning models deployed on the ESP32-S3 microcontroller. The system targets the challenge of real-time detection in resource-constrained environments by employing MobileNetV1 and MobileNetV2 architectures, optimized through post-training quantization to produce low-complexity models suitable for low-power devices. The detection workflow includes image capture, preprocessing, and classification of the driver's eye state (drowsy or awake). On the microcontroller, the model is run with the TensorFlow Lite For Micro (LiteRT) library. In baseline evaluation, MobileNetV1 achieved an accuracy of 88% with an average inference time of 81.5 milliseconds per frame, requiring only 89,488 bytes of memory. MobileNetV2 reached 94% accuracy in baseline evaluation, and when deployed on the ESP32-S3, demonstrated an inference time of 329 milliseconds per image, utilizing 238,258 bytes of PSRAM and 31,608 bytes of internal RAM for image capture and classification. By balancing performance and computational efficiency, this research contributes to the development of embedded systems aimed at enhancing driver safety.

Keywords: Drowsiness detection, ESP32-S3, MobileNetV2, Edge Computing, Real-time Detection.