

ABSTRACT

The challenges faced by visually impaired individuals in performing daily activities, particularly indoors, are a serious concern. Their limited ability to effectively recognize their surroundings can be dangerous and pose significant threats. Developing assistive technology for the visually impaired is a crucial solution to this problem, particularly in enhancing their independence and mobility. This study proposes an object detection system based on speech recognition using the YOLOv8 (You Only Look Once version 8) algorithm. The system is designed to assist visually impaired individuals in recognizing and identifying objects around them in real-time through voice instructions. The system is developed utilizing YOLOv8, optimized with the Adam optimizer and Particle Swarm Optimization for parameter tuning. Additionally, the system integrates speech recognition technology, voice synthesis generation, and an object distance estimation model capable of receiving and processing voice commands from users. This implementation includes training the YOLOv8 model on relevant datasets to ensure optimal object detection capabilities under various environmental conditions. Experiments were conducted using two datasets: the MS COCO dataset for comparison with existing systems and a custom dataset for real-time testing. Experimental results on the system trained on the MS COCO dataset yielded mAP_{50} and mAP_{50-95} values of 66.1% and 44.7%, respectively. Furthermore, real-time testing using the custom dataset showed evaluation results of mAP_{50} at 74.7%, mAP_{50-95} at 48.6%, inference speed at 14.4 ms, a frame rate in the range of 24-30 frames/s, and object distance estimation accuracy at 70.1%. Further testing in real-world scenarios also demonstrated the significant potential of this system to be adopted as a mobility aid for visually impaired individuals, providing them with greater independence and safety in their daily activities.

Keywords: *Adam Optimizer, Object Detection, Particle Swarm Optimization, Speech Recognition, Text-to-Speech, YOLOv8.*

ABSTRAK

Kesulitan penyandang tunanetra dalam melakukan aktivitas sehari-hari khususnya di dalam ruangan menjadi permasalahan serius. Keterbatasan dalam mengenali lingkungan sekitar dengan baik dapat membahayakan dan menciptakan ancaman bagi mereka. Pengembangan teknologi asistif untuk penyandang tunanetra merupakan salah satu solusi untuk permasalahan tersebut, khususnya dalam konteks meningkatkan kemandirian dan mobilitas mereka. Penelitian ini mengusulkan sebuah sistem deteksi objek berbasis *speech recognition* menggunakan algoritma YOLOv8 (*You Only Look Once version 8*). Sistem ini dirancang untuk membantu penyandang tunanetra dalam mengenali dan mengidentifikasi objek di sekitar mereka secara *real-time* melalui instruksi suara. Sistem dikembangkan dengan memanfaatkan YOLOv8 yang dioptimasi oleh *Adam optimizer* dan *Particle Swarm Optimization* untuk optimasi parameter. Selain itu, sistem dikombinasikan dengan teknologi pengenalan suara, pembangkitan sintesis suara, dan model perkiraan jarak objek yang mampu menerima dan memproses perintah suara dari pengguna. Implementasi ini mencakup pelatihan model YOLOv8 pada *dataset* yang relevan untuk memastikan kemampuan deteksi objek yang optimal dalam berbagai kondisi lingkungan. Eksperimen dilakukan dengan menggunakan dua *dataset*, yaitu MS COCO *dataset* untuk perbandingan dengan sistem yang ada saat ini dan *custom dataset* untuk pengujian secara *real-time*. Hasil eksperimen pada sistem yang dilatih pada MS COCO *dataset* menghasilkan nilai mAP₅₀ dan mAP₅₀₋₉₅ masing-masing sebesar 66.1% dan 44.7%. Selanjutnya hasil eksperimen pada situasi *real-time* yang menggunakan *custom dataset* menunjukkan hasil evaluasi mAP₅₀ sebesar 74.7%, mAP₅₀₋₉₅ sebesar 48.6%, kecepatan inferensi sebesar 14.4 ms, *frame rate* berada pada rentang 24-30 frame/s, dan akurasi perkiraan jarak objek sebesar 70.1% . Pengujian lebih lanjut dalam situasi nyata juga menunjukkan potensi besar sistem ini untuk diadopsi sebagai alat bantu mobilitas bagi penyandang tunanetra, memberikan mereka lebih banyak kemandirian dan keamanan dalam aktivitas sehari-hari.

Kata Kunci: *Adam Optimizer*, Deteksi Objek, *Particle Swarm Optimization*, *Speech Recognition*, *Text-to-Speech*, YOLOv8.