BAB I
PENDAHULUAN

1.1 Latar Belakang

Pengelolaan ruang kelas yang efisien merupakan salah satu aspek penting dalam pengelolaan fasilitas pendidikan modern. Seiring dengan meningkatnya kebutuhan terhadap sistem pembelajaran yang adaptif dan berbasis teknologi, institusi pendidikan dituntut untuk mampu mengelola penggunaan ruang kelas secara optimal. Informasi mengenai tingkat okupansi ruang kelas memiliki peran penting dalam penyusunan jadwal perkuliahan, optimalisasi pemanfaatan ruang, serta peningkatan efisiensi penggunaan energi, khususnya pada sistem pencahayaan dan pendingin ruangan. Dalam konsep smart building dan smart classroom, data okupansi yang akurat menjadi komponen penting dalam mendukung pengambilan keputusan berbasis data sehingga penggunaan sumber daya dapat dilakukan secara lebih efektif dan efisien(Monti et al., 2022).

Sistem pemantauan okupansi ruang kelas tidak hanya berkaitan dengan efisiensi penggunaan ruang, tetapi juga berhubungan dengan kenyamanan pengguna dan efektivitas proses pembelajaran. Ketidaksesuaian antara kondisi okupansi aktual dan sistem pengelolaan yang digunakan dapat menyebabkan pemborosan energi, penggunaan ruang yang tidak optimal(Mohottige et al., 2022), serta menurunnya kualitas layanan fasilitas pendidikan. Oleh karena itu, diperlukan sistem yang mampu menyediakan informasi mengenai jumlah penghuni ruang kelas secara akurat, otomatis, dan berkelanjutan.

Dalam praktiknya, pemantauan penggunaan ruang kelas masih banyak dilakukan secara manual atau mengandalkan asumsi berdasarkan jadwal penggunaan. Pendekatan tersebut memiliki keterbatasan karena tidak dapat merepresentasikan kondisi aktual di lapangan secara *real-time*. Selain membutuhkan keterlibatan pengguna atau pencatatan administratif, metode manual juga rentan terhadap kesalahan pencatatan dan tidak mampu memberikan informasi mengenai jumlah penghuni secara aktual. Beberapa penelitian telah memanfaatkan sensor sederhana, seperti *sensor Passive Infrared* (PIR), untuk mendeteksi keberadaan penghuni di dalam ruangan(Lin et al., 2021). Namun, sensor tersebut

hanya mampu mendeteksi adanya gerakan tanpa memberikan informasi mengenai jumlah penghuni, tingkat kepadatan, maupun distribusi penghuni di dalam ruang. Keterbatasan tersebut menyebabkan pendekatan berbasis sensor sederhana kurang sesuai untuk sistem *occupancy monitoring* yang memerlukan estimasi jumlah penghuni secara akurat.

Perkembangan teknologi computer vision dan deep learning memberikan peluang yang signifikan dalam mengatasi keterbatasan metode konvensional. Dengan memanfaatkan kamera sebagai perangkat akuisisi citra, sistem dapat mendeteksi keberadaan manusia secara otomatis berdasarkan citra digital. Pendekatan berbasis visi memungkinkan sistem tidak hanya mendeteksi keberadaan seseorang di dalam ruangan, tetapi juga mengestimasi jumlah penghuni melalui teknik object detection. Berbagai model object detection berbasis *convolutional neural network* (CNN), seperti SSD, YOLOv8n, dan Faster R-CNN, telah menunjukkan performa yang baik dalam mendeteksi manusia pada lingkungan *indoor*(Olorunshola et al., 2023). Meskipun demikian, setiap model memiliki karakteristik yang berbeda dalam hal akurasi deteksi, kompleksitas komputasi, kecepatan inferensi, dan ketahanan terhadap variasi kualitas citra.

Berbagai penelitian telah menerapkan pendekatan *vision-based deep learning* untuk deteksi okupansi ruang kelas menggunakan model *object detection* seperti YOLOv8n, SSD, dan Faster R-CNN. Hasil penelitian tersebut menunjukkan bahwa model *deep learning* mampu memberikan akurasi deteksi yang tinggi pada lingkungan *indoor*(Zhang et al., 2024). Selain itu, penggunaan konfigurasi multi-kamera juga dilaporkan dapat meningkatkan cakupan area pengamatan dan akurasi estimasi okupansi. Namun demikian, pemilihan model yang paling sesuai untuk sistem deteksi okupansi berbasis kamera berbiaya rendah masih menjadi tantangan karena adanya *trade-off* antara akurasi deteksi, kecepatan inferensi, dan kebutuhan komputasi.

Meskipun berbagai penelitian telah menunjukkan bahwa model *deep learning* mampu memberikan performa yang baik dalam deteksi okupansi ruang kelas, masih terdapat beberapa keterbatasan yang belum terjawab. Pertama, sebagian besar penelitian hanya berfokus pada evaluasi satu model tertentu atau melakukan perbandingan model pada dataset dan protokol evaluasi yang berbeda,

sehingga hasil yang diperoleh sulit dibandingkan secara objektif. Kedua, implementasi sistem deteksi okupansi menggunakan perangkat kamera berbiaya rendah seperti ESP32-CAM dengan pendekatan *server-based inference* masih relatif terbatas. Ketiga, evaluasi terhadap konfigurasi *single-camera* dan *multi-camera* pada platform kamera berbiaya rendah juga masih jarang dilakukan, padahal konfigurasi kamera berpotensi memengaruhi cakupan area pengamatan, tingkat *occlusion*, dan akurasi *occupancy counting*.

Berdasarkan *research gap* tersebut, penelitian ini mengembangkan sistem deteksi okupansi ruang kelas berbasis *computer vision* menggunakan ESP32-CAM sebagai perangkat akuisisi citra dan *server-based inference* sebagai pusat pemrosesan deteksi. Penelitian ini membandingkan tiga model *object detection* yang memiliki karakteristik berbeda, yaitu SSD MobileNet V2, YOLOv8n, dan Faster R-CNN, pada konfigurasi *single-camera* dan *multi-camera* dengan berbagai skenario pengujian yang meliputi variasi jumlah penghuni, variasi jarak, dan variasi kondisi pencahayaan. Hasil penelitian diharapkan dapat memberikan kontribusi empiris dalam pemilihan model *object detection* yang paling efektif untuk sistem *occupancy counting* berbasis ESP32-CAM pada lingkungan *smart classroom*.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana performa YOLOv8n, SSD MobileNet V2, dan Faster R-CNN pada konfigurasi *single-camera*?
2. Bagaimana performa YOLOv8n, SSD MobileNet V2, dan Faster R-CNN pada konfigurasi *multi-camera*?
3. Bagaimana pengaruh variasi jumlah *occupant*, jarak, dan pencahayaan terhadap performa ketiga model?
4. Model object *detection* manakah yang paling efektif untuk sistem *occupancy counting* berbasis ESP32-CAM dengan *server-based inference*?

1.3 Tujuan Penelitian

Berdasarkan latar belakang di atas, tujuan dari penelitian ini Adalah sebagai berikut :

1. Menganalisis performa YOLOv8n, SSD MobileNet V2, dan Faster R-CNN pada konfigurasi *single-camera*.
2. Menganalisis performa YOLOv8n, SSD MobileNet V2, dan Faster R-CNN pada konfigurasi *multi-camera*.
3. Menganalisis pengaruh variasi jumlah *occupant*, jarak, dan pencahayaan terhadap performa ketiga model.
4. Menentukan model *object detection* yang paling efektif untuk sistem *occupancy counting* berbasis ESP32-CAM dengan *server-based inference*.

1.4 Manfaat Penelitian

Berdasarkan latar belakang di atas didapatkan manfaat dari penelitian ini sebagai berikut:

1. Memberikan referensi mengenai penerapan *computer vision* dan *object detection* pada sistem *occupancy counting* berbasis ESP32-CAM.
2. Memberikan informasi mengenai performa YOLOv8n, SSD MobileNet V2, dan Faster R-CNN pada konfigurasi *single-camera* dan *multi-camera*.
3. Menjadi acuan dalam pemilihan model *object detection* yang efektif untuk sistem deteksi dan penghitungan jumlah *occupant* pada ruang kelas.
4. Mendukung pengembangan sistem *smart classroom* berbasis kamera yang lebih akurat, efisien, dan aplikatif.

1.5 Batasan Masalah

Penelitian ini memiliki poin utama tersendiri yang difokuskan, sehingga terdapat beberapa batasan sebagai berikut:

1. Penelitian menggunakan ESP32-CAM sebagai perangkat akuisisi citra dengan pendekatan *server-based inference*.
2. Model *object detection* yang dievaluasi adalah YOLOv8n, SSD MobileNet V2, dan Faster R-CNN.

3. Evaluasi dilakukan pada dua konfigurasi sistem, yaitu *single-camera* dan *multi-camera*.
4. Pengujian dilakukan pada empat skenario, yaitu *single object*, *multiple object*, *distance variation*, dan *lighting condition variation*.
5. Objek yang dideteksi terbatas pada kelas *person* untuk keperluan deteksi dan penghitungan jumlah *occupant* di dalam ruang kelas.
6. Seluruh model menggunakan bobot pretrained dan dievaluasi tanpa proses *fine-tuning* atau pelatihan ulang.
7. Evaluasi performa dilakukan berdasarkan metrik *precision*, *recall*, *F1-score*, *mean absolute error* (MAE), *root mean square error* (RMSE), dan waktu pengujian.