

BAB I

PENDAHULUAN

1.1 Latar Belakang

Manajemen persediaan adalah komponen kunci dalam pengelolaan gudang yang efektif, berperan penting dalam kelancaran operasional dan pengembangan bisnis. Sistem persediaan mencakup berbagai proses penting seperti verifikasi stok, serta pencatatan barang masuk dan keluar (Yohanes Mbiri et al., 2023). Pengangkutan kardus dari mobil box ke area penyimpanan gudang merupakan salah satu tahapan penting dalam proses ini. Namun, masalah sering muncul ketika jumlah kardus yang tercatat tidak sesuai dengan jumlah yang sebenarnya sehingga menyebabkan kerugian finansial dan menurunkan kepercayaan konsumen. Jumlah barang yang tidak sesuai sering disebabkan oleh kesalahan pencatatan manual, yang terjadi karena kelelahan karyawan selama proses penghitungan (Teja et al., 2023).

Untuk mengatasi masalah ini, diperlukan teknologi yang modern dan akurat. Teknologi deteksi objek menggunakan *computer vision* merupakan solusi yang. Dengan kemampuannya mengidentifikasi objek secara otomatis dari citra digital (Masril & Caniago, 2023), *computer vision* memungkinkan penghitungan kardus secara *real-time* serta meminimalkan kesalahan yang sering terjadi pada pencatatan manual (Anjali Dompeipen et al., 2021).

Meskipun teknologi *computer vision* menawarkan solusi yang menarik, tantangan besar tetap ada. Salah satu tantangan utama adalah tingginya kesamaan visual pada beberapa kardus yang memiliki ukuran, warna, atau desain kemasan yang mirip (Firmandicky & Susetyo, 2024). Selain itu, pola yang seragam pada tekstur kardus yang berulang dapat mengaburkan sistem deteksi, terutama ketika kardus ditumpuk mengakibatkan kesalahan dalam penghitungan atau kebingungan pada model (Yang et al., 2022).

Salah satu solusi untuk mengatasi kemiripan visual antar objek adalah dengan menggunakan *instance segmentation*, yaitu metode yang menggabungkan segmentasi piksel dan deteksi objek secara simultan untuk membedakan setiap objek dalam citra, termasuk yang berasal dari kelas yang sama, di mana *Mask R-CNN* merupakan model pertama yang dirancang khusus untuk menangani

permasalahan ini (Goel & Koundal, 2025). Penelitian oleh Naumann et al., (2024) menunjukkan bahwa *Mask R-CNN* (*Mask Region-Based Convolutional Neural Network*) dengan *ResNet-50-FPN* sebagai *backbone* memiliki kinerja tinggi pada deteksi *keypoint* dengan *keypoint AP* sebesar 75.76 pada data nyata. Selain itu penelitian oleh Fang et al., (2023) menunjukkan bahwa *Mask R-CNN* dengan *backbone ResNeXt* menghasilkan *mAP* (*mean Average Precision*) sebesar 62.62% untuk deteksi objek dan 57.58% untuk segmentasi. Namun, penelitian ini menekankan bahwa model tersebut membutuhkan komputasi tinggi yang diukur dalam jumlah *FLOPs*, sehingga tidak optimal untuk perangkat dengan sumber daya terbatas. Kemudian Penelitian oleh Rubin et al., (2022) membandingkan *Mask R-CNN* dengan *backbone ResNet-50*, *ResNet-101*, dan *MobileNet-V1*. Hasilnya, *ResNet-101* mencapai *mAP* tertinggi sebesar 98.3%. Selanjutnya Penelitian oleh Huang et al., (2024) menunjukkan bahwa mengganti *backbone ResNet101* pada *Mask R-CNN* dengan *MobileNetV2* berhasil mengurangi jumlah parameter sebesar 92,4% (dari 44,5 juta menjadi 3,4 juta), meningkatkan kecepatan pemrosesan lebih dari 50%, dan mempertahankan akurasi deteksi yang tinggi, menjadikannya pilihan ideal untuk aplikasi *real-time* pada perangkat dengan sumber daya terbatas. Terakhir Penelitian lain oleh Cao et al., (2023) mengembangkan algoritma *YOLOv5* untuk deteksi *real-time* ringan dengan mengganti *backbone* dengan *MobileNetV3* mampu meningkatkan akurasi deteksi objek kecil sekaligus mengurangi parameter hingga 87,4%.

Penelitian ini akan menggunakan *Mask R-CNN* dengan mengganti *backbone* dengan *MobileNetV3* untuk mengatasi permasalahan pencatatan manual serta mengurangi beban komputasi agar dapat diterapkan pada perangkat dengan sumber daya terbatas.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka di ambil perumusan masalah yaitu sebagai berikut.

1. Bagaimana Implementasi deteksi kardus menggunakan *Mask R-CNN* dengan *backbone* default (*ResNet-50*) dan *Mask R-CNN* dengan *backbone MobileNetV3 Small*?

2. Bagaimana konfigurasi *Mask R-CNN* dengan *MobileNetV3 Small* yang paling optimal untuk meningkatkan akurasi deteksi kardus?
3. Bagaimana tingkat akurasi, ukuran model dan kecepatan inferensi (*FPS*) antara *Mask R-CNN* default dibandingkan dengan *Mask R-CNN* yang menggunakan *backbone MobileNetV3 Small* dalam deteksi kardus?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk mengevaluasi dan mengimplementasikan modifikasi pada arsitektur *Mask R-CNN* dalam mendeteksi kardus. Adapun tujuan spesifik dari penelitian ini meliputi:

1. Mengimplementasi deteksi kardus menggunakan *Mask R-CNN* dengan *backbone* default (*ResNet-50*) dan *Mask R-CNN* dengan *backbone MobileNetV3 Small*
2. Menentukan konfigurasi *Mask R-CNN* dengan *MobileNetV3 Small* yang paling optimal untuk meningkatkan akurasi deteksi kardus
3. Membandingkan tingkat akurasi, ukuran model dan kecepatan inferensi (*FPS*) *Mask R-CNN* default dengan *Mask R-CNN* yang menggunakan *backbone MobileNetV3 Small* dalam deteksi kardus?

1.4 Manfaat

Penelitian ini diharapkan dapat memberikan kontribusi baik secara praktis maupun akademis. Adapun manfaat yang dapat diperoleh dari hasil penelitian ini adalah sebagai berikut:

1. Praktis: Sistem deteksi kardus otomatis yang dihasilkan akan meningkatkan proses pengelolaan persediaan dengan meminimalkan ketergantungan pada pencatatan manual, sehingga mengurangi risiko kesalahan dan potensi kerugian finansial.
2. Akademis: Penelitian ini menjadi referensi yang bernilai bagi pengembangan teknologi *computer vision*, khususnya dalam integrasi *Mask R-CNN* dan *MobileNetV3* untuk segmentasi *instance* pada deteksi objek.