Development Of Bag Truck And Box Annotation Dataset For Vision-Based Tracking System Implementation

Pembuatan Dataset Anotasi Bag Truk Dan Box Untuk Penerapan Sistem Vision-Based Tracking

M Shofi Mubarrok[1], Cindy Taurusta[2]

{Shofimubarrok7@gmail.com[1]; Cindy@umsida.ac.id }

[1,2]Program Studi Informatika, Fakultas Sains dan Teknologi, Universitas Muhammadiyah Sidoarjo, Indonesia

Abstract. *General Background Logistics operations increasingly rely on automated visual perception to monitor cargo handling and streamline material movement across industrial facilities. Specific Background Manual loading and unloading of truck bags and boxes remain susceptible to delays, human error, and inconsistent record keeping across conveyors and storage zones. Knowledge Gap Existing computer vision and logistics tracking studies often focus on general object detectors or broad warehouse benchmarks rather than curated bounding box datasets capturing truck bags and boxes under real unloading, conveyor transfer, occlusion, and lighting variations. Aims This study develops an annotated dataset of truck bags and boxes using bounding box labeling on operational logistics imagery to support real-time object detection and tracking models. Results Field images from truck loading and conveyor transfer scenes were annotated by object category and position, validated by two annotators across bounding box placement, label accuracy, and completeness, and structured into training, validation, and testing splits with metadata for detection and tracking formats. Novelty This research compiles a domain-specific truck bag and box bounding box dataset from operational logistics scenes covering conveyor motion, stacking inside truck beds, overhead camera angles, and variable lighting. Implications Providing validated multi-condition annotations supports model training for automated cargo counting, movement monitoring, and reduced recording errors in logistics automation.*
Keywords *Image Annotation, Object Detection, Bounding Box, Logistics Automation, Computer Vision*

Abstrak. Latar Belakang Umum Operasi logistik semakin mengandalkan persepsi visual otomatis untuk memantau penanganan kargo dan memperlancar pergerakan material di fasilitas industri. Latar Belakang Spesifik Pemuatan dan pembongkaran manual tas dan kotak truk tetap rentan terhadap keterlambatan, kesalahan manusia, dan pencatatan yang tidak konsisten di seluruh konveyor dan zona penyimpanan. Kesenjangan Pengetahuan Studi visi komputer dan pelacakan logistik yang ada sering kali berfokus pada detektor objek umum atau tolok ukur gudang luas daripada dataset kotak pembatas yang dikurasi yang menangkap tas dan kotak truk di bawah variasi pembongkaran nyata, transfer konveyor, oklusi, dan pencahayaan. Tujuan Studi ini mengembangkan dataset beranotasi tas dan kotak truk menggunakan label kotak pembatas pada citra logistik operasional untuk mendukung model deteksi dan pelacakan objek secara real-time. Hasil Citra lapangan dari adegan pemuatan truk dan transfer konveyor diberi anotasi berdasarkan kategori dan posisi objek, divalidasi oleh dua anotasi di penempatan kotak pembatas, akurasi label, dan kelengkapan, serta disusun menjadi pembagian pelatihan, validasi, dan pengujian dengan metadata untuk format deteksi dan pelacakan. Novelty Penelitian ini mengumpulkan dataset tas truk dan kotak pembatas khusus domain dari adegan logistik operasional yang mencakup gerakan konveyor, penumpukan di dalam bak truk, sudut kamera overhead, dan pencahayaan variabel. Implikasi Menyediakan anotasi multi-kondisi yang tervalidasi mendukung pelatihan model untuk penghitungan kargo otomatis, pemantauan pergerakan, dan pengurangan kesalahan pencatatan dalam otomasi logistik.
Kata Kunci: Anotasi Gambar, Deteksi Objek, Kotak Pembatas, Otomasi Logistik, Visi Komputer

I. PENDAHULUAN

Kegiatan logistik seperti proses bongkar muat barang masih banyak mengandalkan tenaga manusia sehingga berpotensi menimbulkan ketidakefisienan serta kesalahan dalam identifikasi dan pencatatan barang. Ketergantungan ini dapat menyebabkan keterlambatan distribusi, tingginya risiko human error, serta meningkatnya biaya operasional. Seiring meningkatnya kebutuhan pengiriman barang dalam industri logistik modern, terutama pada sektor e-commerce dan supply chain global, diperlukan solusi otomatis yang mampu meningkatkan produktivitas dan akurasi dalam sistem monitoring barang.

Vision-based tracking merupakan salah satu solusi yang saat ini berkembang pesat sebagai bagian dari teknologi automasi. Dengan memanfaatkan kamera dan algoritma pendeteksian objek berbasis kecerdasan buatan, sistem ini mampu

melakukan pelacakan objek secara real-time tanpa intervensi manual. Implementasi teknologi ini telah diterapkan pada berbagai bidang seperti surveillance, transportasi, dan manajemen gudang[1]. Namun, keberhasilan sistem vision-based tracking sangat dipengaruhi oleh kualitas data pelatihan yang digunakan, khususnya dataset anotasi yang relevan dengan kondisi lingkungan nyata.

Dalam konteks *logistik*, objek yang dilacak memiliki karakteristik visual yang sangat beragam, salah satunya bag truk dan box yang dipindahkan dari kendaraan ke conveyor ataupun area penyimpanan. Kondisi lingkungan seperti variasi pencahayaan, pergerakan cepat, tumpukan barang, dan kemiripan warna seringkali menjadi tantangan dalam proses deteksi otomatis. Oleh karena itu, dibutuhkan dataset anotasi yang mencakup beragam kondisi dan sudut pandang agar model visi komputer mampu beradaptasi dan bekerja secara optimal di area operasional logistik.

Berdasarkan permasalahan tersebut, penelitian ini berfokus pada pembuatan dataset anotasi objek bag truk dan box yang diperoleh dari lingkungan logistik nyata. Dataset ini dibuat menggunakan teknik bounding box untuk menandai setiap objek pada citra. Hasil penelitian ini diharapkan dapat mendukung pengembangan model deteksi dan pelacakan objek secara otomatis pada industri logistik, sehingga mampu meningkatkan efisiensi proses bongkar muat, mempercepat alur distribusi barang, serta meminimalkan kesalahan dalam sistem monitoring inventaris.

II. METODE

Penelitian ini dilakukan dengan tujuan untuk menghasilkan dataset anotasi objek berupa *bag* truk dan *box* yang dapat mendukung pengembangan sistem vision-based tracking pada lingkungan industri logistik. Metode penelitian dibagi ke dalam beberapa tahap yang meliputi pengumpulan data citra, proses anotasi manual, validasi anotasi oleh ahli, pengelompokan kategori objek, serta persiapan dataset untuk model deteksi dan pelacakan objek. Setiap tahap dirancang untuk memastikan bahwa dataset yang dihasilkan memiliki kualitas yang baik dan layak digunakan sebagai data pelatihan dalam sistem visi komputer.

Tahap pertama merupakan **pengumpulan data citra**, di mana sumber data diambil dari area operasional logistik seperti proses bongkar muat pada truk dan jalur conveyor. Kamera diposisikan pada beberapa sudut pandang untuk memperoleh variasi data yang cukup, termasuk variasi kondisi pencahayaan, jarak kamera terhadap objek, serta tingkat kepadatan barang. Citra yang dikumpulkan kemudian dipilah dan dipilih berdasarkan kejelasan objek serta relevansi terhadap kebutuhan pelatihan model.

Tahap selanjutnya adalah **proses anotasi manual** menggunakan perangkat lunak anotasi seperti LabelImg atau software anotasi serupa. LabelImg adalah sebuah perangkat lunak (software) *open-source* yang digunakan untuk melakukan anotasi citra. Proses anotasi dilakukan dengan memberikan *bounding box* pada setiap objek target, yaitu bag truk dan box. Setiap objek diberi label sesuai kategori dan posisinya dalam citra, apakah berada di dalam truk, pada conveyor, atau dalam proses perpindahan. Tahapan ini menjadi sangat penting karena ketepatan anotasi akan berpengaruh langsung terhadap performa model deteksi objek yang digunakan pada tahap selanjutnya.

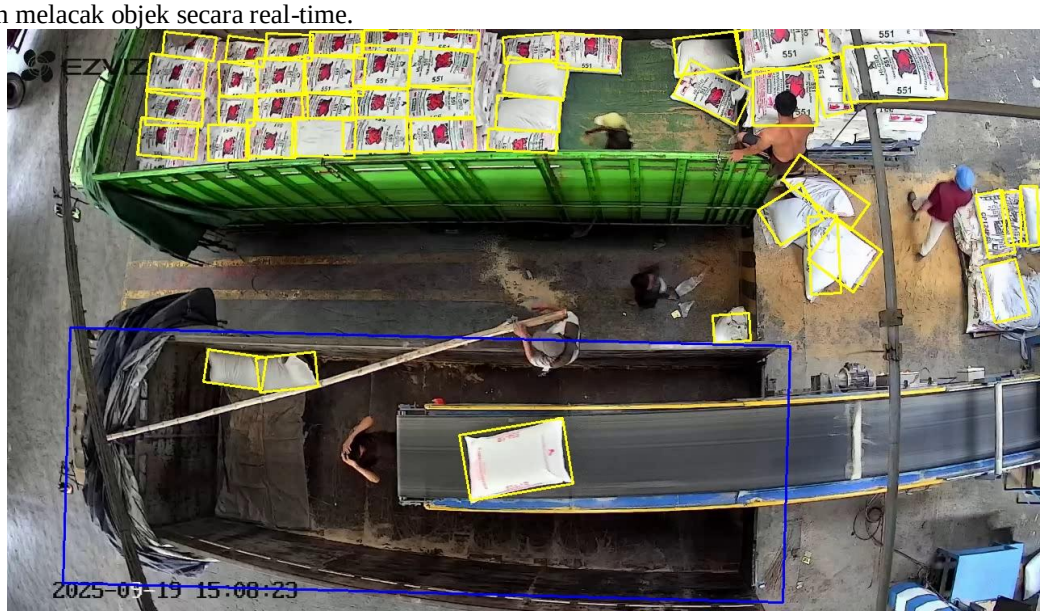
Pada tahap ketiga dilakukan **validasi anotasi**, yang bertujuan memastikan konsistensi dan kualitas penandaan objek. Validasi ini dilakukan oleh dua annotator atau ahli yang memahami karakteristik objek logistik. Parameter penilaian mencakup ketepatan posisi bounding box, kesesuaian label kategori objek, serta kelengkapan anotasi pada citra. Jika ditemukan data yang tidak konsisten atau kurang akurat, dilakukan revisi untuk memastikan dataset sesuai standar ilmiah.

Tahap terakhir adalah **persiapan dataset**, di mana dataset yang telah divalidasi dikonversi ke dalam format yang sesuai untuk pelatihan model kecerdasan buatan seperti YOLO annotation format atau Pascal VOC XML. Dataset kemudian dibagi ke dalam tiga subset yaitu data pelatihan (training), validasi, dan pengujian (testing) dengan proporsi tertentu agar model yang dikembangkan nantinya mampu melakukan generalisasi secara optimal. Dataset akhir juga dilengkapi dengan dokumentasi struktur file, daftar label objek, dan informasi metadata lainnya agar mudah diimplementasikan dalam berbagai model deteksi dan pelacakan objek.

III. HASIL DAN PEMBAHASAN

Dataset anotasi yang dikembangkan pada penelitian ini memuat variasi kondisi lingkungan operasional logistik untuk mendukung penerapan sistem *vision-based tracking*. Fokus utama anotasi terletak pada objek bag dan box yang menjadi komponen utama dalam alur distribusi barang. Pembuatan dataset dilakukan berdasarkan pengambilan data lapangan yang secara langsung merepresentasikan proses logistik, mulai dari bongkar muat pada truk hingga tahapan pemindahan ke conveyor untuk didistribusikan ke titik sortir. Variasi sudut kamera, intensitas cahaya, serta kondisi objek yang saling tumpang tindih, dipertimbangkan sebagai faktor penting agar dataset dapat mencerminkan situasi nyata industri logistik.

mendeteksi dan melacak objek secara real-time.



Gambar 1: hasil anotasi pada area bongkar muat dengan conveyor

Gambar 1 menunjukkan objek bag dan box yang sedang dalam proses perpindahan dari truk ke conveyor oleh tenaga kerja lapangan. Kondisi ini menggambarkan mobilitas objek yang tinggi dan sering mengalami perubahan orientasi maupun tumpang tindih. Proses anotasi dilakukan dengan menerapkan bounding box berwarna kuning, yang disesuaikan dengan bentuk objek secara presisi meskipun objek mengalami distorsi visual dan pergerakan cepat. Selain itu, area conveyor turut dianotasi menggunakan garis pemisah untuk kebutuhan analisis tracking path serta untuk mempermudah model dalam membedakan antara objek statis, bergerak, dan area kerja.

Keberhasilan anotasi pada kondisi ini membuktikan bahwa dataset mampu menangkap karakteristik kompleks yang umum terjadi pada sektor logistik manual. Tantangan seperti motion blur, occlusion, dan penumpukan barang teratasi melalui konsistensi penandaan pada setiap frame yang direkam sebagai bagian dari data video.



Gambar 2. hasil anotasi pada objek bag di dalam bak truk

Berbeda dengan kondisi sebelumnya, Gambar 2 menampilkan objek bag yang berada dalam posisi statis di dalam truk. Meskipun objek tidak bergerak, proses anotasi masih menghadapi tantangan berupa penataan objek yang bertumpuk dan sudut kamera menyudut dari sisi atas, yang dapat menyebabkan distorsi perspektif. Namun, bounding box tetap diterapkan secara akurat mengikuti batas visual objek sehingga membantu model dalam mempelajari bentuk bag dari berbagai

orientasi.

Selain itu, kondisi pencahayaan di dalam bak truk yang cenderung redup juga menjadi pertimbangan penting, sehingga dataset ini dapat meningkatkan ketahanan model terhadap variasi intensitas cahaya. Kualitas hasil anotasi terbukti tetap konsisten dalam kondisi pencahayaan yang minim maupun bercampur cahaya luar yang masuk dari sisi truk.

Secara keseluruhan, kedua contoh gambar tersebut menunjukkan bahwa dataset hasil anotasi memiliki kedalaman variasi data yang dapat berkontribusi secara signifikan terhadap *generalization ability* model deteksi objek. Keberagaman posisi, orientasi, dan kondisi lingkungan menjadikan dataset lebih representatif terhadap situasi logistik sebenarnya, sehingga model dapat berfungsi optimal saat diimplementasikan secara real-time di lapangan.

Proses anotasi yang konsisten dan akurat akan mendukung peningkatan performa model dalam mendeteksi dan melacak objek pada area kerja yang dinamis. Hal ini akan berdampak langsung terhadap efisiensi proses penanganan barang, seperti identifikasi jumlah bag yang sedang dipindahkan, pelacakan pergerakan barang dari truk ke conveyor, serta monitoring potensi kesalahan dalam proses distribusi.

Dengan demikian, hasil dan pembahasan pada penelitian ini tidak hanya menegaskan kualitas dataset yang dihasilkan, tetapi juga memberikan gambaran kuat mengenai kontribusi nyata dalam mendukung pengembangan sistem automasi logistik berbasis visi komputer di lingkungan industri yang terus berkembang.

IV. KESIMPULAN

Penelitian ini berhasil menghasilkan dataset anotasi objek bag dan box yang diperoleh langsung dari lingkungan operasional logistik nyata. Proses anotasi dilakukan menggunakan teknik *bounding box* untuk menandai objek pada kondisi yang beragam, seperti pergerakan dinamis di conveyor, penataan barang dalam truk, sudut pandang kamera yang tinggi, serta variasi pencahayaan dan *occlusion*. Variasi kondisi ini memberikan nilai tambah signifikan karena mampu meningkatkan kemampuan generalisasi model deteksi objek yang akan dilatih menggunakan dataset tersebut.

Hasil yang diperoleh menunjukkan bahwa dataset dapat mendukung pengembangan sistem vision-based tracking secara real-time dalam proses bongkar muat logistik. Implementasi sistem ini diharapkan mampu memberikan kontribusi terhadap peningkatan efisiensi operasional, pengurangan kesalahan pencatatan barang, dan automasi proses monitoring. Dengan demikian, penelitian ini memberikan fondasi penting bagi pengembangan lebih lanjut dalam bidang automasi logistik berbasis teknologi *Computer Vision*.

Ke depannya, pengembangan dataset dapat diarahkan pada penambahan jenis objek, variasi kondisi lingkungan yang lebih ekstrem, serta integrasi metode anotasi berbasis segmentasi untuk menghasilkan representasi objek yang lebih detail. Selain itu, evaluasi performa model yang dilatih dari dataset ini juga diperlukan sebagai langkah lanjutan untuk menilai kelayakan implementasi sistem di lingkungan industri secara penuh.

REFERENSI

[1] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” *arXiv Preprint arXiv:1804.02767*, 2018.

[2] T.-Y. Lin *et al.*, “Microsoft COCO: Common Objects in Context,” in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2014, pp. 740–755.

[3] M. Everingham *et al.*, “The Pascal Visual Object Classes Challenge: A Retrospective,” *Int. J. Comput. Vis.*, vol. 111, no. 1, pp. 98–136, 2015.

[4] W. Liu *et al.*, “SSD: Single Shot MultiBox Detector,” in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 21–37.

[5] M. Chen and Y. Zhang, “Vision-Based Tracking in Logistics Automation: A Review,” *IEEE Access*, vol. 9, pp. 1–18, 2021.

[6] S. Zang *et al.*, “Real-Time Object Detection for Industrial Logistics,” *Sensors*, vol. 20, no. 12, 2020.

[7] H. Fang *et al.*, “Bounding Box Annotation Strategies for Machine Learning-Based Inspection Systems,” *J. Manuf. Syst.*, vol. 62, pp. 1–14, 2022.

[8] J. J. Losada del Olmo, E. P. Ballesteros, Á. L. Perales Gómez *et al.*, “Fog Computing-Driven Logistics:

Leveraging Few-Shot Learning and Foundational Computer Vision Models," *Clust. Comput.*, vol. 28, no. 14, p. 938, 2025.

[9] P. Kadam, G. Fang, and J. J. Zou, "Object Tracking Using Computer Vision: A Review," *Computers*, vol. 13, no. 6, p. 136, 2024.

[10] J. Rutinowski, H. Youssef, S. Franke *et al.*, "Semi-Automated Computer Vision-Based Tracking of Multiple Industrial Entities: A Framework and Dataset Creation Approach," *EURASIP J. Image Video Process.*, vol. 2024, no. 1, p. 8, 2024.

[11] J. Rutinowski *et al.*, "Semi-Automated Computer Vision-Based Tracking of Multiple Industrial Entities: A Framework and Dataset Creation Approach," *EURASIP J. Image Video Process.*, vol. 2024, no. 1, p. 8, 2024, doi: 10.1186/s13640-024-00623-6.

[12] P. Kadam, G. Fang, and J. J. Zou, "Object Tracking Using Computer Vision: A Review," *Computers*, vol. 13, no. 6, p. 136, 2024, doi: 10.3390/computers13060136.

[13] M. Cheng and H. Shahriar, "A Deep Learning Framework for Object Detection and Tracking: CNN-Based Parcel Tracking on Conveyor Belts," in *Proc. 6th Int. Conf. Sustain. Technol. Ind. 5.0 (STI)*, IEEE, 2024, doi: 10.1109/STI64222.2024.10951144.

[14] J. J. Losada del Olmo, E. P. Ballesteros, Á. L. Perales Gómez *et al.*, "Fog Computing-Driven Logistics: Leveraging Few-Shot Learning and Foundational Computer Vision Models," *Clust. Comput.*, vol. 28, no. 14, p. 938, 2025, doi: 10.1007/s10586-025-05600-2.

[15] K. Guo and L. Zhang, "Detection and Classification of Finished Parts on a Conveyor Belt Using an Improved YOLOv8n and a Task-Specific Multi-Scene Dataset," *J. Adv. Mech. Des. Syst. Manuf.*, vol. 20, no. 7, p. JAMDSM0065, 2026.