

BAB 3

PELAKSANAAN KERJA MAGANG

3.1 Kedudukan dan Koordinasi

Pelaksanaan kegiatan magang di KinetixPro Pte. Ltd. berada di bawah Divisi *Quality Assurance* (QA) dengan penempatan pada posisi *Quality Assurance Engineer Intern*. Berdasarkan struktur organisasi perusahaan, Divisi *Quality Assurance* berada dalam lingkup tanggung jawab langsung *Chief Technology Officer* (CTO), dengan pengawasan umum proyek oleh *Chief Executive Officer* (CEO). Dalam pelaksanaannya, kegiatan magang berada di bawah bimbingan *Senior Artificial Intelligence & Machine Learning Engineer*, Thomas Clark selaku supervisor, yang memastikan kegiatan berjalan sesuai rencana serta mendukung kolaborasi lintas fungsi bersama tim teknis utama perusahaan.

Supervisor memberikan pendampingan secara reguler untuk memastikan setiap tugas terselesaikan tepat waktu, serta memberikan pengarahan teknis dan bantuan dalam penyelesaian permasalahan yang kompleks. Selain melakukan pemantauan, supervisor juga menjadi penghubung antara tim *Quality Assurance* dan pihak manajemen teknis, memastikan hasil evaluasi serta perkembangan proyek tersampaikan secara terstruktur pada setiap sesi rapat formal.

Proses kerja dilaksanakan sepenuhnya secara daring karena KinetixPro menerapkan sistem kerja *remote*. Seluruh komunikasi proyek, pelacakan tugas, dan pelaporan progres dilakukan melalui kombinasi beberapa *platform* digital seperti Google Meet, Discord, dan Jira. Google Meet digunakan untuk rapat rutin dua kali setiap minggu, yang terdiri atas sesi *developer catchup* dan *weekly interim meeting*. Discord dimanfaatkan untuk komunikasi harian melalui teks maupun panggilan singkat bersifat *impromptu*, sedangkan Jira digunakan sebagai sistem pelacakan tugas dan penugasan proyek. Pertemuan rutin sebagaimana dijelaskan sebelumnya, berfungsi untuk memantau pencapaian target mingguan serta menyelaraskan hasil evaluasi dengan rencana pengembangan berikutnya. Sementara itu, komunikasi harian difokuskan pada sinkronisasi progres tim dan pelaporan hasil pengujian secara berkelanjutan. Pola ini menjaga efektivitas kerja lintas divisi serta memastikan proses evaluasi berjalan terstruktur dan terdokumentasi dengan baik.

3.2 Tugas yang Dilakukan

Selama pelaksanaan magang di KinetixPro Pte. Ltd., kegiatan difokuskan pada proses analisis performa sistem deteksi visual berbasis YOLOv7 yang telah diterapkan pada lingkungan produksi. Setiap tugas dirancang untuk mendukung tujuan utama, yaitu meningkatkan akurasi dan stabilitas sistem melalui analisis akar penyebab penurunan performa serta implementasi langkah perbaikan yang berkelanjutan. Kegiatan dilaksanakan secara kolaboratif bersama tim *Artificial Intelligence & Machine Learning (AI/ML)* dan *Product Development* dengan pendekatan yang terstruktur serta berorientasi pada hasil evaluasi model. Adapun rincian tugas yang dilakukan adalah sebagai berikut.

1. Melakukan pengumpulan dan pengorganisasian data visual (gambar maupun video) dari hasil deteksi sistem untuk dianalisis lebih lanjut sebagai bagian dari proses *performance monitoring*.
2. Melakukan visualisasi kesalahan deteksi menggunakan model yang telah di-*deploy* untuk melakukan pengujian langsung terhadap performa model itu sendiri, sehingga pola kesalahan dapat diamati secara kontekstual dan lebih mudah ditelusuri.
3. Melakukan validasi terhadap hasil deteksi model YOLOv7 dengan membandingkannya terhadap *ground truth* atau data referensi guna mengidentifikasi potensi kesalahan klasifikasi maupun deteksi ganda.
4. Melakukan analisis akar penyebab (*root cause analysis*) terhadap kasus penurunan performa, termasuk penelusuran aspek data, model, dan integrasi sistem yang berpotensi memengaruhi akurasi deteksi.
5. Mendokumentasikan hasil analisis, temuan, dan rekomendasi perbaikan ke dalam laporan internal yang digunakan sebagai dasar dalam proses peninjauan dan pengambilan keputusan oleh tim pengembang.

Melalui rangkaian kegiatan tersebut, kontribusi terhadap peningkatan kualitas dan keandalan sistem berbasis kecerdasan buatan dapat diwujudkan secara berkelanjutan. Pelaksanaan tugas yang terarah serta berbasis evaluasi komprehensif ini juga mendukung terciptanya sistem deteksi visual yang lebih adaptif, stabil, dan sesuai dengan standar mutu yang diterapkan perusahaan.

3.3 Uraian Pelaksanaan Magang

Kegiatan magang dilaksanakan secara bertahap dan berkelanjutan, dengan fokus utama pada proses *quality assurance* (QA), analisis performa model deteksi visual, serta peningkatan akurasi sistem berbasis YOLO. Setiap minggu memiliki cakupan tugas yang berbeda, mengikuti alur peningkatan tanggung jawab teknis dan kedalaman analisis terhadap hasil deteksi. Seluruh aktivitas dilakukan secara kolaboratif dengan tim *Artificial Intelligence & Machine Learning* (AI/ML) dalam konteks evaluasi model pasca-*deployment*, guna memastikan performa sistem tetap konsisten dan memenuhi standar kualitas. Rincian kegiatan magang selama enam belas minggu disajikan pada Tabel 3.1 berikut.

Tabel 3.1. Pekerjaan mingguan selama pelaksanaan kerja magang

Minggu Ke -	Pekerjaan yang dilakukan
1	Melakukan <i>deployment</i> model deteksi orang dan memantau hasil deteksinya untuk memastikan keluaran sistem berjalan stabil di lingkungan produksi.
2	Memperbaiki anotasi pada data orang, memvisualisasikan kesalahan deteksi, serta menyusun laporan analisis mengenai <i>false detection</i> yang ditemukan pada hasil inferensi model.
3	Berdiskusi dengan tim <i>Artificial Intelligence & Machine Learning</i> (AI/ML) mengenai penyesuaian ambang deteksi (<i>threshold</i>) pada aturan tertentu untuk mengoptimalkan performa model, melakukan pemeriksaan statistik setiap versi model, serta mendokumentasikan permasalahan deteksi yang terulang.
4	Melakukan pengujian visualisasi terhadap <i>false detection</i> pada deteksi <i>Personal Protective Equipment</i> (PPE), kemudian memfilter dan memindahkan hasil deteksi yang benar ke antarmuka pengguna (UI).
Lanjut pada halaman berikutnya	

Tabel 3.1 Pekerjaan mingguan selama pelaksanaan kerja magang (lanjutan)

Minggu Ke -	Pekerjaan yang dilakukan
5	Melakukan analisis kesalahan deteksi dari hasil visualisasi dan pengujian model pada lebih banyak rekaman, melatih ulang model untuk kelas yang lebih spesifik, memisahkan model untuk keperluan <i>debugging</i> dan pemantauan performa, serta memperbaiki label kendaraan sebelum melakukan pelatihan ulang dan penyesuaian <i>threshold</i> .
6	Melakukan penyaringan dan pra-pemrosesan dataset untuk model deteksi orang yang telah diperbarui, melakukan pemeriksaan kesalahan deteksi per kamera untuk menentukan akar penyebab, serta menguji performa deteksi <i>Personal Protective Equipment</i> (PPE) pada tingkat model dan tingkat waktu.
7	Memperluas area <i>bounding box</i> pada gambar berlabel, menyiapkan data baru, melakukan pelabelan, dan menjalankan iterasi pelatihan model baru berdasarkan hasil pembaruan tersebut.
8	Menyusun laporan statistik hasil deteksi untuk aturan kendaraan, melakukan pemeriksaan <i>recall</i> menggunakan data rekaman aktual, serta memvisualisasikan nilai <i>recall</i> melalui pengujian model yang sedang diimplementasikan.
9	Berkolaborasi dengan tim <i>Artificial Intelligence & Machine Learning</i> (AI/ML) untuk menganalisis kasus <i>false detection</i> berulang, menyiapkan fase awal pengembangan model <i>housekeeping</i> , membuat visualisasi kesalahan dari model terbaru, dan menjalankan beberapa iterasi pelatihan <i>housekeeping</i> beserta versi dataset-nya.
10	Melakukan <i>deployment</i> model <i>housekeeping</i> , menguji performanya secara <i>real-time</i> , menyusun aturan deteksi baru, serta mengonfigurasi ambang deteksi (<i>threshold</i>) agar sesuai dengan kondisi produksi.
Lanjut pada halaman berikutnya	

Tabel 3.1 Pekerjaan mingguan selama pelaksanaan kerja magang (lanjutan)

Minggu Ke -	Pekerjaan yang dilakukan
11	Melakukan pengumpulan data baru untuk model <i>housekeeping</i> , melabeli data tersebut bersama tim, menjalankan beberapa iterasi model dengan kondisi berbeda, serta menguji hasilnya secara langsung untuk menilai performa aktual.
12	Mengekstraksi rekaman hasil <i>housekeeping</i> , menganalisis penyebab kesalahan deteksi, membandingkan hasilnya dengan lingkungan produksi, dan memberikan umpan balik kepada tim <i>Artificial Intelligence & Machine Learning</i> (AI/ML) terkait potensi kesalahan konfigurasi sistem.
13	Memantau nilai <i>recall</i> pada aturan kendaraan, meninjau ulang persyaratan <i>true positive</i> untuk aturan boomlift, serta memvisualisasikan kesalahan deteksi yang terjadi pada masing-masing aturan.
14	Mengekstraksi rekaman dari seluruh kamera yang memiliki isu deteksi, melakukan visualisasi lokal untuk menelusuri akar penyebab, memperbaiki kesalahan <i>bounding box</i> pada tingkat dataset, dan melatih ulang model dengan data yang telah diperbaiki.
15	Memeriksa seluruh aturan untuk menilai presisi, menganalisis aturan dengan performa rendah seperti kendaraan dan forklift, serta berkoordinasi dengan tim <i>Artificial Intelligence & Machine Learning</i> (AI/ML) untuk pengujian di lingkungan produksi sambil mengekstraksi kesalahan deteksi baru dari <i>pipeline</i> untuk evaluasi QA lokal.
16	Mengidentifikasi sumber permasalahan pada anotasi data, menyiapkan <i>pipeline</i> untuk pengumpulan data berkelanjutan, serta melakukan iterasi cepat dalam skala kecil untuk menyesuaikan model terhadap kendaraan atau objek baru yang belum pernah terdeteksi sebelumnya.

3.3.1 User Requirement

User requirement pada proyek ini disusun untuk menggambarkan kebutuhan dan ekspektasi pengguna terhadap performa sistem deteksi visual yang dikembangkan oleh KinetixPro Pte. Ltd. Kebutuhan ini menjadi dasar bagi perusahaan dalam memastikan kualitas model berbasis YOLO tetap konsisten dan memenuhi standar performa perusahaan. Dengan demikian, seluruh proses analisis dan peningkatan model difokuskan pada pemenuhan kebutuhan deteksi yang akurat, stabil, dan adaptif terhadap variasi kondisi di lapangan.

Secara umum, kebutuhan pengguna dalam konteks proyek ini mencakup aspek ketepatan deteksi, keandalan sistem, serta waktu respons model. Adapun rincian kebutuhan tersebut dijabarkan sebagai berikut.

1. Akurasi Deteksi Tinggi

Sistem mampu memberikan hasil deteksi dengan tingkat *precision* yang tinggi pada berbagai kondisi operasional. Setiap objek yang relevan, seperti pekerja, kendaraan, dan alat bantu kerja, harus dapat dikenali dengan benar tanpa menghasilkan *false detection* yang signifikan.

2. Responsivitas dan Ketepatan Deteksi

Sistem memiliki tingkat *recall* yang tinggi agar mampu merespons perubahan kondisi di lapangan secara andal, memastikan setiap objek relevan terdeteksi tanpa terlewat. Hal ini menjadi aspek penting dalam menjaga efektivitas pemantauan keselamatan kerja secara waktu nyata.

3. Konsistensi dan Adaptivitas Model

Model perlu menjaga performanya secara stabil di berbagai kondisi lingkungan, termasuk perubahan pencahayaan, posisi kamera, maupun kemunculan objek baru yang belum pernah dikenali sebelumnya. Kemampuan beradaptasi terhadap variasi tersebut menjadi faktor utama untuk mempertahankan keandalan sistem dalam jangka panjang.

Berdasarkan kebutuhan tersebut, kegiatan magang difokuskan pada proses analisis performa dan identifikasi penyebab penurunan akurasi model. Setiap temuan dan perbaikan yang dilakukan diarahkan untuk memastikan sistem deteksi visual memenuhi standar kualitas yang diharapkan klien serta mendukung peningkatan keandalan produk perusahaan secara berkelanjutan.

3.3.2 Pengumpulan dan Pengorganisasian Data Visual

Tahap awal kegiatan difokuskan pada proses pengumpulan dan pengorganisasian data visual sebagai dasar untuk analisis performa sistem. Proses ini dilakukan melalui pemantauan hasil deteksi yang muncul di antarmuka pengguna (UI). Ketika ditemukan kesalahan deteksi atau perilaku model yang tidak sesuai, hasil tersebut dipindahkan ke dalam satu *bucket* khusus untuk ditandai sebagai data evaluasi. Seluruh data yang ditandai ini kemudian akan dicocokkan dengan rekaman video aktual yang tersimpan di *server* maupun hasil tangkapan gambar otomatis (*auto-labeled capture*) yang diambil menggunakan *pipeline* internal.

Proses ini memastikan setiap kesalahan deteksi dapat dianalisis dalam konteks visual yang sesuai, sehingga penyebabnya dapat ditelusuri secara akurat. Dengan pengorganisasian data yang sistematis, tim *Quality Assurance* dapat menelusuri pola kesalahan yang berulang sekaligus mempersiapkan data yang relevan untuk tahap validasi dan analisis akar penyebab.

3.3.3 Visualisasi Kesalahan Deteksi Model

Langkah berikutnya adalah melakukan visualisasi kesalahan deteksi dengan menggunakan model yang telah di-*deploy*. Visualisasi ini bertujuan untuk memahami konteks di balik kesalahan deteksi, karena hasil yang ditampilkan pada UI hanya berupa keluaran akhir tanpa menunjukkan proses inferensi secara menyeluruh. Oleh karena itu, model yang digunakan di lingkungan produksi diambil dan dijalankan ulang secara lokal untuk mengamati hasil inferensi secara langsung terhadap rekaman yang sama. Proses visualisasi ini dilakukan dengan memanfaatkan *pipeline* YOLO yang diadaptasi melalui modul Python dan berkas konfigurasi YAML yang telah disesuaikan agar dapat mensimulasikan deteksi pada produksi.

Salah satu contoh visualisasi ditunjukkan pada Gambar 3.1, yang menampilkan keluaran deteksi person beserta nilai *confidence score*. Visualisasi seperti ini membantu tim *Quality Assurance* melihat pola kesalahan secara lebih jelas, termasuk bagaimana model memberikan skor keyakinan pada objek yang terdeteksi, serta apakah terdapat fluktuasi deteksi antar-*frame* yang tidak terlihat pada UI.



Gambar 3.1. Contoh visualisasi hasil inferensi model secara lokal, menampilkan deteksi *person* beserta nilai *confidence score*. [1]

Melalui pendekatan ini, tim dapat menilai bagaimana model mendeteksi objek dalam kondisi sebenarnya serta melihat perbedaan antara inferensi di lingkungan lokal dan sistem produksi. Meski terdapat perbedaan kecil dalam parameter teknis seperti kecepatan pemrosesan, *frame per second*, maupun konfigurasi kamera, proses ini memberikan wawasan penting mengenai karakteristik kesalahan yang terjadi. Contohnya seperti deteksi yang tidak stabil, kehilangan objek tertentu, atau perbedaan nilai ambang (*confidence threshold*) antar lingkup inferensi.

3.3.4 Validasi Hasil Deteksi dan Lingkungan Produksi

Setelah proses visualisasi, tahap selanjutnya adalah validasi hasil deteksi untuk memastikan kesesuaian antara hasil inferensi model dan kondisi nyata di lapangan. Validasi ini dilakukan secara manual oleh tim *Quality Assurance* dengan membandingkan hasil deteksi terhadap data referensi (*ground truth*) maupun observasi langsung. Dalam beberapa kasus, personel lapangan turut dilibatkan untuk mengonfirmasi apakah hasil deteksi sesuai dengan kondisi aktual di area produksi.

Proses ini tidak hanya digunakan untuk menilai kebenaran hasil deteksi, tetapi juga untuk membedakan apakah kesalahan terjadi akibat model YOLO atau konfigurasi aturan (*object rule*) pada sistem produksi. Apabila model menunjukkan hasil yang benar di lingkungan lokal namun keliru di sistem produksi, maka permasalahan kemungkinan besar berasal dari konfigurasi aturan yang tidak selaras dengan keluaran model. Dengan demikian, tahap validasi ini berperan penting

dalam memastikan bahwa penurunan performa dapat diidentifikasi secara tepat dan tidak salah atribusi.

3.3.5 Analisis Performa Model Deteksi

Setelah pengumpulan, visualisasi, dan validasi hasil deteksi, tahap berikutnya adalah analisis performa model secara kuantitatif. Evaluasi ini bertujuan untuk memastikan bahwa model tetap bekerja dengan akurat dan konsisten di lingkungan produksi. Metrik yang digunakan mencakup *precision*, *recall*, dan *F1-score*, yang merepresentasikan keseimbangan antara ketepatan dan kelengkapan deteksi. Metrik-metrik ini juga menjadi indikator utama untuk memantau kualitas model pasca-*deployment* dan mendeteksi penurunan performa secara dini.

Data performa dikumpulkan secara rutin dari hasil inferensi model yang telah di-*deploy*, kemudian dianalisis per kategori deteksi, misalnya orang, kendaraan, dan *Personal Protective Equipment* (PPE). Perbandingan antar versi model membantu menilai apakah ada peningkatan atau penurunan performa dari iterasi pelatihan terbaru.

Untuk menelusuri penyebab penurunan performa, digunakan metode *Root Cause Analysis* (RCA) dengan pendekatan 5 *Why*. Menurut Ateng et al., metode ini mempermudah pencarian akar penyebab secara mendalam dengan mengulang pertanyaan "mengapa" sebanyak lima kali [6]. Dengan cara ini, setiap gejala performa yang menurun dapat ditelusuri hingga penyebab dasarnya, sehingga perbaikan yang dilakukan lebih tepat sasaran.

Sebagai contoh, apabila performa menurun, analisis 5 *Why* dapat dilakukan sebagai berikut.

1. Mengapa performa sistem menurun?
Karena sistem menghasilkan lebih banyak kesalahan deteksi dibanding sebelumnya, baik berupa *false positive* maupun *false negative*.
2. Mengapa *false positive* dan *false negative* semakin banyak?
Karena model gagal mengenali sebagian objek dan sebaliknya mendeteksi objek yang tidak relevan.
3. Mengapa model gagal mengenali atau mendeteksi objek yang tidak relevan?
Karena terdapat pola atau variasi objek baru di lapangan yang tidak tercakup dalam data pelatihan, atau aturan deteksi (*rule*) berubah akibat kondisi operasional yang dinamis.

4. Mengapa perubahan data atau aturan tidak terakomodasi dalam model?

Karena proses pembaruan model dan *pipeline* dilakukan tanpa evaluasi ulang yang terstandarisasi, serta dokumentasi versi model belum lengkap.

5. Mengapa evaluasi ulang dan dokumentasi belum berjalan konsisten?

Karena belum terdapat mekanisme peningkatan kualitas model yang memastikan data baru dan perubahan konteks operasional masuk kembali ke proses pengembangan.

Pendekatan ini membantu tim menentukan langkah perbaikan yang tepat, baik berupa pembaruan model, penyesuaian aturan, maupun pengayaan dataset, sehingga performa model tetap optimal.

A Penurunan Performa pada Model Deteksi

Penurunan performa pada model deteksi dapat muncul secara bertahap seiring perubahan kondisi data, lingkungan, aturan, maupun versi model yang digunakan. Berdasarkan hasil pengamatan dan analisis QA, degradasi performa biasanya disebabkan oleh tiga faktor utama, yaitu data yang tidak konsisten, kesalahan konfigurasi model, dan perubahan konteks visual di lapangan.

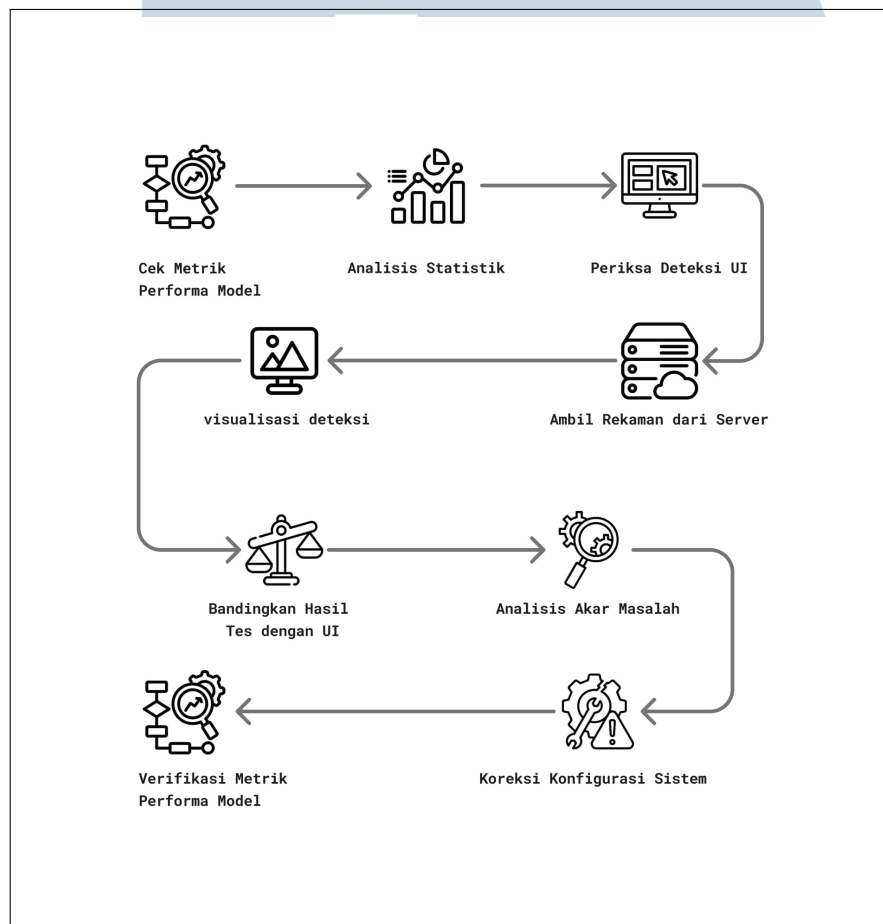
Dari sisi data, penurunan akurasi sering kali berkaitan dengan kualitas anotasi yang tidak seragam atau kurang representatif terhadap kondisi nyata. Model yang dilatih dengan data terbatas atau bias terhadap sudut pandang tertentu cenderung gagal mendeteksi objek baru dengan bentuk atau posisi yang berbeda. Sementara itu, dari sisi konfigurasi model, ketidaktepatan pengaturan ambang deteksi (*threshold*) atau ketidaksesuaian kelas pada aturan (*object rule*) dapat menyebabkan model salah mengklasifikasikan objek.

Selain faktor teknis, penurunan performa juga dipengaruhi oleh perubahan lingkungan operasional, seperti variasi pencahayaan, kondisi kamera, atau munculnya objek baru di area deteksi yang belum pernah dikenali model sebelumnya. Faktor-faktor ini menyebabkan model yang sebelumnya bekerja stabil menjadi kurang akurat dalam mengenali konteks tertentu. Oleh karena itu, diperlukan pendekatan evaluatif dan iteratif yang berkelanjutan untuk mendeteksi, menganalisis, dan memperbaiki sumber penurunan performa secara sistematis.

Sebagai bentuk penerapan dari analisis tersebut, dilakukan serangkaian studi kasus untuk menelusuri penyebab spesifik penurunan performa pada beberapa kategori deteksi di lingkungan produksi, yaitu *housekeeping*, boomlift, dan forklift.

A.1 Kasus Housekeeping: Kesalahan Konfigurasi Kelas

Proses identifikasi permasalahan pada kelas *housekeeping* dilakukan melalui serangkaian evaluasi sistematis. Alur investigasi divisualisasikan pada Gambar 3.2, yang mencakup langkah pemeriksaan metrik model, analisis ketidakseimbangan hasil deteksi, validasi melalui UI, perbandingan dengan rekaman lapangan, hingga verifikasi lokal dan penelusuran akar permasalahan pada konfigurasi sistem.

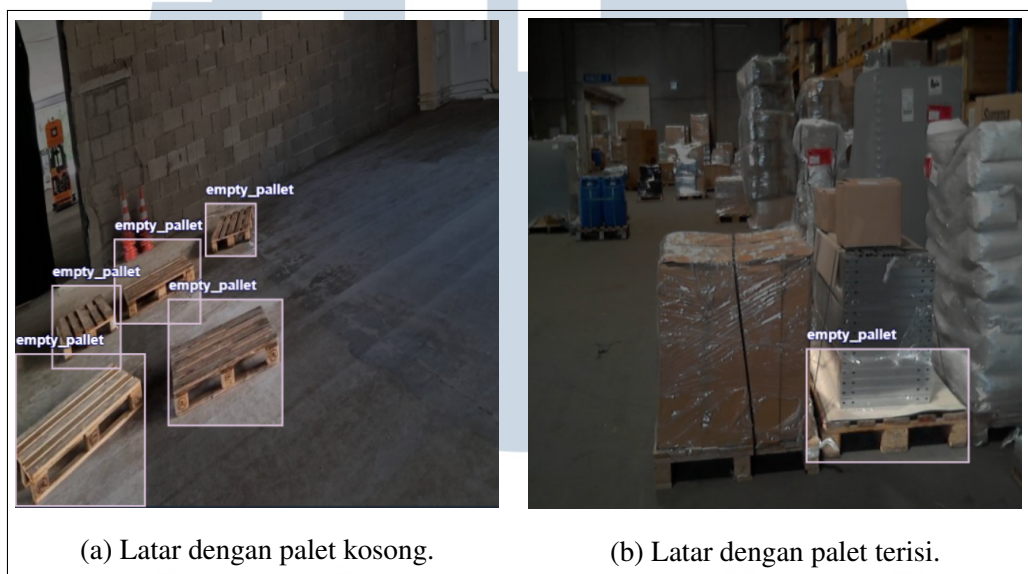


Gambar 3.2. Alur penanganan kasus *housekeeping*.

Permasalahan pada model *housekeeping* pertama kali teridentifikasi dari hasil analisis kualitas deteksi yang menunjukkan ketidakseimbangan signifikan antara nilai *precision* dan *recall*. Pada evaluasi awal setelah pembaruan versi model, *precision* berada pada kisaran 98%, sedangkan *recall* hanya sekitar 30% berdasarkan perbandingan antara objek yang muncul pada antarmuka pengguna (UI) dan objek yang sebenarnya terlihat pada rekaman lapangan. Ketidakseimbangan ini menghasilkan nilai *F1-score* sekitar 0.45, yang

mengindikasikan adanya permasalahan dalam proses pemetaan hasil model ke sistem produksi.

Contoh deteksi pada UI ditunjukkan pada Gambar 3.3. Gambar 3.3 (a) menampilkan contoh objek *empty_pallet* yang muncul pada UI secara benar dan responsif, sementara Gambar 3.3 (b) menunjukkan *loaded_pallet* yang salah diklasifikasikan sebagai *empty_pallet* dan beberapa lainnya terlewat pada UI. Kedua kondisi tersebut memperlihatkan bahwa hasil yang tampil pada UI memiliki bias untuk lebih responsif pada satu *class* dibanding yang lain.



Gambar 3.3. Contoh hasil deteksi pada UI untuk kategori *housekeeping*.

Untuk memastikan apakah ketidaksesuaian berasal dari model atau dari sistem produksi, evaluasi tambahan dilakukan menggunakan pengujian lokal. Hasil pengujian lokal menunjukkan bahwa objek *loaded_pallet* terdeteksi dengan benar dan diberi nilai *confidence* tinggi sebagaimana ditunjukkan pada Gambar 3.4. Temuan ini mengindikasikan bahwa kemampuan model YOLO berada dalam kondisi baik dan tidak menunjukkan gejala *degradation*.



Gambar 3.4. Deteksi *loaded_pallet* pada pengujian lokal.

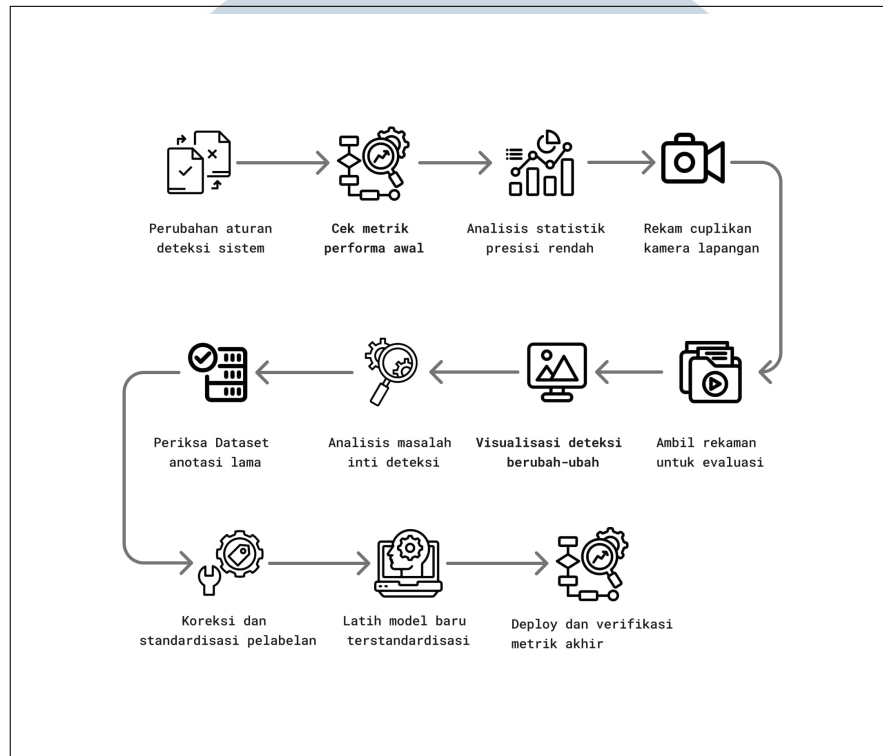
Pemeriksaan lebih lanjut terhadap sistem mengungkap bahwa sumber permasalahan terletak pada konfigurasi kelas di tingkat aturan (*rule level*) dalam *pipeline* deteksi. Meskipun model memiliki dua kelas berbeda, yaitu *empty_pallet* dan *loaded_pallet*, konfigurasi sistem hanya memasukkan kelas *empty_pallet*. Akibatnya, seluruh hasil deteksi untuk *loaded_pallet* tidak pernah diproses atau ditampilkan pada UI, sehingga menghasilkan *recall* rendah dan sejumlah kesalahan klasifikasi.

Setelah konfigurasi diperbaiki dan *pipeline inference* dijalankan ulang, sistem menunjukkan peningkatan performa yang signifikan. Nilai *recall* meningkat menjadi 100%, sementara *precision* sedikit menurun menjadi sekitar 89% karena beberapa objek dengan bentuk mirip *loaded_pallet* ikut terdeteksi. Meskipun demikian, nilai *F1-score* meningkat drastis dari 0.45 menjadi 0.94, menandakan bahwa keseimbangan antara ketepatan dan kelengkapan deteksi telah tercapai. Sistem kemudian kembali beroperasi secara konsisten dan sesuai dengan kondisi lapangan.

A.2 Kasus Boomlift: Kesalahan Anotasi dan Ketidaksesuaian Rule

Proses investigasi terhadap permasalahan pada kategori boomlift dilakukan secara bertahap untuk memastikan sumber ketidaksesuaian hasil deteksi. Alur analisis ditunjukkan pada Gambar 3.5, yang mencakup tahapan mulai dari adanya

perubahan aturan deteksi, pemeriksaan performa setelah perubahan, peninjauan cuplikan kamera lapangan, analisis dataset, hingga proses standarisasi anotasi, dan akhirnya pembuatan serta *deployment* dari model baru.



Gambar 3.5. Alur penanganan kasus boomlift.

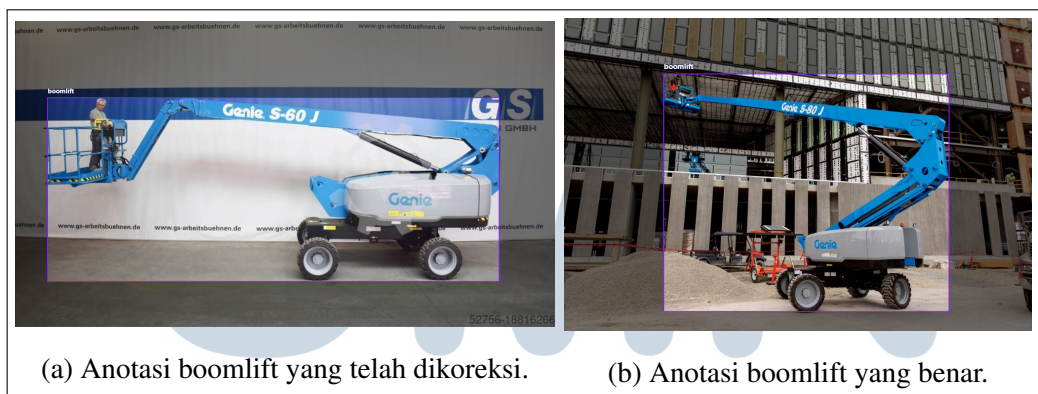
Permasalahan pada kategori boomlift mulai terlihat setelah diterapkannya perubahan definisi aturan deteksi (*rule*) yang mengharuskan sistem hanya menampilkan boomlift yang sedang digunakan di lokasi kerja, sementara boomlift yang terparkir atau tidak aktif harus diabaikan. Meskipun perubahan tersebut bertujuan meningkatkan relevansi deteksi, penerapannya justru menimbulkan inkonsistensi karena boomlift terparkir tetap terdeteksi sebagai aktif. Kondisi ini menyebabkan peningkatan jumlah *false positive* dan menurunkan nilai *precision* hingga sekitar 48%.

Evaluasi silang terhadap kelas dan aturan lain yang sudah stabil kemudian dilakukan untuk memastikan bahwa anomali hanya terjadi pada kategori boomlift. Peninjauan rekaman kamera menunjukkan bahwa *bounding box* yang dihasilkan model berubah-ubah meskipun objek tidak bergerak, sehingga boomlift terparkir sering terbaca sebagai sedang digunakan.



Gambar 3.6. Dataset yang digunakan sebelum perubahan anotasi. [2, 3]

Penyelidikan lanjutan terhadap dataset pelatihan mengungkap bahwa sumber utama permasalahan berasal dari ketidakkonsistenan anotasi. Sebagian data diberi label hanya pada bagian dasar boomlift, sementara sebagian lainnya mencakup keseluruhan struktur. Variasi bentuk anotasi ini ditunjukkan pada Gambar 3.6. Ketidakteraturan tersebut membuat model kesulitan mengenali representasi objek secara konsisten dan berkontribusi terhadap tingginya *false positive*.



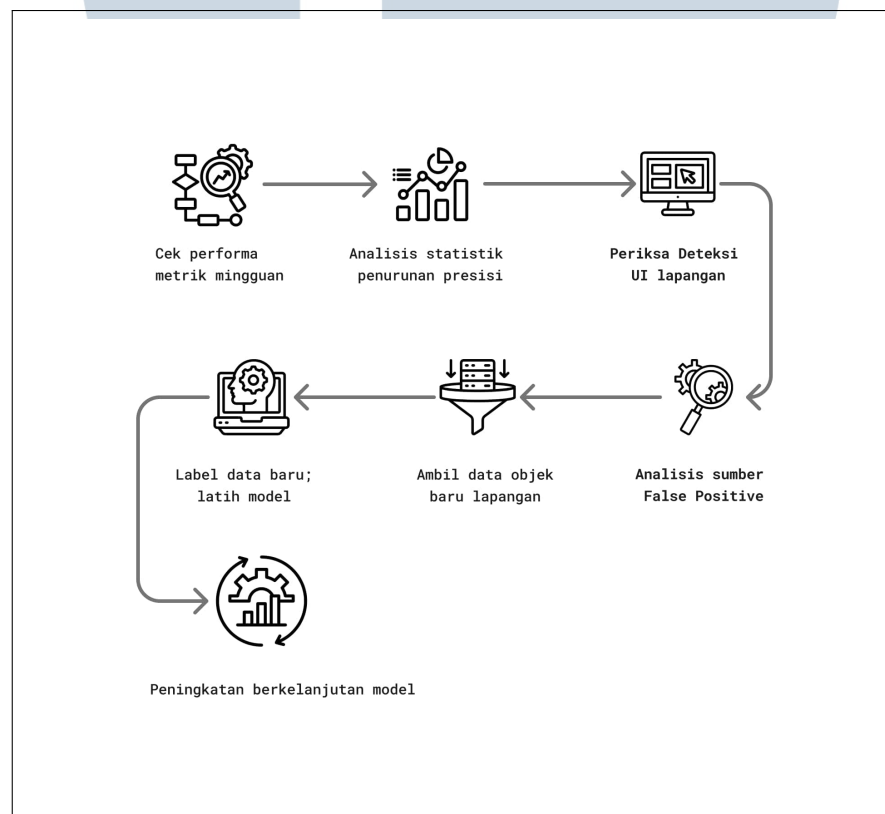
Gambar 3.7. Dataset yang digunakan setelah standardisasi anotasi kelas.[2, 3]

Setelah seluruh anotasi dikoreksi dan distandardisasi agar setiap boomlift diberi label secara seragam mencakup keseluruhan struktur, model dilatih ulang dan diuji kembali pada berbagai kondisi, termasuk variasi pencahayaan, posisi kamera, dan jumlah objek pada satu adegan. Hasil pengujian menunjukkan peningkatan signifikan, dengan nilai *precision* meningkat dari sekitar 48% menjadi 90%. Sisa kesalahan hanya muncul pada kasus khusus seperti objek yang tertutup atau

terhalang kendaraan, namun tidak memberikan dampak berarti terhadap performa keseluruhan.

A.3 Kasus Forklift: Adaptasi terhadap Objek Baru

Proses identifikasi permasalahan pada kategori forklift dilakukan melalui pemantauan performa mingguan, analisis statistik, serta pemeriksaan langsung terhadap hasil deteksi di lapangan. Gambar 3.8 memperlihatkan alur analisis yang meliputi evaluasi metrik awal, penelusuran sumber *false positive*, pengumpulan data objek baru, hingga pembaruan model untuk memastikan kemampuan adaptasi terhadap perubahan kondisi operasional.



Gambar 3.8. Alur analisis kasus forklift.

Pada kondisi awal, model forklift menunjukkan performa tinggi dengan nilai *precision* sekitar 95% dan *recall* 100%. Namun, pemantauan mingguan mengungkapkan adanya penurunan bertahap pada *precision*, yaitu menjadi 85% pada minggu pertama, 81% pada minggu kedua, dan turun hingga 60% pada bulan berikutnya, sementara *recall* tetap stabil. Pola ini mengindikasikan bahwa model

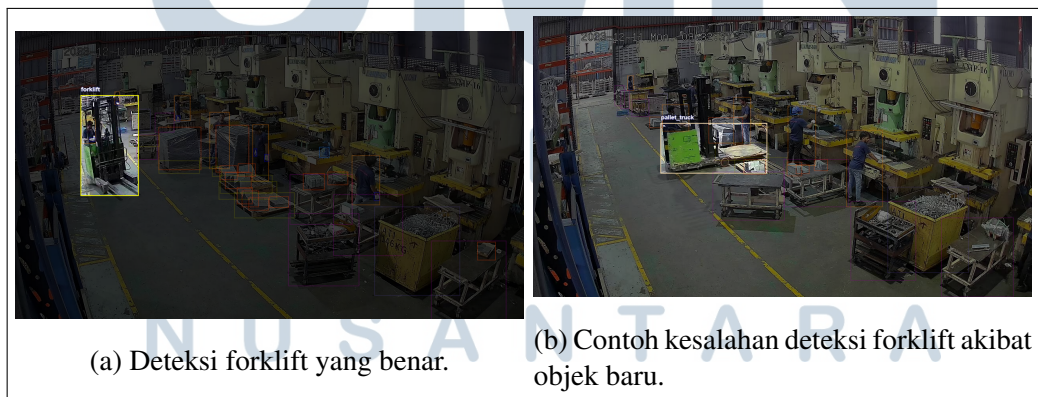
mengalami peningkatan *false positive* tanpa kehilangan kemampuan mendeteksi forklift yang sebenarnya.

Pemeriksaan hasil deteksi pada antarmuka pengguna (UI) menunjukkan bahwa penurunan presisi dipicu oleh kemunculan objek baru di area operasi yang secara visual mirip dengan forklift. Salah satu objek tersebut adalah *pallet truck*, yang sebelumnya tidak terdapat dalam dataset pelatihan. Contoh objek tersebut diperlihatkan pada Gambar 3.9 dan menjadi sumber dominan terjadinya kesalahan deteksi.



Gambar 3.9. Contoh *pallet truck* dalam dataset. [4]

Untuk menganalisis dampaknya, dilakukan perbandingan hasil deteksi sebelum dan sesudah kemunculan objek baru. Gambar 3.10 menampilkan perbedaan tersebut, dengan bagian (a) menunjukkan deteksi forklift yang benar dan bagian (b) menunjukkan contoh kesalahan deteksi yang muncul setelah *pallet truck* mulai beroperasi di area produksi.



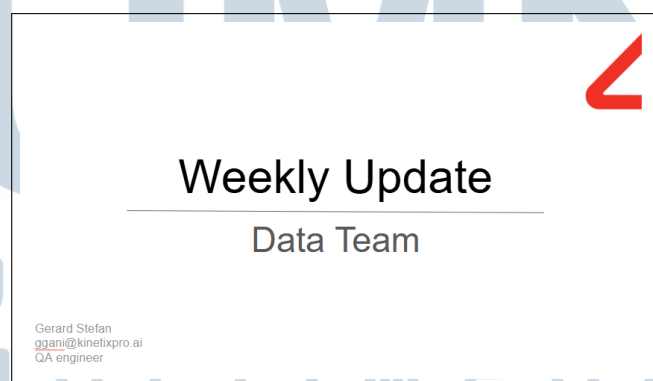
Gambar 3.10. Perbandingan hasil deteksi forklift sebelum dan sesudah kemunculan objek baru.

Seluruh hasil deteksi yang salah kemudian dikumpulkan dan ditambahkan ke dataset pelatihan untuk melakukan pembaruan model. Objek baru yang mirip forklift namun bukan forklift dimasukkan ke dalam kelas *unknown* agar tidak tercampur dengan kelas forklift. Proses pelabelan dan penambahan data baru tersebut disertai pelatihan ulang model secara berkala sebagai bagian dari strategi peningkatan berkelanjutan.

Setelah model diperbarui dengan data terbaru, nilai *precision* meningkat kembali menjadi sekitar 90%, sementara *recall* tetap stabil. Meskipun peningkatan ini signifikan, proses pembaruan model dan penambahan data secara berkelanjutan tetap diperlukan untuk menjaga ketahanan model terhadap perubahan kondisi operasional di lapangan.

3.3.6 Dokumentasi dan Pelaporan Hasil QA

Setiap hasil analisis dan perbaikan yang dilakukan didokumentasikan secara terstruktur dalam bentuk laporan QA. Dokumentasi ini mencakup catatan kesalahan deteksi, tangkapan visual, ringkasan metrik performa, serta rekomendasi tindak lanjut yang perlu dilakukan oleh tim pengembang. Proses dokumentasi tersebut tidak hanya berfungsi sebagai pelaporan, tetapi juga sebagai media komunikasi lintas tim. Setiap temuan penting dirangkum dalam bentuk presentasi internal menggunakan *Google Slides* atau *PowerPoint*. Contoh halaman sampul untuk pelaporan rapat ditunjukkan pada Gambar 3.11.



Gambar 3.11. Contoh halaman sampul laporan dalam presentasi internal.

Dokumentasi ini kemudian disampaikan pada sesi *sharing* rutin agar setiap tim terkait memahami konteks dan implikasi dari temuan tersebut. Selain menopang proses komunikasi, dokumentasi QA juga berfungsi sebagai arsip teknis yang dapat

dijadikan dasar evaluasi performa model dan referensi untuk iterasi pengembangan pada masa mendatang.

3.3.7 Kendala yang Dihadapi

Selama upaya peningkatan kualitas deteksi berbasis YOLO di KinetixPro Pte. Ltd., ditemukan beberapa kendala sebagai berikut.

1. Model yang telah diperbaiki melalui pengayaan data dan iterasi tambahan tetap menunjukkan penurunan performa setelah beberapa waktu, meskipun metrik *precision* dan *recall* sempat meningkat setelah pembaruan.
2. Kesalahan pelabelan (*mislabeled*) pada objek baru yang ditambahkan oleh tim anotator eksternal, sehingga beberapa objek terdeteksi secara tidak akurat atau model cenderung keliru dalam klasifikasi.
3. Ketidakkonsistenan dalam penandaan bounding box oleh tim anotator eksternal, terutama karena ukuran dan posisi bounding box untuk objek yang sama tidak seragam antar variasi objek. Perbedaan ini menyebabkan model mempelajari representasi objek yang kurang stabil, sehingga menurunkan kualitas deteksi secara keseluruhan.

3.3.8 Solusi yang Diterapkan

Untuk mengatasi kendala-kendala tersebut, langkah-langkah berikut diterapkan.

1. Melanjutkan pengayaan dataset dengan menambahkan data baru yang relevan berdasarkan *root cause* dari penurunan performa model, sehingga model dapat belajar dari variasi yang lebih lengkap.
2. Melakukan pengecekan dan koreksi *mislabeled* yang berasal dari tim anotator eksternal secara rutin untuk memastikan anotasi sesuai ketentuan dan akurat untuk kebutuhan pelatihan model.
3. Menetapkan dan mensosialisasikan standar *bounding box* yang konsisten, termasuk pedoman ukuran, posisi, dan cakupan objek, sehingga penandaan objek tidak bervariasi antar anotator maupun antar *instance* objek yang memiliki variasi bentuk atau orientasi.