

Permasalahan yang dialami selama mengerjakan challenge dan solusinya

Dalam proses pengumpulan dan anotasi dataset pada Challenge 3, tantangan utama yang dihadapi adalah ketidaksesuaian antara dataset open-source yang tersedia dengan skema labeling yang telah dirancang. Dataset SH17 awalnya dipertimbangkan karena sifatnya yang terbuka dan luas. Namun, setelah ditelusuri lebih dalam, dataset ini ternyata memiliki banyak sekali kelas yang tidak relevan dengan kebutuhan sistem pendeteksi PPE yang difokuskan hanya pada enam label utama.

Akibatnya, anotasi yang sudah tersedia dalam dataset SH17 sebagian besar tidak dapat digunakan secara langsung. Gambar-gambar yang relevan perlu dipilah satu per satu secara manual, dan anotasinya harus dibuat ulang agar sesuai dengan skema labeling yang telah ditetapkan. Proses ini tentu memakan waktu dan tenaga tambahan.

Selain itu, banyak gambar dalam dataset SH17 yang tidak menunjukkan situasi di lingkungan industri sebagaimana yang dibutuhkan dalam challenge ini. Untuk mengatasi keterbatasan tersebut, dilakukan pencarian data tambahan secara manual dari sumber lain, terutama melalui Pinterest. Kata kunci yang digunakan antara lain: “person wear gloves,” “person wear hard hat,” “person wear safety glasses,” “person wear safety vest,” dan “person wear steel toe boots.”

Gambar-gambar hasil pencarian tersebut diseleksi berdasarkan kesesuaiannya dengan konteks industri. Setelah itu, seluruh gambar dianotasi secara manual menggunakan tools Roboflow. Meskipun memerlukan upaya ekstra, pendekatan ini memastikan bahwa dataset yang dihasilkan benar-benar relevan dan sesuai dengan kebutuhan sistem pendeteksi pelanggaran PPE yang dikembangkan.