

Deskripsi Dataset

1. Tujuan Dataset

Dataset ini disusun untuk mendukung pengembangan sistem deteksi pelanggaran penggunaan alat pelindung diri (Personal Protective Equipment / PPE) berbasis AI. Gambar-gambar yang dipilih mencerminkan skema pelabelan yang telah ditetapkan sebelumnya, yaitu mendeteksi keberadaan objek berikut: person, hard_hat, safety_vest, gloves, safety_glasses, dan steel_toe_boots.

2. Sumber Gambar

Gambar dalam dataset ini dikumpulkan dari berbagai sumber terbuka yang relevan:

- **SH17 Dataset**
Dataset publik yang mengandung banyak gambar pekerja di lingkungan kerja, namun tidak seluruhnya sesuai dengan skema pelabelan PPE yang ditetapkan. Oleh karena itu, hanya sebagian gambar yang dipilih dan dianotasi ulang secara manual.
- **Pinterest**
Pencarian dilakukan menggunakan kata kunci spesifik seperti “person wearing hard hat”, “person with safety vest”, dan lainnya, untuk mendapatkan gambar real-world yang sesuai dengan konteks industri.
- **Roboflow Universe**
Beberapa gambar tambahan yang sesuai dengan label dan konteks juga diambil dari proyek open-source Roboflow Universe. Gambar-gambar ini diperiksa ulang dan dianotasi kembali sesuai dengan format dan kebutuhan sistem deteksi PPE.

3. Distribusi Gambar

Sumber	Jumlah Gambar
SH17 Dataset	9
Pinterest	36
Roboflow Universe	5
Total	50

4. Proses Anotasi

Semua gambar yang terpilih dianotasi ulang secara manual menggunakan platform Roboflow Annotation Tool. Proses anotasi mengikuti panduan format YOLO, dengan label yang konsisten terhadap skema labeling yang ditetapkan di Challenge 1. Anotasi dilakukan hanya pada objek PPE dan person yang terlihat dengan jelas, menghindari noise dari objek yang buram atau tidak relevan.