

Challenge 2: Dataset Collection

Latar Belakang

PT. Qwerty sedang merancang sistem AI berbasis Computer Vision untuk mendeteksi pelanggaran penggunaan alat pelindung diri (PPE) melalui rekaman CCTV. Untuk membangun model AI yang efektif dan akurat, dibutuhkan dataset gambar yang relevan dan representatif.

Challenge 2 ini berfokus pada tahap awal dalam pengumpulan data, yaitu mengumpulkan gambar yang menggambarkan personil di lingkungan kerja yang mengenakan atau tidak mengenakan PPE, sebagai dasar pelatihan sistem AI.

Kriteria Pemilihan Gambar

Gambar dipilih berdasarkan asumsi kebutuhan sistem, dengan mempertimbangkan hal-hal berikut:

- Jenis PPE yang terlihat: minimal salah satu dari 5 jenis berikut:
 - Hard Hat
 - Safety Vest
 - Gloves
 - Safety Glasses
 - Steel-toe Boots
- Variasi kondisi: indoor/outdoor, posisi kamera, jumlah orang, angle tubuh.
- Kualitas visual: cukup tajam untuk anotasi, tidak terlalu blur.
- Keberagaman: mencakup pekerja yang:
 - Mengenakan PPE secara lengkap
 - Mengenakan sebagian PPE

Sumber Dataset

Dataset yang digunakan dikumpulkan dari berbagai sumber legal dan terbuka, yaitu:

No	Sumber	Keterangan
1	SH17 Dataset for PPE Detection	Dataset resmi dari publikasi jurnal <i>ScienceDirect</i> (2024), berisi ribuan gambar pekerja konstruksi dan industri dengan anotasi PPE.
2	Google Images	Gambar pekerja konstruksi dengan berbagai jenis PPE.

Rincian Dataset

- Total gambar terkumpul: 40 gambar
- Format gambar: .jpg dan .jpeg
- Nama file: image1 hingga image40

Jenis PPE yang Tercover dalam Dataset

Jenis PPE	Gambar
Hard Hat	image1-image40
Safety Vest	image2-image13
Gloves	image14-image37
Safety Glasses	image1-image40
Steel-toe Boots	image9-image35

Alasan Pemilihan Dataset

Dataset ini dipilih karena:

1. Relevan langsung dengan sistem yang akan dibangun (deteksi PPE).
2. Beragam secara visual, mencerminkan kondisi nyata di lapangan.
3. Cukup tajam dan jelas untuk digunakan dalam anotasi bounding box.

Potensi Penggunaan

Dataset ini akan digunakan pada tahap berikutnya, yaitu:

1. Proses anotasi bounding box untuk tiap jenis PPE
2. Training dan evaluasi model object detection (misalnya YOLOv5, Faster R-CNN)
3. Pembangunan sistem deteksi pelanggaran PPE secara otomatis

Penutup

Pengumpulan gambar dalam challenge ini merupakan fondasi penting untuk sistem deteksi pelanggaran PPE. Dengan gambar yang dipilih secara strategis dan representatif, diharapkan proses anotasi dan pelatihan model AI dapat dilakukan secara efisien dan menghasilkan sistem yang akurat serta dapat diandalkan di lingkungan kerja nyata.