

PPE Detection and Assessment of Correct Wearability

Devyansh Choudhary, Mannan Gupta, Mohak Agarwala

The Problem

- In 2021, **more than 162 workers** died in manufacturing industries.
- Between 2017 and 2020, the **average number of deaths** in registered factories was **1,109 per year**.
- **Personal Protective Equipment (PPE)** is **essential for worker safety**, especially in hazardous environments like construction sites, factories, and industrial settings.
- Non-compliance with proper safety equipment is **very common in India**.
- **Inspired by construction at Plaksha University.**

Existing Solutions

- Studies on detection of non-compliance of safety equipment have been extensively done.
- Researchers have used YOLO (you only look once), deep learning type approaches to detect whether the workers in the image are compliant with the safety protocols.

The Gap

- Existing solutions don't take into account the proper wearability of the equipment.
- For eg. if in an image there is a worker and they hold a helmet in their hand, it will be detected as correct but it is wrong as the helmet should go onto the head.

Our Solution

- We will bridge this gap by making a model which will check if the person is wearing the equipment properly in addition to detecting the safety equipment present in the image.
- The solution enhances workplace safety by ensuring not only the detection but also the correct usage of PPE, reducing the risk of injuries and improving compliance.



Existing Models:



Our Model:



Literature Survey

<u>S.No</u>	Title	Year	Journal/Conference	Methods	Accuracy
1	PPE Detector: A YOLO-Based Architecture to Detect Personal Protective Equipment (PPE) for Construction Sites	2022	<u>PeerJ</u> Computer Science	YOLOX (anchor-free)	86.4% <u>mAP @ 0.5</u>
2	Deep Learning Detection of Personal Protective Equipment to Maintain Safety Compliance on Construction Sites	2020	Construction Research Congress 2020	Deep learning-based image similarity	91% (hardhat F1-score), 84% (vest recall), 83% (overall compliance)
3	Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based Deep Learning Techniques	2020	Frontiers in Built Environment	YOLOv3 with transfer learning	79.1% <u>mAP</u>

Paper 1

Title: PPE Detector: A YOLO-Based Architecture to Detect Personal Protective Equipment (PPE) for Construction Sites

Source: [PeerJ](#)

Authors: Md. Ferdous, S.M. Ahsan, et al.

Classes: Colored hardhats (white, blue, red, yellow), safety vest, person body, person head, safety glasses.

Methodology:

- YOLOX (anchor-free) with Darknet53 backbone.
- Trained on CHVG dataset (1,699 images, 11,604 annotations) with mosaic augmentation and photometric changes.

Final Accuracy:

- mAP50-95: 44.7%
- mAP50: 66.3%
- Colored hardhats: 82.5%
- Safety vest: 76.8%
- Safety glasses: 58.2%

Limitations:

- Low accuracy for small/occluded objects (e.g., safety glasses).
- Performance drops in artificial rainy conditions (-15.2% mAP).
- Misclassification with yellow machinery parts.
- Doesn't check for correct wearability of the PPE.

Paper 2

Title: Deep Learning Detection of Personal Protective Equipment to Maintain Safety Compliance on Construction Sites

Source: [ASCE Library](#)

Authors: Pingbo Tang, J. Chen, et al.

PPE Classes: Hard hats, Safety vests, PPE colors

Methodology:

- Custom deep learning model with dynamic feature matching
- Compared query images with a gallery of PPE configurations
- Dataset: 3,309 in-house construction worker images

Accuracy:

- Hardhat (F1-score): 91%
- Vest (F1-score): 84%
- PPE color recognition: 77%
- Compliance classification: 83%

Limitations:

- Moderate color accuracy
- Issues with complex backgrounds, partial occlusions
- Works on static images only, not real-time
- Doesn't check for correct wearability of the PPE.

Paper 3

Title: Detection of PPE Compliance Using Deep Learning Techniques

Source: [ResearchGate](#)

Authors: V. Delhi, R. Sankarlal, T. Thomas

Compliance Classes: SAFE, NOT SAFE, NoHardHat, NoJacket

Methodology:

- YOLOv3 with transfer learning
- Dataset: 2,509 images from videos & web sources
- Integrated with real-time alarms & timestamped alerts

Accuracy:

- Overall mAP: 79.1%
- SAFE: 84.5%
- NOT SAFE: 73.7%
- NoHardHat: 81.2%
- NoJacket: 77.4%

Limitations:

- Only detects hardhats and jackets
- Biased by lighting and camera angles
- High computational load for real-time use
- Doesn't check for correct wearability of the PPE.

Dataset

- We are using the SH17 Dataset for PPE Detection
- Contains 8099 images
- Has 75,994 instances across 17 distinct classes, including both PPE items and relevant human body parts.
- **Classes:** Person, Head, Face, Glasses, Face-mask-medical, Face-guard, Ear, Earmuffs, Hands, Gloves, Foot, Shoes, Safety-vest, Tools, Helmet, Medical-suit, Safety-suit
- IDs- 0: person, 1: ear, 2: ear-mufs, 3: face, 4: faceguard, 5: face-mask-medical, 6: foot, 7: tool, 8: glasses, 9: gloves, 10: helmet, 11: hand, 12: head, 13: medical-suit, 14: shoes, 15: safety-suit, 16: safety-vest
- The images were sourced from Pexels, a platform that offers clear usage rights for all its content.
- No ethical concerns.



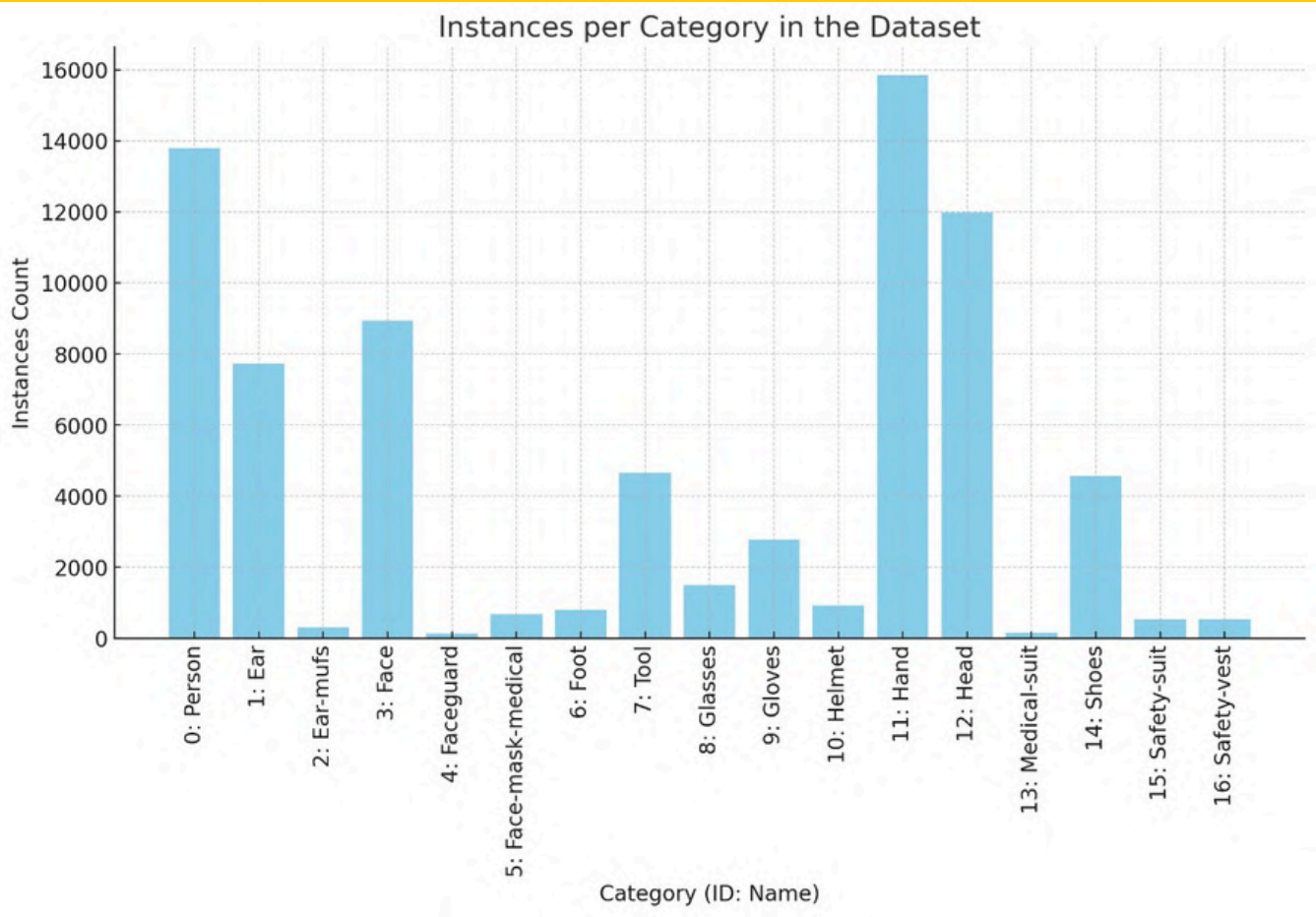
```
11 0.358417 0.177500 0.071167 0.219000
14 0.116667 0.811125 0.233000 0.364750
15 0.581750 0.524625 0.489167 0.355250
10 0.536750 0.393375 0.249833 0.274750
0 0.211750 0.488500 0.415833 0.976500
7 0.313833 0.727250 0.138333 0.545500
```

[class_id] [x_center] [y_center] [width] [height]

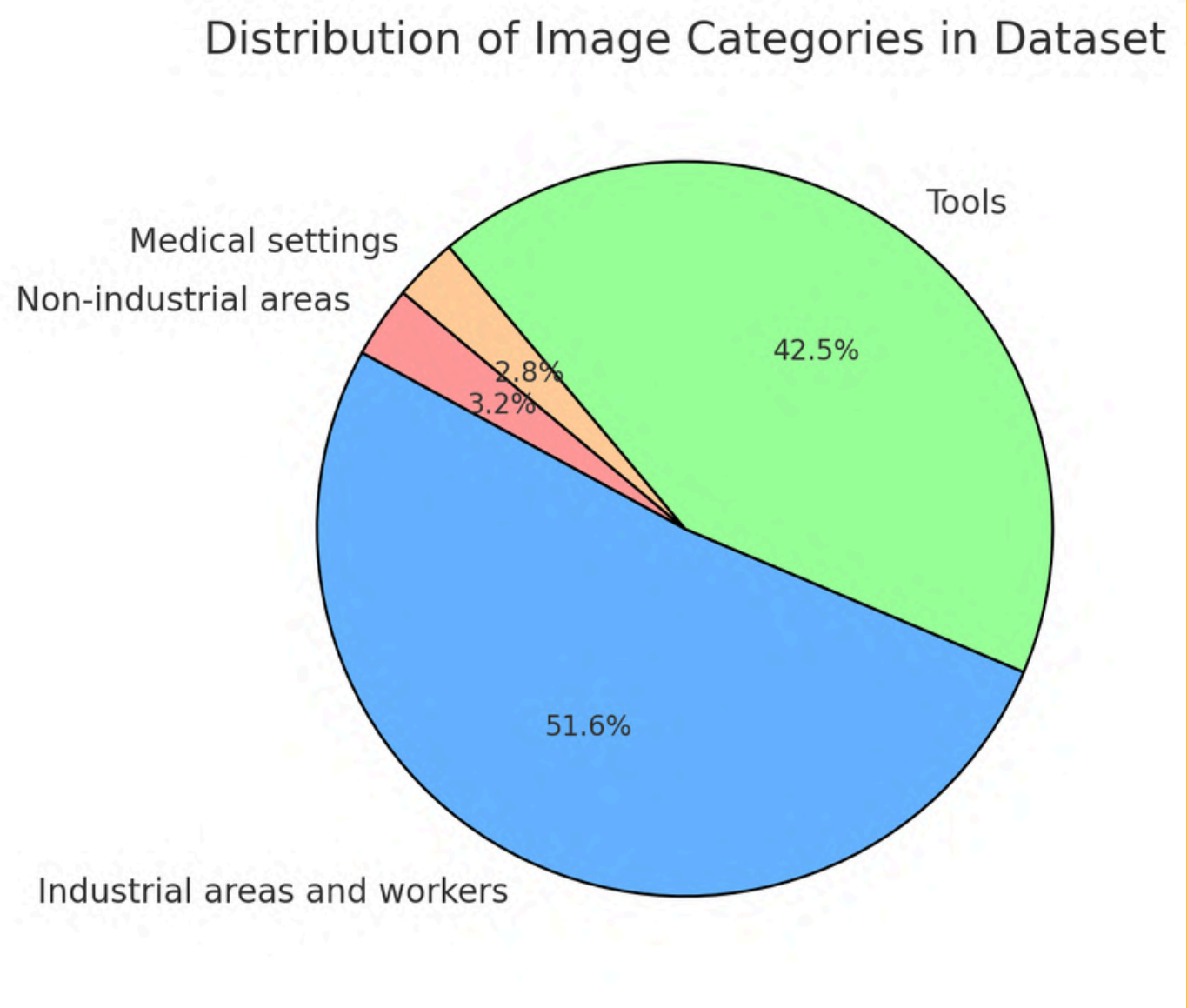
Features Preprocessing

- No missing features
- Dataset contains images of industrial areas, non-industrial areas, people working in industrial areas, industrial tools.
- Dataset also contains txt files for all images with bounding box coordinates for all the images for all the classes defined.
- We removed 2 features from the dataset, reducing the dimensionality.
- Removed features: 'face-mask-medical', 'medical-suit'

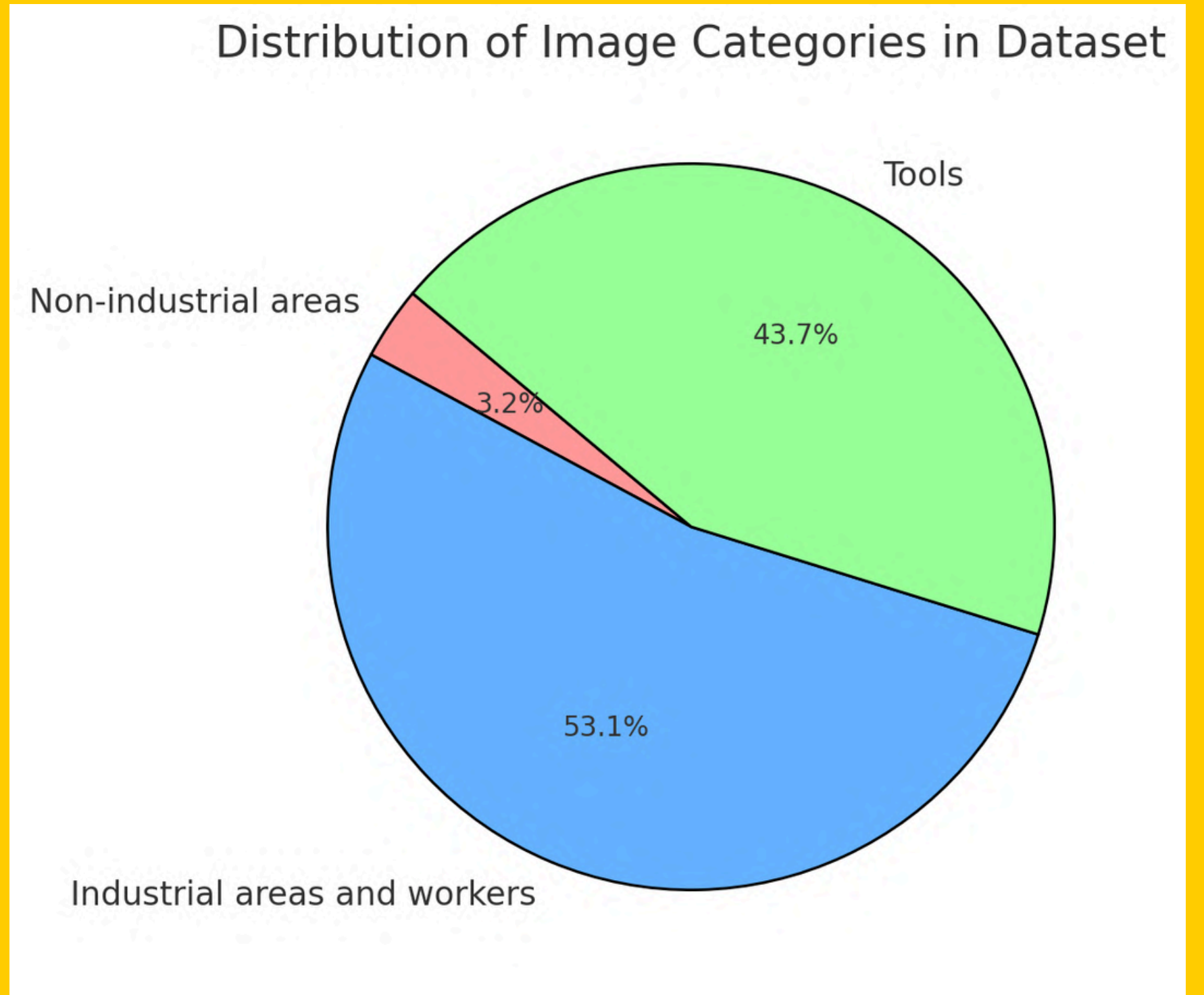
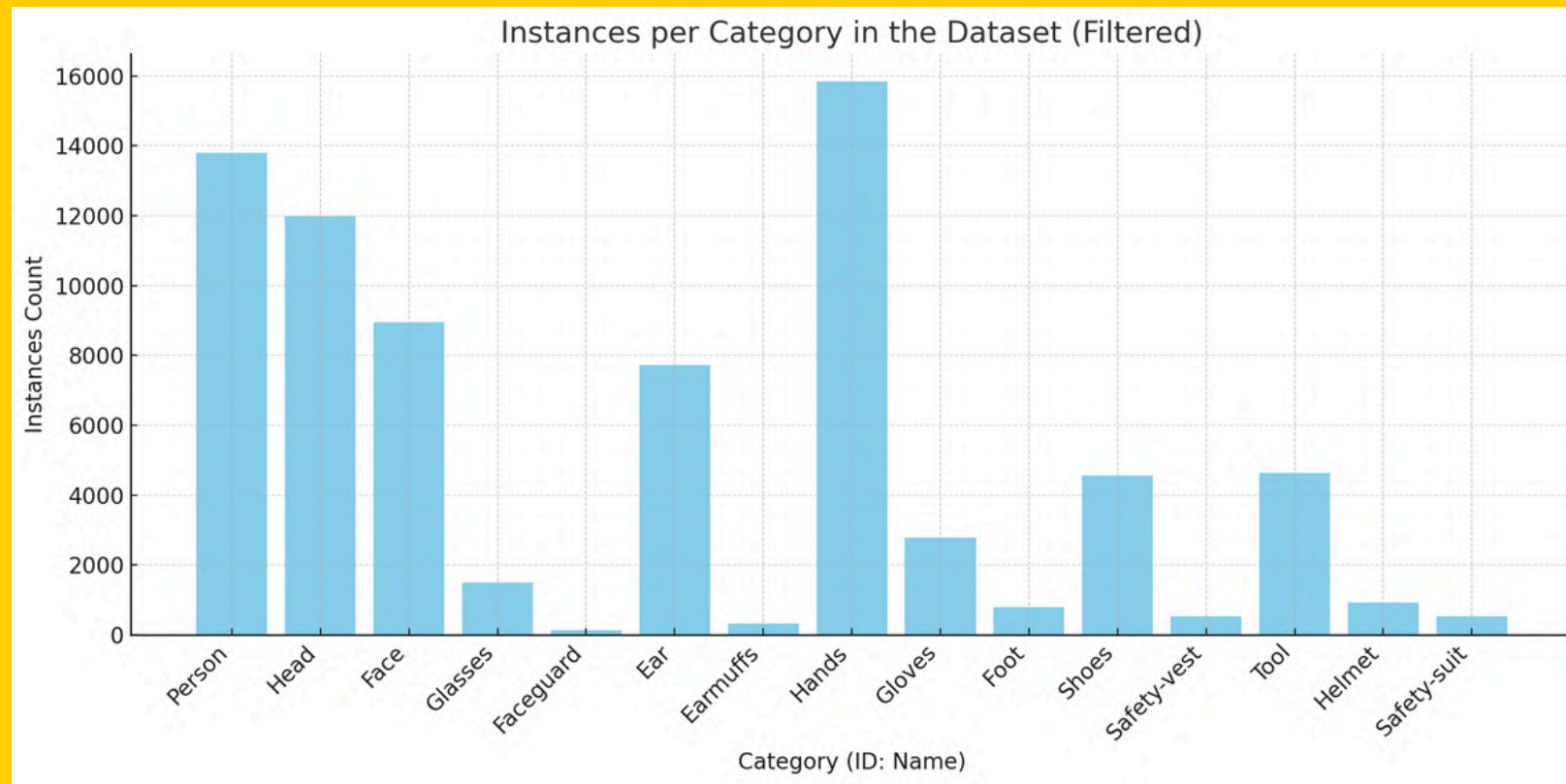
Dataset before removing medical images



ID	Name	Additional tags	Instances
1	Person	male, female, children	13802
2	Head	-	11985
3	Face	-	8950
4	Glasses	on, off, safety, vision	1945
5	Face-mask-medical	on, off	669
6	Face-guard	on, off	134
7	Ear		7730
8	Earmuffs	on, off	318
9	Hands		15850
10	Gloves	on, off	2790
11	Foot		796
12	Shoes	on, off, safety, other	4560
13	Safety-vest	on, off	530
14	Tools	on, off	4647
15	Helmet	on, off, white, red, black, yellow, blue	927
16	Medical-suit	on, off	155
17	Safety-suit	on, off	530



Dataset after removing medical images



Methodology Used

Overview: Combines object detection and pose estimation for real-time PPE verification.

YOLOv8 for Object Detection:

- It is a real-time object detection model that predicts bounding boxes and class labels in a single forward pass through a neural network, making it fast and accurate.
- High-speed, accurate detection of PPE items like helmets, gloves, and vests.

MediaPipe for Pose Estimation:

- It uses lightweight machine learning models to detect human body landmarks (keypoints) in real-time, enabling accurate pose estimation and body part localization.
- Real-time keypoint detection for correct PPE placement (e.g., helmet on head).
- Detects 33 body keypoints to verify body posture and motion.

Integration:

- YOLOv8 detects PPE, MediaPipe verifies correct wearability.
- Parallel processing for immediate alerts in dynamic environments.



Detected violations: ['Safety vest detected but not worn properly']



Detected violations: ['Safety vest missing from image']

Performance Metrics for Object Detection

- **mAP (Mean Average Precision):** 0.734 overall, 0.915 person, 0.935 face & more.
- **Accuracy:** The proportion of correctly predicted PPE presence labels (present or not) over all predictions.
 - 0.97
- **Precision:** The proportion of correctly detected PPE items among all detections.
 - 0.98
- **Recall:** The proportion of correctly detected PPE items among all actual PPE items in the dataset.
 - 0.97
- **F1-score**
 - 0.97

Performance Metrics for Pose Estimation

- **Accuracy:** The proportion of correctly predicted PPE compliance labels (compliant or not) over all predictions.
 - 0.93
- **Precision:** The proportion of correctly identified compliant PPE items (e.g., helmets worn on the head) among all items predicted as compliant.
 - 0.90
- **Recall:** The proportion of correctly detected PPE items among all actual PPE items in the dataset.
 - 0.96
- **F1-score**
 - 0.92

Performance Metrics

```

Metrics for 'person_present':
  Accuracy : 1.00
  Precision: 1.00
  Recall   : 1.00
  F1-Score : 1.00

Metrics for 'helmet_present':
  Accuracy : 1.00
  Precision: 1.00
  Recall   : 1.00
  F1-Score : 1.00

Metrics for 'helmet_compliant':
  Accuracy : 0.93
  Precision: 0.91
  Recall   : 1.00
  F1-Score : 0.95

Metrics for 'gloves_present':
  Accuracy : 0.93
  Precision: 0.91
  Recall   : 1.00
  F1-Score : 0.95

Metrics for 'gloves_compliant':
  Accuracy : 0.93
  Precision: 0.88
  Recall   : 1.00
  F1-Score : 0.93

Metrics for 'safety_vest_present':
  Accuracy : 1.00
  Precision: 1.00
  Recall   : 1.00
  F1-Score : 1.00

Metrics for 'safety_vest_compliant':
  Accuracy : 0.93
  Precision: 1.00
  Recall   : 0.86
  F1-Score : 0.92

Metrics for 'safety_suit_present':
  Accuracy : 0.93
  Precision: 1.00
  Recall   : 0.83
  F1-Score : 0.91

Metrics for 'safety_suit_compliant':
  Accuracy : 0.93
  Precision: 0.80
  Recall   : 1.00
  F1-Score : 0.89

== PRESENCE (OBJECT DETECTION) METRICS ==
Average Accuracy : 0.97
Average Precision: 0.98
Average Recall   : 0.97
Average F1-Score : 0.97

== COMPLIANCE (POSE ESTIMATION) METRICS ==
Average Accuracy : 0.93
Average Precision: 0.90
Average Recall   : 0.96
Average F1-Score : 0.92

```

Class	Images	Instances	Box(P	R	mAP50
all	1513	13765	0.79	0.682	0.734
person	1412	2504	0.871	0.889	0.915
ear	919	1485	0.925	0.769	0.846
face	1119	1809	0.953	0.892	0.935
tool	405	861	0.535	0.327	0.359
glasses	287	351	0.809	0.715	0.746
gloves	212	439	0.741	0.56	0.634
helmet	82	134	0.789	0.641	0.729
hand	1222	3009	0.887	0.842	0.896
head	1215	2204	0.944	0.884	0.936
shoes	298	850	0.822	0.616	0.707
safety-suit	19	32	0.598	0.5	0.542
safety-vest	42	87	0.606	0.547	0.567

Overcoming Key Challenges

Challenge: Low Initial Detection Accuracy (mAP = 0.389)

Solution:

- **Hyperparameter Tuning:** Adjusted learning rate, batch size, and number of epochs for better model convergence.
- **Data Augmentation:** Applied image transformations (rotation, scaling, color variations) to increase model robustness.
- Changed model from YOLOv8n to YOLOv8l.

Scaling Real-World Impact

Target Areas:

- High-risk environments like construction sites, manufacturing plants, and warehouses.

Future Improvements:

- **Mobile or Low-Resource Deployment:** Optimize the model for mobile devices or edge computing for on-site monitoring.
- **Cross-Domain Transfer:** Extend the model for other sectors like healthcare (e.g., PPE for doctors and nurses) or manufacturing.
- **Integration with Smart Surveillance Systems:** Leverage existing CCTV infrastructure for real-time, automated PPE monitoring.

Application at Plaksha?

- Yes, the solution is highly applicable as extensive construction work is ongoing on campus, where workers are often seen without proper safety equipment.
- CCTV footage from construction sites can be integrated with the PPE detection and compliance system for real-time monitoring.
- Potential challenge with this would be the requirement of a high resolution and good quality camera.
- The model can be deployed on a central server to alert supervisors when non-compliance is detected.

Challenges while Scaling?

- **High Computational Requirements:** Real-time PPE compliance monitoring across multiple sites will require powerful hardware or cloud infrastructure, which can be costly to maintain and scale.
- **Dataset Generalization Issues:** The model might struggle with unfamiliar PPE designs, cultural attire, or environmental conditions unless continuously retrained with diverse data.
- **Integration and Maintenance Overhead:** Integrating the system with different existing security or safety monitoring setups and maintaining consistent performance across all locations may pose logistical and technical challenges.

THANK YOU!

