International Journal of Research Publication and Reviews, Vol (5), Issue (12), December (2024), Page – 5254-5257

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Object Detection Via Independent Dataset Training

Sai Ananya J[1], Shreya C Mathad[2], Tejashwini V[3], Vinutha H V[4], Manasa Sandeep[5]

[1] Dayananda Sagar Academy of Technology and Management saiananyajana92@gmail.com
[2] Dayananda Sagar Academy of Technology and Management shreyacm7@gmail.com
[3] Dayananda Sagar Academy of Technology and Management vteju328@gmail.com
[4] Dayananda Sagar Academy of Technology and Management vidyavinutha909@gmail.com
[5] Dayananda Sagar Academy of Technology and Management manasa-cs@dsatm.edu.in

ABSTRACT :

Worker safety in industrial and construction settings is a must for their safety. Compliance with Personal Protective Equipment (PPE) requirements, such as gloves, footwear, goggles, jackets, and helmets, is necessary to prevent injuries at the workplace. Manual monitoring of PPE compliance is usually laborious, error-prone, and inefficient, especially in large-scale operations. This project introduces an Intelligent PPE Detection System based on the You Only Look Once (YOLO) model to address these challenges.

Object recognition has been used as a highly powerful method for the automatic processing of visual data so that PPE items are found real-time. A custom dataset was prepared and trained precisely for this application, incorporating critical stages such as image annotation, dataset generation, and fine-tuning of the YOLO architecture for optimal performance. Thus, relying on a large and diversified dataset ensures high precision with a minimal number of false positives and negatives, with particular emphasis on model reliability and robustness.

The real-time detection system provides visual outputs with bounding boxes and confidence scores, automating PPE compliance checks with minimal human error. This enhances workplace safety, offering an efficient and scalable solution for safety monitoring. Moreover, the project demonstrates the flexibility of YOLOV8 in being applied to various real-world applications, such as traffic monitoring, inventory management, and autonomous navigation. By showing the practical implementation of YOLO for PPE detection, the project indicates its potential in addressing more extensive domain-specific challenges.

Introduction :

Working in industrial and construction industries requires proper care regarding the issue of safety in the work environment. Following Personal Protective Equipments (PPEs) is the basic need for averting accidents and injuries;PPE includes helmets, goggles, gloves, boots, and jackets, so it protects individuals from a different type of hazards. Its traditional method of inspection on PPEs involves people, which are resource time-consuming, inconsistent, and prone to human error. This challenge is further amplified in large-scale operations where the sheer number of workers makes real-time monitoring impractical. To address these limitations, there is a growing need for automated systems that ensure accurate and efficient compliance checks.Recent advances in computer vision and artificial intelligence have made object recognition revolutionize many applications, such as autonomous vehicles, surveillance, and industrial automation. This paper discusses how these advances can be leveraged to create an Intelligent PPE Detection System based on the You Only Look Once (YOLO) .YOLO is a real-time object detection framework that is known for its speed and accuracy in object detection, which makes it an ideal choice for detecting PPE items in dynamic environments.

The project focuses on the development of a custom dataset that caters to the specific needs of PPE detection. Unlike pre-trained models that rely on generic datasets, a custom dataset ensures that the system is optimized for the unique challenges of workplace safety. The dataset preparation includes the collection of images of PPE items in different scenarios, annotation to define bounding boxes, and division of the data into training, validation, and testing subsets. This process enhances the model's robustness, enabling it to perform reliably in diverse conditions, such as varying lighting and occlusions. The use of Custom Dataset will help to train the model according to specific real world requirements such as workers safety in this case and thus to get accuracy for the training purpose usage of high- quality and more number of images are gathered and used.Real- time PPE detection is realized by integrating the trained YOLOV8 model in a system that offers visual outputs such as bounding boxes, labels, and confidence scores for the detected items. This auto approach reduces human intervention considerably, minimizes errors, and increases the efficiency of compliance monitoring. It ensures safety for workers while also becoming a scalable solution for high-risk operations industries.The significance of this project transcends workplace safety.

Literature survey :

Object detection using YOLO,challenges, architectural successors, datasets, and applications

The YOLO algorithm is superior in real-time object detection
because it treats the problem as a single regression problem, predicting the bounding boxes and class probabilities simultaneously. Custom datasets improve the utility of YOLO in domain-specific applications but face challenges such as class imbalance, annotation accuracy, and the detection of small objects. Successors like YOLOv3, YOLOv4, and YOLOv8 overcome these limitations through architectural improvements, such as multi-scale predictions and advanced activation functions. The proper collection and annotation of the dataset, along with appropriate validation and testing splits, are essential for effective custom training. These developments enable YOLO to fulfill various, real-world object detection needs.

Deep Residual Learning for Image Recognition

Deep Residual Networks significantly improved object detection by allowing the training of deeper networks through skip connections, thus avoiding the vanishing gradient problem. ResNets are robust backbones in object detection models, providing hierarchical feature extraction critical for custom datasets. They generalize well even with limited data, reducing over-fitting and improving accuracy. ResNets excel at detecting small or subtle objects, a common challenge in custom datasets. Their integration into frameworks such as YOLO ensures efficient and precision detection tailored to domain-specific requirements

Exploring Data Augumentation for Deep Learnning Model Genralization

The paper on exploring data augmentation for deep learning model generalization highlights how augmentations enhance performance in object detection tasks, especially with custom datasets. It generates variations in the training data through techniques like flipping, cropping, and color adjustments, which ensures that models learn robust and generalized features. This approach helps in mitigating over fitting, so the model performs well on unseen data. For custom datasets, where the diversity is usually limited, it simulates real-world cases very effectively, enhancing accuracy and reliability in detection. It's an important step toward developing adaptable, high-performance object detection systems.

Methodology :

It includes Requirement Gathering and Analysis

Data Collection

To create a robust and efficient PPE detection system, the dataset was curated meticulously from diverse images collected from various publicly available sources, like Google Images. The dataset had to represent a wide scope of workplace scenarios to cover crucial PPE items, such as gloves, footwear, goggles, jackets, and helmets, under different lighting conditions, angles, and environmental settings. The images were manually annotated with labeling tools, which gave the exact location and class of each PPE item in the images. Then, the dataset was split into three subsets: training, validation, and testing. The training set, comprising 70% of the dataset, was used to train the deep learning model, enabling it to learn feature representations of PPE items. The validation set was applied for the training of models by using 20% of the dataset to fine-tune hyper-parameters and avoid over-fitting. The other 10% constituted the testing set, which was saved to test the final model on new, unseen data. In this way, data was collected and prepared according to a structured methodology in such a way that a well-balanced dataset would lead to generalization and strengthen the model to recognize PPE in the real world.

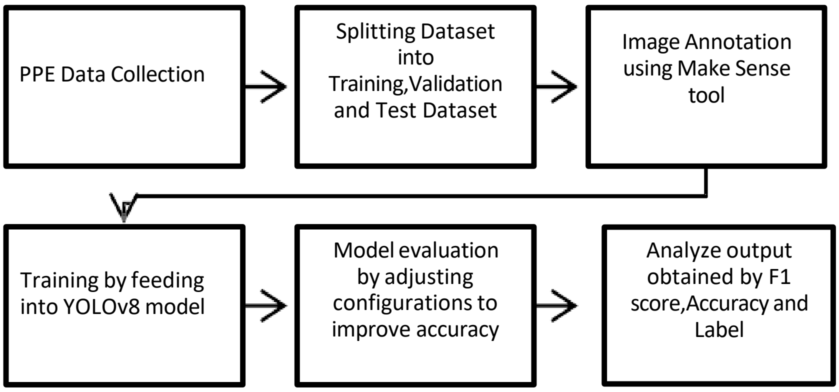
Data Annotation

The dataset collected was enhanced and annotated using the Make-Sense platform, a user-friendly web based tool for preparing datasets. This tool is used to label images and define bounding box parameters: x-axis and y-axis coordinates, width, and height for each object. To increase the diversity of the dataset, data augmentation techniques, such as flipping, rotation, and brightness adjustments, were applied, simulating various real-world conditions. The Make-Sense tool efficiently annotated all images of the PPE items: gloves, footwear, goggles, jackets, and helmets to be in YOLO format compatible. The annotations were then exported with the augmented dataset in text files in the YOLO format. Each line corresponds to an object and thus specifies a class label and normalized bounding box parameters. This streamlined approach resulted in better quality of the dataset and also reduced efforts for manual annotation. This augmented dataset really improved the generalization ability of YOLOv8 that enabled it to work perfectly in various workplace scenarios robustly. It was critical in preparing the model before it could achieve precise real- time PPE detection and compliance monitoring
Total Annotations = Number of Images * Average number of Objects per Image

Training

The YOLOv8 model was trained using Google Colab, which provided free GPU resources to train the model efficiently and cost-effectively. The custom dataset was prepared by downloading PPE images from Google Images and annotated using Labeling in a format compatible with YOLO. The dataset was divided into three subsets: training, validation and testing, which ensured a balanced distribution across classes such as gloves, footwear, goggles, jackets, and helmets. A configuration file named "data.yaml" was created in order to mention the dataset paths, number of classes, and class names. Transfer learning was used to initialize the pre- trained weights of the YOLOv8 model in order to utilize existing feature representations to improve training efficiency. Fine-tuning hyperparameters for training included an input image size of 640×640 pixels, a batch size of 16, and an adaptive learning rate scheduler. The model was trained for 25 epochs with validation to monitor and curb over-fitting. Training was done with metrics like training loss, validation loss, precision, recall, and mean Average Precision (mAP) for monitoring performance. The Google Colab's GPU made the computation faster and faster convergence. The trained model is tested on the test dataset with accurate detection of PPE items under different conditions. The complete training approach will make the model effective for real- time monitoring in the workplace for compliance and manual oversight.

Implementation :



The flow diagram describes about the implementation of project.

Algorithms :

Neural Network Architecture

The architecture of YOLOv8 is built upon previous iterations but incorporates major improvements on the way to higher accuracy and speed. It consists of three main components: the backbone, neck, and head— designed in a modular approach. The backbone is based on CSPNet, which improves the efficiency of feature extraction by partitioning the feature maps into two parts and only partially propagating them through the network to reduce the computational cost without loss of representation power. The neck employs a PANet, which facilitates the fusion of features across different scales, ensuring robust multi-scale predictions. Finally, the head predicts bounding boxes, class labels, and, if present, segmentation masks by summing up all the learned features into actionable outputs.

Loss Functions

The architecture of YOLOv8 is built upon previous iterations but incorporates major improvements on the way to higher accuracy and speed. It consists of three main components: the backbone, neck, and head— designed in a modular approach. The backbone is based on CSPNet, which improves the efficiency of feature extraction by partitioning the feature maps into two parts and only partially propagating them through the network to reduce the computational cost without loss of representation power. The neck employs a PANet, which facilitates the fusion of features across different scales, ensuring robust multi-scale predictions. Finally, the head predicts bounding boxes, class labels, and, if present, segmentation masks by summing up all the learned features into actionable outputs.

Conclusion :

The performance evaluation of our PPE detection model clearly shows its high accuracy and reliability in identifying essential safety gear. With an impressive overall F1 score and robust performance across all PPE classes, our implementation of real-time detection using bounding boxes with confidence percentages and labels provides a practical solution for enhancing workplace safety by enabling immediate identification of compliance and non-compliance.The automation of monitoring reduces dependency on human inspection and mitigates the aspect of human error. The level of vigilance is always consistent with a higher-risk environment. This technology supports compliance with safety standards while fostering a safer and more productive work environment. Potential enhancements may be in extending the ability of the system to add other PPE types as well as making the system adaptable to varying environmental conditions.

REFERENCES :

1. Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *arXiv:1804.02767*.

2. Redmon, J., & Farhadi, A. (2016). YOLO9000: Better, Faster, Stronger. *arXiv:1612.08242*.

3. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *arXiv:1506.02640*.
4. Wang, C., & Zhang, Y. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv:2004.10934*.
5. Bhalodi, A., & Kaur, A. (2021). YOLOv5: A New Approach to Object Detection. *arXiv:2104.10934*.
6. Bertasius, G., Yang, J., & Shi, J. (2020). Object Detection with a Single Deep Neural Network. *arXiv:2004.10934*.
7. Zhou, X., Wang, D., & Tang, X. (2019). Objects as Points. *arXiv:1904.07850*.
8. Liu, S., & Wang, Y. (2020). A Comprehensive Review on Object Detection. *arXiv:2004.10934*.