DETECTION OF PPE & FACIAL RECOGNITION OF INDUSTRY WORKERS

Srikar Mukkamala, Elavartha Nikhil Reddy, Boppudi Sai Ashrith, Ankadala Hemanth
CSD Major Project
Instructor : Gagan Raj Gupta

***Abstract*—Surveillance Application to verify PPE in industry In this project, we have to process feeds from video surveillance cameras, detect and track workers and alert with name of the person who is not wearing proper "personal protection equipment" for the specified area/zone. Eg. Construction worker, mill worker, miner etc.**

I. Introduction

Factories can be risky if people don't have the right safety gear and training. New computer programs with smarts (AI) are being used more and more in places where safety is important.

This project is building a system to make sure people wear their Personal Protective Equipment (PPE) like goggles, gloves or masks. It uses special cameras to see what people are wearing and who they are. If someone has the right gear, the system acts like a keycard and lets them enter a special area. This could prevent many accidents and injuries every year.

Every year, a shocking number of people - around 350,000 - die from accidents at work, according to the International Labour Office. Additionally, millions more are injured badly enough to miss work for several days.

Good safety practices in the workplace are linked to greater productivity. However, some jobs are riskier than others. In the United States alone, construction sees the most work-related deaths, with over 5,000 fatalities between 2016 and 2020. This is significantly higher than other industries, like agriculture or finance.

Research by the National Institute for Occupational Safety and Health (NIOSH) points out that a quarter of all construction deaths involve head injuries, and most of these (84%) could have been prevented by wearing a hardhat.

Construction sites often involve heavy machinery and trucks, making it crucial for workers to be easily seen to avoid accidents. Wearing high-visibility vests with reflective material can significantly decrease the risk of serious injuries or death.

II. Goal

This project aims to boost safety for workers in high-risk environments. We're building a powerful system that uses image recognition to check if workers are wearing their Personal Protective Equipment (PPE). This system will combine with an siamese neural network for identifying faces, all while running fast and smoothly in real-time for a user-friendly experience.

We'll be focusing on detecting these specific PPE items:
1. Hard Hat
2. Safety Vest
3. Humans

This project explores a new approach to workplace safety by combining facial recognition with object detection for PPE (Personal Protective Equipment). While both technologies are commonly used in monitoring systems separately, this study is one of the first to combine them. This approach allows the system to identify individuals and verify they are wearing the required PPE simultaneously, potentially enhancing safety and security in hazardous workplaces.

The project involved detecting five different PPE items (list them here if provided) along with facial recognition. This combined approach offers a unique contribution to the field of workplace safety.

III. Preliminaries

A. *Yolo v9*

YOLOv9 is a recent advancement in real-time object detection, aiming to improve upon its predecessor, YOLOv8. It achieves this through innovative techniques like Programmable Gradient Information (PGI) and the Generalized Efficient Layer Aggregation Network (GELAN). PGI tackles the challenge of information loss within the neural network, crucial for accurate object detection. GELAN contributes by enhancing information flow and allowing the network to retain crucial details throughout the detection process. These features, combined with YOLOv9's efficient architecture, contribute to its superior performance in terms of accuracy, speed, and efficiency compared to previous versions. Although not yet as widely adopted as earlier YOLO models, YOLOv9 holds promise for various real-world applications requiring high-performance object detection.

B. *One Shot Learning*

One-shot learning is a learning network with a single sample image. Let us take an example at the security gate of the company where the employees are allowed after verifying their face in their database. Suppose there might be a problem where an employee does not have more than ten images and building a convolutional neural network becomes difficult. Suppose a new member is added or removed from the company then the whole network is to be trained again. Here comes the one-shot learning, where we build a similarity function that compares two images and tells us whether they matched or not.

C. Many Shot Learning

In this paper we want to explore what happens if we use multiple positive images to compare the photo and say whether they match or not. For that we create a new function quadraples where we include two positive images along an anchor image and negative image.

D. Transfer Learning

Transfer Learning (TL) is a machine learning method that stores knowledge gained while learning in one problem and applies it to a similar but different problem. For example, you want your system to recognize human faces in an image. By using models that have already been trained on millions of faces (also known as pre-trained models), we can solve related problems without the need for large amounts of data.

E. Face Recognition

It is a method to verify or identify a person from a picture or from a frame in the video. It works on the basis of comparing that person's face to that of faces present in the dataset. The model is trained on the images in the pubfig dataset.

IV. Methodology

A. PPE Detection

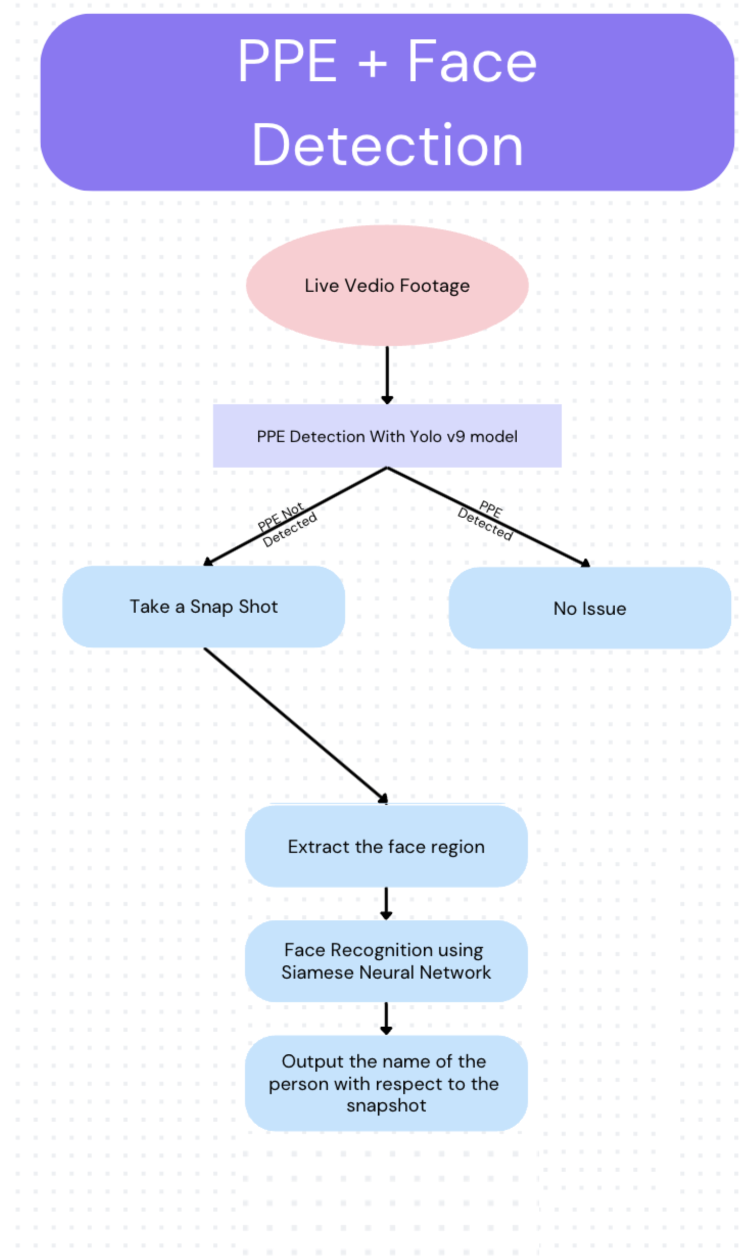
This project leverages YOLOv9, a cutting-edge object detection model known for its speed and accuracy, to identify PPE (Personal Protective Equipment) kits. YOLOv9's efficiency makes it suitable for real-time applications. To ensure accurate detection, we employ a combination of three loss functions: box loss, class loss, and DFL. These functions help the model refine its predictions during training. Additionally, we utilize three metrics: precision, recall, and MAP, to evaluate the model's performance and effectiveness in detecting PPE kits accurately.

B. Face Detection

To hone in on facial features, we perform a precise zoom on the image, extracting only the facial region. For facial recognition, we plan to explore the Siamese neural network architecture. This approach excels at comparing similarities between inputs, making it well-suited for identifying individuals. We will test the effectiveness of two loss functions from my previous research project: triplet loss and quadruple loss. These will guide the network's learning process by penalizing it for incorrect comparisons and rewarding it for accurate ones.

C. High Level Design

The system prioritizes safety. When someone is detected wearing their PPE correctly, the system focuses on that individual and no facial recognition occurs. However, if someone is not wearing proper PPE, the system takes a snapshot of their face and extracts the facial region. This information is then used for facial recognition to identify the individual, likely for further action or training purposes.



V. Individual Contribution

Srikar - PPE Detection Model & Backend Deployment
Nikhil - Face Recognition Model & Testing
Hemanth - Backend Deployment of the Model
Ashrith - Testing

VI. Dataset

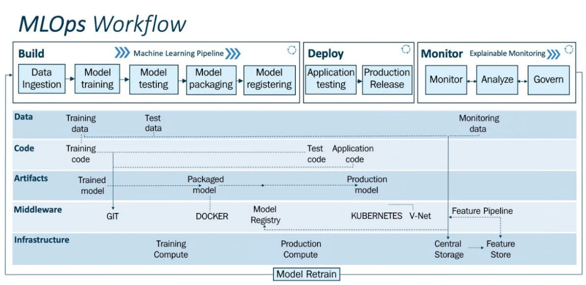
We've talked to Gagan sir about the dataset and we're in touch with a senior PHD student. We'll keep them updated on our progress in the first phase.

VII. Metrics

Latency: Measure the time taken for the system to process and respond to incoming video feeds.
Throughput: Evaluate the number of video feeds processed per unit of time.
Error Rate: Monitor the accuracy of PPE detection and facial recognition to identify potential issues.
Resource Utilization: Track CPU and memory usage to optimize resource allocation and identify bottlenecks.

VIII. Operation Strategies

Continuous Monitoring: Implement monitoring tools to track system health, performance, and potential issues in real-time.
Regular Maintenance: Schedule regular maintenance tasks, such as model updates, to ensure the system stays up-to-date and continues to perform effectively.
Incident Response: Develop a robust incident response plan to address unexpected issues promptly and minimize downtime.
Backup and Recovery: Implement regular backup mechanisms to recover from data loss or system failures efficiently.
Deployment and Inference: We will try to implement a forward proxy where we can also have a rate limiter and load balancer where users can be directed to different models according to CI/CD and Model improvement.



IX. Time Line

A. Phase 1

By the end of phase 1, we will have completed the model for the project, including the development of the machine learning (ML) model. This means we'll have created the algorithm and trained it to perform the tasks required for the project. Once phase 1 is finished, we'll be ready to move on to the next steps, like testing and refining the model.

B. Phase 2

During phase 2, we'll be taking the next step by deploying the model for real-time use. This means that the model will be put into action immediately, allowing us to utilize its capabilities in real-life situations as they happen. This deployment will enable us to assess its performance and effectiveness in real-time scenarios, providing valuable insights for further optimization and refinement.

X. References

1. https://arxiv.org/abs/2402.13616
2. https://www.cs.cmu.edu/ rsalakhu/papers/oneshot1.pdf
3. S. M. M, A. Geroge, A. N and J. James, "Custom Face Recognition Using YOLO.V3," in 2021 3rd International Conference on Signal Processing and Communication (ICPSC), 2021.