International Journal of Advance Research in Science and Engineering
Volume No. 13, Issue No. 06, June 2024
www.ijarse.com
IJARSE
ISSN 2319 - 8354

# WORKERS SAFETY HELMET DETECTION USING YOLO

Ms. T L K Prasnna[1],Ms. P. Sirisha[2],Ms. B. Priyanka[3], Ms. A. Venkata Ramana[4], Mr. P. Gopi[5]
[1] Assistant professor, Department of ECE, Tirumala Engineering College

*Abstract*— **Computer vision technology like You Only Look Once (YOLO) can help make construction sites safer. These tools can automatically spot people without helmets and identify who's in the area. Putting these technologies into a Flask framework makes it easy for anyone to use. We can also set up alerts to let the right people know if someone isn't wearing their helmet. Testing everything well and keeping an eye on it ensures the system works reliably when it matters most. Surveillance systems offer a promising solution to enhance safety protocols in construction sites .So, to do that activity using whether the person wearing a helmet or not in critical places. For these we are using deep learning algorithm construction using yolo model. In this helmet detection will be done by simple yolo and rcnn models with better pre- trained models.**

*Keywords—Safety helmet wearing detection,YOLOv8, four detection scales, attention*

*chanism,, RCNN*

## 1. INTRODUCTION:

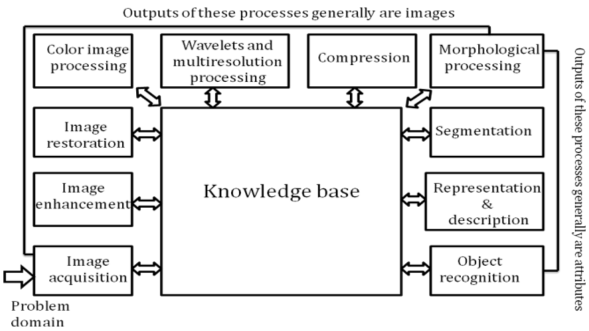



**Fig.1:Basic fundamental steps of image processing**

An image is a two-dimensional picture, which has a similar appearance to some subject usually a physical object or a person. Image is a two-dimensional, such as a photograph, screen display, and as well as a three-dimensional, such as a statue. They may be captured by optical devices—such as cameras, mirrors, lenses, telescopes, microscopes, etc. and natural objects and phenomena, such as the human eye or water surfaces.

The clever bit is to interpret collections of these shapes as single objects, e.g. cars on a road, boxes on a conveyor belt or cancerous cells on a

microscope slide. One reason this is an AI problem is that an object can appear very different when viewed from different

angles or under different lighting. Another problem is deciding what features belong to what objectand which are

background or shadows etc. Thehuman visual system performs these tasks mostly unconsciously but a computer requires skilfull programming and lots of processing power to approach human performance

Manipulation of data in the form of an image through several possible techniques. An image is usually interpreted as a two-dimensional array of brightness values, and is most familiarly represented by such patterns as those of a photographic print, slide, television screen, or movie screen. An image can be processed optically or digitally with a computer

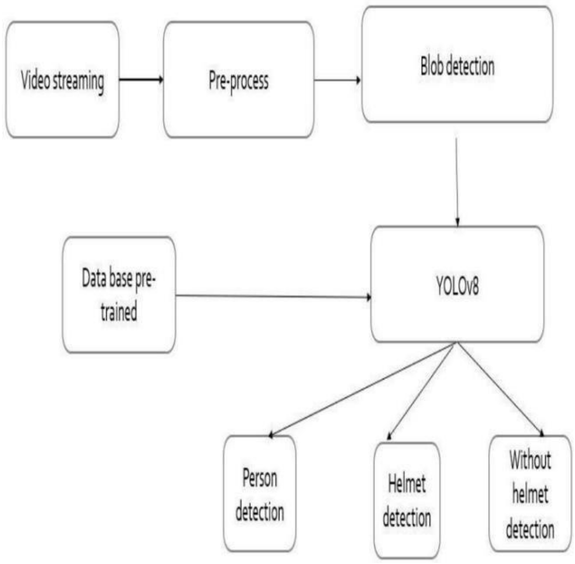
The main purpose of this paper is to propose a fast, high precision and easy-to-use network to realize the automatic detection of helmet wearing status in the construction site. At the same time, the model should have strong generalizationability, not limited to helmet detection in a specific scene, and can well deal with small targets and multi-scale challenges. Above factors are taken into fully consideration in the dataset used in this paper, and various scenes including helmet images of various challenges are collected.

Currently, the deep learning-based methods are data driven, which means the performance of object detection algorithms depends on the datasets used for training, but the number of safety helmet images is limited. In order to improve the performance of YOLOv8 on the task of safety helmet wearing detection.

## 2. RELATED WORK

Object Detection algorithms find application in various fields such as defence, security, and healthcare. In this paper various Object Detection Algorithms such as face detection, skin detection, colour detection, shape detection, target detection are simulated and implemented using MATLAB 2017b to detect various types of objects for video surveillance applications with improved accuracy. Further, various challenges and applications of Object Detection methods are elaborated. Index Terms— Colour Detection, Face Detection, Object Detection Algorithms, Skin Detection, Target Detection, Video Surveillance.[1].

In this paper, we propose a method to detect abandoned object from surveillance video. In first step, foreground objects are extracted using background subtraction in which background modeling is done through running average method.




In second step, static objects are detected by using contour features of foreground objects of consecutive frames. In third step, detected static objects are classified intohuman and non-human objects by using edge based object recognition method which is capable to generate the score for full or partial visible object. Nonhuman static object is analyzed to detect abandoned object. Experimental results show that proposed system is efficient and effective for real-time video surveillance, which is tested on IEEE Performance Evaluation of Tracking and Surveillance data set (PETS 2006, PETS 2007) and our own dataset. Keywords— Background subtraction Foreground Abandoned object detection.[2].

Fig.2. Proposed System Block Diagram

Object detection in a road scene has received a significant attention from research fields of developing autonomous vehicle and automatic road monitoring systems. However object occlusion problems frequently occur in generic road scenes. Due to such occlusion problems, previous object detection methods have limitations of not being able to detect objects accurately. In this paper, we propose a novel object detection network which is robust in occlusions. For effective object detection even with occlusion, the proposed network mainly consists of two parts;
Object detection framework, 2) Multiple object bounding box (OBB)-Critic network for predicting a BB map which estimates both object region and occlusion region. Comprehensive experimental results on a KITTI Vision Benchmark Suite dataset showed that the proposed object detection network outperformed the state-of-the-art methods. Index Terms— Object detection, adversarial learning, actor-critic network, plug-in, occlusion [3].

Object Detection algorithms find application in various fields such as defence, security, and healthcare. In this paper various Object Detection Algorithms such as face detection, skin detection, colour detection, shape detection, target detection are simulated and implemented using MATLAB 2017b to detect various types of objects for video surveillance applications with improved accuracy. Further, various challenges and applications of Object Detection methods are elaborated. Keywords— Colour Detection, Face Detection, Object Detection Algorithms, Skin Detection, Target Detection, Video Surveillanc[4]

## 3. Methodology

### 3.1 PROPOSED SYSTEM

Video Streaming: Video streaming technology is one way to deliver video over the Internet. Using streaming technologies, the delivery of audio and video over the Internet can reach many millions of customer using their personal computers, PDAs, mobile smartphones or other streaming devices.
Pre-processing: Pre-processing is a common name for operations with images at the lowest level of abstraction -- both input and output are intensity images. The aim of pre- processing is an improvement of the image data that suppresses unwanted distortions or enhances some image features important for further processing.
Blob Detection: In computer vision, blob detection methods are aimed at detecting regions in a digital image that differ in properties, such as brightness or

IJARSE
ISSN 2319 - 8354

color, compared to surrounding regions. Informally, a blob is a region of an image in which some properties are constant or approximately constant; all the points in a blob can be considered in some sense to be similar to each other. The most common method for blob detection is convolution
YOLO (You Only Look Once): YOLO is a real-time object detection algorithm known for its speed and efficiency. It divides the input image into a grid and predicts bounding boxes and class probabilities directly from the entire image. RCNN (Region-based Convolutional Neural Network): RCNN-based approaches, such as Faster R-CNN or Mask R- CNN, propose regions of interest (ROI) and then classify these regions. They are known for their accuracy but may be slower than YOLO. Convolutional neural networks. Sounds like a weird combination of biology and math with a little CS sprinkled in, but these networks have been some of the most influential innovations in the field of computer vision. 2012 was the first year that neural nets grew to prominence as Alex Krizhevsky used them to win that year's ImageNet competition (basically, the annual Olympics of computer vision), dropping the classification error record from 26% to 15%, an astounding improvement at the time.Ever since then, a host of companies have been using deep learning at the core of their services. Facebook uses neural nets for their automatic tagging algorithms, Google for their photo search, Amazon for their product recommendations, Pinterest for their home feed personalization, and Instagram for their search infrastructure.

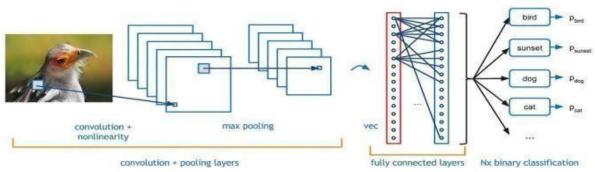


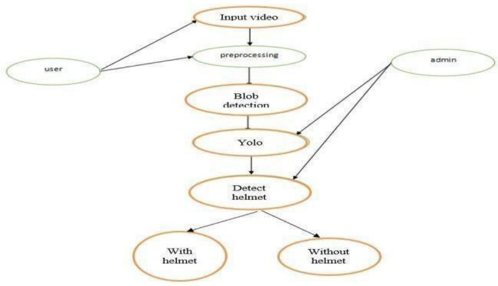

Fig:3.Convolution Neural Network

### 3.3 YOLO Algorithm :

YOLO is an abbreviation for the term 'You Only Look Once'. This is an algorithm that detects and recognizes various objects in a picture (in real-time). Object detection in YOLO is done as a regression problem and provides the class probabilities of the detected images.

YOLO algorithm employs convolutional neural networks (CNN) to detect objects in real-time. As the name suggests, the algorithm requires only a single forward propagation through a neural

network to detect objects.
This means that prediction in the entire image is done in a single algorithm run. The CNN is used to predict various class probabilities and bounding boxes simultaneously. The YOLO algorithm consists of various variants. Some of the common ones.In fig 4 shows the flow chart of proposed system.




### 3.4 Introduction to computer vision:

Fig 4 Flowchart of Proposed System

Computer Vision is the process of using machines to understand and analyze imagery (both photos and videos). While these types of algorithms have been around in various forms since the 1960's, recent advances in Machine

Learning, as well as leaps forward in data storage, computing capabilities, and cheap high-quality input devices, have driven major improvements in how well our software can explore this kind of content.

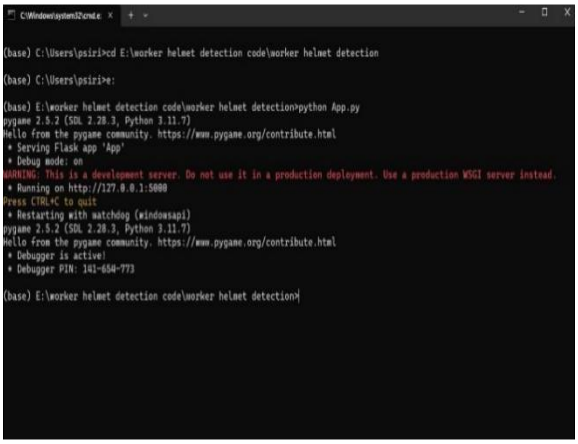
Computer Vision is the broad parent name for any computations involving visual content – that means images, videos, icons, and anything else with pixels involved. But within this parent idea, there are a few specific tasks that are core building blocks:

**WORKS:**

Machines interpret images very simply: as a series of pixels, each with their own set of color values. Consider the simplified image below, and how grayscale values are converted into a simple array of numbers.
Think of an image as a giant grid of different

squares, or pixels (this image is a very simplified version of what looks like either Abraham Lincoln or a Dementor). Each pixel in an image can be represented by a number, usually from 0 – 255. The series of numbers on the right is what software sees when you input an image. For our image, there are 12 columns and 16 rows, which means there are 192 input values for this image.

Fig.5: Grayscale conversion




## 4. Results & Discussion

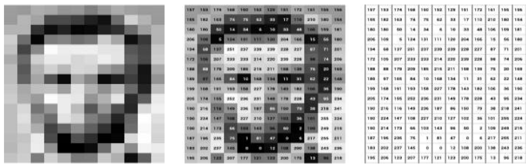
Fig. 6: Simulation results1

In this fig 6 HTML, we have a table that displays the simulation results. Each row of the table represents an image processed during the simulation, with the image displayed in the first column and the detection result (whether a helmet was detected or not) in the second vcolumn. You can replace "path/to/image1.jpg" and "path/to/image2.jpg" with the actual paths to your images, and add more rows as needed for additional images.

In fig 7 shows the wheather the person is wearing the helmet or not Using software to parse the world's visual content is as big of a revolution in computing as mobile was 10 years ago, and will provide a major edge for developers and businesses to build amazing products.

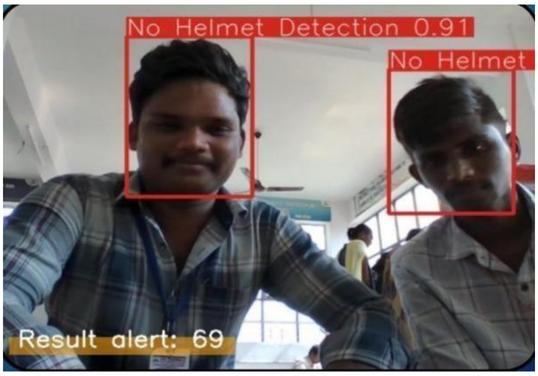
## Conclusion & Future scope

Helmet and person detection in construction sites using the YOLO object detection framework. By leveraging the capabilities of YOLO, we have developed a real-time surveillance system capable of identifying individuals without helmets and detecting peoplewithin construction site environments.



In future work, we plan to further refine and optimize our system to improve detection accuracy and reliability. This may involve exploring advanced techniques such as multi- scale object detection and incorporating contextual information to enhance scene understanding.

The future scope of worker helmet detection using


OpenCV and Python is quite promising, especially considering the increasing emphasis on safety in various industries such as construction, manufacturing, mining, and logistics.

**Enhanced Accuracy**: Continued advancements in machine learning algorithms and computer vision techniques can lead to even more accurate detection of helmets and other safety gear. This includes the development of deep learning models for object detection that can handle various lighting conditions, occlusions, and helmet types.

**Integration With Safety Protocols**: Worker helmet detection can be integrated with broader safety protocols and systems within workplaces. For example, triggering alarms or notifications when a workeris detected without proper head protection, ensuring compliance with safety regulations.

## References

1. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real- time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) (pp. 779- 788).

2. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R- CNN: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems (pp. 91- 99).

3. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In Proceedings of the IEEE international conference on computer vision (ICCV) (pp. 2961-2969).

4. Zhang, Y., Yang, S., & Zhao, X. (2020). Industrial safety helmet detection using deep learning for construction site monitoring. Automation in Construction, 113, 103129.

5. Chandran, S., & Narayanan, S. S. (2020). Deep learning- based approach for detection of safety helmets on construction workers. Automation in Construction, 110, 103010.

6. Huang, C. M., Wu, Y. J., Tsai, M. J., & Lin, W. J. (2020).
Intelligent detection system for improper wearing of safety helmets using deep learning. IEEE Access, 8, 79438- 79447.

7. Dong, B., Duan, Z., & Xu, Z. (2019). Helmet detection in construction images with convolutionalneural networks. Visualization in Engineering, 7(1), 1-15.

8. Islam, M. S., Sujan, M. S. H., Rahman, M. A., & Ferdous,
M. S. (2021). Smart helmet detectionsystem using YOLOv3 for industrial safety. In 2021 International Conference on Sustainable.

## Biography

Ms. T L K Prasanna received her M.Tech (VLSI Design) from Joganapally Institute of Techonology, Andhra Pradesh, India in the year 2014, She is currently working as Assistant Professor in ECE, Tirumala Engineering College,Narasaraopet,Andhra Pradesh.Her research interests are VLSI & Computer Networks.

P. Sirisha pursing B. Tech in Electronics &

Communication Engineering at Sai Tirumala NVR Engineering College ,Narasaraopet, Andhra Pradesh .

B. priyanka pursing B. Tech in Electronics & Communication Engineering at Sai Tirumala NVR Engineering College ,Narasaraopet, Andhra Pradesh.

A. Venkataramana pursing B. Tech in Electronics & Communication Engineering at Sai Tirumala NVR Engineering College ,Narasaraopet, Andhra Pradesh.