

# Person Detection

Zalak Shah <sup>1</sup>   Freyal Shah <sup>2</sup>   Khush Soni <sup>3</sup>   Dhruvam Bhalodiya <sup>4</sup>

<sup>1</sup>Ahmedabad University



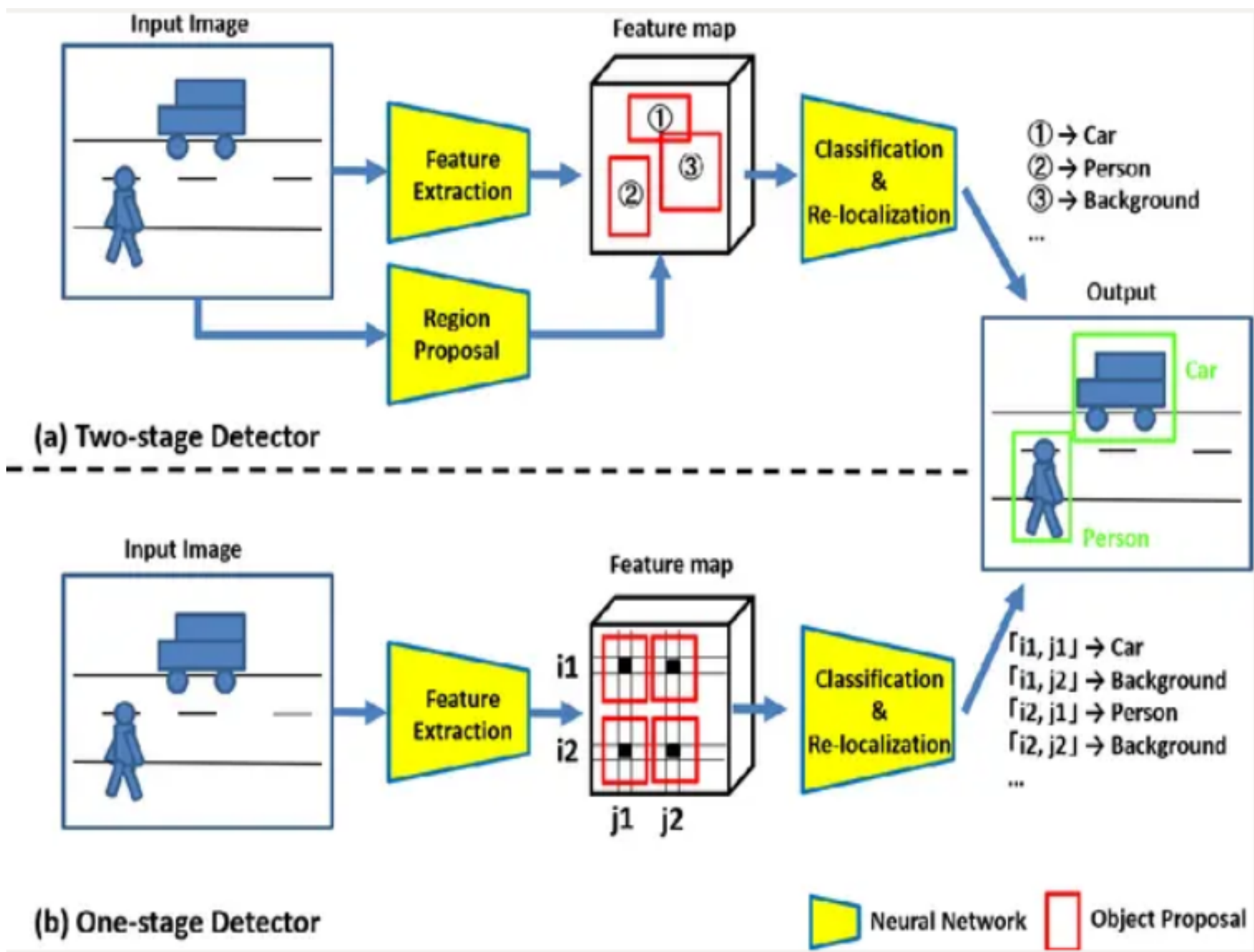
## Abstract

Object detection is a subfield of computer vision that involves training models to recognize and locate specific objects in images or videos, such as people, animals, vehicles, and other objects of interest. Person detection is a specific case of object detection, where the object of interest is a human body. This task has numerous applications in various fields, including surveillance, robotics, and human-computer interaction. The main objective of person detection is to locate the presence of people accurately and efficiently in images or videos, regardless of variations in pose, appearance, and background. The aim of this report is to provide an overview of the different object detectors, challenges faced in detecting objects, implementation using YOLOv8 object detector and the results produced when tested against a test data.

## Literature review

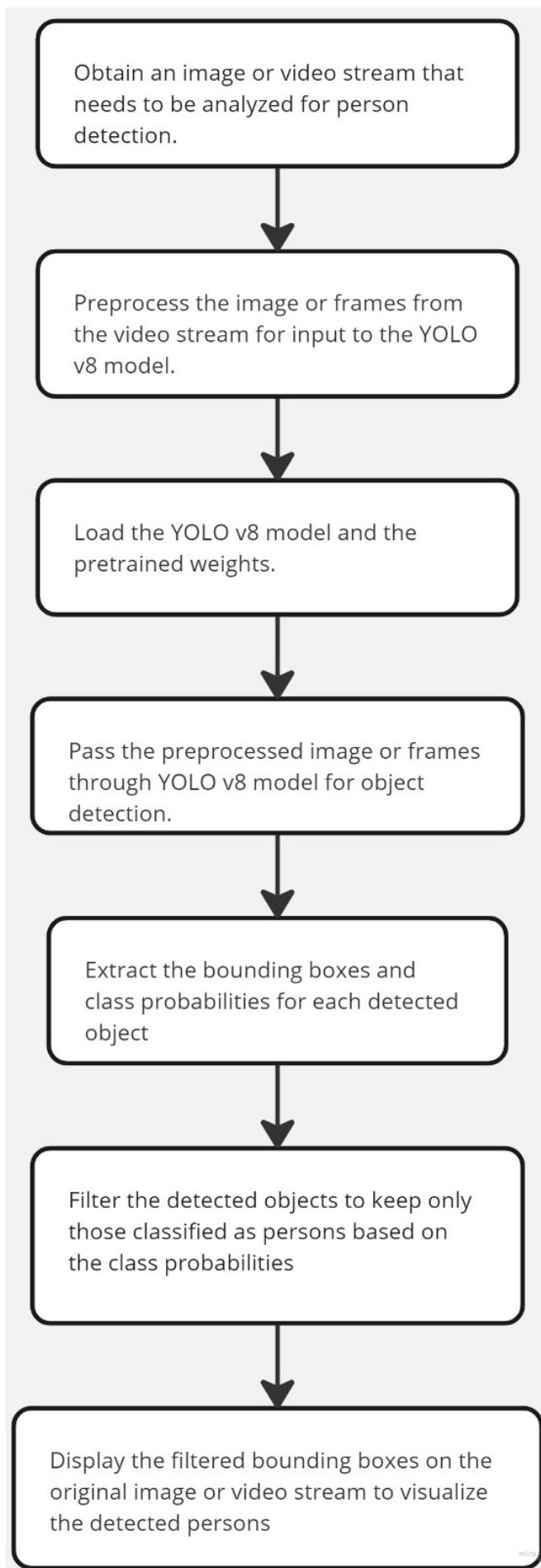
Over the past few decades, significant progress has been made, and numerous object detection algorithms have been proposed. In this literature review, some of the most prominent approaches to object detection are being explored. Viola-Jones Detector: Viola-Jones detector is a classic object detection algorithm [1-2] that uses sliding window approach. This includes going through all the possible locations in the image and seeing if any window has an object. This algorithm is widely used for face detection and has been extended to other object detection tasks. However, though the process was simple it was not a good idea to use this approach because the calculation was far beyond the computer's power.

- R-CNN (Regions with Convolutional Neural Network):
- Faster R-CNN:
- YOLO (You Only Look Once):



## Study methodology

The following shows the methodology of Person Detection.



## Dataset

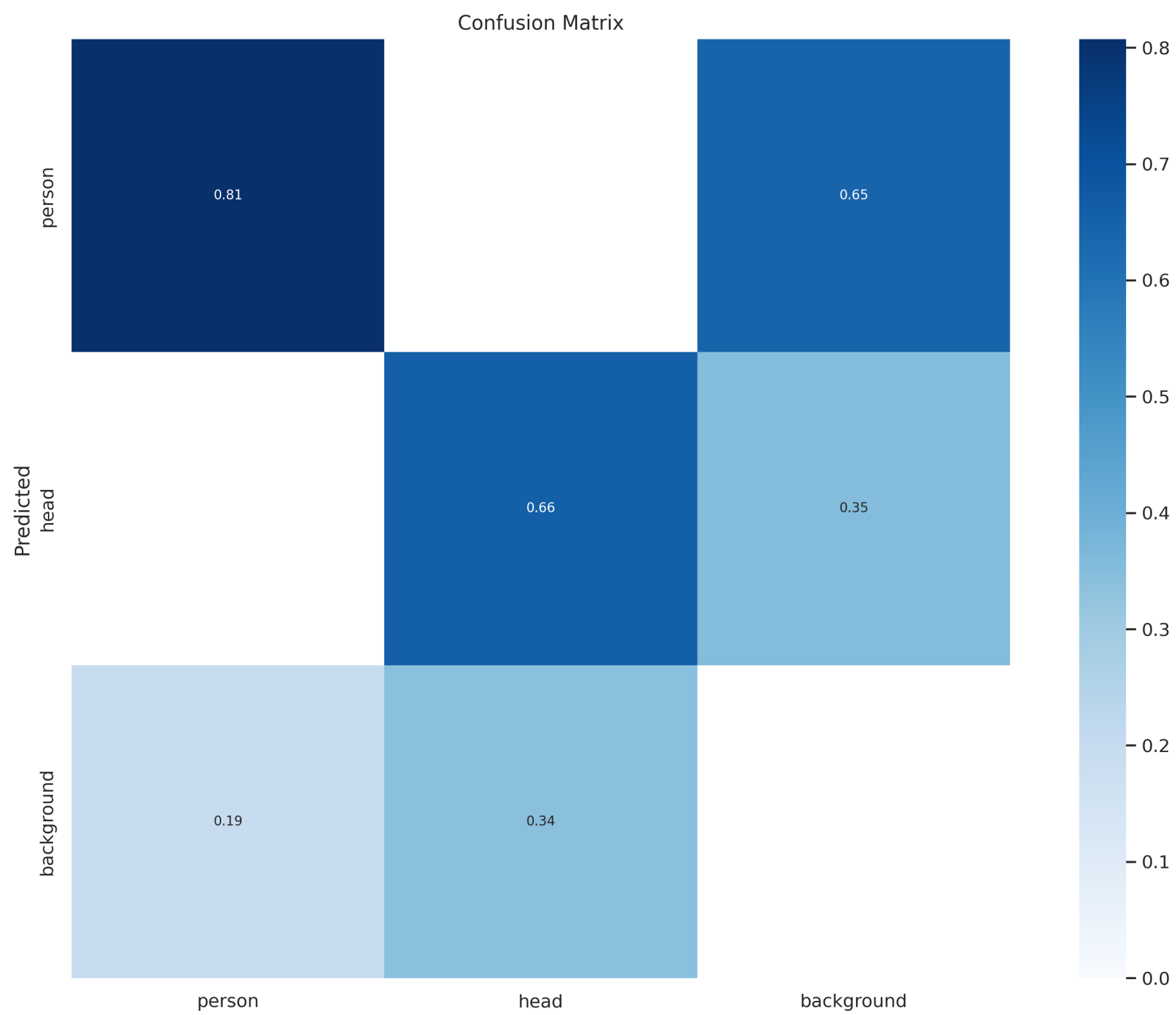
The dataset contained a combination of pascal VOC and crowd dataset. Of the 9523 images, 80 percent of the images were used for training 20 percent for testing.



## Results and discussion

### Confusion Matrix

A confusion matrix is a table used to evaluate the performance of a classification model. It compares the predicted labels of a model with the true labels to measure its accuracy. The true positive value of the person parameter is 0.81, which means that the model correctly identified 81% of the actual person instances.



## Conclusions

Object detection is an important computer vision task that involves identifying objects of interest in an image or video stream and localizing them with bounding boxes. It is a challenging task, as it requires both accurate classification and precise localization of objects in the image. By increasing the quantity of training data and fine-tuning the model parameters, the model's accuracy may be increased even more.



## References

- [1] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in CVPR, vol. 1. IEEE, 2001, pp. 1-1.
- [2] P. Viola and M. J. Jones, "Robust real-time face detection," International journal of computer vision, vol. 57, no. 2, pp. 137-154, 2004.
- [3] J. R. Uijlings, K. E. Van De Sande, T. Gevers, and A. W. Smeulders, "Selective search for object recognition," International journal of computer vision, vol. 104, no. 2, pp. 154-171, 2013.
- [4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in CVPR, 2016, pp. 779-788.
- [5] L. V. G. K. I. W. and W. Z. Mark Everingham, "The PASCAL Visual Object Classes (VOC) Challenge," in International Journal of Computer Vision, vol. 88, no. 2, pp. 303-338, Jun. 2010, doi: 10.1007/s11263-009-0275-4.