

**CSE 541 Computer Vision**

**Project 3: Enhanced Deployment of YOLO V8 on Nvidia Jetson Nano through Model**

**Compression Techniques**

**Group: Visionary minds**

**Team Members**

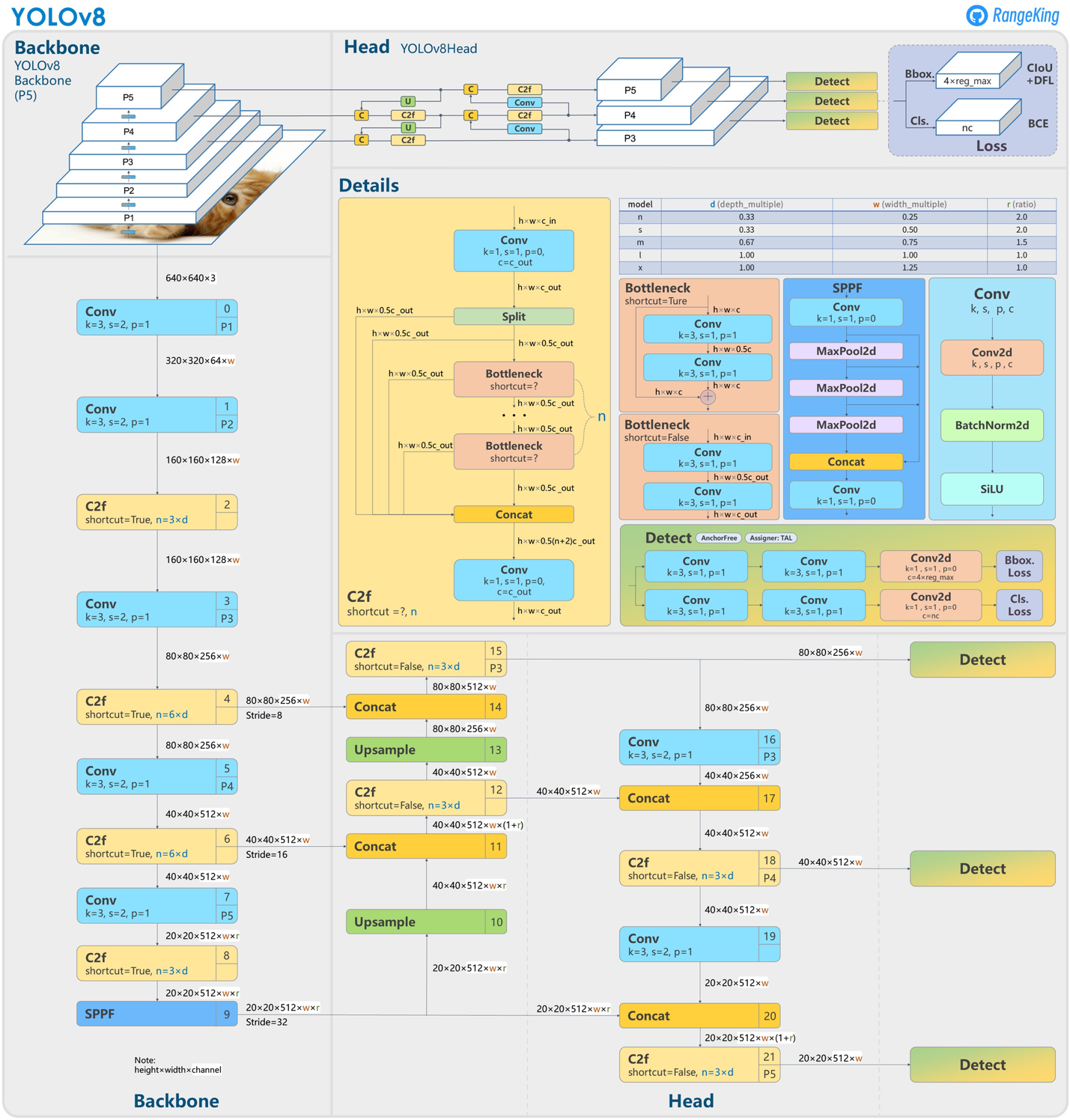
|  |  |
| --- | --- |
| Enrollment Number | Name |
| AU2140154 | Harsh Loriya |
| AU2140179 | Urmil Jagad |
| AU2140205 | Rahul Patel |

**Week Five: Understanding the Yolo v8 Architecture briefly and finding trained models on Yolo v8 that can be utilized in the project.**

**Brief Understanding of the model.**

* YOLOv8, short for “You Only Look Once version 8,” is a powerful object detection model used in computer vision tasks. Imagine it as a smart detective that can quickly spot objects in images or videos.
* Backbone: YOLOv8’s backbone is like the foundation of a building. It uses a modified version of the CSPDarknet53 architecture, which consists of 53 convolutional layers. These layers extract essential features from the input image, like identifying edges, textures, and shapes.
* Cross-Stage Partial Connections: Think of these connections as bridges in a city. They allow smooth communication between different layers in the backbone. Just like bridges connect neighbourhoods, these connections help information flow efficiently.
* Head: The head of YOLOv8 follows the backbone. It’s the decision-making centre. Here, multiple convolutional layers and fully connected layers work together. Their job? Predicting object locations (where things are) and object classes (what those things are). It’s like the detective’s brain making sense of clues.
* Anchor-Free Split Ultralights Head: Instead of traditional anchors (like fixed-sized boxes), YOLOv8 splits the detection process into smaller chunks. This clever approach makes it more accurate and efficient. Imagine breaking down a big puzzle into smaller pieces—it’s easier to solve!
* Performance: YOLOv8 strikes a balance between accuracy and speed. It’s great for real-time tasks, whether you’re spotting cats, cars, or anything else.

**Image of Yolo v8 architecture.**



<https://blog.roboflow.com/whats-new-in-yolov8/>

**Different model in Yolo v8**

1. **COCO Dataset**:
   * Purpose: The COCO (Common Objects in Context) dataset is a widely used benchmark for object detection and segmentation.
   * Data Used: It contains 1.7 million training images and 42,000 validation images, covering diverse object categories.
   * What It Does: YOLOv8 models pre-trained on COCO excel at detecting and segmenting objects in complex scenes.
2. **SKU-110K**:
   * **Purpose**: Focused on dense object detection in retail environments.
   * **Data Used**: Contains over **11,000** images with **1.7 million** bounding boxes.
   * **What It Does**: YOLOv8 models pre-trained on SKU-110K are adept at detecting objects in crowded retail scenes.
3. **Pothole Detection Dataset**:
   * **Purpose**: Specifically designed for detecting potholes in road surfaces.
   * **Data Used**: This dataset comprises more than **7,000** images from various sources, including research papers, YouTube videos, and manually annotated images.
   * **What It Does**: YOLOv8 models trained on this dataset can identify potholes, aiding road maintenance and safety.

References:

* Ultralytics. (2024, February 3). *Ultralytics YOLOV8 Modes*. Ultralytics YOLOv8 Docs. https://docs.ultralytics.com/modes/
* Solawetz, J. (2024, February 19). *What is YOLOv8? The Ultimate Guide. [2024]*. Roboflow Blog. https://blog.roboflow.com/whats-new-in-yolov8/
* Kundu, R. (2023, April 20). YOLO: Algorithm for Object Detection Explained [+Examples]. *V7*. https://www.v7labs.com/blog/yolo-object-detection#:~:text=does%20YOLO%20work%3F-,YOLO%20Architecture,of%20YOLO%20is%20shown%20below.