



# Ahmedabad University

**CSE 641 Computer Vision**

**Weekly Report - Week 5**

**Project Title:** Evaluate Performance of YOLO Family Models in Small Object Detection (HBB)

**Section - 1**

**Group - 02**

<b>Name</b>	<b>Roll No</b>
Romir Bedekar	AU2240254
Hetav Shah	AU2240253

Namra Maniar	AU2240080
--------------	-----------

On finding the map values we observed that the value for YOLO-R model is less compare to YOLO-NAS, YOLOX and PP-YOLO. Thus, there can be certain improvement to optimized that algorithm using the blocks that we saw during the study of different architectures.

To start off we are Implementing the Anchor Box Optimization for VisDrone using K-Means Clustering on the YOLO-R model.

### **Objective**

The goal of this experiment was to improve the performance of YOLOR on the VisDrone dataset by customizing the anchor boxes using K-Means clustering. VisDrone is a dataset with small and dense objects, making default COCO-trained anchors suboptimal.

### **Methodology**

#### **1. Dataset Preparation**

We extracted width and height of all ground-truth bounding boxes from VisDrone .txt annotation files. Each box was represented as a (w, h) pair.

#### **2. K-Means Clustering**

We applied KMeans from sklearn.cluster with k=12 to find 12 optimal anchor box sizes. The clustering was done in the width-height space, and anchors were sorted by area.

#### **3. Anchor Box Result**

The newly computed anchor boxes from VisDrone dataset were:

12,14, 19,32, 40,24, 36,56, 72,41, 62,92, 113,64, 102,146, 173,97, 271,142, 176,249, 450,256

These anchors capture the dense, small object distribution better than the default anchors derived from MS-COCO.

#### **4. Configuration Update**

We modified the anchor settings in the YOLOR configuration file:

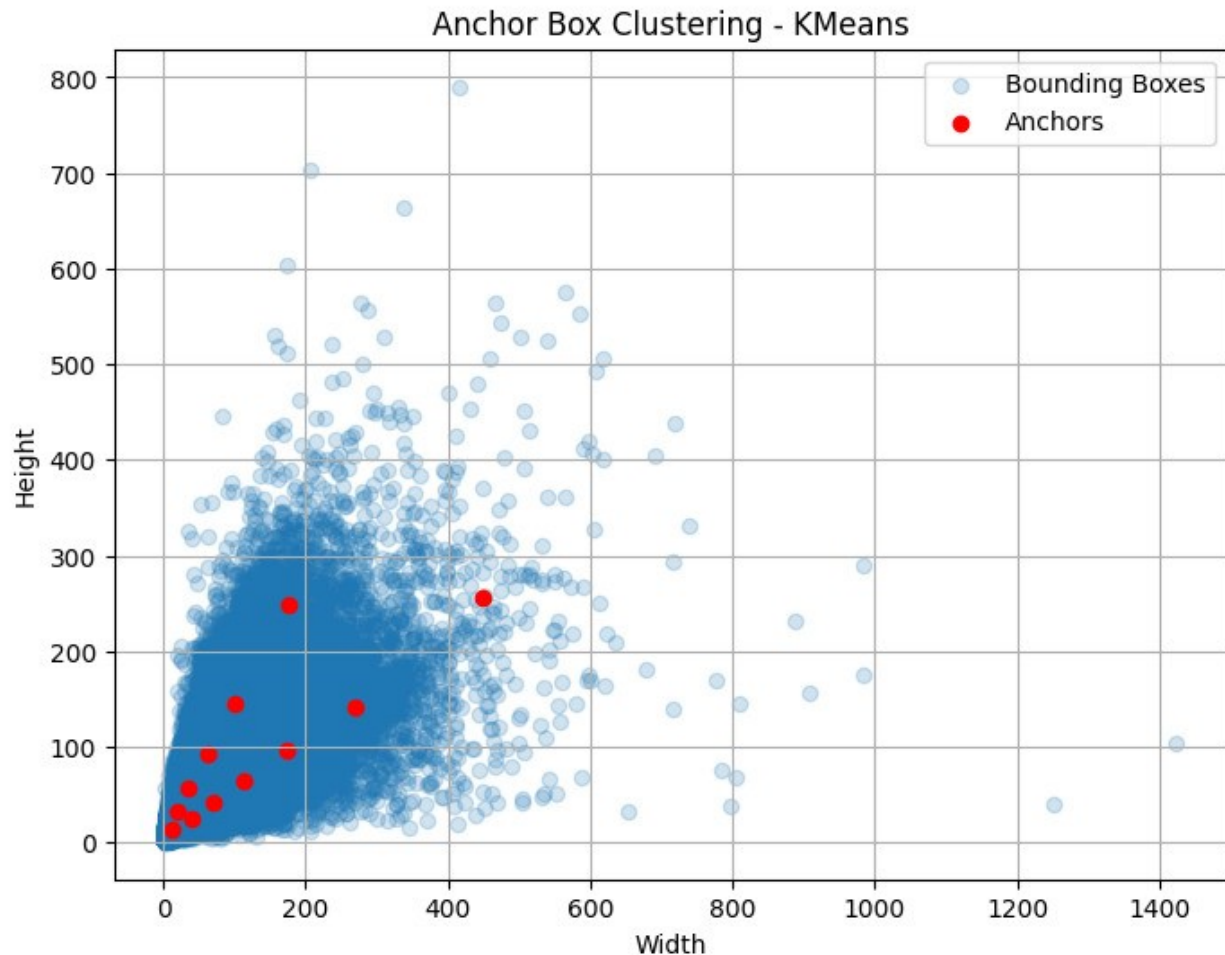
/cfg/yolor\_p6.cfg

Changed the anchor line to:

anchors = 12,14, 19,32, 40,24, 36,56, 72,41, 62,92, 113,64, 102,146, 173,97, 271,142, 176,249, 450,256

## Visualization

We plotted the original bounding boxes against the 12 clustered anchor boxes. The red points represent the optimized anchors, showing a good spread across common object scales in VisDrone.



## Conclusion

Customized anchors from K-Means clustering are more tailored to the VisDrone dataset, especially for smaller objects.

This simple change in `yolor_p6.cfg` can improve recall, precision, and convergence speed.