Ahmedabad University

CSE 641 Computer Vision

Weekly Report - Week 2

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

Section - 1
Group - 02

Name	**Roll No**
Romir Bedekar	AU2240254
Namra Maniar	AU2240080

Hetav Shah	AU2240253

Challenges in Small Object Detection

Detecting small objects using diverse YOLO models presents an array of challenges, which include:

1. **Limited Feature Representation:** The small size means these contain a limited number of pixels, and these pixels were not always easily detectable by the convolutional layers to find meaningful features.

2. **Anchor Box Mismatch:** Default anchor boxes do not always work well with small object sizes. This leads to poor localization and classification.

3. **Low Resolution in Deep Layers:** As images are propagating forward through the deeper layers of the CNN model, the spatial resolution decreases, and this could lead to the loss of fine-grained details necessary for the detection of small-sized objects.

4. **High Overlap in Crowded Scenes:** Since most small objects are usually clustered within crowded scenes, chances for incidence of false positives or missed detections increase.

5. **Background Confusion:** Being small objects, they can blend easily with complex backgrounds, decreasing model confidence and leading to misclassification.

6. **Loss Function Bias:** Commonly defined loss functions may in practice favour larger objects, with a small object contributing less to model learning.

7. **Limited Context Information:** This explains the very little context information related to such small-sized objects within the surroundings, which should enable the model to discriminate between small objects from noise.

YOLO-Based Approaches to Address These Challenges

In order to improve detection of small objects, different YOLO models include specialized techniques:

1. **Multi-Scale Feature Fusion (FPN, PANet, BiFPN):** This process accentuates small-object representation by combining high- and low-resolution features.

2. **Anchor-Free Detection (YOLOX, YOLO-NAS):** This technique abolishes fixed anchor boxes, thus dynamically predicting object locations to improve detection accuracy.

3. **Larger Input Resolutions:** With larger training sizes, better feature retention of small objects can be reaped.

4. **Backbone Improvements (CSPDarknet53, RepVGG, and EfficientNet):** Enhances the ability to extract fine-grained details and improves feature extraction concerning small objects.

5. **Adaptive Spatial Attention Mechanisms:** Helps the model focus on relevant regions in the image which improves detection of small objects.

6. **IoU-Aware Loss Functions:** This modifies the standard loss function to give more weight to small objects and guarantees better localization accuracy.

7. **Training with Augmentations (Mosaic, MixUp, CutMix):** Increases variabilities of the objects in training, which also helps better representations of small objects learned by models.

These strategies together improve the effectiveness of YOLO models in detecting small objects while maintaining real-time inference capability.