Loss Functions in YOLOv5x for Football Analysis System

Your Name

June 2, 2025

Abstract

This report explains the three loss functions used in training a YOLOv5x model for a football analysis system: box loss, classification loss, and distribution focal loss (DFL). These losses guide the model to accurately detect and classify players, referees, and footballs in video frames. We detail their definitions, calculations, and trends from training logs, highlighting their role in the project.

1 Introduction

The YOLOv5x model, fine-tuned on the `football-players-detection-2` dataset, uses three loss functions to optimize object detection: box loss, classification loss, and distribution focal loss (DFL). These losses ensure accurate localization and classification of objects, critical for tracking, team assignment, and movement analysis in the football analysis system.

2 Loss Functions

2.1 Box Loss

Box loss measures the error in predicting bounding box coordinates for objects (e.g., players, footballs). It uses the Complete Intersection over Union (CIoU) loss, defined as:

$$\text{CIoU Loss} = 1 - \text{IoU} + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \tag{1}$$

where $\text{IoU} = \frac{\text{Area of Intersection}}{\text{Area of Union}}$, $\rho^2(b, b^{gt})$ is the squared distance between box centers, c is the diagonal of the smallest enclosing box, and v accounts for aspect ratio consistency. In the training log, box loss decreased from 1.511 (Epoch 1) to 0.9904 (Epoch 18), with a weight of 7.5, indicating improved localization.

2.2 Classification Loss

Classification loss (cls loss) measures the error in predicting object classes (e.g., player, referee, football). It uses Binary Cross-Entropy (BCE) loss:

$$\text{BCE Loss} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)\right] \tag{2}$$

where y_i is the ground truth label and $\hat{y}_i$ is the predicted probability. With 4 classes (`nc=4`), cls loss dropped from 2.015 to 0.5373, with a weight of 0.5, showing reliable classification.

2.3 Distribution Focal Loss (DFL)

DFL loss refines bounding box predictions by modeling coordinates as distributions. It uses Cross-Entropy Loss over discretized bins:

$$\text{DFL Loss} = -\sum_{i} [p_i \log(\hat{p}_i)] \tag{3}$$

where p_i and $\hat{p}_i$ are ground truth and predicted distributions. DFL loss decreased from 0.9084 to 0.8195, with a weight of 1.5, improving precision for small objects like footballs.

3 Training Progress

The training log (18 epochs) shows:

- **Box Loss**: 1.511 $\rightarrow$ 0.9904, indicating better bounding box alignment.
- **Cls Loss**: 2.015 $\rightarrow$ 0.5373, showing improved classification.
- **DFL Loss**: 0.9084 $\rightarrow$ 0.8195, refining box predictions.

These trends correlate with improved metrics (mAP50: 0.231 $\rightarrow$ 0.757), supporting the model's effectiveness for the football analysis system.

4 Project Relevance

The losses enable:

- **Box Loss**: Precise localization for tracking and spatial measurements.
- **Cls Loss**: Accurate classification for team assignment via KMeans.
- **DFL Loss**: Enhanced detection of small objects (e.g., footballs) in crowded scenes.

5 Conclusion

The box, cls, and DFL losses guide YOLOv5x to achieve accurate object detection (mAP50=0.757) for the football analysis system. Continued training and dataset improvements can further reduce losses.