Elephant Project

September 2025

1 Proposed Methodology

In this work, we extend the YOLOv13 object detection framework with two key improvements tailored for imbalanced and occlusion-prone datasets such as animal intrusion monitoring: (i) Focal Loss for classification and (ii) Distance/Complete IoU (DIoU/CIoU) loss for bounding box regression. These modifications aim to mitigate class imbalance, suppress easy negatives, and enhance localization accuracy.

1.1 Focal Loss for Classification

Conventional YOLO models employ Binary Cross-Entropy (BCE) for classification, which assigns equal weight to all samples. In wildlife intrusion datasets, negative (background) anchors significantly outnumber positive (animal) anchors, leading to bias toward background predictions. We replace BCE with Focal Loss [**?**], defined as:

$$\mathcal{L}_{\text{focal}}(p_t) = -\alpha(1 - p_t)^\gamma \log(p_t), \tag{1}$$

where p_t is the predicted probability for the true class, $\alpha \in [0, 1]$ balances positive vs. negative examples, and $\gamma \geq 0$ focuses learning on hard samples. We adopt $\alpha = 0.25$ and $\gamma = 2.0$ in our experiments.

1.2 DIoU/CIoU for Bounding Box Regression

Standard YOLO box regression minimizes IoU or ℓ_1 losses, which may fail when predicted and target boxes have no overlap. We adopt Distance-IoU (DIoU) and Complete-IoU (CIoU) losses [**?**].

Given predicted box B_p and ground truth B_{gt} in (x_1, y_1, x_2, y_2) format:

$$\text{IoU}(B_p, B_{gt}) = \frac{|B_p \cap B_{gt}|}{|B_p \cup B_{gt}|}. \tag{2}$$

The DIoU loss penalizes center distance:

$$\mathcal{L}_{\text{DIoU}} = 1 - \text{IoU} + \frac{\rho^2(\mathbf{b}_p, \mathbf{b}_{gt})}{c^2}, \tag{3}$$

where ρ is the Euclidean distance between box centers, and c is the diagonal length of the smallest enclosing box.

The CIoU loss further considers aspect ratio consistency:

$$\mathcal{L}_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(\mathbf{b}_p, \mathbf{b}_{gt})}{c^2} + \alpha v, \tag{4}$$

with

$$v = \frac{4}{\pi^2}\left(\arctan\frac{w_{gt}}{h_{gt}} - \arctan\frac{w_p}{h_p}\right)^2, \quad \alpha = \frac{v}{(1 - \text{IoU}) + v}. \tag{5}$$

1.3 Overall Loss Function

The final training objective is a weighted sum:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{cls}}\mathcal{L}_{\text{focal}} + \lambda_{\text{box}}\mathcal{L}_{\text{CIoU/DIoU}} + \lambda_{\text{obj}}\mathcal{L}_{\text{obj}}, \tag{6}$$

where $\mathcal{L}_{\text{obj}}$ is objectness loss (BCE), and $\lambda.$ are balancing weights. In practice, $\lambda_{\text{box}} = 5.0$, $\lambda_{\text{cls}} = 1.0$, and $\lambda_{\text{obj}} = 1.0$.

1.4 Algorithmic Description

Algorithm 1 summarizes the modified YOLOv13 training loop.

Algorithm 1 YOLOv13 Training with Focal + CIoU/DIoU Losses

Require: Training images I, annotations Y

Initialize YOLOv13 model parameters θ
for each training epoch **do**
    for mini-batch (I_b, Y_b) **do**
        Predict $\hat{B}, \hat{C}, \hat{O} \leftarrow f_\theta(I_b)$
        Compute classification loss: $\mathcal{L}_{cls} \leftarrow \text{FocalLoss}(\hat{C}, Y_b^{cls})$
        Compute regression loss: $\mathcal{L}_{box} \leftarrow \text{CIoU/DIoU}(\hat{B}, Y_b^{box})$
        Compute objectness loss: $\mathcal{L}_{obj} \leftarrow \text{BCE}(\hat{O}, Y_b^{obj})$
        Combine: $\mathcal{L}_{total} \leftarrow \lambda_{cls}\mathcal{L}_{cls} + \lambda_{box}\mathcal{L}_{box} + \lambda_{obj}\mathcal{L}_{obj}$
        Update $\theta \leftarrow \theta - \eta\nabla_\theta\mathcal{L}_{total}$
    end for
end for