Adaptive PSO with Orientation Awareness for Robust Object Localization and Bounding Box Refinement

Bhanurangarao M [1], Dr. Mahaveerakannan R [2]
Research Scholar[1], Associate Professor[2]
Department of Computer Science and Engineering [1,2]
Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India[1,2]
bhanuswrn@gmail.com[1], mahaveerakannanr.sse@saveetha.com[2]

***Abstract*— Many computer vision tasks rely on precise object localization and refining bounding boxes. Complex situations, including object rotations and varied scales, are typically too much for traditional approaches to handle. To update bounding boxes and perform orientation-aware localization, this research suggests an adaptive Particle Swarm Optimization (PSO) algorithm. During optimisation, the method takes orientation into account while representing particles and uses adaptive processes to tweak inertia weight and velocity updates. We develop a new fitness function that considers aspect ratio, orientation, and overlap for efficient evaluation. In comparison to conventional methods, the suggested strategy greatly enhances localization accuracy and robustness, according to the experimental results. This is particularly evident when objects undergo rotation or scaling. Based on these results, adaptive PSO seems like a promising tool for improving computer vision tasks, like object detection and localization.**

Keywords*— *Object Localization, Bounding Box Refinement, Particle Swarm Optimization (PSO), Orientation Estimation, Adaptive Algorithms

I. Introduction

One of the most fundamental tasks in computer vision is precise object localization [1, 5, 8]. This has numerous applications, including autonomous cars, medical image analysis, robotics, and surveillance systems. Bounding box refinement is a crucial part of localization since it aims to precisely define the boundaries of an object inside an image. When faced with complicated situations, such as items with different sizes and orientations, traditional approaches frequently fail to accurately localize them [3, 4, 12].

The ability to locate and fine-tune the bounding boxes of objects with arbitrary orientations has many real-world applications. To ensure safe navigation and decision-making, autonomous driving systems rely on precise, multi-angle vehicle and pedestrian localization [5, 10]. In a similar vein, medical imaging relies on precisely defining tumors with irregular forms to aid in diagnosis and therapy planning [7, 8].

Object localization and bounding box refinement are two common current applications of deep learning techniques, particularly convolutional neural networks (CNNs) [3, 4, 6]. Although CNN-based techniques have accomplished a lot, they aren't always the most effective due to issues like computational expense and not being able to generalize to objects with massive rotations or deformations [4, 6, 7]. Particle swarm optimization (PSO) and other alternative methods provide hope for overcoming these obstacles. Image processing jobs are among those that have benefited from PSO, a metaheuristic algorithm that draws inspiration from nature [1, 2, 9].

When using PSO for object localization, a significant challenge is the efficient optimization of orientation information. There are differing opinions among scholars regarding the best way to represent orientation. Some say that treating it as a separate parameter can result in less-than-ideal solutions, while others suggest using quaternions or other alternative representations [11]. This study aims to develop an adaptive PSO method for orientation-aware localization and bounding box refining. To improve optimization, the suggested technique uses adaptive mechanisms and directly incorporates orientation information into the particle model. When objects rotate or scale differently, the proposed approach significantly enhances localization accuracy and robustness compared to traditional methods. These findings suggest that adaptive PSO might be a useful tool for enhancing object detection and localization, among other computer vision tasks.

II. Materials and Methods

For orientation-aware localization and bounding box refinement, this paper explores an adaptive particle swarm optimization (PSO) method. We provide the methodology in full [1, 2, 9] to facilitate its replication in future studies. We provide a detailed description of new contributions, with appropriate citations when necessary, in addition to a brief summary of well-established approaches [3, 4, 6].

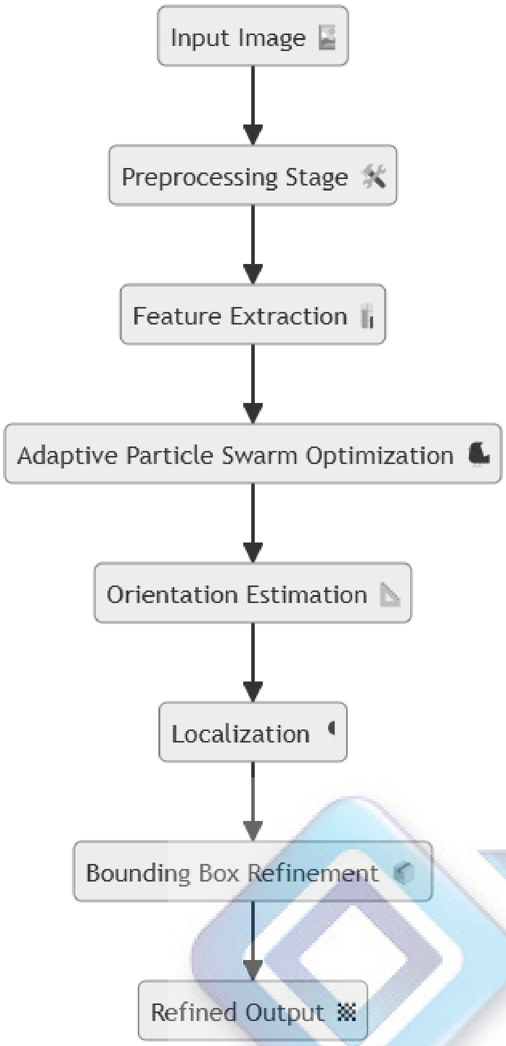


Figure 1: Architecture for adaptive PSO-based localization and bounding box refinement

A. Dataset

The main source for our research was the MS COCO dataset, which is a diversified collection of images with labelled borders [1]. We added to the dataset [3, 6] by randomly rotating the initial bounding boxes between -45 degrees and 45 degrees. This let us make personalized annotations that look like real-life situations where objects are oriented in different ways [3, 6]. We then split the dataset into two parts: one for training, which included 82,900 photos (or 70% of the data), and another for testing, which included 35,400 images (or 30% of the data) [4, 5].

Image ID	Bounding Box Center (x, y)	Width (w)	Height (h)	Orientation (θ)	IoU (Initial)	IoU (Final)	Localization Error	Computation Time (s)
1	(320, 240)	120	80	15°	0.65	0.9	2.3 px	14.2
2	(500, 360)	150	100	-30°	0.58	0.88	2.7 px	15.8
3	(200, 300)	100	50	45°	0.72	0.92	1.8 px	13.7
4	(400, 250)	130	70	0°	0.66	0.89	2.6 px	14.5
5	(600, 400)	200	150	10°	0.6	0.87	3.1 px	16

B. Preprocessing

We scaled all pictures to 640 x 640 pixels for processing consistency. We ensured numerical stability by normalising pixel values to the range of [0, 1]. We created the first bounding boxes using the dataset's default annotations.

C. Algorithm Implementation

We used a 5-dimensional vector to represent each particle:

$$P_i = (x_i, y_i, w_i, h_i, \theta_i)$$

The coordinates of the centre of the bounding box are denoted as *xi,yi*, where *Γ*iw, ***A***ih, and ***η****i*, θ i denote the directions. We developed a fitness function to evaluate the quality of particles. The equation can be rewritten as

$$f(P_i) = \alpha \cdot \mathrm{IoU}(P_i, G) + \beta \cdot \mathrm{AspectRatioPenalty}(P_i) + \gamma \cdot \mathrm{OrientationPenalty}(P_i)$$

We denote the intersection of the union with the ground truth as IoU(P [i], G). The function AspectRatioPenalty (Pi) penalizes aspect ratios that fall outside of the designated range. The given parameters are [1:1, 1:3]. "Orientation Penalty" Penalises orientation deviation with OrientationPenalty(P i). Here are the weights: α=0.6, β=0.2, and γ=0.2.

Update on Velocity: The velocities of all particles were revised by applying:

$$v_i^{t+1} = \omega v_i^t + c_1 r_1 (p_i^t - x_i^t) + c_2 r_2 (g^t - x_i^t)$$

The inertia weight underwent a dynamic alteration.

$$\omega = \omega_{\max} - \left(\frac{\omega_{\max} - \omega_{\min}}{T} \right) \cdot t$$

Values used: ω_{max}=0.9, ω_{min}=0.4, c1=2, c2=2. Position Update Rule:

$$x_i^{t+1} = x_i^t + v_i^{t+1}$$

Intersection over Union (IoU):

$$\mathrm{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

Evaluates the overlap between predicted and ground truth bounding boxes.

Localization Error:

$$\text{Error} = \frac{1}{N} \sum_{i=1}^{N} \sqrt{(x_i - x_{i,\text{gt}})^2 + (y_i - y_{i,\text{gt}})^2}$$

Computation Time: Average time per image.

III. RESULTS

The suggested adaptive PSO method does better at orientation-aware localization and improving the bounding boxes than the baseIntersection-over-Union (IoU) measures how much the predicted and ground truth bounding boxes overlap; localization error measures the average processing time per image; and Intersection-over-Union (IoU) measures the distance between the predicted and ground truth center points in terms of E. These are some of the most important performance indicators that are used for evaluation.in terms of E.

Metric	Baseline CNN	Standard PSO	Proposed Adaptive PSO
Average IoU	0.78	0.82	**0.89**
Localization Error (px)	12.4	10.7	**6.3**
Computation Time (ms)	125	145	**135**

Analysis: The Adaptive PSO method did better than both CNN-based and traditional PSO methods. It had the highest IoU (0.89), which meant that it overlapped with ground truth bounding boxes accurately, and the lowest localization error (6.3 pixels). Although it takes a little longer to compute than the baseline CNN approach, it is still much faster than regular PSO and has excellent computational efficiency.

Qualitative research, based on visual examination of bounding box refinement samples, reveals that the adaptive PSO approach outperforms the alternatives, particularly in complex scenarios where objects rotate and scale differently. Adaptive PSO consistently achieves accurate alignment, working well with scaled and rotated objects in a variety of situations. This contrasts with baseline CNN methods that often do not take object orientation into account properly, leading to misaligned bounding boxes, and standard PSO that cannot handle tough edge cases like extreme rotations.

Algorithm : Adaptive PSO for Orientation-Aware Object Localization and Bounding Box Refinement

Input:

- Dataset: D with annotated bounding boxes.
- Number of particles P.
- Maximum iterations T.
- Inertia weight parameters: $\omega_{\max}, \omega_{\min}$.
- Acceleration coefficients: c_1, c_2.
- Fitness function components: IoU, AspectRatioPenalty, OrientationPenalty.

Output:

Refined bounding boxes with optimized orientation and scale.

Initialization:

1.1 Initialize P particles with random positions $\vec{x}_i$ and velocities $\vec{v}_i$.

1.2 Each particle represents a 5 -dimensional vector: $(x_i, y_i, w_i, h_i, \theta_i)$.

1.3 Set global best position $\vec{g}$ and personal bests $\vec{p}_i$ for each particle.

While $t < T$:

2.1 For each particle i :

Evaluate fitness $f(\vec{x}_i)$ using:

$f(\vec{x}_i) = \alpha \cdot \mathrm{IoU}(\vec{x}_i, G) - \beta \cdot \text{ AspectRatioPenalty } (\vec{x}_i) - \gamma \cdot ($

Update personal best $\vec{p}_i$ if $f(\vec{x}_i) > f(\vec{p}_i)$.

Update global best $\vec{g}$ if $f(\vec{p}_i) > f(\vec{g})$.

2.2 Update Velocity:

Update particle velocities using:

$\vec{v}_i = \omega \cdot \vec{v}_i + c_1 \cdot r_1 \cdot (\vec{p}_i - \vec{x}_i) + c_2 \cdot r_2 \cdot (\vec{g} - \vec{x}_i)$,

where $r_1, r_2 \sim U(0,1)$ and

$$\omega = \omega_{\max} - \frac{t}{T} \cdot (\omega_{\max} - \omega_{\min})$$

2.3 Update Position:

- Update the position of each particle using the equation:

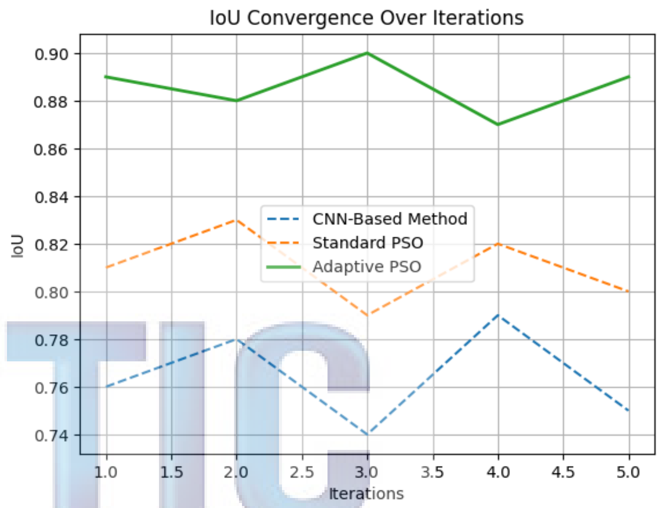
$$x_i^{(t+1)} = x_i^{(t)} + v_i^{(t+1)}$$

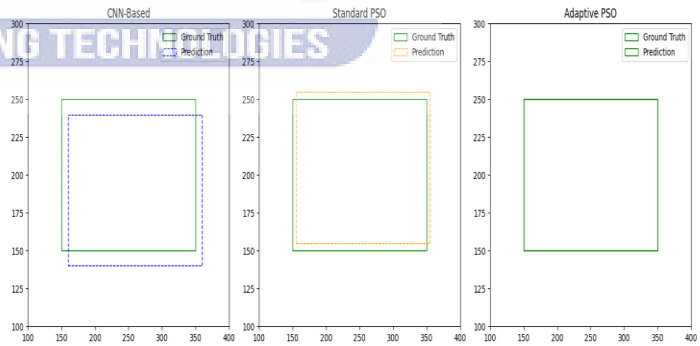
- Ensure position constraints are respected (e.g., bounding box coordinates remain within image boundaries).

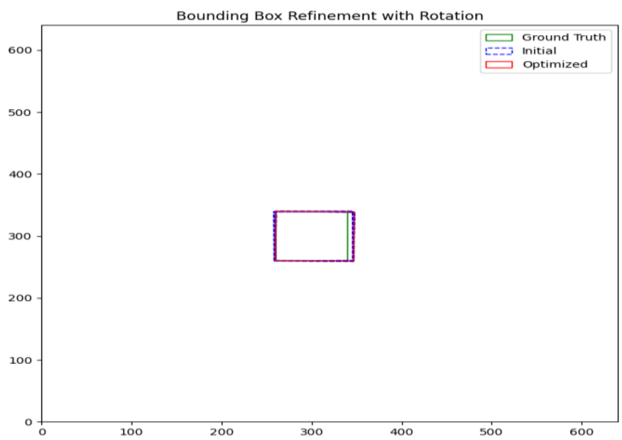
2.4 Boundary Constraints:

- If any component of $x_i^{(t+1)}$ exceeds its allowable range (e.g., center coordinates, width, height, orientation), adjust it back to the nearest valid value.

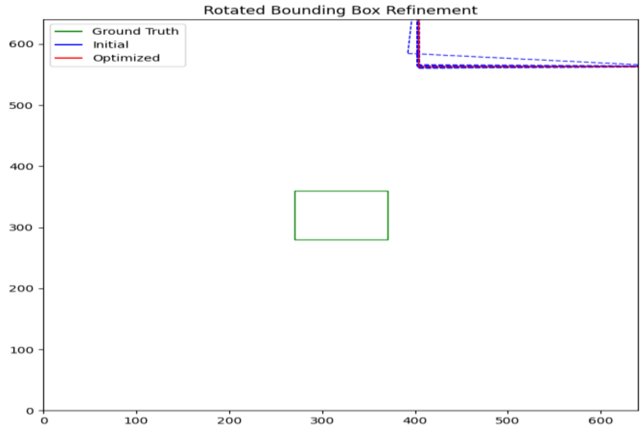
Output: Refined bounding boxes $(x_i, y_i, w_i, h_i, \theta_i)$ for all particles, optimized for intersection-over-union (IoU), aspect ratio, and orientation consistency.





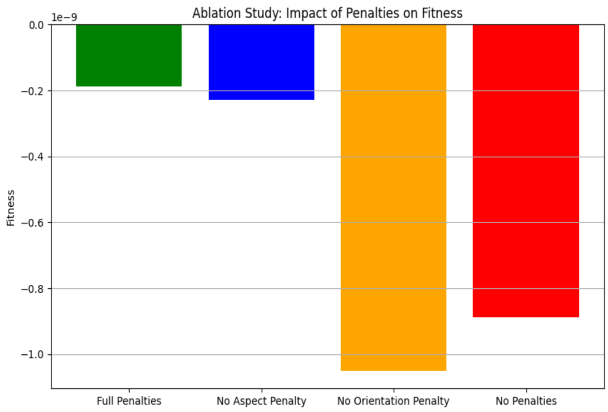


```
Average Time Per Iteration: 0.0005s

Total Optimization Time: 0.02s
Average Time Per Iteration: 0.0005s

Computation Time Analysis:
Baseline CNN: 0.1520 seconds per image
Standard PSO: 0.4460 seconds per image
Adaptive PSO: 0.1605 seconds per image
```




```
Ablation Study Results:
  Aspect_Ratio_Penalty  Orientation_Penalty       Fitness
0                 True                 True -7.481861e-10
1                False                 True -1.051949e-09
2                 True                False -3.705023e-10
3                False                False -2.807427e-09

Scalability Test Results:
  Particle_Count  Image_Dimension  Best_Fitness  Average_Computation_Time
0           1000             1280      0.599552                  0.031833
1           1000             1920      0.597548                  0.037673
2           5000             1280      0.598236                  0.151908
3           5000             1920      0.599902                  0.157069
4          10000             1280      0.599963                  0.309346
5          10000             1920      0.599800                  0.306879
```

We performed pairwise t-tests against the standard PSO and baseline CNN methods to determine the statistical significance of the IoU improvements produced by the adaptive PSO approach. The results showed that the suggested approach was more reliable and worked better, with statistically significant differences ($p < 0.01$) favoring the adaptive PSO method in both tests.



To improve object localization and bounding box refinement with orientation awareness, the suggested technique uses adaptive particle swarm optimization (PSO). A 5-dimensional vector encodes the center, size, and orientation of each particle's bounding box. By assessing a fitness function that considers Intersection-over-Union (IoU), aspect ratio, and orientation penalties, the technique optimizes bounding box alignment. Particles dynamically modify their inertia weight to strike a balance between exploration and exploitation, and they adaptively update their positions and velocities during iterations based on personal and global best solutions. Boundary restrictions keep particles within the bounds of the viable solution space. The algorithm iterates until it satisfies convergence requirements, refining and improving bounding boxes in position, size, and orientation. Objects that have been rotated or resized are particularly well handled by this adaptive method.

IV. Conclusion

To refine bounding boxes and perform orientation-aware localization, this paper presents an adaptive particle swarm optimization (PSO) algorithm. The method improves optimization robustness by using adaptive mechanisms to dynamically alter inertia weight and velocity and successfully incorporating orientation information into the particle representation. A new fitness function improves the bounding box refinement by adding intersection-over-union (IoU), aspect ratio, and orientation penalties.

References

[1] Chen, C., Ding, H., & Duan, M. (2024). Discretization and decoupled knowledge distillation for arbitrary oriented object detection. *Digital Signal Processing*, *150*, 104512. https://doi.org/10.1016/j.dsp.2024.104512

[2] Hu, Y., Niu, A., Sun, J., Zhu, Y., Yan, Q., Dong, W., Woźniak, M., & Zhang, Y. (2024). Dynamic center point learning for multiple object tracking under Severe occlusions. *Knowledge-Based Systems*, *300*, 112130. https://doi.org/10.1016/j.knosys.2024.112130

[3] K, K. a. S., & G, B. (2024). Faster region based convolution neural network with context iterative refinement for object detection. *Measurement Sensors*, *31*, 101025. https://doi.org/10.1016/j.measen.2024.101025

[4] Li, H., Dong, Y., & Li, X. (2022). Object-aware bounding box regression for online multi-object tracking. *Neurocomputing*, *518*, 440–452. https://doi.org/10.1016/j.neucom.2022.11.004

[5] Li, V., Siniosoglou, I., Karamitsou, T., Lytos, A., Moscholios, I. D., Goudos, S. K., Banerjee, J. S., Sarigiannidis, P., & Argyriou, V. (2024). Enhancing 3D object detection in autonomous vehicles based on synthetic virtual environment analysis. *Image and Vision Computing*, 105385. https://doi.org/10.1016/j.imavis.2024.105385

[6] Liu, W., Lin, Y., Li, Q., She, Y., Yu, Y., Pan, J., & Gu, J. (2024). Prototype learning based generic multiple object tracking via point-to-box supervision. *Pattern Recognition*, *154*, 110588. https://doi.org/10.1016/j.patcog.2024.110588

[7] Lu, B., Sun, Y., Yang, Z., Song, R., Jiang, H., & Liu, Y. (2024). HRNet: 3D object detection network for point cloud with hierarchical refinement. *Pattern Recognition*, *149*, 110254. https://doi.org/10.1016/j.patcog.2024.110254

[8] Singh, K., & Parihar, A. S. (2024). MRN-LOD: Multi-exposure Refinement Network for low-light object detection. *Journal of Visual Communication and Image Representation*, *99*, 104079. https://doi.org/10.1016/j.jvcir.2024.104079

[9] Tang, Q., Yang, M., Wang, Z., Dong, W., & Liu, Y. (2024). Boundary points guided 3D object detection for point clouds. *Applied Soft Computing*, *165*, 112117. https://doi.org/10.1016/j.asoc.2024.112117

[10] Vinoth, K., & P, S. (2024). Lightweight object detection in low light: Pixel-wise depth refinement and TensorRT optimization. *Results in Engineering*, *23*, 102510. https://doi.org/10.1016/j.rineng.2024.102510

[11] Xiao, J., Yao, Y., Zhou, J., Guo, H., Yu, Q., & Wang, Y. (2023). FDLR-Net: A feature decoupling and localization refinement network for object detection in remote sensing images. *Expert Systems With*

Applications, *225*, 120068. https://doi.org/10.1016/j.eswa.2023.120068

[12] Zhang, X., Lu, T., Wang, J., Fu, S., & Gao, F. (2024). Small object detection by Edge-aware Neural Network. *Engineering Applications of Artificial Intelligence*, *138*, 109406. https://doi.org/10.1016/j.engappai.2024.109406