International Journal of
Applied Data Science in Engineering and Health
https://ijadseh.com

Enhanced Pedestrian Tracking using Grey Wolf Optimizer

Negar Raeisidehkordi[1], Farshad Kumarci[2], Saeid Raeisidehkordi[3], Alireza Hadipour[4]

[1] *Department of Computer Engineering, Shahrekord Branch, Islamic Azad University, Iran*

(Email: Negarreisi15@gmail.com)

[2] *Department of Computer Engineering, Shahrekord Branch, Islamic Azad University, Iran*

[3] *Department of Humanities, Dehaghan Branch, Islamic Azad University, Iran*

[4] *Department of Electric Engineering, Najaf Abad Branch, Islamic Azad University, Iran*

Received date: October 5, 2024; Accepted date: January 3, 2025

Abstract

Object tracking is a challenging task due to variations in lighting, occlusions, and interactions with other objects, which can impact the appearance model and cause tracking failures. Accurate feature selection is critical for building robust models that can handle noise and sudden changes without drifting. This study introduces an intelligent tracking algorithm, GWO-OF (Grey Wolf Optimizer with Optical Flow), which leverages the hierarchical leadership and hunting behavior of grey wolves for feature optimization and integrates optical flow for motion tracking. The proposed method was evaluated using standard tracking metrics, achieving superior results across all benchmarks. It recorded the highest accuracy (MOTP: 80.53%), precision (MOTA: 48.35%), and F1-Score (73.54%), outperforming state-of-the-art methods such as Memetic-Adaboost and SVM. The GWO-OF algorithm demonstrated robustness in handling occlusions and noise, making it highly effective for real-world tracking applications. These findings highlight the potential of the proposed method as a reliable and precise solution for object tracking.

Keywords*:* Object Tracking; Grey Wolf Optimizer (GWO); Optical Flow.

Introduction

Pedestrian tracking is identifying and continuously following individuals across a sequence of video frames. This task is critical in numerous areas, including video surveillance, crowd monitoring, traffic management, and sports analytics. The ability to track pedestrians enables the detection of unusual activities, facilitates real-time decision-making, and provides insights into human behavior. For instance, in video surveillance, pedestrian tracking helps identify unauthorized access, monitor crowd movements, and enhance public safety. Similarly, in sports, tracking athletes provide valuable performance analysis and tactical planning data. The importance of pedestrian tracking extends to urban planning, where understanding pedestrian flow can inform the design of more efficient transportation systems and public spaces. Given its wide-ranging applications, pedestrian tracking is a vital area of research in computer vision and artificial intelligence [1,2].

Despite its significance, pedestrian tracking presents considerable challenges, especially in dynamic and complex environments. One major challenge is the variability in environmental conditions, such as changes in background, lighting, and camera movement. These factors can significantly affect the accuracy of tracking systems. Occlusion is another common problem, where pedestrians are partially or fully obscured by other objects, such as vehicles, trees, or other people, making it difficult to maintain consistent tracking. Additionally, poses and appearance changes pose significant hurdles, as individuals may change their orientation, clothing, or behavior during a tracking sequence. Crowded scenes introduce further complexity, requiring systems to distinguish between multiple individuals with similar appearances while managing a high volume of data. Lastly, real-time (online) pedestrian tracking is computationally demanding, requiring instant processing and analyzing video data. These challenges necessitate the development of adaptive tracking methodologies capable of handling noise, occlusions, and diverse scenarios effectively [3,4].

Metaheuristic algorithms are optimization techniques inspired by natural processes, such as evolutionary biology, animal behavior, and physical phenomena. These algorithms are particularly well-suited for solving complex, nonlinear, and high-dimensional optimization problems where traditional methods may falter. In the context of pedestrian tracking, metaheuristic algorithms offer a powerful framework for addressing challenges related to feature selection, model optimization, and motion prediction [5-8]. Feature selection is one of the most critical aspects of pedestrian tracking, as it determines the ability of the system to distinguish between targets and background noise. Metaheuristic algorithms, such as Genetic Algorithms (GA) and Particle Swarm Optimization (PSO), have been widely used to identify the most relevant features, improving the accuracy and robustness of tracking systems. Moreover, these algorithms can optimize tracking models by fine-tuning parameters to enhance their adaptability to diverse environments. For example, in scenarios with occlusions or varying lighting conditions, metaheuristics can help refine the model to maintain tracking accuracy. Another significant application of metaheuristic algorithms in pedestrian tracking is motion estimation. By leveraging algorithms like Ant Colony Optimization (ACO) or Grey Wolf Optimizer (GWO), tracking systems can predict the future positions of pedestrians more accurately, even in complex and dynamic environments. This capability is particularly valuable in crowded scenes where interactions between individuals make tracking more challenging. Additionally, the flexibility and adaptability of metaheuristic algorithms make them ideal for real-time applications, where computational efficiency is crucial. The growing use of metaheuristic algorithms in pedestrian tracking underscores their potential to overcome longstanding challenges and enhance system performance. By integrating these techniques, tracking systems can achieve higher accuracy, robustness, and scalability, enabling their application in a broader range of real-world scenarios.

In this study, we leverage the Grey Wolf Optimizer (GWO) [9] algorithm to enhance pedestrian tracking by addressing key challenges such as noise, occlusions, and feature selection. Effective tracking begins with the identification and separation of the moving object from the background, which requires a robust model capable of adapting to dynamic environments. The process of selecting the most relevant features is critical to ensure the model’s ability to track the target accurately under varying conditions. The GWO algorithm, inspired by the social hierarchy and cooperative hunting strategies of grey wolves, is employed for feature selection in this research. By assigning weights to features based on their significance, GWO helps create a precise and reliable target model, ensuring robustness against environmental noise and occlusions.

Figure 1 Few samples from the dataset

Additionally, optical flow is used to handle the motion tracking phase, complementing the feature selection process and enabling efficient tracking in complex scenarios. This combination of GWO and optical flow provides an innovative solution to overcome the inherent challenges of pedestrian tracking.

Methods and Materials

In this section, we present the materials and explain the overall methods. The process begins with data preprocessing, which includes three key steps: understanding the data, cleaning the data, and transforming it to ensure readiness for feature extraction and model training. The Grey Wolf Optimizer (GWO) algorithm is employed for intelligent and dynamic feature selection. In GWO, the hierarchical leadership structure of grey wolves—consisting of alpha, beta, delta, and omega roles—is utilized to model the optimization process. The algorithm iteratively performs three main steps: searching for prey, encircling prey, and attacking prey, to identify and retain the most relevant features for tracking. After identifying the specific target (or unique identifier) using GWO, the optical flow algorithm is applied for motion tracking. This method evaluates multiple key points from the target's motion model, predicts the next position based on the overlap of these points with the previous location, and confirms the target’s movement if a high overlap is detected. This combination of GWO for feature selection and optical flow for tracking ensures accurate and efficient pedestrian tracking, even in complex scenarios.

Dataset

The dataset used in this study is one of the most comprehensive and challenging benchmarks for multi-pedestrian tracking, consisting of 350,000 labeled pedestrian bounding boxes extracted from 250,000 video frames with a resolution of 640×480 pixels [10]. Spanning 137 minutes of video footage and involving information from 1,900 individuals, the dataset provides a rich and diverse collection of scenarios for evaluating tracking methodologies. It is organized into 22 sections, with 11 sections allocated for training and evaluating general pedestrian detectors and the remaining sections designated for multi-pedestrian tracking evaluation. The training dataset includes three primary components: the det folder, which contains general pedestrian detectors with an initial frame and a corresponding .txt file specifying the positions and features of the detectors across frames; the gt folder, which provides precise positions of specific target identifiers; and the img1 folder, which contains all video frames for training purposes. In contrast, the testing dataset comprises the det folder and the img1 folder, used to predict the positions of targets based on general detectors. The accuracy of the system is then assessed by comparing the predicted positions with the actual positions from the ground truth (gt) to calculate error rates. This systematic structure allows the dataset to address the complexities of real-world scenarios, including occlusions, dynamic environments, and multi-pedestrian interactions, making it an excellent benchmark for evaluating the proposed pedestrian tracking approach.

Data Preprocessing

Data preprocessing is a crucial step to ensure the dataset is clean, balanced, and ready for use in model training. In this study, the preprocessing process involves three primary stages: data understanding, data cleaning, and data transformation.

The first step in preprocessing focuses on analyzing the dataset and identifying relationships between selected features. Redundant or irrelevant features are removed during this stage, often with the help of a correlation matrix. This analysis ensures that only the most relevant features are retained, optimizing the model's performance. Another critical aspect is balancing the dataset. Imbalanced datasets, where one class significantly outweighs others, can undermine the effectiveness of model training, leading to biased predictions. To address this, the dataset is balanced so that each class has an equal number of samples, ensuring fairness and reliability in the learning process. The second stage involves cleaning the data by identifying and handling outliers and inconsistent data points. Outliers are extreme values that deviate significantly from the rest of the dataset, often appearing near the boundaries of the data range or contradicting general trends. These outliers may indicate errors in data entry or anomalies that can adversely impact the model's performance. Depending on their nature, outliers are either removed or replaced with the mean value of the corresponding feature to stabilize the dataset. Additionally, inconsistent data points that do not align with the overall patterns are identified and eliminated to enhance the dataset's reliability. The final step in preprocessing is normalizing the data to ensure uniform scaling of all features. In this study, the Min-Max normalization algorithm is used to scale the values of all features to a range between 0 and 1. This step is particularly important for datasets where features have varying scales, as it ensures that all features contribute equally to the model's learning process. During normalization, categorical data, such as letters or class labels, are also converted into numerical values within the same range, making them compatible with machine learning algorithms.

By following these preprocessing steps, the dataset is prepared to facilitate accurate, efficient, and unbiased model training, addressing potential issues that could compromise the performance of the proposed pedestrian tracking system.

Grey Wolf Optimization

The Grey Wolf Optimizer (GWO) is a metaheuristic optimization algorithm inspired by the hunting behavior and social hierarchy of grey wolves. Proposed by Mirjalili et al. in 2014 [9], GWO models the leadership structure of grey wolf packs, categorizing wolves into four hierarchical levels: alpha, beta, delta, and omega. This hierarchy governs the pack's hunting strategy, which includes tracking, encircling, and attacking prey. The algorithm leverages these natural behaviors to balance exploration and exploitation in the search space, making it effective for solving complex optimization problems. The key concepts and algorithm structure are as follows:

1. *Population Initialization:*
   GWO begins by initializing a random population of candidate solutions (wolves) within the search space. Each wolf's position represents a potential solution, and its fitness value corresponds to the objective function being optimized.

2. *Leadership Hierarchy:*
   The wolves are ranked based on their fitness values, with the top three solutions designated as alpha (best solution), beta (second best), and delta (third best). These three wolves guide the search process, while the remaining wolves (omega) follow their directions.

3. *Encircling Prey:*
   Grey wolves encircle their prey during hunting. Mathematically, this behavior is modeled as:

   $$\vec{D} = |\vec{C} \cdot \vec{X}_p - \vec{X}(t)|, \quad \vec{X}(t+1) = \vec{X}_p - \vec{A} \cdot \vec{D}$$

   Where *X(t)* is the current position of the wolf, X_p is the position of the best Solution, *A* and *C* are the coefficient vectors that control the encircling behavior, and *D* is the distance vector between the wolf and prey. The coefficients *A* and *C* are calculated as:

   $$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a}, \quad \vec{C} = 2 \cdot \vec{r}_2$$

   where *r* is a random vector in [0, 1], and *a* decreases linearly from 2 to 0 over iterations, balancing exploration and exploitation.

4. *Hunting and Attacking Prey:*
   The positions of all wolves are updated based on the alpha, beta, and delta wolves' positions, simulating the pack's hunting strategy:

   $$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3}$$

   Where X_1, X_2, and X_3 are updated positions based on the alpha, beta, and delta wolves, respectively. This ensures the wolves converge toward the most promising regions in the search space.

5. *Exploration and Exploitation:*
   GWO balances exploration (searching new areas) and exploitation (refining solutions in promising areas) through parameter *a*. When $|A|>1|$, wolves explore new regions; when $|A|<1$, they exploit the current region to refine the solution.

6. *Termination:*
   The algorithm continues iterating until a stopping criterion is met, such as a maximum number of iterations or a convergence threshold. The position of the alpha wolf at the end of the process represents the optimal or near-optimal solution.

Application in Pedestrian Tracking

In this study, GWO is used to optimize feature selection for pedestrian tracking. The algorithm assigns weights to features based on their significance, enabling the identification of the most relevant features while discarding irrelevant ones. This step improves the robustness of the tracking model against noise and occlusions. The optimized features are then integrated into the pedestrian tracking framework, enhancing accuracy and computational efficiency. By mimicking the cooperative and adaptive behavior of grey wolves, GWO offers an effective solution for tackling the complex challenges of pedestrian tracking.

The Grey Wolf Optimizer (GWO) was implemented in this study to address the challenges of feature selection and target identification in pedestrian tracking. In a typical video frame, both foreground (target) and background elements are present. While most tracking systems rely primarily on foreground information, certain portions of the background can be useful in distinguishing the target from noise. To enhance tracking accuracy, preprocessing was employed to retain relevant background elements while discarding unnecessary regions. A significant challenge in pedestrian tracking lies in constructing a robust model to reliably handle noise and occlusions. The effectiveness of this model depends largely on the selection of relevant features, which is a critical and challenging step. In this study, GWO was used to develop a dynamic and intelligent feature selection algorithm that automatically identifies the most important features for tracking. Implementation of GWO in pedestrian tracking is as follows:

1. *Population Initialization:*
   The process begins with initializing a random population of grey wolves, where each wolf represents a potential solution—a subset of features selected from the dataset. These wolves are categorized into hierarchical groups: alpha (best solution), beta, delta, and omega, reflecting their roles in the optimization process.

2. *Feature Weighting and Selection:*
   Each wolf is assigned weights based on the significance of the features it represents. The alpha wolf, being the most optimal solution at a given iteration, receives the highest weight. Features with weights below a predefined threshold are discarded, while the most significant features are retained. This ensures the model's robustness against noise and occlusions by focusing only on discriminative features.

3. *Population Enhancement:*
   To ensure rapid convergence and maintain diversity, the omega wolves (least fit solutions) are replaced with either higher-performing wolves or randomly selected wolves from the population. This approach ensures that the optimization process consistently moves toward better solutions while avoiding local optima.

4. *Target Identification:*
   Once feature selection is completed, all pedestrians within a frame are identified, and the specific target is distinguished from the general detectors using the optimized feature set. The GWO ensures that the feature selection process is dynamic and adaptive, enabling precise identification of the target even in the presence of multiple objects or distractions.

5. *Target Tracking with Optical Flow:*
   After the target is identified, the optical flow algorithm is used for motion tracking. This algorithm identifies several key points based on the motion model of the target and predicts its next location by analyzing the overlap between these points and the target's previous position. If the overlap surpasses a predefined threshold, the location is confirmed as the target's next position. Otherwise, the system reverts to the training phase to refine the model and feature selection process.

6. *Iterative Optimization:*
   The tracking system iteratively evaluates whether the target has been correctly identified and tracked. If the tracking is accurate, the algorithm terminates; otherwise, it returns to the training phase and continues optimizing until the best tracking performance is achieved.

The proposed system also accommodates scenarios with multiple targets within a single video frame. GWO dynamically identifies specific targets among general detectors, treating each target as a unique identifier. This capability enables robust and efficient multi-target tracking, ensuring accurate identification and tracking even in crowded and complex scenes. By combining the adaptive feature selection of GWO with the predictive capabilities of optical flow, this study presents a comprehensive framework for pedestrian tracking. The integration of these methodologies ensures high accuracy and robustness, addressing key challenges such as noise, occlusions, and dynamic environments effectively.

Experiments

In the initial phase, real-world data and general detectors were used to train the objects targeted for tracking. The proposed algorithm successfully completed the training phase, accurately identifying and learning the objects of interest. Figure 2 illustrates the detection and tracking of objects during the training and testing stage. In the testing phase, the algorithm identifies general detectors in each frame and simultaneously tracks objects frame by frame. The system handles challenges such as noise, occlusions, or instances where objects temporarily exit the video. Figure 3 demonstrates the process of identifying the target object and matching it based on the defined threshold. For example, in this figure, target number 5 is detected in frame 57. To estimate the target's next position, the algorithm analyzes the last ten frames and evaluates the threshold for each frame. If the threshold exceeds 50%, the next position is predicted accurately. This iterative approach ensures robust tracking even in dynamic and challenging environments.

Results

In this section, tracking accuracy was evaluated using the provided data and the MOTP (Multiple Object Tracking Precision) metric.

$$MOTP = \frac{\sum_t d_t^i}{\sum_t c_t}$$

The MOTP formula calculates the accuracy based on the distance between the actual trajectory and the predicted trajectory of object i in frame t, divided by the total number of trajectories in all frames.

Figure 2 Detection and object tracking in training (left) and testing phase (right)

```
target 5: bb overlaps 1.00 0.91 0.91 0.95 0.86 0.90 0.96 0.87 0.91 0.92

ftracked: 1.00 0.92
anchor 1
frame 57, state 2

target 5: frame ids 31 54 52 51 50 56 53 55 49 57
target 5: medFB 2.12 0.17 0.35 0.33 0.36 0.10 0.34 0.10 0.29 0.04
target 5: medFB left 1.35 0.18 0.35 0.33 0.31 0.08 0.25 0.13 0.32 0.03
target 5: medFB right 3.15 0.17 0.35 0.35 0.42 0.10 0.45 0.10 0.29 0.06
target 5: medFB up 4.72 0.12 0.26 0.26 0.25 0.03 0.16 0.03 0.16 0.02
target 5: medFB down 1.28 1.00 0.41 0.73 2.76 0.38 0.83 0.30 1.65 0.07
target 5: medNCC 0.70 0.87 0.86 0.82 0.77 0.95 0.86 0.90 0.81 0.98
target 5: overlap 0.71 0.77 0.79 0.81 0.82 0.79 0.74 0.82 0.78 0.82
target 5: detection score 60.30 60.30 60.30 60.30 60.30 60.30 60.30 60.30 60.30 60.30
target 5: flag 1 1 1 1 1 1 1 1 1 1
target 5: angle 1.00 0.99 1.00 1.00 1.00 1.00 1.00 0.98 1.00 0.93
target 5: ncc 0.73 0.88 0.89 0.87 0.83 0.89 0.85 0.91 0.80 0.97
```

Figure 3 Examining the threshold of target object number 5 in the last ten frames, current frame: 57

According to the simulation results in Table 1, the proposed method, GWO-OF, achieved the highest tracking accuracy with an MOTP of 80.53%, outperforming other state-of-the-art methods. The closest competing method, Memetic-Adaboost, recorded an MOTP of 77.40%, while the lowest-performing method, Support Vector Machine (SVM), achieved an MOTP of only 63.86%. These findings highlight the superior performance of the GWO-OF approach, which effectively enhances tracking precision compared to alternative methods. This demonstrates the robustness and reliability of the proposed algorithm in addressing challenges associated with multi-object tracking.

The tracking accuracy in this study was evaluated using the MOTA (Multiple Object Tracking Accuracy) metric, which measures the performance of tracking algorithms by considering false negatives (missed detections), false positives (incorrect detections), and identity switches in relation to the total number of ground truth objects across all frames.

$$MOTA = 1 - \frac{\sum_t (fn_t + fp_t + Idsw_t)}{\sum_t g_t}$$

Table 1 Performance comparison of different methods

Method	MOTP (%)	MOTA (%)	F1-Score (%)
CNN	74.32	36.65	61.76
FUZZY-Ada	73.12	37.60	60.65
RMOT	67.24	35.80	59.12
Memetic	72.40	35.12	61.09
SVM	63.86	15.23	52.43
Memetic-Adaboost	77.40	41.56	68.86
GWO-OF (ours)	80.53	48.35	73.54

The simulation results in Table 1 demonstrate that the proposed method, GWO-OF (Grey Wolf Optimizer with Optical Flow), achieved the highest accuracy with a MOTA of 48.35%, outperforming all other methods. The closest competing algorithm, Memetic-Adaboost, achieved a MOTA of 41.56%, while the weakest performer, Support Vector Machine (SVM), recorded a significantly lower MOTA of 15.23%. These findings highlight the superior capability of the GWO-OF method in minimizing false detections, missed detections, and identity switches, establishing its effectiveness in enhancing tracking accuracy compared to existing methods.

The F1-Score is used as a balanced metric to evaluate the performance of tracking algorithms by considering both precision and recall. While precision measures the proportion of correctly identified positives out of all positive predictions, recall focuses on identifying the total number of true positives. Sometimes, a model may achieve high precision by minimizing false positives but at the cost of a low recall, or vice versa. To address this trade-off, the F1-Score provides a harmonic means of precision and recall, offering a more holistic evaluation of the model's performance. The formula for calculating the F1-Score is:

$$\mathrm{F1-Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

According to the results provided in Table 1, the proposed method, GWO-OF (Grey Wolf Optimizer with Optical Flow), achieved the highest F1-Score of 73.54%, outperforming all other methods. The closest competitor, Memetic-Adaboost, achieved an F1-Score of 68.86%, while the weakest performer, Support Vector Machine (SVM), recorded an F1-Score of only 52.43%. These findings emphasize the robustness of the GWO-OF approach, which effectively balances precision and recall, demonstrating its superiority in tracking tasks.

Conclusion

Object tracking is a complex task due to challenges like lighting variations, occlusions, and interactions with other objects, which can impact the appearance model and tracking accuracy. This study proposed an intelligent tracking algorithm, GWO-OF (Grey Wolf Optimizer with Optical Flow), to address these challenges by optimizing feature selection and leveraging optical flow for robust motion tracking. The findings highlight the superiority of the proposed method, achieving the highest accuracy (MOTP: 80.53%), precision (MOTA: 48.35%), and F1-Score (73.54%), outperforming methods like Memetic-Adaboost and SVM. The GWO-OF algorithm effectively handled occlusions and noise, maintaining high performance under challenging conditions. These results demonstrate their robustness and potential for real-world applications requiring reliable and precise tracking.

Conflict of Interest

The authors imply no conflict of interest.

References

[1] Kokul, Thanikasalam, Amirthalingam Ramanan, and U. A. J. Pinidiyaarachchi. “Online multi-person tracking-by-detection method using ACF and particle filter.” In *2015 IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS)*, pp. 529-536. IEEE, 2015.
[2] Mantini, Pranav, and Shishir K. Shah. “Multiple people tracking using contextual trajectory forecasting.” In *2016 IEEE Symposium on Technologies for Homeland Security (HST)*, pp. 1-6. IEEE, 2016.
[3] Shaamili, R., and A. Ruhan Bevi. “An Improved Parallel Feature Fusion-CBAM method for occluded pedestrian identification.” (2023).
[4] Yang, Bin, Junjie Yan, Zhen Lei, and Stan Z. Li. “Aggregate channel features for multi-view face detection.” In *IEEE international joint conference on biometrics*, pp. 1-8. IEEE, 2014.
[5] Kolluri, Johnson, and Ranjita Das. "Intelligent multimodal pedestrian detection using hybrid metaheuristic optimization with deep learning model." *Image and Vision Computing* (2023): 104628
[6] Shen, Zheng, Wenxiang Wu, Shuaiya Sun, and Kan Hu. "A Pedestrian Detection Method Based on Improved ResNet for Container Terminal." In *Human Centered Computing: 7th International Conference, HCC 2021, Virtual Event, December 9–11, 2021, Revised Selected Papers*, pp. 119-132. Cham: Springer Nature Switzerland, 2023.
[7] Qi, Zhouming, Mian Zhou, Guoqiang Zhu, and Yanbing Xue. "Multiple Pedestrian Tracking in Dense Crowds Combined with Head Tracking." *Applied Sciences* 13, no. 1 (2023): 440.
[8] Praveenkumar, S. M., Prakashgouda Patil, and P. S. Hiremath. "Real-time multi-object tracking of pedestrians in a video using convolution neural network and Deep SORT." In *ICT Systems and Sustainability: Proceedings of ICT4SD 2021, Volume 1*, pp. 725-736. Springer Singapore, 2022
[9] Mirjalili, Seyedali, Seyed Mohammad Mirjalili, and Andrew Lewis. "Grey wolf optimizer." *Advances in engineering software* 69 (2014): 46-61.
[10] https://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians.