2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

Swarm-ReID: Decentralized Self-Adaptive Gallery Construction for Multi-Robot Open-World Person Re-Identification

1st Marios Kaplanis
R&D
Toyota Motor Europe NV/SA
associated partner by contracted service
Brussels, Belgium
marios.kaplanis@external.toyota-europe.com

2nd Miquel Kegeleirs
IRIDIA
Université libre de Bruxelles
Brussels, Belgium
miquel.kegeleirs@ulb.be

3rd Lorenzo Garattoni
R&D
Toyota Motor Europe
Brussels, Belgium
lorenzo.garattoni@toyota-europe.com

4th Mauro Birattari
IRIDIA
Université libre de Bruxelles
Brussels, Belgium
mauro.birattari@ulb.be

5th Gianpiero Francesca
R&D
Toyota Motor Europe
Brussels, Belgium
gianpiero.francesca@toyota-europe.com

***Abstract*—Swarm perception enables a robot swarm to collectively sense and understand the environment by integrating sensory inputs from individual robots. We explore its application to person re-identification (re-id), the task of recognizing previously observed individuals. Traditional re-id systems rely on static offline galleries, which restricts their use in open-world scenarios where new identities appear over time. In robotics, most methods address single-robot re-id in person-following tasks, limiting scalability to multi-person settings, while swarm perception studies largely overlook the role of re-id algorithms. To address these gaps, we propose `Swarm-ReID`, an unsupervised method for decentralized swarm re-identification. Our method introduces mechanisms for robot-to-robot communication and informed movement strategies, enabling the swarm to collaboratively construct adaptive galleries online without centralized control. Simulations across diverse environments, number of people, swarm sizes, communication protocols, and exploration behaviors show that `Swarm-ReID` consistently outperforms existing swarm perception methods. Our results highlight how communication and informed movement improve recognition performance, establishing `Swarm-ReID` as a state-of-the-art method for open-world multi-robot person re-identification.**

***Index Terms*—Swarm robotics, Person re-identification, Open-world perception, Decentralized communication, Informed exploration, Human–robot interaction**

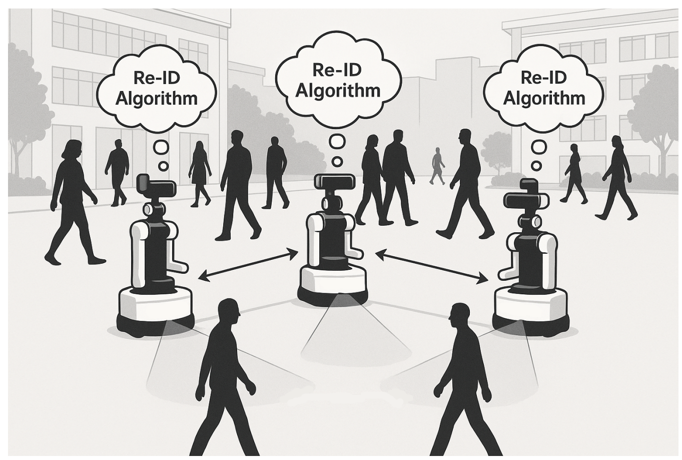

Fig. 1. Conceptual illustration of `Swarm-ReID`. A team of robots collaboratively track and re-identify people in a crowded environment. During encounters, robots exchange observations and update a shared identity database, preserving consistency across the swarm.(Generated with OpenAI ChatGPT.)

I. Introduction

Swarm perception refers to the ability of a robot swarm [1], [2] to integrate the sensory input of individual robots and build a collective understanding of the environment [3], [4]. Due to their distributed nature, swarms can collectively gather, share, and update information about their surroundings in a scalable, flexible, and fault-tolerant manner [5], [6]. This property is particularly advantageous in person (re-)identification and tracking scenarios, especially in environments with unknown or dynamic structure where static approaches—which rely on strategic sensor placement or predefined paths—are shown to be ineffective [7], [8]. These challenges motivate the exploration of swarm perception as a foundation for decentralized and adaptive person re-identification in open-world robotics.

In computer vision, most person re-identification methods assume a closed-world setting with static galleries defined *a priori*. Even recent advances remain benchmark-driven and

MB acknowledges support from the Belgian *Fonds de la Recherche Scientifique*–FNRS of which he is a Research Director.

focused on surveillance scenarios, which limits their applicability to open-world robotics settings where identities continuously appear and disappear [9], [10]. In robotics, re-id has been primarily studied in single-robot person following, where it is used to re-acquire a designated target after occlusions or appearance changes [11], [12]. More recently, swarm perception has also been investigated for re-id, where robots exchange features to improve recognition performance in decentralized settings [13]. Although these works indicate strong potential for re-id on robotics applications, they typically rely either on online continual learning or on distance-based clustering, which are not sufficient for adaptive, open-world scenarios that require robustness and long-term persistence of numerous identities.

To address the limitations above, we introduce `Swarm-ReID`, an unsupervised method for open-world person re-identification in swarm robotics (see Figure 1). `Swarm-ReID` builds on the concept of self-adaptive gallery construction [14] and extends it with mechanisms that enable decentralized operation in multi-robot settings. Each robot extracts multimodal descriptors by combining appearance features with skeleton joints, which provide complementary cues for robust re-identification across diverse viewpoints. Observations are compared against the current gallery using a similarity-based classification process: confident matches update existing classes, while uncertain cases trigger dynamic expansion. Unlabeled samples that cannot be matched are processed by an unknown-data manager, which filters, clusters, and initializes new identities. A gallery optimization module maintains compact and representative models by discarding redundant or low-quality samples. This single robot logic is complemented by robot-to-robot communication that allows for exchange of information about the previously seen people.

We validate `Swarm-ReID` through experiments in simulation across diverse environments, varying the number of people and robots, communication protocols, and exploration behaviors. The results show that `Swarm-ReID` consistently outperforms existing swarm perception methods in terms of cumulative matching characteristics (CMC) and mean average precision (mAP). To complement these findings, we present a real-world demonstration with physical robots, illustrating the feasibility of deploying `Swarm-ReID` beyond simulation. These findings confirm that decentralized communication and informed movement significantly enhance swarm-based re-identification, establishing `Swarm-ReID` as a state-of-the-art method for open-world multi-robot person re-identification.

II. Related Work

A. Person Re-ID in Computer Vision

Person re-identification (re-id) has long been a prominent research area in computer vision, originally developed to identify individuals in video surveillance systems. Early approaches relied on hand-crafted features, such as combinations of color and texture descriptors [15], or attribute-based representations that encoded semantic properties of individuals [16]. With the success of deep learning, CNN-based methods became dominant by learning end-to-end representations that capture complex appearance variations [17].

Recent works have introduced targeted innovations to address these challenges. Pose2ID leverages human pose estimation to normalize identity features and improve matching under strong pose variation [9]. AG-VPReID proposes a benchmark and multi-stream architecture for aerial–ground re-id, highlighting the difficulty of matching across extreme viewpoint discrepancies [10]. Transformer-based architectures such as TransReID employ part-aware attention and long-range feature modeling to improve robustness in complex scenarios [18], while OSNet integrates omni-scale feature learning into a compact network, achieving a favorable trade-off between accuracy and efficiency [19].

These approaches demonstrate significant progress in handling pose, viewpoint, and scale variations in static camera networks. Yet they remain designed for centralized, offline systems and do not directly address the decentralized, adaptive requirements of robot swarms.

B. Person Re-ID in Robot Person Following

In robotics, person re-identification has primarily been investigated in the context of robot person following. In these scenarios, a mobile robot must maintain persistent identification of a designated person despite occlusions, visual distractions, or appearance changes. Re-id is therefore employed as a recovery mechanism whenever the tracked person is lost.

Several recent approaches illustrate this direction. One line of work uses online continuous learning to adapt the feature extractor during operation [11]. The system updates its model using short- and long-term observations, allowing it to remain effective even when the person's appearance changes gradually over time. Another method introduces a continually adaptable re-id framework for personalized assistance [12]. This approach integrates continual updates with mechanisms to maintain identity consistency, enabling a robot to robustly follow the correct person in scenarios where multiple visually similar individuals are present.

Beyond these, other strategies have been proposed to increase robustness. For instance, lightweight multimodal metrics that combine visual appearance with complementary cues such as height or gait are used to make the re-id process more resilient in outdoor and cluttered environments [20]. Interactive frameworks have been explored, where the robot uses gestural input or user calibration to initialize re-id models before following a target [21]. Although these works demonstrate that re-id can significantly enhance person following performance, they remain limited in scope. They generally assume a single robot following a single target, relying on continual learning or distance-based models that do not scale to multi-person or multi-robot settings. As such, their applicability to open-world swarm robotics remains limited.

C. Person Re-ID in Robot Swarms

The integration of re-identification into swarm robotics has only recently been considered. Swarms offer unique ad-

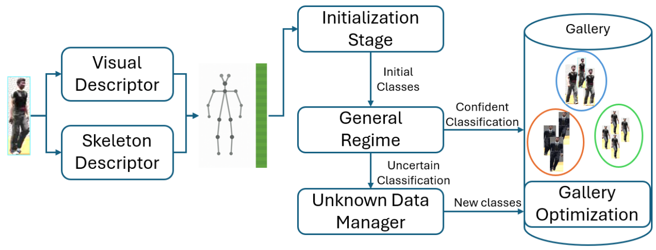

Fig. 2. `Swarm-ReID` individual data acquisition process.

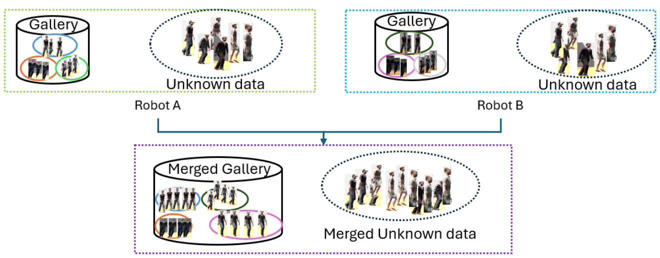

Fig. 3. `Swarm-ReID` collective data sharing process.

vantages for perception, as robots can collectively sense an environment from multiple viewpoints and share information to improve recognition. An initial study evaluated the effect of feature communication on swarm perception for re-id [13]. In this work, locally extracted features are exchanged across robots, enabling distributed, decentralized clustering and matching. The results demonstrated that decentralized communication improved recognition performance compared to isolated robots and even exceeded some centralized configurations under specific conditions.

Preliminary studies have also explored multi-robot systems [22] and swarm-based approaches [3], [4], highlighting promising directions for future advancements. Unfortunately, most prior efforts emphasize the performance of individual models rather than their integration into robust and scalable systems, and are typically validated only in simulation. In contrast, our work focuses on the system-level effectiveness of decentralized swarm perception, combining communication, movement, and re-id. Such effectiveness is also supported by a real-world proof-of-concept.

III. Method

A. Problem Statement and Overview

We formalize the problem of open-world, multi-person re-identification in the context of swarm robotics. Each robot is equipped with onboard vision sensors and must autonomously recognize individuals that may appear, disappear, or change their visual appearance. The swarm operates without a central coordinator: each robot maintains a local gallery, communicates selectively with peers, and adapts its movement to improve collective perception. We introduce `Swarm-ReID`, a decentralized method that integrates unsupervised re-identification with swarm-level communication and behavior adaptation.

B. Open-World Multi-Person Re-ID Pipeline

The re-identification pipeline of `Swarm-ReID` builds on the self-adaptive gallery construction paradigm [14], and extends it to multi-robot settings. It is composed of four main stages illustrated on Figure 2:

1) Feature Extraction and Initialization Stage: Yolov8 is used to detect each person [23]. Each detected person is represented by a multimodal feature vector

$$f_j^i = (x_j^i, s_j^i), \tag{1}$$

where x_j^i is the appearance descriptor extracted from an OSNet backbone and s_j^i is the set of visible skeleton joints obtained from Yolov8-Pose [23]. At initialization, the system forms candidate classes $\mathbb{B} = \{\mathbb{B}_1, \dots, \mathbb{B}_k\}$, where the first class is seeded by the first observed sample. A new sample x_q is assigned to an existing candidate class if the cosine similarity with its members is higher than a threshold ϵ; otherwise, a new candidate class is created. Once a candidate class reaches a minimum size l, it is promoted to a gallery class $\mathcal{C}_i$.

2) General Regime: After initialization, each new sample is compared with the gallery. The probability that a sample x_q belongs to class $\mathcal{C}_i$ is

$$p_i(x_q) = \frac{\exp(\bar{x}_i^\top x_q/\nu)}{\sum_{j=1}^N \exp(\bar{x}_j^\top x_q/\nu)}, \tag{2}$$

where ν is a temperature parameter, N is the number of gallery classes, and $\bar{x}_i$ is the weighted centroid of class $\mathcal{C}_i$:

$$\bar{x}_i = \frac{\sum_{j=1}^m r_j^i x_j^i}{\sum_{j=1}^m r_j^i}, \quad r_j^i = \frac{s_j^i}{s_T}, \tag{3}$$

with s_j^i the number of detected joints in the sample and s_T the total number of skeleton joints. A sample is confidently assigned to class $\mathcal{C}_i$ if

$$\frac{\max_i p_i(x_q)}{\max_{j \neq i} p_j(x_q)} \geq \tau, \tag{4}$$

where τ is a confidence threshold. Otherwise, it is sent to the Unknown Data Manager.

3) Unknown Data Manager: Samples not meeting (4) are clustered using DBSCAN [24] to discover new classes. Clusters are validated using a minimum distance criterion:

$$D(\mathcal{C}_w, \mathcal{C}) = \min_{\mathcal{C}_i \in \mathcal{C}} \left(\min_{x_j \in \mathcal{C}_i,\, x \in \mathcal{C}_w} (1 - x^\top x_j) \right). \tag{5}$$

If $D(\mathcal{C}_w, \mathcal{C}) > \alpha$, the cluster $\mathcal{C}_w$ is added as a new identity; otherwise, it is merged with the closest existing class.

4) Optimization: To maintain a bounded gallery, two criteria are used: (1) Intra-class diversity of a feature $x \in \mathcal{C}_i$:

$$D_i(x) = \min_{x_j \in \mathcal{C}_i \setminus x} (1 - x^\top x_j), \tag{6}$$

with class diversity

$$D(\mathcal{C}_i) = \min_{x_j, x_k \in \mathcal{C}_i, j \neq k} (1 - x_j^\top x_k). \tag{7}$$

A new sample x_q is retained only if $D_i(x_q) \geq D(\mathcal{C}_i)$.

(2) Uncertainty of x, measured with Shannon entropy:

$$H(x) = -\sum_{i=1}^{N} p_i(x) \log p_i(x). \tag{8}$$

When memory is exceeded, the feature to drop is

$$x^* = \arg\max_{x \in \mathcal{C}_i} \big(\gamma H(x) \log(1/N) - (1-\gamma) D_i(x)\big), \tag{9}$$

where $\gamma \in [0, 1]$ balances entropy and diversity. This ensures the gallery remains compact, diverse, and informative.

C. Communication

In addition to local gallery construction, `Swarm-ReID` relies on decentralized communication to ensure that knowledge about identities propagates across the swarm (see Figure 3). We investigate three communication strategies:

- *Gallery exchange:* robots periodically share compact representations of their adaptive galleries.
- *Unknown data exchange:* robots share uncertain samples to increase the chances of discovering new identities collaboratively.
- *Hybrid exchange:* a combination of gallery and unknown data sharing.

Communication is constrained by a radius R_c, which defines the spatial range within which two robots can exchange information. Varying R_c allows us to evaluate the trade-off between communication density and redundancy.

To further characterize the swarm's communication patterns, we recorded the number and duration of pairwise meetings among robots. Figure 4 illustrates these statistics, highlighting meetings frequencies and durations in the swarm.

D. Robotics Pipeline

`Swarm-ReID` complements the re-id pipeline with swarm-level exploration behaviors that influence data collection. We investigate 3 different behaviors.

1) Random Walk: Robots explore the environment using a random strategy designed to provide unbiased coverage of people and locations. We adopt the ballistic motion with obstacle avoidance [25], [26]. Each robot moves forward at a constant speed until it detects an obstacle; upon detection, it turns away from the obstacle for a random number of control steps and then resumes forward motion.

TABLE I
OVERVIEW OF EXPERIMENTAL SCENARIOS. WE VARY THE NUMBER OF PEOPLE, ROBOTS, COMMUNICATION RANGES, ENVIRONMENTS, AND ROBOT BEHAVIORS.

Factor	**Values**
Nb. of People	25, 50
Nb. of Robots	4, 8
Communication range	1.5, 5, 10
Environment	with / without obstacles
Behaviors	3 exploration strategies

2) Naive Attraction–Repulsion: Robots regulate their inter-robot spacing using the attraction and repulsion modules of AutoMoDe-Chocolate [25]. The controller is tuned to favor pairs while avoiding both isolation and groups larger than two. When no peer is detected, a robot continues random exploration; upon sensing the first peer, it is attracted to form a temporary pair, increasing the chance that both capture the same scene. If more than two peers are within range, robots apply repulsion to disperse, preventing congestion and redundant sampling.

3) Closest–Neighbor Attraction–Repulsion: Robots broadcast their positions to nearby peers to improve coordination. They explore via a simple random walk until a peer is detected within an attraction radius, at which point each robot selects the closest neighbor and steers toward it. This nearest-neighbor rule promotes rapid pair formation while discouraging larger clusters. When the inter-robot distance falls below a separation threshold, a brief, fixed-duration repulsion phase pushes the robots apart before they resume exploration. This cycle avoids congestion and fosters repeated, transient meet-ups, increasing the likelihood of co-observing the same scene.

IV. EXPERIMENTS

We evaluate `Swarm-ReID` across a set of simulated scenarios designed to test its performance under varying conditions. In our experiments we systematically vary the number of people and robots, the presence of obstacles, and robot exploration behaviors. The experiments run for 20 minutes. Table I summarizes the experimental configurations.

A. Experimental Setup

To evaluate the swarm's ability to perform re-identification using `Swarm-ReID`, we design a series of simulated experiments in which Toyota HSR [27] robots navigate in human-populated environments. Our setup follows the simulation framework introduced by [13], where human motion is simulated in Unity and robot control in ARGoS3 [28], with synchronization via ROS. We consider environments both with and without obstacles, and vary the number of people, robots, communication range, and behaviors as summarized in Table I.

1) Evaluation Metrics: For each experiment, we compute the cumulative matching characteristics (CMC) curve and the mean average precision (mAP) [29]. Results are aggregated across all robots in the swarm.

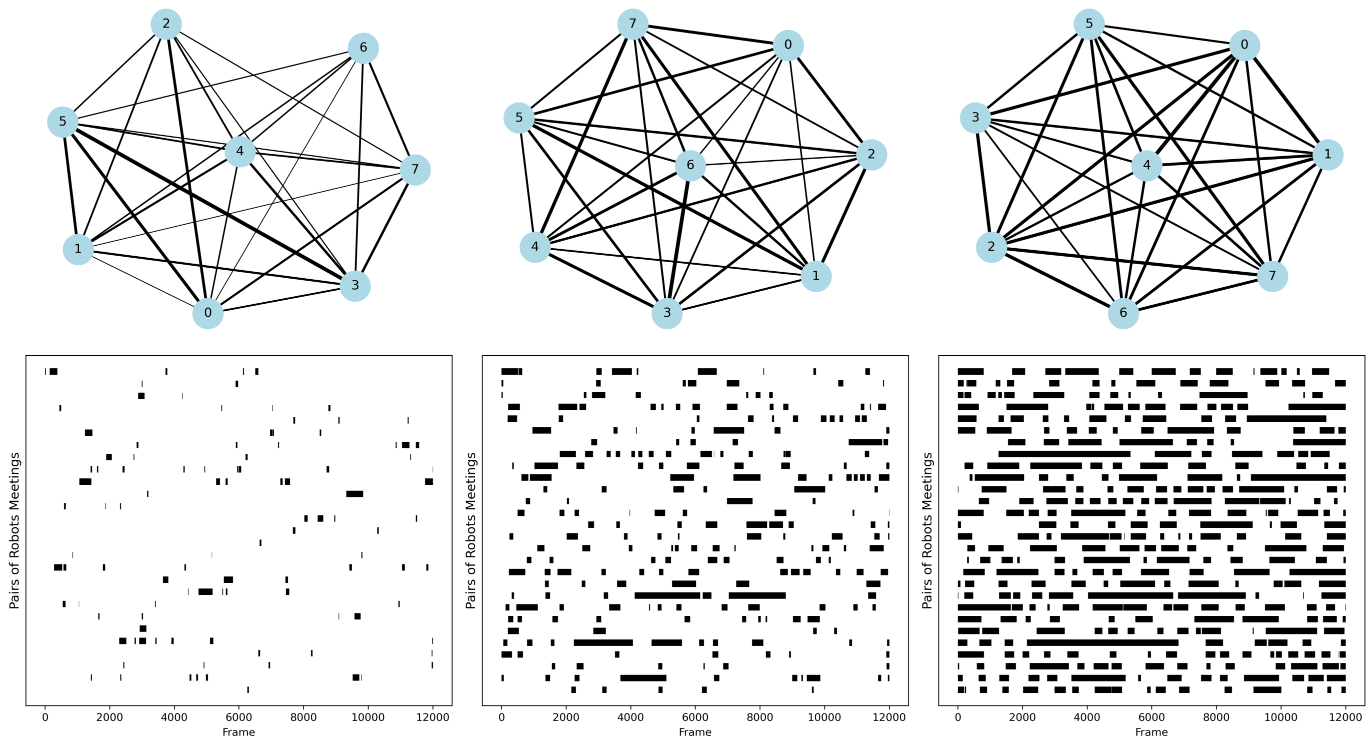

Fig. 4. Communication patterns in simulation with 50 people and 8 robots. Results are shown for three communication radii $R_c \in \{1.5, 5, 10\}$ m from left to right. Top: pairwise connections between robots, where edge width increases with the number of meetings. Bottom: timelines of meetings, with horizontal bars marking meeting intervals.

CMC@k evaluates top-k identification accuracy, i.e., how often the first correct identity appears within the top-k ranks.

mAP evaluates retrieval quality over the full ranking, quantifying how completely and how early all correct matches are returned. Both CMC and mAP are computed using 10% of all images observed by the robots during the experiment as queries, with a minimum of 1000 and a maximum of 5000 queries.

2) Dataset and Re-ID Features: Bounding boxes are extracted with YOLOv8 [23], and appearance descriptors are obtained using state-of-the-art re-id model OSNet [19]. Each robot runs the `Swarm-ReID` pipeline, incrementally building and updating its gallery of observed people.

3) Communication Strategies: We evaluate five communication strategies. In the *No Communication* case, each robot maintains its gallery independently. In the *Comm Galleries* and *Comm Unknowns* settings, robots exchange only gallery entries or only unknown samples, respectively. The *Comm Both* setting enables the exchange of both, while the *Comm Both + Optimization* variant additionally applies the gallery optimization described in Section III.

4) Communication Radius: We investigate the impact of communication range on swarm re-identification by varying the communication radius between robots. As summarized in Table I, we test three settings: $1.5\,\mathrm{m}$, $5\,\mathrm{m}$, and $10\,\mathrm{m}$. These values reflect different levels of connectivity within the swarm: short-range communication leads to fragmented information sharing, while larger radii allow denser information exchange and faster consensus across the swarm.

5) Exploration Behaviors: We further compare different exploration strategies for the swarm, as summarized in Table I. These include *Random Walk*, *Naive Attraction–Repulsion*, and *Closest-Neighbor Attraction–Repulsion*, the latter relying on inter-robot communication to share robot poses.

6) Baseline: We compare against the swarm-based collective perception approach of Kegeleirs et al. [13], which represents the closest prior benchmark for swarm-based people tracking. Each robot independently detects people with YOLOv8, tracks them with BoT-SORT, and extracts embeddings via OSNet fine-tuned on standard re-id datasets. These embeddings are clustered into identities using a distance-based merging strategy, resulting in a local gallery for each robot.

When robots meet, they exchange their databases of people, and new clusters are integrated using the same distance-based merging procedure. If the distance between clusters is below a predefined threshold, they are merged; otherwise, a new cluster is created. Unlike `Swarm-ReID`, this method relies solely on distance-based fusion and does not exploit specialized communication of galleries, unknown data, or optimization. We refer the reader to [13] for a detailed description of the full pipeline.

TABLE II
COMPARISON WITH BASELINE [13]. METRICS: CMC@1 AND MAP ACROSS COMMUNICATION SETTINGS. BEST VALUES PER COLUMN ARE IN BOLD.

	No Comm		Comm 1.5		Comm 5		Comm 10	
Method	CMC@1	mAP	CMC@1	mAP	CMC@1	mAP	CMC@1	mAP
Kegeleirs et al.	0.288	0.189	0.534	0.213	0.602	0.187	0.644	0.201
`Swarm-ReID`	**0.784**	**0.676**	**0.827**	**0.731**	**0.843**	**0.759**	**0.859**	**0.784**

TABLE III
RESULTS UNDER DIFFERENT BEHAVIORS FOR `Swarm-ReID` (AVERAGED). METRICS: CMC@1 / MAP.

	No Comm		Comm 1.5		Comm 5		Comm 10	
Behavior	CMC@1	mAP	CMC@1	mAP	CMC@1	mAP	CMC@1	mAP
Random Walk	0.784	0.676	0.827	0.731	0.843	0.759	0.859	0.784
Naive A-R	0.821	0.597	0.884	0.715	0.885	0.765	0.890	0.778
Closest N. A-R	**0.840**	**0.681**	**0.896**	**0.803**	**0.891**	**0.783**	**0.894**	**0.812**

B. Results

We present the results of our experiments, following the structure described in Section IV-A. As our analysis revealed no significant differences in performance across variations in population size (25 vs. 50 individuals), number of robots (4 vs. 8), or the presence of obstacles, we report all subsequent results aggregated across these conditions. Each subsection then focuses on isolating specific factors—namely communication strategy, robot behavior, and communication radius—to examine their individual contributions to the performance of `Swarm-ReID`.

1) Comparison with Baseline: We compare `Swarm-ReID` with the baseline approach of Kegeleirs et al. [13] on simulations with 25 and 50 people and 4 and 8 robots, both with and without obstacles. Table II summarizes the averaged results across all scenarios and communication ranges.

`Swarm-ReID` consistently outperforms the baseline, even without communication (CMC@1 of 0.784 vs. 0.288). Although the baseline benefits from communication, its improvement saturates quickly, with CMC@1 remaining below 0.65 even at the largest communication radius. In contrast, `Swarm-ReID` scales more effectively with communication, reaching up to 0.859 CMC@1 and 0.784 mAP at 10 m. These results highlight the robustness of `Swarm-ReID` and its ability to better exploit shared information in large-scale long-term settings.

2) Impact of Exploration Behaviors: We analyze the impact of different exploration behaviors, as mentioned in Section III-D. Table III summarizes the averaged results across all scenarios for the `Swarm-ReID` method.

The results highlight three main insights. First, communication improves re-identification performance regardless of behavior: both CMC@1 and mAP steadily increase from the no-communication setting to the largest communication range (10 m). This confirms that prolonged interaction enables the swarm to converge to a more consistent shared gallery.

Second, the type of behavior strongly influences performance. Random Walk yields the lowest accuracy, as robots tend to explore inefficiently and exchange limited information. Naive Attraction–Repulsion improves results by promoting more frequent encounters between robots, while Closest Neighbor Attraction–Repulsion achieves the highest mAP values, demonstrating the benefit of more purposeful motion strategies leveraging information sharing. Indeed, the average communication rate (meetings per second) increases across strategies: Random Walk 0.29, Naive Attraction–Repulsion 0.34, and Closest-Neighbor Attraction–Repulsion 0.40.

Finally, even though the improvements saturate at higher communication ranges (5–10 m), the combination of structured behaviors and communication allows the swarm to sustain high accuracy over extended runs. This validates the robustness of `Swarm-ReID` under long-term conditions, where maintaining reliable identity tracking is particularly challenging.

3) Impact of Communication Radius: We investigate the effect of the communication range (1.5 m, 5 m, 10 m) on swarm-level performance. As reported in Table II (baseline comparison runs) and Table III (exploration behavior runs), increasing the communication radius consistently improves both CMC@1 and mAP across all scenarios. Even a modest range of 1.5 m already yields significant gains over the no-communication baseline, while larger ranges (5 m and 10 m) further enhance consistency by enabling more frequent gallery exchanges.

Importantly, the improvements saturate as the range increases: the jump from no communication to 1.5 m is the largest, while gains from 5 m to 10 m are comparatively smaller. This trend is visible across all behaviors and environments, demonstrating that while extended communication helps, the primary benefit lies in enabling robots to exchange data at all.

4) Impact of Communication Strategies: To further disentangle the contributions of each component, we ablate the communication design by selectively enabling gallery sharing, unknown data sharing, or both, with and without the optimization. Table IV reports the results averaged across all simulations and across the various communication ranges.

We observe that communication of either galleries or unknowns already improves over the no-communication baseline. Exchanging galleries provides the largest gain, while exchanging unknown data alone also leads to significant improvements. Combining both yields the best overall performance, highlighting the complementary nature of the two information streams. Interestingly, adding optimization slightly reduces both metrics compared to plain gallery+unknown exchange (and even compared to exchanging galleries only), suggesting that optimization trades recall for precision and may be more beneficial in settings with higher noise. Therefore, optimization should be applied selectively depending on scenario constraints.

To better understand swarm dynamics, we analyze how performance evolves over time as communication is enabled.

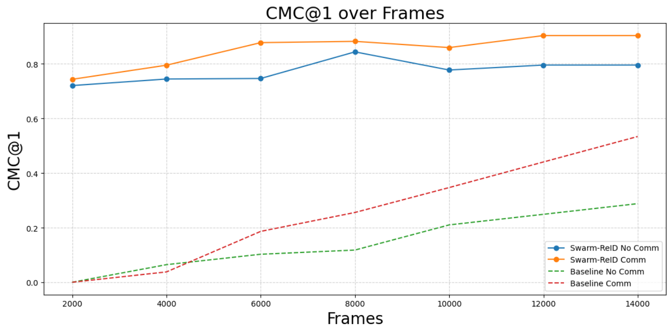

Fig. 5. CMC@1 over time for Swarm-ReID and Baseline [13] under different communication settings (with vs. without communication).

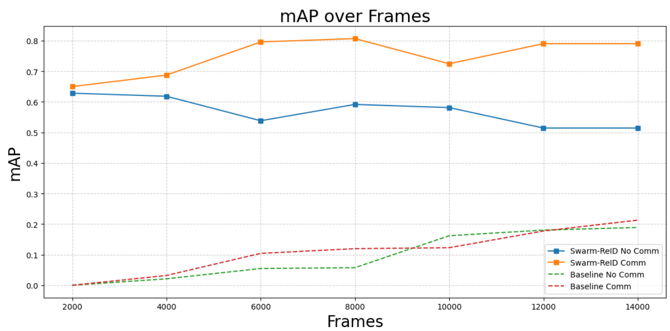

Fig. 6. mAP over time for Swarm-ReID and Baseline [13] under different communication settings (with vs. without communication).

Fig. 5 and Fig. 6 plot the temporal progression of CMC and mAP, respectively, for the scenario with 50 people and 8 robots under different communication settings. Two key observations emerge. First, `Swarm-ReID` consistently maintains a large margin over the baseline. Second, communication improves CMC across both methods, but its effect on mAP is strong for `Swarm-ReID` and negligible for the baseline. In summary, these results demonstrate that while communication is generally beneficial, its impact is significantly amplified in `Swarm-ReID`.

Overall, our study confirms that communication is the key driver of swarm-level performance, while optimization should be applied selectively depending on scenario constraints. In addition to these results, we validate `Swarm-ReID` using a real-world demonstration with the 4 Mercator robots. The Mercator platform is described in [30]. The video with the accompanying results is provided as supplementary material.

TABLE IV
ABLATION STUDY OF `Swarm-ReID` COMMUNICATION STRATEGIES. METRICS: CMC@1 AND MAP (AVERAGED ACROSS ALL SIMULATIONS AND THE COMMUNICATION RANGES CONSIDERED).

Method	CMC@1	mAP
No Communication	0.700	0.654
Comm Galleries	0.849	0.764
Comm Unknowns	0.816	0.714
Comm Both	**0.850**	**0.766**
Comm Both + Opt	0.833	0.730

V. CONCLUSION

In this work, we introduced `Swarm-ReID`, a novel method for cooperative person re-identification in robot swarms. In contrast to existing approaches that treat robots as independent observers, `Swarm-ReID` leverages inter-robot communication to maintain a shared and coherent gallery of identities in dynamic, open-world scenarios. Through a variety of experiments, we demonstrated that communication substantially improves performance across all metrics, environments, and robot behaviors. Our experiments highlight consistent gains over the baseline [13], with improvements increasing with communication radius. We showed that swarm-level robustness depends not only on communication but also on robot motion strategies: closest-neighbor attraction–repulsion consistently outperforms naive behaviors, while random walk was the least effective. Finally, our ablation study shows that exchanging both galleries and unknown data is the most beneficial, although optimization must be applied with care.

In addition to simulation, we presented a real-world proof-of-concept demonstration with robots, providing first evidence that `Swarm-ReID` can be deployed beyond simulation.

In summary, our results establish communication as a key enabler for scalable and reliable swarm re-identification.

While the current study provides valuable insights into the performance of `Swarm-ReID`, several directions remain open for future investigation. In particular, exploring the relationship between the performance of individual robot models and the role of communication presents a promising avenue. Additionally, further studies on scalability of the system across swarm sizes and number of people is an important direction for future work. Further research is also needed to assess the impact of bandwidth constraints and communication errors on system performance, and to develop mechanisms that enable the swarm to remain robust under such conditions. In this context, adaptive communication strategies that balance bandwidth usage and accuracy could offer an effective solution.

From a benchmarking perspective, we believe that the task of multi-robot open-world person re-identification opens up a broad spectrum of research opportunities that extend beyond the re-identification problem itself. This includes challenges related to distributed perception, collaborative learning, and resilient communication in dynamic environments.

REFERENCES

[1] G. Beni, "From swarm intelligence to swarm robotics," in *SAB 2004*, ser. LNCS, E. Şahin and W. M. Spears, Eds., vol. 3342. Berlin, Germany: Springer, 2005, pp. 1–9.
[2] E. Şahin, "Swarm robotics: from sources of inspiration to domains of application," in *SAB 2004*, ser. LNCS, E. Şahin and W. M. Spears, Eds., vol. 3342. Berlin, Germany: Springer, 2005, pp. 10–20.
[3] M. Kegeleirs, D. Garzón Ramos, G. Legarda Herranz, I. Gharbi, J. Szpirer, O. Debeir, K. Hasselmann, L. Garattoni, G. Francesca, and M. Birattari, "Collective perception for tracking people with a robot swarm," in *ICRA@40*. IEEE, 2024.

[4] M. Kegeleirs, D. Garzón Ramos, G. Legarda Herranz, I. Gharbi, J. Szpirer, K. Hasselmann, L. Garattoni, G. Francesca, and M. Birattari, "Leveraging swarm capabilities to assist other systems," https://arxiv.org/abs/2405.04079, 2024.

[5] M. Brambilla, E. Ferrante, M. Birattari, and M. Dorigo, "Swarm robotics: a review from the swarm engineering perspective," *Swarm Intell.*, vol. 7, no. 1, pp. 1–41, 2013.

[6] M. Dorigo, M. Birattari, and M. Brambilla, "Swarm robotics," *Scholarpedia*, vol. 9, no. 1, p. 1463, 2014.

[7] C. Robin and S. Lacroix, "Multi-robot target detection and tracking: taxonomy and survey," *Auton. Robots*, vol. 40, no. 4, pp. 729–760, 2016.

[8] M. Kegeleirs and M. Birattari, "Towards applied swarm robotics: current limitations and enablers," *Front. Robot. AI*, vol. 12, 2025.

[9] C. Yuan, G. Zhang, C. Ma, T. Zhang, and G. Niu, "From poses to identity: Training-free person re-identification via feature centralization," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 24 409–24 418.

[10] H. Nguyen, K. Nguyen, A. Pemasiri, F. Liu, S. Sridharan, and C. Fookes, "Ag-vpreid: A challenging large-scale benchmark for aerial-ground video-based person re-identification," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 1241–1251.

[11] H. Ye, J. Zhao, Y. Zhan, W. Chen, L. He, and H. Zhang, "Person re-identification for robot person following with online continual learning," *IEEE Robotics and Automation Letters*, 2024.

[12] F. Rollo, A. Zunino, N. Tsagarakis, E. M. Hoffman, and A. Ajoudani, "Continuous adaptation in person re-identification for robotic assistance," in *ICRA 2024*. IEEE, 2024, pp. 425–431.

[13] M. Kegeleirs, I. Gharbi, M. Kaplanis, L. Garattoni, G. Francesca, and M. Birattari, "Assessing the impact of feature communication in swarm perception for people re-identification," *Frontiers in Robotics and AI*, vol. Volume 12 - 2025, 2025.

[14] S. Casao, P. Azagra, A. C. Murillo, and E. Montijano, "A self-adaptive gallery construction method for open-world person re-identification," *Sensors*, vol. 23, no. 5, p. 2662, 2023.

[15] M. Hirzer, C. Beleznai, P. M. Roth, and H. Bischof, "Person re-identification by descriptive and discriminative classification," in *Scandinavian conference on Image analysis*. Springer, 2011, pp. 91–102.

[16] R. Layne, T. M. Hospedales, S. Gong, and Q. Mary, "Person re-identification by attributes." in *Bmvc*, vol. 2, no. 3, 2012, p. 8.

[17] W. Li, R. Zhao, T. Xiao, and X. Wang, "Deepreid: Deep filter pairing neural network for person re-identification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 152–159.

[18] S. He, H. Luo, P. Wang, F. Wang, H. Li, and W. Jiang, "Transreid: Transformer-based object re-identification," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 15 013–15 022.

[19] K. Zhou, Y. Yang, A. Cavallaro, and T. Xiang, "Omni-scale feature learning for person re-identification," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 3702–3712.

[20] K. Koide and J. Miura, "Identification of a specific person using color, height, and gait features for a person following robot," *Robotics and Autonomous Systems*, vol. 84, pp. 76–87, 2016.

[21] F. Rollo, A. Zunino, G. Raiola, F. Amadio, A. Ajoudani, and N. Tsagarakis, "Followme: A robust person following framework based on visual re-identification and gestures," in *2023 IEEE International Conference on Advanced Robotics and Its Social Impacts (ARSO)*. IEEE, 2023, pp. 84–89.

[22] M.-A. Popovici, I. E. Gil, E. Montijano, and D. Tardioli, "Distributed dynamic assignment of multiple mobile targets based on person re-identification," in *Iberian Robotics conference*. Springer, 2022, pp. 211–222.

[23] R. Varghese and S. M., "Yolov8: A novel object detection algorithm with enhanced performance and robustness," in *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, 2024, pp. 1–6.

[24] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *kdd*, vol. 96, no. 34, 1996, pp. 226–231.

[25] G. Francesca, M. Brambilla, A. Brutschy, L. Garattoni, R. Miletitch, G. Podevijn, A. Reina, T. Soleymani, M. Salvaro, C. Pinciroli, F. Mascia, V. Trianni, and M. Birattari, "AutoMoDe-Chocolate: automatic design of control software for robot swarms," *Swarm Intell.*, vol. 9, no. 2–3, pp. 125–152, 2015.

[26] M. Kegeleirs, D. Garzón Ramos, and M. Birattari, "Random walk exploration for swarm mapping," in *TAROS 2019*, ser. LNCS, K. Althoefer, J. Konstantinova, and K. Zhang, Eds., vol. 11650. Cham, Switzerland: Springer, 2019, pp. 211–222.

[27] U. Yamaguchi, F. Saito, K. Ikeda, and T. Yamamoto, "Hsr, human support robot as research and development platform," *The Abstracts of the international conference on advanced mechatronics : toward evolutionary fusion of IT and mechatronics : ICAM*, vol. 2015.6, pp. 39–40, 2015.

[28] C. Pinciroli, V. Trianni, R. O'Grady, G. Pini, A. Brutschy, M. Brambilla, N. Mathews, E. Ferrante, G. A. Di Caro, F. Ducatelle, M. Birattari, L. M. Gambardella, and M. Dorigo, "ARGoS: a modular, parallel, multi-engine simulator for multi-robot systems," *Swarm Intell.*, vol. 6, no. 4, pp. 271–295, 2012.

[29] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian, "Scalable person re-identification: A benchmark," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1116–1124.

[30] M. Kegeleirs, D. G. Ramos, J. Szpirer, C. L. Vera, G. L. Herranz, I. Gharbi, G. Francesca, and M. Birattari, "Mercator: A modular swarm-dedicated robot platform," *HardwareX*, vol. 25, p. e00751, 2026.