A collaborative real-time object detection and data association framework for autonomous robots using federated graph neural network

Feriel Talbi, Samir Ouchani, Yohan Dupuis, Mimoune Malki

A Collaborative Real-Time Object Detection and Data Association Framework for Autonomous Robots Using Federated Graph Neural Network

Feryal Batoul TALBI[1,2], Samir OUCHANI[2], DUPUIS Yohan[3], and Mimoun MALKI[1]

[1] École Supérieure en Informatique, Sidi Bel-abbes, Algeria,
[2] CESI LINEACT, Aix-en-Provence, France,
[3] CESI LINEACT, Paris La Défense, France.

Abstract. Autonomous robotics require secure and decentralized decision-making systems that ensure data privacy and computational efficiency, especially in critical areas. Current centralized models or human input are associated with data breaches and security vulnerabilities. To counter these, we propose **CoRODDA**, a dedicated framework combining federated learning and graph neural networks. **CoRODDA** enhances object detection and data association in autonomous robots, enabling them to learn from local data while preserving privacy and interpreting graph-structured associated data to understand the surrounding environments. The experiments showed the effectiveness of **CoRODDA** compared to the state-of-the-art, particularly in non-detected objects, improving data privacy and decision-making capabilities.

Keywords: Federated Learning · Graph Neural Network · Autonomous Robots · Data Association · Object Detection · Security.

1 Introduction

Motivation. Although significant advancements have been made in robotic systems, a key challenge that persists is the execution of complex tasks during censorious situations in critical environments. These tasks, based on object detection and data association, need to be performed in a decentralized manner while ensuring data privacy [1]. Conventional approaches often fall short in situations where centralized data aggregation is either unfeasible or poses a threat to privacy [2]. However, in the current era of rapid technological evolution, autonomous robotic systems are recognized as a driving force behind numerous industry transformations [3, 4]. Yet, a global challenge remains to ensure these robotic systems not only carry out complex tasks (such as object detection and data association) but also maintain data privacy, especially in outdoor environments [5, 6].

Related Work. Existing solutions often grapple with challenges when centralized data processing becomes infeasible, or as data privacy takes center stage [7]. Recognizing these challenges, the adoption of Graph Neural Networks (GNNs) and Federated Learning (FL) has significant advancements in autonomous robotic systems, notably in intricate tasks like object detection and data association [8]. GNNs, renowned for their adeptness in managing graph-structured data, have revolutionized robotic interactions and understanding of their surroundings [9]. In parallel, FL has ushered in transformative methodologies, allowing robots to harness local data insights while upholding rigorous data privacy and refining computational efficiency [10].

The field of Fed and GNN has seen extensive research, covering areas such as privacy preservation, non-IID data challenges, decentralization, and FL personalization [11]. Though innovative techniques such as local differential privacy and homomorphic encryption have been seminal, they also underscore computational and model convergence dilemmas. Persistent challenges revolve around data heterogeneity and the nuances of managing non-IID data. Decentralization, despite its merits in distributed control, contends with obstacles in real-time responsiveness and model scalability [12]. While numerous strategies address spatial-temporal dependencies, aligning them with the dynamic needs of stakeholders is a complex endeavor [13]. Striking an equilibrium between instantaneous adaptability and stringent data privacy remains a significant hurdle [14].

Contributions. In the realm of autonomous robotics, the integration of GNN and FL (FedGNN) holds great promise. GNNs are pivotal for handling graph-structured data, empowering robots to

understand better and navigate their environment. Concurrently, FL ensures efficient decentralized learning, emphasizing data privacy. We develop a Collaborative Real-Time Object Detection and Data Association (**CoRODDA**) framework for autonomous robots using FedGNN that synergistically leverages the capabilities of both FL and GNNs, specially tailored for object detection and data association tasks. Emphasizing real-world critical applications, **CoRODDA** stands out, marking a new frontier in data privacy and decision-making. Drawing from an extensive review of the related literature and building on the state-of-the-art, this work presents the following contributions:

1. Suggest **CoRODDA**, an amalgamation of FL and GNN strengths, optimized for object detection Section 2.
2. An in-depth assessment of **CoRODDA**, highlighting its preeminence in maintaining data privacy and decision-making Section 3.

2 CoRODDA Framework

In this section, we present a detailed overview of our proposed framework, highlighting the significance and operation of each component. Figure 1 depicts the structure of **CoRODDA**. It progresses through a series of local and centralized stages, with each stage serving a pivotal role in realizing the overarching objective of detecting objects with low scores or those missed by object detection algorithms.

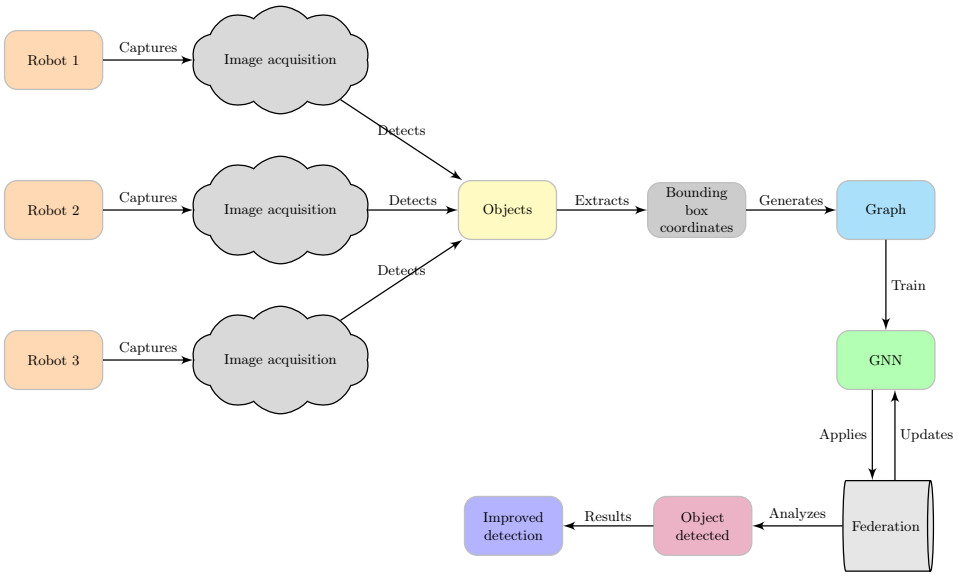

Fig. 1: Detailed Object Detection Process in **CoRODDA**.

1. **Local Operations:**
   (a) **Data Acquisition and Preprocessing:** Robots capture images, serving as the primary input for **CoRODDA**. All raw image data remains localized to each robot, ensuring the confidentiality of the initial dataset.
   (b) **Object Detection and Extraction:** Advanced object detection techniques and predefined models are utilized to identify objects within the processed images. This step also meticulously extracts the spatial boundaries of identified objects.
   (c) **Graph Generation and Training:** A graph that represents detected objects and their spatial relationships is produced. FedGNN model is then fed with this data. Robots individually refine their models without transmitting any data elsewhere.
2. **Centralized Operations:**

(a) **Model Aggregation:** The server accumulates updates from all participating robots, refining a global model to improve object detection capabilities and confidence scores.The strength of the federated learning approach is evident here, as only model updates, devoid of any raw or derivative data, are shared.
(b) **Real-time Analysis:** The global model is employed for real-time object analysis. Feedback is dispatched to the robots, amplifying their ability to recognize not just common objects but also those that might be minute or situated closely — objects previously undetected by initial techniques.

3 Experimental Results

In this section, we analyze **CoRODDA**'s performance in object detection with YOLOv5 across 80 classes, using the Stanford Drone Dataset[4]. **CoRODDA** uses a spatial approximation to segment insights across a federated network. By incorporating the GNN model, GraphSAGE, **CoRODDA** interprets graph-structured data, emphasizing cooperative learning and data privacy.

3.1 GNN Model Architecture

Central to **CoRODDA** is the GraphSAGE GNN model, known for its ability to generalize larger unseen graphs. Its selection is based on attributes apt for our application. Below, we detail the model's architecture and features.

- **Node Features**: Nodes are defined by features forming the base for further analysis. These features capture the essential attributes and characteristics of the nodes,
- **Network Depth**: The 16-layer deep model emphasizes critical data patterns through multiple nonlinear transformations.
- **Optimization**: We use the Adam optimizer with a 0.01 learning rate, known for its adaptive properties, to strike a balance between convergence speed and precision.
- **Dropout**: A 0.5 dropout rate is applied to reduce overfitting by randomly deactivating neurons during training.
- **Training Duration**: The model is trained over 10 epochs to optimize learning without risking overfitting.
- **Metrics**: Post-training, the model achieves a 73% accuracy rate for the training dataset, demonstrating its proficiency in handling graph data.

This intricate architecture, characterized by its depth, underscores our commitment to achieving unparalleled results in graph data processing and analysis. The model is adept at capturing contextual information and relational dependencies, as illustrated in Figure 2. Over the span of 10 epochs, the training accuracy begins at 70.31% and peaks at 73.09%, while the validation accuracy commences at 71.31% and culminates at 74.09%. In tandem, the F1 Score for training initiates at 65.18% and reaches 69.00%, and for validation, it starts at 66.18% and ascends to 70.00%.

3.2 Enhancements via FL

In his approach, we have married the strengths of GNNs with the decentralization and privacy preservation features of FL. Key attributes and advantages of this federated model are:

- **Decentralized Training**: Our approach empowers each robot with its own GNN model, promoting learning from real-time interactions. This reduces communication needs, optimizes bandwidth, and notably bolsters data privacy by retaining data locally.
- **Genetic Algorithms**: We utilize GAs in the FL workflow, refining model weights by mimicking biological evolution. As a result, models are iteratively optimized based on performance with local datasets.

[4] https://cvgl.stanford.edu/projects/uav_data/

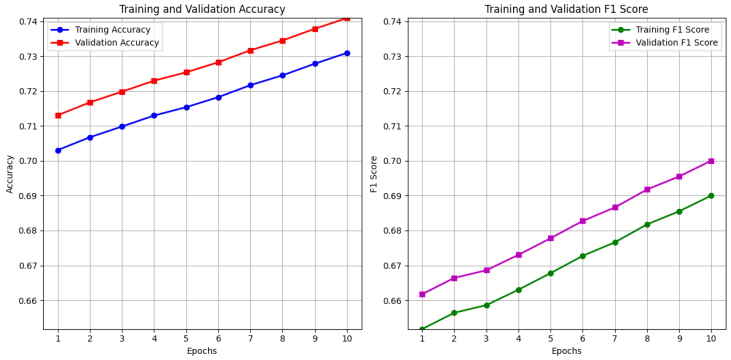

Fig. 2: Accuracy and F1 Score achieved by the GNN Model.

– **Federated Genetic Algorithm [15–17]**: Our advanced FedGA method amalgamates weights from all robot models, crafting a holistic global model that benefits from the entire network's collective intelligence.
– **Collaborative Data Utilization**: Robots contribute unique insights, and when combined, these diverse updates form a global model that's adaptable and represents various environments.
– **Accuracy Gains**: Post a 50-epoch training, our federated and genetically enhanced approach achieved a significant 78% rise in accuracy.

The trajectory of accuracy over training epochs is illustrated in Figure 3 which shows the distinct improvements brought by FedGA. Table 1 methodically contrasts the standalone GNN model with the enhanced FedGA and FedAVG models. The most notable observations from this table are:

– FedGA has undergone a significant increase in training epochs compared to the 10 epochs of the standalone GNN model.
– In terms of accuracy, FedGA achieved a notable 5% improvement over the GNN model, registering at 78% for the validation training. Meanwhile, FedAVG slightly underperformed with an accuracy of 70.8% training dataset.
– Furthermore, the F1 Score for FedGA outshines both the GNN model and the FedAVG, boasting a score of 80%, a marked increase from the GNN model's 69% and a substantial leap from FedAVG's 0.59%.

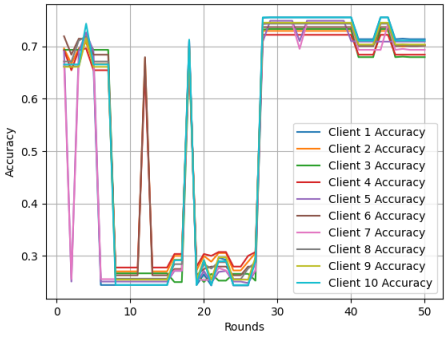

Fig. 3: Accuracy improvement over epochs in the federated GNN Model.

3.3 Integration with YOLOv5 and Performance Insights

In **CoRODDA**, we combine the real-time object detection capabilities of YOLOv5 with the sophisticated data interpretation potential of FedGNN, resulting in significant performance enhancements.

3.3.1 YOLOv5 Object Detection YOLOv5 efficiently detects objects in images using confidence scores; while high scores signify precise detection, low scores can result in missed objects, represented by empty brackets.

– **Predefined Classes in YOLOv5.** YOLOv5 is trained in specific classes. Objects outside these classes, may be present but remain undetected due to the predefined scope.

Metrics/Features	GNN Model	FedGa	FedAVG
Input features per node	3	3 (Unchanged)	3 (Unchanged)
Hidden layers	16	16 (Unchanged)	16 (Unchanged)
Optimizer's learning rate	0.01	0.01 (Unchanged)	0.01 (Unchanged)
Dropout rate	0.5	0.5 (Unchanged)	0.5 (Unchanged)
Training epochs	10	50	50
Achieved accuracy	73%	78%	70.8%
F1 Score	69%	80%	0.59%
Inference time	Real-time	Real-time	Real-time
Machine Specifications	MacBook Pro M1 chip, 16GB RAM		

Table 1: Comparative Summary of Performance Metrics and Model Features.

– **FedGNN's Role in Refinement.** FedGNN adopts a graph-based approach that leverages spatial relationships between objects for a deeper understanding. It refines YOLOv5's initial detections and improves accuracy.
– **Performance Indicators. CoRODDA** provides the following performance indicators:
  1. **Detection Accuracy.** It evaluates the correctness compared to ground truth annotations. Collaboration between YOLOv5 and FedGNN enhances the accuracy.
  2. **F1 Score.** An harmonic mean of precision and recall, providing a balanced detection performance measure.

3.3.2 Results To shed light on the confidence differences between the YOLOv5 and **CoRODDA** object detection algorithms, Figure 4 was charted. In this figure, the x-axis represents different object regions (or images), while the y-axis quantifies their corresponding scores. Blue bars denote YOLO's confidence, whereas green bars depict the scores from **CoRODDA** for images overlooked by YOLOv5.

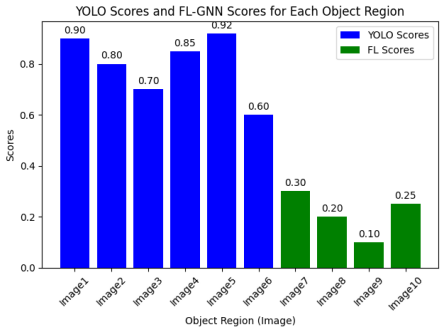

Fig. 4: Comparison of YOLO and FL-GNN Scores for Object Regions. Green bars correspond to objects missed by YoloV5.

The notable drop in confidence scores for **CoRODDA**, evident from the green bars, suggests that while it identifies objects YOLOv5 misses, it is often less confident in its detections. Objects may be missed by YOLOv5 due to factors like subtle variations in lighting, orientation, or occlusions, which might be more perceptible to **CoRODDA**. However, the framework's lower confidence could also arise from the challenges in dealing with such nuances. This comparison emphasizes the unique strengths and potential gaps of both algorithms, underscoring their combined potential in providing a holistic object detection system.

4 Conclusion

This paper has highlighted the significance of enhancing robotic responsiveness in dynamic settings. Our research journey began with an in-depth examination of the complexities of decision-making, which subsequently uncovered distributed object detection as a pivotal solution. Through **CoRODDA**, we illustrated the mechanism of robots capturing images, detecting objects, and systematically representing the detected object's coordinates on a graph. Intriguingly, the subsequent processing by FedGNN, which updates a central server in real-time, offers insights into potential optimizations in the realm of object detection. As the server meticulously analyzes the derived data, the prospects of enhancing object detection precision emerge more clearly, aligning seamlessly with the overarching goals we set out with. This work underscores the potential for refining robotic actions, adaptability, and interactions, especially in dynamic contexts. As we continue to push the boundaries of automated capabilities, the findings from this study provide a roadmap for future endeavors in this domain and we target to improve the federation process and apply **CoRODDA** on more benchmarks.

Acknowledgements

This work is partially funded under the ANR-21-ASRO-0005-01 agreement attached to the SCOPES project (ASTRID ASRO 2021 scheme funded by the Agence de l'Innovation de Défense (AID)).

Bibliography

[1] Samir Ouchani and Gabriele Lenzini. Generating attacks in sysml activity diagrams by detecting attack surfaces. *Journal of ambient intelligence and humanized computing*, 6:361–373, 2015.

[2] Keith Bonawitz, Hubert Eichner, Wolfgang Grieskamp, Dzmitry Huba, Alex Ingerman, Vladimir Ivanov, Chloe Kiddon, Jakub Konečnỳ, Stefano Mazzocchi, H Brendan McMahan, et al. Towards federated learning at scale: System design. In *Proceedings of the 2nd SysML Conference*, 2019.

[3] Erico Guizzo. The rise of the robot worker. *IEEE Spectrum*, 48(10):34–41, 2011.

[4] Fahem Zerrouki, Samir Ouchani, and Hafida Bouarfa. Quantifying security and performance of physical unclonable functions. In *2020 7th International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*, pages 1–4. IEEE, 2020.

[5] Yan Zhou and Oncel Tuzel. Towards safe autonomous driving: Capture uncertainty in the deep neural network for lidar 3d vehicle detection. *arXiv preprint arXiv:1804.05132*, 2018.

[6] Samir Ouchani. A security policy hardening framework for socio-cyber-physical systems. *Journal of Systems Architecture*, 119:102259, 2021.

[7] Philip N Howard and Muzammil M Hussain. Big data and the future of business. *Management Information Systems Quarterly*, 38(2):625–638, 2014.

[8] Jiahui Zhou. Convolutional neural networks explained. *Towards Data Science*, 2018. URL https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939.

[9] Jie Zhou, Ganqu Cui, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. Graph neural networks: A review of methods and applications. In *AI Open*, volume 1, pages 57–81, 2018.

[10] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3):50–60, 2020.

[11] Yue Tan, Yixin Liu, Guodong Long, Jing Jiang, Qinghua Lu, and Chengqi Zhang. Federated learning on non-iid graphs via structural knowledge sharing, 2022.

[12] Kai Hu, Jiasheng Wu, Yaogen Li, Meixia Lu, Liguo Weng, and Min Xia. Fedgcn: Federated learning-based graph convolutional networks for non-euclidean spatial data. *Mathematics*, 10 (6):1000, 2022.

[13] Chuizheng Meng, Sirisha Rambhatla, and Yan Liu. Cross-node federated graph neural network for spatio-temporal data modeling. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 1202–1211, 2021.

[14] Samir Ouchani, Yosr Jarraya, Otmane Ait Mohamed, and Mourad Debbabi. Probabilistic attack scenarios to evaluate policies over communication protocols. *J. Softw.*, 7(7):1488–1495, 2012.

[15] Souhila Badra Guendouzi, Samir Ouchani, and Mimoun Malki. Genetic algorithm based aggregation for federated learning in industrial cyber physical systems. In *The 15th International Conference on Computational Intelligence in Security for Information Systems (CISIS 2022) Proceedings*, pages 12–21. Springer, 2022.

[16] Souhila Badra Guendouzi, Samir Ouchani, and Mimoun Malki. Aggregation using genetic algorithms for federated learning in industrial cyber-physical systems. In *2022 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, pages 1–6. IEEE, 2022.

[17] Souhila Badra Guendouzi, Samir Ouchani, and Mimoune Malki. Enhancing the aggregation of the federated learning for the industrial cyber physical systems. In *2022 IEEE International Conference on Cyber Security and Resilience (CSR)*, pages 197–202. IEEE, 2022.