Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

Bonn-Aachen
International Center for
Information Technology

Sensor Intelligence.

Master Thesis Proposal

Federated Learning for Object Detection Using 3D Depth Images

Kevin Patel

Supervised by

Prof. Dr.-Ing. Sebastian Houben
Prof. Dr. Robert Lange
Dr. Markus Hammes
Dr. Nikolaus Mayer

September 2024

1 Introduction

1.1 Topic of This Thesis

- Provide reasonably detailed description of what you intent to do in your thesis project.
- You may also discuss the challenges that you have to address.
- Reflect on the profile of the reader and PLEAAAASE, tell a story here and refrain from bombarding the readers with details which they may not be able to appreciate.
- TODO: Put a story or a very basic scenario here to make the reader understand the problem.
- Keep it in industrial context
- Start with privacy issue with the centralized training especially in the consumer AI context
- Issue with scaling DL models
- Then how FL solves that problem, put a FL workflow diagram
- FL settings, what is our focus in that
- Industrial context
- Applicable to broader fields, medical, finance, surveillance
- Laws enforcing FL, put those papers
  - [1]
- From [2]
  - Large-scale machine learning models deployed in real-world scenarios often require training data sharing on a centralized server, where the

actual model optimization takes place. Federated Learning (FL) was initially introduced in [3] as an alternative approach to train a global model across multiple devices while preserving the privacy and decentralization of their respective data. [2]

- In particular, machine learning methods for computer vision heavily rely on collecting and storing a huge amount of annotated image data on a central server. Centralizing such data necessitates the transfer of a significant volume of information, resulting in substantial communication overhead. Moreover, centralized data storage poses risks to user privacy and confidentiality, and recent regulations on data privacy prohibit the uploading of sensitive local data to centralized data centers [1].
- Federated learning has emerged as a promising solution to address these challenges, as it enables on-device training of visual models. In FL, data remains localized on individual devices, and the collaborative training process involves exchanging model parameters instead of raw data.

1.2 Relevance of This R&D Project

- Who will benefit from the results of this R&D project?
- What are the benefits? Quantify the benefits with concrete numbers.

2 Related Work

2.1 Survey of Related Work

- What have other people done to solve the problem?
- You should reference and briefly discuss at least the "top twelve" related works
- Rising research trend for FL, put that fig with papers over years
- But mostly theoretical aspect on benefits of FL and its comparison with some toy datasets

- FL+CV, only classification is covered in most of the papers, show that paper published graphs
- A very few FL+OD papers, that to even in practical and application oriented are close to none
- The first seminal paper [3] , talks about Federated Learning(FL) as compared to centralized model training. This paper coined the term federated learning. The key focus of this paper privacy preserving machine learning models for consumer mobile devices. The object is to how to leverage the real world user data to train an efficient machine learning model without risking user's privacy. Also discussed difference between decentralized and federated learning.
  - The experiments proves that FL achieves comparable results to centralized training. The paper presented a basic CNN-based models on two datasets, MNIST and CIFAR10. The results are promising and show that FL learning paradigm can be used to take advantage of the large data and maintain privacy.
  - The paper presented federated aggregation algorithm called FedAvg, which aggregates individual clients models at the server-end.
- Limitations:
  - the presented results are on the tiny models and using toy datasets.
  - the main task was image classification only
  - Entire FL workflow implemnted on a single GPU-powered machine and FL is performed in a simulated fashion.

2.2 Limitation and Deficits in the State of the Art

- List the deficits that you have discovered in the related work and explain them such that a person who is not deep into the technical details can still understand them.
- For each deficit, provide at least two references

- You should reference and briefly discuss at least the "top twelve" related works
- None in the 3D depth image domain or intensity images
- Here, input data itself prioritize privacy especially when it comes to person recognition
- Most of the existing work focusing on the image classification task and very few of works explore the object detection task
  - Even in object detection, the models are old and recent advancement in the detection models are not addressed well
- Regarding the dataset, many of the work use toy datasets and lack the exploration on the real world dataset.
- Most studies on federated learning focus on its theoretical and communication aspects [4], [5], [6], and [7].

3 Problem Statement

- Which of the deficits are you going to solve?
  - Most of the existing work focusing on the image classification task and very few of works explore the object detection task
- What is your intended approach?
- How will you compare you approach with existing approaches?
- FL+OD working pipeline, on a real world dataset
- A non-RGB based dataset
- Implementation on edge-devices
- Compare FL frameworks and choose 1 or 2

- Check FL aggregation methods
- Analyze compute resource constrained OD models
- Use real world 3D depth images for object detection
- As seen in the Figure (TODO: add a figure with computer vision task papers), most of the literature work is done in federated learning (FL) for computer vision is for classification. FL with object detection is not yet well explored with the new edge hardware and the state-of-the-art models.

4 Project Plan

4.1 Work Packages

WP1 Literature Study

- WP1.1 Conduct a comprehensive literature review on state-of-the-art FL-based object detection methods
- WP1.2 Analyze existing centralized and FL-based object detection frameworks, focusing on their architectures and performance metrics
- WP1.3 Identify key models and techniques to replicate and compare, and document best practices for FL in the context of object detection on 3D depth images

WP2 Data Collection and Preparation

- WP2.1 Generate a custom dataset consisting of 3D depth images suitable for training and testing object detection models
- WP2.2 Preprocess the dataset, including data cleaning, augmentation, and handling any data imbalance issues
- WP2.3 Develop tools for managing and preparing the dataset for Federated Learning experiments (e.g., partitioning the dataset across simulated nodes)

WP3 Model Development and Initial Testing

- WP3.1 Replicate and test existing models from the literature to establish a baseline performance on the custom dataset
- WP3.2 Implement the FL-based object detection pipeline, starting with a centralized approach for comparison
- WP3.3 Test and evaluate initial models to ensure functionality and establish initial performance benchmarks

WP4 Comparative Analysis and Performance Evaluation

- WP4.1 Perform a comparative analysis between centralized and FL-based object detection methods, focusing on performance metrics such as accuracy, inference time, and system profile
- WP4.2 Evaluate models on the custom dataset and analyze the strengths and weaknesses of each approach
- WP4.3 Optimize model performance by tuning hyperparameters, modifying architecture, or adjusting the dataset for improved FL-based results

WP5 Implementation on Edge Devices

- WP5.1 Implement the FL-based object detection model on edge devices, ensuring it runs efficiently under resource constraints (e.g., limited computation, bandwidth)
- WP5.2 Test and evaluate the performance of the model on edge devices, measuring factors such as accuracy, inference time, and system profile
- WP5.3 Integrate any necessary optimizations for FL-based object detection specifically tailored to edge deployment scenarios

WP6 Advanced Development and Optional Extensions

- WP6.1 Develop a containerized workflow to facilitate easy deployment of the FL-based object detection pipeline across different environments

- WP6.2 If time permits, extend the project by integrating object tracking with FL-based object detection or adding uncertainty estimation capabilities to the models

WP7 Project Report and Finalization

- WP7.1 Write a comprehensive project report detailing the research objectives, methodology, results, and findings
- WP7.2 Present a critical analysis of the performance comparison between centralized and FL-based approaches, highlighting contributions, limitations, and future research directions
- WP7.3 Prepare the final thesis draft, ensuring clarity and coherence in presenting the experimental outcomes and their implications for FL-based object detection

4.2 Milestones

- M1 Comprehensive literature review on state-of-the-art FL-based object detection methods completed, with key models and techniques identified for replication
- M2 Replication of existing models from relevant literature completed
- M3 Custom 3D depth image dataset generated and preprocessed (e.g., cleaning, augmentation)
- M4 Comparative analysis of at least two object detection methods on the custom dataset completed, with detailed performance metrics recorded
- M5 Development of a working pipeline for FL-based object detection (simulated) on 3D depth images completed
- M6 Implementation of FL-based object detection on edge devices completed
- M7 Comparative analysis of centralized vs. FL-based object detection completed, with performance results and key findings documented

M8 Final model development and testing completed, including optimization and an in-depth evaluation of strengths and weaknesses

M9 Extension of the project with the development of a containerized workflow for easy deployment completed

M10 Final project report completed, covering methodology, experimental setup, results, and thorough analysis.

4.3 Project Schedule

WP ID	Task Description	Duration (Weeks)	Start Date	End Date	September	October	November	December	January	February
1	Literature Study	8	01.09.2024	31.10.2024						
2	Data Collection and Preparation	6	01.10.2024	15.11.2024						
3	Model Development and Initial Testing	8	15.10.2024	15.12.2024						
4	Comparative Analysis and Performance Evaluation	6	01.12.2024	17.01.2025						
5	Implementation on Edge Devices	6	15.12.2025	31.01.2025						
6	Advanced Development and Optional Extensions	4	01.02.2025	28.02.2025						
7	Project Report and Finalization	10	20.12.2024	03.03.2025						

Figure 1: Gantt chart of the project schedule

4.4 Deliverables

Minimum Viable

- Conduct a comprehensive literature review on state-of-the-art Federated Learning (FL)-based object detection methods
- Develop and test existing models to reproduce results from relevant literature
- Generation of a custom dataset consisting of 3D depth images
- Perform a comparative analysis of at least two methods on a custom dataset
- Produce a detailed project report that summarizes the work done and the obtained results

Expected

- Fulfill all minimum viable deliverables
- Perform a comparison of centralized versus FL-based object detection, focusing on performance analysis
- Complete the final development and testing of the model, and a detailed analysis of the strengths and weaknesses of each approach
- Develop a working pipeline for FL-based object detection on 3D depth images (simulated)
- Produce a more extensive project report detailing the methodology, experimental setup, results, and in-depth analysis

Desired

- Fulfill all expected deliverables
- Implement FL-based object detection on edge devices
- Develop a containerized workflow for easy deployment

- Additional objectives (if time permits):
  - Integrate object tracking with FL-based object detection
  - Implement FL-based object detection with uncertainty estimation

References

[1] Paul Voigt and Axel Von dem Bussche. The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, 10(3152676):10–5555, 2017.

[2] Donald Shenaj, Giulia Rizzoli, and Pietro Zanuttigh. Federated learning in computer vision. *IEEE Access*, 2023.

[3] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[4] Durmus Alp Emre Acar, Yue Zhao, Ramon Matas Navarro, Matthew Mattina, Paul N Whatmough, and Venkatesh Saligrama. Federated learning based on dynamic regularization. *arXiv preprint arXiv:2111.04263*, 2021.

[5] Nan Lu, Zhao Wang, Xiaoxiao Li, Gang Niu, Qi Dou, and Masashi Sugiyama. Federated learning from only unlabeled data with class-conditional-sharing clients. *arXiv preprint arXiv:2204.03304*, 2022.

[6] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. Federated optimization in heterogeneous networks. In *Proceedings of Machine Learning and Systems (MLSys)*, volume 2, pages 429–450, 2020.

[7] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank Reddi, Sebastian Stich, and Ananda Theertha Suresh. SCAFFOLD: Stochastic controlled averaging for federated learning. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*,

volume 119 of *Proceedings of Machine Learning Research*, pages 5132–5143. PMLR, July 2020.

[8] Author Name. Book title. *Lecture Notes in Autonomous System*, 1001:900–921, 2003.

[9] Bingyan Liu, Nuoyan Lv, Yuanchun Guo, and Yawen Li. Recent advances on federated learning: A systematic survey. *Neurocomputing*, page 128019, 2024.

[10] Deepthi Jallepalli, Navya Chennagiri Ravikumar, Poojitha Vurtur Badarinath, Shravya Uchil, and Mahima Agumbe Suresh. Federated learning for object detection in autonomous vehicles. In *2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService)*, pages 107–114. IEEE, 2021.

[11] Vinit Hegiste, Tatjana Legler, and Martin Ruskowski. Federated ensemble yolov5–a better generalized object detection algorithm. In *2023 Eighth International Conference on Fog and Mobile Edge Computing (FMEC)*, pages 7–14. IEEE, 2023.

[12] Hao Guan, Pew-Thian Yap, Andrea Bozoki, and Mingxia Liu. Federated learning for medical image analysis: A survey. *Pattern Recognition*, page 110424, 2024.

[13] Keval Doshi and Yasin Yilmaz. Federated learning-based driver activity recognition for edge devices. In *Proceedings of the IEEE/CVF Conference on computer Vision and Pattern Recognition*, pages 3338–3346, 2022.

[14] GeeksforGeeks. Collaborative learning - federated learning, Mar 2024.