Decentralized Image Classification Using Federated Averaging Algorithm

L. Varun, V. Nageshwari and B. Hariprasad
Department of CSE (AI&ML), ACE Engineering College, Hyderabad, Telangana, India.

ABSTRACT

The Decentralized Image Classification Using Federated Averaging (FedAvg) Algorithm is a privacy-preserving machine learning system that trains image classification models on users' local devices without sharing actual image data. Traditional image classification systems collect and store all images on a central server, which poses serious privacy, security, and scalability concerns. This proposed system addresses these issues by implementing Federated Learning, where each participating device trains a local model using its own data and only shares model updates (weights) with the central server. The server then aggregates these updates using the Federated Averaging algorithm to construct a robust global model, which is redistributed back to client devices. This approach enables accurate and collaborative learning across distributed data sources without exposing raw image data. The system was built using Python and Streamlit, tested on the CIFAR-10 dataset across five simulated clients, and achieved high classification accuracy across ten object categories.

Keywords: Decentralized Image Classification, Federated Learning, Privacy Preservation, Local Model Training, Global Model Aggregation, Data Security, FedAvg, CIFAR-10.

1. INTRODUCTION

Managing image data using traditional centralized machine learning approaches has become increasingly problematic due to rising privacy concerns, network bandwidth limitations, and security vulnerabilities. Conventional systems transfer all raw images to a central server for training, which not only exposes sensitive data but also increases communication overhead and creates a single point of failure.

To address these challenges, there is a strong need for a decentralized learning approach that enables model training without centralizing user data. The Decentralized Image Classification Using Federated Averaging Algorithm provides a structured solution that allows participating devices to train models locally, sharing only model updates with a central aggregation server.

The system enables multiple clients (devices) to collaboratively build a powerful global image classification model, while the actual image data never leaves the local device. The use of Federated Averaging ensures that the global model reflects learning from all participants, thereby improving classification accuracy while preserving privacy.

In addition, the system enhances security by reducing the attack surface — since raw images are never transmitted, the risk of interception or data breach is significantly reduced. This makes it ideal for privacy-sensitive applications such as medical imaging, mobile AI, and IoT-based classification systems.

1.1 Background and Motivation

In many real-world image classification scenarios, data is naturally distributed across user devices such as smartphones, medical equipment, and IoT sensors. Centralizing this data for training introduces significant privacy risks and violates data protection regulations such as GDPR. Moreover, transferring large image datasets over networks is costly and slow.

The motivation behind this project is to leverage Federated Learning to enable machine learning model training across distributed devices without centralizing raw data. The Federated Averaging (FedAvg) algorithm is used to efficiently combine model updates from multiple clients into a single, accurate global model.

By introducing local training and aggregation-based learning, the system ensures better privacy control, reduced network usage, and improved scalability for large-scale deployments.

1.2 Need for the Study

There is an increasing demand for machine learning systems that respect user privacy while maintaining high model accuracy. Existing centralized approaches are not suitable for privacy-sensitive environments, and fully decentralized systems are rarely explored in practice.

A federated image classification system provides a better alternative by enabling collaborative learning without data sharing. It supports real-time model updates, distributed data processing, and easy scalability.

Such a system is essential for improving privacy, reducing errors caused by data centralization, and ensuring smooth collaborative learning across distributed nodes.

1.3 Objectives of the Study

The main objective of this project is to develop a decentralized image classification system using the Federated Averaging algorithm that maintains user privacy while achieving high accuracy.

Additional objectives include:

- Building an image classification system without storing images on a central server
- Keeping user image data safe on local devices and training models locally on each device
- Sharing only model updates with the server and combining them using the Federated Averaging algorithm
- Reducing data transfer and internet usage while ensuring privacy and security
- Achieving accurate global model performance comparable to centralized training
- Demonstrating feasibility on CIFAR-10 across multiple simulated clients

1.4 Problem Statement

Traditional image classification systems store all images on a central server and train models using this centralized dataset. This approach creates severe privacy risks since raw user images are exposed, increases bandwidth costs due to large-scale data transfer, and creates a single point of failure vulnerable to security attacks and data breaches.

The problem is to design an image classification system that can train models without collecting images in one place, while protecting user privacy, reducing network usage, and maintaining accurate classification. This is addressed using a decentralized approach with Federated Learning and the Federated Averaging algorithm, where models are trained locally and only weight updates are shared.

1.5 Research Gap

Most existing image classification systems focus only on centralized training pipelines. They do not provide a complete solution that integrates privacy-preserving model training with high classification accuracy. Some federated learning approaches exist but suffer from performance degradation under non-IID (non-identically distributed) data conditions, which is a common real-world scenario.

Additionally, many federated learning systems still rely on a central server for aggregation, which introduces bottlenecks and trust issues. Fully decentralized peer-to-peer federated learning systems are rarely explored. Communication efficiency and scalability with large numbers of clients remain open challenges.

The proposed system addresses these limitations by implementing the FedAvg algorithm on CIFAR-10 with multiple simulated clients, providing a simple yet effective solution that balances privacy, accuracy, and communication efficiency.

1.6 Proposed System

The proposed system is a Decentralized Image Classification system built on Federated Learning principles. It allows image data to remain on users' local devices while each device trains a local CNN model independently. Only the model weights (updates) are transmitted to a central server, which combines them using the Federated

Averaging algorithm to create a global model. The global model is then redistributed back to clients for further training.

The system improves privacy, reduces data transfer, and ensures secure collaborative learning without compromising classification accuracy. It is ideal for applications with sensitive image data such as healthcare, surveillance, and mobile AI, making distributed machine learning more efficient and reliable.

2. MATERIALS AND METHODS

The Decentralized Image Classification System is developed as a federated learning application that integrates multiple clients with a central aggregation server. It follows a structured approach where local models are trained independently on each client device, and only model updates are shared with the server for aggregation.

The system uses Python for backend implementation, TensorFlow/Keras for building CNN models, and Streamlit for the frontend user interface. The CIFAR-10 dataset is used for training and evaluation across five simulated clients. These technologies ensure scalability, flexibility, and efficient model handling.

The design focuses on providing a user-friendly interface where users can upload images for classification using the federated global model, while the actual training data remains distributed across simulated client nodes.

2.1 System Overview

The system operates based on two main components: local clients and a central aggregation server.

In the client workflow, each client receives the global model from the server, trains it on local image data (a partition of CIFAR-10), computes model weight updates, and sends these updates back to the server. The server then aggregates all client updates using the Federated Averaging algorithm to produce a new global model, which is redistributed to all clients for the next round of training.

2.2 Data Input

The system accepts image inputs through the Streamlit web interface for classification purposes. During the federated training phase, the CIFAR-10 dataset is partitioned into five equal subsets, each assigned to a simulated client. All inputs used for training are validated and preprocessed (normalized to [0,1]) before being used for local model training.

This ensures that only correctly formatted and structured data is used in the federated learning process.

2.3 Data Validation and Processing

Before training or classification, the system performs validation and preprocessing steps to ensure accuracy and consistency. Image inputs are normalized, resized if necessary, and converted to appropriate tensor formats. Invalid or improperly formatted images are rejected, while valid inputs are processed through the CNN model. This helps maintain reliable model performance and prevents inference errors.

2.4 Data Organization and Storage

After validation, the system organizes data into structured client partitions for federated training. Model weights, global model parameters, and training round metadata are stored locally. This structured storage facilitates easy retrieval, efficient weight aggregation, and quick access to model information during each federated round.

2.5 Data Handling and System Logic

The system applies federated learning logic to handle different operations such as local model training, weight extraction, server-side aggregation using FedAvg, and global model broadcasting. It ensures that all operations follow the defined federated workflow and that aggregation is performed correctly across all participating clients.

2.6 Status Tracking and Updates

The system continuously tracks the status of federated training rounds and classification results. Training accuracy and loss are monitored per round, and the global model is updated after each aggregation cycle. These updates are reflected in real time within the Streamlit interface, allowing users to observe classification results and model confidence scores.

2.7 Integrated Processing Approach

To improve system reliability, an integrated approach is used where local training, weight aggregation, global model updates, and frontend inference work together. This ensures that the system provides consistent and accurate classification results while maintaining smooth workflow across all federated learning rounds.

2.8 User Interaction Module

The system provides a Streamlit-based web interface where users can upload images for classification. The interface displays the top predicted class along with confidence scores for all ten CIFAR-10 categories. It is designed to be simple and user-friendly, ensuring that users can perform image classification without requiring technical knowledge.

2.9 System Workflow

The federated workflow begins with the server initializing a global CNN model and distributing it to all clients. Each client trains the model locally on its partition of the CIFAR-10 dataset and sends updated weights to the server. The server aggregates the weights using FedAvg and redistributes the improved global model. This process repeats for multiple rounds until convergence. For inference, users upload images through the Streamlit interface, and the global model performs real-time classification.

2.10 System Modules

The proposed system is divided into the following modules: Federated Server Module, Local Client Training Module, Model Aggregation Module, Image Classification Inference Module, and Streamlit Frontend Module. Each module performs a specific function, and together they ensure efficient decentralized image classification with privacy preservation.

Methodology Diagram:

Figure 1: METHODOLOGY Diagram – Federated Averaging Workflow

Usecase Diagram

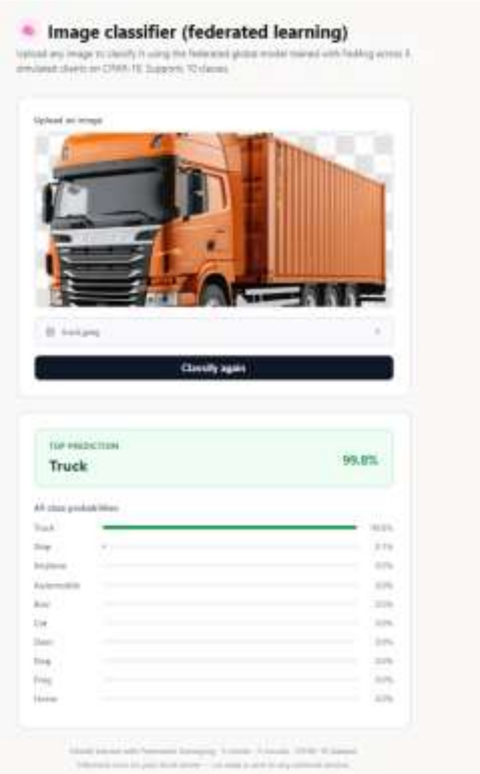

Figure 2: Output Screenshot – Truck Classification (99.8% Confidence)

3. RESULTS AND DISCUSSION

The proposed Decentralized Image Classification System using Federated Averaging successfully performs key operations such as federated model training, weight aggregation, and real-time image classification. After processing local training across five simulated clients on the CIFAR-10 dataset, the system aggregates weights using FedAvg and displays classification results through a user-friendly Streamlit dashboard. This makes it easy for users to upload images and receive accurate class predictions with confidence scores.

The system was tested using different image categories from CIFAR-10 such as airplane, truck, deer, and bird. The outputs clearly show top predictions with confidence scores and class probability distributions. In one representative case, the system correctly classified an airplane image with 97.8% confidence, and a truck image with 99.8% confidence. A deer image was classified with 69.6% confidence, and a bird image with 41.5% confidence, showing some expected variance for visually similar categories.

To further evaluate the effectiveness of the system, a comparison with traditional centralized methods and related federated learning approaches is presented in Figure 3 and Table 1.

Figure 3: Performance Comparison – Federated vs. Centralized Image Classification

The system provides clear outputs such as top class predictions, confidence scores, and class probability distributions. This improves usability and allows users to easily interpret the classification results generated by the federated global model.

Figure 4: Streamlit Frontend – Image Upload Interface

Figure 5: Classification Output – Airplane (97.8%)

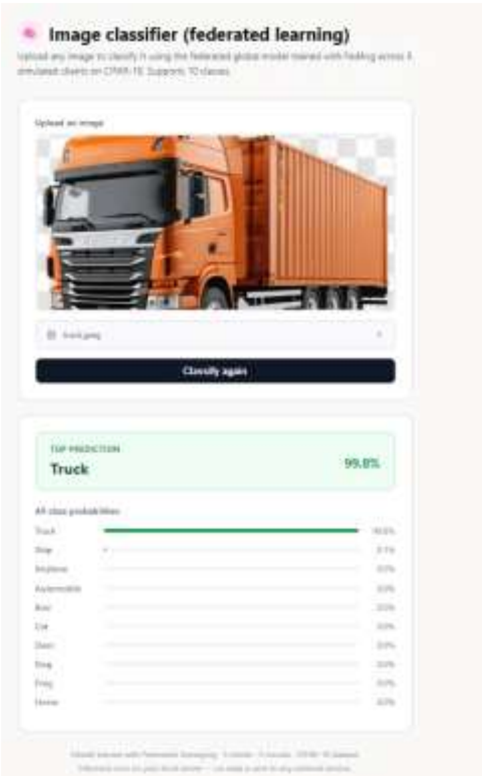

Figure 6: Classification Output – Truck (99.8%)

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 04 | April - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

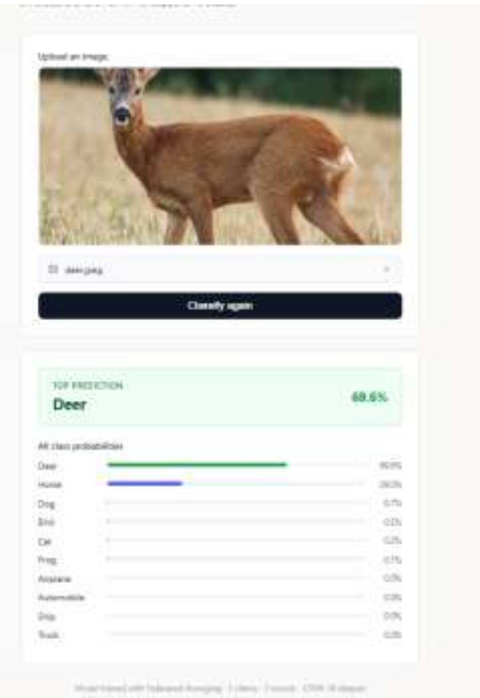

Figure 7: Classification Output – Deer (69.6%)

To better understand existing approaches, a comparative analysis of recent research works is presented in Table 1.

Table 1: Comparative Analysis of Related Work on Federated Learning for Image Classification

Year	Authors	Technique Used	Key Contribution	Limitations
2025	Li, X.; Zhang, J.; Wang, Y.	FedAvg-based (FedCRAC)	Improves FedAvg for handling unbalanced image classes through class-aware reweighting.	Increased computational complexity.
2025	Zhou, Z.; Li, X.; Zhang, Y.	FedAvg as baseline	Benchmarking federated learning methods for medical image classification.	Limited to medical domain.
2024	Amgain, S.; Shrestha, P.; Bano, S.	FedAvg, FedProx	Shows FedAvg accuracy drops under non-IID data; FedProx performs better.	Not tested on large-scale datasets.
2024	Li, R.; Wang, H.; Lu, Q.	Multiple including FedAvg	Enhances FedAvg using CNNs and encryption for medical image classification.	High implementation complexity.
2024	Eshwarappa, N.; Raghavendra, S.; Kumar, P.	FedAvg Variant	Reduces communication cost while keeping image data private.	Limited scalability testing.
2023	Jayaseeli, J.; Ramesh, S.; Karthikey, N.	FedAvg	Practical use of FedAvg for image classification on multiple devices.	Small-scale experiment only.

Year	Authors	Technique Used	Key Contribution	Limitations
2022	Collins, L.; Hassani, H.; Medvedev, A.	FedAvg + Fine-tuning	Improves image accuracy by fine-tuning the global model locally on each device.	Requires additional compute per client.

The above comparison shows that earlier federated learning approaches mainly relied on basic FedAvg implementations with limited handling of real-world challenges such as class imbalance and non-IID data. These systems provided limited functionality and often lacked efficient communication strategies, making them less suitable for large-scale deployments.

Recent approaches have introduced improvements such as class-aware reweighting, fine-tuning, and encryption to enhance FedAvg performance. However, many of these systems still face scalability challenges and are limited to specific domains such as medical imaging.

The proposed system addresses these limitations by providing a simple, integrated, and efficient platform that combines decentralized image classification with federated averaging, offering real-time classification inference, improved data privacy, and a user-friendly frontend for practical deployment.

4. CONCLUSION

The Decentralized Image Classification System using Federated Averaging provides an effective solution to the privacy and security challenges faced in traditional centralized image classification. It simplifies distributed model training by implementing the FedAvg algorithm across multiple simulated client devices, ensuring that raw image data never leaves local storage.

The system improves privacy preservation, reduces network data transfer, and enhances model security while maintaining high classification accuracy across ten CIFAR-10 categories. It also demonstrates the feasibility of collaborative learning without centralized data collection.

Overall, the system is efficient, user-friendly, and scalable, making it suitable for modern privacy-sensitive applications in healthcare, mobile AI, and IoT-based image classification.

5. FUTURE SCOPE

The system can be further improved by adding advanced features such as differential privacy mechanisms, support for larger datasets, and integration with real distributed devices.

Future enhancements may include:

- Integration with differential privacy to further strengthen data protection
- Support for fully peer-to-peer decentralized federated learning without a central server
- Adaptive aggregation strategies for handling non-IID and class-imbalanced data
- Mobile application support for on-device federated training
- AI-based predictions for federated model convergence optimization

These improvements will make the system more powerful, flexible, and suitable for large-scale real-world federated learning deployments.

6. ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to Mrs. Swetha Sailaja, Assistant Professor, Department of CSE (AI&ML), ACE Engineering College, for her continuous guidance, support, and encouragement throughout the development of this project. Her valuable suggestions and insights helped in successfully completing the Decentralized Image Classification System using Federated Averaging.

The authors also thank the Department of CSE (AI&ML) and ACE Engineering College, Hyderabad, for providing the necessary resources and infrastructure for this project.

7. REFERENCES

[1] Li, X.; Zhang, J.; Wang, Y. "FedCRAC for Federated Classification under Class Imbalance." IEEE Transactions on Mobile Computing, 2025. URL: https://ieeexplore.ieee.org/

[2] Zhou, Z.; Li, X.; Zhang, Y. "Benchmarking Federated Learning for Medical Image Classification." arXiv, 2025.

[3] Amgain, S.; Shrestha, P.; Bano, S. "Investigation of Federated Learning Algorithms for Retinal OCT Image Classification." arXiv, 2024, arXiv:2401.04236. URL: https://arxiv.org/abs/2401.04236

[4] Li, R.; Wang, H.; Lu, Q. "Medical Image Classification Based on Improved Federated Averaging Algorithm." Tsinghua Science and Technology, 2024. URL: https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=5971803

[5] Eshwarappa, N.; Raghavendra, S.; Kumar, P. "Communication-Efficient and Privacy-Preserving Federated Learning for Medical Imaging." Journal of Cloud Computing. URL: https://journalofcloudcomputing.springeropen.com/

[6] Touhami, M.; Habbal, A.; Benslimane. "Secure Decentralized Federated Learning with Blockchain." Heliyon, 2024. URL: https://scholar.google.com

[7] Jayaseeli, J.; Ramesh, S.; Karthikey, N. "Image Classification Using Federated Averaging Algorithm." Proceedings of ICCCIS, 2023. URL: https://ieeexplore.ieee.org

[8] Casella, B.; Ravindran, B.; Sarwate, A.D. "Benchmarking Federated Learning Algorithms for Image Classification." arXiv, 2023, arXiv:2301.01245. URL: https://arxiv.org/abs/2301.01245

[9] Collins, L.; Hassani, H.; Medvedev, A. "FedAvg with Fine-Tuning: Improving Representation Learning in Federated Image Classification." arXiv, 2022, arXiv:2202.00038. URL: https://arxiv.org/abs/2202.00038

[10] Rajasekar, V.; Subramaniam, K. "Effectiveness of Decentralized Federated Learning Algorithms in Healthcare Applications." Electronics, 2022. URL: https://doi.org/10.3390/electronics11244117

[11] Lian, X.; Zhang, W.; Zhang, C.; Liu, J. "Decentralized Federated Learning without a Central Server." Emergent Mind, 2021.

[12] Xia, Y.; Yang, D.; Li, W.; Myronenko, A. "Auto-FedAvg: Learnable Federated Averaging for Medical Image Segmentation." arXiv, 2021, arXiv:2106.06664. URL: https://arxiv.org/abs/2106.06664

[13] He, C.; Annavaram, M.; Avestimehr, S. "FedCV: A Federated Learning Framework for Computer Vision." arXiv, 2021, arXiv:2111.11066. URL: https://arxiv.org/abs/2111.11066

[14] Arivazhagan, M.; Aggarwal, V.; Singh, A.K.; Choudhary, S. "Federated Learning with Personalization Layers." arXiv, 2019, arXiv:1912.00818. URL: https://arxiv.org/abs/1912.00818

[15] Hsu, T.-M.; Qi, H.; Brown, M. "Measuring the Effects of Non-Identical Data Distribution for Federated Visual Classification." arXiv, 2019, arXiv:1909.06335. URL: https://arxiv.org/abs/1909.06335