

Center for Research in Computer Vision

UNIVERSITY OF CENTRAL FLORIDA

FINAL ORAL EXAMINATION

OF

Matías Andrés Mendieta
B.S., University of North Carolina at Charlotte, 2019
M.S., University of North Carolina at Charlotte, 2020

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

(COMPUTER SCIENCE)

1, JULY, 2024, 11:00 A.M.
HEC-356

DISSERTATION COMMITTEE

Dr. Chen Chen, *Chairman*	*chen.chen@ucf.edu*
Dr. Mubarak Shah	*shah@crcv.ucf.edu*
Dr. Ser-Nam Lim	*sernam@ucf.edu*
Dr. Mingjie Lin	*milin@ucf.edu*

DISSERTATION RESEARCH IMPACT

In recent years, the application of deep learning in computer vision has significantly transformed the global technological landscape. The introduction of AlexNet marked a pivotal moment, driving a rapid expansion in the adoption and utilization of computer vision technologies across numerous industries. However, as deep learning methodologies have become more prevalent and diversified in their applications, there has been a corresponding increase in the complexity and resource demands of these approaches. This includes the development of larger model architectures, the need for greater volumes of training data, and the requirement for enhanced computational power to effectively train and deploy these models.

To address these challenges and ensure the continued applicability and real-world impact of deep learning systems, it is essential to develop innovative methods that enhance the efficiency of these applications without compromising their effectiveness and performance. In this dissertation, various methods are proposed to tackle these issues head-on in both distributed and centralized settings. Novel training paradigms and frameworks are developed to overcome current limitations in memory, computation, labeled data, and communication. These advancements aim to pave the way for more accessible and efficient deep learning solutions.

SELECTED PUBLICATIONS

Total Citations: 882, h-index 12

1. **Local Learning Matters: Rethinking Data Heterogeneity in Federated Learning,** Matías Mendieta, Taojiannan Yang, Pu Wang, Minwoo Lee, Zhengming Ding, Chen Chen, *in CVPR 2022, Oral, Best Paper Finalist*.
2. **Towards Geospatial Foundation Models via Continual Pretraining,** Matías Mendieta, Boran Han, Xingjian Shi, Yi Zhu, Chen Chen. in *ICCV, 2023*.
3. **PGFed: Personalize Each Client's Global Objective for Federated Learning,** Jun Luo, Matías Mendieta, Chen Chen, Shandong Wu, *in ICCV 2023, Oral*.
4. **FedPerfix: Towards Partial Model Personalization of Vision Transformers in Federated Learning.** Guangyu Sun, Matías Mendieta, Jun Luo, Shandong Wu, Chen Chen, *in ICCV 2023.*
5. **FeatER: An Efficient Network for Human Reconstruction via Feature Map-Based TransformER.** Ce Zheng, Matías Mendieta, Taojiannan Yang, Guo-Jun Qi, Chen Chen, *in CVPR 2023*.
6. **Navigating Heterogeneity and Privacy in One-Shot Federated Learning with Diffusion Models,** Matías Mendieta, Guangyu Sun, Chen Chen, *submitted*.

DISSERTATION

EFFICIENT AND EFFECTIVE DEEP LEARNING METHODS FOR COMPUTER VISION IN CENTRALIZED AND DISTRIBUTED APPLICATIONS

In the rapidly advancing field of computer vision, the application of deep learning has sparked significant technological transformations. However, the broad deployment of these technologies often faces efficiency challenges, including high memory usage, demanding computational resources, and extensive communication overhead, particularly in distributed and centralized learning systems.

Therefore, efficiency has emerged as a crucial factor for both centralized and distributed applications of deep learning technologies. This is essential for ensuring scalability, real-world applicability, and ultimately facilitating its democratization and widespread accessibility. For example, in distributed settings, federated learning (FL) presents a powerful strategy by enabling collaborative model training across multiple clients while upholding the privacy of their data. This approach, while promising, underscores pivotal challenges due to each client's operation under significant constraints, including limited memory, computational power, and bandwidth. These constraints become particularly severe in large-scale FL systems where communication overhead escalates into a critical bottleneck, underscoring the necessity for enhanced efficiency in memory, computation, and communication to render distributed paradigms like federated learning both viable and effective. Conversely, centralized training systems also demand high efficiency, particularly focusing on optimizing compute resources during training and inference phases. In specialized fields such as remote sensing, where labels are often scarce for downstream tasks, label efficiency is critical for success. Enhancing both compute and label efficiency can significantly influence the performance and practicality of centralized training models. Addressing these efficiency challenges holistically in both federated learning and centralized training systems can lead to significant advancements, supporting more extensive and effective deployment of machine learning models across varied domains.

To this end, this dissertation first tackles the challenges inherent in the distributed setting of federated learning, where the decentralized training process and data distributed across multiple clients often result in significant inefficiencies in communication, memory, and computational power. To combat these issues, a novel method is introduced to improve handling of data heterogeneity through distillation-based regularization. This optimizes local model performance while reducing memory and computational demands, facilitating more scalable and practical deployments of federated learning. Further addressing communication inefficiencies, a novel approach is presented that integrates diffusion models within a one-shot federated learning framework, focusing on reducing communication costs by minimizing the frequency of model updates needed across clients. This method not only remains robust against varied data distributions but also incorporates differential privacy to protect user data, enhancing its suitability for privacy-sensitive applications. In the centralized domain, the dissertation develops a novel training paradigm for geospatial foundation models. A multi-objective continual pretraining strategy is designed that significantly enhances training efficiency on geospatial tasks, improving label efficiency and reduce the computational resources required for training large-scale models. This approach provides a robust framework for the effective deployment of deep learning technologies in geospatial applications, where data can be particularly challenging to label and process.

Overall, this dissertation provides substantive advancements in deep learning system efficiency, emphasizing improvements in memory utilization, computational demands, and communication efficiency. These contributions are crucial for the scalable and effective application of deep learning in both distributed and centralized environments.



Matías Andrés Mendieta

2016-2019	B.S., University of North Carolina at Charlotte
2019-2020	M.S., University of North Carolina at Charlotte
2020-2021	Ph.D. student, University of North Carolina at Charlotte
2022	Applied Scientist Intern, Amazon
2023	Machine Learning Intern, Apple
2021-2024	Ph.D., University of Central Florida