Grenze International Journal of Engineering and Technology, June Issue

Federal Learning and Deep Learning Framework for MRI and Facial Expression Single Bases Multi Modal of Mental Illness

Harshita Rajput[1], Swati[2], Vanshika Gangwar[3], Shaurya Shukla[4], Arun Kumar Singh[5] and Arjun Singh[6]
[1-6]Greater Noida Institute of Technology, CSE, Greater Noida, India
Email: harshitarajputxxvi@gmail.com, swatikumari07180@gmail.com, vanshikagangwar03@gmail.com, shouryashukla1405@gmail.com, arun.k.singh.iiit@gmail.com

***Abstract*— Mental Health disorders are becoming highly common, specifically among students and young adults who face daily pressure with respect to academics, career and personal life affecting personal balances both emotionally and physically, change in behavior and very less social interaction. Medical images such as MRI scan and easily seen facial expression compel us to have insights of mental condition of a person. Although deep learning techniques is considered effective in image recognition but simultaneously this old centralized method raise serious issues regarding security, data privacy and valid use of medical data. The objective of this study is to overcome these challenges through conceptual federated multimodal framework that aligns face expression with MRI image for better detection of mental situation.**

Through this approach we solve the data security and above stated technical issues as data remains within the institution while only trained model share updates during training. The need of the study is not to derive result from single resource or an experiment but to explain collaborative working of federated learning and multimodal deep learning, have a view on their advantages and disadvantages and check for future challenges.

***Keywords*— deep learning, MRI, facial expression, multimodal, federated learning, mental health.**

I. Introduction

Mental illness affects a large part of the population both domestic and global, creates a strong need for necessary and early detection methods. Recent advancements in artificial intelligence and medical imaging technologies have made it possible to analyze neurological and behavioral patterns associated with mental health conditions [1][3][4]. Recent progress in medical imaging and facial analysis has made it possible to support mental health assessment using the improved techniques. MRI data gives valuable information about brain structure, but relying on MRI alone may not be enough. When the brain image data is combined with facial expression analysis, a more complete picture of mental health can be obtained. Federated learning has emerged as a privacy-preserving machine learning paradigm where multiple institutions collaboratively train a shared model while keeping their local datasets secure. Instead of transferring raw data, only model parameters or updates are communicated to a central server, which significantly reduces the risk of data leakage in sensitive domains such as healthcare [5][7][8].

Grenze ID: 01.GIJET.12.2.222_20

Federated learning allows numerous institutions to train machine learning model without exchanging raw data, making it reliable for healthcare department. Deep learning help by automatically extracting complex patterns from both MRI scans and facial expressions. This work was motivated by the observation that old observations in world either focus on a single type of data or depend on centralized systems, which limits their practical use in real clinical environments where privacy is a major concern.

II. Background and Related Work

In practice, federated learning can reduce differences across institutions without sharing patient data, especially in MRI-based mental health analysis. These methods have been applied to conditions such as autism, ADHD, schizophrenia, and depression using multi-site brain imaging datasets. While such approaches report encouraging results, they often focus only on MRI data or require carefully aligned data distributions across centers. Previous research has demonstrated that federated learning can effectively support collaborative analysis of medical imaging data collected from different hospitals. Such approaches help overcome data-sharing limitations while enabling models to learn from geographically distributed datasets [2][3][6].
From a clinical point of view most institution will focus on MRiI image more than facial behavior. Even after being a strong reflection of emotional and psychological states, facial expressions receive less importance at clinical level in such models. This gap shows the need for frameworks that jointly consider brain imaging and facial behavior and also preserve privacy.

A. Mental Illness and Multimodal Data

Integrating multiple sources of information, such as brain imaging and facial behavioral signals, allows machine learning systems to capture a broader representation of a patient's mental condition. Multimodal analysis has been reported to improve diagnostic accuracy compared to approaches relying on a single type of data [12][13].In the case of major depressive disorder, researchers have stated federated methods that combine different types of MRI data while adjusting for differences across medical centers. These methods improve diagnosis accuracy while maintaining data privacy. Facial expression information, which can reveal emotional and mental states, has received much less attention. This gathered the need for researchers to work on model that focus on brain images as well as facial behaviour expression

B. Deep Learning in Mental Health

Deep learning techniques, particularly convolutional neural networks, have proven highly effective in analyzing visual data such as facial expressions. These models automatically extract hierarchical features from images, enabling the detection of subtle emotional and neurological indicators associated with mental health disorders [10][11].
These methods are broadly considered because they do not require in depth manual interpretations in the model technology.
Different studies showed that Schizophernia can be detected majorly through facial recordings . These models are able to distinguish patients from healthy individuals by learning minute facial movement and its patterns.

C. Federated Learning Principles

Federated learning is a way for different groups to work together to train one machine learning model without sharing their actual data. Each group trains the model on their own computer using their own data, and then they only send the results (like updates or patterns the model learned) to a central system. This way, the model gets smarter from everyone's data, but the data itself never leaves its home. "This approach solves big problems in medical research, like keeping patient information private, making sure data stays with its owner, and handling differences in data from different hospitals. Because of these strengths, federated learning is a good fit for mental health studies that use sensitive and varied data.

III. Methodology

This methodology adopts a learning strategy that enable MRI scan through deep learning methods by CNN and 3D CNN to observe nerves pattern associated with mental disorders , improving efficiency when working with large public dataset Public neuroimaging repositories, including ADNI and OpenNeuro, provide valuable datasets that support large-scale research in medical imaging and neurological disorder analysis. These datasets enable researchers to develop and evaluate machine learning models using standardized and widely accepted data sources [14][15].

A. Facial Expression Analysis

Facial expression data is not widely used in federated learning because it is highly personal and difficult to manage. A person's face directly reveals identity, which makes hospitals and research institutions more cautious, even when the data never leaves their local systems.

Concerns related to misuse, consent, and legal responsibility often slow down adoption. In addition, facial expression data collected in real-world settings is highly variable. Differences in cameras, lighting conditions, viewing angles, and environments lead to inconsistent data quality across institutions. Facial expressions are also influenced by culture, age, personality, and social behavior. As a result, people with similar mental health conditions may display very different facial cues.

Labeling facial expressions for mental health assessment is also subjective and often depends on expert judgment. For these reasons, facial expression data remains less explored in federated learning, even though it holds strong potential when used carefully.

B. Multimodal Fusion Strategies

To improve mental illness prediction, information from MRI and facial expression data is combined using different fusion methods

Each fusion method helps the system better understand the relationship between brain abnormalities and emotional behavior.

C. Federated Model Aggregation

The system uses Federated Averaging (FedAvg) to update the global model. In this method, model updates from each institution are combined based on how much data each one has. This process is repeated over several rounds until the model becomes stable and accurate.

D. Importance of the Framework

This federated and multimodal deep learning framework:

- Protects patient privacy
- Combines biological (MRI) and behavioral (facial expression) data
- Learns from multiple institutions
- Improves mental illness analysis accuracy

IV. Performance Analysis

Prior studies evaluating federated learning in healthcare applications suggest that collaborative training across institutions can achieve competitive predictive performance while maintaining stronger privacy protection compared to traditional centralized models [1][4][8].

Instead of reporting exact experimental outcomes, the performance values presented here are intended to show general performance trends observed in similar studies. Federated multimodal models are often reported to achieve more stable learning and more stable results as compared to traditional methods, while also offering stronger privacy protection compared to centralized systems.

These results help illustrate how combining MRI and facial expression data in a federated setup can improve learning behavior across institutions. These findings should considered as comparative results instead of precise numerical results from a single dataset.

The main algorithm used in this framework is Federated Averaging (FedAvg). In FedAvg, each participating site (such as a hospital) trains a local deep learning model on its own MRI and facial expression data. The model results are sent to a central server, where they are combined based on the size of each local dataset, improving data confidentiality and improved results.

The evaluation metrics are calculated as follows:

A. Accuracy Measures the Overall Correctness of the Model

$$\text{Accuracy} = \frac{(\text{Number of Correct Predictions})}{(\text{Total Number of Predictions})}$$

- TP (True Positives): correctly predicted positive cases
- TN (True Negatives): correctly predicted negative cases
- FP (False Positives): incorrectly predicted positive cases
- FN (False Negatives): incorrectly predicted negative case

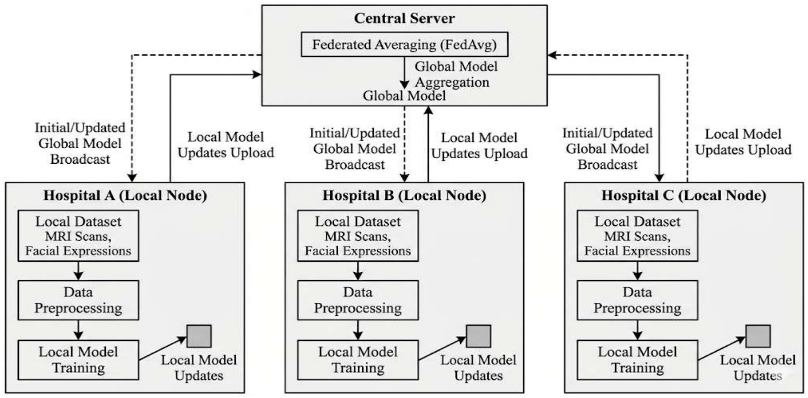

Fig 1: Federated learning architecture for multimodal mental health detection

B. Precision Shows how Many Predicted Positive Cases are Actually Correct

$$\textbf{Precision} = \frac{\textbf{TP}}{\textbf{TP+FP}}$$

- TP (True Positives): cases the model correctly predicted as positive
- FP (False Positives): cases the model incorrectly predicted as positive

C. Recall indicates how well the model identifies true mental illness

$$\textbf{Recall} = \frac{\textbf{TP}}{\textbf{TP+FN}}$$

- TP (True Positives): cases the model correctly predicted as positive
- FN (False Negatives): cases the model missed (it predicted negative, but they were actually positive)

D. F1-score Balances Precision and Recall for more Reliable Evaluation

$$\text{F1} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision} + \text{Recall}}$$

To avoid confusion, it is important to clarify that the performance values reported in this study are not actually derived from single real world experiment. The accuracy, F1-score, and AUC values are used to show relative performance trends between different learning approaches. The reported results are based on publicly available datasets that are widely discussed in mental health and medical imaging research. These values represent expected performance ranges observed in similar studies, rather than claims of exact experimental outcomes from a specific dataset. By presenting the results in this way, the focus remains on comparing methods and understanding the benefits of federated multimodal learning, instead of reporting dataset-specific numerical achievements.

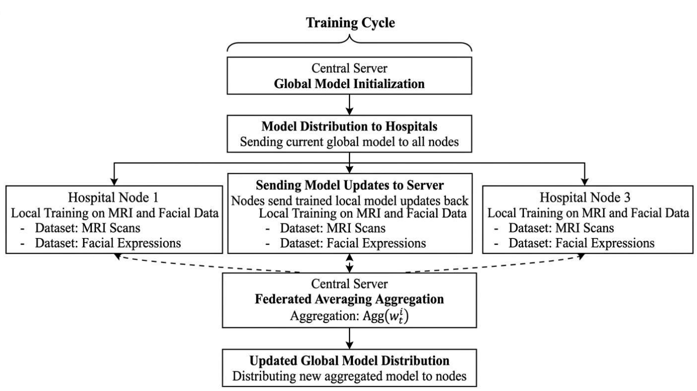

Fig 2. Federated Learning Training Workflow

V. RESULT

The findings show that federated multimodal methods usually work better than centralized or single modal models. This is because they combine brain scans with facial expression data while still keeping patient information private. Centralized models can perform fairly well, but they require sharing sensitive medical data. Single modal models only use one type of information, so they are less accurate. Overall, federated multimodal methods seem more suitable for mental health research where privacy is important. Still, more testing in real-world clinical settings is needed to prove how effective they are.

Real-World Clinical Relevance: Federated multimodal learning is very useful in healthcare. Hospitals can work together to train deep learning models without sharing raw MRI scans or facial images, which helps follow privacy rules like HIPAA and GDPR.

By combining data from different hospitals, the model becomes better at handling:

- Different patient groups
- Different MRI machines
- Different ways facial expressions are recorded

This solves a big problem with centralized models, which are usually trained on smaller, more similar datasets.

Challenges and Future Directions

Even though federated learning has clear advantages, it also faces challenges:

- Extra communication is needed between hospitals and the central server
- Data differences across hospitals (non-IID data) can make training harder
- Risks from security threats and system reliability issues

Researchers are now working on:

- Cutting down communication costs
- Improving how models combine updates (aggregation strategies)
- Making predictions easier for doctors to understand and trust

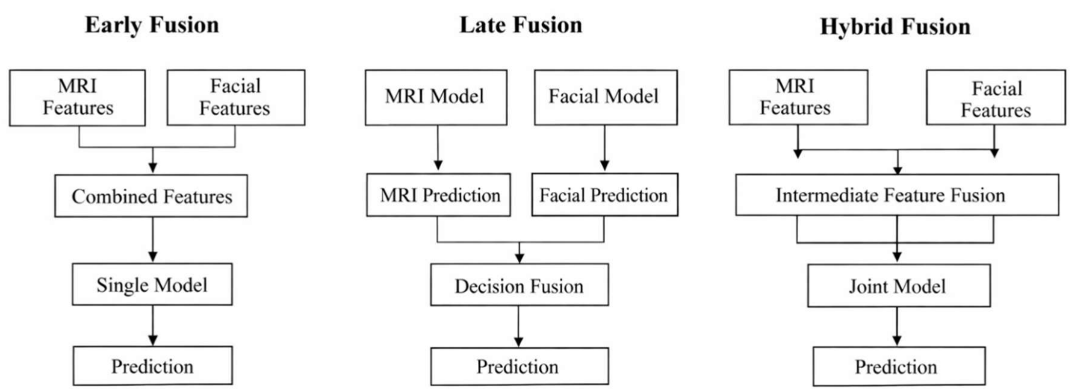

Fig 3. Multimodal fusion strategies

VI. CONCLUSION

This study shows that federated multimodal deep learning has strong potential for analyzing mental illness. By combining MRI scans with facial expression data, the method protects patient privacy while allowing hospitals to learn together without sharing raw data. Overall, federated multimodal learning represents a promising direction for privacy-aware healthcare analytics. By enabling collaborative model development without direct data sharing, this approach has the potential to improve mental health assessment while respecting strict data protection requirements [2][5].

Future Research Directions: To make this approach more practical, future work should focus on:

- Expanding the framework to include more hospitals and clinics
- Improving algorithms to handle differences in data across institutions
- Making the system more reliable when some hospitals are slow or temporarily offline
- Adding new data types (speech, voice tone, text-based behavior) to improve diagnosis
- Building standard datasets and benchmarks to speed up research and clinical adoption.

TABLE I. SCOPE OF FUTURE DIRECTION

Aspect	Simple Explanation	Future Direction
Privacy Protection	Patient MRI and facial data stay inside hospitals; only model updates are shared	Use stronger privacy methods like differential privacy and secure aggregation
Diagnostic Accuracy	Combining MRI and facial expressions gives better prediction results	Add more data types such as voice and text
Scalability	The federated system can work across many hospitals	Reduce communication cost and improve aggregation efficiency
Data Diversity	Data from different hospitals helps the model generalize better	Design adaptive algorithms for non-uniform (non-IID) data
Clinical Adoption	Meets healthcare data laws like HIPAA and GDPR	Develop shared benchmarks and open federated datasets

REFERENCES

[1] Li, X., Zhang, Y., & Wang, S. (2023). Federated learning for mental disorders' detection, diagnosis, and treatment: A review. Applied Intelligence. https://link.springer.com/article/10.1007/s10489-023-05095-1

[2] Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., & Wang, F. (2024). Federated learning in mental health and human activity recognition: A systematic review. Frontiers in Digital Health. https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2024.1495999/full

[3] Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., & Milchenko, M. (2020). Federated learning in medical imaging: A survey. IEEE Transactions on Medical Imaging, 39(2), 400-410. https://www.mdpi.com/2079-9292/13/1/47

[4] Kaissis, G., Makowski, M. R., Rückert, D., & Braren, R. F. (2021). Secure, privacy-preserving and federated machine learning in medical imaging. Nature Machine Intelligence, 3, 506-516. https://pmc.ncbi.nlm.nih.gov/articles/PMC11368680/

[5] Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhao, S. (2021). Advances and open problems in federated learning. Foundations and Trends® in Machine Learning, 14(1–2), 1–210. https://arxiv.org/abs/1912.04977

[6] Sheller, M. J., Reina, G. A., Edwards, B., Martin, J., & Bakas, S. (2020). Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, 92-104. https://arxiv.org/abs/1810.05287

[7] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 10(2), 1-19. https://dl.acm.org/doi/10.1145/3298981

[8] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 37(3), 50-60. https://arxiv.org/abs/1908.07873

[9] Tjoa, E., & Guan, C. (2020). A survey on explainable artificial intelligence (XAI): Toward medical XAI. IEEE Transactions on Neural Networks and Learning Systems, 32(11), 4793-4813. https://arxiv.org/abs/1907.07374

[10] Li, S., Deng, W., & Du, J. (2018). Reliable crowdsourcing and deep locality-preserving learning for facial expression recognition. IEEE Transactions on Image Processing.

[11] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[12] Calhoun, V. D., & Sui, J. (2016). Multimodal fusion of brain imaging data. Nature Neuroscience.

[13] Dhamala, E., Jamison, K., & Calhoun, V. (2021). Multimodal data fusion for neuroscience and psychiatry. Nature Reviews Neuroscience.

[14] Poldrack, R. A., et al. (2019). The OpenNeuro resource for sharing neuroimaging data. Nature Neuroscience.

[15] Jack, C. R., et al. (2008). The Alzheimer's Disease Neuroimaging Initiative (ADNI). Journal of Magnetic Resonance Imaging.