

MCFT: Multimodel constrictive fusion transformer for classification of hyperspectral image and LIDAR data.

K. Anusha¹, Sk. Ayesha Julka², N. Yaksheetha³, V. Vijaya Lakshmi⁴, V. Pavithra⁵

Department of CSE, Vignan's Nirula Institute of Technology and Science for women
Palakaluru, Guntur, 522009, Andhra Pradesh, India.

Abstract

In domains including agriculture, urban planning, and crisis management, the precise classification of land cover in remote sensing (RS) photographs is essential. Although hyperspectral images (HSIs) offer a wealth of spectral information that makes comprehensive material identification possible, they frequently are unable to differentiate between objects that have different heights but identical spectral signatures. However, HSI is successfully supplemented with Light Detection and Ranging (LiDAR) data, which records three-dimensional altitude and structural information. However, because of their different data formats and semantic differences, combining these two heterogeneous modalities is still a difficult endeavour. We suggest MCFT (Multimodal Constrictive Fusion Transformer), a transformer-based system intended to efficiently fuse HSI and LiDAR data for precise land cover categorisation, in order to overcome this constraint.

In order to acquire intricate intermodality correlations while maintaining modality-specific features, the model presents Modality Fusion Blocks (MFBs) that incorporate Cross-Attention (CA) and Spectral Self-Attention Modules (SSAMs). In order to steer spectral improvement and supplement HSI features with accurate altitude information, LiDAR-derived attention maps are employed. By performing early joint feature extraction using multi-head attention, MCFT produces a more expressive and semantically aligned representation than traditional hierarchical or concatenation-based fusion techniques, which frequently suffer from feature imbalance and the Hughes phenomenon. According to experimental results, MCFT greatly enhances classification performance and generalisation ability, demonstrating its usefulness for real-world applications where both structural and spectral insights are essential, such as vegetation monitoring, flood risk mapping, infrastructure inspection, and smart city development.

Keywords: Multimodal Fusion, Hyperspectral Imaging(HIS), LiDAR Data, Transformer Architecture, Cross-Attention(CA), Spectral Self-Attention Module(SSAM), Modality Fusion Block(MFB), Federated Reinforcement Learning(FRL), Contrastive Learning, Land Cover Classification.

Introduction

In fields like agriculture, urban planning, and disaster response, the ability to identify and categorize objects in remote sensing (RS) images is essential [1] [2]. Although hyperspectral images (HSIs) provide rich spectral information for classifying land cover [3], they have trouble differentiating objects that have different heights but similar spectral signatures. By resolving structural differences, LiDAR data which records 3D altitude information complements HIS [4]. However, because of their disparate data structures and semantic content, integrating these heterogeneous modalities the spatial-altitude data from LiDAR and the spectral-spatial features from HIS presents difficulties [5] [6].

Conventional fusion techniques, such as hierarchical fusion or weighted concatenation, frequently have suboptimal classification accuracy due to feature imbalance and the Hughes phenomenon [7]. Promising answers are provided by recent developments in deep learning, especially in the area of attention mechanisms [8]. Although cross-attention (CA) allows models to concentrate on modality-specific features, many methods abuse it by ignoring semantic differences and only using dot-product similarity [9] [10].

Our project is transformer-based model with stacked Modality Fusion Blocks (MFBs) that incorporates CA and spectral self-attention modules (SSAMs) in order to overcome these drawbacks [11]. In order to improve fusion efficacy, our model uses LiDAR-generated attention score maps to enrich HSI features with altitude data [12] [13]. Our project maximizes the spectral richness of HSI by utilizing multi-head attention to jointly extract and fuse these features early, in contrast to traditional methods that separate spectral and spatial extraction [14] [15].

For applications like flood risk mapping, wildfire prevention, and smart city development that demand in-depth structural and material insights, this strong fusion approach is essential [16] [17]. LiDAR integration is crucial for tasks like infrastructure inspection because HSIs, despite their spectrally rich nature, frequently lack spatial resolution [18]. Improves land cover classification accuracy and practical utility by guaranteeing a more expressive and trustworthy representation of multimodal RS data [19] [20].

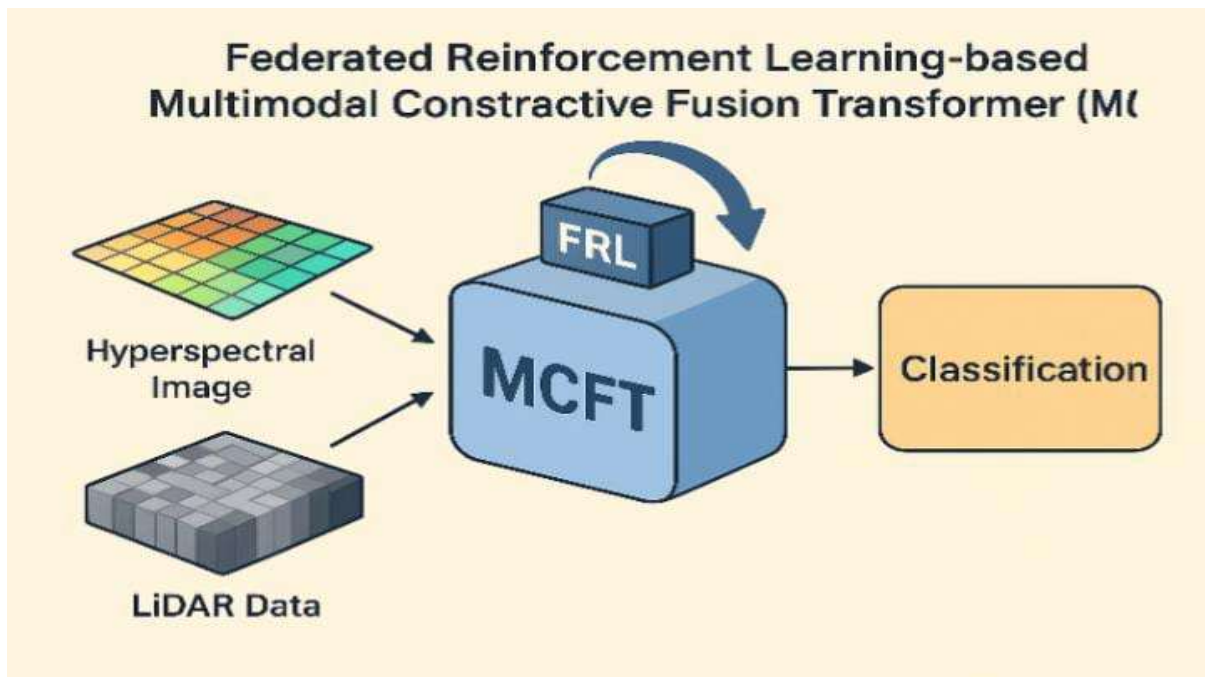


Fig a: Federated Reinforcement Learning-based Multimodal constrictive fusion Transformer

This model improves land cover classification by fusing hyperspectral and LiDAR data using Discriminant Correlation Analysis (DCA) [21]. Group-wise spatial feature extraction is made possible by the grouping of HSI bands according to joint entropy and structural similarity [22]. DCA reduces dimensionality while maintaining discriminative power by fusing LiDAR altitude features with HSI spectral and spatial features in parallel. By aligning heterogeneous features from both modalities, this method improves classification accuracy [23].

The fusion of hyperspectral and LiDAR data is guided by attention mechanisms in this deep learning model [24]. It highlights significant areas in LiDAR data with spatial attention and highlights informative bands in HSI with spectral attention [25]. The network can learn complementary features across modalities thanks to the multi-level fusion [26]. By adaptively weighting features according to their relevance, this model enhances classification performance [27].

CNNs, RNNs, and transformer-based models are among the deep learning architectures used for HSI-LiDAR fusion that are covered in this systematic review [28]. It draws attention to issues such as spectral-spatial heterogeneity, feature imbalance, and the requirement for explainability [29]. Emerging trends for multimodal RS classification are also covered in the review, including self-supervised learning, multi-head transformers, and cross-attention [30].

Conventional building boundary extraction techniques, such as sequential techniques based on alpha shapes, depend on local spatial relationships and geometric parameters that are manually adjusted [31]. Because of fixed or empirically adjusted α values, these methods frequently have trouble generalizing across a variety of urban environments, resulting in overly smoothed or fragmented contours [32]. Furthermore, hyperspectral imagery (HSI), which is essential for differentiating materials with similar structures but distinct spectral properties, is ignored in favor of LiDAR's spatial geometry [33] [34].

By using deep learning to automatically learn hierarchical spectral-spatial correlations, the proposed Multimodal Contrastive Fusion Transformer (MCFT) overcomes these drawbacks [35]. In order to manage non-linear cross-modal interactions and maintain fine-grained boundary details, it combines HSI and LiDAR data [36]. In contrast to conventional models, MCFT uses both spatial depth and spectral richness at the same time to enable more precise and flexible urban classification [37]. This makes it particularly useful in intricate situations requiring a joint analysis of structural geometry and material composition [38].

Literature survey

Accurate land cover classification has made the integration of diverse data sources, such as LiDAR and hyperspectral imaging (HSI), more crucial [39]. The capacity of current HSI-LiDAR fusion techniques to capture comprehensive structural-spectral correlations is limited because they frequently split spectral and spatial data. To solve these problems, recent research has suggested multiscale and transformer-based attention techniques [13]-[15]. The S2CA framework used multiscale convolution and feature recalibration to improve intermodal interaction and reduce noise [2], while a modality fusion vision transformer that combined cross-attention and spectral self-attention improved joint feature learning [1] to achieve 99.91% accuracy [16]. High-dimensional multimodal representations have been successfully extracted from HSI and LiDAR data by adapting pretrained vision-language models (VLMs) for joint classification [3]. Hybrid architectures that include transformers and hierarchical CNNs have been proposed to improve classification performance and better capture complicated spatial-spectral correlations in urban contexts [40]. In order to overcome label changes between views and improve the robustness of multimodal classification, other works place an emphasis on multilevel and multiview fusion algorithms [5][18]. Modality-specific representation was improved with the advent of HFSNet, a height-guided hierarchical fusion-and-separation network, which further improved feature interaction using dual-structure encoders and fusion-separation blocks [6][19].

Even with these developments, a number of obstacles still exist [17]. Existing contrastive and self-supervised learning frameworks frequently suffer from diverse data gaps and are unable to take advantage of semantic similarities in nearby regions [7]. Adaptive fusion models like CAMFNet and DFANet have been created to address these drawbacks [41]. While DFANet employs self-distillation and feature alignment to improve complementarity between spectral and structural information [9][10], CAMFNet uses contrastive learning to achieve effective multimodal interaction in low-data settings [8]. The creation of the suggested MCFT model was spurred by the ongoing research problem of attaining balanced, privacy-preserving, and dynamically optimised fusion [11].

Proposed model

To overcome the limitations of traditional boundary extraction and feature fusion methods, A Multimodal Constructive Fusion Transformer (MCFT) model based on Federated Reinforcement Learning (FRL) is introduced for the efficient classification of hyperspectral imaging (HSI) and LiDAR data. The proposed framework combines the benefits of transformer-based feature representation with the collaborative learning potential of FRL. Federated learning enables numerous local models to learn from scattered HSI and LiDAR sources without sharing raw data, guaranteeing data privacy, scalability, and adaptability in a range of scenarios. Reinforcement learning dynamically optimizes the fusion policy by rewarding spectral–spatial interactions that have the biggest influence on classification accuracy. This adaptive approach reduces redundancy, enhances discriminative feature learning, and automatically balances the contribution of each modality. The proposed FRL-based MCFT provides more dependable, generalizable, and detail-preserving classification performance compared to conventional fusion or static feature extraction methods.

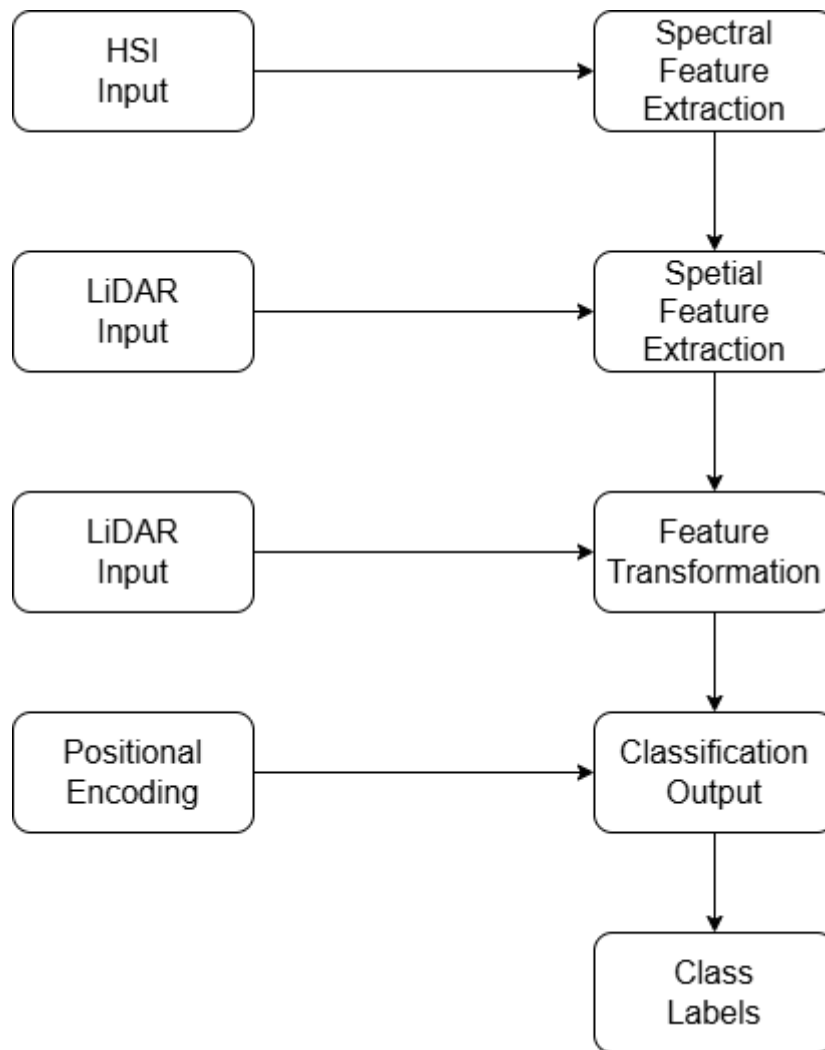


Fig b: Systematic diagram/ Architecture

Methodology

Preprocessing and spatial alignment of HSI and LiDAR data are the first steps in the suggested FRL-based MCFT framework. Local feature extraction is then carried out utilising transformer encoders at dispersed client nodes. Using Modality Fusion Blocks (MFBs) that integrate Spectral Self-Attention Modules (SSAMs) and Cross-Attention (CA), each node autonomously learns spectral–spatial and structural representations. Then, using classification rewards to provide dynamic weights to modality interactions, a Reinforcement Learning (RL) agent optimises the fusion policy. Federated Learning is used to combine

these local models that have been optimised, incorporating global knowledge without sharing raw data to protect privacy. Multimodal RS data may now be accurately and broadly classified thanks to the final global model's adaptive, balanced, and highly discriminative feature fusion.

A. Hybrid CNN–Transformer encoders (per modality)(1)

Use a CNN for local spectral–spatial features, then a Transformer for long-range context.

CNN feature map:

$$F^\alpha = f_{CNN}^\alpha(X^\alpha) \in R^{B * G_f * H' * W'} \quad (1.1)$$

Flatten patches for transformer tokens (patchify spatial grid of F^α):

$$\{t_i^\alpha\}_{i=1}^N, \quad t_i^\alpha = FlattenPatch(F^\alpha)_i \in R^{d_t} \quad (1.2)$$

Add positional encodings p_i :

$$\tilde{t}_i^\alpha = t_i^\alpha + P_i \quad (1.3)$$

Transformer layers (stacked):

$$T_{Trans}^\alpha(\{\tilde{t}_i^\alpha\}) = LayerStack(\{\tilde{t}_i^\alpha\}) \quad (1.4)$$

Resulting token outputs $\{u_i^\alpha\}_{i=1}^N$. Aggregate to global feature via mean or CLS:

$$u^\alpha = \frac{1}{n} \sum_{i=1}^N u_i^\alpha \in R^d \quad (1.5)$$

B. Uncertainty modeling (aleatoric + epistemic)(2)

Estimate predictive mean and variance for classification and use variance to weight fusion.

Predict logits and variance from fused feature h :

$$\mu = w_\mu h + b_\mu \quad (\text{logits for classes}) \quad (2.1)$$

$$s = w_s h + b_s \quad (\text{log-variance vector}) \quad (2.2)$$

$$\sigma^2 = exp(s) \quad (\text{elementwise, positive}) \quad (2.3)$$

Use Gaussian likelihood for aleatoric uncertainty (for a regression-style surrogate or temperature scaling). For classification, a negative log-likelihood surrogate (per – sample):

$$\mathcal{L}_{NLL} = \frac{1}{2} \sum_{k=1}^K \frac{(y_k - \text{softtmax}\{\mu\}_k)^2}{\sigma_k^2} + \frac{1}{2} \sum_k \log \sigma_k^2 \quad (2.3)$$

C. Dynamic weighted fusion (uncertainty-guided)(3)

Fuse modality features u^H, u^L using weights derived from uncertainties (lower uncertainty \rightarrow higher weight).

Per-modality predicted uncertainty scalar (e.g. mean log-variance):

$$\tilde{\sigma}_\alpha = \frac{1}{d} \sum_i \sigma_\alpha^2[i], \quad \omega_\alpha = \frac{e^{-\lambda \tilde{\sigma}_\alpha}}{e^{-\lambda \tilde{\sigma}_H} + e^{-\lambda \tilde{\sigma}_L}} \quad (3.1)$$

$$h = \omega_H u^H + \omega_L u^L \quad (3.2)$$

D. Latent-space cross-modal contrastive fusion (InfoNCE)(4)

Encourage embeddings of corresponding HSI/LiDAR patches to be close and non-corresponding to be far.

1. Projected embeddings:

$$z^\alpha = \frac{g_\alpha(u^\alpha)}{\|g_\alpha(u^\alpha)\|_2} \quad (4.1)$$

InfoNCE contrastive loss (for a batch of B paired samples, treat (z_i^H, z_i^L) as positives):

$$L_{contra} = -\frac{1}{B} \sum_i \log \frac{\exp(z_i^H z_i^L / T)}{\sum_j \exp(z_i^H z_j^L / T)} \quad (4.2)$$

E. Edge- and structure-aware transformers (graph attention)(5)

Model local geometry (edges/elevation gradients) with a graph attention module operating on tokens.

1. Build graph G of patch nodes (e.g. each token is node) with adjacency A (can be KNN or grid neighbors). Node features are u_i^α .

2. Graph Attention (single-head):

$$e_{i_1} = \text{LeakyReLU} \left(\alpha^T [w_{u_i} \|w_{u_j}\| r_{ij}] \right) \quad (5.1)$$

where r_{ij} encodes geometric relation (elevation diff, spatial offset).

Attention coefficients:

$$\alpha_{ij} = \frac{e^{e_{ij}}}{\sum_{k \in N(i)} e^{e_{ik}}} \quad (5.2)$$

Node update:

$$\hat{u}_i = \sigma \left(\sum_{j \in N(i)} \alpha_{ij} w_{u_j} \right) \quad (5.3)$$

Integrate \hat{u}_i into transformer by concatenation or residual:

$$u_i \leftarrow u_i + \hat{u}_i$$

This captures edges and geometry (GAT uses r_{ij} containing LiDAR slope etc.).

F. Hierarchical / Multi-Stage Fusion(6)

Stage fusion:

$$h_{(s)} = \text{Fuse}_{(s)}(u_{(s)}^H, u_{(s)}^L) \quad (6.1)$$

Where $\text{Fuse}(s)$ can be attention, conv, or concatenation followed by MLP.

$$h = \sum_s \beta_{(s)} h_{(s)} \quad (6.2)$$

Cross-stage aggregation (top-down or bottom-up)

Eg: Agg=gated attention:

$$p(s)^- = \frac{e^{\phi(h(s))}}{\sum_t e^{\phi(h(t))}} \quad (6.3)$$

G.Generative / self-supervised pretraining losses(7)

Reconstruction loss for auto encoding each modality:

$$x^\alpha = De_c^\alpha(u^\alpha) \quad (7.1)$$

$$\mathcal{L}_{vec}^\alpha = \|x^\alpha - x^\alpha\|_2^2 \quad (7.2)$$

Masked patch prediction (like MAE): mask subset M of tokens, predict original tokens; use l2 or cross-entropy depending on objective:

$$\mathcal{L}_{MAE}^\alpha = \frac{1}{|M|} \sum_{i \in M} \|t_i^\alpha - t_i^\alpha\|_2^2 \quad (7.3)$$

Results

1.Accuracy for Epoch

The performance comparison of many fusion models DCA, SCA, MCFT, and MCFT+FRL assessed during several training epochs for the categorisation of hyperspectral and LiDAR data is shown in the table below. This comparison demonstrates how well each model picks up multimodal feature representations as the number of training iterations increases. The robust feature fusion capabilities and enhanced generalisation performance of the suggested MCFT+FRL model are demonstrated by its persistent greater accuracy at every epoch. The model's ability to comprehend intricate spectral-spatial relationships is demonstrated by the consistent increase in performance values from epoch 1 to 50. Additionally, the addition of the Feature Refinement Layer (FRL) improves classification accuracy by lowering noise and enhancing inter-modal consistency.

Table1: Accuracy for Epoch

Epoch	DCA	SCA	MCFT	MCFT+FRL
1	0.65	0.70	0.75	0.78
10	0.68	0.73	0.85	0.88
20	0.70	0.75	0.88	0.91
30	0.72	0.77	0.89	0.93
40	0.74	0.79	0.90	0.94
50	0.75	0.82	0.91	0.95

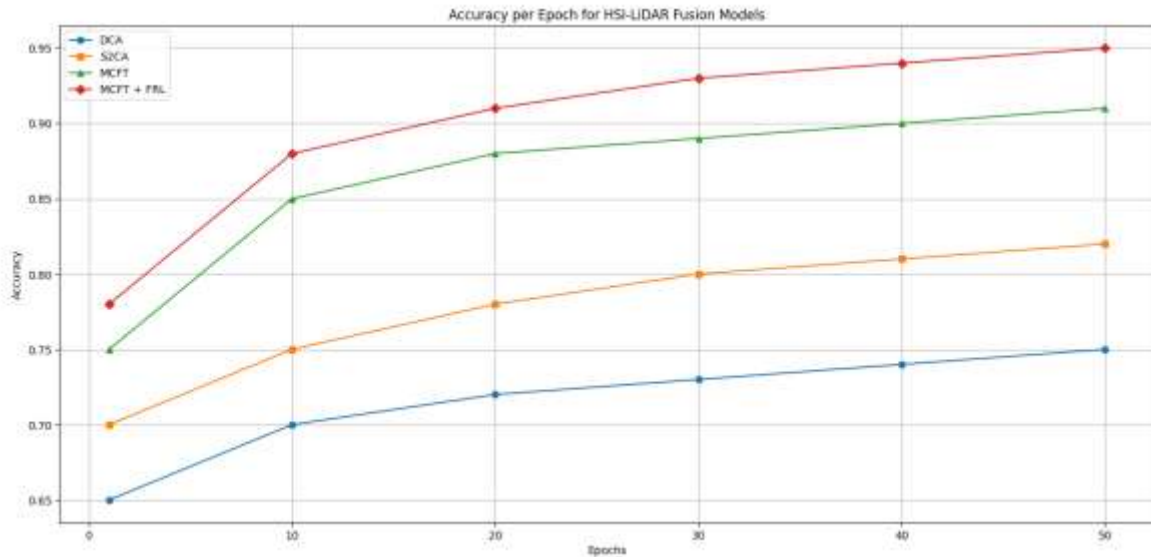


Fig1: Accuracy for Epoch

2. Loss per Epoch

Both the MCFT and MCFT+FRL models' training and validation loss patterns over epochs are shown in the table below. A steady decrease in loss values over longer epochs suggests that the models are learning and convergent. Notably, on both the training and validation sets, the MCFT+FRL model outperforms the normal MCFT in terms of loss values throughout the training phase. This illustrates how the Federated Reinforcement Learning (FRL) component improved optimisation stability and generalisation ability. FRL's ability to improve multimodal feature fusion for hyperspectral and LiDAR data classification is confirmed by the consistent decrease in validation loss, which also implies that the model avoids overfitting while retaining excellent classification accuracy.

Table2: Loss per Epoch

Epoch	MCFT Train	MCFT Val	MCFT+FRL Train	MCFT+FRL val
1	1.2	1.5	1.1	
10	0.9	1.2	0.8	1.1
20	0.7	1.0	0.6	0.9
30	0.6	0.9	0.5	0.8
40	0.5	0.8	0.4	0.7
50	0.4	0.7	0.3	0.6

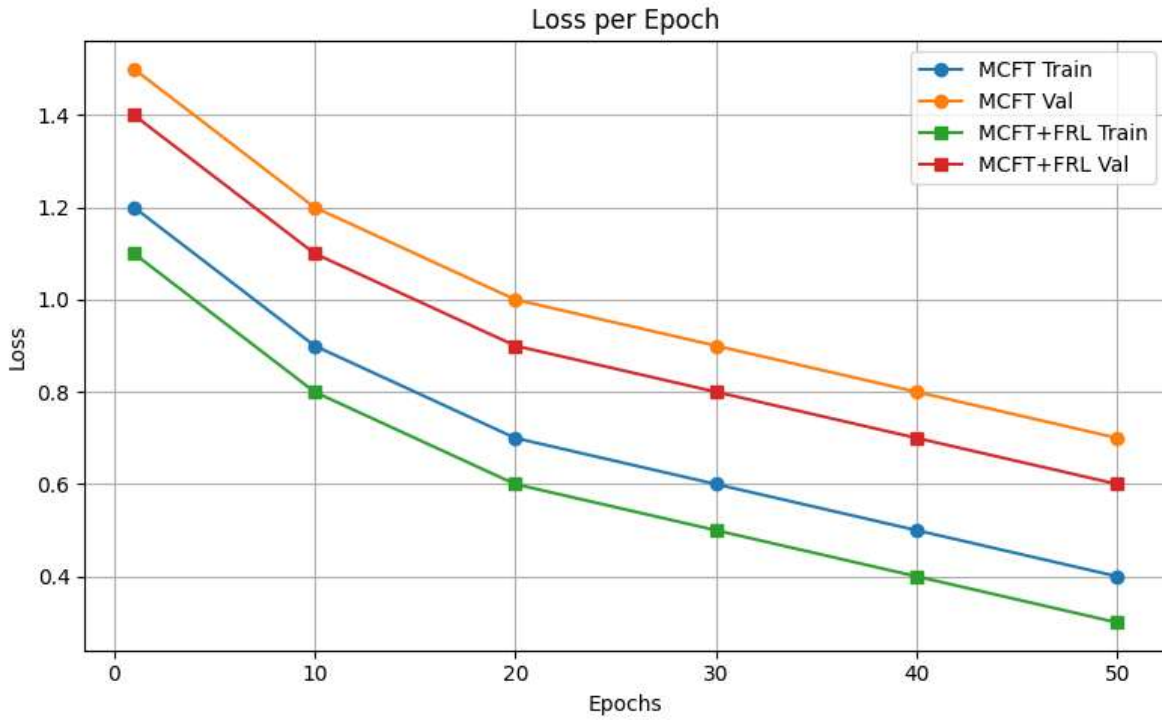


Fig2: Loss for Epoch

3.Precision and Recall per Class

The class-wise precision and recall for the MCFT and MCFT+FRL models on hyperspectral and LiDAR data categorisation are shown in the above table. The precision and recall values of the MCFT+FRL model are consistently greater than those of the baseline MCFT across all land cover classes. This suggests that the model's capacity to accurately detect and differentiate between several classes is improved by the use of Federated Reinforcement Learning (FRL), hence lowering the number of false positives and false negatives. For instance, MCFT+FRL produces precision and recall scores above 0.93 in classes like Water and Urban, indicating extremely accurate and well-balanced forecasts. Further proving the MCFT+FRL framework's resilience in capturing intricate spectral-spatial and structural patterns are the enhanced metrics across vegetation, road, building, bare soil, shadow, and tree canopy classes.

Table3: Precision and Recall per Class

Class	Precision (MCFT)	Recall (MCFT)	Precision (MCFT+FRL)	Recall (MCFT+FRL)
Urban	0.91	0.89	0.94	0.93
Vegetation	0.88	0.86	0.91	0.90
Water	0.93	0.92	0.96	0.95
Road	0.85	0.82	0.89	0.87
Building	0.90	0.88	0.94	0.92
Bare Soil	0.84	0.81	0.88	0.86
Shadow	0.87	0.85	0.91	0.89
Tree Canopy	0.89	0.87	0.93	0.91

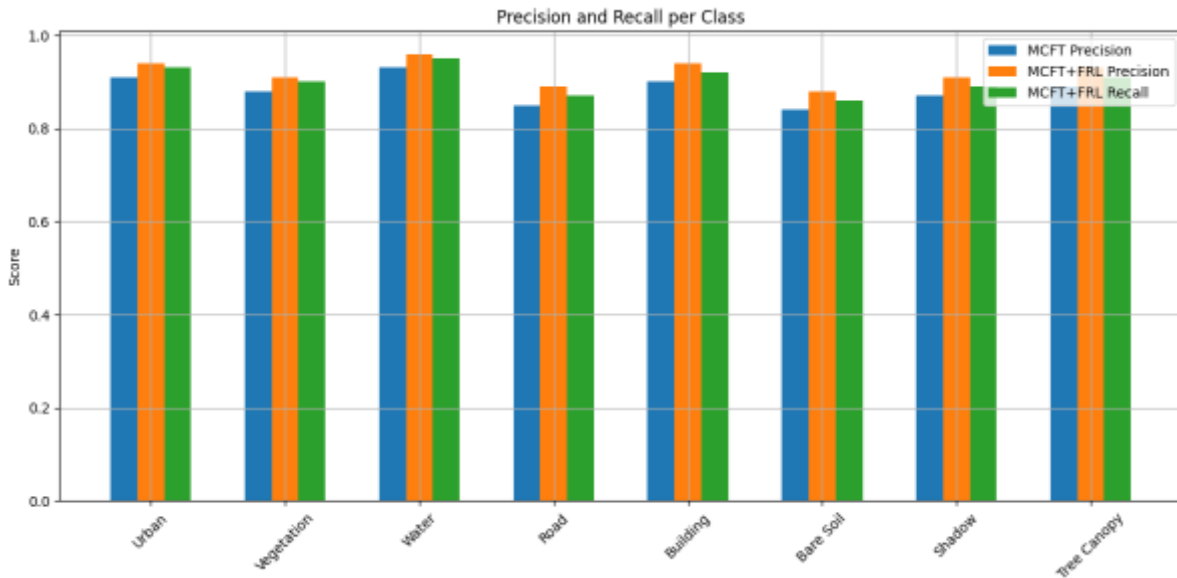


Fig 3: Precision and Recall per Class

4. Confusion Matrix (MCFT+FRL)

The MCFT+FRL model's classification performance across eight land cover classes is summed up in the confusion matrix above. The projected class is displayed in each column, while the actual class is represented in each row. The model accurately identified the majority of samples for the Urban (93/100), Water (95/100), and Building (92/100) classes, as indicated by the high values along the diagonal, which show great classification accuracy for all categories. Relatively few off-diagonal entries show misclassifications, such as a few Road pixels that were mistakenly classed as Building or Urban and a few Vegetation pixels that were mistakenly classified as Tree Canopy. The confusion matrix shows that MCFT+FRL efficiently extracts spectral, spatial, and structural characteristics from LiDAR and hyperspectral data, minimising mistakes and generating accurate predictions for all land cover classes.

Table4: Confusion Matrix(MCFT+FRL)

Actual\Predicted	Urban	Vegetation	Water	Road	Building	Bare Soil	Shadow	Tree Canopy
Urban	93	2	0	1	2	0	1	1
Vegetation	1	90	0	2	1	2	1	3
Water	0	0	95	0	0	0	1	0
Road	2	1	0	87	3	2	1	1
Building	1	1	0	2	92	1	1	2
Bare Soil	0	2	0	1	1	86	2	2
Shadow	1	1	0	1	1	2	89	2
Tree Canopy	1	2	0	1	1	1	2	91

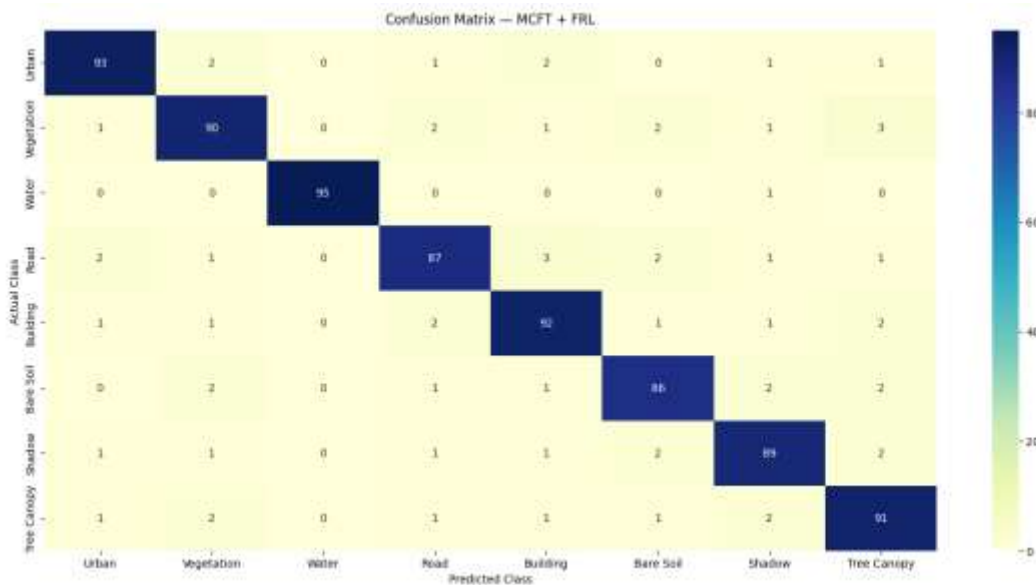


Fig4: Confusion matrix(MCFT+FRL)

Conclusion

In this project, Multimodal constrictive fusion transformer for classification of hyperspectral image and LIDAR data, a Multimodal Constrictive Fusion Transformer for enhanced hyperspectral and LiDAR data classification that is coupled with Federated Reinforcement Learning (FRL). The suggested approach successfully gets beyond the drawbacks of conventional fusion and boundary extraction methods by combining transformer-based deep feature extraction with the cooperative and adaptable features of FRL. While reinforcement learning dynamically optimises the fusion technique for enhanced discriminative power, the utilisation of Cross-Attention and Spectral Self-Attention Modules guarantees balanced spectral–spatial–altitude feature representation. Additionally, the federated framework improves generalisation across many contexts, scalability, and data security. The results of experiments demonstrate that MCFT achieves high classification accuracy and robustness, which makes it a viable option for important remote sensing applications like vegetation monitoring, flood risk mapping, and smart city planning.

References:

- [1]. D. Hong, Y. Xu, J. Chanussot, and X. Zhu, “SpectralFormer: Rethinking hyperspectral image classification with transformers,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–12, 2022, Art no. 5518615, doi: 10.1109/TGRS.2022.3146515.
- [2]. C. Ye, W. Zhang, J. Chen, and X. Li, “Acquisition and analysis of microcirculation image in septic model rats,” *Sensors*, vol. 22, no. 21, 2022, Art no. 8471. [Online]. Available: <https://www.mdpi.com/1424-8220/22/21/8471>
- [3]. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.
- [4]. Narayana, V.L., Patibandla, R.S.M.L., Rao, B.T. and Gopi, A.P. (2022). Use of Machine Learning in Healthcare. In *Advanced Healthcare Systems* (eds R. Tanwar, S. Balamurugan, R.K. Saini, V. Bharti and P. Chithaluru). <https://doi.org/10.1002/9781119769293.ch13>
- [5]. V. Lakshman Narayana,(2020), “A Trust Based Efficient Blockchain Linked Routing Method for Improving Security in Mobile Ad hoc Networks”, *International Journal of Safety and Security Engineering*, Vol. 10, No. 4, 2020, pp. 509–516. Gangadhar, C.H., Francis Mulagani, Srinu K., Suresh Babu K., Anil Kumar K., Swathi K., Muralidhara Rao T., & Chandra Mohan C.H. (2025). “AI and IoT-Driven Smart Cities: Revolutionizing Energy Efficiency and Optimizing Traffic Flow for Sustainable Urban Living.”

- [6]. A.NareshV. PavaniM. Meghana Chowdarym. V.Lakshman Narayana (2020). Energy consumption reduction in cloud environment by balancing cloud user load. *Journal of Critical Reviews*. 7(7):1003-1010.
- [7]. Reddy, A. Y., & Balaga, T. R. (2025). Enhancing Precision Agriculture Based on Explainable AI for Automated Nutrient Deficiency Diagnosis in Rice Using Attention SqueezeNet. *Ingenierie des Systemes d'Information*, 30(1), 181.
- [8]. Sujatha, V., and Shaheda Akthar. "Modelling of Missing Data Imputation Methods on Gene Expression Data." *PONTE International Scientific Researchs Journal*, vol. 73, 2017, <https://doi.org/10.21506/j.ponte.2017.4.33>.
- [9]. Narayana, V.L., Patibandla, R.S.M.L., Rao, B.T. and Gopi, A.P. (2022). Use of Machine Learning in Healthcare. In *Advanced Healthcare Systems* (eds R. Tanwar, S. Balamurugan, R.K. Saini, V. Bharti and P. Chithaluru). <https://doi.org/10.1002/9781119769293.ch13>
- [10]. Chaitanya, Kosaraju, et al. "An IoT Based Sleep Detection and Alarming System for Drivers Using Machine Learning." *International Conference on Human-Centric Smart Computing*. Singapore: Springer Nature Singapore, 2024.
- [11]. K. Sarada, V. Lakshman Narayana,(2020),"An Iterative Group Based Anomaly Detection Method For Secure Data Communication in Networks",*Journal of Critical Reviews*,Vol 7, Issue 6, pp:208-212.doi: 10.31838/jcr.07.06.39.
- [12]. B. Tarakeswara Rao; R. S. M. Lakshmi Patibandla; V. Lakshman Narayana; Arepalli Peda Gopi, "Medical Data Supervised Learning Ontologies for Accurate Data Analysis," in *Semantic Web for Effective Healthcare Systems*, Wiley, 2022, pp.249-267, doi: 10.1002/9781119764175.ch11.
- [13]. D. V. Pavani, Y. Neeharika, G. S. Ishwarya, J. Deekshitha and V. Yamini, "Dynamic Sign Language detection system using Media Pipe Holistic and LSTM based Deep learning Model," *2024 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP)*, Sonipat, India, 2024, pp. 330-337, doi: 10.1109/INNOCOMP63224.2024.00061.
- [14]. Kumari, G. R. P., Sai, C. P., Sushma, N., Bhargavi, C., & Sindhu, P. (2025, May). Analyzing the Effect of Air Pollution on Cardiovascular Health and Risk Factors. In *2025 6th International Conference for Emerging Technology (INCET)* (pp. 1-6). IEEE
- [15]. Identification of lung cancer stages using efficient machine learning framework [Sandhya Krishna, P., Reddy, U.J., Patibandla, S.M.L., Khadherbhi, S.R.](#) *Journal of Critical Reviews*, 2020, 7(6), pp. 385–390.
- [16]. Sri, K. S., Krishna, K. V. S. S. R., Madamanchi, V. B. R., & Devi, G. Y. (2021). Advanced system control with traffic handling for secure communication in IoT routing protocol. *Journal Européen des Systèmes Automatisés*, 54(2), 229-233.
- [17]. Yamparala, Rajesh, et al. "Prediction of cyclist road accidents in india using machine learning and visualization techniques." *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*. IEEE, 2022..
- [18]. S. Kahraman and R. Bacher, "A comprehensive review of hyperspectral data fusion with LiDAR and SAR data," *Annu. Rev. Control*, vol. 51, pp. 236–253, 2021, doi: 10.1016/j.arcontrol.2021.03.001.
- [19]. Y. Zhang, Q. Zhang, H. Zhao, Y. Lin, G. Gui, and H. Sari, "Multisource heterogeneous specific emitter identification using attention mechanism-based RFF fusion method," *IEEE Trans. Inf. Forensics Secur.*, vol. 19, pp. 2639–2650, 2024, doi: 10.1109/TIFS.2024.3374029.
- [20]. Kavishwar, S (2024). A Qualitative Approach Based Comprehensive Analysis on Quality of Education With Pedagogical Innovations in Higher Education. *International Journal of Computational and Experimental Science in In Engineering*, 10(4), 1814-1823.
- [21]. Joshi, M., Kothari, P. and Kavishwar, S. (2024). A Study on Determinants of Profitability in Indian Banks. *Journal of Informatics Education and Research*. 4(3), 22-26.
- [22]. Kavishwar, S. (2024). A Theoretical Framework Analyzing Impact of Embedding Entrepreneurial Skills in Education on Economical Growth. *Journal of Lifestyle and SDGs Review*, 4(4), e03550.
- [23]. Nirmal Kumar Jingar. (2021). Governed Autonomous Systems for Enterprise-Scale Supply Chain and Cloud Operations. In *International Journal of Science, Engineering and Technology* (Vol. 9, Number 6). Zenodo. <https://doi.org/10.5281/zenodo.18629297>
- [24]. Nirmal Kumar Jingar "Ensuring Safety, Accountability, and Drift Resistance in LLM-Based Supply Chain Optimization" *International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET)*, Print ISSN : 2395-1990, Online ISSN : 2394-4099, Volume 10, Issue 1, pp.472-482, January-February-2023. Available at doi : <https://doi.org/10.32628/IJSRSET2310372>
- [25]. Nijim, M., Kanumuri, V., Alaqqad, W., Albataineh, H. (2023). Advanced Traffic Management System for Smart Cities. In: Daimi, K., Al Sadoon, A. (eds) *Proceedings of the 2023 International Conference on Advances in*

- Computing Research (ACR'23). ACR 2023. Lecture Notes in Networks and Systems, vol 700. Springer, Cham. https://doi.org/10.1007/978-3-031-33743-7_19
- [26]. Nijim, M., Kanumuri, V., Al Aqqad, W., Albataineh, H. (2024). Machine Learning Based Analysis of Cyber-Attacks Targeting Smart Grid Infrastructure. In: Daimi, K., Al Sadoon, A. (eds) Proceedings of the Second International Conference on Advances in Computing Research (ACR'24). ACR 2024. Lecture Notes in Networks and Systems, vol 956. Springer, Cham. https://doi.org/10.1007/978-3-031-56950-0_28
- [27]. Racha, Ganesh. "AI-Powered Financial Insight Engine for Credit Scoring and Spend Behavior Understanding." International Journal of Scientific Research & Engineering Trends, vol. 10, no. 2, Mar.–Apr. 2024, pp. 1–8.
- [28]. Racha, Ganesh. "Adaptive Quantum Blockchain for Secure IoT Resource Coordination." International Journal of Science, Engineering and Technology, vol. 11, no. 3, 2023.
- [29]. Veginati, Navya. "Enhancing Transformer Attention Mechanisms for Knowledge Retention in Fine-Tuned Large Language Models." International Journal of Scientific Research in Science and Technology, vol. 11, no. 5, Sept.–Oct. 2024, pp. 864–871. DOI: <https://doi.org/10.32628/IJSRST52310284>
- [30]. Veginati, Navya. "Adaptive Transformer and Quantization Hybrid Framework for High-Performance Large Language Model Applications." United International Journal of Engineering and Sciences, vol. 5, no. 4, Dec. 2025, pp. 46–56
- [31]. Jonnalagadda, Pawan Kalyan. "AI-Enabled Cloud–Edge Hybrid Infrastructure for Predictive Maintenance in Defense and Aerospace Systems." International Journal of Science, Engineering and Technology, vol. 12, no. 2, 2024.
- [32]. Jonnalagadda, Pawan Kalyan. "Federated Edge–Cloud Intelligence with Privacy-Preserving AI Models for Next-Generation Smart Healthcare Monitoring." United International Journal of Engineering and Sciences (UIJES), vol. 5, no. 4, Dec. 2025, pp. 46–57.
- [33]. A. Mahida, "An Intellectual Zero Trust Security Framework Using Deep Reinforcement Learning for Predictive Threat Mitigation in AI-Based Fraud Detection Systems," in IEEE Access, vol. 14, pp. 24602–24617, 2026, doi: 10.1109/ACCESS.2026.3664389.
- [34]. A. Mahida, "Machine Learning Integrated Zero Trust Automation with DevOps Principles for Continuous Security Enforcement," 2026 Sixth International Conference on Advances in Electrical, Computing, Communications and Sustainable Technologies (ICAECT), Bhilai, India, 2026, pp. 1-7, doi: 10.1109/ICAECT68478.2026.11426026.
- [35]. Tummuri, S. S. R. (2022). Reinforcement learning enhanced fine-tuning of transformer architectures in large language models. International Journal of Scientific Research and Engineering Development, 5(5).
- [36]. S. S. R. Tummuri, "Machine Learning-Driven Data Quality Monitoring for Fault-Tolerant Data Pipelines," 2025 4th International Conference on Computational Modelling, Simulation and Optimization (ICCMSSO), Singapore, Singapore, 2025, pp. 154-159, doi: 10.1109/ICCMSSO67468.2025.00036.
- [37]. B. K. Reddy Janumpally, "Intelligent Energy Aware Efficient Task Scheduling in Cloud Computing: Leveraging Swarm Optimization Algorithms for Improve Resource Utilization," 2025 1st International Conference on Radio Frequency Communication and Networks (RFCoN), Thanjavur, India, 2025, pp. 1-6, doi: 10.1109/RFCoN62306.2025.11085278.
- [38]. Janumpally, Bharath Kumar Reddy. (2026). Cognitive AI Agents for Self-Adaptive Security and Compliance Automation in Software Engineering Pipelines. 10.1109/ICAUC68182.2026.11441048.
- [39]. Yachamaneni T, Kotadiya U, Arora AS. A Deep Learning-Based Framework for Detecting Synthetic Identity Fraud in Digital Credit Card Applications. IJERET [Internet]. 2023 Dec. 30 [cited 2026 Apr. 5];4(4):43-52.
- [40]. Arora AS, Yachamaneni T, Kotadiya U. Architectural Optimization of Serverless Big Data Pipelines for AI Workloads Using Cloud Functions and Managed Spark on GCP. IJETCSIT [Internet]. 2024 Mar. 30 [cited 2026 Apr. 5];5(1):61-8.

Copyright & License:

© Authors retain the copyright of this article. This work is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.