

# A Deep Multimodal Contrastive Transformer Framework for Hyperspectral Image and LiDAR Classification

Dr. V. Lakshman Narayana<sup>1</sup>, Ch. Bindhu<sup>2</sup>, U. Aadhilakshmi<sup>3</sup>, S. Vasantha<sup>4</sup>, P. Venkata Lavanya<sup>5</sup>

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 Contrastive 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).

## 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]. Although hyperspectral images (HSIs) provide rich spectral information for classifying land cover, they have trouble differentiating objects that have different heights but similar spectral signatures [2]. By resolving structural differences, LiDAR data which records 3D altitude information complements HIS [3]. 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 [4] [5].

Conventional fusion techniques, such as hierarchical fusion or weighted concatenation, frequently have suboptimal classification accuracy due to feature imbalance and the Hughes phenomenon [6]. Promising answers are provided by recent developments in deep learning, especially in the area of attention

mechanisms [7]. 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 [8] [9].

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 [10] [11]. In order to improve fusion efficacy, our model uses LiDAR-generated attention score maps to enrich HSI features with altitude data [12]. 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 [13] [14].

For applications like flood risk mapping, wildfire prevention, and smart city development that demand in-depth structural and material insights [15], this strong fusion approach is essential. LiDAR integration is crucial for tasks like infrastructure inspection because HSIs, despite their spectrally rich nature, frequently lack spatial resolution [16]. Improves land cover classification accuracy and practical utility by guaranteeing a more expressive and trustworthy representation of multimodal RS data [17] [18].

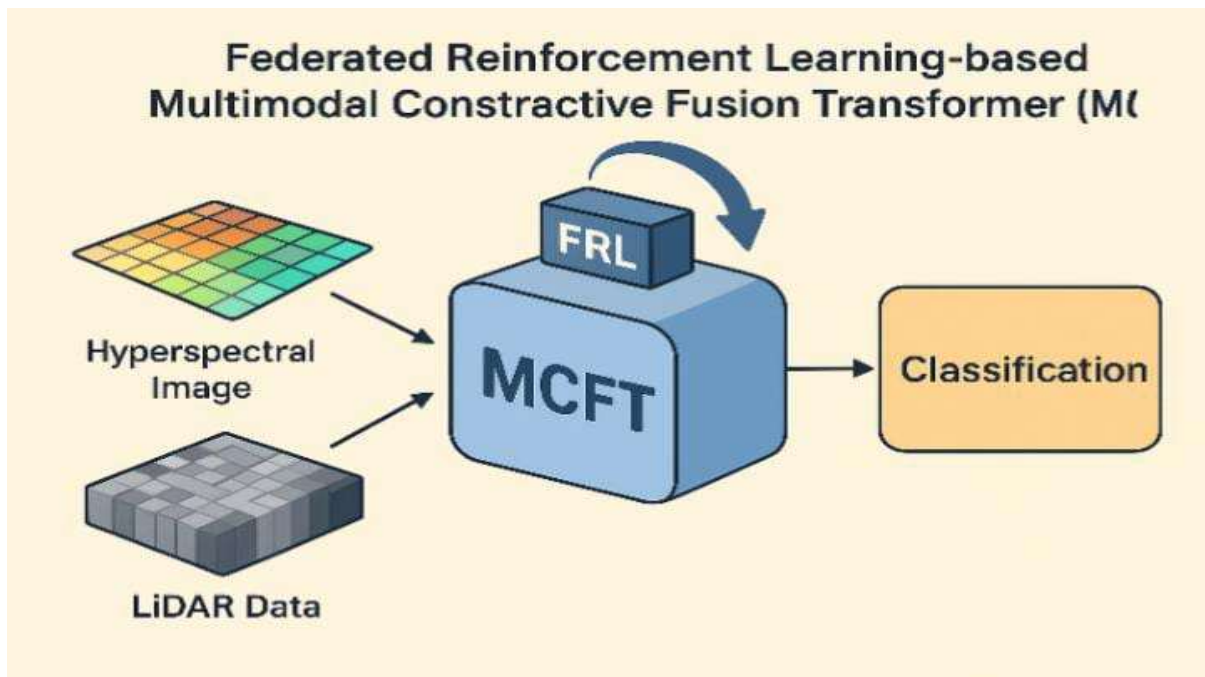


Fig a: Federated Reinforcement Learning-based Multimodal constrictive fusion Transformer

This model improves land cover classification by fusing hyperspectral and LiDAR data using Discriminant Correlation Analysis (DCA) [19]. Group-wise spatial feature extraction is made possible by the grouping of HSI bands according to joint entropy and structural similarity [20]. DCA reduces dimensionality while maintaining discriminative power by fusing LiDAR altitude features with HSI spectral and spatial features in parallel [21]. By aligning heterogeneous features from both modalities, this method improves classification accuracy [22].

The fusion of hyperspectral and LiDAR data is guided by attention mechanisms in this deep learning model. It highlights significant areas in LiDAR data with spatial attention and highlights informative bands in HSI with spectral attention [23]. The network can learn complementary features across modalities thanks to the multi-level fusion. By adaptively weighting features according to their relevance, this model enhances classification performance [24] [25].

CNNs, RNNs, and transformer-based models are among the deep learning architectures used for HSI-LiDAR fusion that are covered in this systematic review [26]. It draws attention to issues such as spectral-spatial heterogeneity, feature imbalance, and the requirement for explainability [27]. Emerging trends for multimodal RS classification are also covered in the review, including self-supervised learning, multi-head transformers, and cross-attention [28].

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 [29]. Because of fixed or empirically adjusted  $\alpha$  values, these methods frequently have trouble generalizing across a variety of urban environments, resulting in overly smoothed or fragmented contours [30]. 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 [31].

By using deep learning to automatically learn hierarchical spectral-spatial correlations, the proposed Multimodal Contrastive Fusion Transformer (MCFT) overcomes these drawbacks [32]. In order to manage non-linear cross-modal interactions and maintain fine-grained boundary details, it combines HSI and LiDAR data [33]. 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 [34]. This makes it particularly useful in intricate situations requiring a joint analysis of structural geometry and material composition [35].

## Literature survey

Accurate land cover classification has made the integration of diverse data sources, such as LiDAR and hyperspectral imaging (HSI), more crucial [36]. The capacity of current HSI-LiDAR fusion techniques to capture comprehensive structural-spectral correlations is limited because they frequently split spectral and spatial data [37]. To solve these problems, recent research has suggested multiscale and transformer-based attention techniques [38]. The S2CA framework used multiscale convolution and feature recalibration to improve intermodal interaction and reduce noise [39], while a modality fusion vision transformer that combined cross-attention and spectral self-attention improved joint feature learning to achieve 99.91% accuracy [40]. High-dimensional multimodal representations have been successfully extracted from HSI and LiDAR data by adapting pretrained vision-language models (VLMs) for joint classification [3][17]. 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 [4][18]. 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]. 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 [41].

Even with these developments, a number of obstacles still exist [19]. 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. While DFANet employs self-distillation and feature alignment to improve complementarity between spectral and structural information [9], CAMFNet uses contrastive learning to achieve effective multimodal interaction in low-data settings [8][20]. The creation of the suggested MCFT model was spurred by the ongoing research problem of attaining balanced, privacy-preserving [14], and dynamically optimised fusion [15].

## Proposed model

To overcome the limitations of traditional boundary extraction and feature fusion methods, A Multimodal Constrictive 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.

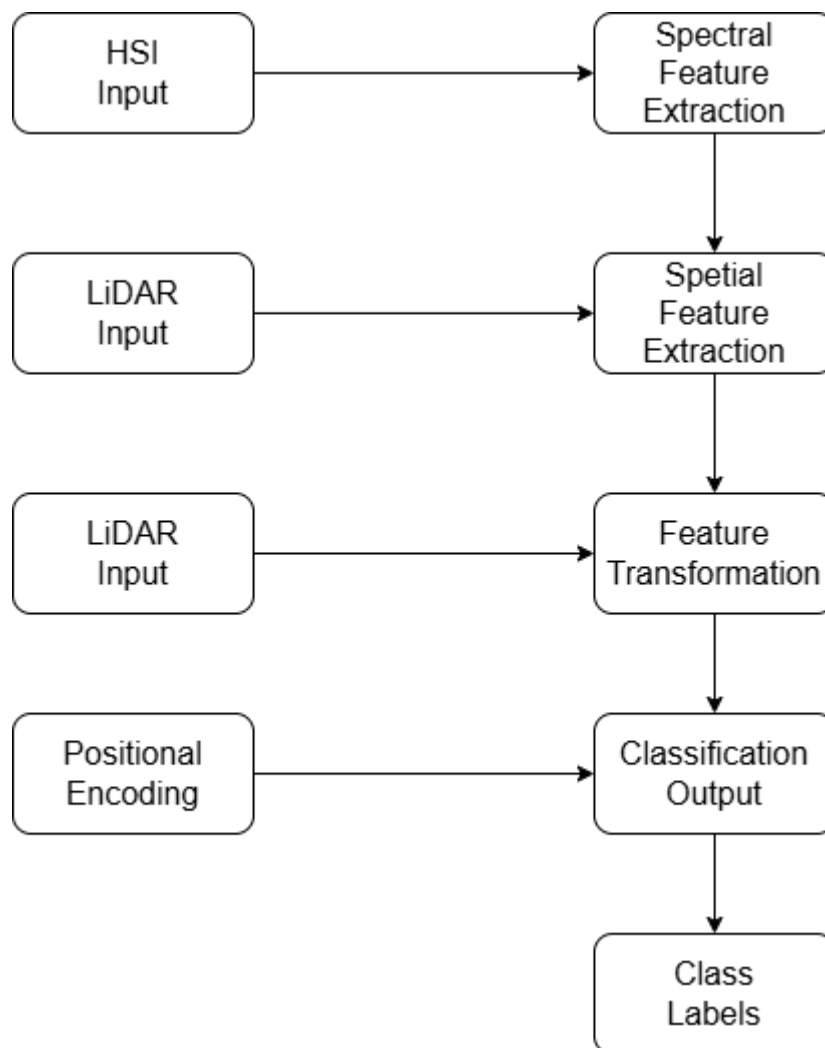


Fig b: Systematic diagram/ Architecture

## 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  $F^\alpha$ ):

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

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):

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

### 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  $u_i^\alpha$ .

2. Graph Attention (single-head):

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

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 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

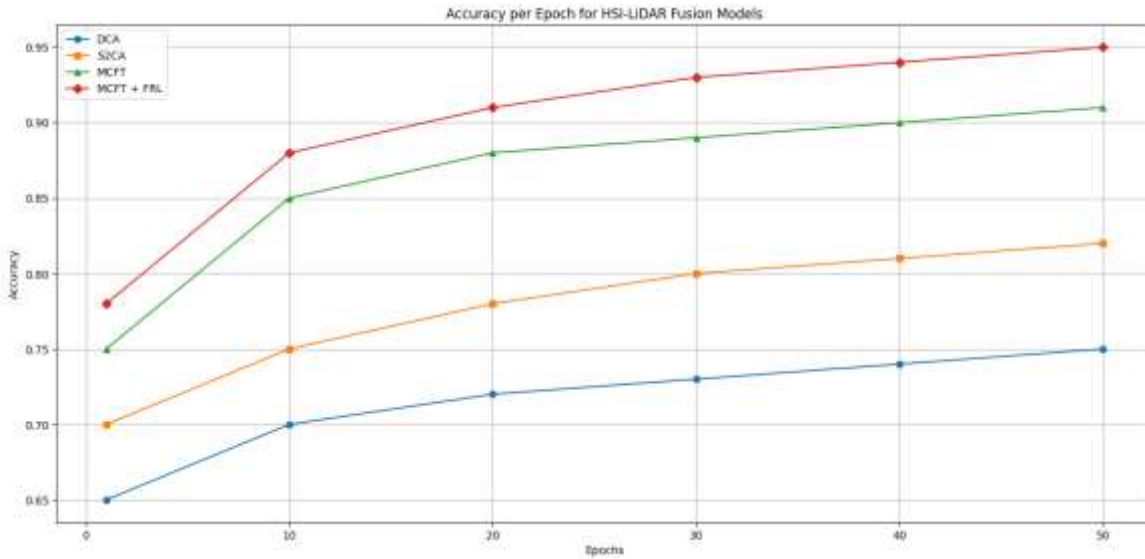


Fig1: Accuracy for Epoch

## 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

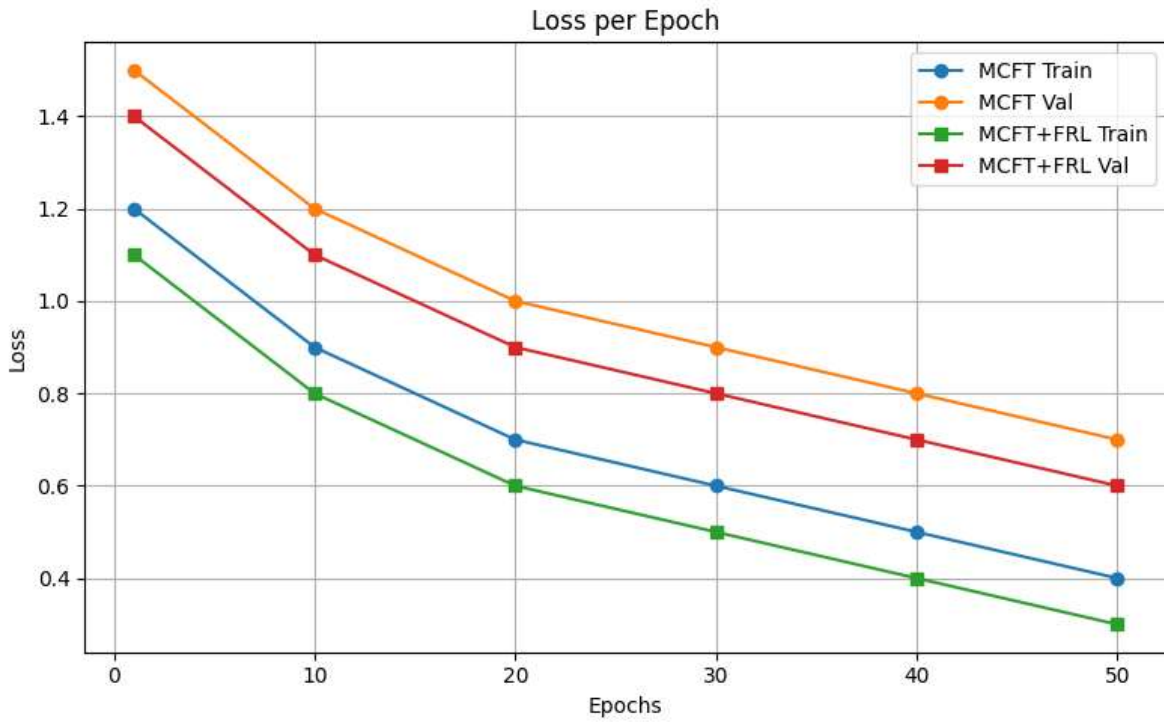


Fig2: Loss for Epoch

### 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

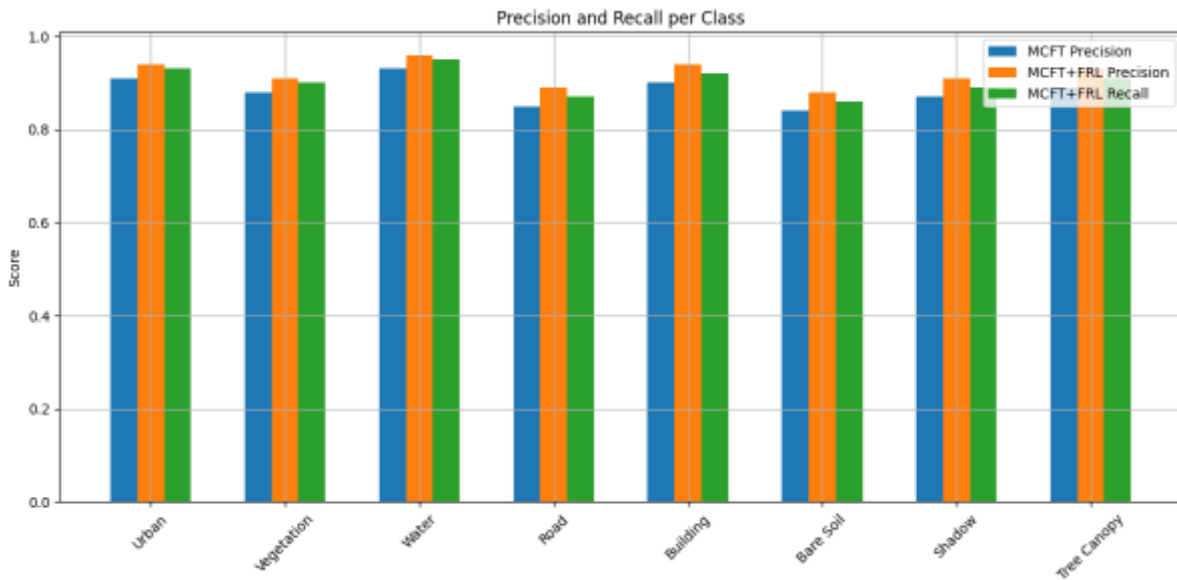


Fig 3: Precision and Recall per Class

#### 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

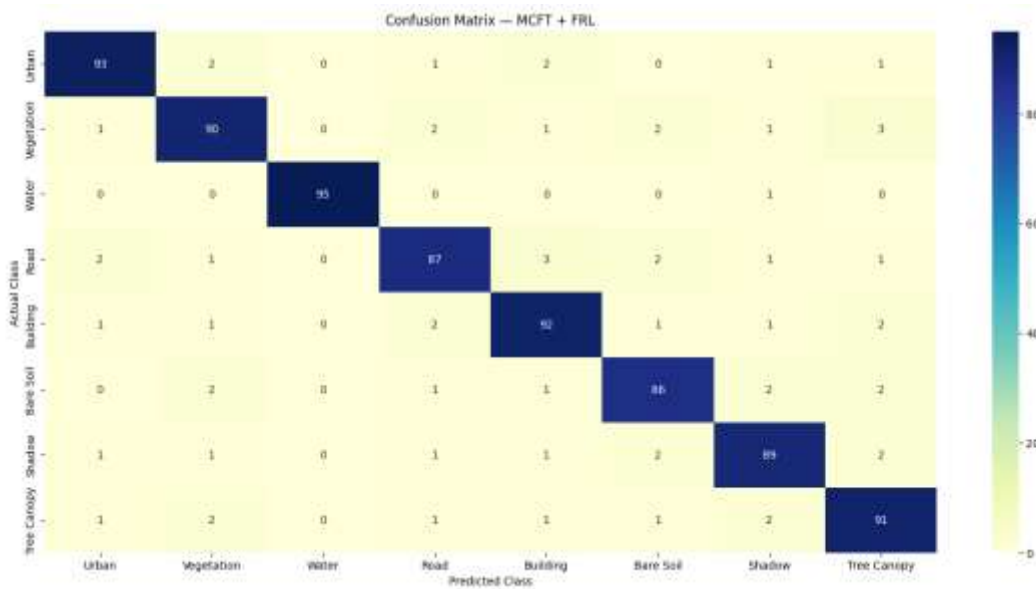


Fig4: Confusion matrix(MCFT+FRL)

### 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]. L. N. Vejjendla, B. Bysani, A. Mundru, M. Setty and V. J. Kunta, "Score based Support Vector Machine for Spam Mail Detection," 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2023, pp. 915-920, doi: 10.1109/ICOEI56765.2023.10125718
- [5]. V. Pavani, S. Sri. K, S. Krishna. P and V. L. Narayana, "Multi-Level Authentication Scheme for Improving Privacy and Security of Data in Decentralized Cloud Server," 2021 2nd International MANETS Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2021, pp. 391-394, doi: 10.1109/ICOSEC51865.2021.9591698.

- [6]. Lakshman Narayana Vejendla and Bharathi C R, (2018), "Effective multi-mode routing mechanism with master-slave technique and reduction of packet droppings using 2-ACK scheme in", *Modelling, Measurement and Control A*, Vol.91, Issue.2, pp.73-76.
- [7]. Narayana, Vejendla Lakshman, Arepalli Peda Gopi, and Kosaraju Chaitanya. "Avoiding Interoperability and Delay in Healthcare Monitoring System Using Block Chain Technology." *Rev. d'Intelligence Artif.* 33.1 (2019): 45-48.
- [8]. Sirisha, A., Chaitanya, K., Krishna, K. V. S. S. R., & Kanumalli, S. S. (2021). Intrusion detection models using supervised and unsupervised algorithms-a comparative estimation. *International Journal of Safety and Security Engineering*, 11(1), 51-58.
- [9]. Suajtha, V. "Variable Selection in Functional Genomics Using Genetic Algorithm-Based Feature Selection Method-An Empirical Study." *Journal of Engineering and Applied Sciences*, 21 Sept. 2022. ISSN Online 1818-7803, ISSN Print 1816-949x.
- [10]. Majety, Vasumathi Devi, V. Sujatha, V. S. Sai Rama Krishna Komanduri, and Satya Sandeep Kanumalli. "Enhanced Secure Communication AODV Routing Protocol Using SVM in MANETS." *AIP Conference Proceedings*, vol. 2724, no. 1, AIP Publishing, 2023. <https://doi.org/10.1063/5.0130170>.
- [11]. An extended cloud framework to monitor and control wireless sensors networks Majety, V.D., Sravanthi, G.L., Didla, D. *International Journal of Innovative Technology and Exploring Engineering*, 2019, 8(11), pp. 3805–3808
- [12]. Kosaraju, Chaitanya, et al. "Mirchi crop yield prediction based on soil and environmental characteristics using modified RNN." *2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*. IEEE, 2023.
- [13]. Patibandla, R.S.M.L., Narayana, V.L., Gopi, A.P. (2021). Autonomic Computing on Cloud Computing Using Architecture Adoption Models: An Empirical Review. In: Choudhury, T., Dewangan, B.K., Tomar, R., Singh, B.K., Toe, T.T., Nhu, N.G. (eds) *Autonomic Computing in Cloud Resource Management in Industry 4.0*. EAI/Springer Innovations in Communication and Computing. Springer, Cham. [https://doi.org/10.1007/978-3-030-71756-8\\_11](https://doi.org/10.1007/978-3-030-71756-8_11)
- [14]. V. Pavani, N. VijayaLakshmi, N. Harika, G. S. Sowjanya and V. Deepthi, "Deep Learning-based Analysis of Brain MRI for Enhanced Diagnosis of Multiple Sclerosis," *2024 5th International Conference on Data Intelligence and Cognitive Informatics (ICDICI)*, Tirunelveli, India, 2024, pp. 1141-1148, doi: 10.1109/ICDICI62993.2024.10810928.
- [15]. Kumari, G. R. P., Reddy, G. A., Nazarana, S., Vanaja, K., Snehitha, V., & Alapati, N. (2025, January). Deep Learning-Based Lung Tumor Analysis for Enhanced Oncology Diagnostics. In *2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI)* (pp. 1401-1408). IEEE.
- [16]. P. S. Krishna and S. R. Peram, "A Brief Survey on Image Denoising based Feature Extraction and Classification Models for Oral Cancer Detection," *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, Erode, India, 2023, pp. 702-708, doi: 10.1109/ICSCDS56580.2023.10104790.
- [17]. Sirisha, A., Chaitanya, K., Krishna, K. V. S. S. R., & Kanumalli, S. S. (2021). Intrusion detection models using supervised and unsupervised algorithms-a comparative estimation. *International Journal of Safety and Security Engineering*, 11(1), 51-58.
- [18]. Vignan's Nirula, I. T. W. "Data outsourcing based on secure association rule mining processes." *International Journal of Security and Its Applications* 9.3 (2015): 41-48.
- [19]. 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.
- [20]. 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.

- [21]. Narlawar, N., Kavishwar, S. (2019). Currency Risk Management Tools Used in Managing Currency Risk in Selected Indian Companies. *Indian Journal of Research and Analytical Reviews*. 6(2), 609-614.
- [22]. Ghangare, A. S., & Kavishwar, S. The Increasing Significance of Green Corporate Finance in India. *Journal of Management & Entrepreneurship*, 277-286.
- [23]. Kavishwar, S., & Shahu, A. (2011). Reporting Intangible Assets-Convergence of Accounting Standard. *Journal of Accounting and Finance*. 26(1), 73-79.
- [24]. Arora AS, Yachamaneni T, Kotadiya U. Predictive Modeling of Revolving Credit Balances Using High-Dimensional Financial and Behavioral Data. *IJAIBDCMS* [Internet]. 2023 Mar. 30 [cited 2026 Apr. 5];4(1):98-107.
- [25]. Kotadiya U, Arora AS, Yachamaneni T. Intelligent Orchestration of Cloud-Native Applications Using Google Cloud Platform and Microservices-Based Architectures. *IJAIBDCMS* [Internet]. 2024 Dec. 30 [cited 2026 Apr. 5];5(4):106-14.
- [26]. Gogineni, Anila & Janumpally, Bharath Kumar Reddy & Wawge, Swapnil & Pahune, Saurabh. (2025). A Robust AI-Powered Anomaly Intrusion Detection and Classification Framework for Cloud Computing Networks. 1-6. 10.1109/INDISCON66021.2025.11253743.
- [27]. A. Joon, B. K. R. Janumpally, A. Gogineni and P. Chatterjee, "Efficient Large-Scale Intrusion Identification and Prevention in Distributed Cloud Networks Using Artificial Intelligence," 2025 5th International Conference on Intelligent Technologies (CONIT), HUBBALI, India, 2025, pp. 1-8, doi: 10.1109/CONIT65521.2025.11167760.
- [28]. S. S. R. Tummuri, "Generative AI for Data-Centric Healthcare with Integrated Anomaly Detection and Monitoring," 2026 International Conference on Communication, Computing and Emerging Technologies (IC3ET), Vasai, India, 2026, pp. 520-526, doi: 10.1109/IC3ET64989.2026.11467187.
- [29]. Tummuri, S. S. R. (2024). Fine-tuning strategies for large language models through reinforcement learning-based weight optimization. *International Journal of Science, Engineering and Technology*. Volume 4, Issue 3.
- [30]. Ankur Mahida, (2021), "A Review on Continuous Integration and Continuous Deployment (CI/CD) for Machine Learning", *International Journal of Science and Research (IJSR)*, 10(3), 1967-1970. <https://dx.doi.org/10.21275/SR24314131827>, <https://www.ijsr.net/getabstract.php?paperid=SR24314131827>
- [31]. "Mahida, A. (2022). Comprehensive Review on Optimizing Resource Allocation in Cloud Computing for Cost Efficiency. *Journal of Artificial Intelligence & Cloud Computing*. SRC/JAICC-249. DOI: [doi.org/10.47363/JAICC/2022\(1\),232,2-4](https://doi.org/10.47363/JAICC/2022(1),232,2-4)."
- [32]. Jonnalagadda, P.K. (2026). Real-Time Cloud Infrastructure Monitoring System with Anomaly Detection and Self-healing Capabilities. In: Kumar, V.N., Senkerik, R., Prasad, V.K., Kumar, T.K. (eds) *Intelligent Computing and Communication*. ICICC 2025. Lecture Notes in Networks and Systems, vol 1839. Springer, Cham. [https://doi.org/10.1007/978-3-032-18349-1\\_43](https://doi.org/10.1007/978-3-032-18349-1_43)
- [33]. 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.
- [34]. Veginati, Navya. "Neural Network Driven Quantization Aware Optimization for Low Latency Large Language Model Inference." *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 10, no. 3, May-June 2024, pp. 1162–1170, doi:10.32628/CSEIT25113584.
- [35]. 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>
- [36]. Racha, Ganesh. "Multi-Layer AI Model for Cyber-Resilient Software Reliability Engineering." *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 11, no. 5, Sept.–Oct. 2025, pp. 507–519. <https://doi.org/10.32628/CSEIT26121364>
- [37]. Racha, Ganesh. "Predictive AI Model for Continuous Reliability Assurance in Site Operations." *International Journal of Scientific Research in Science and Technology*, vol. 12, no. 2, Mar.-Apr. 2025, pp. 1469-78, <https://doi.org/10.32628/IJSRST2613340>.

- [38]. R. Eswarawaka, S. K. Kudikala, S. C. Kuchi and V. Verma K., "The analysis on search engine optimization supported by six sigma methodology," 2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bengaluru, India, 2017, pp. 653-658, doi: 10.1109/ICIMIA.2017.7975544.
- [39]. Albatineh, H., Kanmuri, V., Alaqad, W., Nijim, M. (2024). Utilizing Machine Learning for Intrusion Detection in Smart Grid Systems. In: Daimi, K., Al Sadoon, A. (eds) Proceedings of the Third International Conference on Innovations in Computing Research (ICR'24). ICR 2024. Lecture Notes in Networks and Systems, vol 1058. Springer, Cham. [https://doi.org/10.1007/978-3-031-65522-7\\_44](https://doi.org/10.1007/978-3-031-65522-7_44)
- [40]. Jingar, N. K. (2026, February 13). Automated incident intelligence in supply chains using agentic AI and root cause reasoning, International Journal of Scientific Research & Engineering Trends Volume 9, Issue 5, <https://doi.org/10.5281/zenodo.18162511>
- [41]. Jingar, N. K. (2022). Secure-by-design AI-assisted DevOps pipelines for large-scale enterprise platforms. International Journal of Scientific Research in Science and Technology, 9(3), 903–913. <https://doi.org/10.32628/IJSRST2291348>

**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.