HSLiNets: Evaluating Band Ordering Strategies in Hyperspectral and LiDAR Fusion

I. Experiments

This section describes the experimental methodology, presents the obtained results, and offers a thorough analysis of the findings on the impact of hyperspectral band order on classification performance and feature fusion.

A. Experimental Setup

Our proposed framework is implemented using the PyTorch library. The model begins by fusing raw HSI and LiDAR data before embedding. These fused data patches are then processed through the model's forward and backward non-linear directional blocks, followed by spatial feature extraction modules. For hardware we used NVIDIA RTX 3090 GPU with 24 GB VRAM is used in the experiment. Hyper-paramters for training are: batch size 32; initial learning rate 0.00001, optimiser Adam, and epochs 100.

Table I defines four hyperspectral band order configurations and their LiDAR variants. Each configuration explores how spectral sequencing and LiDAR integration impact feature learning in our fusion framework. The order here is for the primary branch, and the secondary branch is the corresponding reversed order.

Six representative deep learning-based approaches are selected as competitors for comparison, as detailed below: S2FL [1], and CoSpace-L1 [2] algorithms, and the recent popular deep learning methods, such as two branch CNN [3], MAHiDFNet [4], FDCFNet [5], and UCAFNet [6]. These methods are evaluated both visually and quantitatively.

The evaluation is designed to rigorously assess the performance of the proposed parallel feature fusion method and its potential applications in hyperspectral imaging (HSI) and LiDAR data fusion. The experiments are conducted using the Houston 2013 and Trento datasets. The distribution of training and test samples, as detailed in Github, is utilised for this purpose.

These methods are evaluated both visually and quantitatively. Quantitative metrics include overall accuracy (OA), average accuracy (AA), and the statistical Kappa to assess classification performance comprehensively.

We first conducted a systematic investigation of hyperspectral band ordering and its effects on classification performance. The band order impact experiments include pixel level classification, 3D input data with various patch sizes, training with different number of samples, and generalisation on the other architectures. Detailed experiments are explained in the corresponding sections as follows.

B. Impact of Band Order on Single-pixel Analysis

Following the selected band methodology from supervised bands orders in [7], a patch size of 1 is chosen for our experiments to ensure sensitivity to variations in band order first. Smaller patch sizes are more susceptible to external influences [8], and using a patch size of 1 eliminates spatial context, allowing for a more precise evaluation of the band order's impact on classification performance. In contrast, larger patch sizes introduce spatial dependencies, which can obscure the direct effects of band order.

Based on the experimental outcome in Table II, there is a significant variance in classification accuracy across classes 4 to 13. A preliminary analysis of these 10 classes revealed that the normal band order achieved an average accuracy of 0.8824, while the reversed band order slightly outperformed it with an average accuracy of 0.8926. The importance-to-less band order resulted in an average accuracy of 0.8788, whereas the less-to-important band order achieved 0.8864. Among the normal and reversed HSI data band order, the reversed band order yielded the best results. Notably, Classes 4, 7, 8, 9, 10, 11, and 13 exhibited a 1% to 3% variance in accuracy for normal order and reversed; For the selected band two orders, the variance can be reached higher 6%, indicating that band order significantly influences classification performance.

After integrating LiDAR data, we recalculated the average accuracy for Classes 4 to 13. The normal band order achieved 0.9256, while the reversed band order slightly improved to 0.9261. The importance-to-less band order resulted in 0.9199, whereas the less-to-important order performed the best at 0.9277. Although the less-to-important order showed the highest accuracy, the normal and reversed band orders closely followed in second and third place. Only Classes 9, 11, and 13 exhibited noticeable variations, suggesting that LiDAR data enhances classification stability, making the results less sensitive to band order changes.

C. Impact of Band Order with Few Training Samples

We then evaluate the cases with fewer training samples. Specifically, we choose 10 samples per class for training. Table III, shows the results, and it demonstrates the band order impacting the classification performance when using very few training samples.

For HSI-Only Baselines (DB1–DB4), DB2 achieves the highest OA (0.8156), followed by DB3 (0.8113), DB4 (0.8044), and DB1 (0.7830). Band selection/ordering (DB2 vs. DB1) improves OA by 3.26%, highlighting spectral order's importance.

In the fusion task, DB1Li+DB2Li achieves the highest OA (0.8739), outperforming all HSI-only baselines by 5.83% (DB1 $\rightarrow$ DB1Li+DB2Li). Combining complementary band orders (DB1+DB2) with LiDAR yields the largest gain, validating your dual-path fusion design.

TABLE I
HYPERSPECTRAL BAND ORDER CONFIGURATIONS WITH LIDAR INTEGRATION

Data ID	**Band Order Type**	**Description**
DB1	Original Order	Hyperspectral bands arranged in their native wavelength sequence (e.g., 400–2500 nm).
DB2	Reversed Order	Hyperspectral bands inverted from their original sequence (e.g., 2500–400 nm).
DB3	Descending Importance Order	Selected bands ranked by importance (most → least discriminative) using attention weights from pre-trained models.
DB4	Ascending Importance Order	Selected bands ranked by importance (least → most discriminative) to test counterintuitive feature hierarchies.
DB1Li	Original Order + LiDAR	DB1 with LiDAR features (elevation, intensity) appended as pseudo-spectral bands at sequence start and end.
DB2Li	Reversed Order + LiDAR	DB2 with LiDAR features appended as pseudo-spectral bands.
DB3Li	Descending Importance + LiDAR	DB3 with LiDAR features appended as pseudo-spectral bands.
DB4Li	Ascending Importance + LiDAR	DB4 with LiDAR features appended as pseudo-spectral bands.

TABLE II
EXPERIMENTS SUMMARISATION COMPARISON RAW DATA
ORDER AND BAND SELECTION PATCH SIZE=1 (PIXEL LEVEL) STANDARD TRAINING SAMPLES.BLUE IS THE BEST METHOD, RED IS THE SECOND PLACE METHOD

Category	Based (HSI Only)				Based HSI + LiDAR)					
	DB1	DB2	DB3	DB4	DB1Li	DB2Li	DB1Li+DB2Li	DB3Li	DB4Li	DB3Li+DB4Li
Class1	0.9791	0.9772	0.9763	0.9706	0.9867	0.9734	0.9744	0.9848	0.9839	0.9848
Class2	0.9662	0.9709	0.9652	0.9718	0.9831	0.9831	0.9868	0.9887	0.9831	0.9868
Class3	1.0000	0.9980	0.9980	1.0000	1.0000	1.0000	0.9980	0.9960	1.0000	1.0000
Class4	0.9536	0.9678	0.9659	0.9763	0.9801	0.9886	0.9896	0.9848	0.9820	0.9886
Class5	0.9801	0.9773	0.9763	0.9782	0.9867	0.9839	0.9877	0.9811	0.9848	0.9830
Class6	0.9930	0.9860	0.9930	0.9860	0.9790	0.9860	0.9930	0.9860	0.9930	0.9930
Class7	0.8881	0.9188	0.9086	0.9039	0.9263	0.9496	0.9319	0.9142	0.9310	0.9347
Class8	0.8594	0.8784	0.8604	0.8765	0.9421	0.9459	0.9544	0.9506	0.9411	0.9563
Class9	0.8159	0.7866	0.8036	0.7790	0.8924	0.9122	0.9046	0.8857	0.8659	0.8971
Class10	0.8736	0.9073	0.9015	0.9044	0.9286	0.9498	0.9633	0.9556	0.9141	0.9344
Class11	0.8567	0.8662	0.8463	0.8624	0.9061	0.9042	0.9099	0.8861	0.9222	0.8833
Class12	0.8809	0.8905	0.8232	0.8847	0.8934	0.8799	0.9308	0.8972	0.9030	0.9145
Class13	0.7228	0.7474	0.7088	0.7123	0.8211	0.7789	0.8281	0.7579	0.8351	0.7754
Class14	0.9960	0.9960	0.9919	0.9960	0.9960	0.9960	0.9960	1.0000	0.9960	0.9960
Class15	0.9852	0.9873	0.9915	0.9873	1.0000	0.9873	0.9958	0.9915	1.0000	0.9979
OA	0.9110	0.9191	0.9086	0.9155	0.9459	0.9469	0.9568	0.9443	0.9452	0.9482
AA	0.9167	0.9237	0.9140	0.9193	0.9481	0.9467	0.9575	0.9440	0.9490	0.9484
Kappa	0.9035	0.9122	0.9008	0.9082	0.9412	0.9423	0.9531	0.9395	0.9405	0.9437

Class 8, 10, and 7 are low-performance classes. Class 8 (Lowest) is the worst in DB2 (0.4562), best in DB1Li+DB2Li (0.6345), improving 17.83%. LiDAR fusion mitigates spectral ambiguity in highly challenging classes. Class 10 (HSI-only) is worst in DB4 (0.4108), best in DB1Li+DB2Li (0.6204), improving 20.96%, it indicates fusion leverages elevation data to resolve spectral confusion. Class 7 DB1Li+DB2Li (0.9269), improving 28.22% compared with DB4 (0.6447), the fusion dramatically enhances performance for noise-prone classes with limited samples.

DB1Li+DB2Li (OA0.8739) and DB3Li+DB4Li (OA0.8558) outperform single-order fused configurations DB3Li (OA0.8388). Dual-path processing of complementary band orders maximises feature diversity, critical for fewer-data scenarios.

D. Impact of Band Order with Standard Training Samples

To validate our proposed model with different band order inputs on standard training test split, further experiments are conducted (as shown in Table IV): For the baseline (HSI-only: DB1–DB4), DB2 achieves the highest overall accuracy (OA) of 0.9967, followed by DB4 (0.9962), DB3 (0.9955), and DB1 (0.9952). The comparison between DB2 and DB1 shows that band selection and ordering improve OA by 0.15%, suggesting that spectral sequence significantly impacts feature learning.

In the fusion experiment, DB1Li+DB2Li and DB3Li+DB4Li both achieve the highest overall accuracy (OA) of 0.9989, outperforming all HSI-only baselines. Fusion improves OA by 0.37% (from DB1 to DB1Li+DB2Li) and 0.34% (from DB3 to DB3Li+DB4Li). Combining complementary HSI band orders (e.g., DB1+DB2) with

TABLE III
EXPERIMENTS SUMMARISATION COMPARISON RAW DATA
ORDER AND BAND SELECTION PATCH SIZE=9 BASED ON FEW SAMPLES 10

Category	Based on HSI				Based on HSI + LiDAR					
	DB1	DB2	DB3	DB4	DB1Li	DB2Li	DB1Li+DB2Li	DB3Li	DB4Li	DB3Li+DB4Li
Class1	0.9758	0.9774	0.9613	0.9750	0.9629	0.9790	0.9637	0.9790	0.9484	0.9573
Class2	0.8264	0.8376	0.8368	0.8280	0.8296	0.8352	0.8481	0.8280	0.8328	0.8432
Class3	0.9956	0.9971	0.9971	0.9942	0.9971	0.9956	1.0000	0.9971	0.9971	0.9927
Class4	0.7626	0.8630	0.8695	0.8225	0.8995	0.9141	0.9190	0.9109	0.9182	0.9198
Class5	1.0000	0.9992	1.0000	0.9968	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Class6	0.9968	0.9238	0.9175	0.9429	0.9556	0.9905	1.0000	0.9238	0.9968	1.0000
Class7	0.7544	0.6781	0.6987	0.6447	0.8855	0.8490	0.9269	0.8887	0.8967	0.8808
Class8	0.5162	0.4562	0.4765	0.4838	0.5462	0.5665	0.6345	0.5105	0.5721	0.5762
Class9	0.6723	0.7593	0.7295	0.8591	0.9356	0.8977	0.9163	0.9404	0.9122	0.9002
Class10	0.4462	0.5694	0.5062	0.4108	0.5522	0.4117	0.6204	0.4585	0.5440	0.5341
Class11	0.5902	0.8294	0.8408	0.8482	0.8147	0.5118	0.9053	0.7788	0.8180	0.8767
Class12	0.8806	0.8700	0.8937	0.8414	0.7719	0.8152	0.7874	0.8185	0.7073	0.7939
Class13	0.9172	0.8911	0.8301	0.8627	0.8497	0.8105	0.9020	0.9085	0.8824	0.9412
Class14	0.9904	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9976	1.0000	1.0000
Class15	0.9877	0.9985	0.9969	0.9985	0.9985	1.0000	0.9815	0.9846	0.9969	1.0000
OA	0.7830	0.8156	0.8113	0.8044	0.8452	0.8103	0.8739	0.8388	0.8431	0.8558
AA	0.8208	0.8433	0.8370	0.8339	0.8666	0.8385	0.8937	0.8617	0.8682	0.8811
Kappa	0.7658	0.8009	0.7962	0.7889	0.8327	0.7951	0.8637	0.8257	0.8305	0.8442

TABLE IV
EXPERIMENTS SUMMARISATION COMPARISON: RAW DATA
ORDER AND BAND SELECTION BASED ON STANDARD TRAINING

Category	Based on HSI				Based on HSI + LiDAR					
	DB1	DB2	DB3	DB4	DB1Li	DB2Li	DB1Li+DB2Li	DB3Li	DB4Li	DB3Li+DB4Li
Class1	0.9858	1.0000	0.9896	0.9905	0.9924	0.9905	0.9962	0.9877	0.9896	0.9991
Class2	0.9991	1.0000	1.0000	0.9991	1.0000	0.9991	1.0000	1.0000	0.9991	1.0000
Class3	1.0000	1.0000	0.9980	1.0000	1.0000	0.9980	1.0000	1.0000	1.0000	1.0000
Class4	0.9962	0.9991	0.9981	0.9991	0.9991	0.9981	1.0000	0.9991	0.9991	1.0000
Class5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9972	1.0000	0.9991
Class6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Class7	0.9823	0.9851	0.9841	0.9888	0.9916	0.9963	0.9916	0.9897	0.9879	0.9925
Class8	1.0000	0.9991	0.9991	1.0000	1.0000	1.0000	0.9867	1.0000	1.0000	1.0000
Class9	0.9953	0.9972	0.9943	0.9962	0.9981	0.9972	1.0000	0.9981	1.0000	1.0000
Class10	0.9990	1.0000	0.9981	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Class11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Class12	0.9894	0.9837	0.9866	0.9846	0.9914	0.9942	0.9981	0.9942	0.9885	0.9962
Class13	0.9895	0.9930	0.9965	0.9930	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Class14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Class15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
OA	0.9952	0.9967	0.9955	0.9962	0.9976	0.9978	0.9989	0.9975	0.9969	0.9989
AA	0.9958	0.9971	0.9963	0.9968	0.9982	0.9982	0.9992	0.9977	0.9976	0.9991
Kappa	0.9947	0.9964	0.9951	0.9959	0.9974	0.9976	0.9988	0.9968	0.9966	0.9988

LiDAR maximises performance, validating the effectiveness of the dual-path fusion design.

The DB1Li configuration (OA 0.9976) shows a 0.24% improvement compared to DB1 (OA 0.9952). This demonstrates that integrating LiDAR data helps compensate for suboptimal hyperspectral imaging (HSI) band orders, such as those in DB1. Classes 2, 3, 6, and 10–15 achieve near-perfect accuracy across all configurations. These classes are spectrally and spatially well-defined, making them less affected by band order variations or LiDAR fusion. For HSI-only configurations, Class 7 achieves its lowest accuracy in DB1 (0.9823) but improves with DB2 (0.9851). The highest accuracy (0.9925) is achieved using DB3Li + DB4Li, representing a 1.02% improvement over HSI-only configurations. Class 12 reaches its highest accuracy (0.9981) in DB1Li + DB2Li, showing a 1.44% improvement compared to HSI-only setups. This indicates that LiDAR fusion reduces the sensitivity of noise-prone classes to band-order variations. Configurations that combine multiple band orders, such as DB1Li + DB2Li and DB3Li + DB4Li, outperform those that use a single band order with fusion. This supports the effectiveness of your symmetric architecture in leveraging complementary band sequences for improved classification accuracy.

E. Impact of HSI Band Order on Cross-Architecture Validation

In this experiment, we use 10 samples and standard samples with four different band orders to run on another network

TABLE V
ARCHITECTURE PERFORMANCE COMPARISON:
TWO BRANCH AND MODIFIED TWO BRANCH MODELS BASED ON FEW TRAINING SAMPLES

	Two Branch		Modified Two Branch			Two Branch		Modified Two Branch		
Class	**DB1+Li**	**DB2+Li**	**DB1Li**	**DB2Li**	**DB1Li+DB2Li)**	**DB3+Li**	**DB4+Li**	**DB3Li**	**DB4Li**	**DB3i+DB4Li**
Class1	0.9234	0.8453	0.8525	0.8566	0.8558	0.9315	0.8590	0.8566	0.8566	0.8566
Class2	0.7902	0.7741	0.8264	0.6777	0.8312	0.7066	0.8127	0.9960	0.9775	0.9928
Class3	1.0000	0.9869	0.9840	0.9782	0.9592	0.9913	1.0000	0.9898	0.9869	0.9884
Class4	0.9830	0.9335	0.9417	0.9344	0.9425	0.9400	0.9376	0.9400	0.9392	0.9400
Class5	0.9951	0.9870	0.9951	0.9943	0.9919	0.9959	0.9959	1.0000	1.0000	1.0000
Class6	0.9238	0.9587	0.9778	1.0000	0.9905	0.9302	0.9079	0.9429	0.9175	0.9810
Class7	0.7901	0.8983	0.8378	0.8696	**0.8951**	0.8784	0.8601	0.9467	**0.9348**	**0.9595**
Class8	0.6864	0.7342	0.8079	0.6126	0.7699	0.5827	0.7139	0.7601	0.8047	0.7658
Class9	0.9461	0.7520	0.5628	0.6852	0.6538	0.6916	0.9275	0.5145	0.6473	0.6103
Class10	0.3722	0.1832	**0.6294**	**0.6689**	**0.7050**	0.3788	0.3048	**0.7313**	0.6984	**0.7699**
Class11	0.2800	0.4841	0.4269	0.3551	0.3976	0.6596	0.6759	0.4751	0.5910	0.6310
Class12	0.3876	0.1962	**0.9150**	**0.9002**	0.9150	0.3197	0.3017	0.9141	0.8585	0.9141
Class13	0.9455	**0.9608**	0.9499	0.9434	**0.9499**	0.9586	**0.9608**	0.9129	0.9521	0.9346
Class14	0.9833	0.9689	0.9306	0.9306	0.9282	0.8828	0.9641	0.9258	0.9522	0.9545
Class15	1.0000	0.9954	0.9908	0.9846	1.0000	1.0000	0.9954	0.9985	0.9938	0.9954
OA	0.7613	0.7314	**0.8122**	0.7920	**0.8252**	0.7526	0.7802	**0.8390**	0.8546	0.8664
AA	0.8004	0.7776	0.8419	0.8261	**0.8524**	0.7898	0.8145	**0.8603**	0.8740	0.8863
Kappa	0.7419	0.7105	0.7971	0.7752	**0.8110**	0.7326	0.7627	0.8259	0.8428	0.8555

architecture: two branch CNN. We modified the original two branch CNN to fit our overall framework. It demonstrated that our proposed concept works on different network architecture.

Table V presents the results of different architectures based on 10 samples. The results highlight that the modified two-branch architecture demonstrates significant improvements over the original architecture across individual class accuracies, overall accuracy (OA), and average accuracy (AA). The result of DB1+Li is the baseline.

In the original two-branch architecture, although the OA and AA for the reversed band order (DB2+Li) are slightly lower compared to the original band order (DB1+Li), specific classes such as Class 8 (C8) and Class 13 (C13) show improved performance with DB2+Li. Additionally, the performance of DB3+Li and DB4+Li is comparable to that of DB1+Li, indicating the robustness of different band order configurations.

For the modified two-branch architecture, there is a noticeable increase in OA, ranging from 3% to 5% for DB1Li, DB2Li, and the fusion of DB1Li and DB2Li. More remarkably, DB3Li, DB4Li, and the combination of DB3Li+DB4Li exhibit an OA improvement of nearly 10% compared to DB1+Li from the original architecture. This substantial gain demonstrates the effectiveness of the modified architecture in enhancing classification performance.

These findings confirm that band order plays a critical role in model performance. The fusion strategy, which integrates different band orders, effectively leverages the strengths of each configuration, leading to improved overall classification results.

Table VI presents results of different architectures based on standard training samples, with DB1+Li serving as the baseline for comparison. In the original two-branch architecture, the OA and AA for the reversed band order (DB2+Li) are slightly higher compared to the original band order (DB1+Li).

Notably, Class 9 (C9) exhibits a performance improvement of over 6% with DB2+Li. Additionally, except for DB2Li, which shows a slightly lower performance than the baseline, all other configurations outperform the baseline model.

In the modified two-branch architecture, there is a noticeable increase in OA, ranging from 0.1% to 0.8% for DB1Li, DB2Li, and the fusion of DB1Li and DB2Li. Although modest, this improvement is consistent. Moreover, DB3Li, DB4Li, and the combination of DB3Li+DB4Li achieve an OA increase of approximately 0.1% compared to the original DB1+Li architecture. These gains highlight the effectiveness of the modified architecture in enhancing classification performance.

While increasing the number of training samples appears to reduce the impact of band order on performance, its influence remains evident. These findings confirm that band order plays a critical role in model performance. The fusion strategy, which integrates different band orders, effectively leverages the strengths of each configuration, leading to improved overall classification outcomes.

F. Impact of Band Order on Data Fusion with Different Patch Sizes

We continue to explore the band order impacting fusion analysis based on different patch sizes 1, 3, 5, 7, 9, 11, 13, 15 of raw data on Houston 2013 and Treno, respectively.

1) Impact of Band Order on Data Fusion with Different Patch Sizes on Houston Dataset: The classification performance for each patch size was assessed using Overall Accuracy (OA), Average Accuracy (AA), and Kappa coefficient, with the Houston 2013 data set results presented in Table VII.

Classes C1, C3, C5, C6, C14, and C15 exhibit near-perfect or perfect classification accuracy across all patch sizes. Specifically, when the patch size exceeds 5, all classes (C1–C15) achieve perfect classification, indicating that these classes are well-separated and relatively easy to classify.

Classes C2, C10, C12, and C13 perform well overall but exhibit minor variations in accuracy for smaller patch sizes (e.g., P1 and P3). These slight dips in performance suggest that

TABLE VI
ARCHITECTURE PERFORMANCE COMPARISON:
TWO BRANCH AND MODIFIED TWO BRANCH MODELS BASED ON STANDARD TRAINING SAMPLES

	Two Branch		Modified Two Branch			Two Branch		Modified Two Branch		
Class	**DB1+Li**	**DB2+Li**	**DB1Li**	**DB2Li**	**DB1Li+DB2Li)**	**DB3+Li**	**DB4+Li**	**DB3Li**	**DB4Li**	**DB3i+DB4Li**
Class1	0.9953	0.9801	0.9972	0.9117	0.9839	0.9820	0.9801	0.9829	0.9829	0.9829
Class2	0.9915	0.9934	0.9850	0.9981	0.9991	0.9953	0.9878	0.9784	0.9699	0.9699
Class3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Class4	0.9991	0.9972	0.9839	0.9886	0.9896	0.9972	1.0000	0.9962	0.9972	0.9962
Class5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Class6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Class7	0.9897	0.9916	0.9897	0.9935	0.9935	0.9897	0.9674	0.9748	0.9729	0.9748
Class8	0.9753	1.0000	0.9820	0.9772	0.9848	0.9715	1.0000	0.9763	0.9677	0.9801
Class9	0.9216	0.9858	0.9906	0.9849	0.9915	0.9868	0.9603	0.9868	0.9915	0.9877
Class10	1.0000	0.9990	0.9527	0.8861	0.9459	1.0000	0.9981	0.9952	0.9942	0.9942
Class11	0.9943	0.9972	0.9943	1.0000	1.0000	1.0000	1.0000	0.9915	0.9905	0.9924
Class12	0.9962	0.9376	0.9962	0.9885	0.9933	0.9894	0.9827	0.9933	0.9971	0.9962
Class13	0.9860	1.0000	1.0000	0.9860	1.0000	1.0000	1.0000	0.9509	0.9509	0.9649
Class14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Class15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
OA	0.9873	0.9898	0.9889	0.9764	0.9898	0.9924	0.9893	0.9880	0.9870	0.9883
AA	0.9899	0.9920	0.9914	0.9810	0.9921	0.9941	0.9918	0.9884	0.9877	0.9893
Kappa	0.9862	**0.9890**	0.9880	0.9744	0.9890	0.9917	0.9883	0.9870	0.9859	0.9873

TABLE VII
MODEL CLASSIFICATION PERFORMANCE BASED ON HOUSTON 2013 DATASET DIFFERENT PATCHES
BLUE IS THE BEST METHOD RED IS THE SECOND PLACE METHOD

Item	**PatchSize1**	PatchSize3	PatchSize5	PatchSize7	PatchSize9	PatchSize11	PatchSize13	PatchSize15
C1	0.9744	0.9991	1.0000	1.0000	0.9924	0.9905	0.9810	0.9972
C2	0.9868	0.9953	0.9953	1.0000	1.0000	1.0000	1.0000	0.9897
C3	0.9980	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
C4	0.9896	0.9981	1.0000	1.0000	0.9991	1.0000	1.0000	1.0000
C5	0.9877	1.0000	1.0000	0.9981	1.0000	1.0000	1.0000	1.0000
C6	0.9930	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
C7	0.9496	0.9925	1.0000	1.0000	1.0000	0.9916	1.0000	0.9935
C8	0.9554	0.9981	1.0000	1.0000	1.0000	0.9867	0.9744	0.9744
C9	0.9046	0.9792	0.9906	0.9943	0.9962	1.0000	1.0000	1.0000
C10	0.9633	0.9971	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
C11	0.9099	0.9943	0.9981	1.0000	1.0000	1.0000	1.0000	1.0000
C12	0.9308	0.9962	1.0000	1.0000	1.0000	0.9981	0.9962	0.9952
C13	0.8281	0.9895	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
C14	0.9960	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
C15	0.9958	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
OA	0.9568	0.9954	0.9986	0.9993	0.9989	0.9971	0.9958	0.9957
AA	0.9575	0.9960	0.9989	0.9995	0.9992	0.9978	0.9968	0.9967
Kappa	0.9531	0.9950	0.9985	0.9993	0.9988	0.9969	0.9955	0.9953

these classes are somewhat sensitive to specific patch configurations, likely due to their spatial or spectral characteristics. C1, C3, C5, and C6 show near-perfect performance across all patch sizes, with only slight variability, demonstrating strong robustness and stability. Compared to Patch size 3, 5, 7, 9, 11, 13, 15, patch size 1's classes C1, C2, C5, C6, C7, C8, C9, C10, C11, C12, C13 are much lower from 2%(C2) to 16%(C13).

Based on the experimental result, the best classification can be obtained when the patch size is 7 for Houston 2013.

2) Impact of Band Order on Data Fusion with Different Patch Sizes on Trento Dataset: Table VIII presents the classification performance of our proposed model on the Trento dataset using different patch sizes: P1, P3, P5, P7, P9, P11, P13, and P15. The results are evaluated in terms of Overall Accuracy (OA), Average Accuracy (AA), and Kappa coefficient, as well as the classification accuracy for each individual class. The best results are highlighted in blue, while the second-best results are marked in red.

Classes C1, C3, C4, and C5 consistently demonstrate near-perfect or perfect accuracy across most or all patch configurations. C3 and C4 achieve perfect classification under several patch sizes, with C4 attaining flawless performance across all patches. C5 exhibits near-perfect accuracy, with only minor variability observed across different patches.

C1 shows strong overall performance, although slight accuracy dips are noted in P5 (0.9936) and P7 (0.9895). Moderately Robust Classes C2 and C6 display good overall performance but show noticeable variability across patch sizes: C2 shows

TABLE VIII
MODEL CLASSIFICATION PERFORMANCE BASED ON TRENTO DATASET DIFFERENT PATCHES
BLUE IS THE BEST METHOD RED IS THE SECOND PLACE METHOD

Item	PatchSize1	PatchSize3	PatchSize5	PatchSize7	PatchSize9	PatchSize11	PatchSize13	PatchSize15
C1	0.8428	0.9880	0.9936	0.9895	0.9887	0.9926	0.9928	0.9936
C2	0.9611	0.9741	0.9712	0.9748	0.9874	0.9921	0.9946	0.9971
C3	0.9786	1.0000	1.0000	1.0000	1.0000	0.9973	1.0000	0.9973
C4	0.9969	0.9994	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000
C5	0.8741	0.9826	0.9989	0.9999	0.9999	1.0000	0.9994	1.0000
C6	0.9410	0.9817	0.9895	0.9784	0.9898	0.9898	0.9856	0.9626
OA	0.9239	0.9878	0.9949	0.9939	0.9962	0.9972	0.9968	0.9950
AA	0.9324	0.9876	0.9922	0.9904	0.9943	0.9953	0.9954	0.9918
Kappa	0.8986	0.9836	0.9932	0.9919	0.9949	0.9962	0.9958	0.9933

slightly lower accuracy in P5 (0.9712) and P3 (0.9741), while higher performance is observed in P11 (0.9921) and P13 (0.9946).

C6 exhibits the highest variability among all classes, with lower accuracy in P1 (0.9410) and P15 (0.9626), while peaking in P5 (0.9895) and P9 (0.9898). This suggests that C6 may represent a more challenging class due to potential spectral or spatial overlap with other classes. Challenging Class C6 is identified as the most challenging class to classify, with significant variability across patch sizes. The lower performance observed under certain configurations highlights its increased complexity and potential misclassification due to similarity with other classes. Easiest Classes to Classify C3 and C4 emerge as the easiest classes to classify, consistently achieving perfect or near-perfect accuracy across all patch sizes.

P11 emerges as the best-performing patch configuration, achieving: Highest Overall Accuracy (OA) 0.9972, Highest Average Accuracy (AA), 0.9953, and The highest Kappa coefficient was 0.9962. These results indicate that P11 provides the most balanced and consistent classification performance across all classes. P5, P9, and P13 also show strong performance, with high OA, AA, and Kappa values. However, they are slightly outperformed by P11 in all metrics.

In summary, P11 is identified as the optimal patch size for the Trento dataset, delivering the highest classification accuracy and consistency across all classes. While other patch sizes such as P5, P9, and P13 offer competitive performance, they fall slightly short of P11 in terms of overall metrics. Among the classes, C3 and C4 are the easiest to classify, while C6 remains the most challenging due to its variability and lower accuracy under certain configurations.

Figure 1 shows that the highest overall accuracy 0.9993 for the Houston 2013 dataset is achieved when the patch size is 7. The Trento dataset's best OA performance, 0.9772 is observed at a patch size of 11. This analysis highlights that the optimal patch size varies depending on the dataset, likely due to differences in spatial characteristics. Fine-tuning the patch size is, therefore, essential for maximising model performance on specific datasets.

G. Comparison with State-of-Arts and Result Analysis

We present a quantitative comparison of previous state of the art networks, and show the superioty of our method in classification performance.

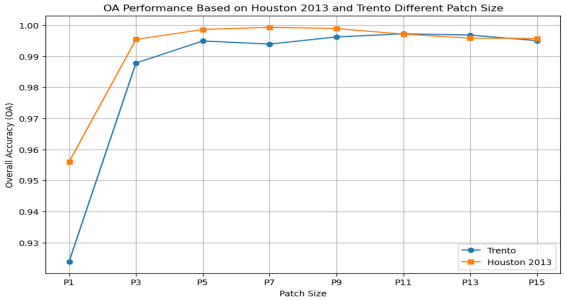

Fig. 1. OA and Patch Size Relationship based on Houston 2013 and Trento Data Sets

1) Quantitative Comparison of Houston 2013 Dataset Experiments: Table IX presents a quantitative comparison of various fusion methods applied to the Houston 2013 dataset. The best classification results for each class are highlighted in BLUE, while the second-best results are indicated in RED.

Our proposed method demonstrated superior performance, achieving the highest classification accuracy in 15 out of 15 classes, which underscores its robustness and reliability across a diverse range of categories. Additionally, it achieved the highest OA (0.9972), AA (0.9953), and Kappa coefficient (0.9962), consistently outperforming the other models.

UCAFNet secured the best classification accuracy in 3 classes (C4, C5, and C15) and ranked second in 6 classes (C2, C6, C8, C11, C13, and C14), demonstrating competitive performance and versatility. FDCFNet performed well, achieving the second-best accuracy in 4 classes (C1, C4, C5, and C9). MAHiDFNet ranked second in 2 classes (C10 and C12), showing its strength in specific categories, though it did not achieve top performance in any class.

Although UCAFNet showed strong performance in several classes, its overall results remained slightly below those of the Dual Directional Networks. FDCFNet also achieved competitive results, ranking among the top methods in certain classes and securing the second-best position for multiple categories.

Overall, the Dual Directional Networks consistently outperformed other models, achieving the highest accuracy across all classes and in the overall metrics. This demonstrates the robustness and effectiveness of our proposed method for hyperspectral image classification on the Houston 2013

TABLE IX
CLASSIFICATION PERFORMANCE COMPARISON BETWEEN OUR FUSION MODEL AND EXISTING FUSION MODELS BASED ON STANDARD HOUSTON 2013 DATASET. BLUE IS THE BEST METHOD, RED IS THE SECOND PLACE METHOD

Class No	S2FL [9]	CoSpace-L1 [2]	TwoBranchCNN [3]	MAHiDFNet [4]	FDCFNet [5]	UCAFNet [6]	DualDirectionalNets	OurBest
C1	0.9057	0.8945	0.9953	0.9853	0.9745	0.9376	0.9905	0.9962
C2	0.9769	0.9705	0.9915	0.9287	1.0000	0.9944	1.0000	1.0000
C3	1.0000	1.0000	0.9627	0.9111	0.9684	0.9880	1.0000	1.0000
C4	0.9839	0.9847	0.9991	0.9810	0.9960	0.9425	1.0000	1.0000
C5	0.9936	1.0000	0.9756	0.9838	1.0000	1.0000	1.0000	1.0000
C6	0.9939	0.9969	1.0000	0.9558	1.0000	0.9815	1.0000	1.0000
C7	0.8052	0.8052	0.9801	0.9897	0.9464	0.9448	0.9916	0.9916
C8	0.6857	0.8939	0.9753	0.8094	0.9711	0.9630	0.9867	0.9867
C9	0.6989	0.6486	0.9216	0.9804	0.9468	0.9449	1.0000	1.0000
C10	0.7001	0.6675	1.0000	0.7281	0.9438	0.9576	1.0000	1.0000
C11	0.8834	0.8842	0.9943	0.7271	0.9830	0.9976	1.0000	1.0000
C12	0.8427	0.8402	0.9962	0.7680	0.9303	0.9278	0.9981	0.9981
C13	0.8358	0.7868	0.9860	0.9580	0.9182	0.9446	1.0000	1.0000
C14	1.0000	1.0000	0.9639	0.9953	0.9977	1.0000	1.0000	1.0000
C15	0.9894	0.9818	1.0000	1.0000	0.9037	1.0000	1.0000	1.0000
OA	0.8681	0.8752	0.9878	0.8958	0.9661	0.9680	0.9971	0.9989
AA	0.8863	0.8899	0.9899	0.9136	0.9653	0.9692	0.9978	0.9992
Kappa	0.8575	0.8651	0.9867	0.8874	0.9634	0.9654	0.9969	0.9988

dataset. Additionally, UCAFNet and FDCFNet showed notable performance, making them viable alternatives for specific classification tasks.

2) Quantitative Comparison of Trento Dataset Experiments: Table X presents a comparative analysis of the classification performance of our proposed networks against six existing models using the Trento dataset. The evaluation includes three primary metrics—Overall Accuracy (OA), Average Accuracy (AA), and Kappa coefficient—as well as detailed accuracy for individual classes.

In terms of class-level performance, our method achieves the highest accuracy in 2 classes (C2 and C4), demonstrating its robustness in these categories. FDCFNet outperformed other methods in 4 classes (C1, C3, C4, and C5), indicating its strength in specific instances. However, the performance gap between FDCFNet and our method in these classes remains relatively minor. Additionally, Dual Directional Networks secured second-best performance in 3 classes (C1, C3, and C6), confirming its consistent and competitive classification capability across multiple categories.

Regarding other methods, UCAFNet shows notable versatility, ranking as the second-best method in 2 classes (C5 and C6). Despite its occasional top performance, its overall results were less competitive compared to our approach.

With respect to overall metrics, our method also achieves the highest scores, with an OA of 0.9972, AA of 0.9953, and Kappa coefficient of 0.9962, consistently outperforming other methods. UCAFNet ranked second overall, achieving an OA of 0.9926, AA of 0.9876, and Kappa coefficient of 0.9909, making it a strong alternative.

H. Classification Map: Visual Comparison

Figs. 2 and 3 visualise the classification maps of different compared methods in the full scene. From the classification map, we can see that all other methods are relatively worse than our method.

Figures 2 and 3 illustrate the classification maps generated by our proposed Model and competing models. Our method's map shows well-defined boundaries and fewer misclassifications in areas where traditional models often show more noises.

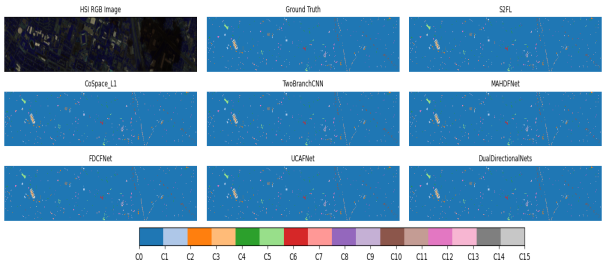

Fig. 2. Visualization and classification maps for the Houston 2013 dataset. Ground-truth map, six comparative methods and Our Dual Directional Nets.

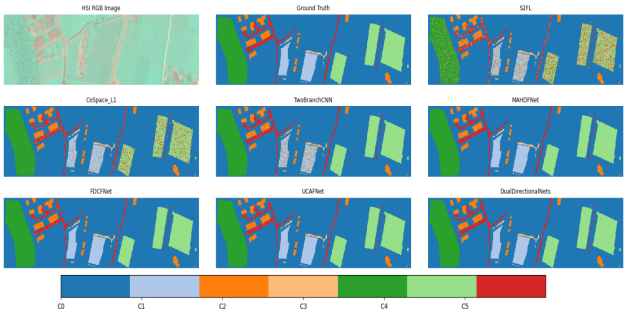

Fig. 3. Visualization and classification maps for the Trento dataset. Ground-truth map, six comparative methods and Our Dual Directional Nets.

In summary, the results indicate that our method not only delivers top-tier performance in several individual classes but also achieves superior overall classification accuracy, underscoring its effectiveness for hyperspectral image classification tasks on the Trento dataset.

I. Ablation Study

The ablation study evaluates the significance of the bidirectional processing mechanism in the SS linear model by

TABLE X
CLASSIFICATION PERFORMANCE COMPARISON BETWEEN OUR FUSION OUTCOME WITH FEWER SELECTED **10** BANDS AND EXISTING FUSION MODELS USING WITH FULL BANDS **63** BASED ON STANDARD TRENTO DATASET. BLUE IS THE BEST METHOD, RED IS THE SECOND PLACE METHOD

Class No	**S2FL [9]**	**CoSpace-L1 [2]**	**TwoBrachCNN [3]**	**MAHiDFNet [4]**	**FDCFNet [5]**	**UCAFNet [6]**	**DualDirectionalNets**
C1	0.7102	0.8604	0.7794	0.9919	1.0000	1.0000	0.9926
C2	0.8257	0.9583	0.8929	0.8892	0.9829	0.9852	0.9921
C3	0.9207	0.9562	0.7003	0.9753	0.9948	0.9916	0.9973
C4	0.8669	0.9873	1.0000	0.9998	1.0000	1.0000	1.0000
C5	0.4890	0.6356	0.9960	0.9990	1.0000	0.9980	1.0000
C6	0.7757	0.8916	0.9774	0.9978	0.9323	0.9505	0.9898
OA	0.7020	0.8348	0.89418	0.9859	0.9911	0.9926	0.9972
AA	0.7647	0.8816	0.8910	0.9755	0.9850	0.9876	0.9953
Kappa	0.6191	0.7866	0.8923	0.9812	0.9881	0.9909	0.9962

systematically removing key components—forward and backward pathways, and the spatial processing block. This analysis highlights the contribution of each component to classification accuracy. Conducted on the Houston 2013 dataset, the study provides valuable insights into the role of these elements in enhancing the model's overall performance.

The full architecture includes 3 modules: Forward, Reversed, and Spatial. In the Table XI, Method1 contains 3 complete modules; Method2 includes forward and backward modules; Method3 includes forward and spatial modules; Method4 includes reversed and spatial moduels; Method5 includes spatial methods only.

Table XI presents the outcomes of the ablation analysis, showcasing the performance metrics for five different methods. The results clearly demonstrate the essential role of each component.

TABLE XI
DIFFERENT METHODS FOR ABLATION ANALYSIS BASED ON HOUSTON 2013 DATA (HSI+LIDAR), PATCH SIZE 11

Methods	Forward	Reversed	Spatial	OA	AA	Kappa
Method1	✓	✓	✓	0.9971	0.9978	0.9969
Method2	✓	✓	×	0.9685	0.9763	0.9650
Method3	✓	×	✓	0.9761	0.9815	0.9732
Method4	×	✓	✓	0.9715	0.9791	0.9682
Method5	×	×	✓	0.9342	0.9434	0.9286

II. CONCLUSION

In this study, a novel framework is introduced, named after HSLiNets. The framework harnesses bidirectional reversed networks to fuse HSI and LiDAR features efficiently, leading to improvements in classification performance. The proposed model, which eliminates the need for self-attention modules, achieves modelling capabilities comparable to state-of-the-art methods.

Experimental evaluations conducted on the Houston 2013 and Trento datasets demonstrate the superior performance of our proposed model. The HSLiNets consistently outperforms existing methods, achieving the highest scores in overall accuracy (OA), average accuracy (AA), and kappa coefficient. These results underscore the effectiveness of our approach in capturing intricate multimodal data characteristics and establishing new performance benchmarks in remote sensing applications.

REFERENCES

[1] J. Yao, D. Hong, L. Gao, and J. Chanussot, "Multimodal remote sensing benchmark datasets for land cover classification," in *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2022, pp. 4807–4810.

[2] D. Hong, J. Chanussot, N. Yokoya, J. Kang, and X. X. Zhu, "Learning-shared cross-modality representation using multispectral-lidar and hyperspectral data," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 8, pp. 1470–1474, 2020.

[3] X. Xu, W. Li, Q. Ran, Q. Du, L. Gao, and B. Zhang, "Multisource remote sensing data classification based on convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 2, pp. 937–949, 2017.

[4] X. Wang, Y. Feng, R. Song, Z. Mu, and C. Song, "Multi-attentive hierarchical dense fusion net for fusion classification of hyperspectral and lidar data," *Information Fusion*, vol. 82, pp. 1–18, 2022.

[5] S. Zhang, X. Meng, Q. Liu, G. Yang, and W. Sun, "Feature-decision level collaborative fusion network for hyperspectral and lidar classification," *Remote Sensing*, vol. 15, no. 17, p. 4148, 2023.

[6] X. Meng, S. Zhang, Q. Liu, G. Yang, and W. Sun, "Uncertain category-aware fusion network for hyperspectral and LiDAR joint classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–15, 2024.

[7] J. X. Yang, J. Zhou, J. Wang, H. Tian, and A. W. C. Liew, "Lidar-guided cross-attention fusion for hyperspectral band selection and image classification," *IEEE Transactions on Geoscience and Remote Sensing*, 2024.

[8] Z. Chen, H. Yang, Q. Liu, Y. Liu, M. Zhu, and X. Liang, "Deep learning for hyperspectral image classification: A critical evaluation via mutation testing," *Remote Sensing*, vol. 16, no. 24, p. 4695, 2024.

[9] D. Hong, J. Hu, J. Yao, J. Chanussot, and X. X. Zhu, "Multimodal remote sensing benchmark datasets for land cover classification with a shared and specific feature learning model," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 178, pp. 68–80, 2021.