Supplementary Material for Pixel-Wise Shuffling with Collaborative Sparsity for Melanoma Hyperspectral Image Classification

A. Effect of Pixel-wise Shuffling (PWS) on model Performance

Figure 1 shows the graphical representation of the proposed method with and without Pixel-wise Shuffling (PWS). As the percentage of training samples increases, there is a clear trend of improvement across all metrics. The quantitative results of the proposed method in terms of overall accuracy (OA) and average accuracy (AA) are significantly lower without the pixel-wise shuffling mechanism, which is highly important for accurate melanoma classification in the proposed approach.

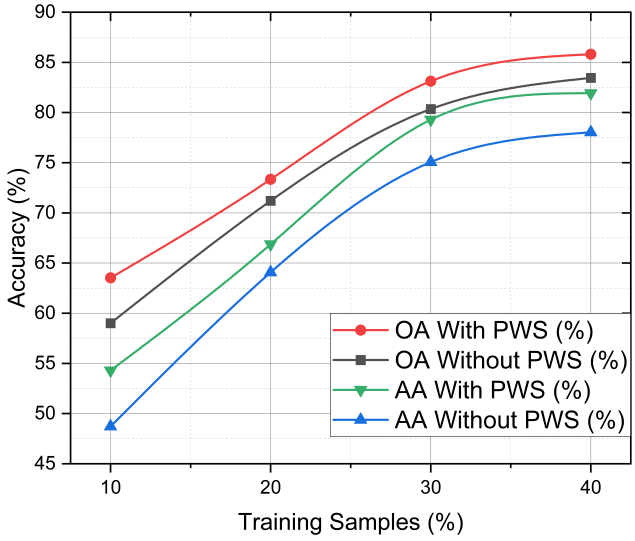

Figure 1. Graphical representation of the proposed method with and without Pixel-wise Shuffling (PWS) using different training samples.

B. Ablation II: Sensitivity Parameters Analysis of the Proposed Model

We evaluate the impact of the collaborative sparse term (φ) and the shuffling window size on the performance of the hyperspectral image classification model. The design parameter φ balances the fidelity of spectral unmixing reconstruction against the sparsity of the abundance matrix, supported by spatial correlation. A lower φ value results in a denser abundance matrix, capturing detailed information but increasing computational complexity, as reflected by the GFLOPs metric in Table 1. The computational requirements and efficiency of the proposed method are important for practical clinical applications. The analysis shows that the optimal configuration ($\varphi = 1 \times 10^{-2}$ and a shuffling window size of 3) achieves the highest classification accuracy with a reasonable computational cost of 24.3472 GFLOPs. This efficiency is achieved by integrating collaborative sparse unmixing and advanced pixel-wise shuffling techniques, which enhance the model's accuracy without significantly increasing computational demand. Despite its streamlined architecture, the method maintains high performance, making it suitable for practical applications in medical diagnostics where computational resources may be limited. This balance is essential for real-time operation and integration with standard medical imaging equipment without extensive upgrades. Lower φ values reduce computational load but compromise classification accuracy, highlighting the need for a trade-off between computational efficiency and model performance. Ensuring compatibility with existing infrastructure and optimizing computational efficiency is vital for seamless integration into clinical workflows, ultimately enhancing melanoma diagnosis accuracy and improving patient outcomes. The study finds the highest precision at $\varphi = 1 \times 10^{-2}$ and a window size of 2, indicating accurate pixel classification. The sensitivity peaks at the same φ value but with a window size of 3, showing the model's effectiveness in identifying relevant pixels. The optimal configuration is $\varphi = 1 \times 10^{-2}$ and a shuffling window size of 3, achieving the highest classification accuracy with reasonable computational cost. While a lower φ reduces computational load, it compromises on classification accuracy.

C. Ablation III: Effect of Hybrid Attention and Spatio-temporal Encoding Blocks on Model Performance

We perform further comprehensive ablation studies on the hybrid attention mechanism (cross and self-attention)

Table 1. Ablation II: Sensitivity Parameters Analysis of the Proposed Model (20% training samples)

Collaborative Sparse Term (φ)	**Shuffling Window size**	**OA (%)**	**AA (%)**	**Kappa**	**PRE (%)**	**SEN (%)**	**STD**	**GFLOPs**
1×10^{-1}	1	68.09	59.08	42.14	67.02	55.30	0.6643	90.9444
	2	69.67	60.97	45.16	68.68	57.95	0.6689	94.6101
	3	67.27	57.95	40.13	65.97	50.61	0.6536	94.8348
	4	65.10	55.53	36.59	64.20	49.97	0.6442	95.5837
	5	65.33	56.18	37.48	63.59	50.54	0.6840	95.8107
1×10^{-2}	1	70.65	62.59	47.06	70.96	58.60	0.6633	22.9670
	2	72.62	65.28	51.25	**73.14**	63.14	0.6737	24.3398
	3 (Proposed)	**73.34**	**66.87**	**53.33**	72.73	**66.40**	0.7051	24.3472
	4	72.33	64.54	50.53	72.04	61.56	0.6924	24.4530
	5	70.90	63.01	48.24	70.81	61.31	0.6930	24.4604
1×10^{-3}	1	68.92	61.36	45.07	67.89	61.01	0.6928	10.7003
	2	69.50	60.91	44.98	68.84	56.31	0.6696	10.7051
	3	71.49	64.11	49.43	71.14	60.25	0.7063	10.7051
	4	71.01	63.56	48.31	70.32	59.02	0.6979	10.9520
	5	71.41	64.13	49.61	70.67	62.66	**0.7175**	10.9570
1×10^{-4}	1	67.47	58.28	41.10	66.93	54.70	0.6544	5.9769
	2	69.73	60.84	45.30	69.53	54.13	0.6904	6.4399
	3	68.47	59.50	43.08	67.68	56.16	0.6600	6.4399
	4	71.32	63.49	49.06	71.07	60.16	0.7057	6.4436
	5	68.04	60.67	43.86	67.28	61.11	0.6854	**6.4436**

Table 2. Ablation III: Effect of Hybrid (Self+Cross) Attention Mechanism and Spatio-temporal Encoding Blocks on Model Performance

Encoding Mechanism		**Attention Mechanism**		**Metrics**					
Temporal	**Spatial**	**Self-Attention**	**Cross-Attention**	**OA**	**AA**	**Kappa**	**PRE**	**SEN**	**STD**
×	✓	✓	×	65.41	56.77	38.97	65.51	53.74	0.6512
✓	×	✓	×	65.86	56.93	39.26	66.91	55.05	0.6536
×	✓	×	✓	67.32	57.71	42.16	68.78	58.19	0.6653
✓	×	×	✓	68.79	59.08	44.36	69.36	59.97	0.6881
✓	✓	✓	×	70.63	61.53	47.25	70.49	62.52	0.6927
✓	✓	×	✓	71.91	64.01	49.51	71.93	64.81	0.6980
✓	✓	✓	✓	**73.34**	**66.87**	**53.33**	**72.73**	**66.40**	**0.7051**

and the spatio-temporal encoding blocks, as seen in Table 2. The table shows that the model's performance improves in the presence of the four components, with the best results achieved when the hybrid attention mechanisms, along with spatial and temporal encoding, are used. It achieves the highest overall accuracy (OA) of 73.34%, average accuracy (AA) of 66.87%, and Kappa coefficient of 53.33%. This configuration also yields the best precision (PRE) at 72.73% and sensitivity (SEN) at 66.40%, demonstrating the model's superior ability to capture complex dependencies and interactions in the data. These results highlight the individual and combined effectiveness of the components in improving the model's accuracy in classifying melanoma in hyperspectral images.

D. Discussion on Potential Real-world Application

For seamless integration, the method must be user-friendly, compatible with existing infrastructure, and capable of real-time operation. This includes optimizing computational efficiency and ensuring compatibility with standard medical imaging equipment without extensive upgrades. An intuitive user interface is essential for clinicians to easily interpret results and make informed decisions. Robust data storage and management systems are needed to handle large volumes of hyperspectral data, and integration with electronic health records is essential for streamlined workflows. Training and support for clinical staff are vital for effective use and accurate interpretation of results, ultimately enhancing accuracy in melanoma diagnosis.