WUHAN UNIVERSITY

CVPR SEATTLE, WA JUNE 17-21, 2024

Unmixing before Fusion: A Generalized Paradigm for Multi-Source-based Hyperspectral Image Synthesis

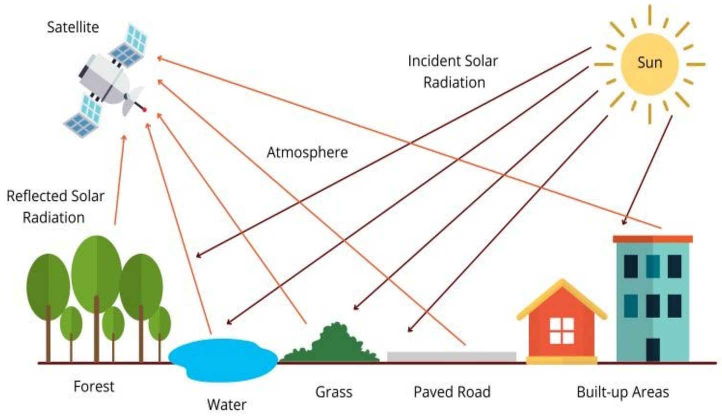
Yang Yu*, Erting Pan*, Xinya Wang, Yuheng Wu, Xiaoguang Mei†, Jiayi Ma

Electronic Information School, Wuhan University, Wuhan, China

1. Introduction

Challenge in hyperspectral image (HSI): Data Scarcity V.S. AI-based application

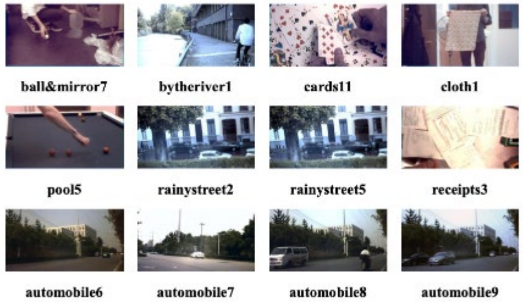
Complex HSI imaging chain



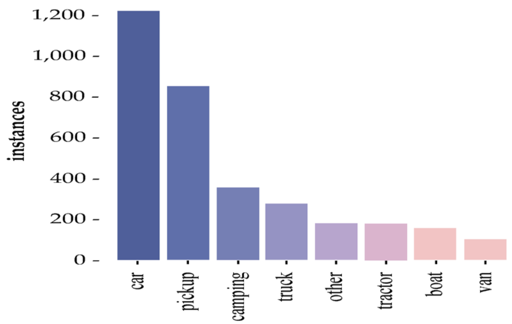
- Unstable process
- Laborious
- Time-intensive

Imperfect Real-world HSIs

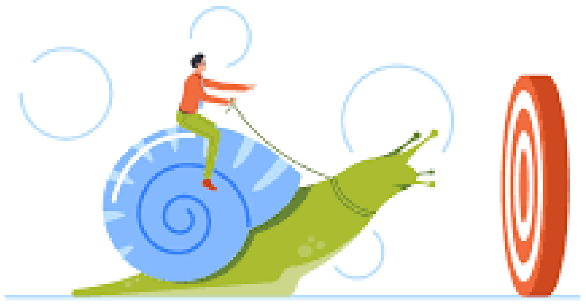
- Limited in quantity & quality



- Severe class imbalance



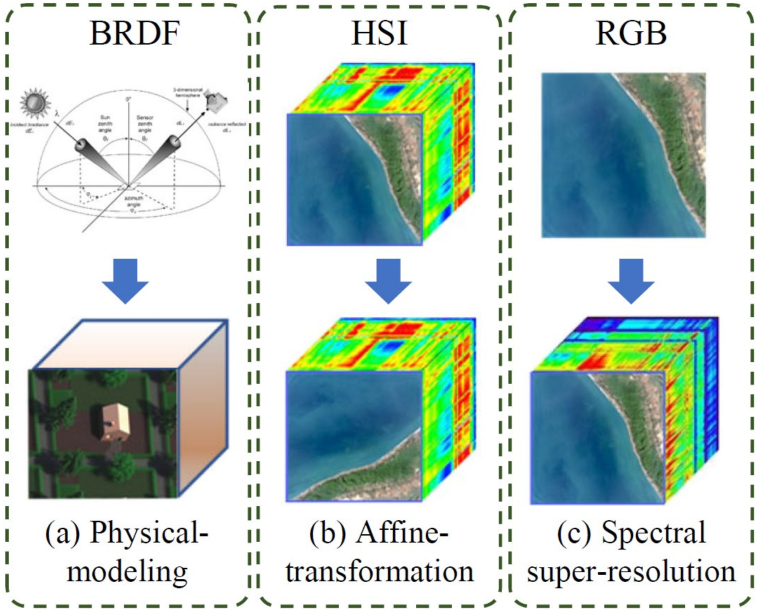
Impede large-scale AI-based applications



Possible solution

HSI Synthesis

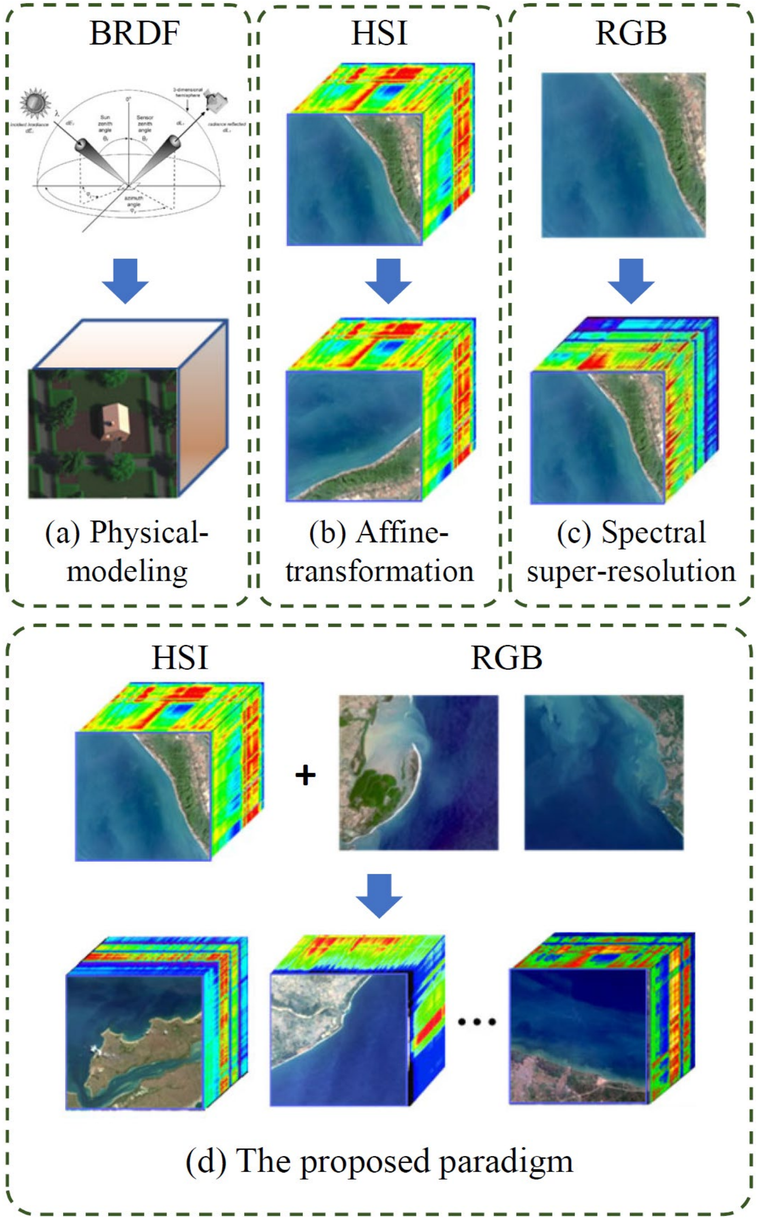
1. Introduction



Prior work related to HSI synthesis

✓ Simulate physical phenomena	✗ Lack of authenticity
✓ Augment the quantity by a series of affine transformations	✗ Restricted diversity
✓ Expand the spectral dimension of existing data	✗ Inability to produce new HSI samples

1. Introduction



Prior work related to HSI synthesis

✓ Simulate physical phenomena	✗ Lack of authenticity
✓ Augment the quantity by a series of affine transformations	✗ Restricted diversity
✓ Expand the spectral dimension of existing data	✗ Inability to produce new HSI samples

The proposed paradigm

- ✓ Customized unmixing across multi-source data to bridge their dimensional gap and release the burden from high-dimensionality
- ✓ Fused unpaired multi-source data (e.g., RGB) to overcome limited sample availability
- ✓ Generalized with various generative AI models for HSI synthesis to generate abundant, diverse, realistic samples

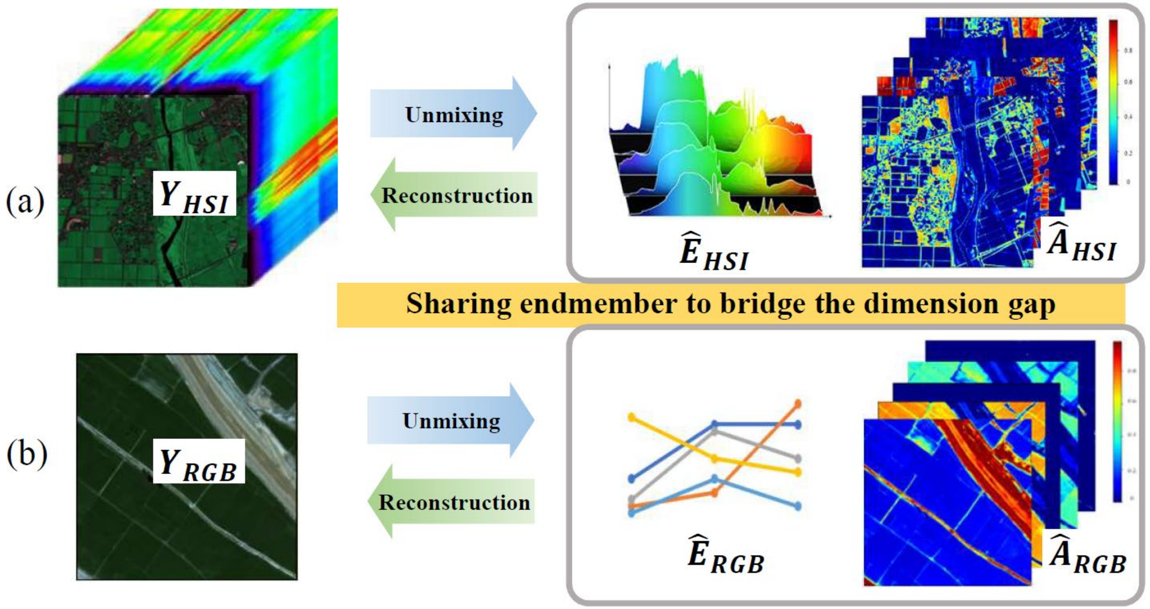
2. Motivation

Why unmixing ?

- Similar scenes share common low-rank features
- Bridging the dimensional gap between RGB images and HSIs in the same low-dimensional abundance space

Why fusion ?

- Data from other source (e.g., RGB) are more easily accessible
- Fusing unpaired multi-source data enables to learn various and realistic spatial distributions



$$Y_{\mathrm{HSI}} = \hat{E}_{\mathrm{HSI}} \cdot \hat{A}_{\mathrm{HSI}} + \epsilon$$

$$Y_{\mathrm{RGB}} = \hat{E}_{\mathrm{RGB}} \cdot \hat{A}_{\mathrm{RGB}} + \epsilon$$

➢ If **endmember sharing**, their abundance maps are in the same low-dimensional space.

3. Method: Unmixing before Fusion

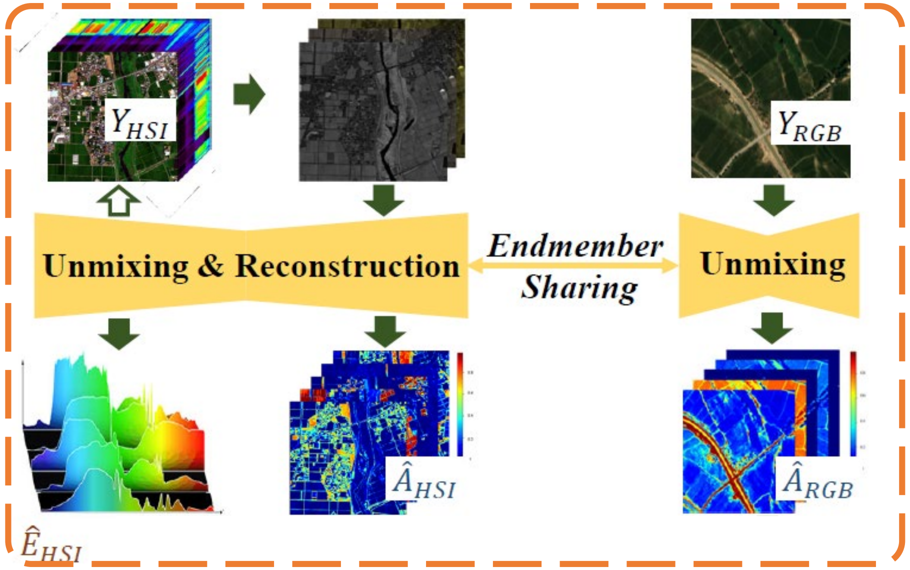
Unmixing across multi-source data

- Training **the unmixing net $\mathcal{U}(\cdot)$** to acquire the endmembers $\hat{E}_{\text{HSI}}$ and abundance maps $\hat{A}_{\text{HSI}}$:

$$\hat{E}_{\text{HSI}}, \hat{A}_{\text{HSI}} = \mathcal{U}(\Psi(Y_{HSI}))$$

- Inferring the abundance maps $\hat{A}_{\text{RGB}}$ from RGB datasets:

$$\hat{A}_{RGB} = \mathcal{U}(Y_{\text{RGB}}; \hat{E}_{\text{HSI}})$$



3. Method: Unmixing before Fusion

Unmixing across multi-source data

- Training **the unmixing net $\mathcal{U}(\cdot)$** to acquire the endmembers $\hat{E}_{\mathrm{HSI}}$ and abundance maps $\hat{A}_{\mathrm{HSI}}$:

$$\hat{E}_{\mathrm{HSI}}, \hat{A}_{\mathrm{HSI}}=\mathcal{U}(\Psi(Y_{HSI}))$$

- Inferring the abundance maps $\hat{A}_{\mathrm{RGB}}$ from RGB datasets:

$$\hat{A}_{RGB}=\mathcal{U}(Y_{\mathrm{RGB}};\hat{E}_{\mathrm{HSI}})$$

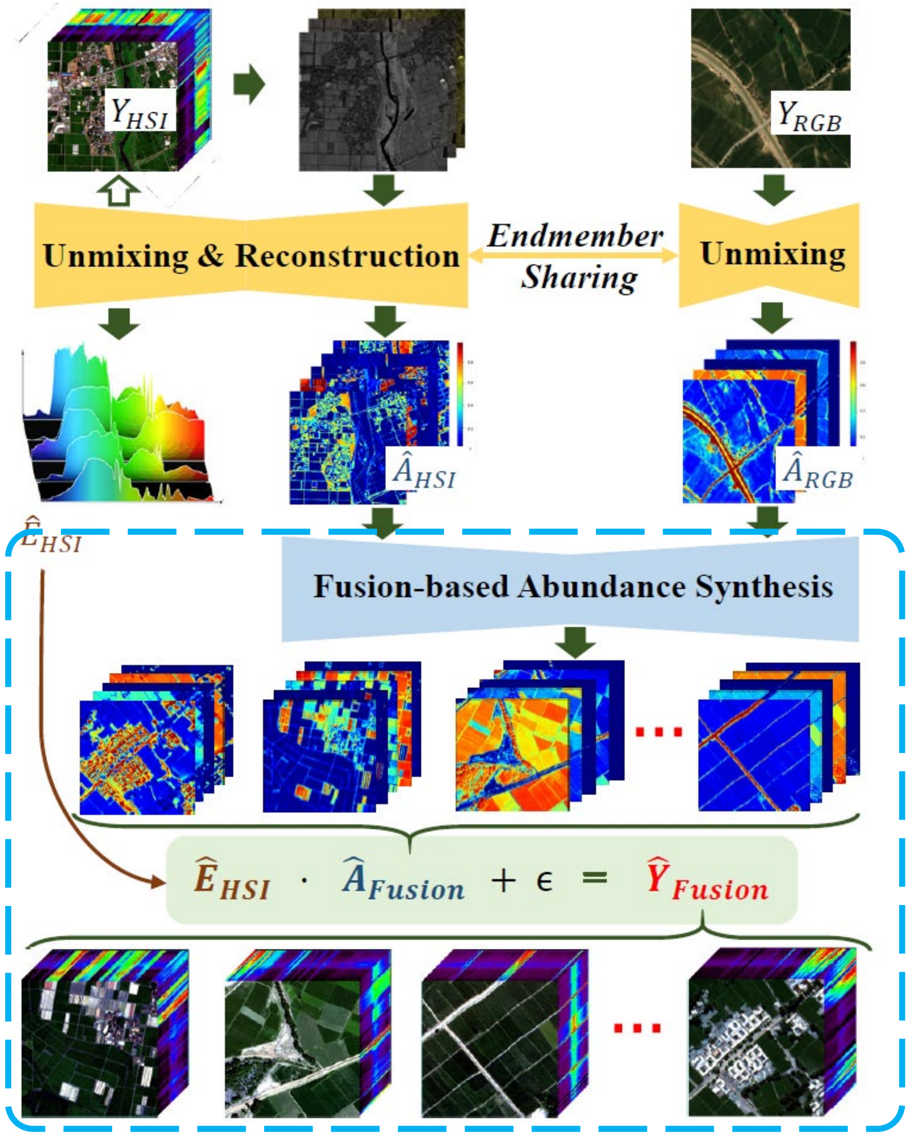
Fusion-based synthesis

- Synthesize abundance utilizing **a generative model $\mathcal{G}(\cdot)$** :

$$\hat{A}_{\mathrm{Fusion}} = \mathcal{G}(\hat{A}_{\mathrm{RGB}}, \hat{A}_{\mathrm{HSI}})$$

- Combining synthetic abundance maps $\hat{A}_{\mathrm{Fusion}}$ and the estimated endmembers $\hat{E}_{\mathrm{HSI}}$:

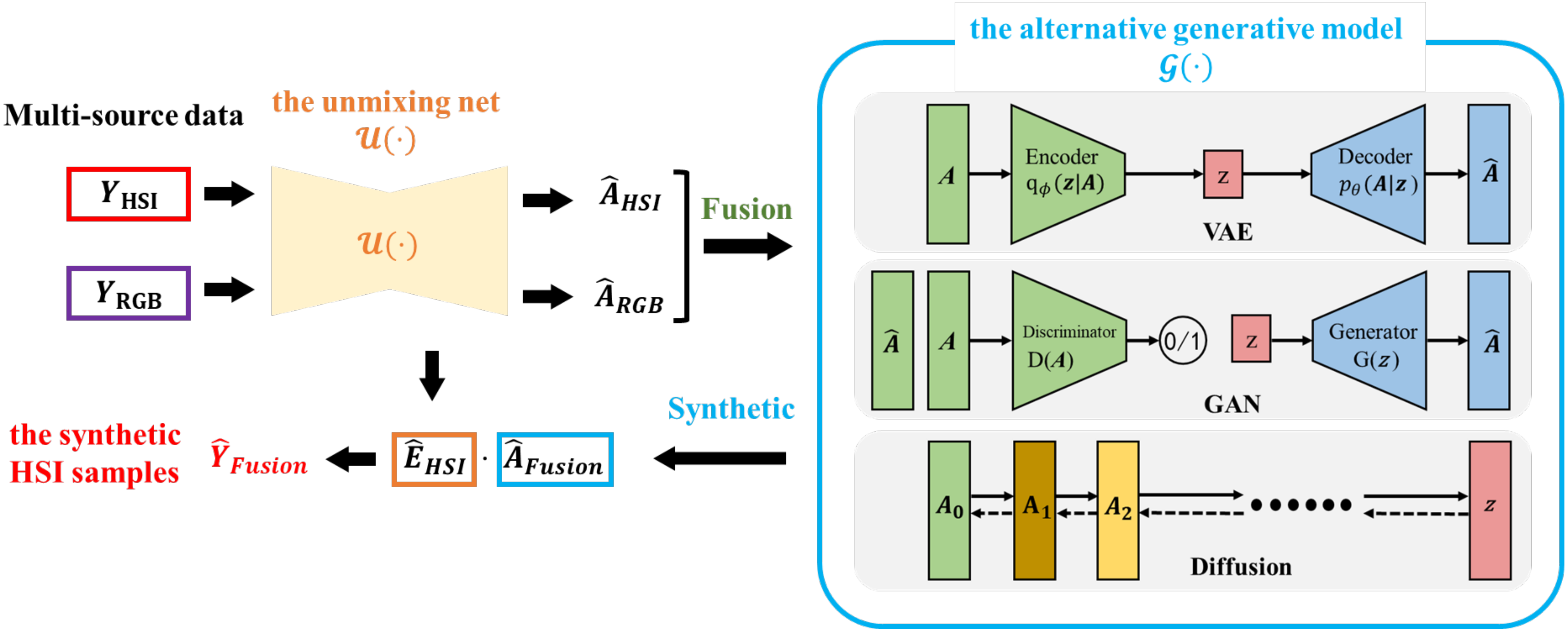
$$\hat{Y}_{\mathrm{Fusion}} = \hat{E}_{\mathrm{HSI}} \cdot \hat{A}_{\mathrm{Fusion}} + \epsilon$$



3. Method: Unmixing before Fusion

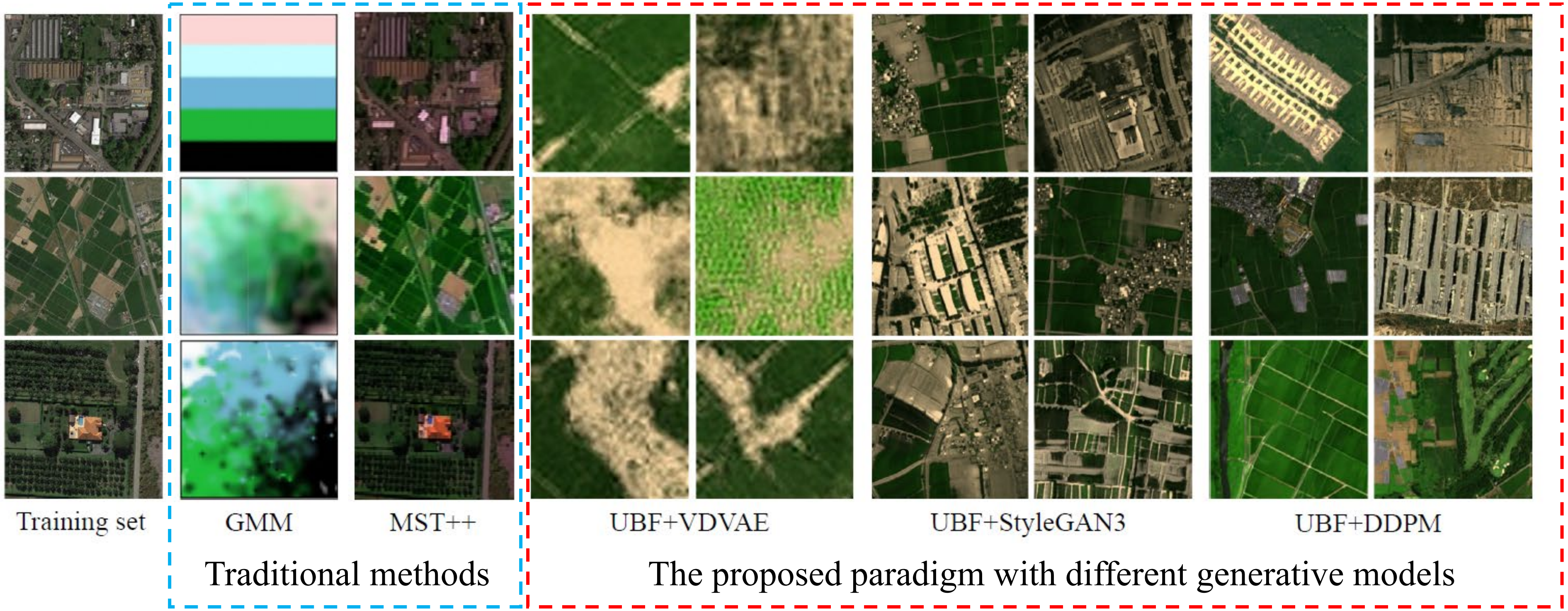
Generalized with various generative AI models

- The UBF paradigm transform high-dimensional data synthesis into low-dimensional abundance synthesis
- The incorporated multi-source data in this paradigm can be unpaired
- Can be generalized with alternative generative models



4. Experiments

Compare with different methods



- The GMM does not mimic the actual distribution of objects, and the MST++ do not generate new data.
- The UBF+VDVAE tends to produce results with blurry textures, low informational content, and poor quality.

  The UBF+StyleGAN3 and UBF+DDPM models can synthesize diverse HSIs with realistic spatial distribution.

Ablation in with/without unmixing

Original HSI

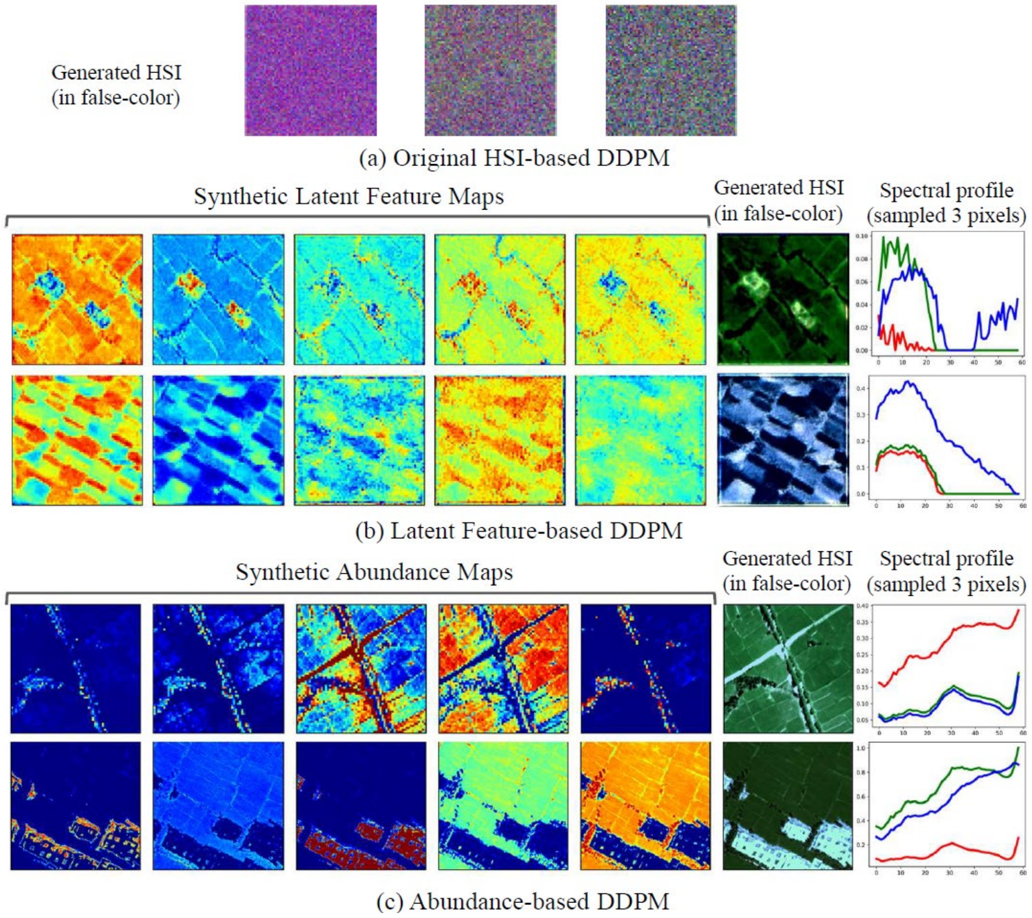
- Meaningless results

Without unmixing

- Unstable latent features
- Unreliable synthetic results with distortion

With unmixing

- clear physical meaning
- High quality synthetic results with reliable spatial-spectral features
- Setting 5 endmembers is most efficient.



Generation space (dimension)	Reconstruction quality				Generation quality		
	RMSE ↓	PSNR ↑	SSIM ↑	SAD ↓	FID ↓	Recall ↑	Precision ↑
HSI Cube(128)	-	-	-	-	-	-	-
Abundance(15)	0.022	36.48	0.944	4.56	49.15	0.11	0.08
Abundance(8)	0.027	35.89	0.938	4.81	15.19	0.36	0.31
Abundance(5)	0.034	34.26	0.931	5.47	8.23	0.50	0.48
Abundance(3)	0.041	30.95	0.927	9.04	8.77	0.50	0.41

4. Experiments

Ablation in without/with fusion

Without fusion

- Monotonous synthetic results.

(a) Synthetic HSI samples with the Chikusei dataset (HSI) alone.

With fusion

- More robust and accurate feature representation.
- Diverse, rich, and reasonable spatial distribution in synthetic HSIs.

(b) Synthetic HSI samples with the Chikusei dataset (HSI) and AID dataset(RGB).

Extension in natural scenario

Synthetic abundance maps:

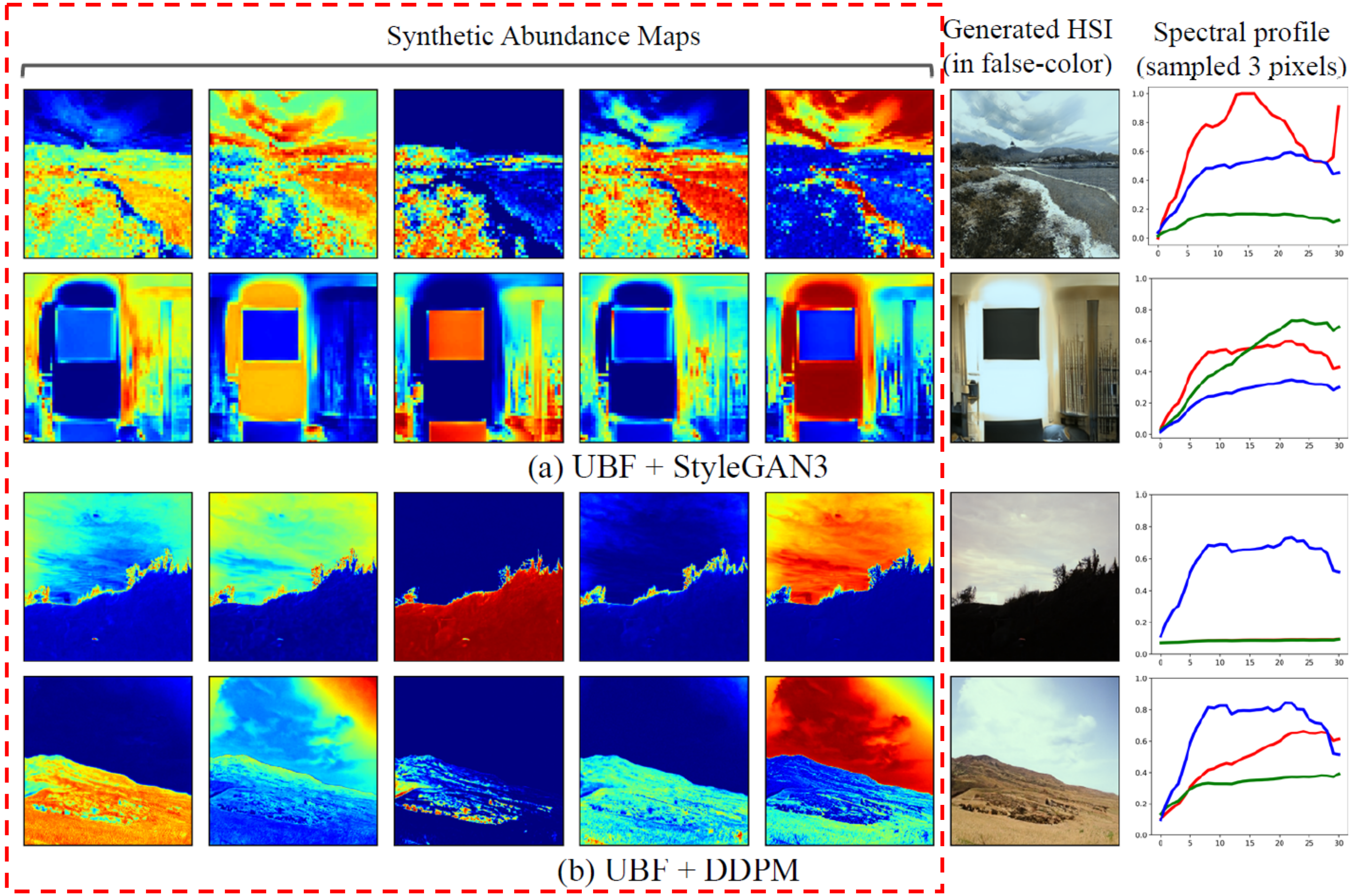
- Accurate spatial distributions

Synthetic spectral profile:

- Distinct spectral features

Synthetic HSIs:

- Convincing visual effects
- Portraying a variety with remarkable fidelity



4. Experiments

Extension in natural scenario

Synthetic abundance maps:

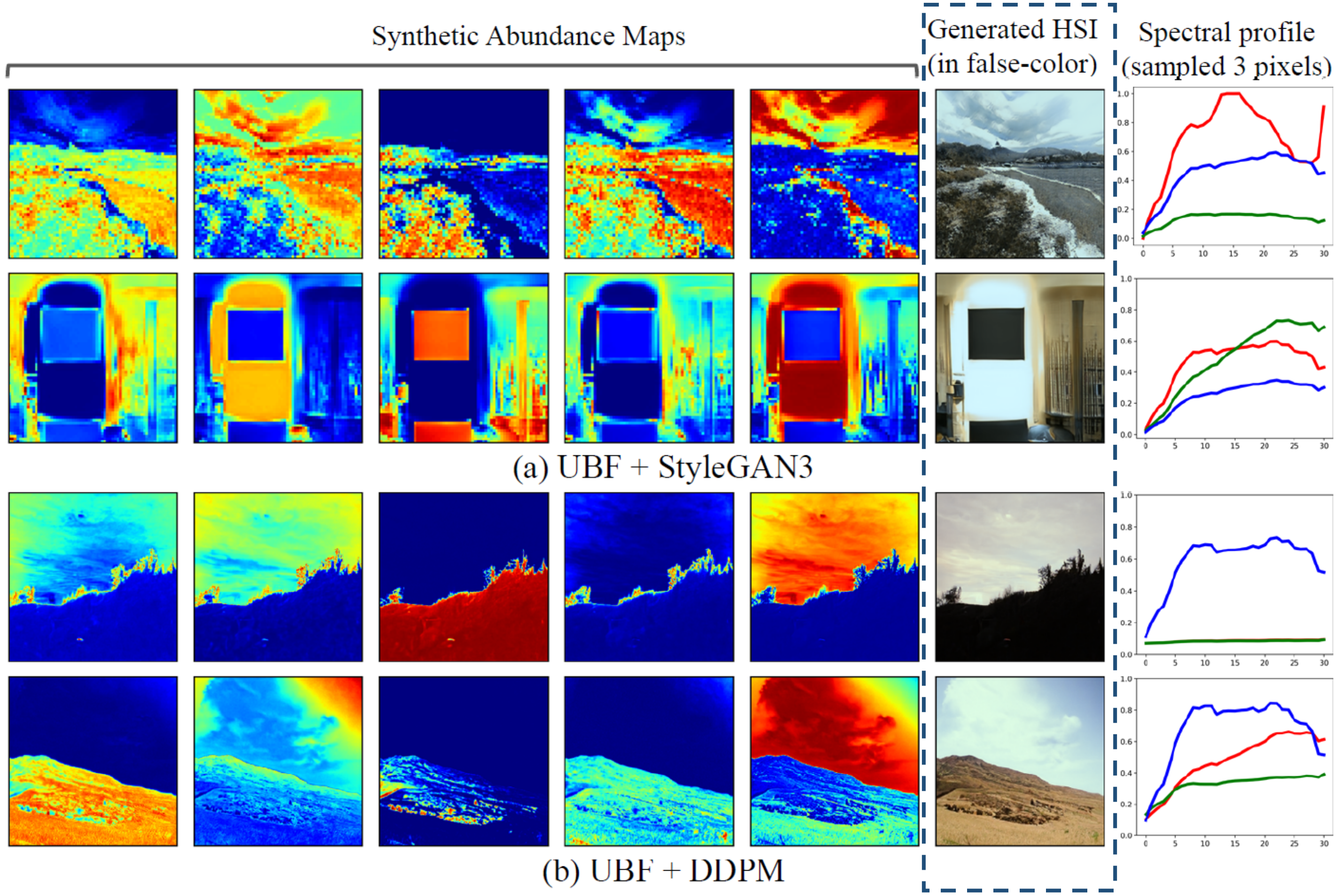
- Accurate spatial distributions

Synthetic spectral profile:

- Distinct spectral features

Synthetic HSIs:

- Convincing visual effects
- Portraying a variety with remarkable fidelity



4. Experiments

Extension in natural scenario

Synthetic abundance maps:

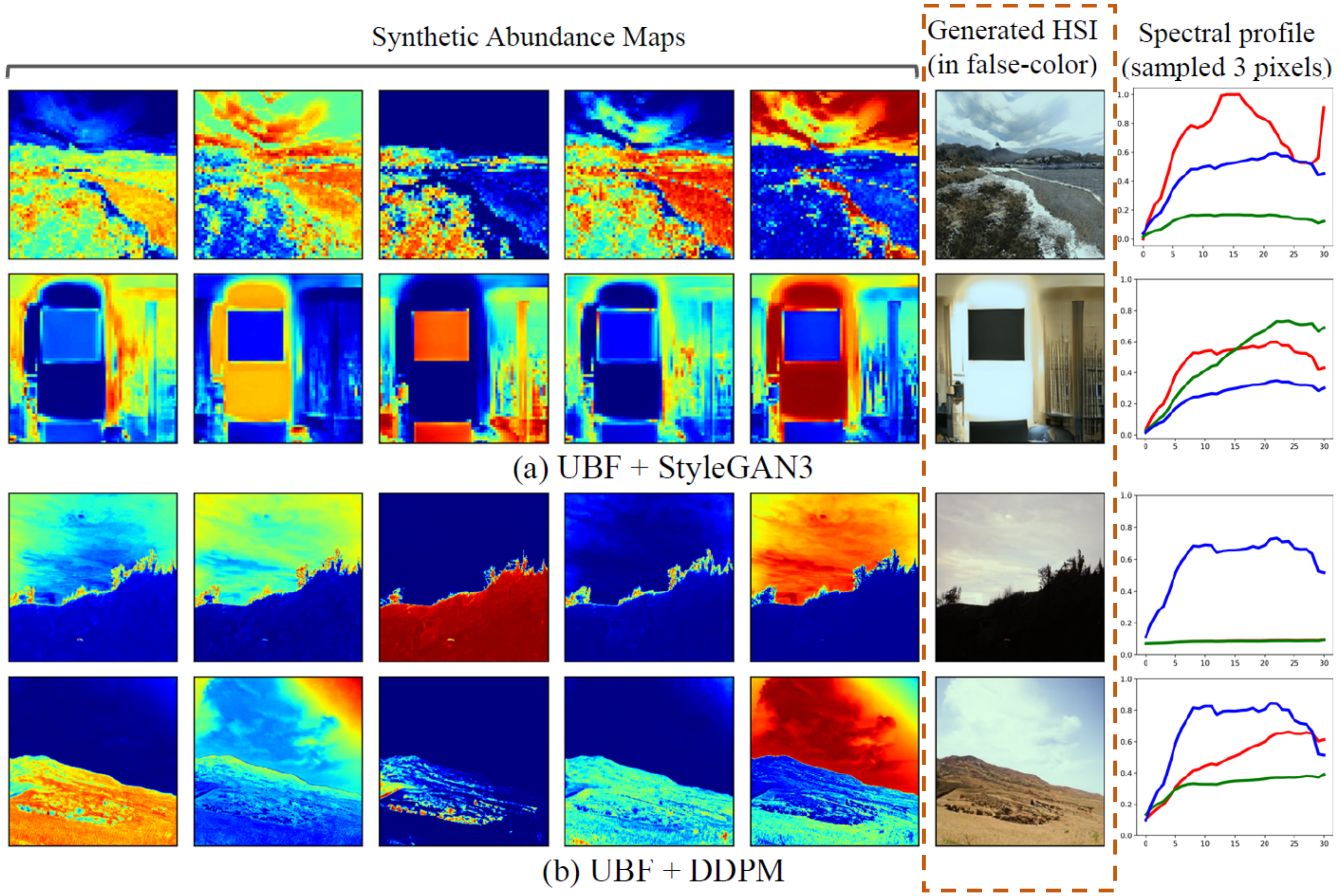
- Accurate spatial distributions

Synthetic spectral profile:

- Distinct spectral features

Synthetic HSIs:

- Convincing visual effects
- Portraying a variety with remarkable fidelity



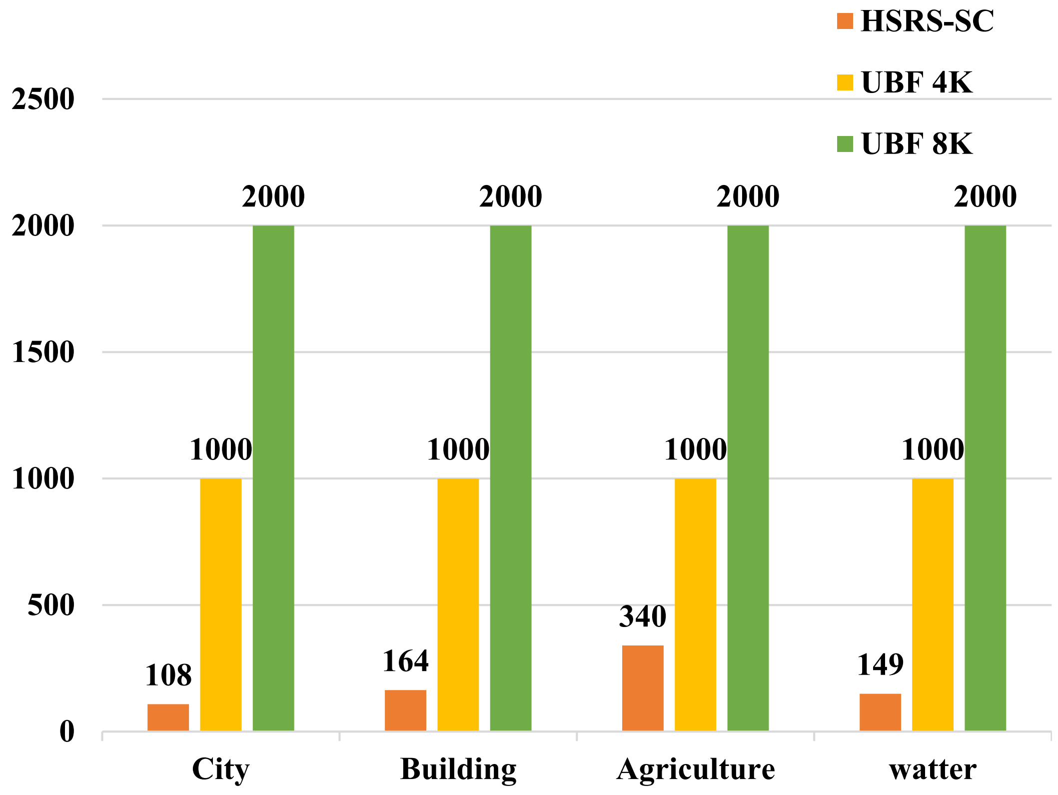
4. Experiments

Extension on downstream tasks

Scene classification

- Mitigating issues like sample scarcity and class imbalance
- Enhancing the diversity and scale of existing limited datasets
- Achieving superior precision opposed to traditional affine transformation.
- Potentially benefit for other downstream tasks.

Augmentation	Training set scale	AlexNet	VGG-16	ResNet-18
✗	761	89.51%	87.30%	37.14%
Affine Trans.	4k	91.11%	88.84%	41.75%
Our UBF	4k	92.70%	93.97%	44.33%
Our UBF	8k	94.29%	94.60%	45.76%



Thank you for listening.

https://hsi-synthesis.github.io/

Acknowledgments: This work was supported by NSFC (U23B200344) and NSFC of Guangdong Province (2023A1515012834).