Fast Hyperspectral Data Augmentation with Spectral-Provenance-Preserving Diffusion

María B. García-Flores[1], Carlos Cañada-Rostro[1], Mercedes E. Paoletti[1], and Juan M. Haut[1]

[1]University of Extremadura, Cáceres, Spain

Hyperspectral image (HSI) synthesis can improve classification when labeled samples are scarce, but diffusion transformers become expensive when fine-grained spatial–spectral tokenization produces thousands of tokens per patch. This work studies how routed-token compression interacts with accelerated diffusion sampling. The evaluated Dynamic Chunking Diffusion Transformer for HSI (DC-DiT-HSI) processes only a fixed token budget through the DiT backbone and reconstructs the dense HSI output afterwards. Because generic reconstruction does not explicitly preserve the correspondence between a routed latent feature and its source spectral band, we introduce Spectral-Provenance-Preserving Chunking (SPPC). SPPC combines spectrally stratified routing with band-preserving dechunking, guaranteeing same-band routed context while keeping the routed-token budget and parameter count unchanged.

Experiments on WHU-Hi HanChuan, Houston 2013, and Xuzhou evaluate downstream augmentation utility, spectral fidelity, distributional structure, and computational cost. In the detailed WHU-Hi analysis at compression factor $N = 8$, SPPC raises full-step DDPM overall accuracy (OA) from 73.82% to 77.58%, while reducing spectral angle from 4.830° to 1.707° and MMD^2 from 0.0634 to 0.0074. The provenance audit confirms complete same-band coverage with zero fallback. Frozen cross-dataset replication also improves full-step OA on all three scenes. Deterministic DDIM-50, however, can strongly degrade augmentation utility under the routed bottleneck, showing that good class-mean spectral similarity alone is insufficient for reliable reduced-step generation.

Stochastic accelerated sampling is substantially more robust. DPM-Solver++ 2M SDE-50 remains within 0.84 percentage points of the corresponding SPPC DDPM-1000 mean across the frozen configurations while requiring $20\times$ fewer denoiser evaluations. On WHU-Hi HanChuan at $N = 8$, SPPC with DPM-Solver++ reaches 77.10% OA, only 0.48 points below full-step SPPC and 3.28 points above the generic full-step architecture, while measured sampling time decreases from 0.930 to 0.0491 s per patch ($\sim 18.9\times$). Token compression provides an additional architecture-level benefit: at $N = 16$, peak training memory decreases from 35.08 to 8.22 GB and DDPM-1000 sampling time from 4.30 to 0.82 s per patch. These results show that spectral provenance and stochastic reduced-step sampling make aggressive token compression practical for fast HSI diffusion augmentation.

Table 1. Frozen five-label pretrained-generator reuse (OA %, mean $\pm$ SD).

Dataset	N	Generic DDPM	SPPC DDPM	SPPC DPM++ SDE-50
WHU-Hi HC	16	76.39 ± 1.12	78.22 ± 0.56	78.03 ± 0.76
Houston 2013	4	78.05 ± 0.31	78.77 ± 0.39	77.93 ± 0.22
Xuzhou	4	87.29 ± 1.82	88.45 ± 0.66	88.94 ± 0.59