PHYSIOLOGY-GUIDED DYNAMIC SPECTRAL SCANNING FOR GENERALIZABLE HYPERSPECTRAL FOREST RECOGNITION

Robertas Damasevicius and Rytis Maskeliunas

Center of Real Time Computer Systems, Kaunas University of Technology, Lithuania

ABSTRACT

We introduce PHYTOSCAN, a physiology-guided state-space architecture that separates physical wavelength coordinates from the computational order in which spectral evidence is aggregated. Its Dynamic Physiological Spectral Scan (DPSS) learns a sample-dependent wavelength traversal regularized by weak radiative-transfer priors and continuous wavelength embeddings, while a crown-aware spatial scan limits recurrent propagation across spectrally dissimilar neighboring crowns. Physics-guided masked pretraining uses a DART model derived forest look-up table before fine-tuning on the 3D3 hyperspectral benchmark. On 3D3 Dataset B, physiologically guided adaptive traversal improves not only clean recognition but more importantly spectral robustness, calibration, and band efficiency at comparable inference cost, supporting the learned computational wavelength order for sensor-resilient forest hyperspectral analysis.

***Index Terms*—** hyperspectral imaging, forest remote sensing, spectral scanning, tree species classification

1. INTRODUCTION

Imaging spectroscopy resolves subtle biochemical and structural differences that are largely invisible to broadband optical sensors. In forests, the differences arise from pigments, leaf water, dry matter, proteins and canopy architecture, whose absorption and scattering effects occupy distinct and often widely separated spectral regions [1]. This makes hyperspectral imagery (HSI) well suited to species identification, stress assessment and functional-trait retrieval, but the same high dimensionality creates strong collinearity, sensor dependence and a pronounced risk of site-specific overfitting. HSI classifiers have progressed from 3-D/2-D convolution [2] to spectral transformers [3, 4]. Selective state-space models (SSMs) offer linear sequence scaling [5], and SpectralMamba and SS-Mamba showed that such models can efficiently process HSI spectral–spatial sequences [6, 7]. PHYTOSCAN adopts this principle at the wavelength level, where the ordering problem has an interpretable physical meaning. Most such approaches inherit a hidden assumption: that after tokenization, the model processes a sequence whose traversal is fixed by wavelength or by a hand-designed spatial raster. In a forest canopy, however, the most informative evidence for one decision may couple chlorophyll-sensitive visible wavelengths with the red edge and SWIR water/dry-matter features.

The PROSPECT models leaf reflectance and transmittance through structure and absorbing constituents; PROSPECT-PRO explicitly separates proteins from other carbon-based constituents and demonstrates informative SWIR domains [1]. DART model extends radiative transfer to heterogeneous 3-D vegetation scenes [8]. Li et al. challenged fixed image scans with Dynamic Adaptive Scan (DAS), which learns data-dependent scan orders while retaining SSM efficiency [9]. We transfer this idea to imaging spectroscopy but make a domain-specific change: *the adaptive scan is regularized by vegetation physiology and absolute wavelength*.

Our contributions: (1) DPSS, which learns a sample-dependent wavelength traversal under soft physiological anchors; (2) a crown-aware spatial scan that discourages state propagation across spectrally dissimilar neighboring crowns; and (3) continuous wavelength conditioning together with a sensor-response consistency regularizer.

2. METHODOLOGY

PHYTOSCAN is illustrated in Fig. 1. The model learns a physiology-regularized spectral order and a crown-aware spatial order before selective state-space aggregation.

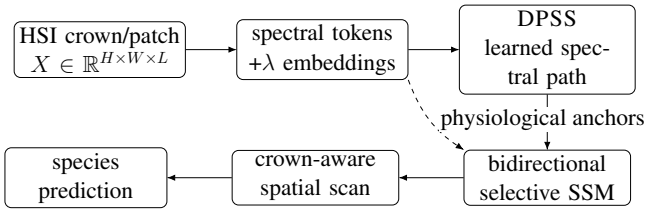


Fig. 1. Concept of PHYTOSCAN.

Let a crown-centered hyperspectral patch be $X \in \mathbb{R}^{H\times W\times L}$ with L bands and band centers $\Lambda = \{\lambda_b\}_{b=1}^{L}$. The goal is to estimate a class y (tree species) while remaining robust to sensor spectral response, missing channels and geographic shift. For spatial location p and band group $\mathcal{N}(b) = \{b-4, \ldots, b+4\}$ (clipped at the spectrum edges), a shallow

spectral encoder produces

$$\mathbf{z}_{p,b} = \phi_\theta\big(X_{p,\mathcal{N}(b)},\ \mathbf{e}(\lambda_b),\ \Delta_\lambda X_{p,b}\big)\,, \tag{1}$$

where Δ_λ is a first spectral derivative and the continuous wavelength embedding uses $K = 8$ log-spaced frequencies,

$$\mathbf{e}(\lambda) = \big[\sin(\omega_1\lambda), \cos(\omega_1\lambda), \ldots, \sin(\omega_K\lambda), \cos(\omega_K\lambda)\big]. \tag{2}$$

We write $\bar{\mathbf{z}}_b = \mathrm{median}_{p\in\mathcal{V}}\, \mathbf{z}_{p,b}$ for the band-b token pooled over the vegetation mask $\mathcal{V}$, and $\mathbf{c} = \mathrm{pool}(\{\bar{\mathbf{z}}_b\}_b)$ for the patch context vector. Each spectral token receives a score

$$q_b = f_q([\bar{\mathbf{z}}_b, \mathbf{c}, \mathbf{e}(\lambda_b)]), \tag{3}$$

and a weak physiological prior built from $R = 5$ Gaussians,

$$a_b = \sum_{r=1}^{R} \kappa_r \exp\left[-\frac{(\lambda_b - \mu_r)^2}{2\varsigma_r^2}\right]. \tag{4}$$

The centers are initialized at $\mu = \{550, 715, 900, 1450, 2100\}$ nm with widths $\varsigma = \{40, 25, 80, 60, 80\}$ nm, i.e. broad regions associated with visible pigments, red edge, NIR structure, leaf water and dry matter/protein absorption. Amplitudes κ_r, widths ς_r, centers μ_r and the global strength γ are trainable. The scan logits are $s_b = q_b + \gamma a_b$. A differentiable soft permutation $P_\tau(\mathbf{s})$ [10] approximates sorting during training, while inference uses the hard permutation $\pi = \mathrm{argsort}(\mathbf{s})$: $\widetilde{Z} = P_\tau(\mathbf{s})Z, \quad \widetilde{\mathbf{z}}_t = \mathbf{z}_{\pi(t)}$. The temperature is annealed geometrically from $\tau = 1.0$ to $\tau = 0.1$, after which soft and hard orderings agree on $> 98\%$ of adjacent pairs. To prevent unstable arbitrary jumps, we penalize wavelength displacement in proportion to latent *dissimilarity*:

$$\mathcal{L}_{\mathrm{path}} = \frac{1}{L-1}\sum_{t=1}^{L-1}\big(1-\rho_{\pi(t),\pi(t+1)}\big)\,|\lambda_{\pi(t+1)} - \lambda_{\pi(t)}|, \tag{5}$$

where $\rho_{ij} = \exp(-\|\bar{\mathbf{z}}_i - \bar{\mathbf{z}}_j\|_2^2/\sigma_z^2)$. A large jump is inexpensive only when the two endpoints carry closely related latent evidence, and expensive otherwise; monotonic wavelength order remains the zero-cost solution. A continuous SSM $\dot{\mathbf{h}}(t) = A\mathbf{h}(t) + B\mathbf{x}(t)$ is discretized with an input-dependent step Δ_t as

$$\bar{A}_t = \exp(\Delta_t A), \tag{6}$$

$$\bar{B}_t = (\Delta_t A)^{-1}(\exp(\Delta_t A) - I)\,\Delta_t B_t, \tag{7}$$

$$\mathbf{h}_t = \bar{A}_t\mathbf{h}_{t-1} + \bar{B}_t\widetilde{\mathbf{z}}_t, \qquad \mathbf{o}_t = C_t\mathbf{h}_t + D\widetilde{\mathbf{z}}_t, \tag{8}$$

with Δ_t, B_t, C_t predicted from the current token as in selective SSMs [5]. Because wavelength has no causal direction, we apply Eq. (8) forward and backward and fuse

$$\mathbf{u}_t = W_u[\mathbf{o}_t^{\rightarrow} \| \mathbf{o}_t^{\leftarrow}]. \tag{9}$$

A fixed image raster can propagate state across neighboring crowns. For spatial tokens i, j at coordinates $\mathbf{r}_i, \mathbf{r}_j$, define

$$A_{ij} = \exp\left(-\frac{\|\mathbf{r}_i - \mathbf{r}_j\|_2^2}{\sigma_r^2} - \frac{d_{\mathrm{SAM}}^2(\mathbf{x}_i, \mathbf{x}_j)}{\sigma_s^2}\right), \tag{10}$$

where d_{SAM} is the spectral angle. A second DAS-like ranker learns a path maximizing local affinity while remaining data dependent. For sensor grids Λ_A and Λ_B, let $\mathcal{R}_{A\to B}$ resample spectra by known response functions. We impose

$$\mathcal{L}_{\mathrm{sensor}} = D_{\mathrm{KL}}[p_\theta(y|X, \Lambda_A)\,\|\,p_\theta(y|\mathcal{R}_{A\to B}(X), \Lambda_B)]\,, \tag{11}$$

where Λ_B is drawn from synthetic band-subsampling and calibration-shift perturbations of the AisaFENIX grid. This acts as a within-sensor regularizer; it is not a cross-sensor transfer experiment. Pretraining uses contiguous spectral masking $\mathcal{M}$ on DART signatures, with reconstruction constrained in both reflectance and derivative domains:

$$\mathcal{L}_{\mathrm{rec}} = \frac{1}{|\mathcal{M}|}\sum_{\lambda\in\mathcal{M}}|\widehat{R}_\lambda - R_\lambda| + \eta_{\mathrm{der}}\sum_{\lambda}|\Delta_\lambda\widehat{R}_\lambda - \Delta_\lambda R_\lambda|. \tag{12}$$

The final objective is

$$\mathcal{L} = \mathcal{L}_{\mathrm{cls}} + \eta_{\mathrm{rec}}\mathcal{L}_{\mathrm{rec}} + \eta_{\mathrm{sen}}\mathcal{L}_{\mathrm{sensor}} + \eta_{\mathrm{path}}\mathcal{L}_{\mathrm{path}}. \tag{13}$$

Algorithm 1 summarizes training.

Algorithm 1 Physics-guided training of PHYTOSCAN

Require: HSI/LUT sample X, wavelength grid Λ, target y

1: Form wavelength-aware tokens using Eq. (1)
2: Predict scan scores and physiological bias (Eq. (4))
3: Compute soft permutation P_τ and reordered sequence
4: Apply bidirectional selective SSM (Eq. (8))
5: Compute affinities (Eq. (10)) and adaptive crown scan
6: Sample contiguous masked bands and a resampled view
7: Predict class and masked reflectance
8: Update θ by minimizing Eq. (13)
9: Anneal permutation temperature τ toward hard scanning

3. DATA AND EXPERIMENTAL PROTOCOL

We used DART LUT dataset containing 3,456,000 Central-European broadleaf forest signatures over 450–2300 nm at 2-nm spacing (926 bands), with leaf chlorophyll, carotenoid, water, dry matter, structure, LAI, canopy cover and sun geometry varied systematically [11]. We also used 3D3 Dataset B, containing SPECIM AisaFENIX imagery from 381.35–2502.38 nm, 364 bands, 60-cm hyperspectral ground sampling distance, with four annotated crown classes: three species (*Pinus nigra*, *Pinus strobus*, *Quercus rubra*) and a pooled *Other* class [12]. To avoid crown leakage, train/ validation/ test sets are separated by crown and, where coordinates permit, by spatial block. All reported numbers use protocol (ii) below unless stated otherwise: (i) the 3D3 author-recommended split [12], used only for the reference comparison in the first row of Table 1; (ii) 5-fold spatial-block cross-validation, one seed per fold, from which all means and

standard deviations are computed; (iii) DART synthetic pretraining followed by fine-tuning on 3D3 training crowns, which is the default initialization for PHYTOSCAN and is ablated in Table 6.

Implementation. AdamW, initial learning rate 3×10^{-4} with cosine decay, weight decay 0.05, batch size 64, 120 fine-tuning epochs, patch size 15×15, SSM state width 16 and 4 bidirectional blocks. Loss weights $\eta_{\text{rec}} = 1.0$, $\eta_{\text{der}} = 0.5$, $\eta_{\text{sen}} = 0.2$, $\eta_{\text{path}} = 0.01$, selected on the validation folds. ECE uses 15 equal-width confidence bins. Pretraining on the DART LUT takes approximately 18 h and fine-tuning 2.5 h on one NVIDIA A100-40GB; latencies in Table 1 are single-crown forward passes at batch size 1 on the same device.

Baselines are RBF-SVM, Random Forest, HybridSN [2], SpectralFormer [3], SSFTT [4], SpectralMamba [6], SS-Mamba [7], a VMamba-style fixed raster scan [13], and a DAMamba-style unrestricted DAS [9].

Robustness is quantified under random missing-band rates $r \in \{0.1, 0.25, 0.5\}$; contiguous outages of 20/40/80 bands; wavelength calibration offsets $\delta\lambda \in \{\pm 2, \pm 5, \pm 10\}$ nm; additive Gaussian radiometric noise at SNR 20/30/40 dB; and resampling to reduced grids of 16/32/64 selected bands. Degradation is

$$\Delta_{\text{rob}}(c) = M_{\text{clean}} - M_c, \tag{14}$$

where M is macro-F1 and c a corruption condition.

4. RESULTS

Table 1 reports the main comparison. PHYTOSCAN (full) OLOGY-Gattains the best Macro-F1, the smallest degradation under 50% band loss and ± 10-nm calibration shift, and the lowest ECE, at a latency within 8% of the fixed-order SpectralMamba and below that of the transformer baselines. The Random-Forest reference in the first row is the macro-average we computed over the four per-class F1 values reported for Dataset B under the published protocol (six repetitions, 70% training ratio, 1000 trees) [12]; it uses protocol (i) and is not directly comparable to the remaining rows.

Table 2 and Fig. 2 detail behaviour under spectral corruption and radiometric noise; Table 3 isolates component contributions; Table 5 tests whether the learned traversal identifies compact wavelength subsets; Tables 4 and 6 report anchor-strength sensitivity and the effect of initialization.

On 3D3-B, PHYTOSCAN (full) reaches Macro-F1 0.931 ± 0.007 and ECE 0.044, improving over the unrestricted learned scan of DAMamba (0.909 ± 0.009, ECE 0.064) and the fixed-order SSM of SpectralMamba (0.898 ± 0.010, ECE 0.069). The advantage widens under corruption: at 50% band loss PHYTOSCAN degrades by 0.043 versus 0.074 for DAMamba and 0.091 for SpectralMamba, and a ± 10-nm shift costs 0.027 versus 0.046 and 0.052 respectively; all differences are significant under the paired Wilcoxon test over folds.

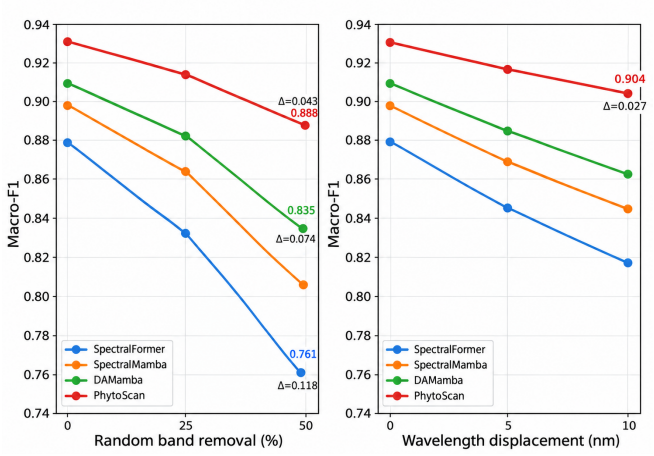


Fig. 2. Robustness to spectral perturbations on 3D3-B. Left: Macro-F1 as an increasing fraction of spectral bands is randomly removed. Right: Macro-F1 under wavelength-calibration displacement.

The ablation attributes $+0.012$ Macro-F1 to the learned spectral scan, $+0.006$ to the continuous λ embedding, $+0.002$ to the crown-aware spatial scan, $+0.001$ to derivative reconstruction, $+0.002$ to DART pretraining and $+0.013$ to the physiological anchors. Anchors are also the largest single robustness contributor, cutting Δ_{50} from 0.061 to 0.043, ahead of the learned scan itself (0.089 to 0.074). OLOGY-G Anchor strength is non-monotonic: $\gamma = 0.50$ maximizes clean F1 and minimizes degradation under both missing bands and wavelength shift, while $\gamma = 0$ and $\gamma = 1.00$ lower both, indicating that the prior helps only as a soft bias and hurts when it dominates the content score. The physiology-regularized traversal also concentrates discriminative bands: with 16 bands PHYTOSCAN retains 0.842 Macro-F1 against 0.784 for DAMamba and 0.742 for SpectralFormer, and reaches 0.916 with 64 of 364 bands, i.e. within 0.015 of full-spectrum performance at 18% of the channels. Varying only the initialization contributes the final $0.925 \rightarrow 0.931$ increment and reduces Δ_{50} from 0.056 to 0.043, with the derivative-domain constraint and the sensor-consistency term each adding a small share. These gaps indicate that the learned scan yields a more transferable spectral sequence rather than a marginal reordering.

5. CONCLUSION

The main value of PHYTOSCAN is greater stability when the spectral representation is incomplete or shifted. Learning the computational spectral order is most useful when it is constrained by physical wavelength information rather than treated as an unrestricted permutation. Replacing fixed wavelength traversal by the complete physiology-guided model raises Macro-F1 from 0.895 to 0.931 while reducing the

Table 1. Comparison on 3D3 Dataset B under 5-fold spatial-block cross-validation. Macro-F1 is mean±std over folds. Lower is better for Δ and ECE. †Published reference under protocol (i); macro-average of the four reported per-class F1 values [12].

Method	Params (M)	Latency (ms)	Macro-F1	$\Delta_{50\%\,\mathrm{band}}$	$\Delta_{\pm 10\,\mathrm{nm}}$	ECE
Random Forest, published † [12]	–	–	0.841	–	–	–
RBF-SVM	–	3.1	0.812 ± 0.016	0.164	0.098	0.121
Random Forest	–	4.0	0.838 ± 0.013	0.151	0.089	0.113
HybridSN [2]	5.1	9.4	0.861 ± 0.012	0.132	0.074	0.094
SSFTT [4]	4.2	11.8	0.872 ± 0.011	0.124	0.068	0.088
SpectralFormer [3]	3.8	12.6	0.879 ± 0.011	0.118	0.061	0.082
VMamba-style fixed raster scan [13]	3.4	8.2	0.886 ± 0.010	0.104	0.057	0.077
SS-Mamba [7]	3.6	8.7	0.893 ± 0.010	0.097	0.054	0.072
SpectralMamba [6]	3.3	8.5	0.898 ± 0.010	0.091	0.052	0.069
DAMamba-style unrestricted DAS [9]	3.9	9.1	0.909 ± 0.009	0.074	0.046	0.064
PHYTOSCAN ($\gamma = 0$, no anchors)	4.0	9.2	0.918 ± 0.008	0.061	0.040	0.056
PHYTOSCAN (full)	4.0	9.2	$\mathbf{0.931 \pm 0.007}$	**0.043**	**0.027**	**0.044**

Table 2. Robustness on 3D3-B. Entries are Macro-F1; parentheses give degradation from the clean condition per Eq. (14). Higher Macro-F1 and smaller degradation are better.

Method	Clean	25% missing	50% missing	40-band outage	±5 nm	±10 nm	SNR 20 dB
SpectralFormer	0.879	0.833 (−0.046)	0.761 (−0.118)	0.794 (−0.085)	0.846 (−0.033)	0.818 (−0.061)	0.775 (−0.104)
SS-MambaOLOGY-G	0.893	0.856 (−0.037)	0.796 (−0.097)	0.821 (−0.072)	0.864 (−0.029)	0.839 (−0.054)	0.801 (−0.092)
SpectralMamba	0.898	0.864 (−0.034)	0.807 (−0.091)	0.829 (−0.069)	0.870 (−0.028)	0.846 (−0.052)	0.811 (−0.087)
DAMamba scan	0.909	0.882 (−0.027)	0.835 (−0.074)	0.850 (−0.059)	0.885 (−0.024)	0.863 (−0.046)	0.834 (−0.075)
PHYTOSCAN ($\gamma = 0$)	0.918	0.896 (−0.022)	0.857 (−0.061)	0.868 (−0.050)	0.898 (−0.020)	0.878 (−0.040)	0.855 (−0.063)
PHYTOSCAN (full)	**0.931**	**0.914** (−0.017)	**0.888** (−0.043)	**0.897** (−0.034)	**0.917** (−0.014)	**0.904** (−0.027)	**0.881** (−0.050)

Table 3. Ablation on 3D3-B, cumulative from top to bottom. Δ_{50} is the Macro-F1 degradation with 50% randomly missing bands. Mean±std over five spatial-block folds. The penultimate row is the PHYTOSCAN ($\gamma = 0$) configuration of Tables 1–5; anchors are added last so the two coincide exactly.

Configuration	Macro-F1	Δ_{50}	ECE
Fixed wavelength scan	0.895 ± 0.010	0.089	0.071
+ learned spectral scan	0.907 ± 0.009	0.074	0.064
+ continuous λ embedding	0.913 ± 0.009	0.067	0.060
+ crown-aware spatial scan	0.915 ± 0.009	0.069	0.058
+ derivative reconstruction	0.916 ± 0.008	0.065	0.057
+ DART pretraining (= γ=0)	0.918 ± 0.008	0.061	0.056
+ PHYTOSCAN (full)	$\mathbf{0.931 \pm 0.007}$	**0.043**	**0.044**

Table 4. Sensitivity to physiological-anchor strength γ, evaluated on clean spectra, under 50% band loss and under ±10-nm wavelength perturbation. The $\gamma = 0$ row is the no-anchor configuration of Table 1.

γ	Clean F1	50% band F1	±10 nm F1
0.00	0.918	0.857	0.878
0.10	0.923	0.869	0.888
0.25	0.928	0.881	0.898
0.50	**0.931**	**0.888**	**0.904**
0.75	0.929	0.885	0.901
1.00	0.924	0.877	0.894

50%-band degradation from 0.089 to 0.043, at essentially unchanged parameter count and inference latency relative to fixed-order SSM baselines. The learned traversal concentrates discriminative information into a comparatively small set of physiologically informative regions rather than relying on dense redundant sampling. Absolute wavelength remains necessary for physical interpretation, but the most useful recurrent path can connect spectrally distant regions associated with pigments, canopy structure, water and dry matter.

Limitations are substantial. The DART LUT represents temperate Central-European broadleaf conditions, healthy mature trees and canopy cover of at least 60%, so synthetic pretraining may be biased [11]. 3D3 does not provide all modalities at every site and the hyperspectral species experiment is confined to Dataset B with four crown classes [12]; claims of biome generalization would require independent HSI datasets that we did not have. Most importantly, the sensor-consistency term is validated only as a within-sensor regularizer against synthetic band-subsampling and calibration shifts: we make no claim of demonstrated transfer between physically different instruments, and a genuine sensor-to-sensor evaluation remains an open question.

Table 5. Performance after reducing the spectral input to the top-K bands. Bands for PHYTOSCAN are ranked by expected early-path occupancy; baseline bands are selected by validation-set importance.

Method	16 bands	32 bands	64 bands	Full
SpectralFormer	0.742	0.801	0.846	0.879
SS-Mamba	0.759	0.821	0.865	0.893
SpectralMamba	0.768	0.829	0.871	0.898
DAMamba scan	0.784	0.847	0.886	0.909
PHYTOSCAN ($\gamma = 0$)	0.808	0.868	0.898	0.918
PHYTOSCAN (full)	**0.842**	**0.892**	**0.916**	**0.931**

Table 6. Effect of initialization on the full PHYTOSCAN. FT is supervised fine-tuning on 3D3 training crowns.

Initialization	Macro-F1	Bal. Acc.	ECE	Δ_{50}
Random	0.925	0.919	0.053	0.056
Masked pretraining, 3D3 only	0.927	0.921	0.050	0.051
DART reflectance reconstruction	0.928	0.922	0.048	0.048
DART reflectance + derivative	0.930	0.924	0.046	0.045
DART + sensor consistency + FT	**0.931**	**0.925**	**0.044**	**0.043**

Declarations

Data Availability Access to DART forest LUT dataset and Dataset B of 3D3 benchmark is provided in [11, 12].

AI Use Writefull for Overleaf and Grammarly were used to enhance the language, clarity, and grammar.

Funding Horizon Europe Framework Programme (HORIZON), call Teaming for Excellence (HORIZON-WIDERA-2022-ACCESS-01-two-stage) - Creation of the centre of excellence in smart forestry "Forest 4.0" No. 101059985.

6. REFERENCES

[1] Jean-Baptiste Féret, Katja Berger, Florian de Boissieu, and Zbyněk Malenovský, "PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constituents," *Remote Sensing of Environment*, vol. 252, pp. 112173, 2021.

[2] Swalpa Kumar Roy, Gopal Krishna, Shiv Ram Dubey, and Bidyut B. Chaudhuri, "HybridSN: Exploring 3-d–2-d cnn feature hierarchy for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 2, pp. 277–281, 2020.

[3] Danfeng Hong, Zhu Han, Jing Yao, Lianru Gao, Bing Zhang, Antonio Plaza, and Jocelyn Chanussot, "Spectralformer: Rethinking hyperspectral image classification with transformers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022, Art. no. 5518615.

[4] Le Sun, Guangrui Zhao, Yuhui Zheng, and Zebin Wu, "Spectral–spatial feature tokenization transformer for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022.

[5] Albert Gu and Tri Dao, "Mamba: Linear-time sequence modeling with selective state spaces," in *Proceedings of the First Conference on Language Modeling*, 2024.

[6] Jing Yao, Danfeng Hong, Chenyu Li, and Jocelyn Chanussot, "Spectralmamba: Efficient mamba for hyperspectral image classification," *arXiv preprint arXiv:2404.08489*, 2024.

[7] Lingbo Huang, Yushi Chen, and Xin He, "Spectral-spatial mamba for hyperspectral image classification," *arXiv preprint arXiv:2404.18401*, 2024.

[8] J.-P. Gastellu-Etchegorry, V. Demarez, V. Pinel, and F. Zagolski, "Modeling radiative transfer in heterogeneous 3-d vegetation canopies," *Remote Sensing of Environment*, vol. 58, no. 2, pp. 131–156, 1996.

[9] Tanzhe Li, Caoshuo Li, Jiayi Lyu, Hongjuan Pei, Baochang Zhang, Taisong Jin, and Rongrong Ji, "DAMamba: Vision state space model with dynamic adaptive scan," *arXiv preprint arXiv:2502.12627*, 2025.

[10] Sebastian Prillo and Julian Martin Eisenschlos, "SoftSort: A continuous relaxation for the argsort operator," in *Proceedings of the 37th International Conference on Machine Learning*, Hal Daumé III and Aarti Singh, Eds. 2020, vol. 119 of *Proceedings of Machine Learning Research*, pp. 7793–7802, PMLR.

[11] Tomáš Hanousek, Terézia Slanináková, Tomáš Rebok, and Růžena Janoutová, "High spatial and spectral resolution dataset of hyperspectral look-up tables for 3.5 million traits and structural combinations of central european temperate broadleaf forests," *Data in Brief*, vol. 57, pp. 111105, 2024.

[12] Lauris Bocaux, Narges Takhtkeshha, Zhenyu Ma, and Fabio Remondino, "A multi-sensor multi-resolution dataset to support forest inventory methods," in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2025, vol. XLVIII-1/W6-2025, pp. 33–40.

[13] Yue Liu, Yunjie Tian, Yuzhong Zhao, Hongtian Yu, Lingxi Xie, Yaowei Wang, Qixiang Ye, Jianbin Jiao, and Yunfan Liu, "VMamba: Visual state space model," in *Advances in Neural Information Processing Systems*, 2024, vol. 37, pp. 103031–103063.