Supplementary Data

Table S1. Summary of hyperparameter configurations and their tested values across three machine learning models and algorithms. Bold fonts highlight the optimal values.

Machine learning algorithms	Hyperparameter settings	Hyperparameter values
Random forest	Mtry	16, 32, **46**, 64
Kfold	5	
Support vector machine	C	1e-03, 1e-02, 1e-01, 1e+00, **1e+01**, 1e+02
	Sigma	1e-06, 1e-05, 1e-04, 1e-03, 1e-02, **1e-01**
Kfold	5	
Self-attention network	Epochs	10, 50, **150**, 200
	Number heads	8, 16, 32, **64**, 124
	Bach size	8, **16**, 32, 64, 124
	Dropout	**0.1**, 0.01, 0.001
	Hidden layer size	8, 16, **32**, 64, 124
	Learning rate	**0.1**, 0.01, 0.001
Kfold	5	
Genetic algorithm	Population size	100, 1000, **2500**
	Chromosome number	2001
	Number generations	100, **1000**, 2000
	Elite percentage	**0.1**, 0.01, 0.001
	Mutation rate	**0.1**, 0.01, 0.001
Kfold	5	

Table S2. Classification performance of the three trained models using selected wavelength data from various wavelength selection methods, optimized with the best-performing pre-processing method.

Models	**Pre – processing methods**	**Datasets**	**Performance measured as Area under curve (AUC)**
Random forest	First derivative	Self-attention networks	0.94
		Random forest	0.92
		Genetic algorithm	0.88
		Full dataset	0.96
Self-attention networks	Support Vector Machine	Self-attention networks	0.97
		Random forest	0.92
		Genetic algorithm	0.94
		Full dataset	0.90
Support vector machine	First derivative	Self-attention networks	0.93
		Random forest	0.90
		Genetic algorithm	0.87
		Full dataset	0.94

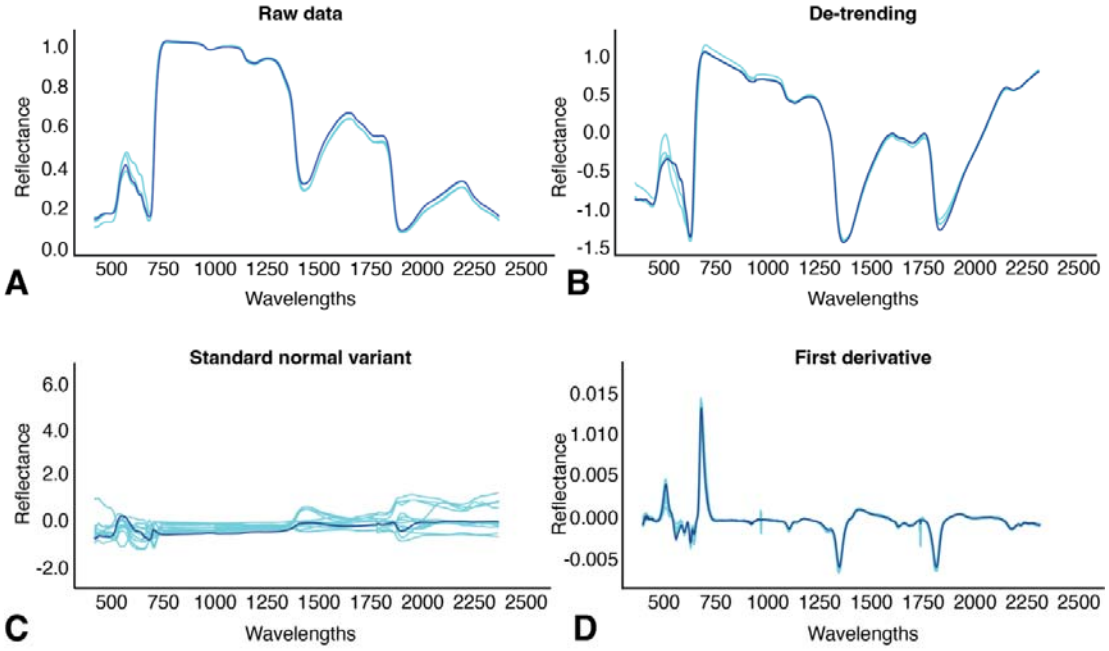

Fig. S1. Data structure of *Eucalyptus benthamii* trees inoculated with *Phytophthora alticola* following different pre-processing methods. The dark blue line represents the average trend of all inoculated trees post-pre-processing, while the light blue lines indicate deviations from the average for specific samples. A) Raw reflectance data, B) wavelength reflectance after de-trending, C) wavelength reflectance after applying standard normal variant normalization, and D) wavelength reflectance after the first derivative values have been calculated.

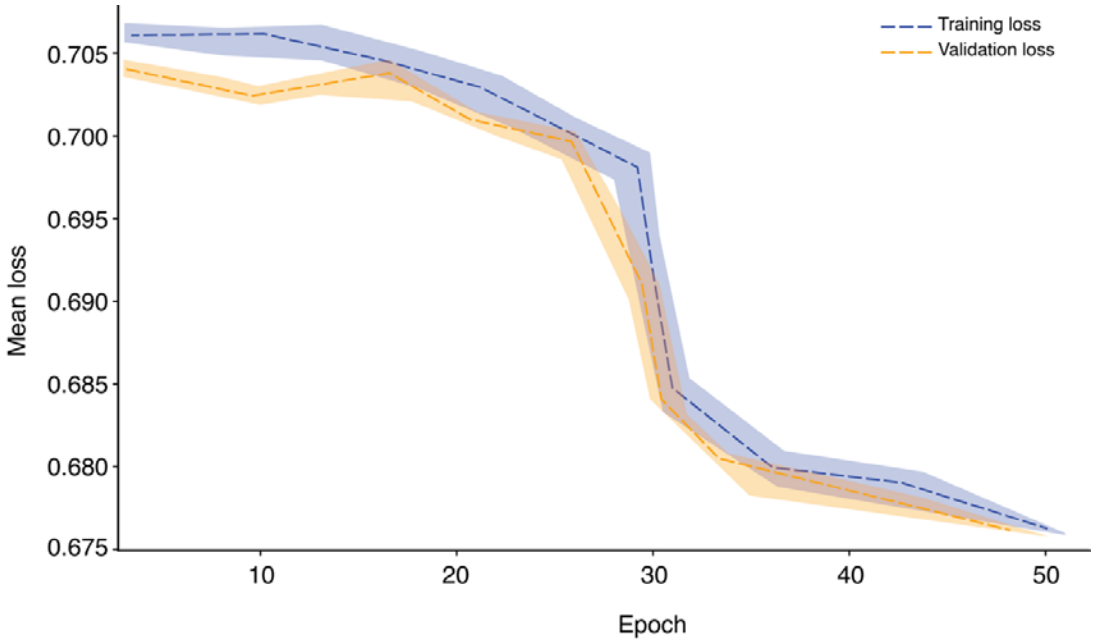

Fig. S2: Learning curve showing average training and validation loss (±1 SD) per epoch across 5-fold cross-validation for SAN model. The graph illustrates the progressive decrease in mean loss over 50 training epochs, indicating improved model performance with continued training. The sharp decline in loss during the initial 20–30 epochs suggest rapid learning, followed by a more gradual convergence toward minimal loss, indicating that the model was able to effectively learn the underlying patterns in the data without signs of overfitting.