Label Semantics for Robust Hyperspectral Image Classification

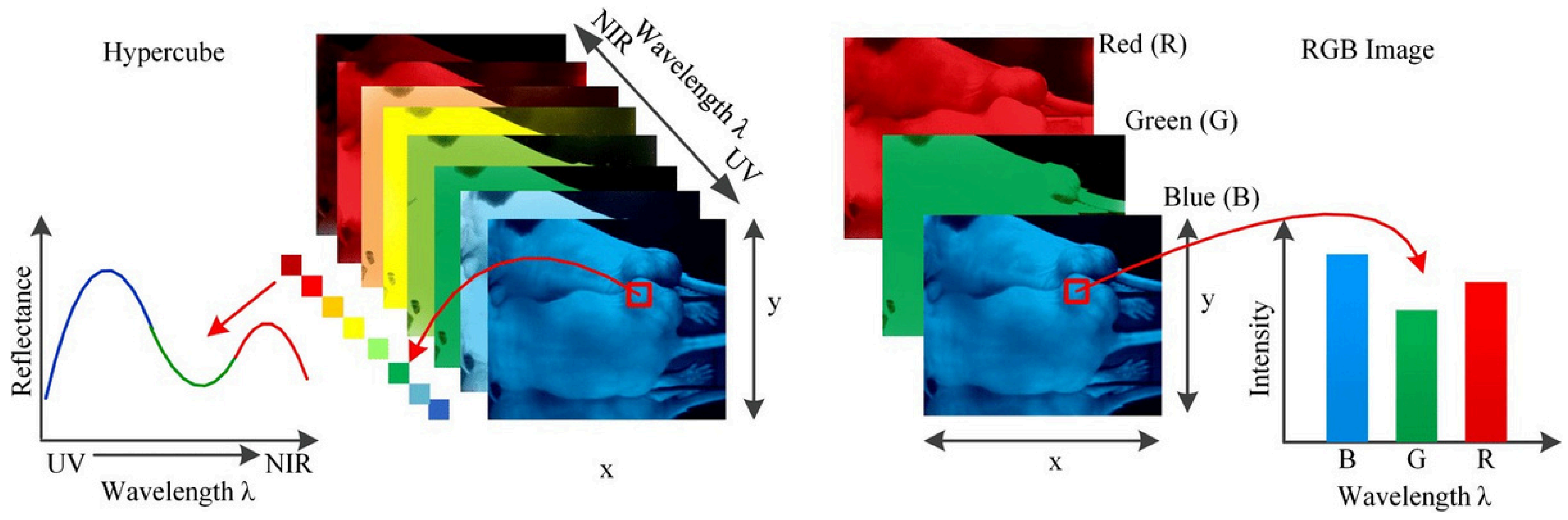
Rafin Hassan **Zarin Tasnim Roshni** **Rafiqul Bari** **Alimul Islam**

North South University
Department of Electrical and Computer Engineering
School of Engineering & Physical Sciences

Supervised by - Dr. Shafin Rahman

Introduction

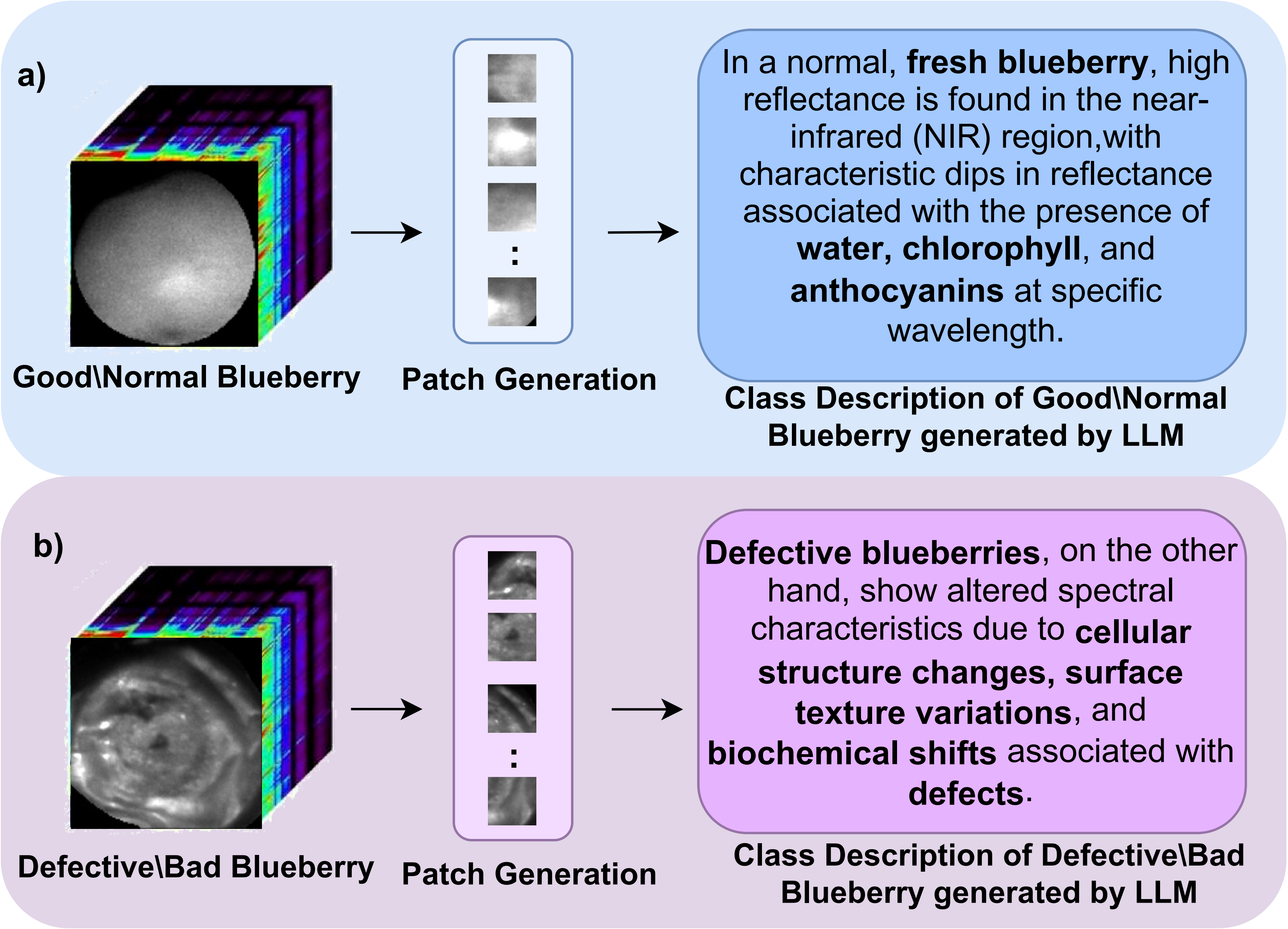
- **What is HSI:**
  - Hyperspectral Imaging (HSI) captures hundreds of spectral bands, far beyond the 3 channels in RGB images.
- **Why it matters:**
  - Applications: agriculture (fruit quality), environmental monitoring, materials science



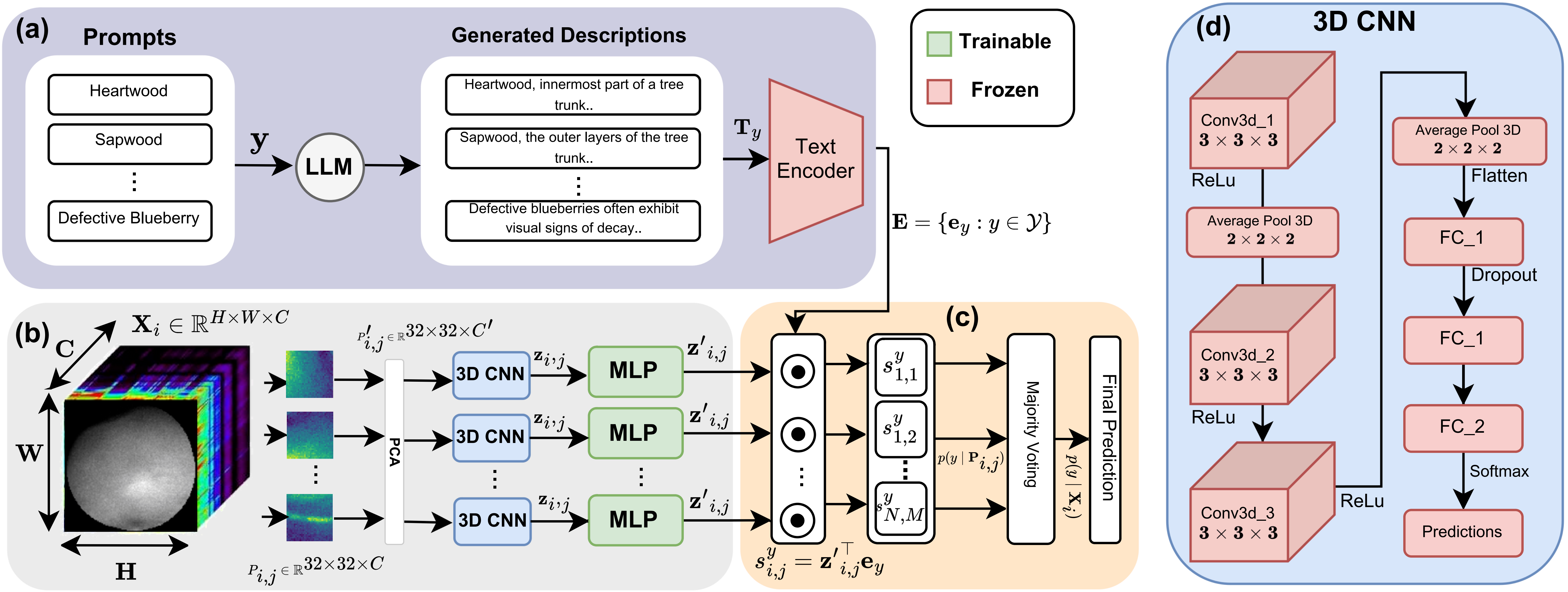
The Challenge

- **Problem Statement:**
  - HSI classification suffers from high-dimensional spectra + scarce labels → overfitting, poor generalization.
  - Traditional models rely only on spectral-spatial features — missing the “why.”

- **Motivation:**
  - Use **LLM-generated class descriptions** to provide semantic guidance, helping the model better understand class-specific characteristics.



Our Solution - S3FN

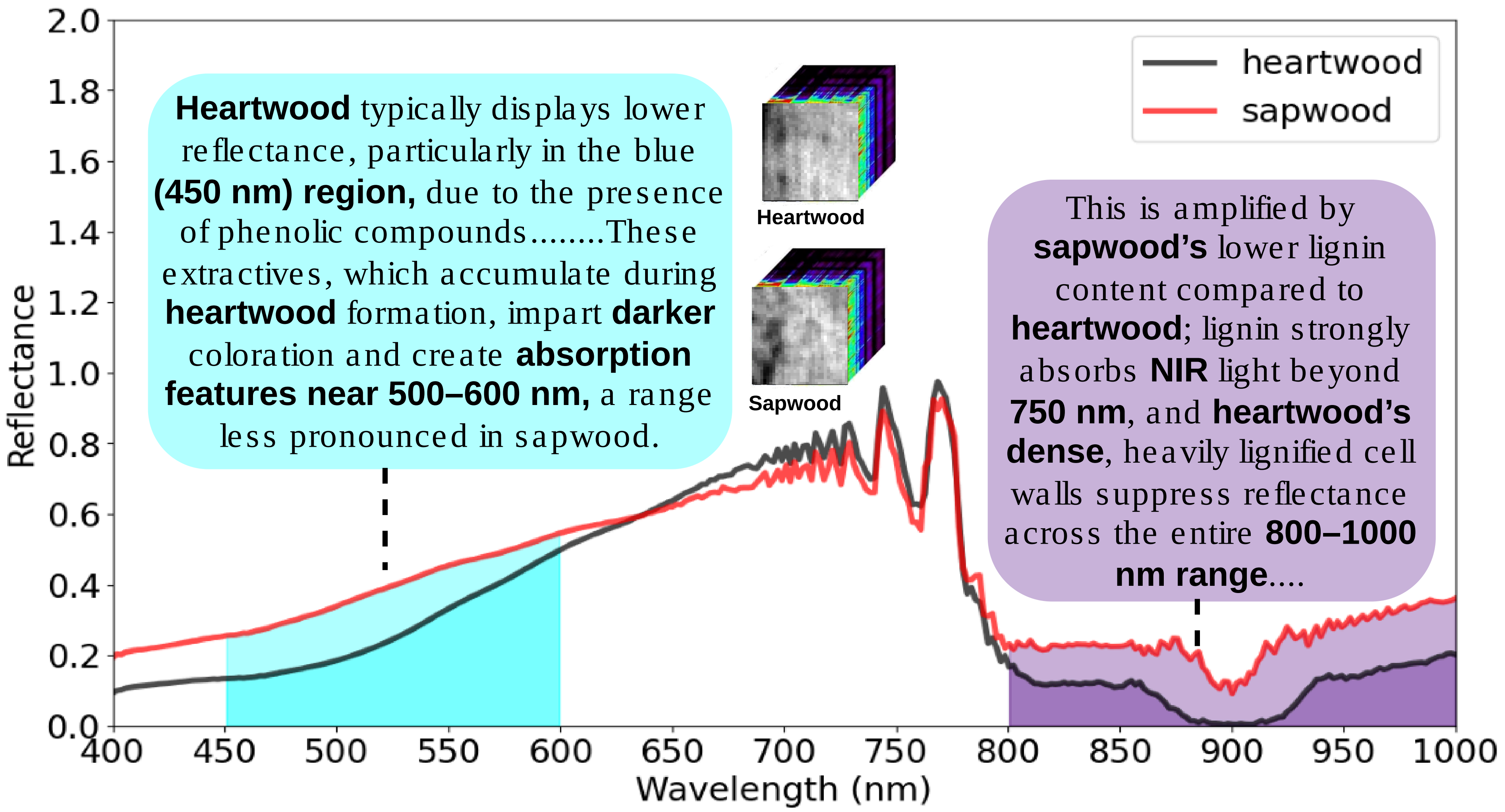


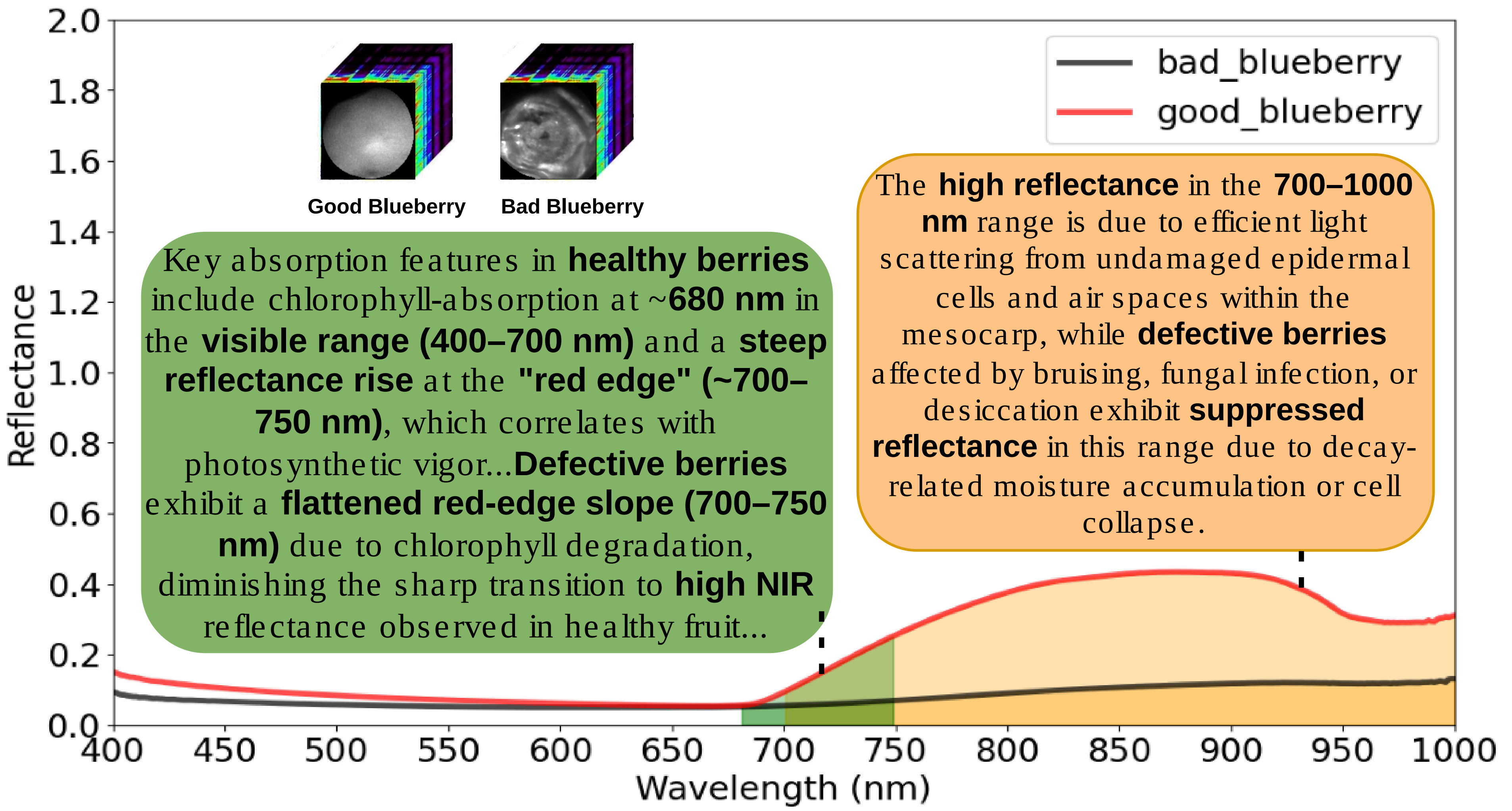
- The **Semantic Spectral-Spatial Fusion Network (S3FN)** combines both image features and label semantics to classify hyperspectral images.
- (**a**) For each class, an LLM generates a detailed text description → encoded into a semantic vector using a text encoder (like BERT/RoBERTa).
- (**b**) Each hyperspectral image is split into patches, then PCA is applied to reduce spectral dimensions → projected via an MLP to match the semantic space.
- (**c**) Feature vectors are compared with class embeddings using a similarity score (dot product), followed by softmax and majority voting to predict the final class..
- (**d**) The 3D CNN architecture which extracts spectral-spatial features from each patch

Contributions:

- We used **LLM-generated class descriptions as textual guidance** for generating semantic label embeddings that describe the Hyperspectral Images.
- A fusion architecture, **Semantic Spectral-Spatial Fusion Network (S3FN)**, aligns the HSI image with the embeddings of the semantic label/description.
- The model is evaluated on three diverse HSI datasets: Hyperspectral Images for Wood Recognition, HyperspectralBlueberries, and DeepHS-Fruit.

Why this Works





- **Mean spectral reflectance** curves for **wood** *(heartwood & sapwood)* and **blueberries** *(healthy/good & bad/defective)* illustrate key **absorption** and **reflectance** features.

- **LLM-generated semantic descriptions** capture **class-specific spectral characteristics**.

- This highlights how the generated textual descriptions align with spectral characteristics.

Results/Experiments

HyperspectralBlueberries

Model	PR	Recall	F1	ACC
SVM	92.0	92.0	92.0	91.7
KNN	76.0	75.0	75.0	75.2
Random forest	77.0	76.0	77.0	76.4
Neural Network	86.0	86.0	86.0	85.8
Decision Tree	81.0	80.0	80.0	80.0
LDA	90.8	78.6	-	85.3
RLDA	**97.7**	93.7	-	95.7
RLDA&LDA	96.5	**96.7**	-	**96.6**
S3FN (Ours)	86.0	86.0	86.0	86.4

DeepHS-Fruit

Model	Ripeness (C1)		Ripeness (C2)	
	Avocado	Kiwi	Avocado	Kiwi
SVM	57.1	55.5	80.0	56.5
KNN	57.1	33.3	86.6	65.2
Random Forest	53.3	57.8	87.0	61.7
Neural Network	80.0	78.9	93.5	76.5
Decision Tree	80.0	42.1	70.9	53.1
ResNet-18	44.4	60.0	66.7	33.3
AlexNet	33.3	33.3	33.3	33.3
HS-CNN	44.4	66.7	33.3	33.3
S3FN (Ours)	**66.7**	**70.4**	**47.1**	**44.8**

Hyperspectral Wood

Model	PR	Recall	F1	ACC
SVM	89.0	89.0	89.0	88.6
KNN	84.0	84.0	84.0	84.0
Random Forest	82.0	82.0	82.0	81.1
Neural Network	88.0	87.0	87.0	87.1
Decision Tree	59.0	58.0	58.0	58.3
Cifar10Net	-	-	-	93.9
S3FN (Ours)	**95.0**	**95.0**	**95.0**	**94.7**

Ablation Studies

Hyperspectral Wood		DeepHS-Fruit			
Configuration	Acc.	Ripeness C1 (Acc.)		Ripeness C2 (Acc.)	
		Avocado	Kiwi	Avocado	Kiwi
Standalone 3D CNN	88.0	55.6	66.7	47.0	43.0
No description	93.9	57.5	68.0	47.0	44.5
S3FN (Ours)	**94.7**	**66.7**	**70.4**	**47.1**	**44.8**

Conclusion

- **LLM-generated descriptions** are used as **semantic guidance**, providing **rich contextual information** to assist hyperspectral image classification.
- Unlike traditional methods that rely only on **spectral-spatial features** or **static embeddings**, this approach dynamically adjusts to improve performance.
- The proposed method, **S3FN**, strengthens the **alignment between features and labels**, making classification more accurate and meaningful.
- By going **beyond predefined labels** and using **contextual embeddings**, it achieves **better classification results** compared to simpler models.

Thank You For Your Attention!