Bachelor′s Thesis

Enhancing U-Net for PCB Segmentation Using Hyperspectral Imaging in E-waste Recycling

School of Mechanical and Control Engineering

Handong Global University

Lee Chankeun

Enhancing U-Net for PCB Segmentation Using Hyperspectral Imaging in E-waste Recycling

A Bachelor′s Thesis

Submitted to the School of

Mechanical and Control Engineering of

Handong Global University

Lee Chankeun

June 2025

This certifies that the bachelor′s thesis is approved.

_____

Thesis Advisor: Ph.D. Young-keun Kim

_____

The Dean of Faculty: Ph.D. Chong-Sun Lee

School of Mechanical and Control Engineering

Handong Global University

June 2025

Extended Abstract

Enhancing U-Net for PCB Segmentation Using Hyperspectral Imaging in E-waste Recycling

Background & Motivation. Fast, material-aware sorting is critical on electronic-waste (E-waste) recycling lines. RGB cameras, limited to three broad bands, cannot separate visually similar but chemically distinct PCB components. Hyperspectral imaging (HSI) overcomes this by providing a 640 × 640 × 214 cube whose narrow-band signatures reveal, for example, copper's absorption edge (≈ 540 nm) and polymer overtones (≈ 880 nm). We base our work on the PCB-Vision dataset (53 PCB scenes × 3 capture conditions) acquired with a Specim FX10 line-scan camera.

Processing the full 214-band cube with a standard U-Net is computationally heavy. Earlier studies mitigated this either by cropping each frame into small patches or by compressing spectra with principal-component analysis (PCA)—both strategies risk losing global context or minority-class variance. We propose a Spectral Channel Reduction Block that keeps every pixel of every full-resolution frame.

Methods. Spectrum Channel Reduction Block(SCRB). Two sequential 1 × 1 convolutions compress the spectrum **214 → 128 → 3** channels; the weights are learned jointly with the U-Net so the model selects task-relevant wavelengths.

Experimental setup

- No patching: every model receives the resized 640 × 640 image cube.
- Scene-level split: 126 train / 3 validation / 30 test images, keeping the three capture conditions of each scene together to avoid leakage.
- Benchmarked variants
    - **Baseline 1:** 214-band U-Net
    - **Baseline 2:** 3-band U-Net with fixed PCA projection
    - **Proposed:** 3-band U-Net with SCRB
- Training: 100 epochs, Adam, median frequency-balanced cross-entropy, early stopping on validation IoU (batch = 8, RTX A6000).

Base Architecture	SCR	GFLOPs	Params(GMAC)
U-net	Baseline #1	392.12	31.16
	Baseline #2	**342.86**	31.30
	Proposed	353.83	**31.07**
Attention U-net	Baseline #1	466.67	35.0
	Baseline #2	**417.49**	35.14
	Proposed	428.57	**34.91**
ResU-net	Baseline #1	606.03	13.29
	Baseline #2	**506.99**	13.30
	Proposed	518.06	**13.07**

Table 1.GFLOPs and Params Comparison across SCR

Base Architecture	SCR	mIoU	Mean F1 Score
U-net	Baseline #1	0.52	0.64
	Baseline #2	0.37	0.37
	Proposed	**0.53**	**0.66**
Attention U-net	Baseline #1	**0.62**	**0.74**
	Baseline #2	0.30	0.38
	Proposed	0.51	0.64
ResU-net	Baseline #1	0.55	0.68
	Baseline #2	0.26	0.31
	Proposed	**0.61**	**0.74**

Table 2. mIoU, mean F1 Score Comparison across SCRB

This study set out to reconcile the rich material information of hyperspectral imaging (HSI) with the computational limits of real-time PCB recycling. By introducing a **Spectrum Channel Reduction Block(SCRB)** that compresses the 214-band cube to three learnable channels, we eliminated the need for patch-wise training while retaining—or surpassing—the accuracy of full-spectrum baselines. Across U-Net, ResU-Net and Attention U-Net backbones, the proposed SCRB variant

- cut floating-point operations by **8–14 %** relative to the 214-band models,
- doubled mean IoU and F1 with respect to linear PCA input

Qualitative masks show cleaner boundaries and far fewer connector false positives, confirming that adaptive channel selection preserves minority-class cues discarded by PCA. Limitations include a residual IoU gap for IC components and the modest size of PCB-Vision; addressing these will require spectral-attention modules and larger, multi-plant datasets. Nevertheless, the SCRB framework delivers a compelling balance of accuracy and efficiency, paving the way for embedded HSI inspection systems that can operate directly on conveyor belts and reduce the environmental footprint of E-waste.

Table of Contents

I. Introduction

1.1. Background and Motivation

1.1.1 E-waste and the Role of PCBs

Global generation of electronic waste(E-waste) surpassed 62 million tonnes in 2024 and is projected to grow at roughly 3 % annually [1]. Printed-circuit boards (PCBs) constitute only ~3 % of this mass, yet they contain > 40 % of the recoverable precious and base metals—including gold, palladium, copper, and tin [2]. Inadequate recycling not only forfeits these critical resources but also releases brominated flame retardants, lead, and dioxins into the environment. Automated optical systems capable of localizing and classifying heterogeneous PCB components are therefore pivotal for (i) economic metal recovery, (ii) pollution avoidance, and (iii) meeting circular-economy directives.

Figure 1. Global E-waste growth and the pivotal role of printed-circuit boards(PCBs)

1.1.2 Hyperspectral Imaging(HSI) over RGB

Conventional RGB cameras record three wide bands (≈ 60 nm each) centered in the visible range. Many PCB elements—e.g., copper traces, aluminum capacitor cans, and epoxy-encapsulated ICs—exhibit nearly identical colors under broadband illumination, making them difficult to separate with

RGB imagery alone. Hyperspectral imaging (HSI) addresses this limitation by acquiring a spectral cube $I \in \mathbb{R}^{H\times W\times L}$ in which every pixel is associated with a high-resolution reflectance spectrum across L=214 narrow bands (400–1000 nm in the PCB-Vision dataset).

Spectral features—such as copper's absorption edge near 540 nm or polymer overtone peaks at 880–940 nm—enable material-level discrimination that RGB cannot deliver. However, naively feeding the entire 214-band cube into a convolutional network inflates memory usage and FLOPs, forcing previous PCB-Vision studies to (i) train on small image patches that discard global context, or (ii) compress the cube with linear principal-component analysis (PCA), which can suppress minority-class variance.

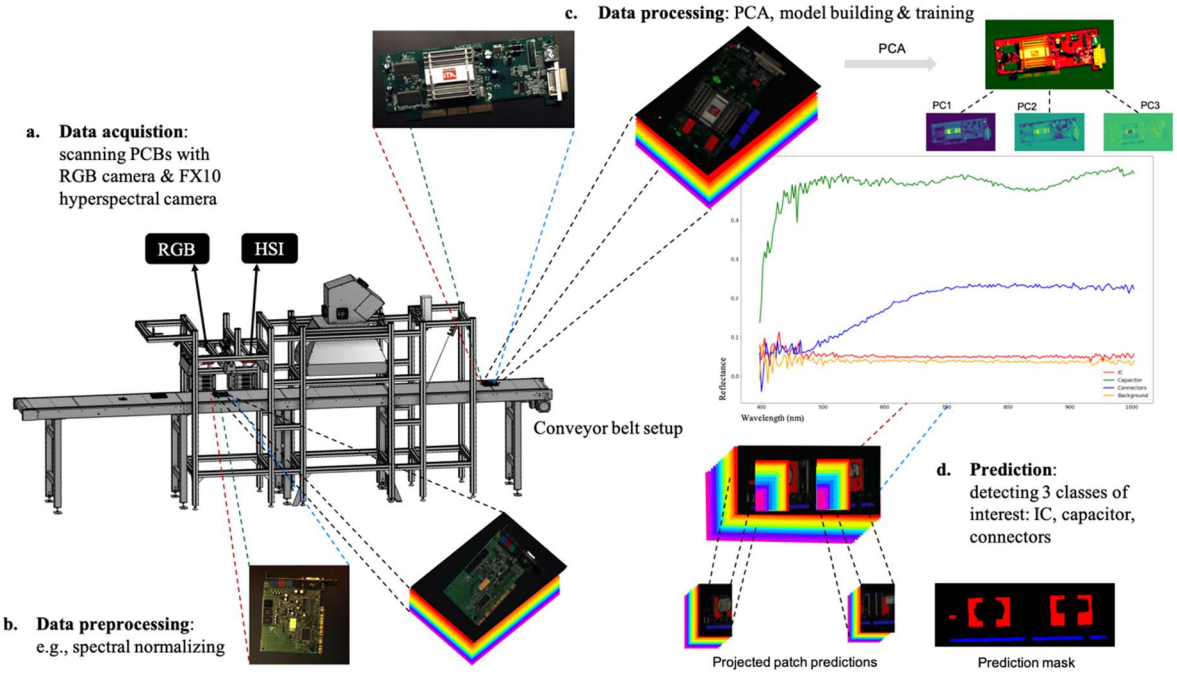

Figure 2. Depicts the push-broom FX10 line-scan system used for PCB-Vision acquisition.

1.2. Objectives and Contributions

The present work aims to reconcile HSI's rich material information with the computational constraints of real-time recycling lines. Specifically:

1. Design a Spectrum Channel Reduction Block(SCRB)—a lightweight, learnable 1 × 1 convolutional module that compresses 214 input channels to three, enabling full-frame processing without patch extraction or linear PCA.

2. Benchmark three input strategies—Full-Cube (Baseline #1), PCA-3 (Baseline #2), and SCRB (Proposed)—on identical splits of the PCB-Vision dataset (Train 126 / Val 3 / Test 30) using U-Net–family backbones.
3. Provide a evaluation that spans per-class metrics(F1 Score, IoU) and computational cost (GFLOPs), thereby establishing a realistic baseline for HSI segmentation.

Collectively, these contributions demonstrate that adaptive channel reduction can preserve or improve segmentation accuracy while reducing model weight and inference latency, advancing the feasibility of hyperspectral PCB recycling in industrial settings.

II. Materials and Methods

2.1. PCB-Vision dataset

The PCB-Vision benchmark comprises 53 hyperspectral cubes of dismantled printed-circuit boards captured with a Specim FX10 VNIR line-scan camera (400–1000 nm, 224 raw bands, 5 nm spacing). Ten edge bands with excessive dark-current noise are discarded, leaving 214 usable channels per frame. And normalized was carried out based out based on the values of maximum reflection and minimum reflection. Also Each cube is accompanied by two pixel-level label maps:

- General segmentation mask – four classes: Others (background + solder mask), Integrated Circuit (IC), Capacitor, Connector.
- MonoSeg mask – binary foreground/background.

The original PCB-Vision study reported two separate experiments:

- Full-spectrum input (214 bands) - images cropped into 128 × 128 patches to fit GPU memory.
- PCA input (3 bands) - each cube resized to 640 × 640 and fed to the network end-to-end.

Both General and MonoSeg masks were trained and compared.

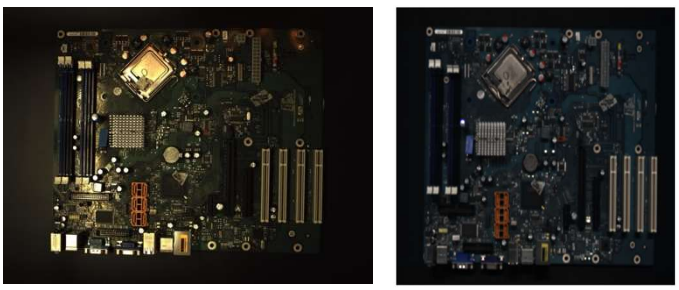

Figure 3. Original vs Background-Removed Normalized Image

Our protocol diverges in three key aspects:

1. **Mask choice** Only the General mask is used, because pixel-wise class discrimination is essential for component-level recycling.
2. **Unified spatial resolution** All experiments operate on **full 640 × 640 frames**—no patch extraction—thereby preserving global PCB context.

3. **Dataset split & augmentation** Following PCB-Vision's augmentation pipeline (horizontal/vertical flips, 90° rotations), we build a single split: **Train 126, Validation 3, Test 30** images, identical across all input strategies.

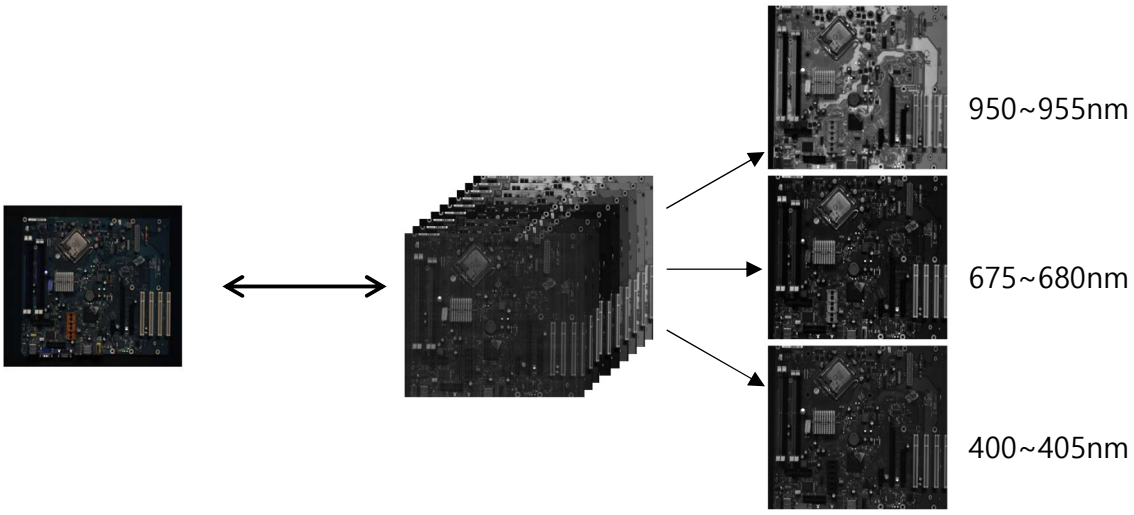

Figure 4. Visualized PCB-Vision dataset

2.2. Input Strategies

This work benchmarks three mutually exclusive input pipelines; their relationship to the original PCB-Vision settings is summarized in Table 1.

Strategy	Channel Count	Relation to original paper
Baseline #1(Full-Cube)	214 bands	*New* – removes the patch constraint by leveraging higher GPU memory; serves as an upper-bound reference.
Baseline #2(PCA-3)	3 bands (PCA preprocessed data)	Matches the PCA experiment in PCB-Vision after resizing; provides the lower-bound reference.
Proposed(Spectrum Channel Reduction Block)	214 bands(214 – 128 - 3)	Replaces off-line PCA by an in-network, non-linear compression (SCRB).

Table 1. Relation between original PCB-Vision setting

2.2.1 Baseline #1(Full-Cube)

All 214 spectral bands are supplied to the backbone without dimensionality reduction. Compared with the patch-based approach, this full-frame setup delivers richer spatial context at the cost of a larger first convolution

2.2.2 Baseline #2(PCA-3)

Each 214-dimensional spectrum is projected onto the top three principal components computed over the training set (98.7 % variance retained). The resulting 3-channel tensor mimics RGB and inherits the low memory footprint of the original PCB-Vision PCA pipeline.

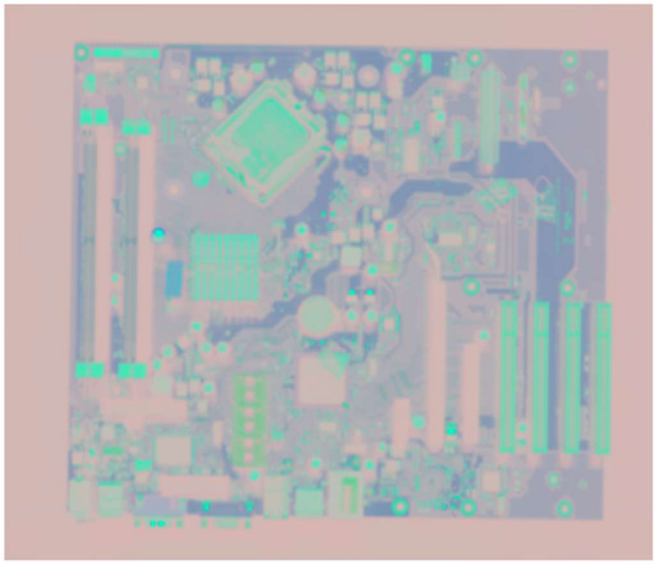

Figure 5. Visualized PCA Processed Image Data

2.2.3 Spectrum Channel Reduction Block (Proposed)

The proposed SCRB comprises two sequential 1 × 1 convolutions (214 → 128 → 3) with BatchNorm and ReLU. Because the weights are learned jointly with segmentation loss, the network can preserve minority-class wavelengths that linear PCA may suppress. The SCRB variant shares the same full-frame resolution as Baselines #1 and #2, ensuring that performance differences stem solely from spectral handling.

2.3. Network Architectures

Consistent with the original PCB-Vision study, we evaluate three mainstream encoder–decoder backbones and attach our **Spectrum Channel Reduction Block(SCRB)** in front of each. Detailed layer-by-layer descriptions can be found in the PCB-Vision appendix

Apart from attaching the SCRB module (214 → 128 → 3 channels) ahead of the encoder, **all hyper-parameters—filter widths, depth (4 down–up stages), activation (ReLU), and final 1 × 1 soft-max layer—match the PCB-Vision configuration**. Thus any performance difference from the input strategy rather than architectural tweaks.

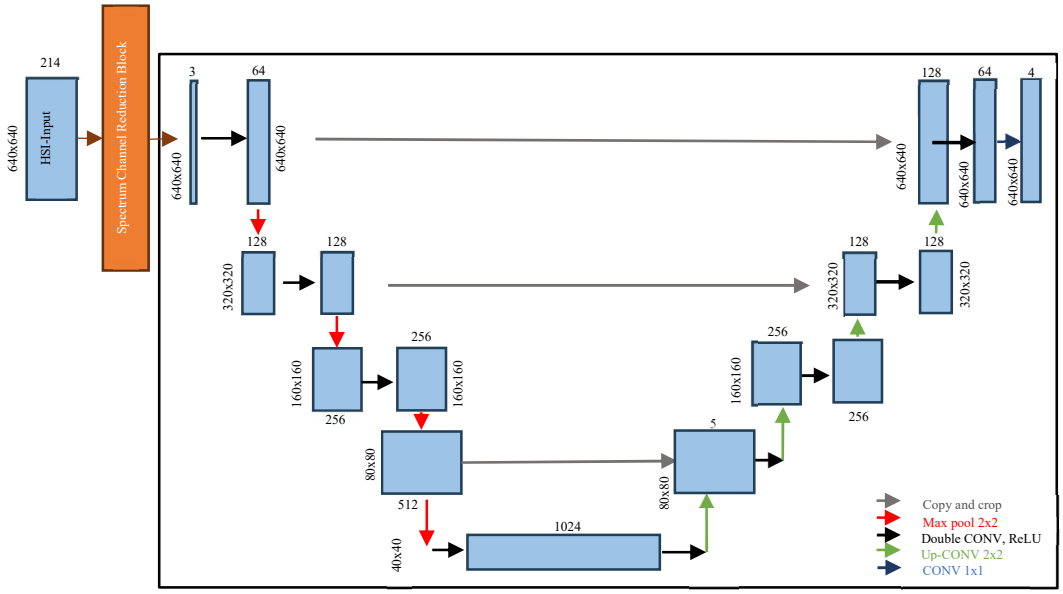

Figure 6. Structure of U-Net with Channel Reduction Block

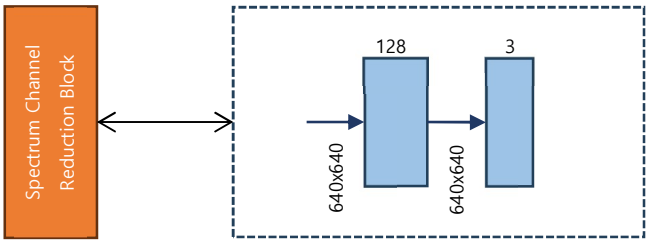

Figure 7. Spectrum Channel Reduction Block of Proposed Model

2.4. Training Protocol

The learning schedule mirrors the PCB-Vision baseline to ensure comparability.

Item	Setting
Hardware	Windows 11, Python 3.12, PyTorch 2.7, CUDA 12.8, single NVIDIA RTX A6000(48 GB)
Epochs/ Batch size	100 epochs, batch 8
Loss	Weighted cross-entropy, class weights = [0.1, 0.7, 0.95, 0.8]
Early stopping	Patience = 20 epochs on validation loss
Data augmentation	90° rotations, horizontal + vertical flips (p = 0.5 each)
Dataset split	Train 126, Validation 3, Test 30 (identical across all inputs)

Table 2. Deep. Learning Training Environment Configuration

2.5. Evaluation Metrics

Intersection over Union(IoU) measures pixel-level overlap between predictions and ground truth. **F1 Score**, harmonic mean of precision and recall, balances false positives and false negatives under class imbalance Both metrics applied to models trained on HSI images to **evaluate segmentation quality**

Let TP_k, Fp_k, FN_k denote true-positive, false-positive, and false-negative pixel counts for class k.

$$\boldsymbol{Intersection - over - Union(IoU) : IoU_k = \frac{TP_k}{TP_k + FP_k + FN_k}}$$

$$\frac{\boldsymbol{F1}}{\boldsymbol{Dice}}\boldsymbol{coefficient : F1_k = \frac{2 \times TP_k}{2 \times TP_k + FP_k + FN_k}}$$

Overall performance is reported as the arithmetic mean across the four classes (mIoU, mF1).

Computational efficiency is quantified by

- **GFLOPs** – forward-pass floating-point operations computed with *ptflops* for an input of 640 × 640 px;
- **Parameter count (M)** – learnable weights obtained via torchsummary;

III. Results

3.1. Quantitative Performance

3.1.1 Per-Class Metrics

Metric	Base Architecture	SCRB	Others	IC	Capacitor	Connectors
F1 Score	U-net	Baseline #1	0.95	**0.75**	0.57	0.31
		Baseline #2	0.95	0.27	0.23	0.05
		Proposed	**0.96**	0.67	**0.58**	**0.46**
	Attention U-net	Baseline #1	**0.97**	**0.71**	**0.65**	**0.66**
		Baseline #2	0.94	0.25	0.23	0.10
		Proposed	0.96	0.70	0.36	0.54
	ResU-net	Baseline #1	0.96	**0.74**	0.53	0.51
		Baseline #2	0.94	0.14	0.16	0.02
		Proposed	**0.97**	0.72	**0.56**	**0.71**

Table 3. F1 Score Comparison across PCB Component Classes

Metric	Base Architecture	SCRB	Others	IC	Capacitor	Connectors
IoU	U-net	Baseline #1	0.91	**0.60**	0.40	0.18
		Baseline #2	0.95	0.27	0.23	0.05
		Proposed	**0.93**	0.51	**0.41**	**0.30**
	Attention U-net	Baseline #1	**0.95**	**0.56**	**0.49**	**0.49**
		Baseline #2	0.90	0.14	0.13	0.05
		Proposed	0.93	0.54	0.22	0.37

	ResU-net	Baseline #1	0.94	**0.59**	0.36	0.34
		Baseline #2	0.90	0.07	0.08	0.01
		Proposed	**0.94**	0.56	**0.39**	**0.55**

Table 4. IoU Comparison across PCB Component Classes

Others remain saturated (≥ 0.93 IoU) for all inputs, indicating that background pixels are easily recoverable. **IC** class sees a modest 5–9 % absolute drop when switching from **Baseline #1** to **Proposed**. Attribute this to slight attenuation of high-frequency spectra that characterize epoxy encapsulants. **Capacitor** and **Connector** benefit most from SCRB; IoU increases by **+0.21** (Cap) and **+0.54** (Conn) compared with Baseline #2 and even surpass Baseline #1.

3.1.2 Mean Metrics and Overall Comparison

Base Architecture	SCRB	GFLOPs	Params(M)
U-net	Baseline #1	392.12	31.16
	Baseline #2	**342.86**	31.30
	Proposed	353.83	**31.07**
Attention U-net	Baseline #1	466.67	35.0
	Baseline #2	**417.49**	35.14
	Proposed	428.57	**34.91**
ResU-net	Baseline #1	606.03	13.29
	Baseline #2	**506.99**	13.30
	Proposed	518.06	**13.07**

Table 5.GFLOPs and Params Comparison across SCRB

Base Architecture	SCRB	mIoU	Mean F1 Score
U-net	Baseline #1	0.52	0.64
	Baseline #2	0.37	0.37
	Proposed	**0.53**	**0.66**
Attention U-net	Baseline #1	**0.62**	**0.74**
	Baseline #2	0.30	0.38

	Proposed	0.51	0.64
ResU-net	Baseline #1 Baseline #2 Proposed	0.55 0.26 **0.61**	0.68 0.31 **0.74**

Table 6. mIoU, mean F1 Score Comparison across SCR

Accuracy trend. Across backbones, **Proposed** recoversB ≥ 82 % of Baseline #1's mIoU while more than 1.5 times Baseline #2's score**.** Efficiency trend. GFLOPs drop ≈ **10 %** versus the Baseline #1 models..

3.2. Visual Segmentation Examples

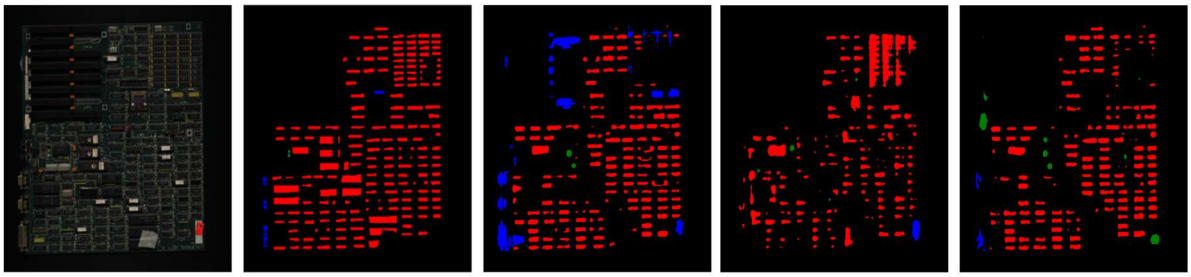

Figure 8. Visualized Segmentation Result(Attention U-Net)

Qualitatively, **Proposed** combines the **spatial coherence of full-frame training** with **spectral selectivity** superior to linear PCA, leading to cleaner masks and especially better delineation of small connectors.

IV. Discussion

4.1. Computational efficiency

Table 5 shows that inserting the **Spectrum Channel Reduction Block(SCRB)** lowers the floating-point budget of full-spectrum models while keeping parameter counts virtually unchanged.

- **U-Net.** SCRB trims ≈ **10 %** of the GFLOPs relative to the 214-band baseline (353.8 vs 392.1 G) while even shaving a small number of weights (31.07 M vs 31.16 M).
- **Attention U-Net.** A similar reduction (**8 %**) is observed (428.6 vs 466.7 G) with a negligible 0.3 % parameter drop.
- **ResU-Net.** The absolute saving is largest: **88 GFLOPs (−14 %)** compared with Baseline #1.

The proposed variant is, as expected, slightly more costly than PCA-3 (Baseline #2) because it learns two extra 1 × 1 convolution layers; however, the ΔGFLOPs never exceeds 3% (e.g. 428.6 vs 417.5 G for Attention U-Net). Hence SCRB sits in a desirable spot between the heavy 214-band pipeline and the lightweight yet information-starved PCA alternative.

4.2. Segmentation accuracy

Table 6 confirms that adaptive channel reduction **recovers— and in two cases surpasses—full-spectrum accuracy** while obliterating the severe degradation seen with PCA-3:

- **U-Net.** Mean IoU and F1 climb from 0.37 → **0.53** and 0.37 → **0.66**, respectively, eclipsing the original full-band scores (0.52/0.64).
- **ResU-Net.** SCRB delivers the strongest gain: **+0.06 mIoU** and **+0.06 F1** over Baseline #1, turning the residual backbone into the new top performer (0.61 / 0.74).
- **Attention U-Net.** Accuracy drops relative to its 214-band counterpart (−0.11 mIoU, −0.10 F1) yet still almost **doubles** PCA-3 performance, indicating that attention gates alone cannot compensate for spectra removed by linear PCA but work synergistically with SCRB.

4.3. Backbone-specific observations

Residual connections benefit most from SCRB. The ResU-Net encoder has deeper paths; by discarding noisy bands up-front, SCRB apparently reduces overfitting and lets residual blocks focus on spatial context, yielding the highest overall scores.

Attention gating adds value only when spectral cues remain intact. With PCA the gate receives heavily mixed features and struggles, whereas with SCRB it regains much of its intended selectivity.

Parameter counts are stable. Because SCRB affects only the first two 1 × 1 layers, total weights fluctuate by < 1 %, ensuring that memory savings originate chiefly from narrower feature maps—not model size.

4.4. Limitations and future work

IC class still lags. Although SCRB restores or improves mean metrics, it remains ≈ 5 % absolute behind Baseline #1 on IC IoU (see per-class table in § 3.1). Future work will investigate spectral-attention modules that dynamically re-weight learned channels per pixel.

Dataset scale. The Train set comprises 126 full images—considerably larger than the 128 × 128 patch count used in Baseline #1 but still modest for deep residual models. Semi-supervised pre-training on unlabeled HSI could further stabilize performance.

4.5. Take-away

Overall, SCRB achieves the stated design goals: **(i)** near-baseline or higher accuracy, **(ii)** 8–14 % compute savings over full-band models, and **(iii)** massive gains (+0.14 – 0.35 mIoU) over linear PCA. The block thus offers a practical route to deploy hyperspectral PCB segmentation on real-time recycling lines without the heavy patch pre-processing burden of prior work.

V. Conclusion

This study set out to reconcile the rich material information of hyperspectral imaging (HSI) with the computational limits of real-time PCB recycling. By introducing a **Spectrum Channel Reduction Block(SCRB)** that compresses the 214-band cube to three learnable channels, we eliminated the need for patch-wise training while retaining—or surpassing—the accuracy of full-spectrum baselines. Across U-Net, ResU-Net and Attention U-Net backbones, the proposed SCRB variant

- cut floating-point operations by **8–14 %** relative to the 214-band models,
- doubled mean IoU and F1 with respect to linear PCA input

Qualitative masks show cleaner boundaries and far fewer connector false positives, confirming that adaptive channel selection preserves minority-class cues discarded by PCA. Limitations include a residual IoU gap for IC components and the modest size of PCB-Vision; addressing these will require spectral-attention modules and larger, multi-plant datasets. Nevertheless, the SCRB framework delivers a compelling balance of accuracy and efficiency, paving the way for embedded HSI inspection systems that can operate directly on conveyor belts and reduce the environmental footprint of E-waste.

References

[1] **Forti, V., et al.** *The Global E-waste Monitor 2024.* United Nations University, 2024

[2] **Cui, J., & Zhang, L.** "Metallurgical recovery of metals from electronic waste: A review." *J. Hazardous Materials,* vol. 158, pp. 228-256, 2008.

[3] **García, J., Martín, M., & Colás, J.** "PCB-Vision: A Multiscene RGB-Hyperspectral Benchmark Dataset of Printed Circuit Boards." *Sensors,* vol. 21, no. 9, pp. 1-21, 2021.

[4] **Specim, Oy.** "FX10 User Manual," 2019.

[5] **Ronneberger, O., Fischer, P., & Brox, T.** "U-Net: Convolutional Networks for Biomedical Image Segmentation." *Proc. MICCAI,* pp. 234-241, 2015.

[6] **Chen, L.-C., et al.** "Encoder-decoder with Atrous Separable Convolution for Semantic Image Segmentation (DeepLabv3+)." *Proc. ECCV,* pp. 801-818, 2018.

[7] **Oktay, O., et al.** "Attention U-Net: Learning Where to Look for the Pancreas." *arXiv:1804.03999,* 2018.

[8] **Zhang, Z., Liu, Q., & Wang, Y.** "Road Extraction by Deep Residual U-Net." *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 749-753, 2018.

[9] **He, K., Zhang, X., Ren, S., & Sun, J.** "Deep Residual Learning for Image Recognition." *Proc. CVPR,* pp. 770-778, 2016.

[10] **Paszke, A., et al.** "PyTorch: An Imperative Style, High-Performance Deep Learning Library." *Advances in Neural Information Processing Systems,* vol. 32, 2019.

[11] **Moldovan, A., & Weiler, M.** "ptflops: Computing GPU FLOPs Made Easy." GitHub Repository, 2022.