

Biorthogonal Tunable Wavelet Unit with Lifting Scheme in Convolutional Neural Network

An Le¹, Hung Nguyen¹, Sungbal Seo², You-Suk Bae², Truong Nguyen¹

¹ Electrical and Computer Engineering Department, University of California San Diego, La Jolla, CA 92093, USA
{d0le,hun004,tqn001}@ucsd.edu

² Department of Computer Engineering, Tech University of Korea, Siheung 15073, Korea
{sungbal,ysbae}@tukorea.ac.kr

Introduction

We present **LS-BiorUwU**, a lifting-scheme biorthogonal tunable wavelet unit that replaces **conv, pooling, and stride operations** to preserve fine detail. Integrated into ResNet-18/34, it improves classification (**+2.12% CIFAR-10; +9.73% DTD**) and yields balanced anomaly segmentation/detection on **MVTec-AD (hazelnut)**.

Contributions

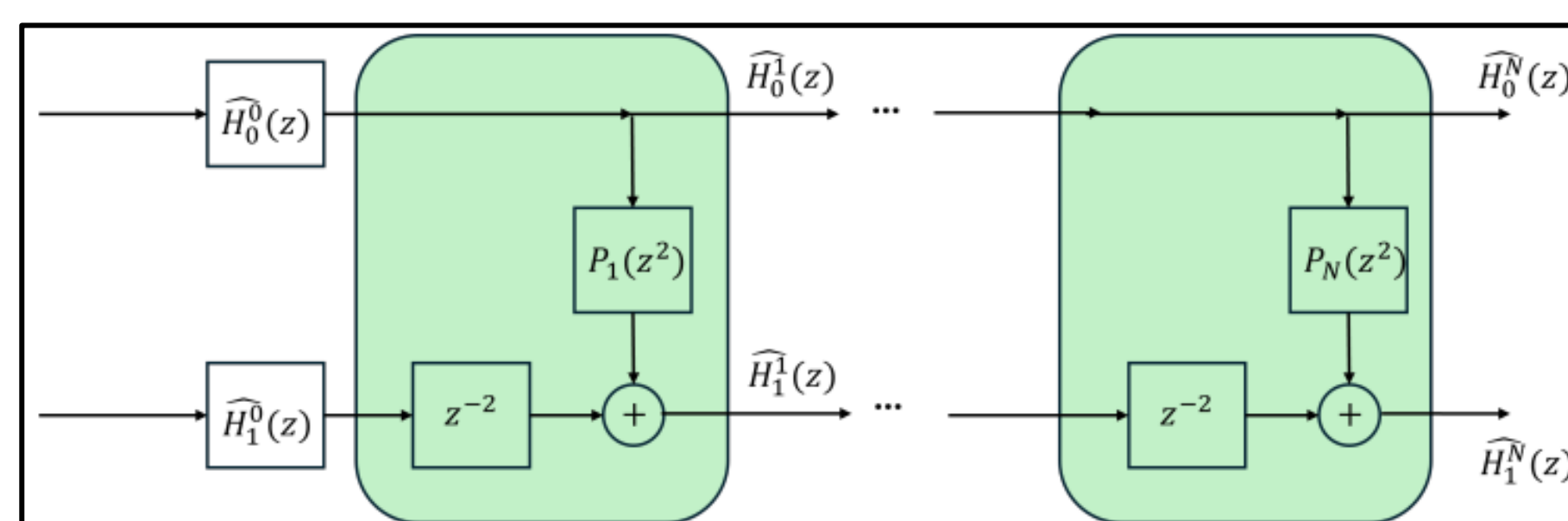
- Learnable lifting-scheme biorthogonal wavelet unit with relaxed constraints.
- Drop-in downsampling upgrade: **Replaces/augments conv, pooling, and stride to retain low/high-frequency detail**.
- Stronger classification: Consistent gains on ResNet-18 (similar trend on ResNet-34).
- Robust anomalies: Balanced segmentation and detection on MVTec-AD (hazelnut).

In future work, we'll scale to larger models/datasets, learn multi-level/adaptive filter banks with spectral regularization, and extend to detection/segmentation and 3D volumes with edge-latency profiling.

LS-BiorUwU

LS-BiorUwU: A **lifting-based, tunable wavelet unit** that replaces strict orthogonality with biorthogonality and permits unequal-length filters, expanding the filter bank's flexibility.

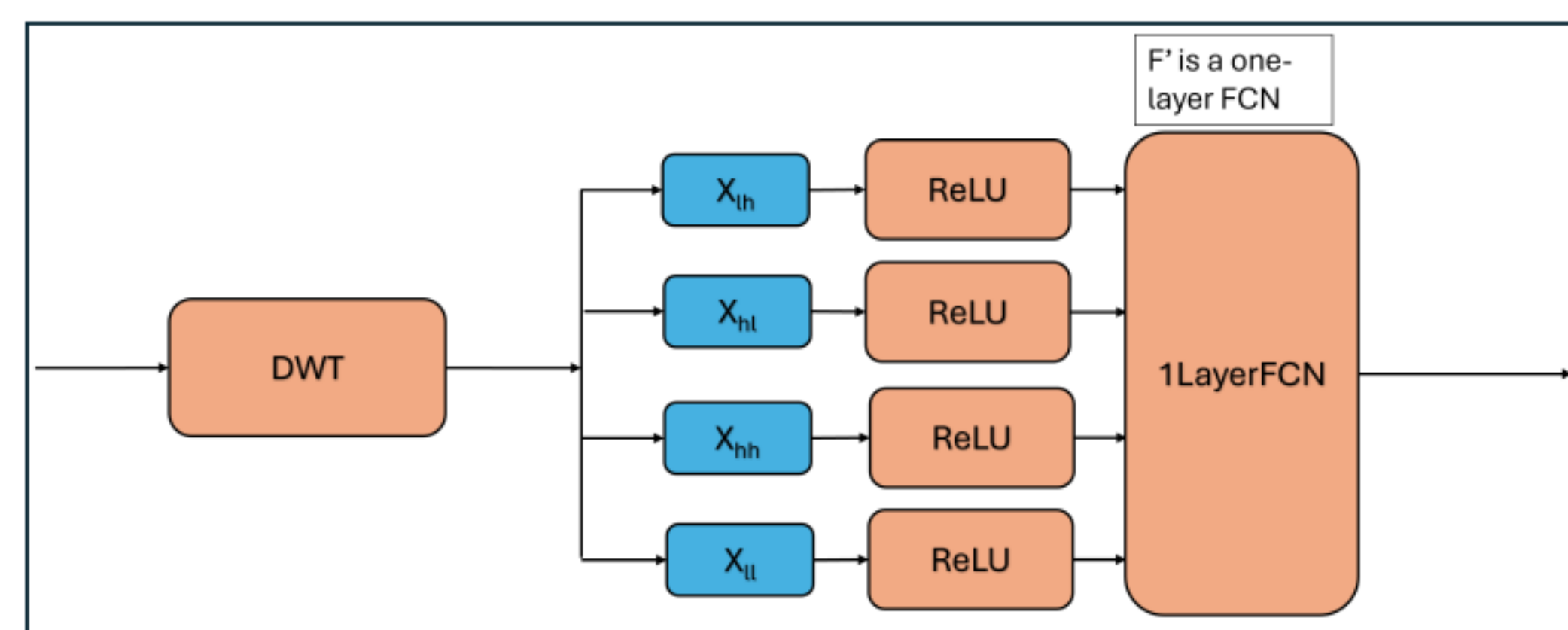
The **analysis** filter bank uses low-pass $\mathbf{H}_0$ and high-pass $\mathbf{H}_1$. While orthogonal wavelets require $\mathbf{h}_0$ and $\mathbf{h}_1$ to be equal-length and related, our **lifting-scheme** construction relaxes this to a **biorthogonal** design with **unequal lengths**. Starting from an initial filter pair ($\mathbf{H}_0^0(z)$, $\mathbf{H}_1^0(z)$), N lifting steps produce the learned analysis pair ($\mathbf{H}_0^N(z)$, $\mathbf{H}_1^N(z)$), where for with trainable a_k each step applies:



$$\begin{bmatrix} H_0(z) \\ H_1(z) \end{bmatrix} = \begin{bmatrix} \widehat{H_0^N}(z) \\ \widehat{H_1^N}(z) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ P_N(z^2) & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & z^{-2} \end{bmatrix} \cdots \begin{bmatrix} 1 & 0 \\ P_1(z^2) & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & z^{-2} \end{bmatrix} \begin{bmatrix} \widehat{H_0^0}(z) \\ \widehat{H_1^0}(z) \end{bmatrix}$$

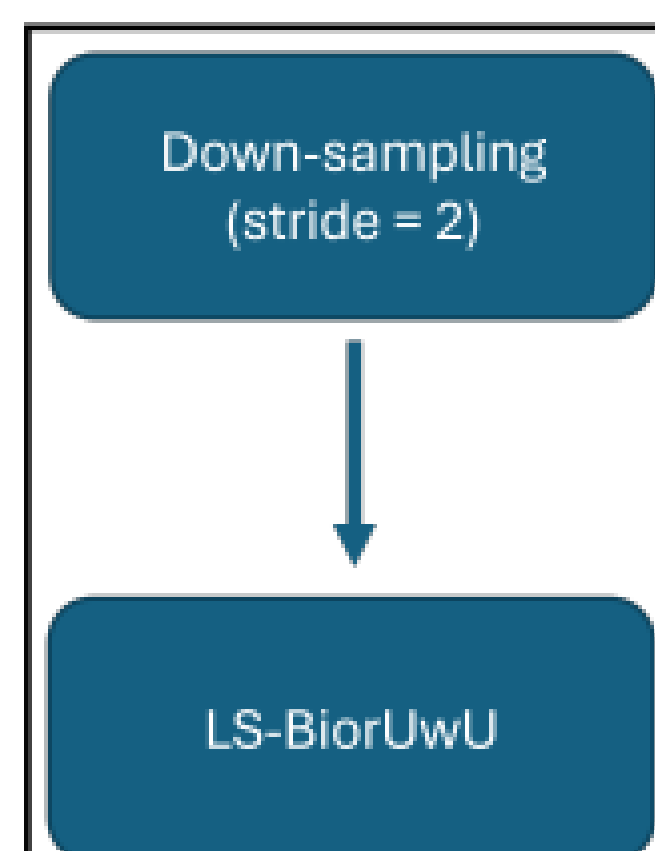
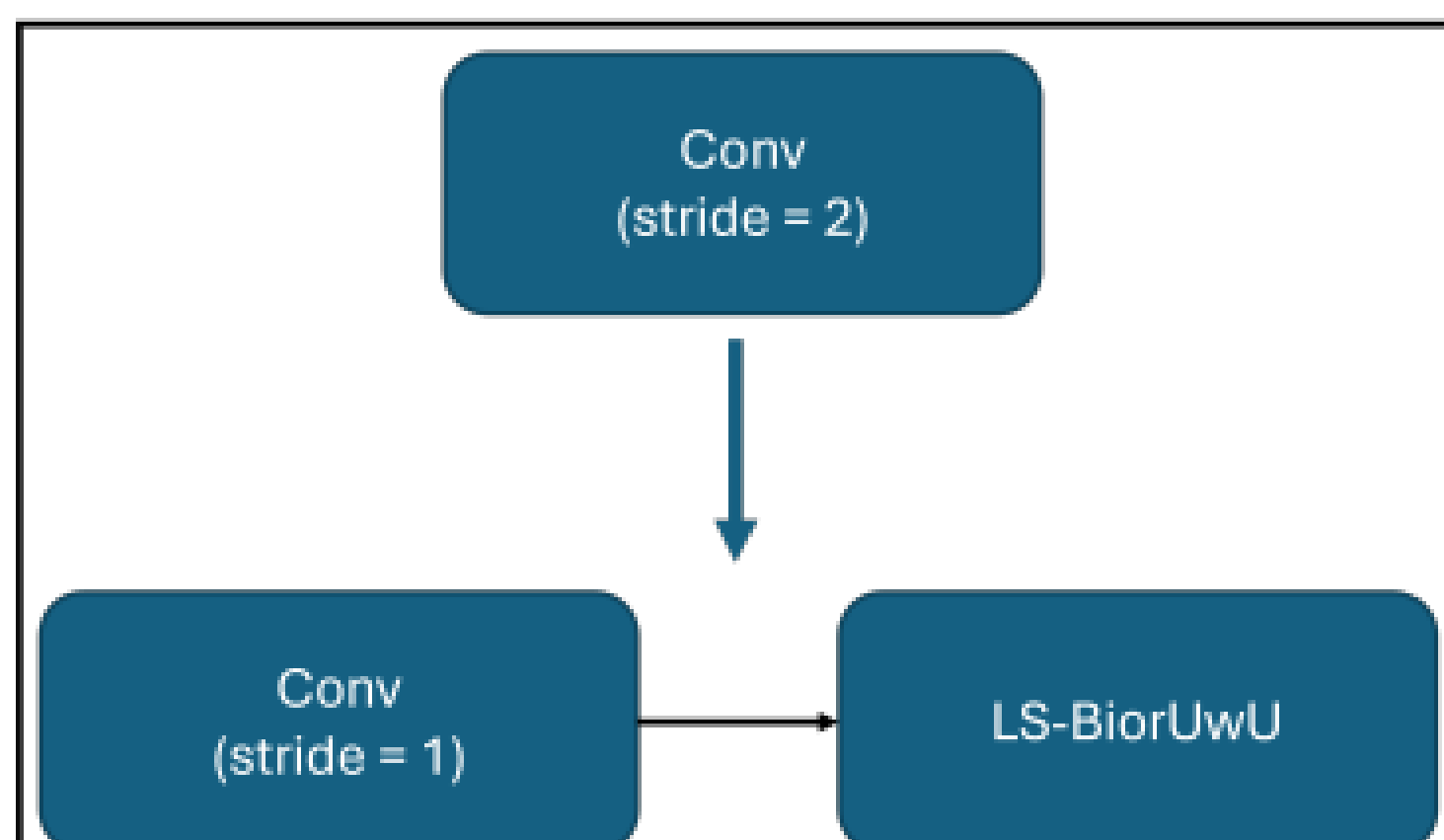
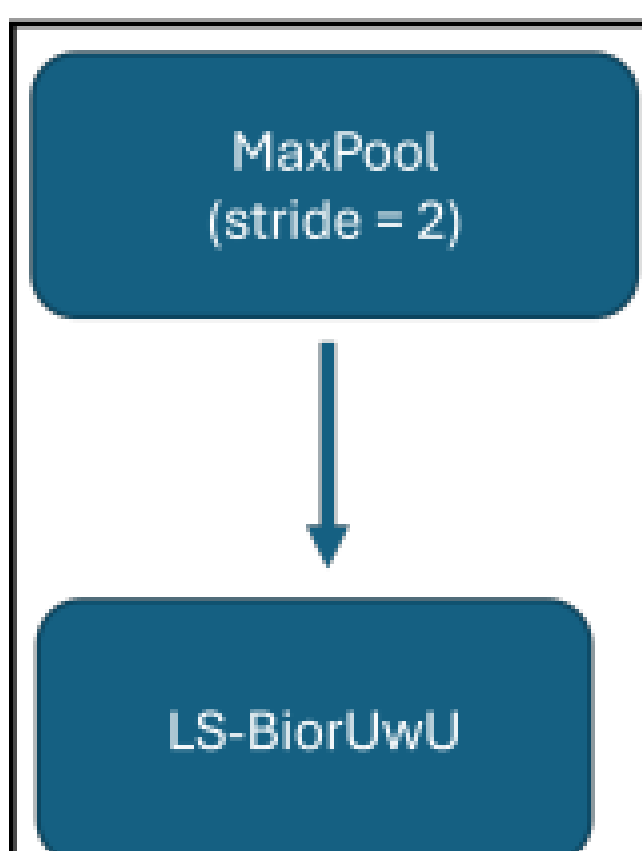
$$P_k(z) = -a_k + a_k z^{-2k} \text{ for } k \text{ in } [1, N]$$

CNN Implementation



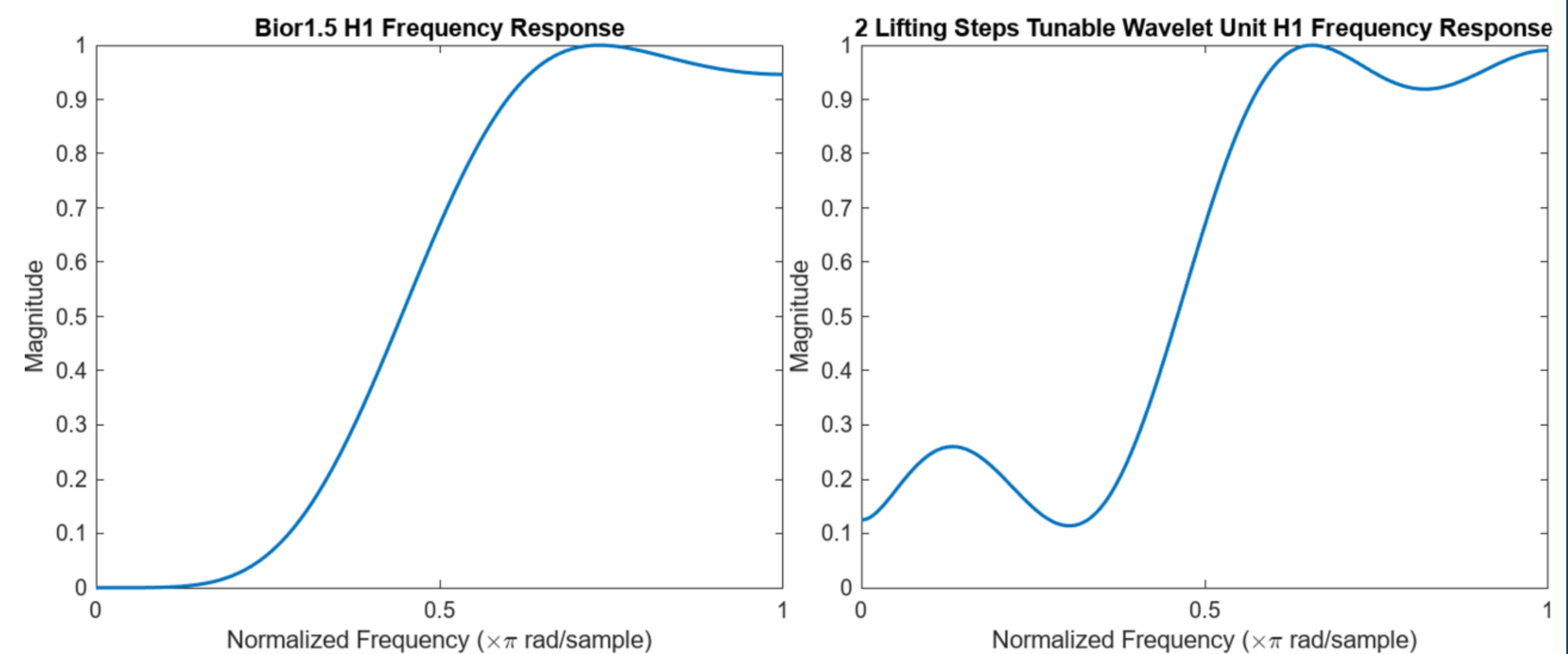
Applying these to a 2D input $\mathbf{X}$ yields the approximation $\mathbf{X}_{ll} = \mathbf{LXL}^T$ and the detail components $\mathbf{X}_{lh} = \mathbf{LXH}^T$, $\mathbf{X}_{hl} = \mathbf{HXL}^T$, and $\mathbf{X}_{hh} = \mathbf{HXH}^T$

Integrated into ResNet backbones, we replace stride-2 convolutions with a non-stride conv followed by **LS-BiorUwU** and append a 1-layer FCN after each UwU in downsampling/pooling.



Lifting Scheme for Tunable Biorthogonal Wavelet Filters Coefficient Analysis

Integrated into a ResNet-18 pooling layer on CIFAR-10, the 2-step LS-BiorUwU shows its tuned $\mathbf{H}_1$ retains high-pass characteristics.



Frequency Response Analysis of the Bior1.5 $\mathbf{H}_1$ (left) and the tuned coefficients of $\mathbf{H}_1$ of LS-BiorUwU after 2 lifting steps (right).

Classification Task

Accuracy of LS-BiorUwU with 1, 2 and 3 Lifting Steps, with OrthLatt-UwW-2Tap Evaluated on ResNet18 for DTD and CIFAR10.

DTD			
Baseline: 33.99%			
OrthLatt-UwU-2Tap	LS-BiorUwU-1Step	LS-BiorUwU-2Step	LS-BiorUwU-3Step
40.37%	42.71%	43.67%	43.72%
CIFAR10			
Baseline: 92.44%			
OrthLatt-UwU-2Tap	LS-BiorUwU-1Step	LS-BiorUwU-2Step	LS-BiorUwU-3Step
94.97%	94.08%	94.56%	94.39%

Accuracy of LS-BiorUwU with 1, 2 and 3 Lifting Steps, with OrthLatt-UwW-2Tap Evaluated on ResNet34 for DTD and CIFAR10.

DTD			
Baseline: 24.47%			
OrthLatt-UwU-2Tap	LS-BiorUwU-1Step	LS-BiorUwU-2Step	LS-BiorUwU-3Step
41.49%	41.76%	41.76%	42.45%
CIFAR10			
Baseline: 94.33%			
OrthLatt-UwU-2Tap	LS-BiorUwU-1Step	LS-BiorUwU-2Step	LS-BiorUwU-3Step
95.44%	94.45%	94.56%	94.36%

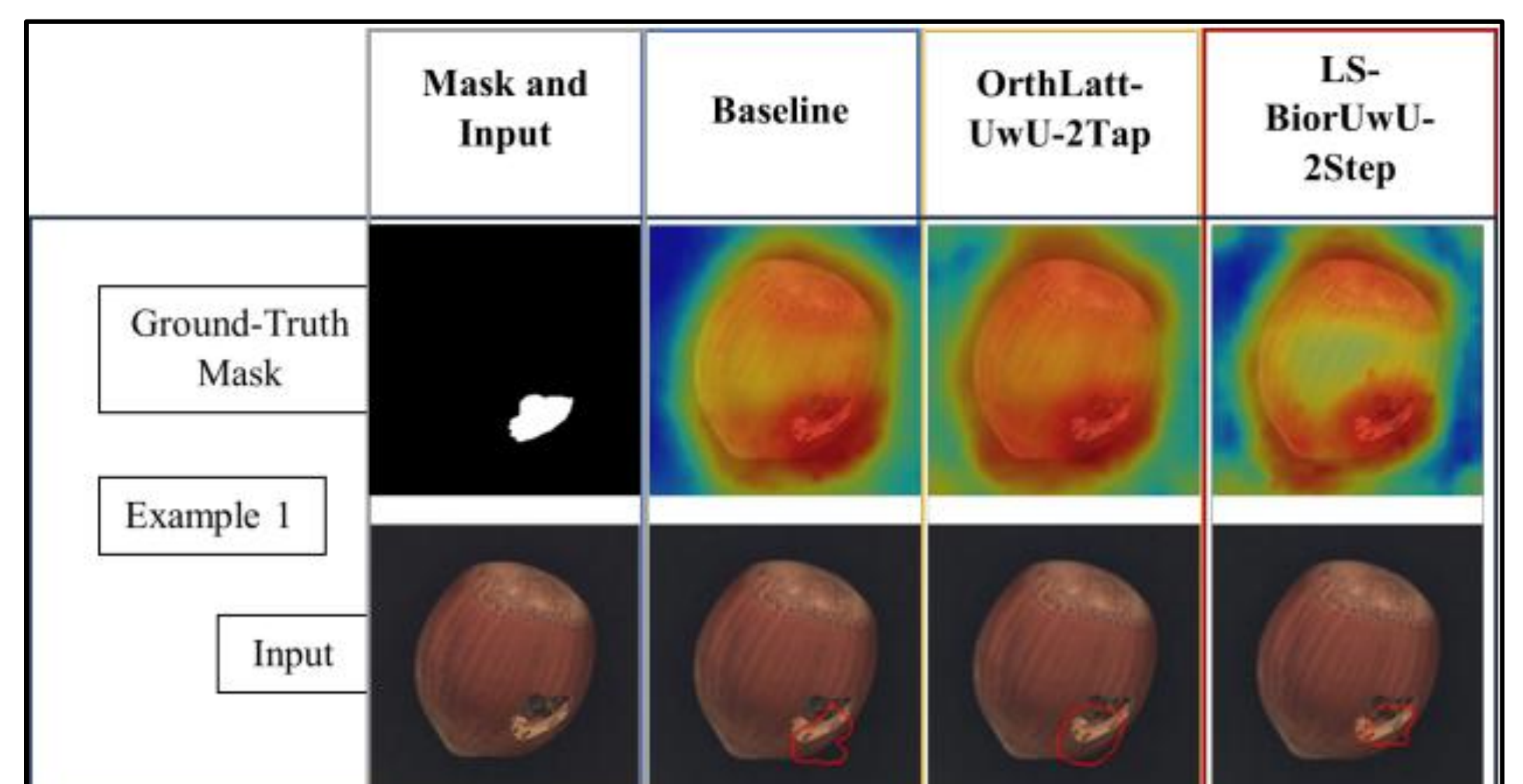
LS-BiorUwU with one to three lifting steps was integrated into ResNet-18 and ResNet-34 and evaluated on CIFAR-10 and DTD. Overall, LS-BiorUwU remains strong on low-res data and excels on texture-rich images, with benefits carrying over to deeper networks.. Increasing the number of steps improved accuracy and the benefits carried to the deeper network.

Anomaly Detection

Segmentation and detection AUROC on the MVTec-AD Hazelnut category for the baseline ResNet-18, OrthLatt-UwU-2Tap, and LS-BiorUwU-2Step encoders.

Models	SegAUROC	DetAUROC
Baseline	96.45%	92.46%
OrthLatt-UwU-2Tap	97.20%	89.21%
LS-BiorUwU-2Step	97.21%	92.11%

Visualization of anomaly detection on hazelnuts from the MVTec Anomaly Detection dataset. From left to right: ground-truth mask with input image, then heatmaps and defect segmentation from the baseline, OrthLatt-UwU-2Tap, and LS-BiorUwU-2Step.



References

- [1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.
- [2] A. D. Le, S. Jin, Y.-S. Bae, and T. Q. Nguyen, "A lattice-structure-based trainable orthogonal wavelet unit for image classification," IEEE Access, vol. 12, pp. 88715–88727, 2024.
- [3] A. Krizhevsky, "Learning multiple layers of features from tiny images," 2009.
- [4] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, and A. Vedaldi, "Describing textures in the wild," in Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2014.
- [5] P. Bergmann, K. Batzner, M. Fauser, D. Sattlegger, and C. Steger, "The mvtec anomaly detection dataset: A comprehensive real-world dataset for unsupervised anomaly detection," International Journal of Computer Vision, vol. 129, no. 4, p. 1038–1059, 2021.