



Tunable Wavelet Unit based Convolutional Neural Network in Optical Coherence Tomography Analysis Enhancement for Classifying Type of Epiretinal Membrane Surgery

An Le¹, Nehal Mehta², William Freeman², Ines Nagel², Melanie Tran², Anna Heinke², Akshay Agnihotri², Lingyun Cheng², Dirk-Uwe Bartsch², Hung Nguyen¹, Truong Nguyen¹, Cheolhong An¹

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

² Jacobs Retina Center, Shiley Eye Institute, University of California San Diego, La Jolla, CA 92093, USA
{nnmehta,wrfreeman,inagel,mdtran,aheinke,aagnihotri,lcheng,dbartsch}@health.ucsd.edu



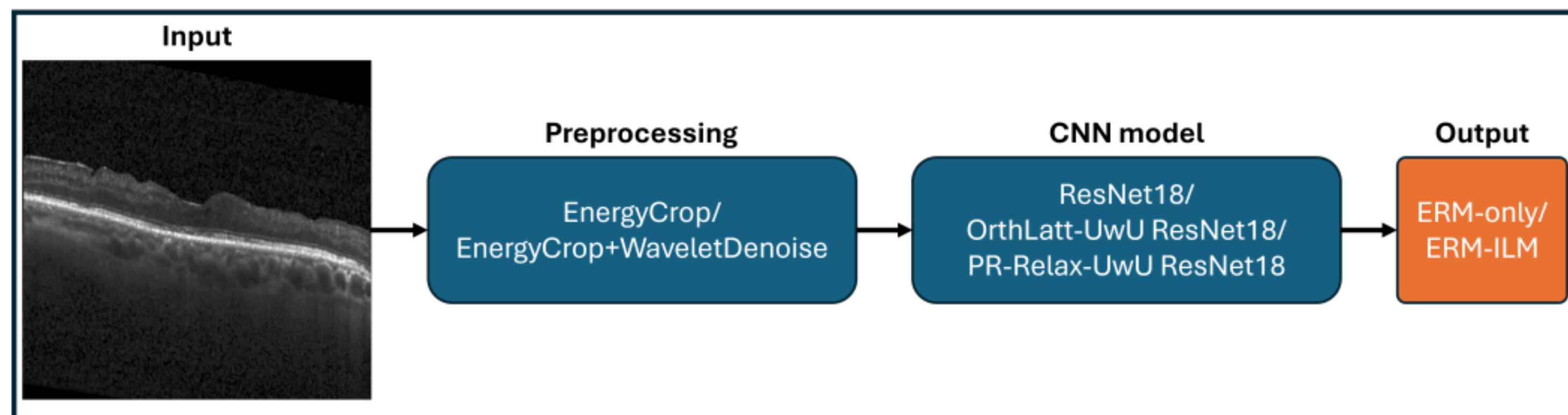
Introduction

We address post-operative OCT classification of ERM surgery type—ERM+ILM removal vs. ERM-only. A ResNet-18 baseline achieves 66% on raw scans and 72% with energy-crop + wavelet-denoising. To preserve high-frequency cues, we embed tunable wavelet units into downsampling/stride-2/pooling.

Contributions

- Baseline setup and preprocessing (with energy-crop + wavelet-denoising.): 66% → 72% accuracy on postoperative OCT.
- Detail-preserving modules: OrthLatt-UwU (76%) and PR-Relax-UwU (78%).
- Human Comparison: Model outperforms a trained human grader (50%) on the same postoperative OCT task.
- **Novelty:** To our knowledge, this is the first application of tunable wavelets to classify ERM surgery type from postoperative OCT.
- **Clinical Relevance:** Demonstrates that detail-preserving CNNs can improve postoperative assessment, surgical optimization, and pathology understanding.
- **Future work:** Extend to OCT volumes and use augmentation/diffusion to enable deeper, more stable models with reduced performance variability.

ERM/ERM-ILM Removal Surgery Classification Pipeline

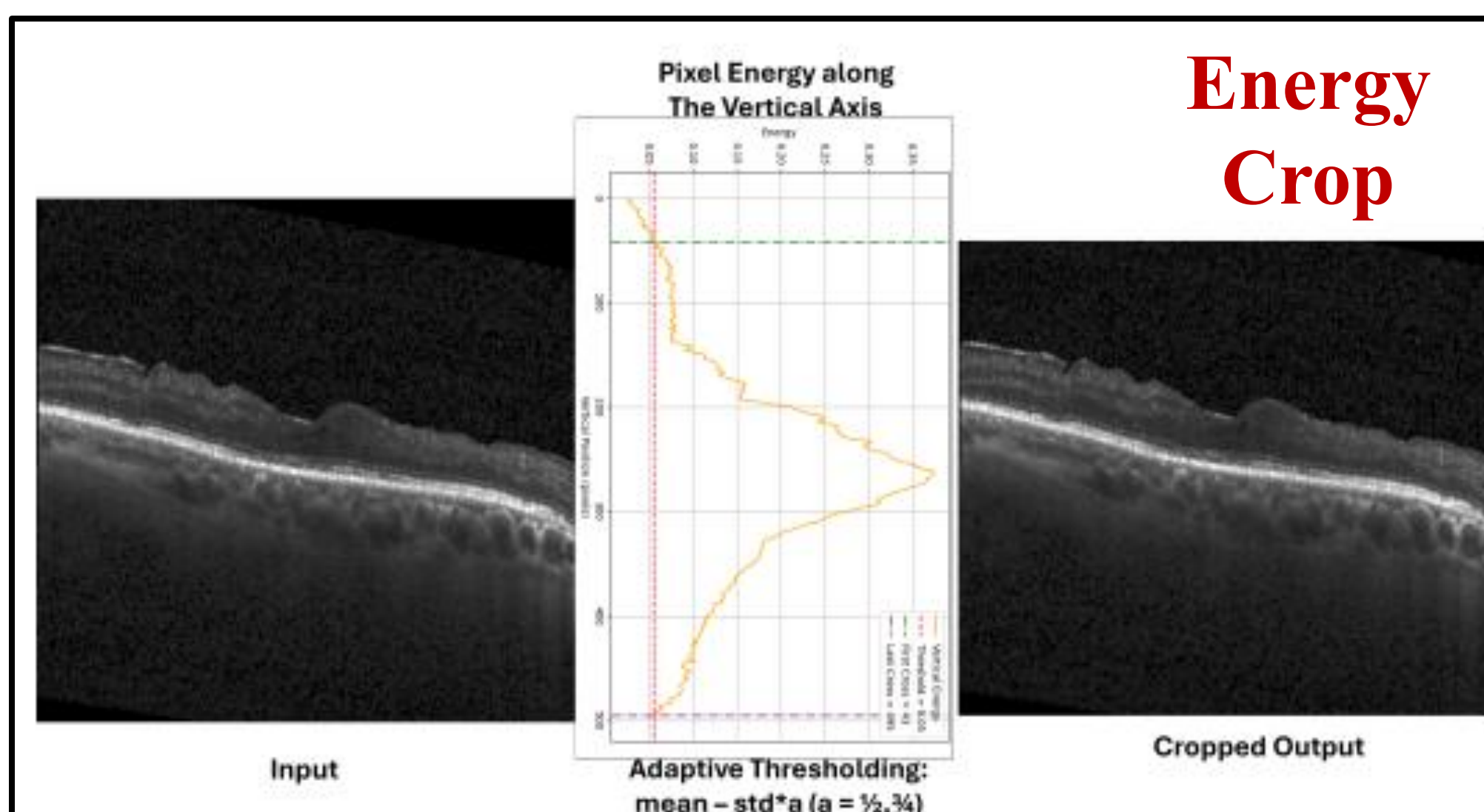


Preprocessing

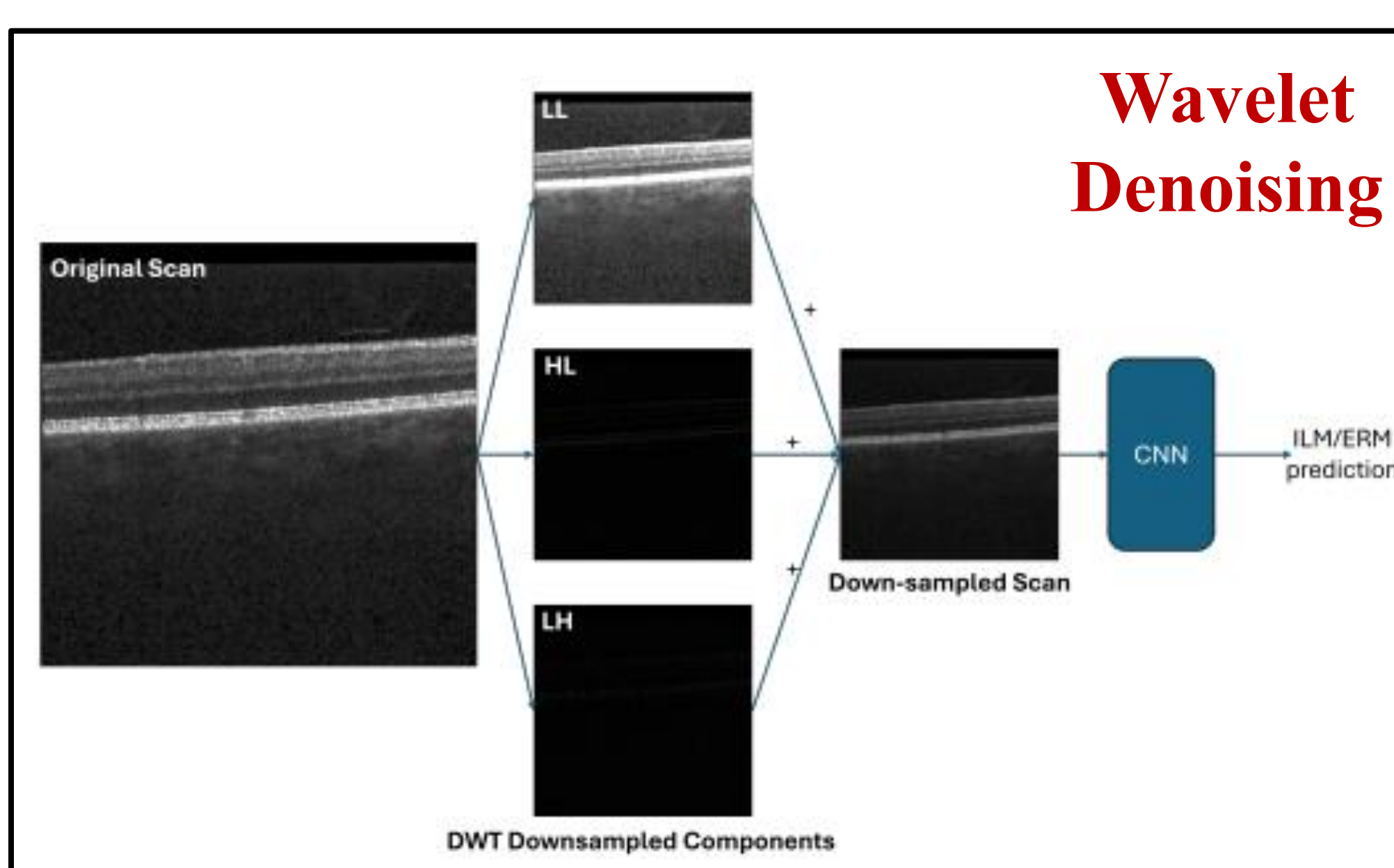
We focus inputs on salient retinal structure and suppress noise. **Energy-Crop** computes a vertical pixel-energy profile and crops between the first/last coordinates exceeding $\text{mean}(E_y) - a \cdot \text{std}(E_y)$ (larger $a \rightarrow$ tighter crops). **Wavelet Denoising** applies a DWT and reconstructs from LL , LH , and HL while discarding HH , removing noise and preserving fine detail for downstream classification.

We crop to the region where vertical pixel-energy E_y exceeds $T = \text{mean}(E_y) - a \cdot \text{std}(E_y)$; larger a yields tighter crops, and a is tuned empirically.

$E_y = \sum |p_{x,y}|$
where $p_{x,y}$ is the pixel value at coordinate (x, y) .



We apply a discrete wavelet transform (DWT) to decompose each image into LL , LH , HL , and HH subbands. To denoise while preserving structure, we reconstruct the image as $I_{\text{denoised}} = LL + LH + HL$ and discard HH , which predominantly contains bi-axial high-frequency noise.



Universal Wavelet Units

We integrate tunable wavelet units into CNNs to mitigate detail loss from pooling. Each unit performs a DWT with learnable coefficients, preserving high-frequency features and optimizing filters end-to-end for classification.

OrthLatt-UwU

OrthLatt-UwU forms $\mathbf{H}_0, \mathbf{H}_1$ by cascading orthogonal rotations $\mathbf{R}_k(\theta_k)$ with delay matrices $\Lambda(z^2)$, yielding an order $N=2K+1$ filter bank. Orthogonality guarantees perfect reconstruction, and the **rotation angles θ_k are learned end to end**, making the unit tunable within the CNN.

$$\begin{aligned} \begin{bmatrix} \mathbf{H}_0(z) \\ \mathbf{H}_1(z) \end{bmatrix} &= \begin{bmatrix} \mathbf{H}_0(z) \\ -z^{-(N-1)}\mathbf{H}_0(-z^{-1}) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{R}_K \Lambda(z^2) \cdots \mathbf{R}_1 \Lambda(z^2) \mathbf{R}_0 \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix} \\ \mathbf{R}_k &= \begin{bmatrix} \cos(\theta_k) & \sin(\theta_k) \\ -\sin(\theta_k) & \cos(\theta_k) \end{bmatrix} = \begin{bmatrix} c_k & s_k \\ -s_k & c_k \end{bmatrix} \\ \Lambda(z) &= \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} \end{aligned}$$

PR-Relax-UwU

Learns $\mathbf{H}_0, \mathbf{H}_1$ while relaxing strict PR by **enforcing alias cancellation** and penalizing deviations from a **half-band response** (via $\mathbf{P}(z) = \mathbf{H}_0(z)\mathbf{H}_0(z^{-1})$). Training minimizes $\mathbf{L} = \mathbf{L}_{\text{CE}} + \alpha \mathbf{L}_{\text{PR}}$, where α controls the PR strength—higher α enforces near-PR; lower α allows more flexible filters.

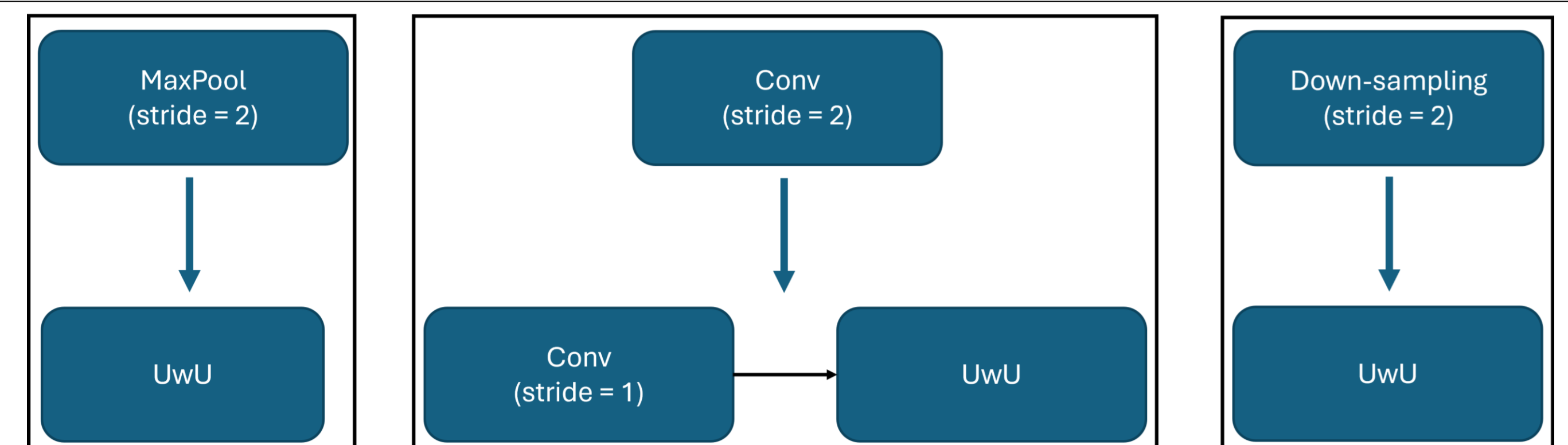
Anti-Aliasing Condition

$$\begin{cases} \text{Order Flip:} & \mathbf{f}_0(n) = \mathbf{h}_0(N-1-n) \\ \text{Alternating Flip:} & \mathbf{h}_1(n) = (-1)^n \mathbf{h}_0(N-1-n) \\ \text{Alternating Sign:} & \mathbf{f}_1(n) = -(-1)^n \mathbf{h}_0(n). \end{cases}$$

Half-Band Condition Relaxation

$$\mathbf{L}_{\text{PR}} = \left| 1 - \sum_{n=0}^{N-1} h(n)^2 \right|^2 + \sum_{l=1}^{N/2} \left(\sum_{0 < n, n+2l < N-1} h(n)h(n+2l) \right)^2$$

Implementation of UwU in CNN architecture, replacing max-pool, stride-convolution, and downsampling functions.



Results

Comparison across ResNet-18 variants: baseline (no preprocessing), with preprocessing (Energy-Crop, Wavelet-Denoising), and with tunable wavelet units (OrthLatt-UwU, PR-Relax-UwU). The trained human grader achieved 50% accuracy

Methods	Best on Test
Baseline	66%
EC-0.5	68%
EC-0.5&Denoise-DB2	72%
OrthLatt-UwU-8Tap EC-0.5&Denoise-DB2	76%
PR-Relax-UwU-8Tap EC-0.5&Denoise-DB2	78%

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] Strang, G., Nguyen, T.Q.: Wavelets and filter banks (1996).
- [3] 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
- [4] A. D. Le, S. Jin, Y. S. Bae, and T. Nguyen, "A novel learnable orthogonal wavelet unit neural network with perfection reconstruction constraint relaxation for image classification," in 2023 IEEE International Conference on Visual Communications and Image Processing (VCIP), pp. 1–5, 2023
- [5] R. Gelman, W. Stevenson, C. Prospero Ponce, D. Agarwal, and J. B. Christoforidis, "Retinal damage induced by internal limiting membrane removal," Journal of Ophthalmology, vol. 2015, no. 1, p. 939748, 2015.