WavPool: A New Block for Deep Neural Networks

New Perspectives 2023, Fermilab

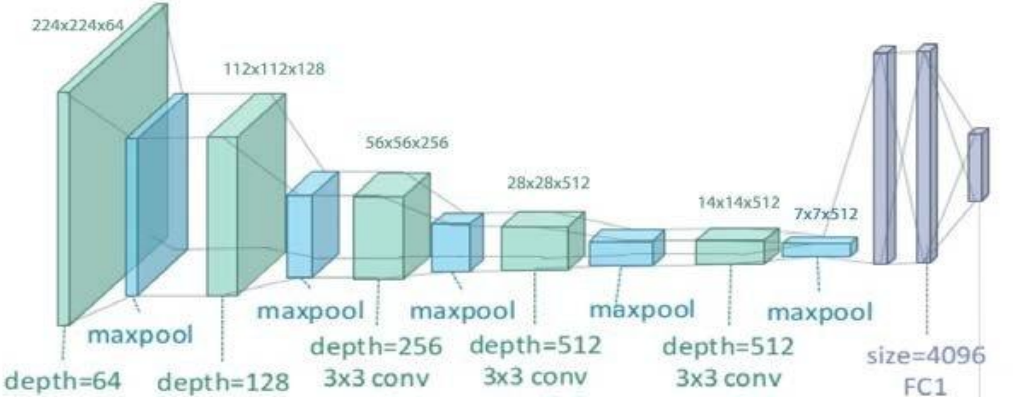
M. Voetberg[1]
Sam D. McDermott[2], Brian Nord[1234]

1. Fermi National Accelerator Laboratory, Scientific Computing Division, Batavia, IL
2. Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL
3. Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL
4. Laboratory for Nuclear Science, MIT, Cambridge, MA

SLIDES-23-100-CSAID

Why Mess With Perfection?

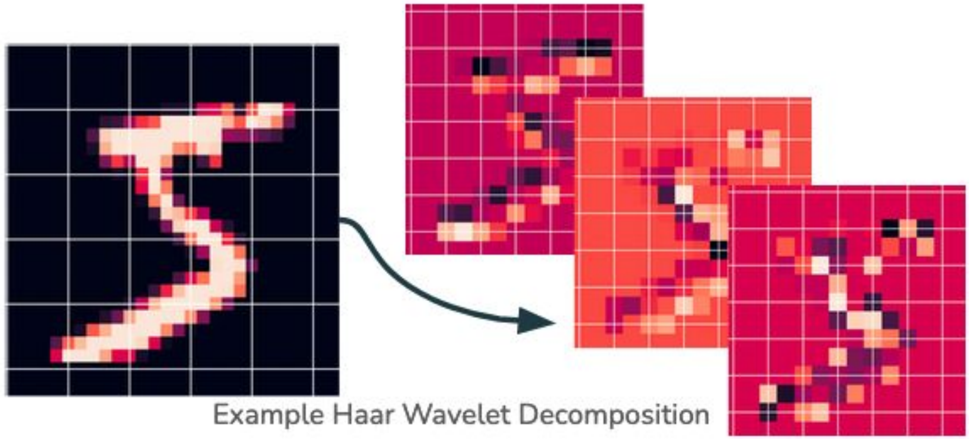
- Trained **kernel convolutions** are the back-bone of most modern neural networks – ruling in image based tasks
- Require large **inefficient blocks to encode spatial data**
  - Trained kernels convolve over inputs to produce large blocks in the model latent space, and explode the size of the model



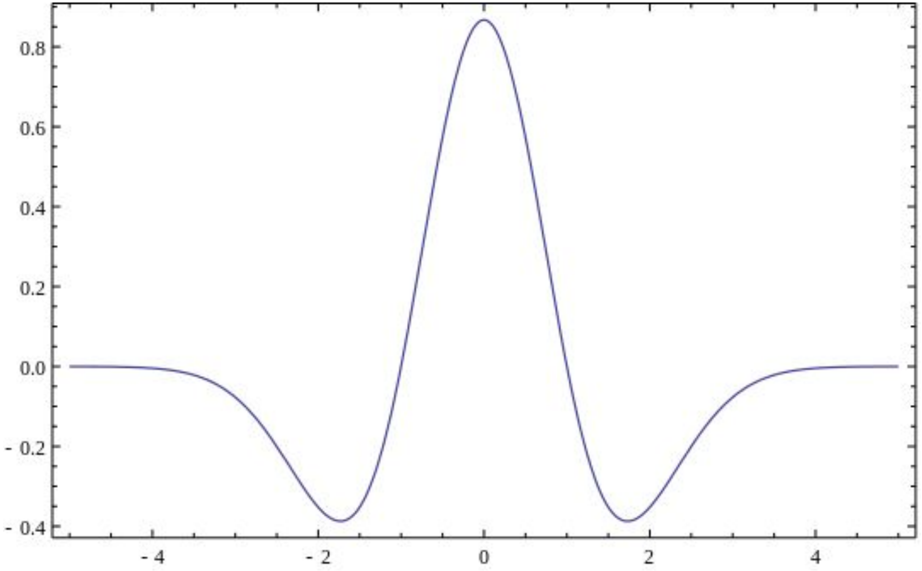
VGG-19, containing 138 million parameters for a 224x224 image

Wavelets and You

- Wavelets commonly used in **signal decomposition tasks** (audio, image)
    - Consider them a generalization of a fourier transform
- Specific operation, convoluted with the input, to produce an encoded output.



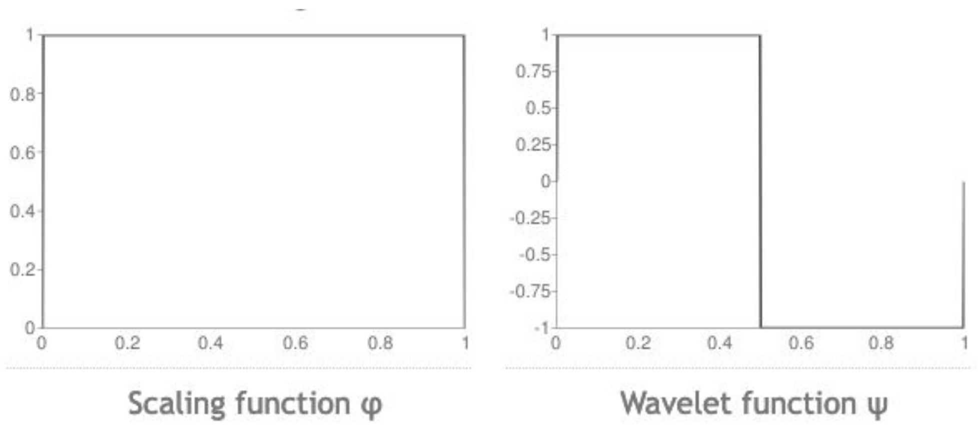
Example Haar Wavelet Decomposition



Example Wavelet - a continuous "Mexican Hat"

Further Background with Discrete Wavelets

- Mathematical operation that captures conjugate data – **sensitive to both position and size**
- Can **encode losslessly**; and without increasing the size of the data vector
- Composition of a “smoothing” (φ) and “differencing” (ψ) wavelet



Haar Wavelet
Source: https://wavelets.pybytes.com/wavelet/haar/

$$\phi_{1,1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix}, \qquad \psi_{1,1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \end{pmatrix}$$

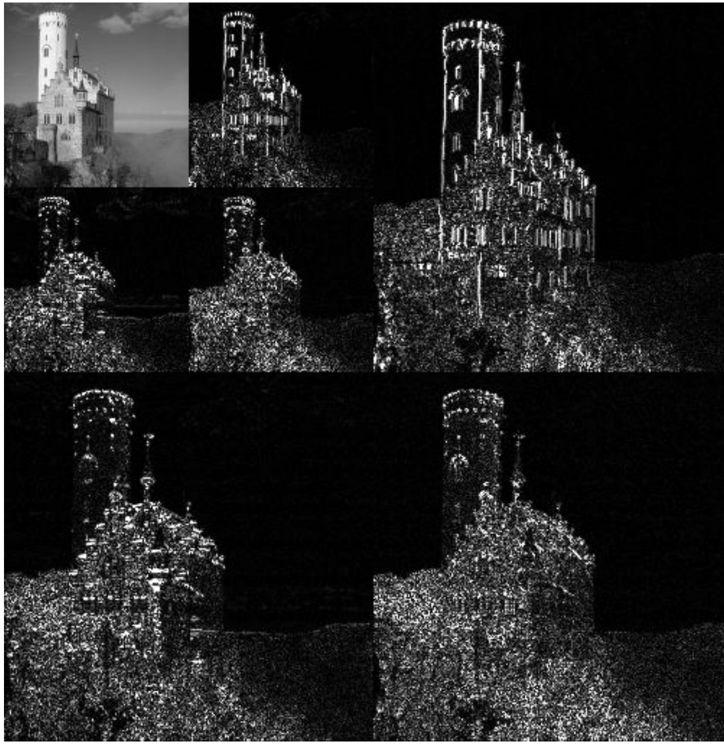
Using Multi-Resolution Decomposition

- Apply wavelets in sequence to **decompose the image at multiple levels**
  - Further decompositions produce a smaller image with less detail
  - Higher L → Less detail
- Use a different sign of ψ^a to **capture vertical, horizontal, diagonal signals**

$$C_\ell(S) = \phi \circ \underbrace{(\phi \circ (\cdots S))}_{\ell\,\text{times}},$$

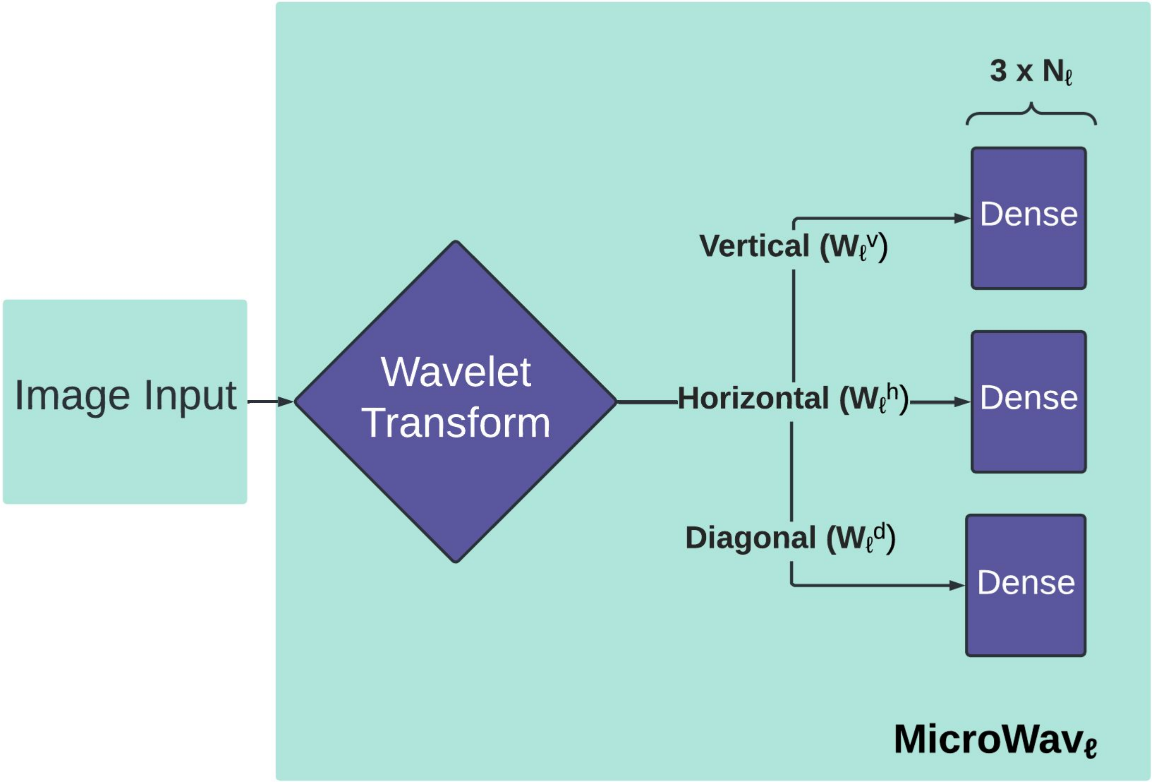
$$W^a_{\ell+1}(S) = \psi^a \circ C_\ell(S) = \psi^a \circ (\phi \circ \underbrace{(\phi \circ (\cdots S))}_{\ell\,\text{times}}).$$

Equation for a Wavelet (W) for direction (a) at level (l+1) over signal (S). The smoothing wavelet (φ) is applied l times to the signal.



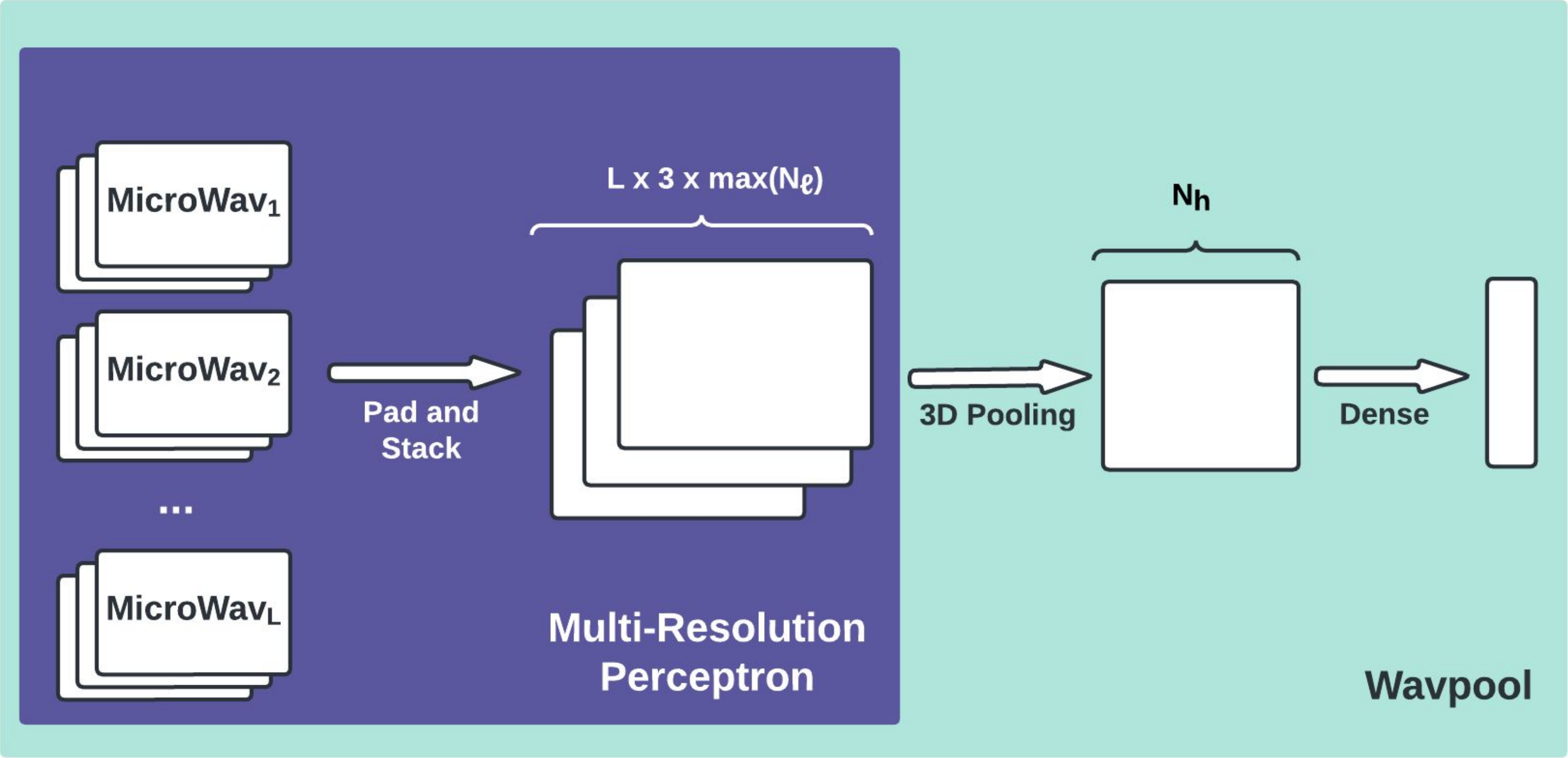
Decomposition using JPG2000, a form of MRD (Source: https://en.wikipedia.org/wiki/Wavelet_transform)

Introducing MicroWav



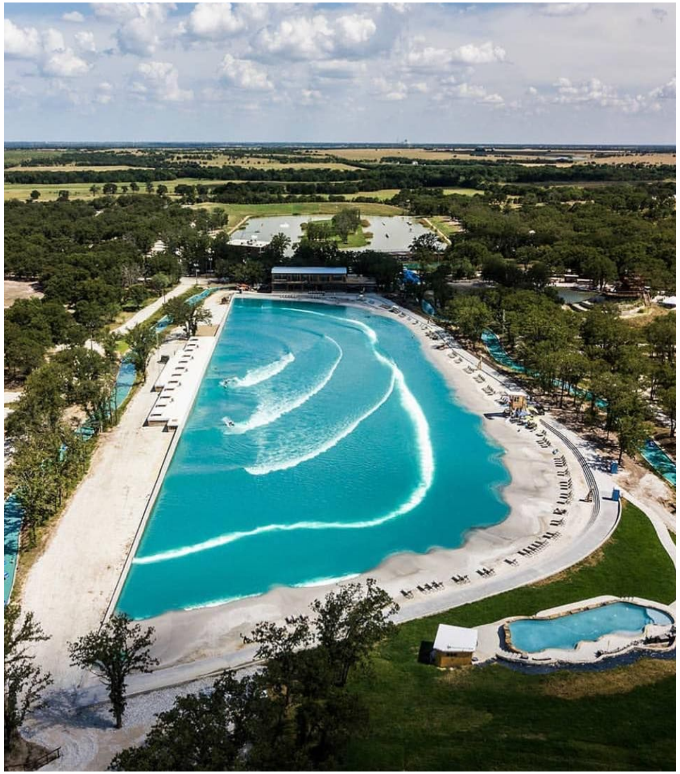
- Utilize the **multiple 'levels' of the MRD** using a Haar wavelet
- Introduce a layer with a 3 tailed output (vertical, horizontal, diagonical)
- Learns the features of the **3 decompositional features independently**

Stacked MicroWavs - WavPool

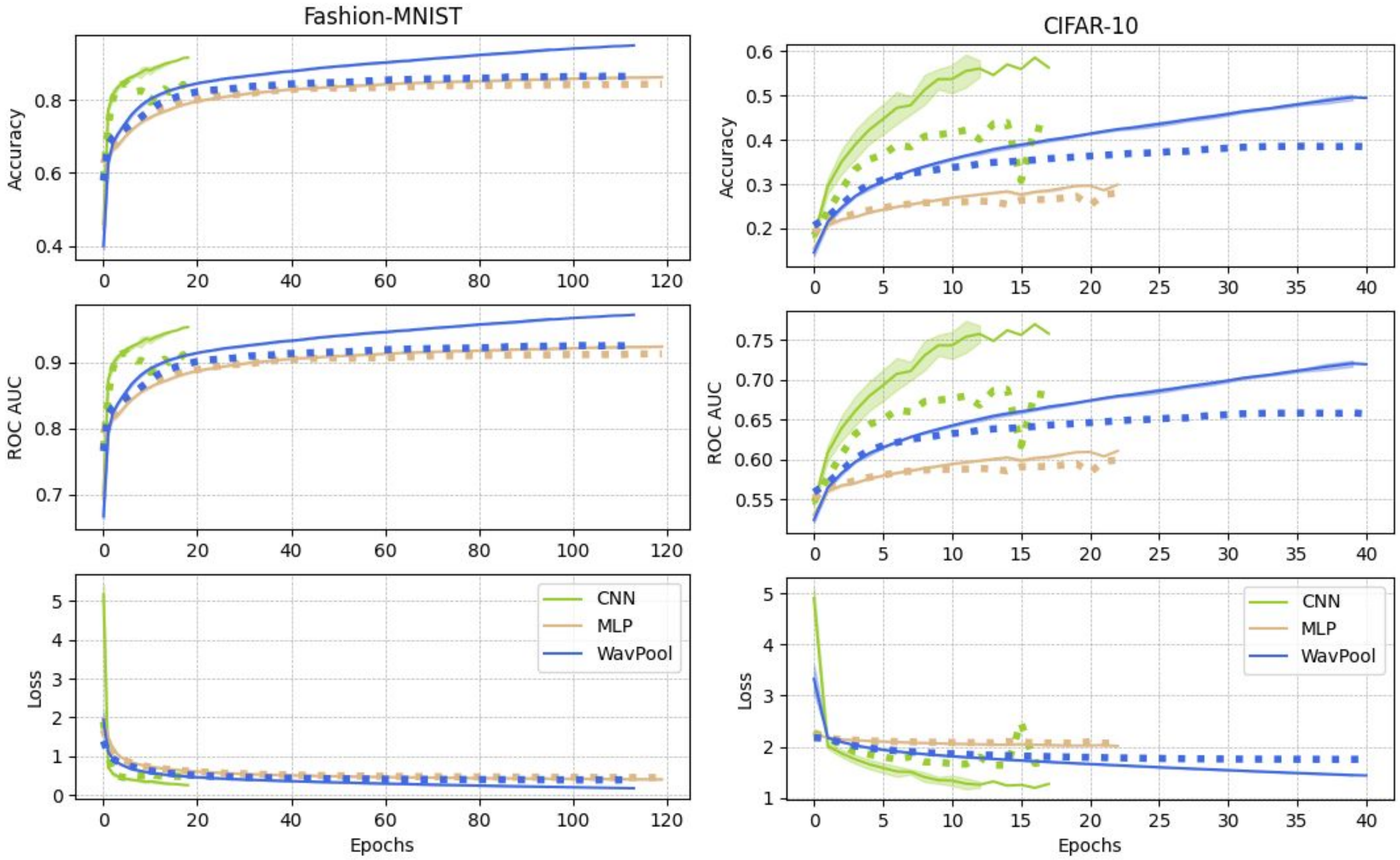


Why WavPool Works

- Wavelets give largely sparse representation of signals
  - **Easier signal to learn**
- Each layer has access to **spacial and size information**
- The transform decomposes images to smaller inputs – The network needs **less dense nodes to encode** them
- The wavelet-dense calculation is less computational intensive than a 3D convolutional block ($O(n+3n^3)$ vs ($O(m_{in}m_{out}n^4)$))



Comparisons



- Trained networks of comparable size and complexity to the WavPool block on benchmark data
- WavPool produces **more consistent solutions**
- Quality far above a comparable MLP, with **fewer parameters**

- Paper - https://arxiv.org/abs/2306.08734
- Code
  - GitHub - https://github.com/deepskies/DeepWavNN
  - PyPi - https://pypi.org/project/wavpool/