Mathematical Cardiology Lecture 10

Local image filters and learned classification

Zan Ahmad

Fall 2023

Specify the output

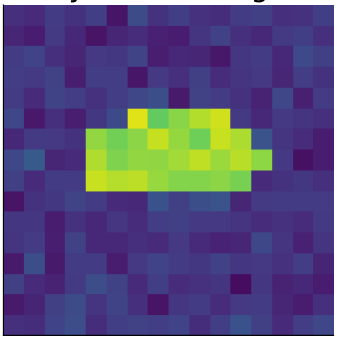
$$I \in \mathbb{R}^{16\times 16}.$$

Classification returns one label: wide or tall. Segmentation returns a label at each location. A correct whole-image label does not establish a correct boundary.

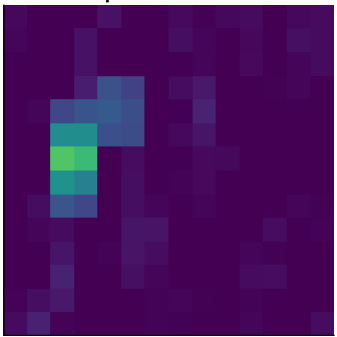
A fixed directional kernel

Fixed filters: no filter weights were learned

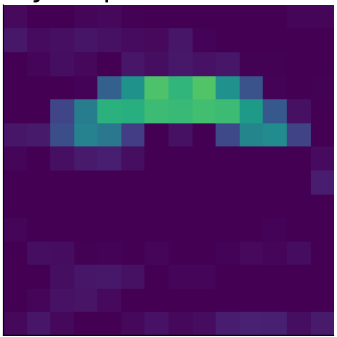
Synthetic image



x-response + ReLU



y-response + ReLU



A fixed directional kernel

$$K_x = \frac{1}{3}\begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}, \qquad K_y = K_x^T.$$

Negative weights are on one side and positive weights on the other.

Calculate one response

For a patch with three identical rows $(0, 0, 1)$,

$$A^x = 3\,[(-1)(0)/3 + 0(0) + (1)(1)/3] = 1.$$

Rows $(1, 0, 0)$ give -1; a constant patch gives zero.

Reuse the local calculation

$$A^x_{ij} = \sum_{a=0}^{2} \sum_{b=0}^{2} (K_x)_{ab} I_{i+a,j+b}.$$

The same kernel is used at every valid window. The code does not reverse the kernel: this is a sliding dot product.

Count the valid window starts

With size N, kernel width k, stride one and no padding,

$$N_{\text{out}} = N - k + 1.$$

$$16 - 3 + 1 = 14.$$

Each output map is 14×14.

Retain only positive responses

$$B_{ij}^{x} = \max(A_{ij}^{x}, 0).$$

ReLU preserves positive rises and discards negative responses. For supplied intensities in $[0, 1]$, $0 \leq B_{ij}^{x} \leq 1$.

Pool the maps into two numbers

$$f_x = \frac{1}{196} \sum_{ij} B^x_{ij}, \qquad f_y = \frac{1}{196} \sum_{ij} B^y_{ij}.$$

The average keeps a response magnitude and loses its location.

An exact transpose check

The kernels satisfy $K_y = K_x^T$. Transposing the image exchanges the pooled features:

$$\boxed{f(I^T) = (f_y(I), f_x(I))}.$$

A wide ellipse becomes tall relative to the axes.

Fit the output rule

From training images only,

$$c_k = \frac{1}{n_k} \sum_{i:y_i=k} f(I_i).$$

Predict the nearer center in squared Euclidean distance. No feature-standardization fit is used in this pipeline.

Weight sharing is not training

One 3×3 kernel has nine weights. They are reused at 196 locations. Separate weights at each location would require

$$196 \cdot 9 = 1764$$

weights for one output map, before biases.

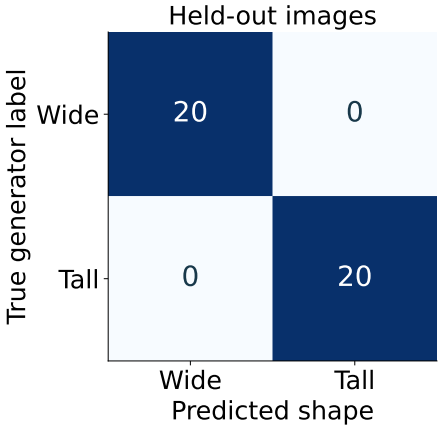
A learned convolution would update the kernels

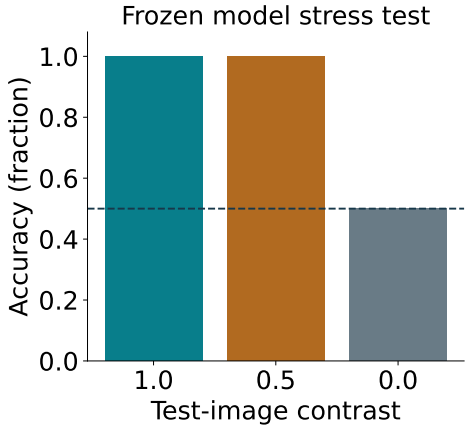
One possible extension has score

$$s_i = w^T f_K(I_i) + b$$

and loss $\frac{1}{n}\sum_i (s_i - y_i)^2$. Training would update selected parameters, including K. The supplied filters remain fixed.

The held-out image counts





The held-out image counts

true\predicted	wide	tall
wide	20	0
tall	0	20

Forty correct predictions on this synthetic test.

Remove the signal without refitting

contrast	1	0.5	0
accuracy	1	1	0.5

At zero contrast the intended ellipse signal disappears. Report all tested levels, not the most favorable one.

An idealized contrast calculation

Let $I' = aI + b\mathbf{1}$, $a \geq 0$, without clipping. Zero-sum filters give $A(I') = aA(I)$. ReLU and averaging then give

$$f(I') = af(I).$$

The frozen class centers do not scale automatically.

An augmentation must respect the target

A ninety-degree rotation swaps wide and tall relative to the axes. Keeping the old label would create a contradictory training example.
A contrast test preserves the intended orientation target while changing its visibility.

The fitted claim and its next test

The current claim concerns synthetic ellipse orientation and a fixed observation process.
A new-patient claim requires patient separation, defensible labels and an appropriate independent test.
More learned weights cannot replace those conditions.