

IEEE Applied Imagery Pattern Recognition Workshop

Session II, Deep Learning, Session Chair: Moody

October 10, 2017

The Cosmos Club, Washington, DC

TITLE: *The Anatomy of a Neural Network*

James P. LaRue, PhD, Jadco Signals (contact: James@JadcoSignals.com) (presenter)

Richard L. Tutwiler, PhD, ARL- Penn State University(Professor Emeritus)/CEO LiveMotion3D LLC

Dennison J. LaRue, Biomedical Engineering, Clemson (2017), Freshman, University of South Carolina Law School

The Anatomy of a Neural Network in Four parts:

Part 1 – Historical path to a (Bi-directional) Convolutional Neural Network,

Part 2 – ‘Typical’ visualization of Objects in Deep CNNs, limited to
‘Convolutional’ Layers 1-3

Part 3 – Our Anatomical point of view looks inside the...

‘Black Box’



- *a little deeper*, by examining Time-Vector representation of each individual neuron comprising the Object, and showing individual neuron convergence, dead or alive
- *a little farther*, covering all Layers 1,2,3,4,5,6,..., (convolutional and decision layers)

Part 4 – Bi-directional Associative Memory to Bi-directional Neural Network

- add a JPAT for intra-layer and inter-layer connections
- develop RAINNE, a Real and Imaginary Neural Network Emulator

Part 1: Historical Path to a Convolutional, and Bi-directional, NN

1. Neural Network - NN

Alexander Bain (1873) and William James (1890)	<i>Neurons interact.</i>
McCulloch and Pitts (1943)	<i>1st computational model.</i>
Rosenblatt (1958)	<i>Feed-forward perceptron, convergence issue.</i>
Werbos (1975)	<i>Fixes the Rosenblatt problem, goes unrecognized.</i>
Rumelhart, Hinton, Williams, McClelland (1984)	<i>Recognize Werbos work.</i>

2. Convolutional Neural Network - CNN

Fukushima (1980)	<i>Neocognitron – hidden layer visual pattern recognition.</i>
LeCun and Bengio (1995)	<i>Convolutional ‘Neocognitron’ (CNN), long training.</i>
Krizhevsky, Sutskever, Hinton (2012)	<i>ImageNet Deep Classification, courtesy of a GPU.</i>

3. Bi-directional Associative Matrix - BAM

Kosko (and Grossberg) (1988)	<i>Bi-directional I/O matrix, no hidden layers, stability issue.</i>
------------------------------	--

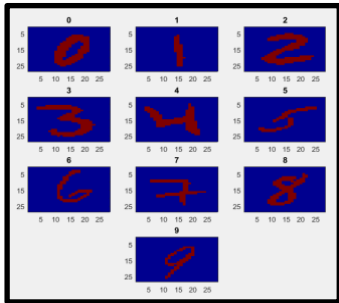
Couple of Comments

Minsky and Papert (1969)	<i>Need at least one hidden layer between I/O to be meaningful.</i>
Cybenko/Hornik (1989)	<i>Universal Approximation Theorem (UAT), one single hidden layer.</i>

4. Bidirectional Neural Network – BNN

LaRue/Luzanov 2011 - AFRL /AFOSR	<i>Stable intra-layer BAM model of CNN.</i>
LaRue/Copeland 2012 - DARPA Innovation House	<i>Met UAT criterion, applied to VMR project.</i>
LaRue/Suarez 2013 - JADCO Signals	<i>Formed inter-layers (two CNNs), Patent app. (Quadri-directional)</i>

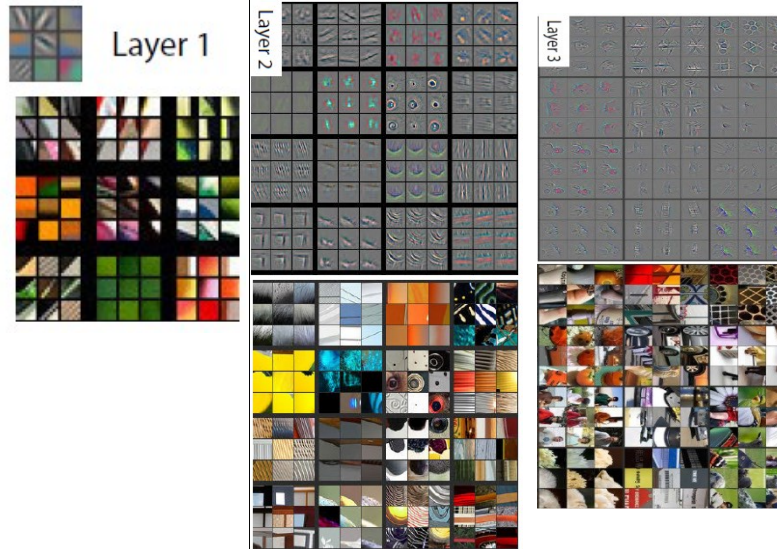
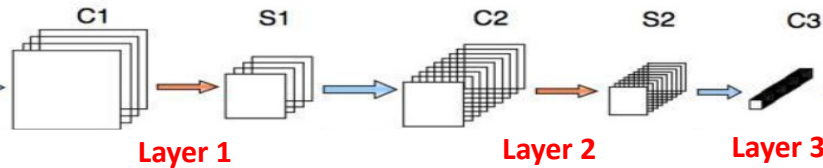
Part 2: A Typical Visualization of: L0, L1, L2, L3



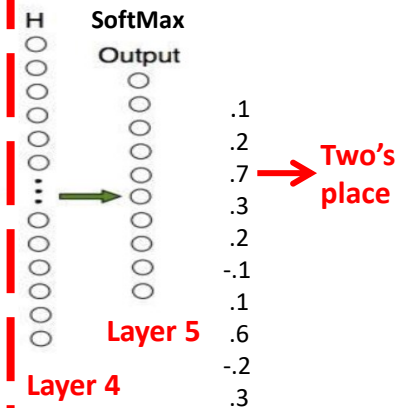
LeCun's MNIST
0,1,2,3,4,5,6,7,8,9



Image
Layer 0



Visualizing and Understanding
Convolutional Networks
Matthew D. Zeiler and Rob Fergus
NYU, Dept of Computer Science



No visualization on
NN decision side by
Zeiler and Fergus

96 learned low-level filters

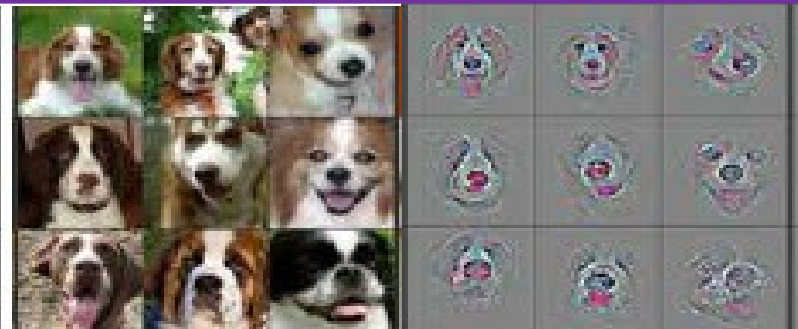


ImageNet Classification with Deep
Convolutional Neural Networks

Krizhevsky, Sutskever, Hinton
University of Toronto

2012 ImageNet Benchmark using GPUs
(dislodging Lowe's SIFT)

Layer Extract
of
Zeiler/Fergus
(just too cute)



Example: Typical Layer 1 visualization with code



for $i = 1:12$ Gabor filters

Apply edge detectors (convolve)

```
a1=ImageIn;  
a2= Gabor #;  
convresults = conv2(a1,a2);
```

Apply activation

```
BoundResults = tanh(convresults);
```

Subsample Once

```
b1=BoundResults(1:2:end,:);  
b2=BoundResults(2:2:end,:);  
Sub1 = b1+b2;
```

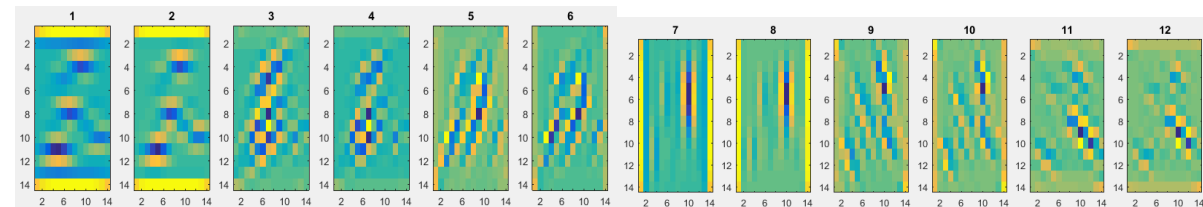
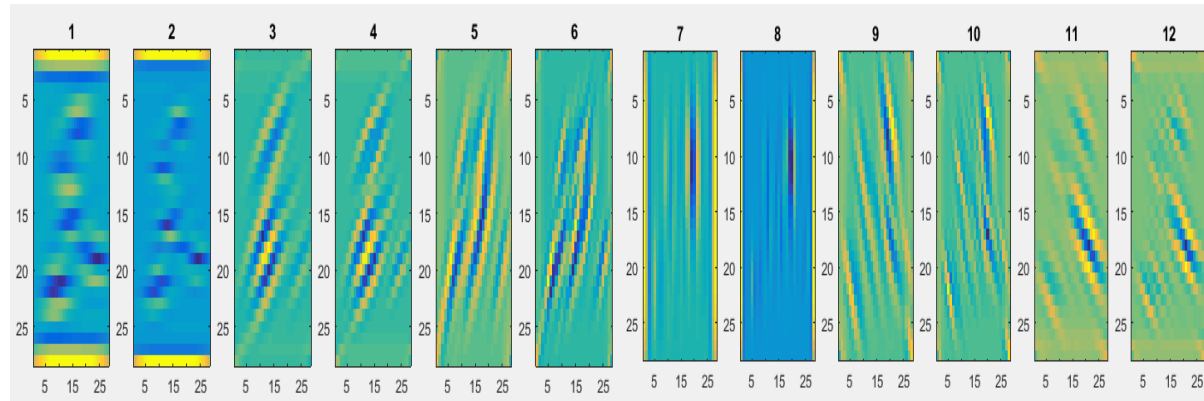
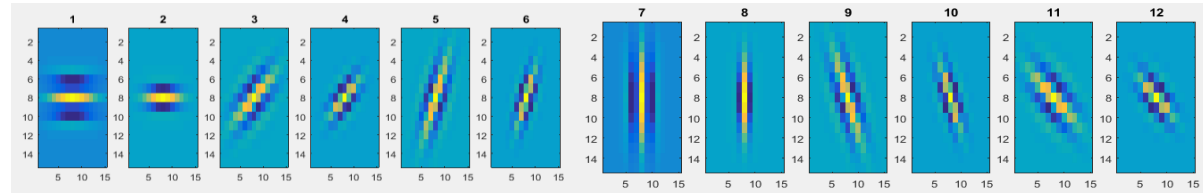
Subsample again

```
b3 = sum12(:,1:2:end);  
b4 = sum34(:,2:2:end))/4;  
sub2 = (c1+c2)/4'
```

Apply weight, bias, and activation

```
temp= subsum2(i)*wgt(i) + biasV1(i)  
Layer1 Out (i)= tanh(temp)
```

End of Layer 1



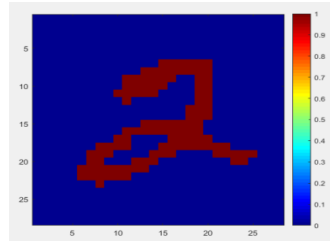
Pause for a moment and talk

Our Anatomical point of view drills down...

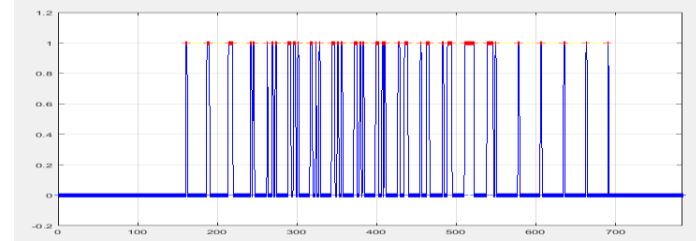
- *a little deeper*, by examining Time-Vector representation of each individual neuron comprising the Object, and **showing individual neuron convergence**, dead or alive
- *a little farther*, **covering all Layers 1,2,3,4,5,6,....**, (convolutional and decision layers)

Part 3: *Our Anatomical point of view begins with,* *Treating each point in the Convolutional Object of interest,* *As a separate neuron with its own unique response*

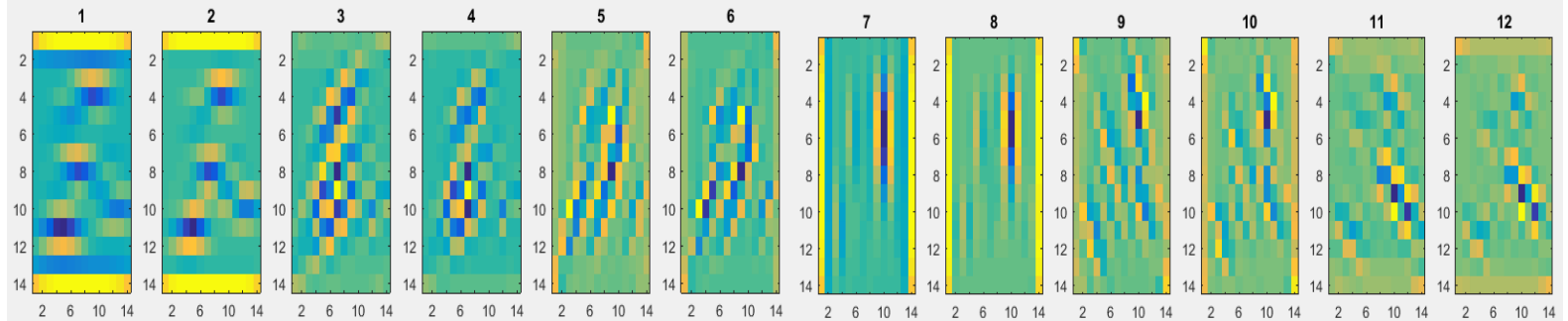
Image
To
vector of
neurons



Reshape
28 x 28 image
to
784 neurons

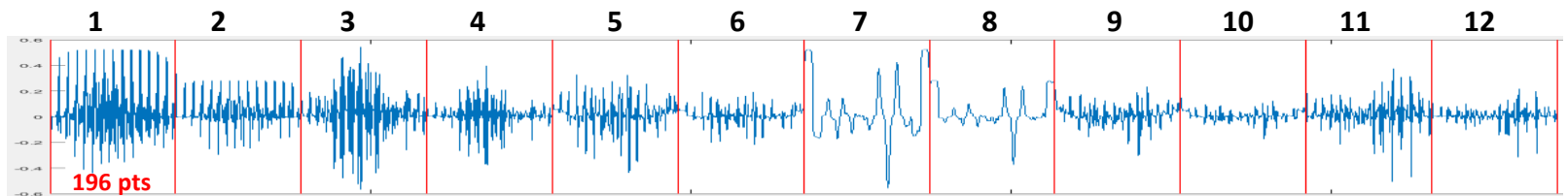


Layer 1
Filtered
to
12 sets of
14 x 14 's



We have $12 \times (14 \times 14) = 12 * 196 = 2352$ 'individual firing neurons'

Image(s)
to
vector of
neurons



1 $\longrightarrow$ $14 \times 14 \times 12 = 196 \times 12 = 2352$ pts. $\longrightarrow$ 2352

Example: Layer 2

```

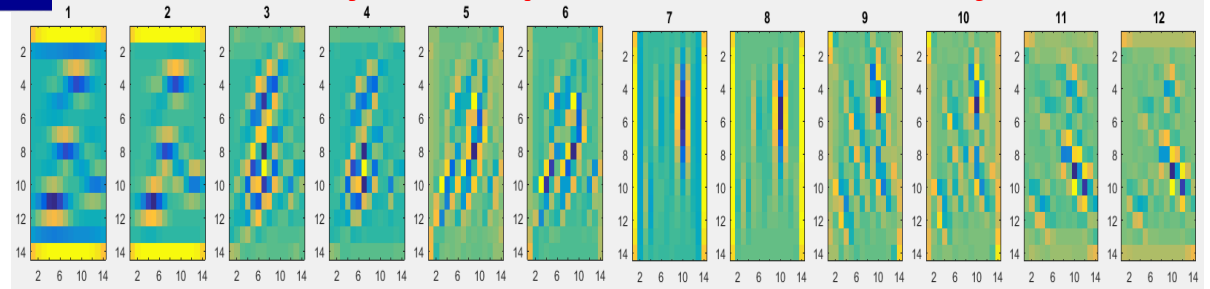
for i=1:nOMsV2 % 16 neurons here
    for j=1:connlistV2
        a5=OMV1(i).list(j)) % from Layer 1
        a6=KV2(connlistV2(i).list(j))
        convResults=temp + conv2(a5,a6);
    end
    % Apply the activation function
    temp = tanh(convResults);
    %subsample #1
    a7=temp(1:2:end,:);
    a8=temp(2:2:end,:);
    temp2 = a7+a8 ;
    %subsample #2
    x=temp2(:,1:2:end)
    y=temp2(:,2:2:end)
    temp3=(x+y)/2
    % bring in wgt, bias, activate
    temp4 = temp3*wgt2(i) + bias2(2,i)
    Layer2Out = tanh(temp4);
end
    
```

KV2 Conv Kernel

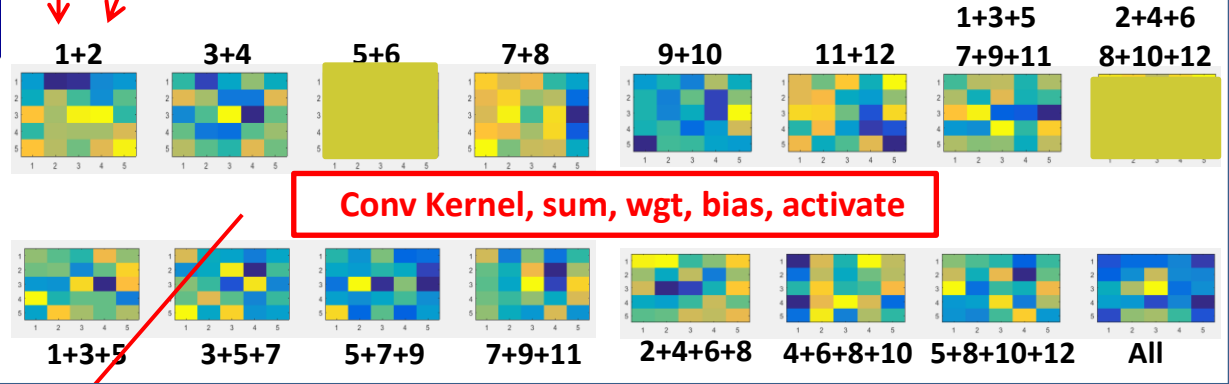
wgt, bias

activate

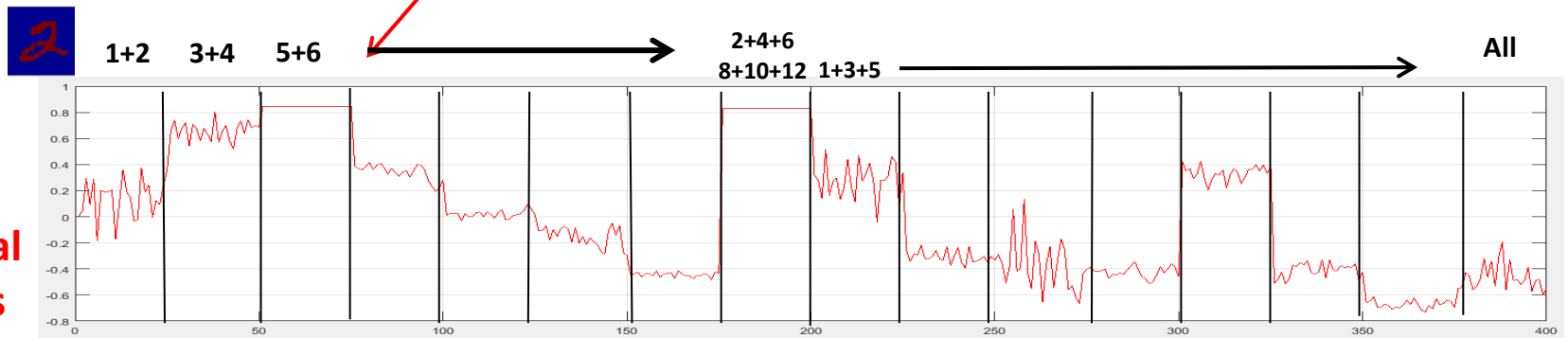
Recall: Layer 1 output of 12 14x14's as objects



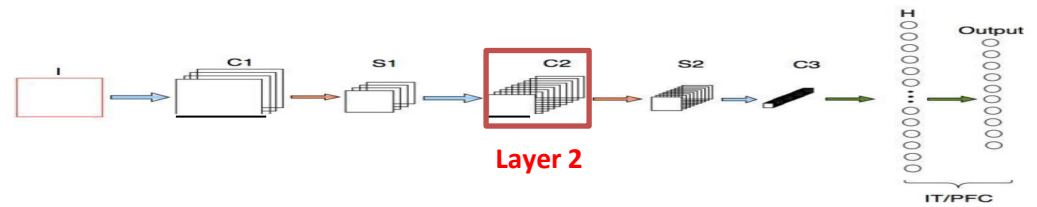
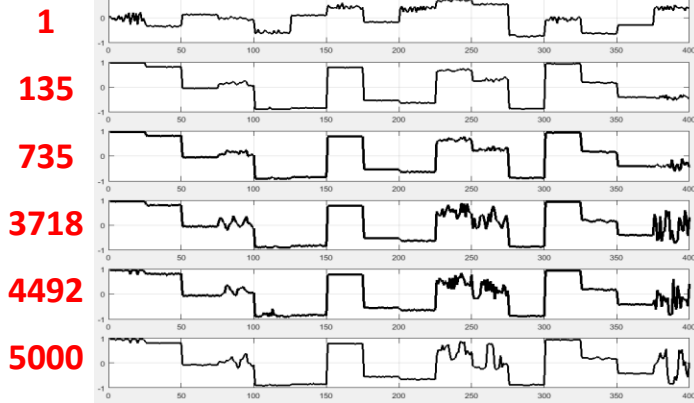
Layer 2 has 16 (5x5) outputs



16
5x5's
equals
400
individual
neurons
firing



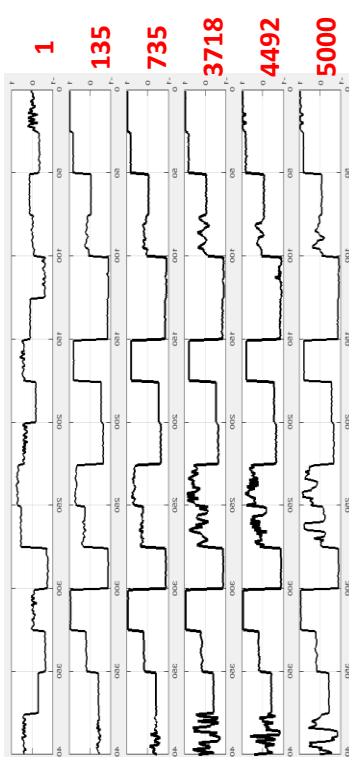
Step 1



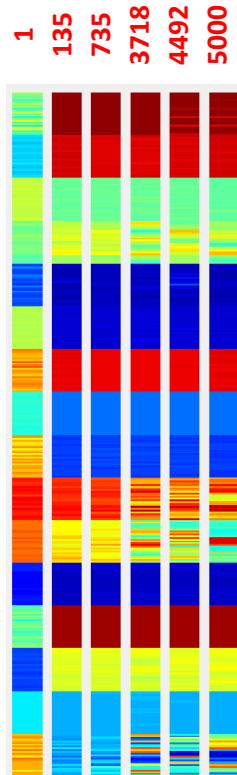
How to see **Layer 2** convergence over 5,000 iterations

1. Plots of iterations: **1, 135, 735, 3718, 4492, 5000**
2. Turn on side (rotation, reflection, (dihedral group))
3. Rotate out as matrix columns, color scale normalized to [- 1, 1]
4. Put all 5,000 together, and ***observe convergence left to right***

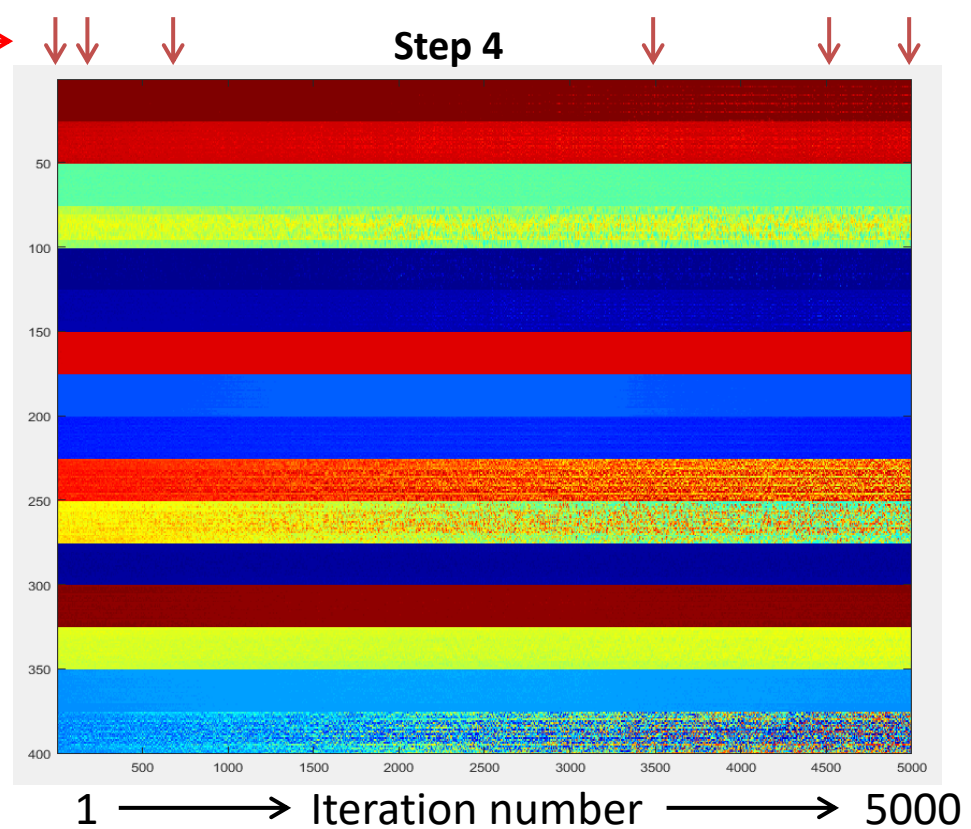
Step 2



Step 3

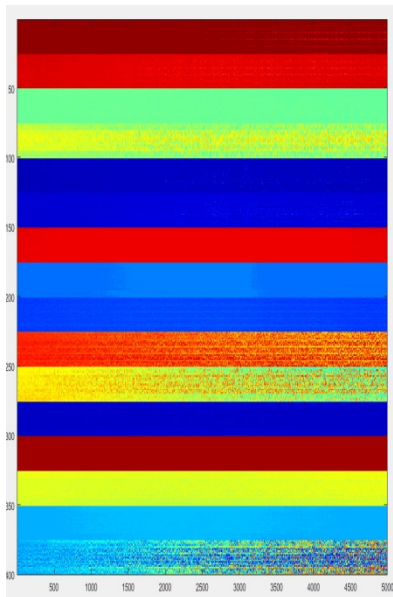


Step 4

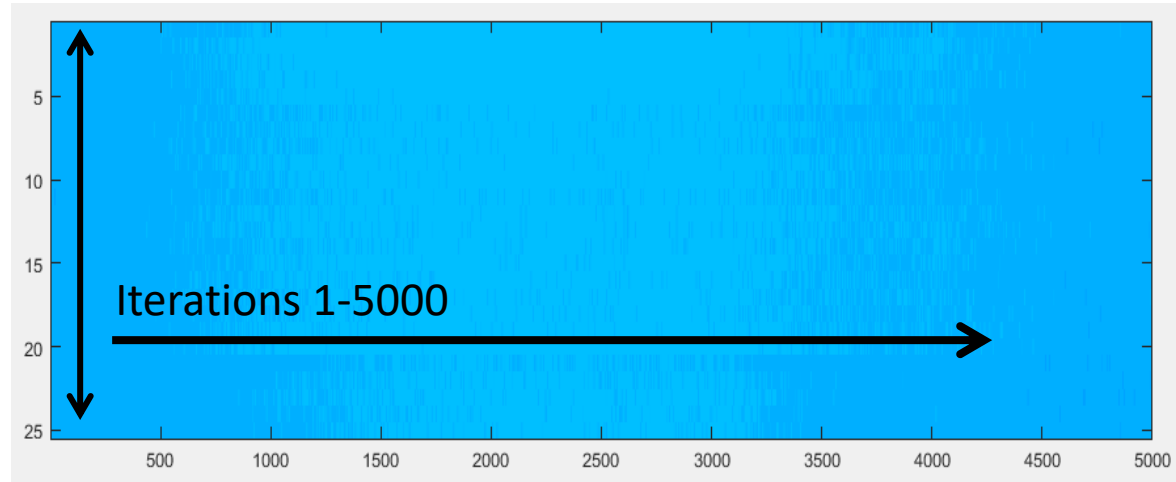


A Drill-down via Patch-2-Vector Neuron conversion across 5000 iterations, to observe neuron activity or, inactivity

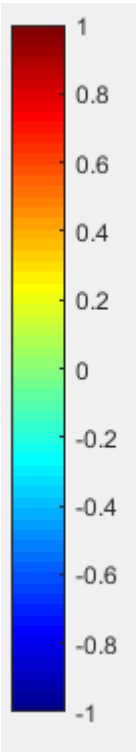
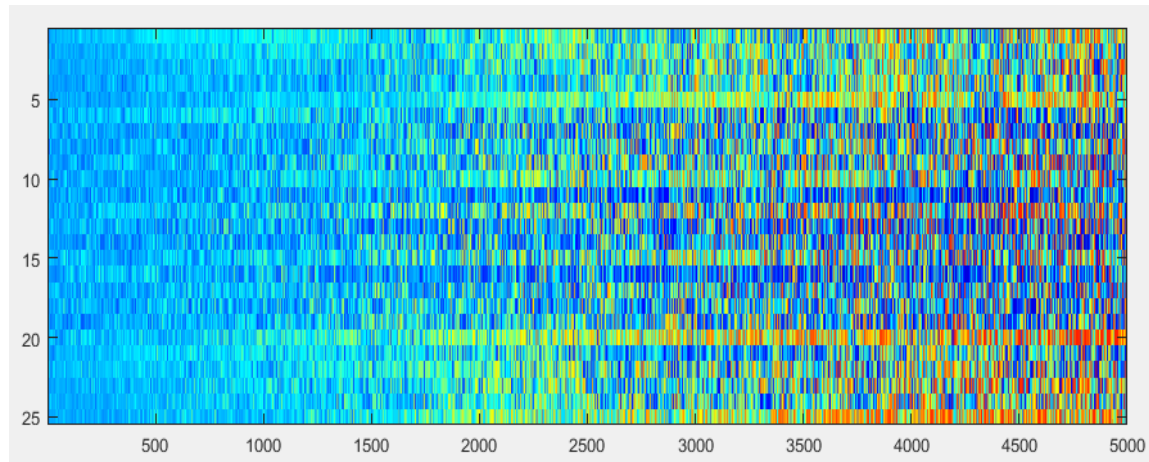
Layer 2
16 5x5 patches
Showing convergence
for Patches 15 and 16



Patch 15 of 5x5 = 25 neurons



Patch 16 of 5x5 = 25 neurons



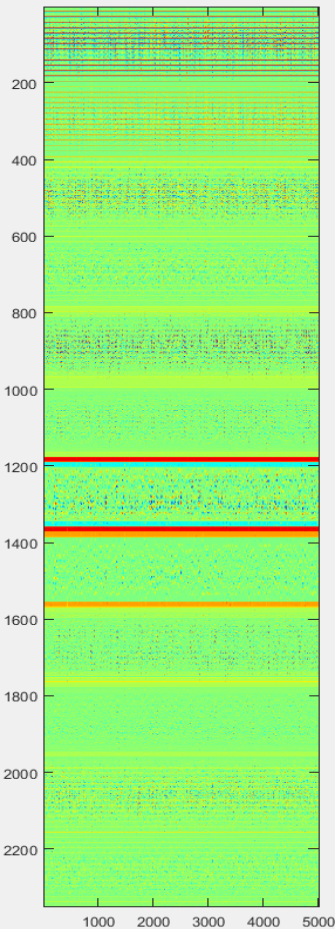
Visualization of Learning for all Layers

Pass in 5000 MNIST digits, 500 each of 0-9, randomly

Each matrix first column is Iteration output 1

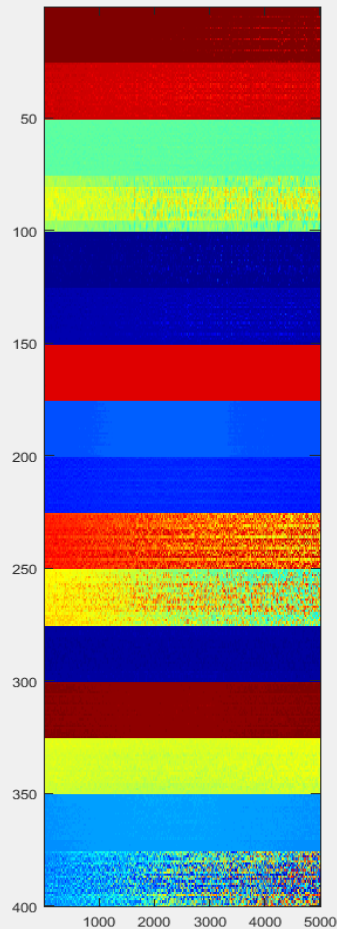
Each matrix last column is Iteration output 5000

Layer 1



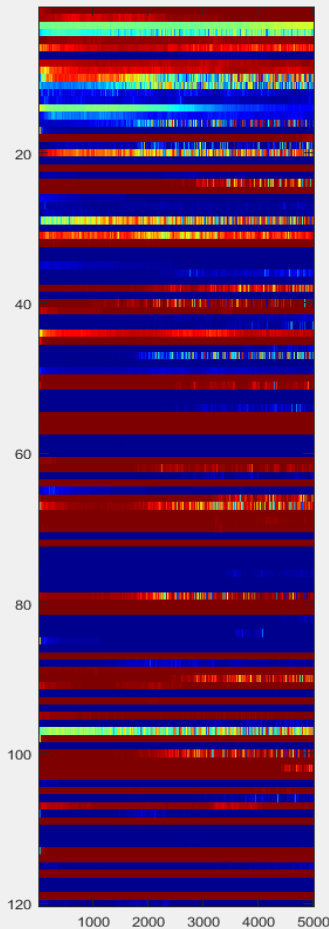
1 --> 5000

Layer 2



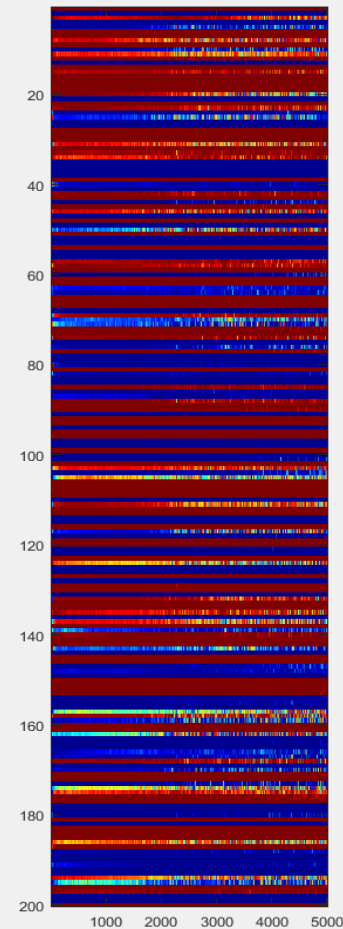
1 --> 5000

Layer 3



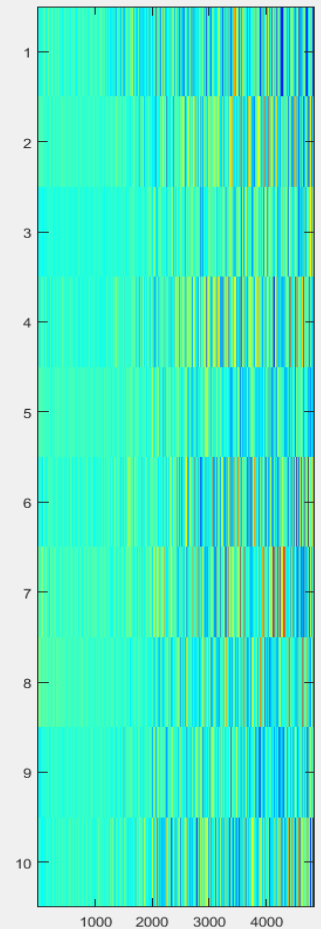
1 --> 5000

Layer 4



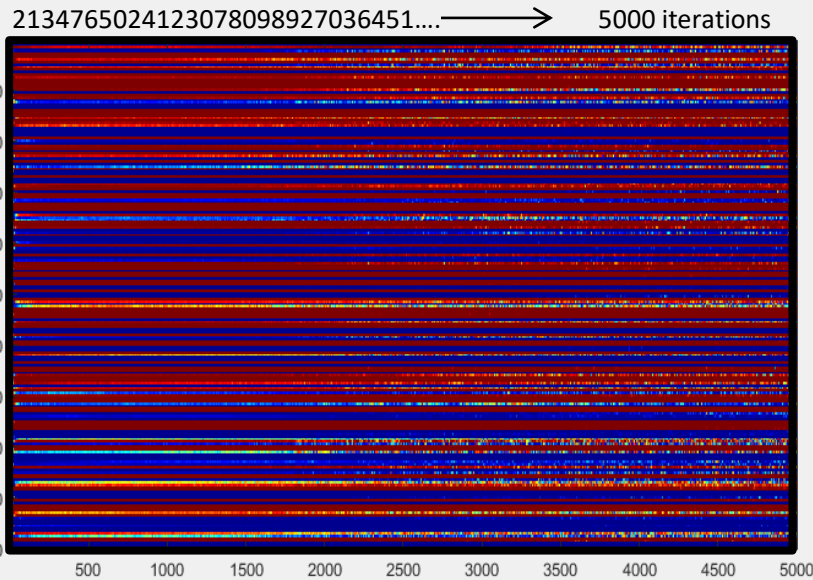
1 --> 5000

Layer 5

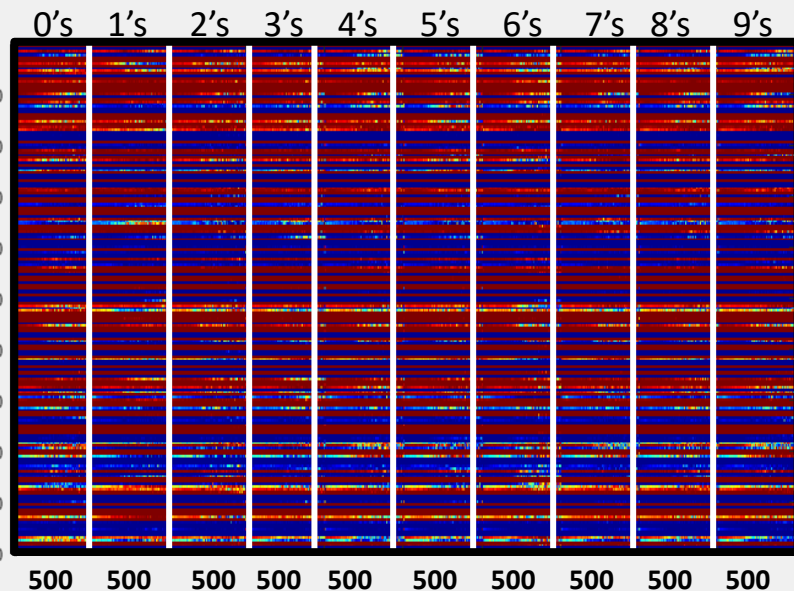


1 --> 5000

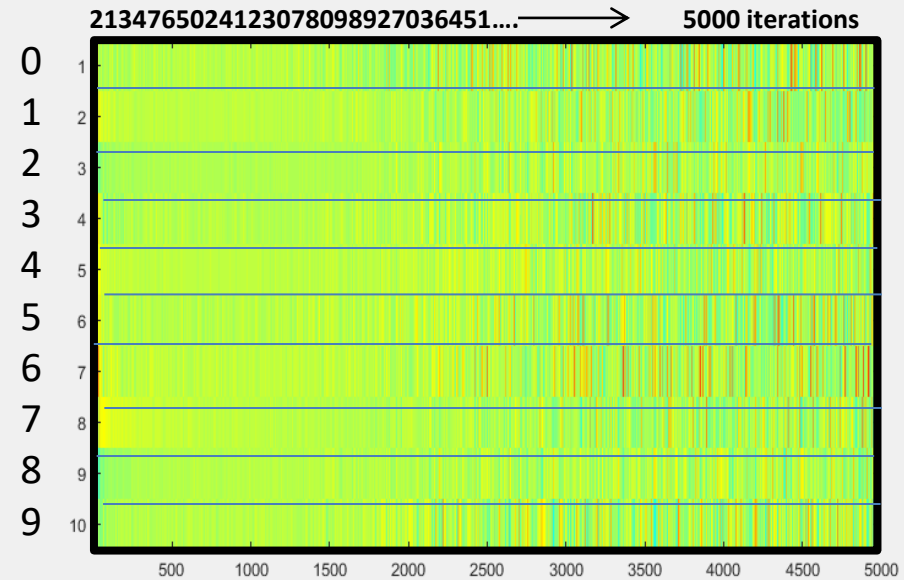
Layer 4



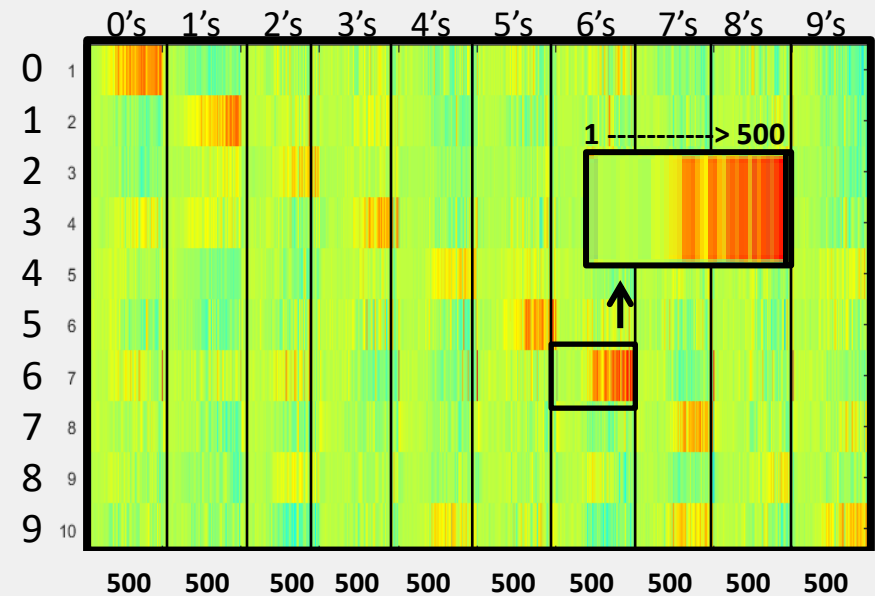
Grouped by



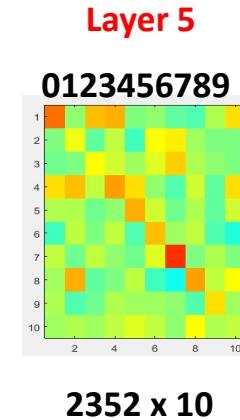
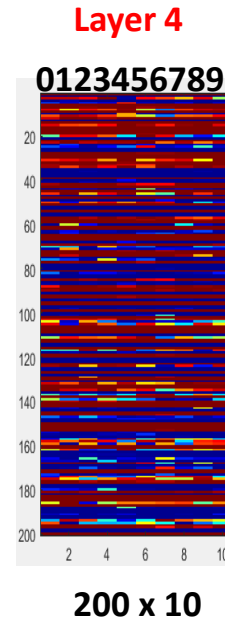
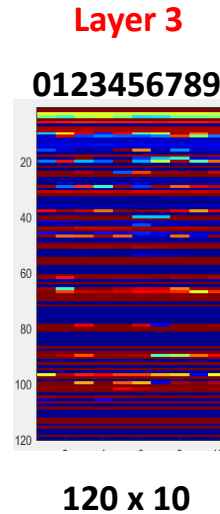
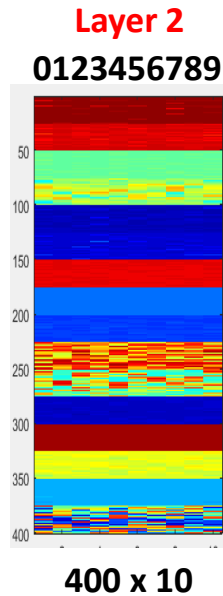
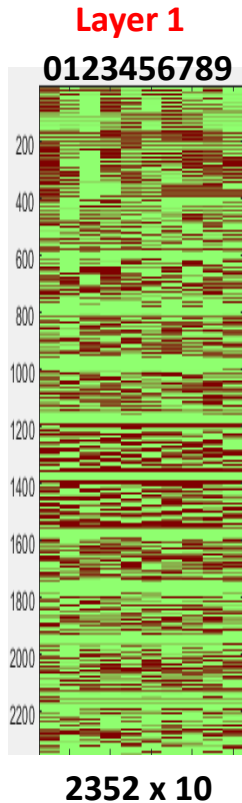
Layer 5 output across 5,000 training selections



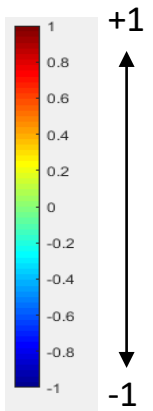
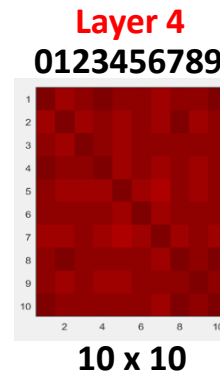
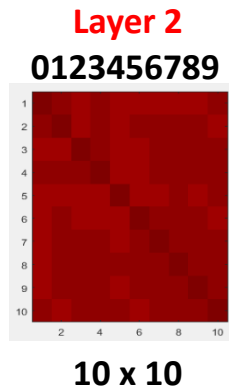
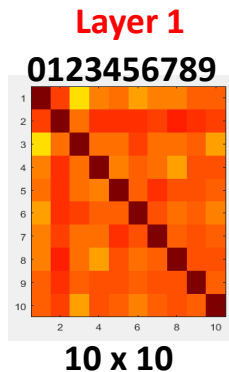
Grouped by digit (500 ea.) to see individual learning



After 5,000 training iterations, keep the last set of 0-9 object vectors from each Layer



Diagnosing Layer Health by Correlation Coefficient



Pause for a loaded question

We applied a correlation measure for Objects within the same layer (same dimension).

So,

can we make the same measure for Objects between different layers?

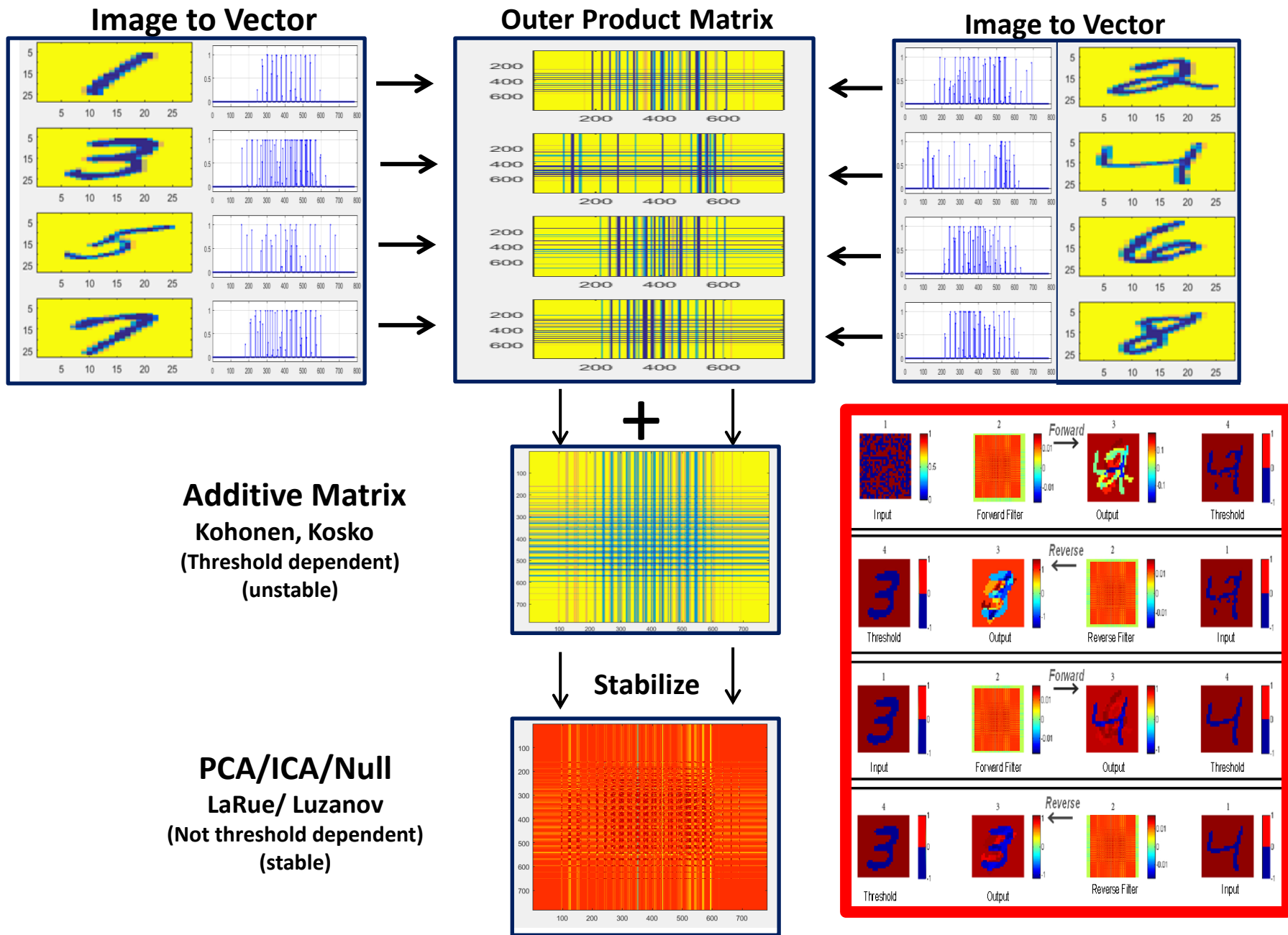
That is,

can we correlate a 2352 dimensional vector with a 400 dimensional vector?

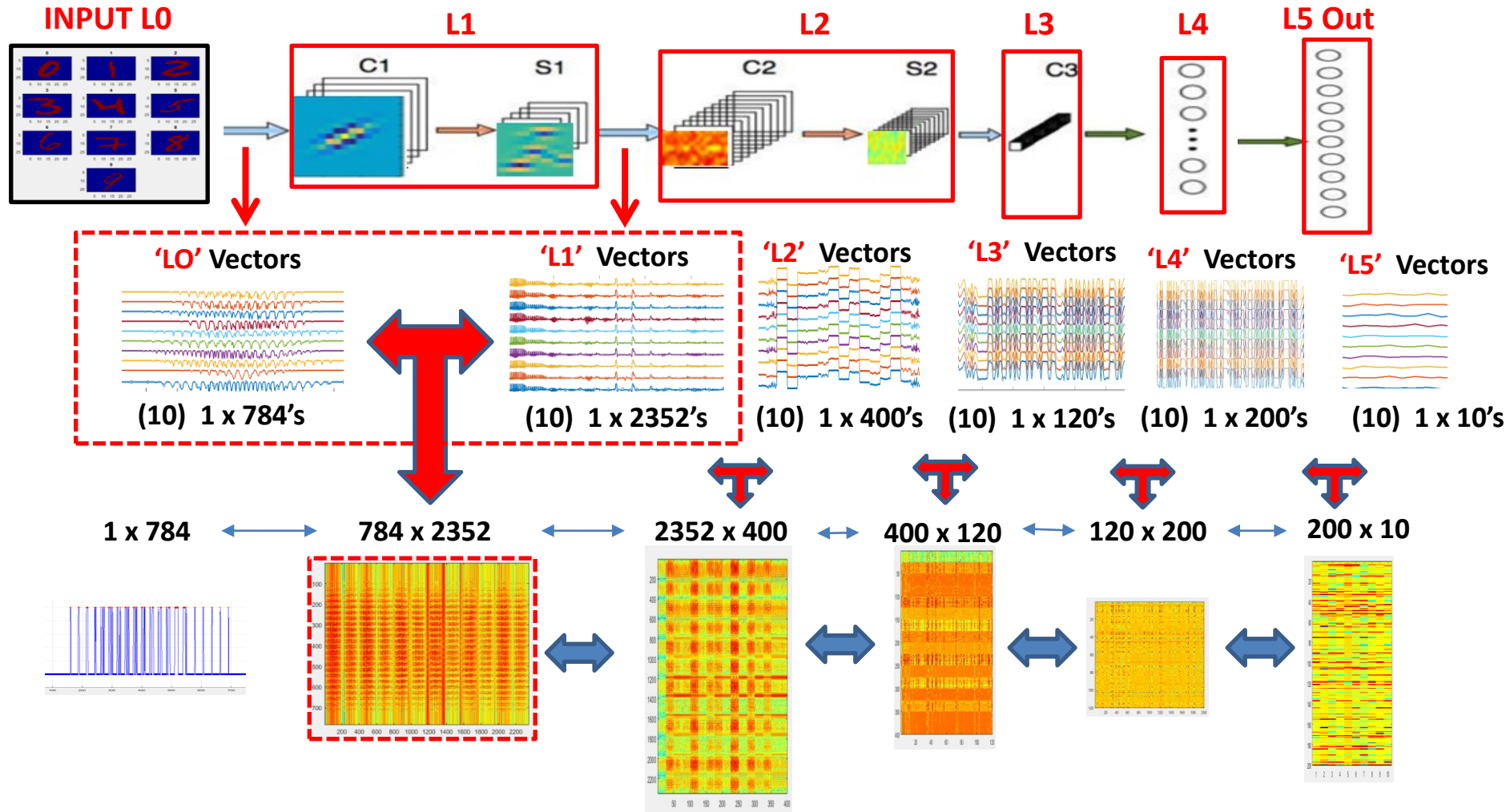
The answer is yes.

Its called a measure of Association

Part 4: Recipe for “Associating” and then Resonating

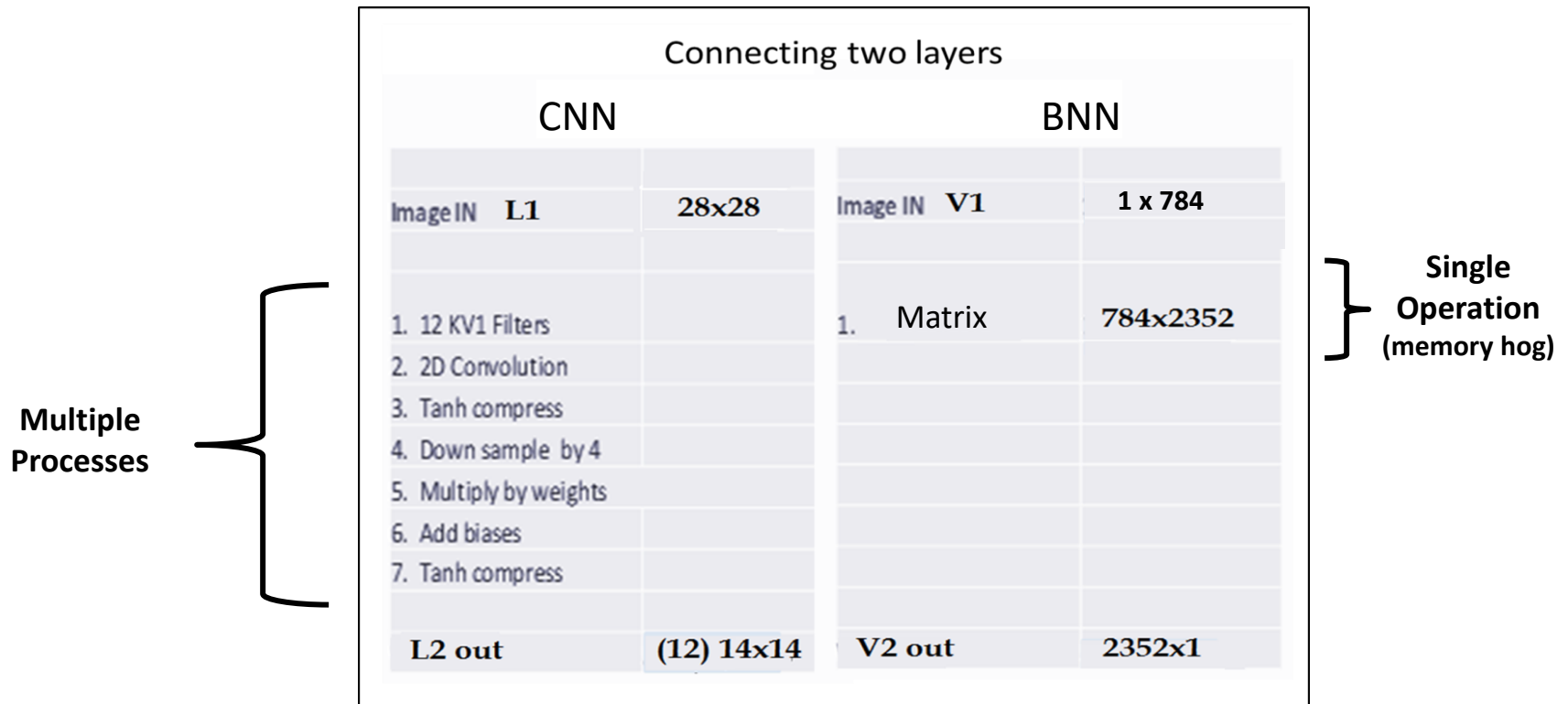


How we convert CNN I/O pairs into BNN I/O pairs



Note: $(784 \times 2352) \times (2352 \times 400) \times (400 \times 120) \times (120 \times 200) \times (200 \times 10) = (784 \times 10)$
 $(A \times B) \times (B \times C) \times (C \times D) \times (D \times E) \times (E \times F) = (A \times F)$

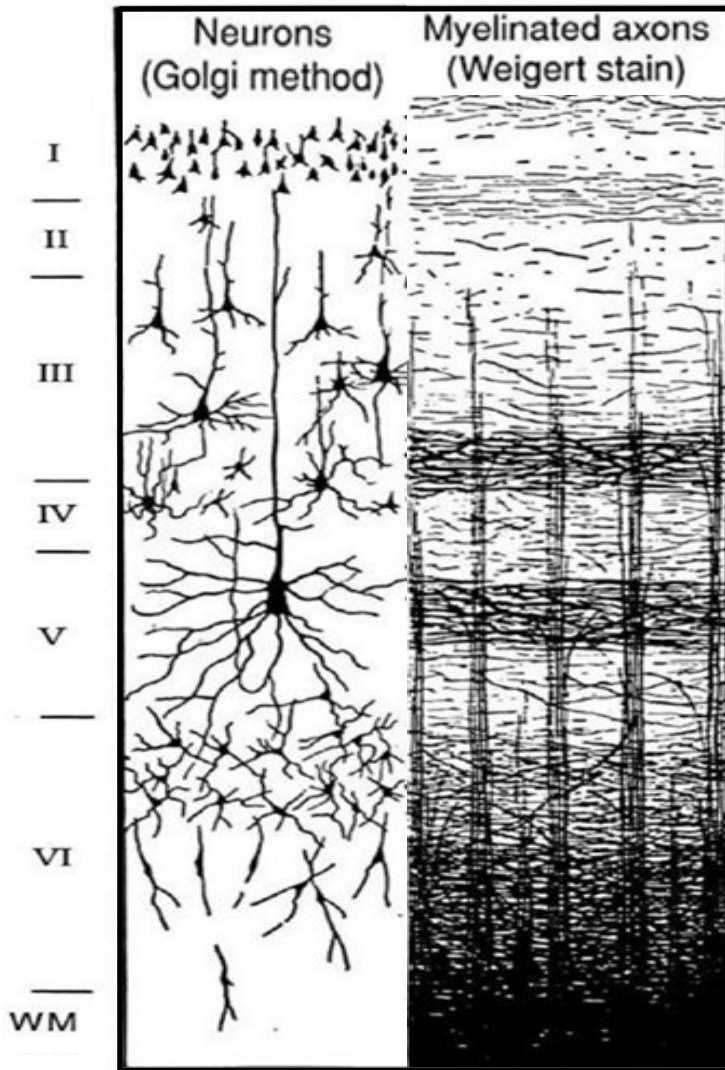
Emphasizing the difference between CNN Process and BNN Matrix



Guess which side is faster?

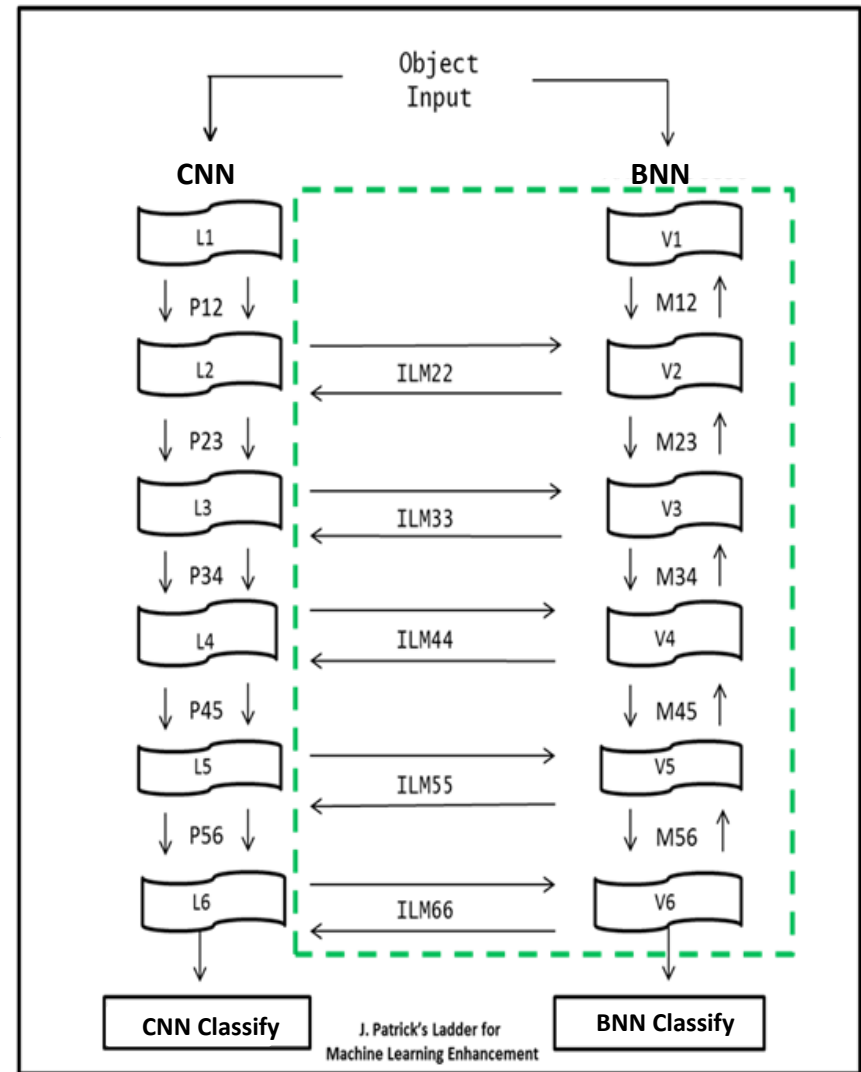
Bridging the gap with RAINNE and JPAT

and back to Alexander Bain (1873) and William James (1890)



Real and Imaginary Neural Network Emulation

L-Layer
V-Vector
P-process
M-Matrix
IL-Inter-layer



Vertical and horizontal connections
of the Ventral Pathways

Joint Proximity Association Template
for (quadri)-directional Neural Networks

Thank-You AIPR 2017 and Audience

Next-step: A Truer Anatomy of a Neural Network

RAINNE with Striatal Beat Frequency using Meditation Music

Request: Looking for 5 second recording of 16x16 EEG nodes, sampled at 4K

