



Hierarchical Neural Network

Deep Learning CS230 Project

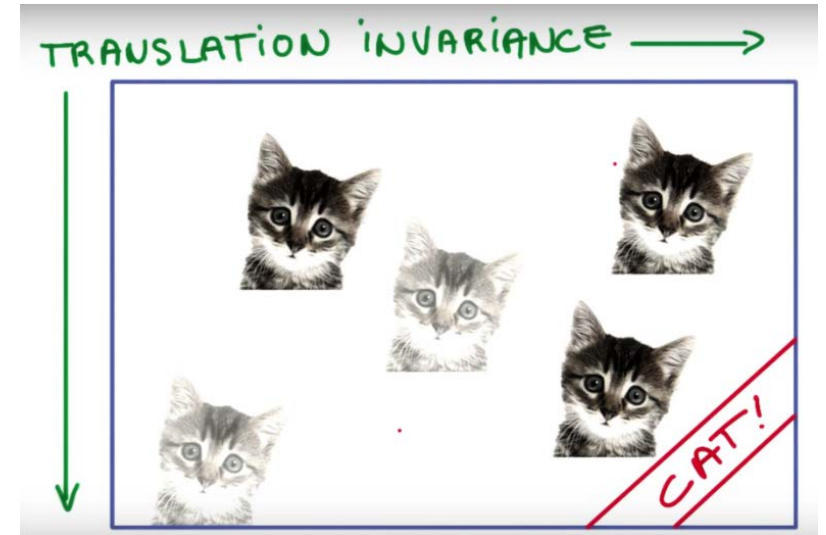
Suresh Sugumar
ssugumar@stanford.edu

Motivation

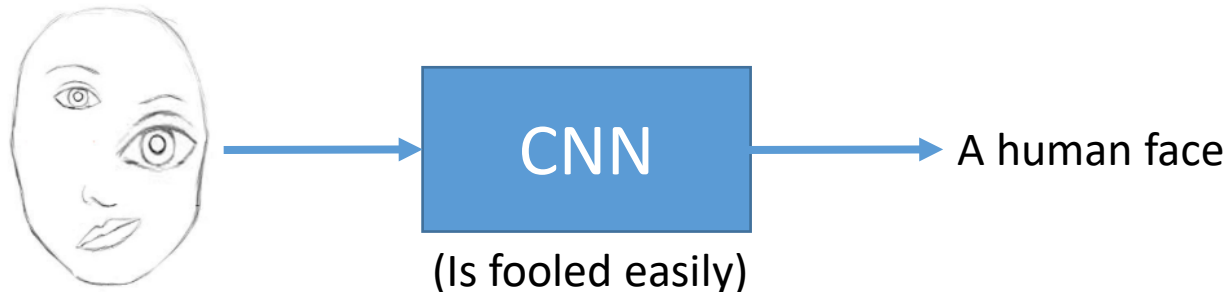


Drawbacks of CNN Architecture

- Pooling layers (max pooling) results in poor translational invariance
 - Loses lot of valuable information during pooling
 - Ignores relation between the parts and the whole
- To improve the translational invariance of CNN
 - Train network with lots of augmented data
 - Inefficient and waste of time and resources
 - But does not solve the problem completely

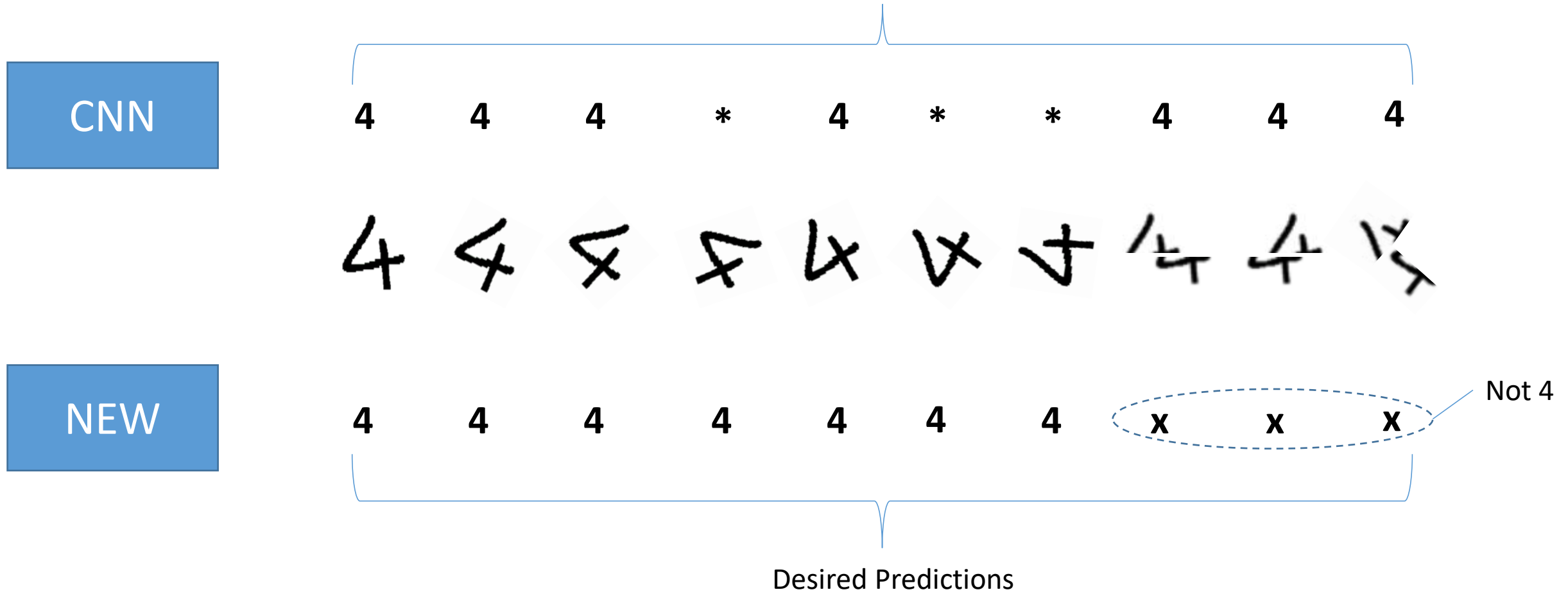


(source)



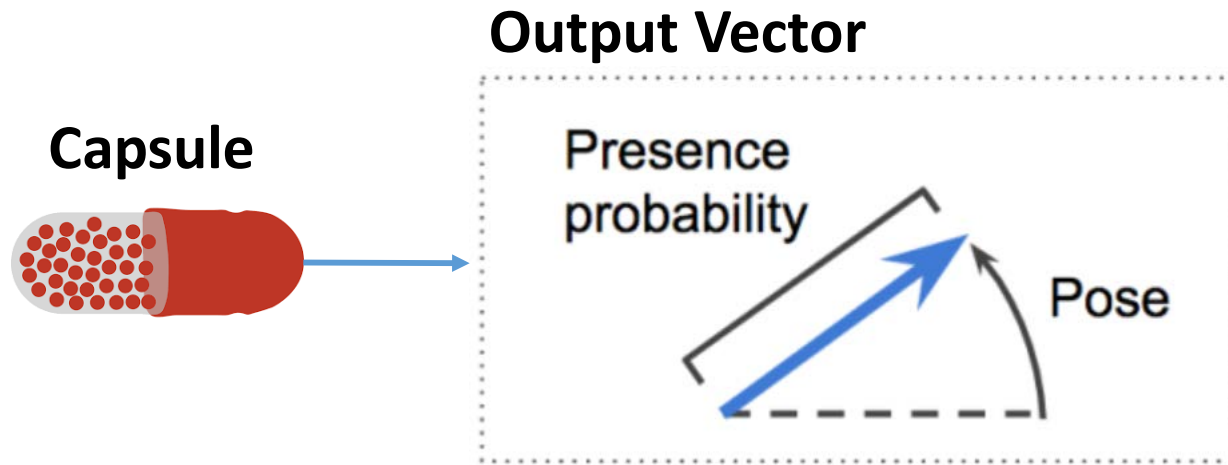
Goal of this Project

Build a neural network that is as good as CNN plus has translational invariance qualities

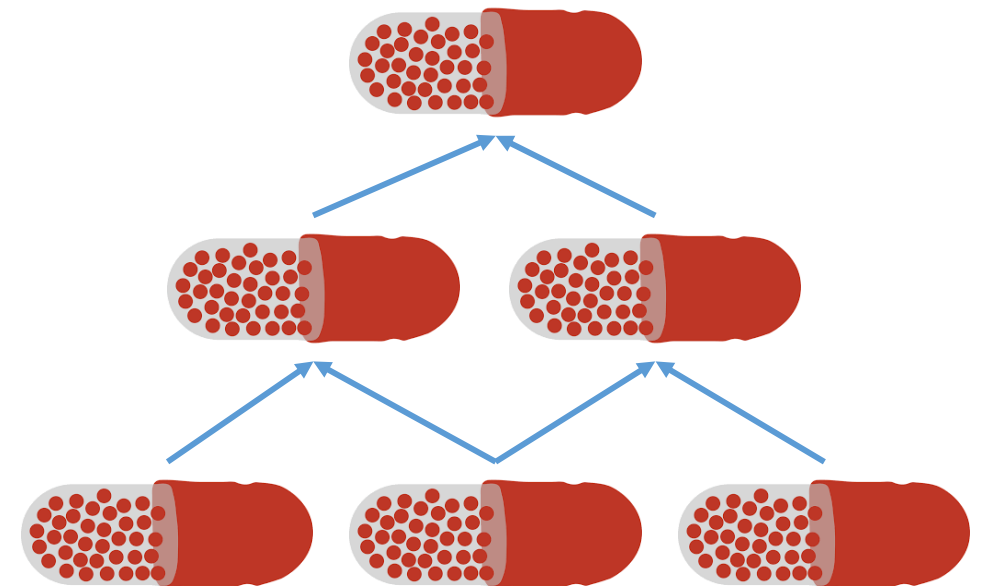


Capsules

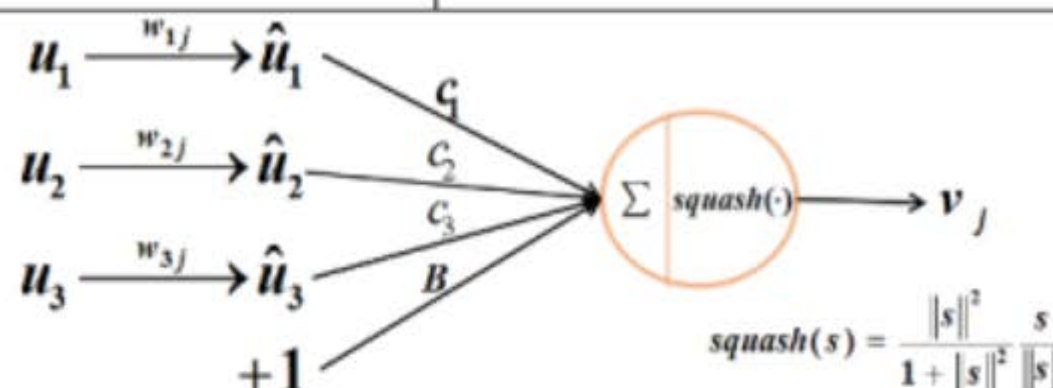
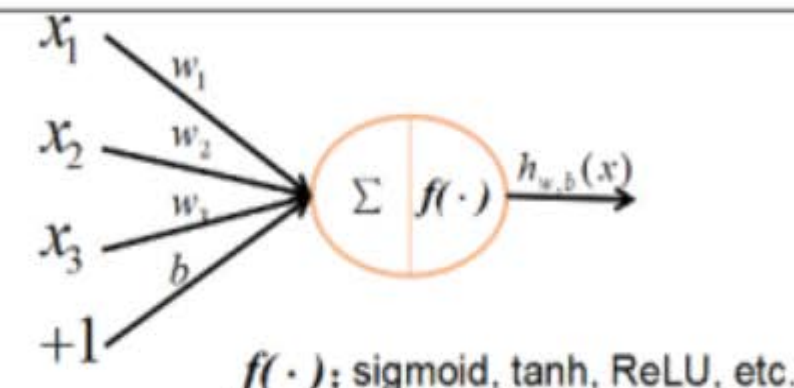
A capsule is a group of neurons whose activity vector represents the instantiation parameters of a specific type of entity such as an object or an object part.



**Hierarchical Neural Network
based on Capsules**

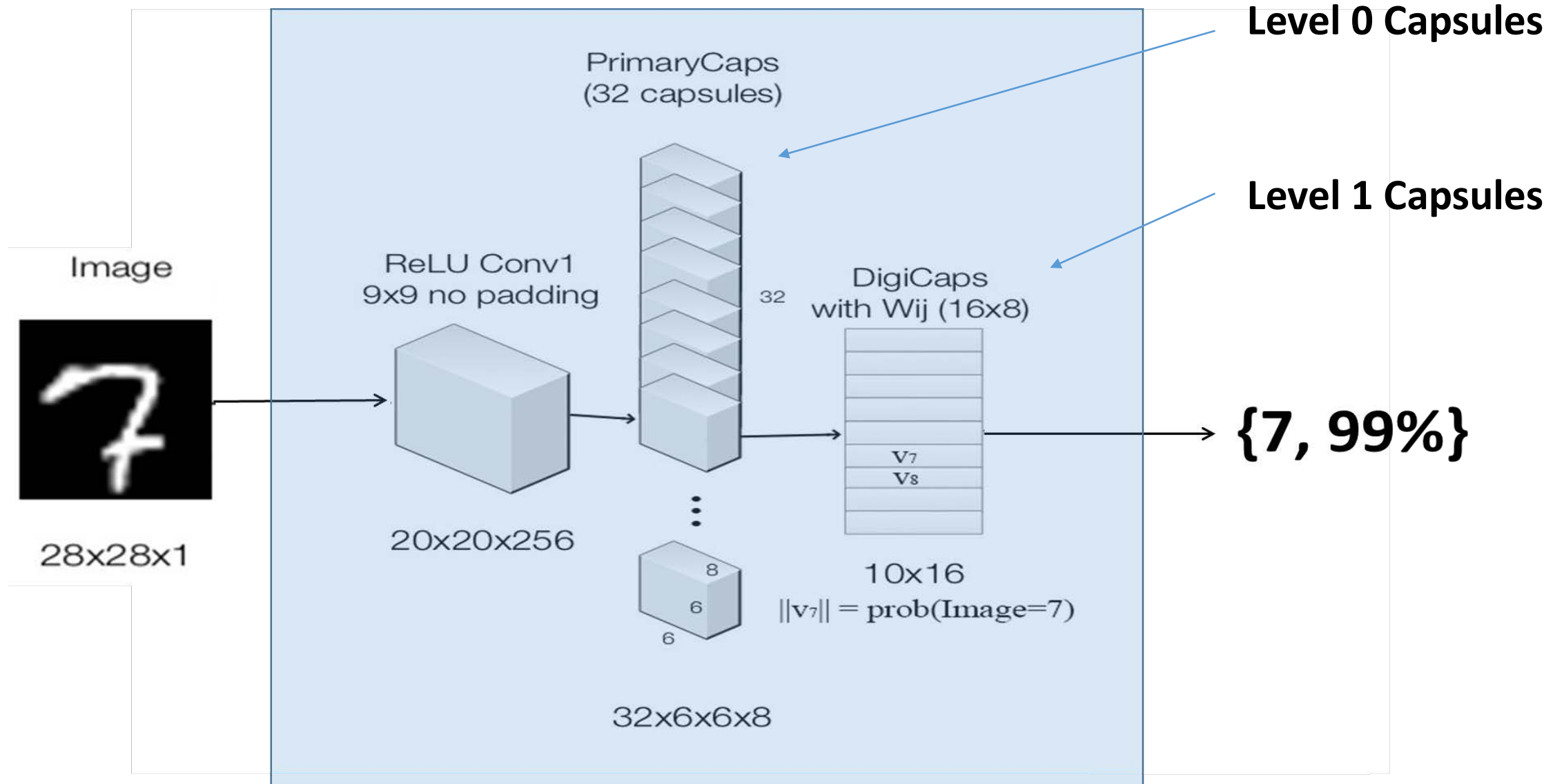


Capsule vs. Traditional Neuron

Input from low-level neurons/capsules		vector(u_i)	scalar(x_i)
Operations	Linear/Affine Transformation	$\hat{u}_{ji} = W_{ij} u_i + B_j$ (Eq. 2)	$a_{ji} = w_{ij} x_i + b_j$
	Weighting	$s_j = \sum_i c_{ij} \hat{u}_{ji}$ (Eq. 2)	$z_j = \sum_{i=1}^3 1 \cdot a_{ji}$
	Summation		
	Non-linearity activation	$v_j = \text{squash}(s_j)$ (Eq. 1)	$h_{w,b}(x) = f(z_j)$
output		vector(v_j)	scalar(h)
			

Capsule Network written for MNIST

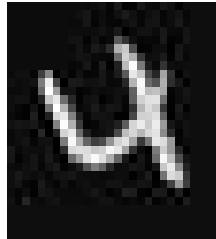
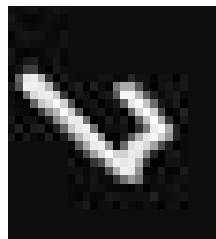
(based on paper by Hinton et., al.)



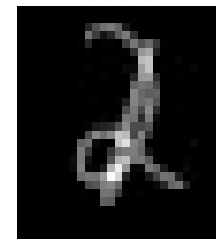
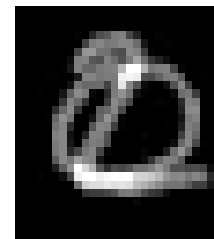
Dataset

- MNIST Handwritten Digits
 - Train set – 60000 examples
 - Test set – 10000 examples
 - Additional 500 images generated for testing, with following operations

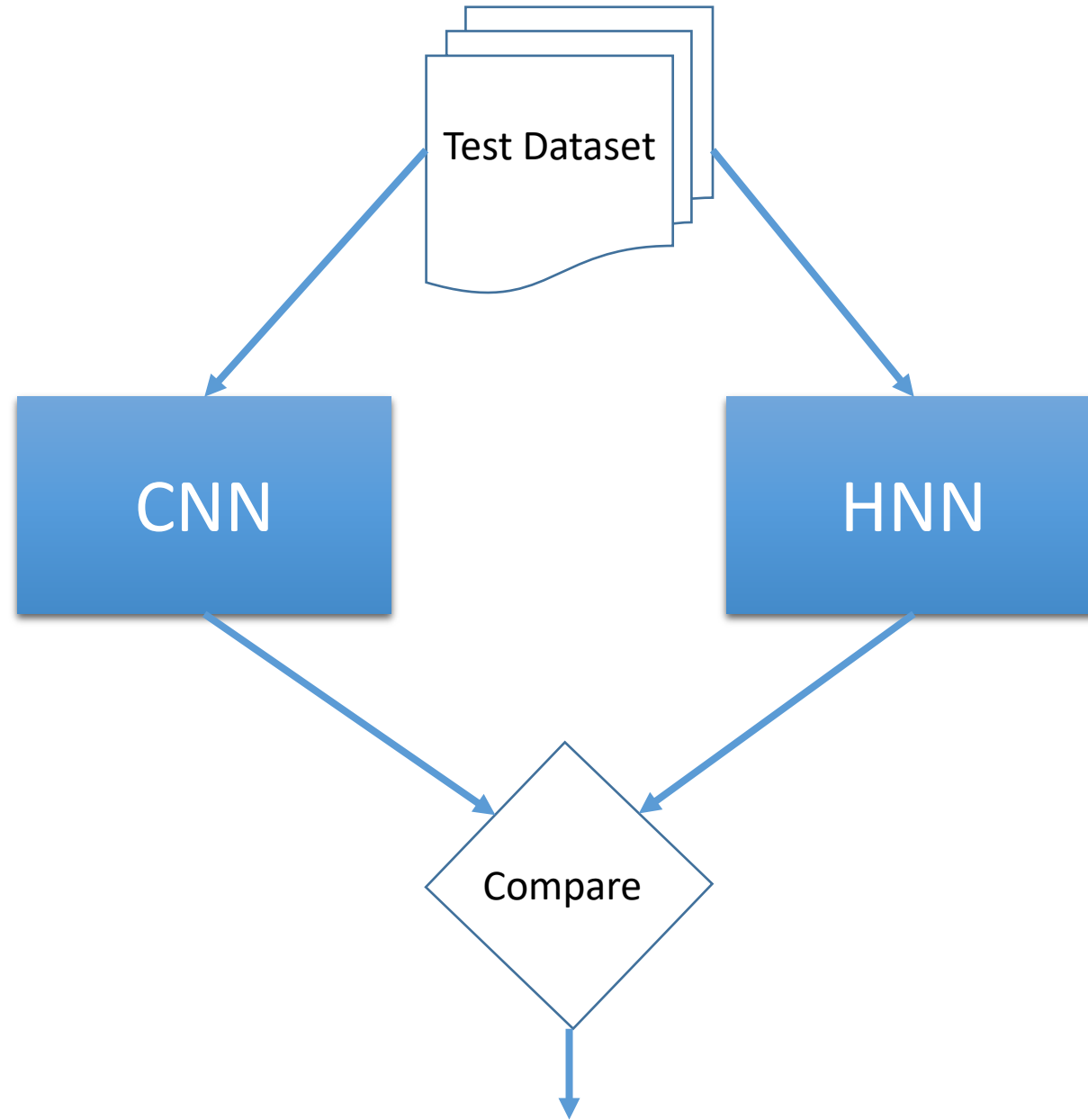
Rotation Operation:



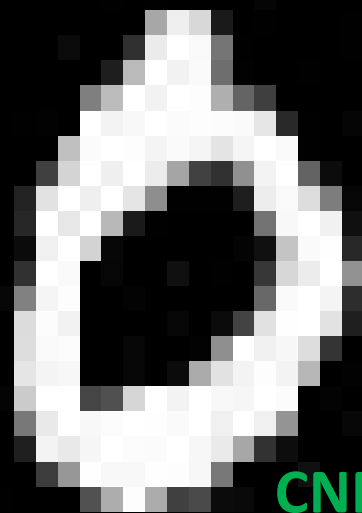
Overlap Operation:



Test Setup



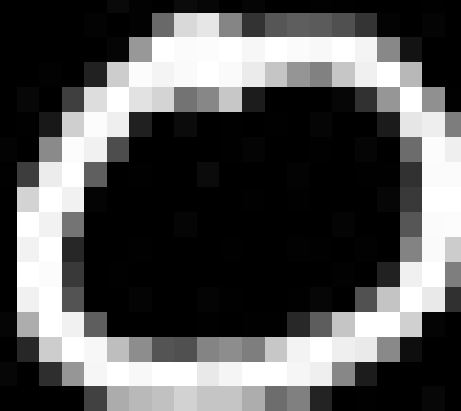
Results



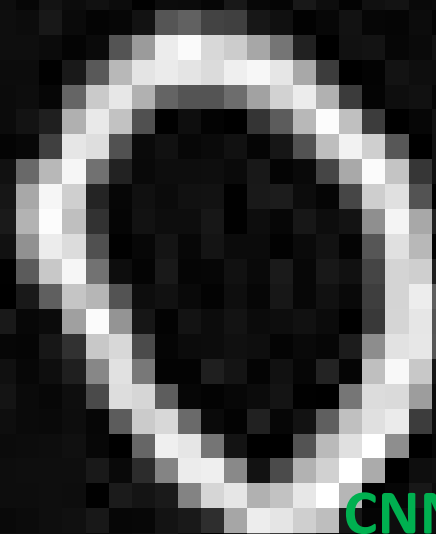
CNN:0
HNN:0



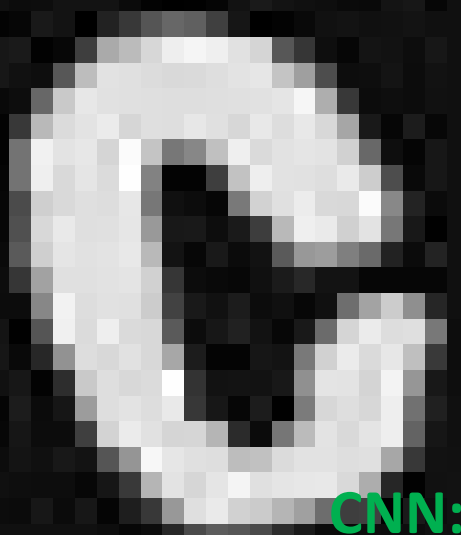
CNN:0
HNN:0



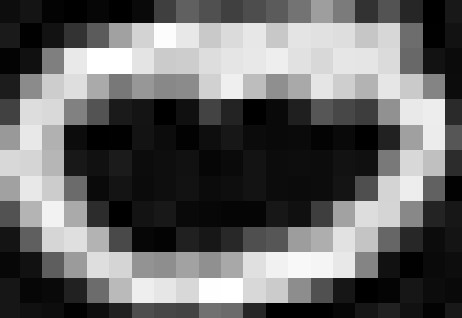
CNN:0
HNN:0



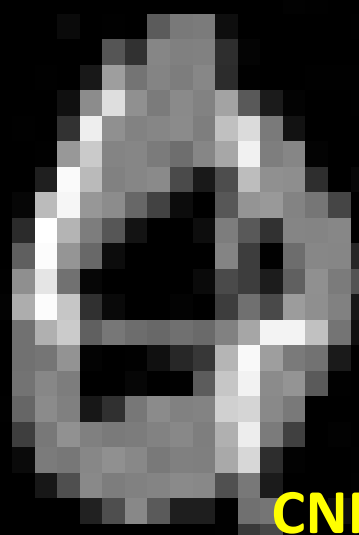
CNN:0
HNN:0



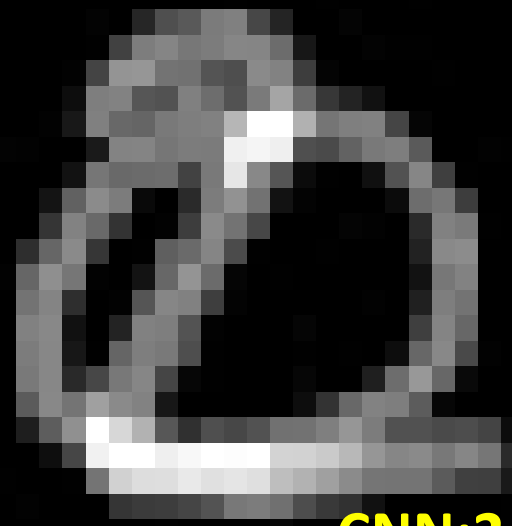
CNN:0
HNN:0



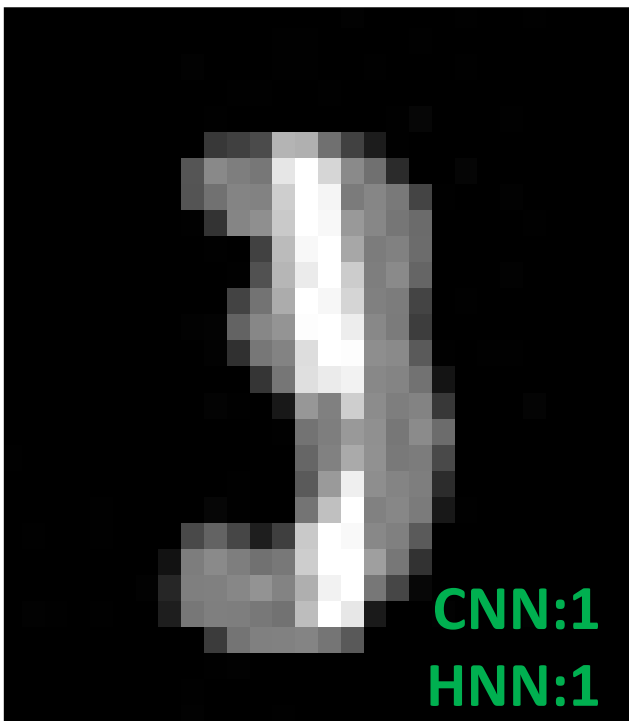
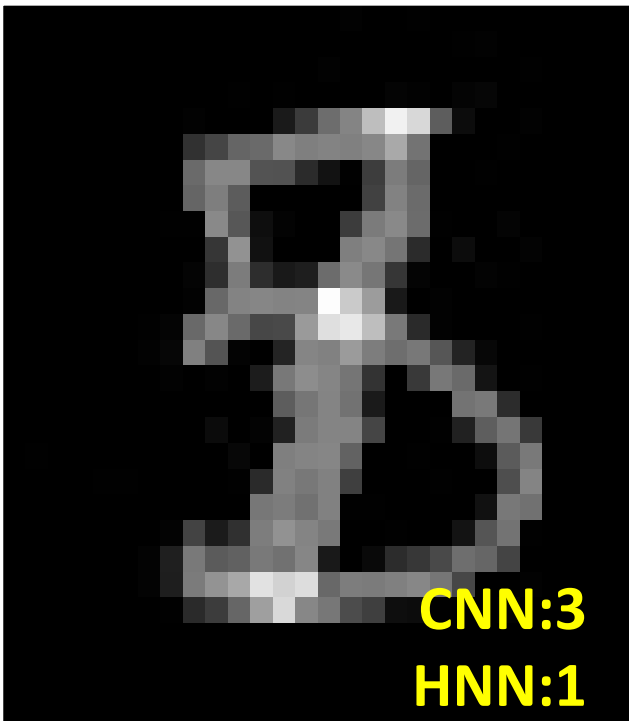
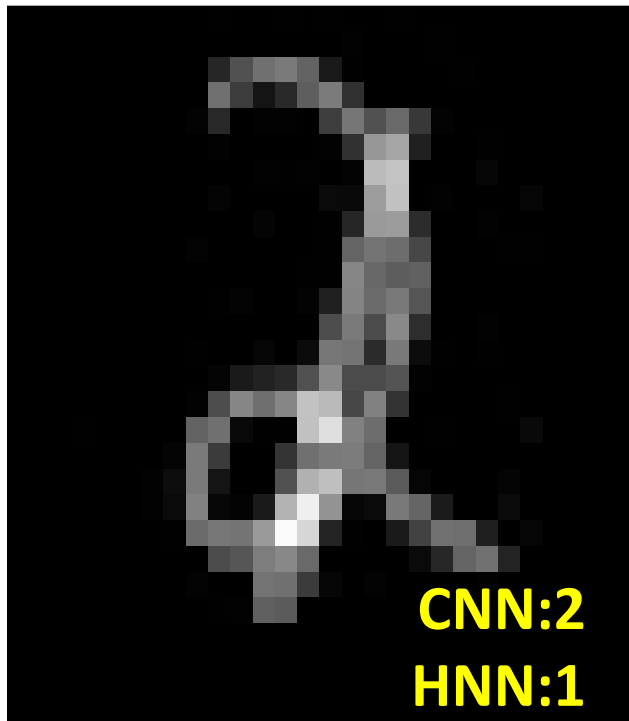
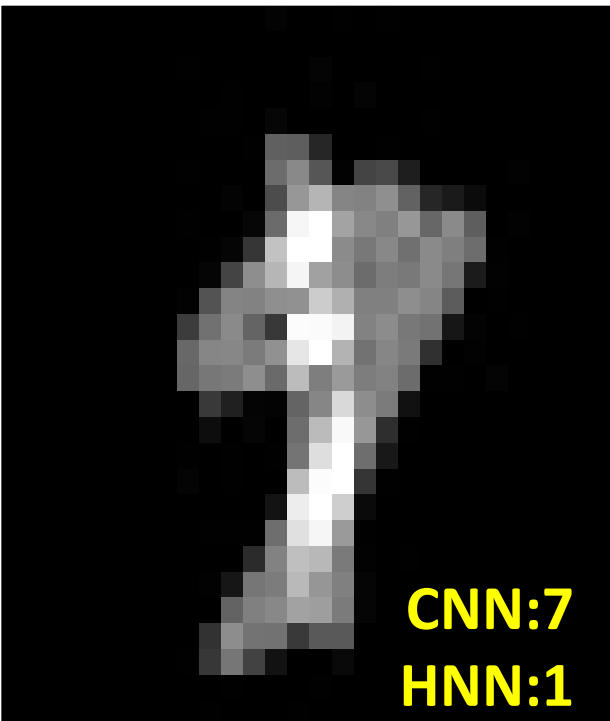
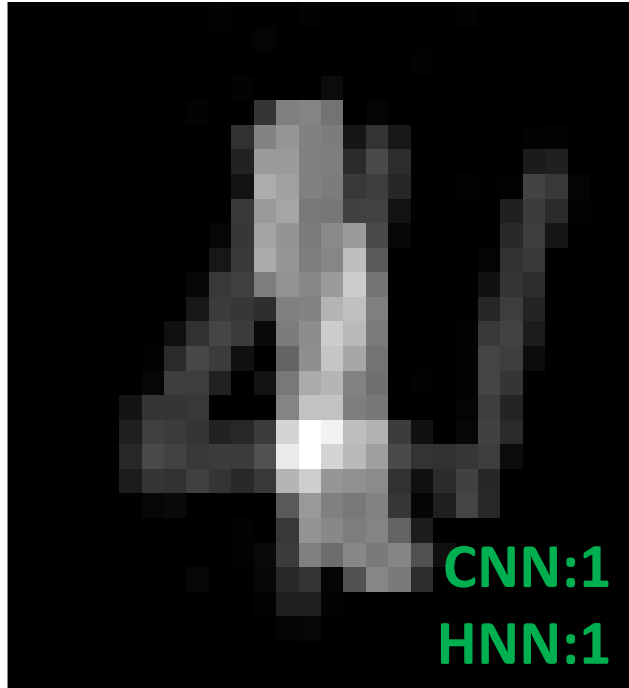
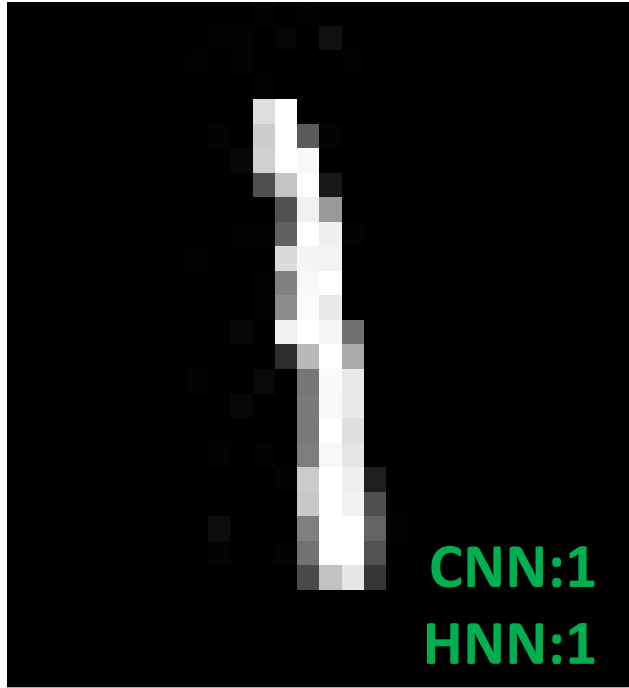
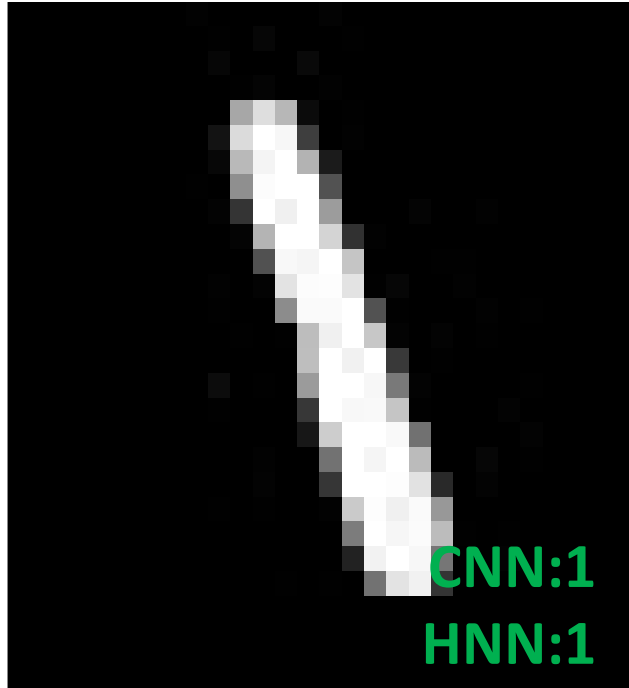
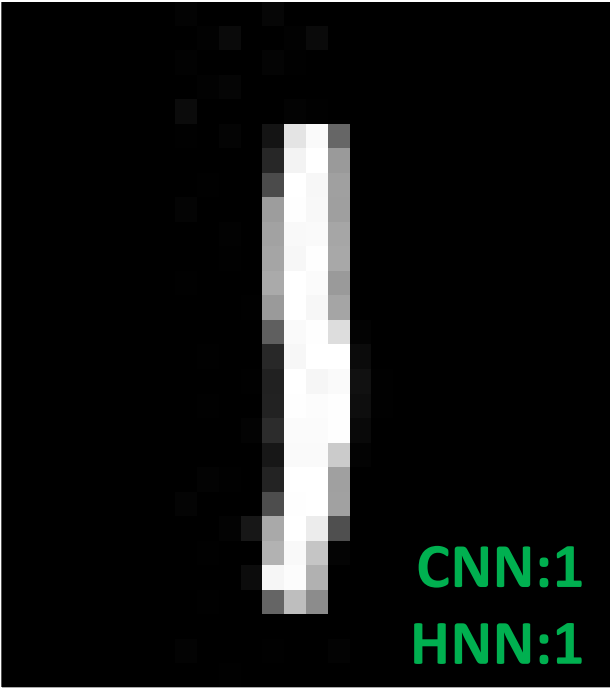
CNN:0
HNN:0

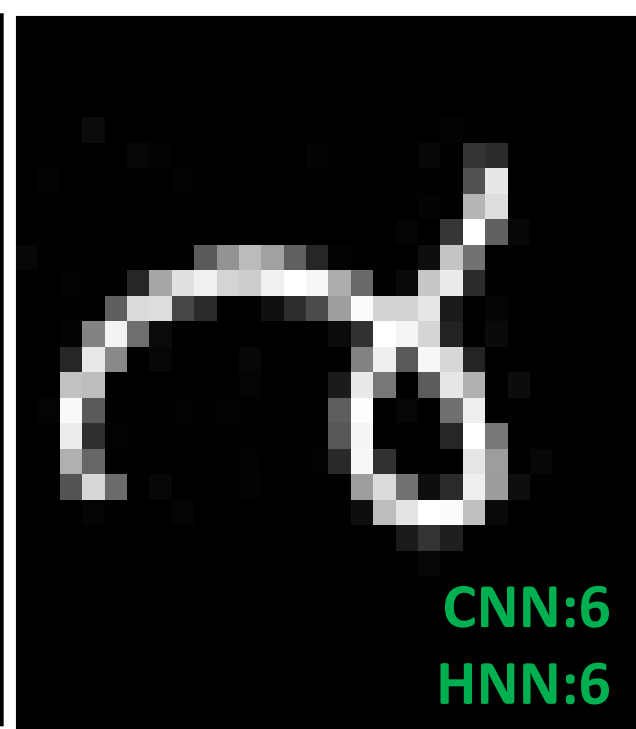
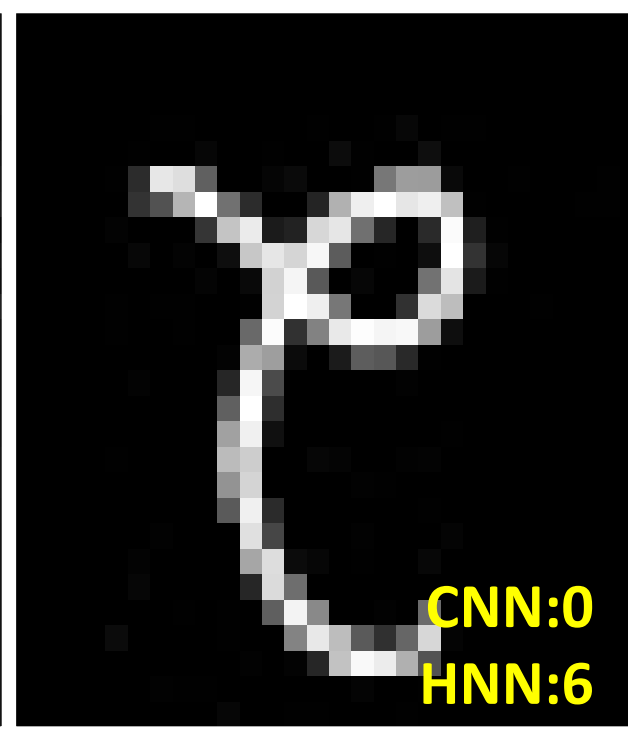
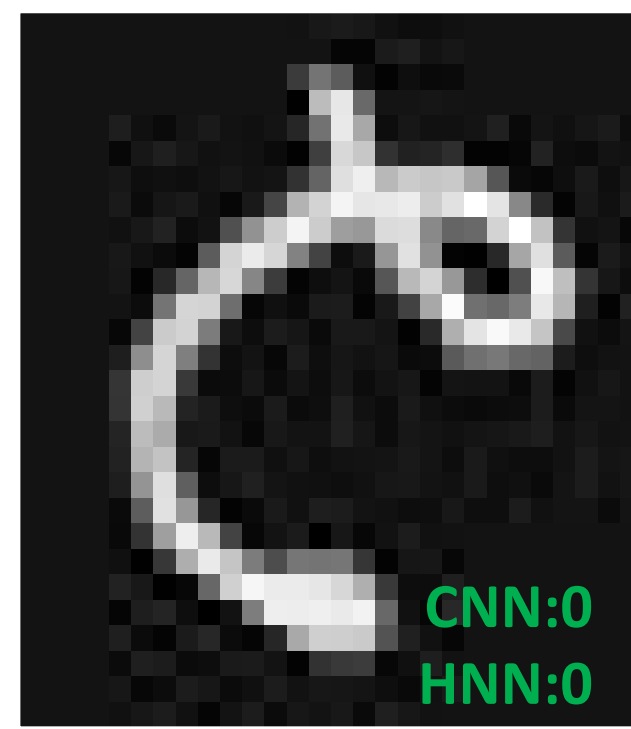
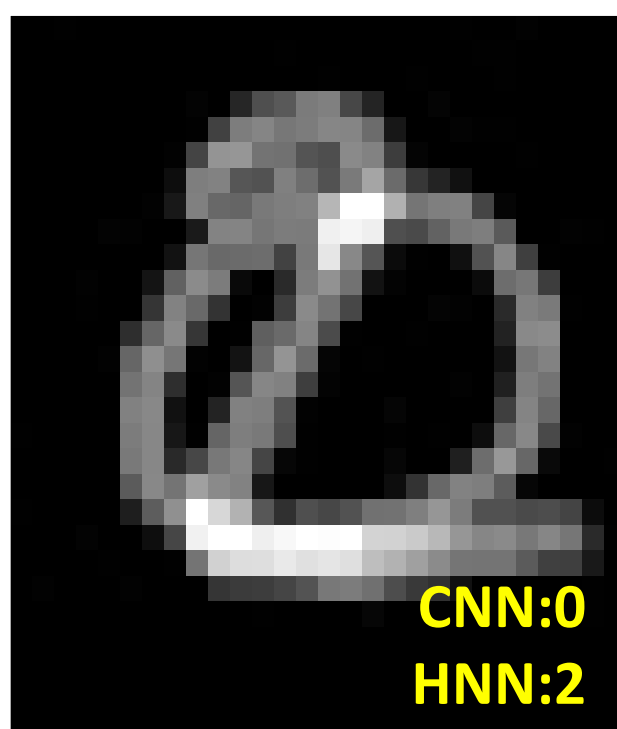
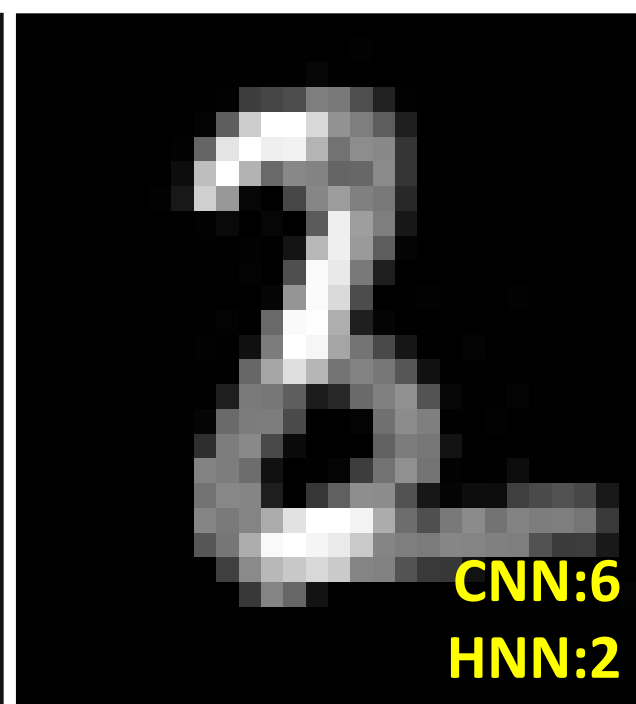
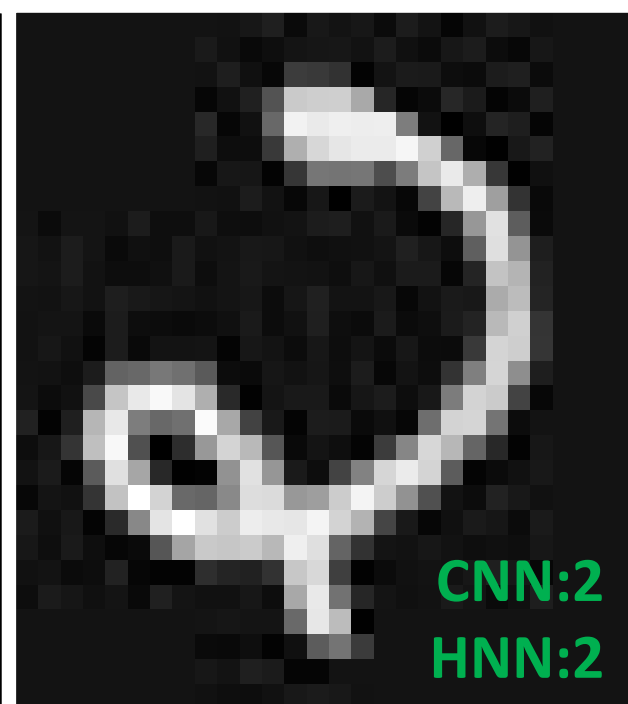
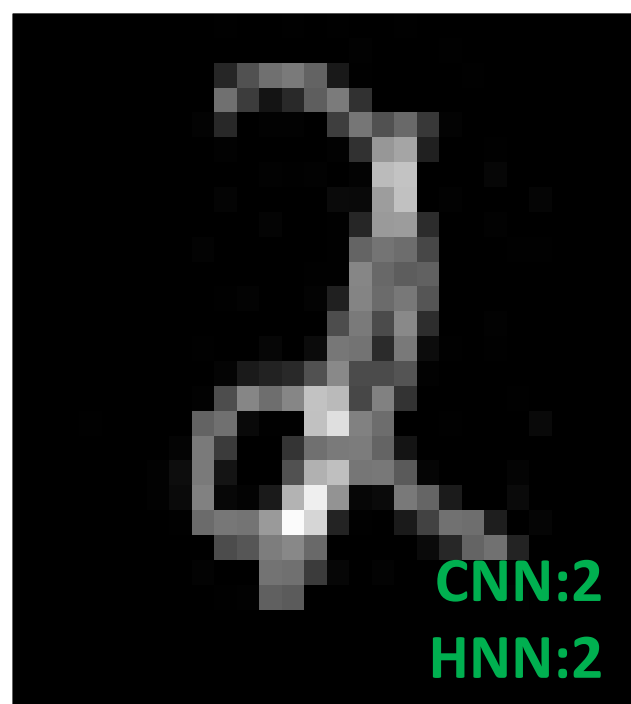
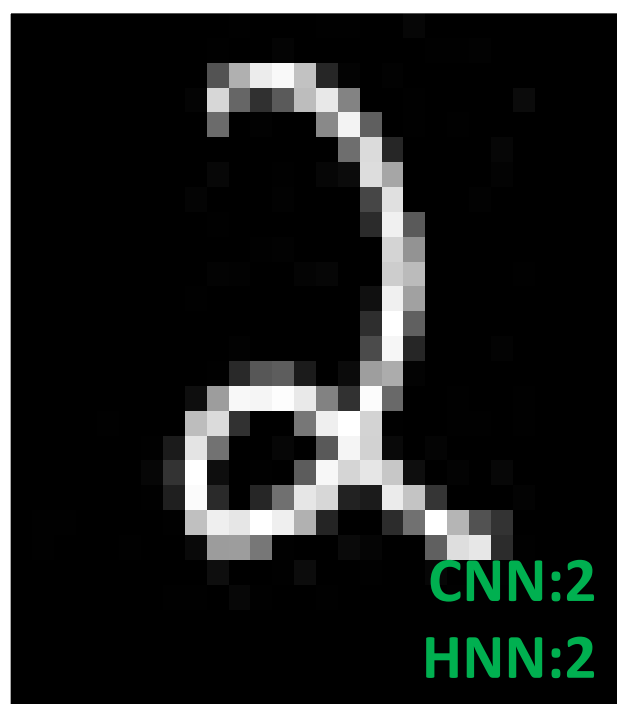


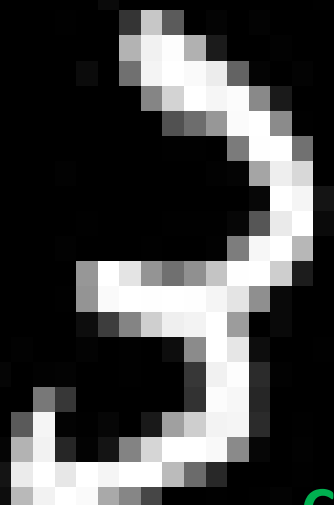
CNN:4
HNN:0



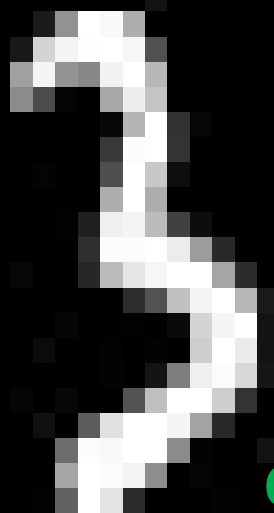
CNN:2
HNN:0



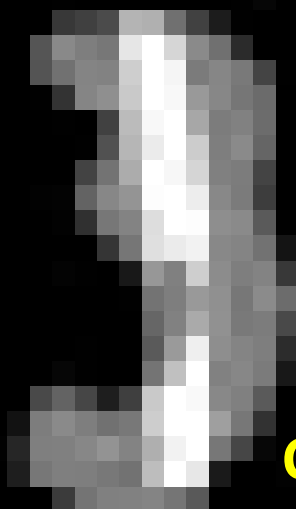




CNN:3
HNN:3



CNN:3
HNN:3



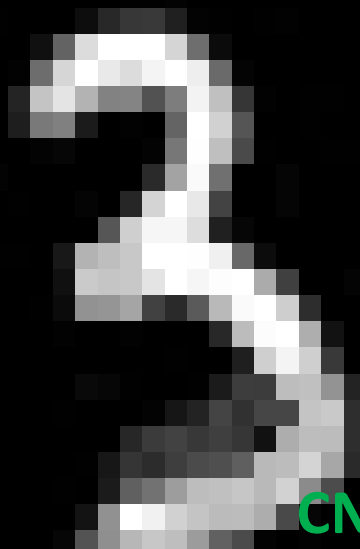
CNN:1
HNN:3



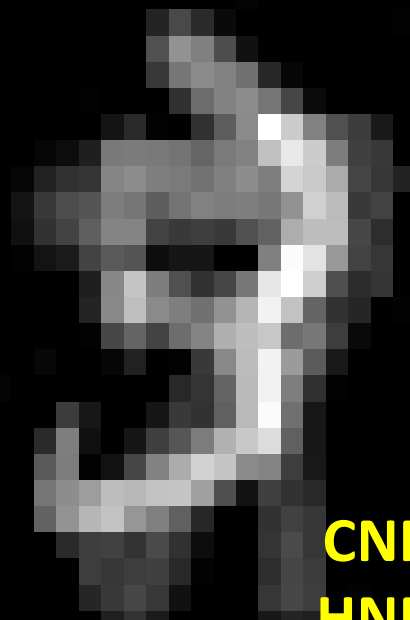
CNN:2
HNN:3



CNN:0
HNN:2



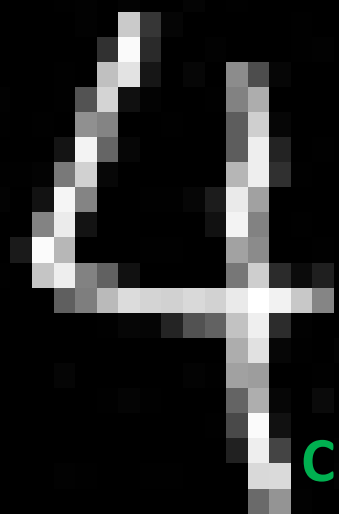
CNN:3
HNN:3



CNN:7
HNN:3



CNN:1
HNN:3



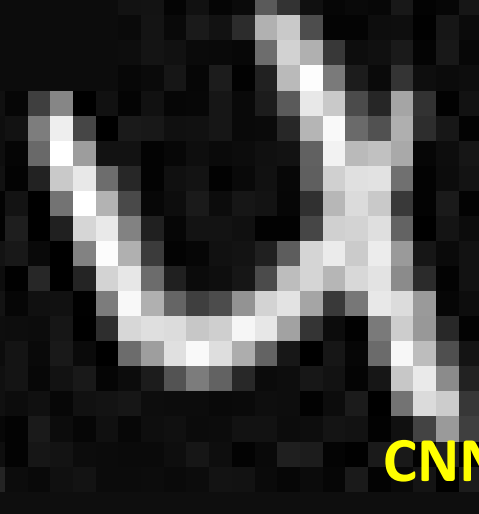
CNN:4
HNN:4



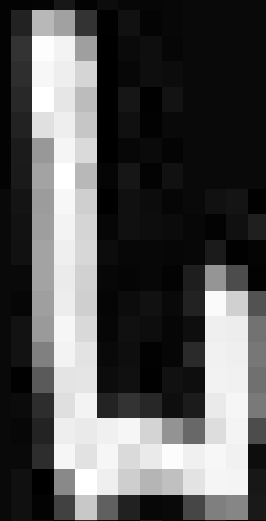
CNN:6
HNN:4



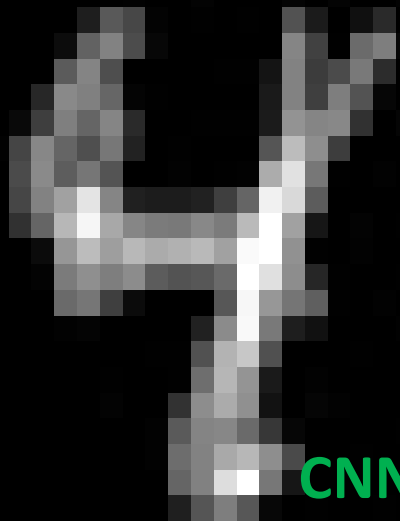
CNN:1
HNN:4



CNN:1
HNN:4



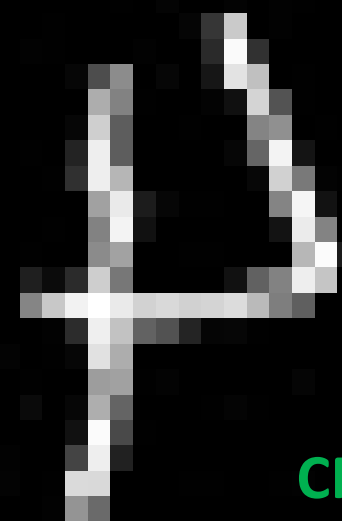
CNN:4
HNN:4



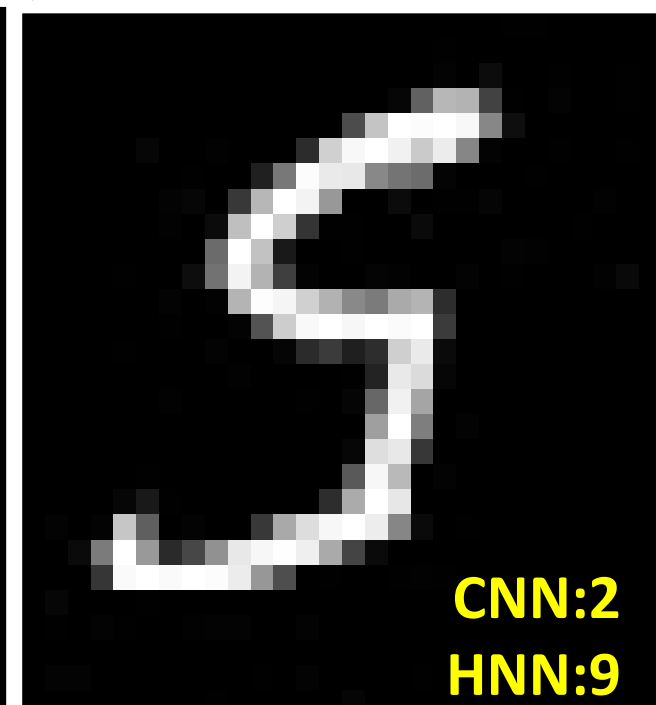
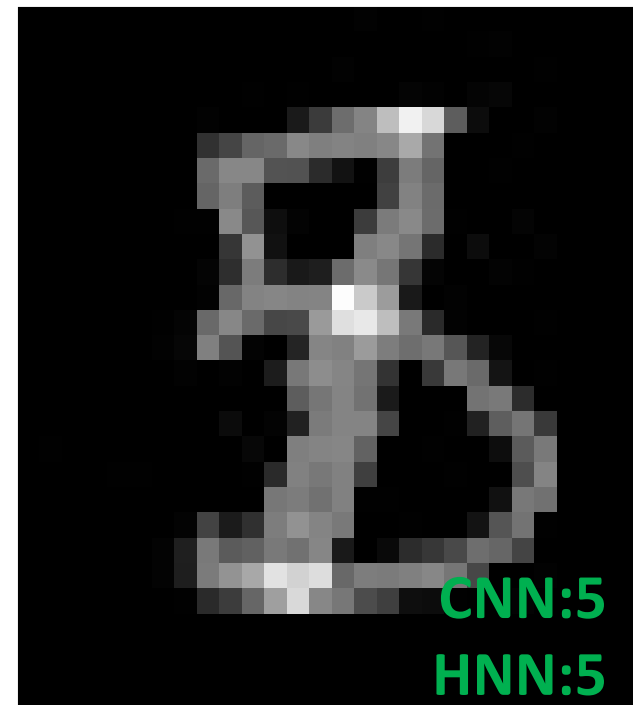
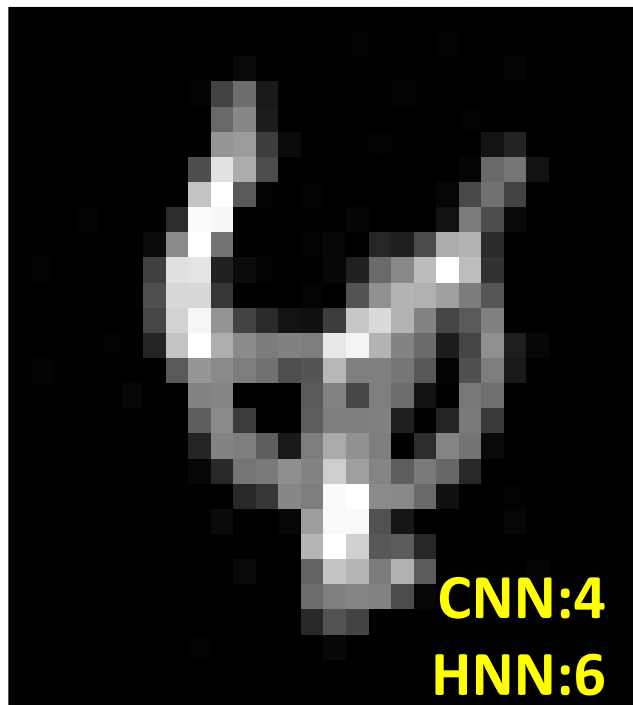
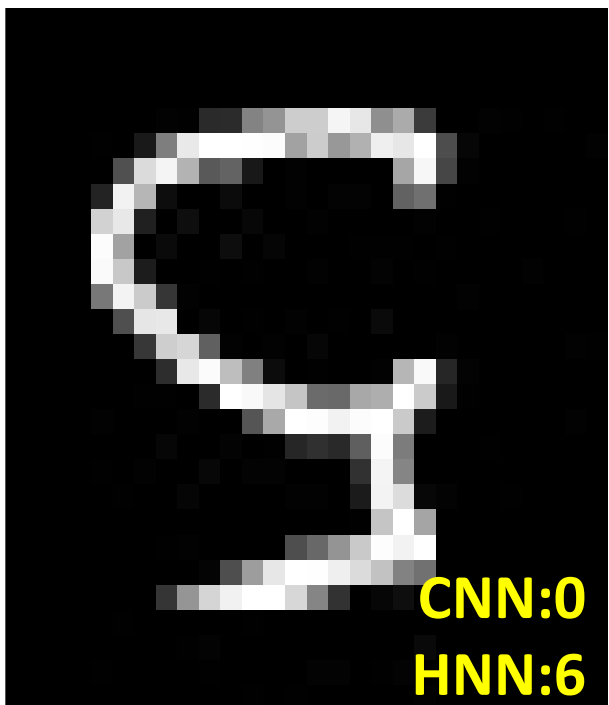
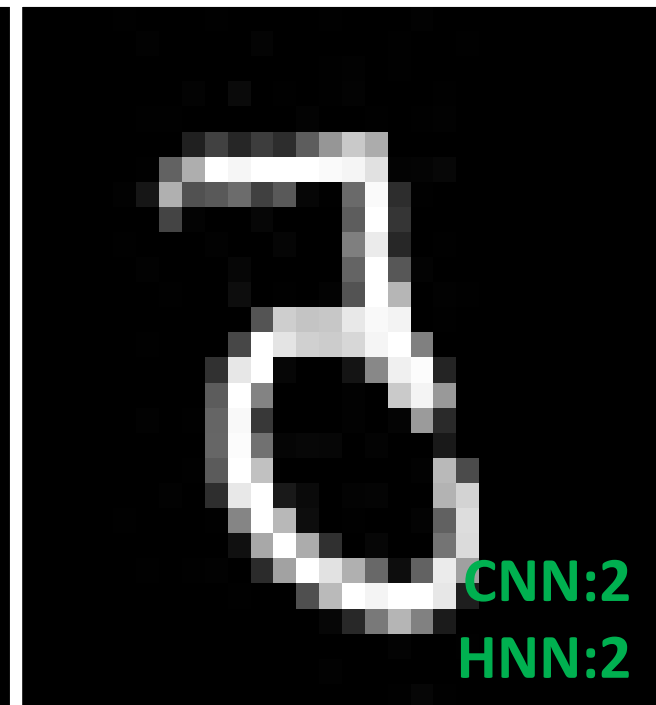
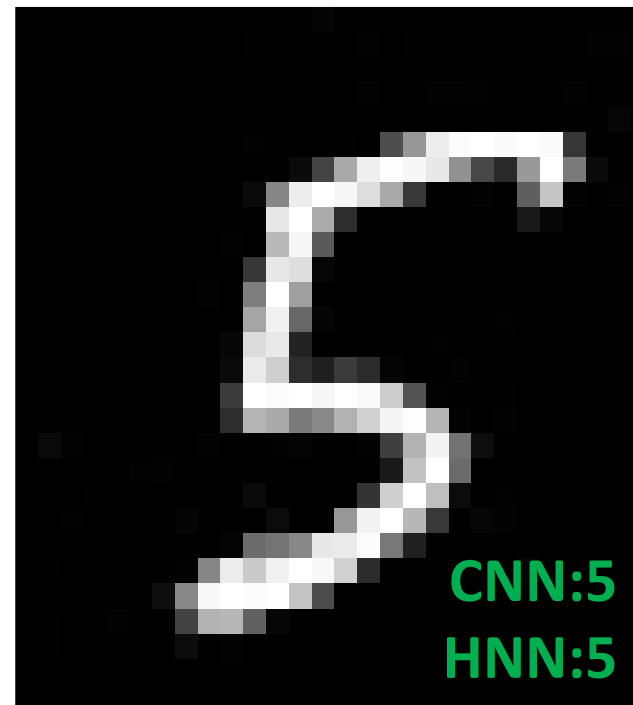
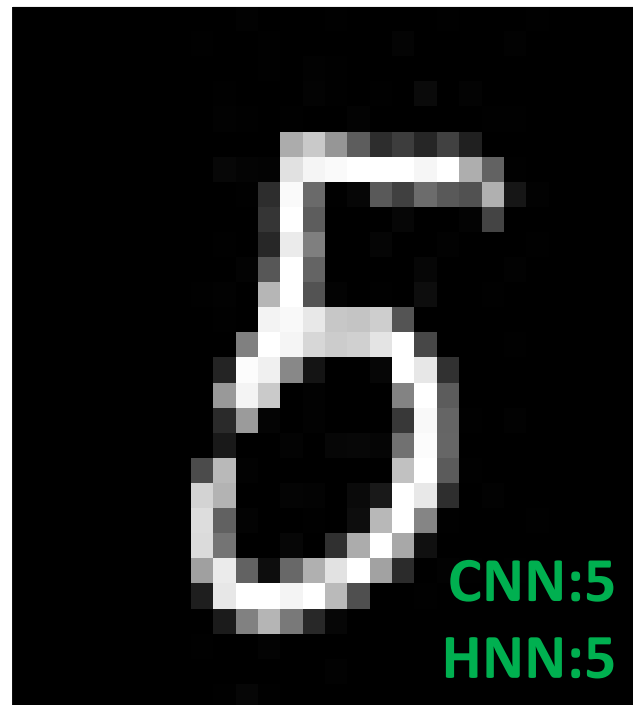
CNN:4
HNN:4

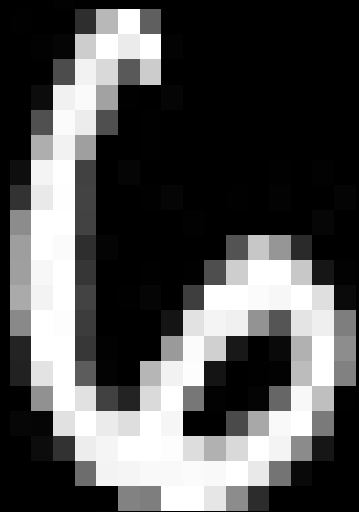


CNN:7
HNN:7

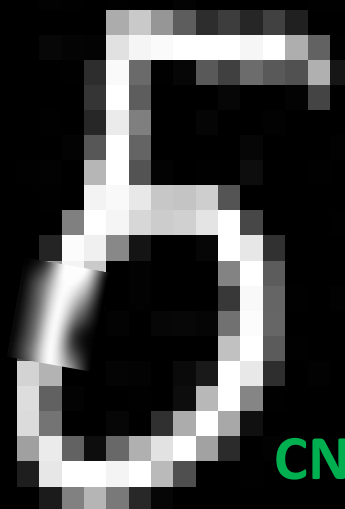


CNN:1
HNN:1





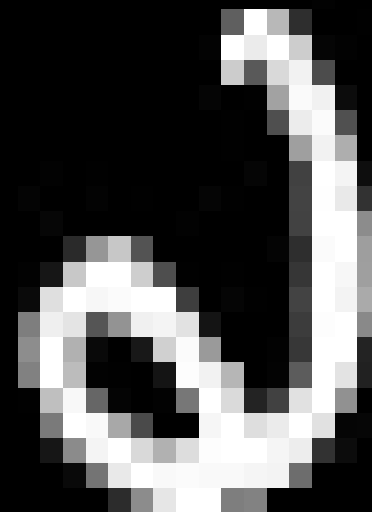
CNN:6
HNN:6



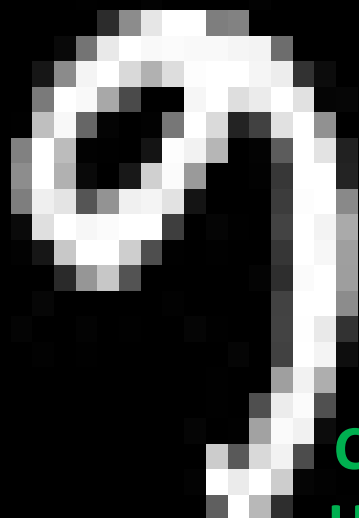
CNN:6
HNN:6



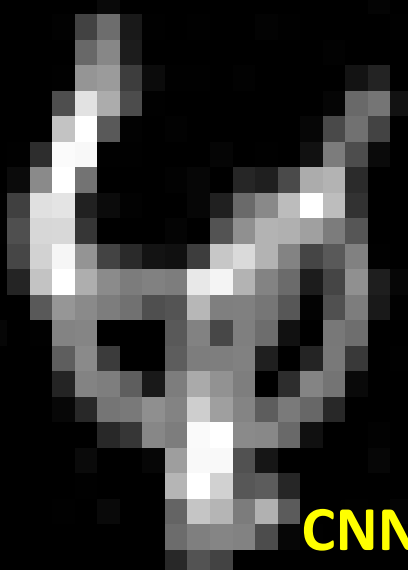
CNN:0
HNN:0



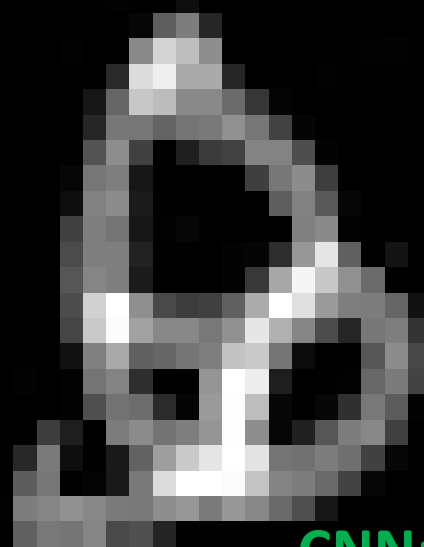
CNN:0
HNN:0



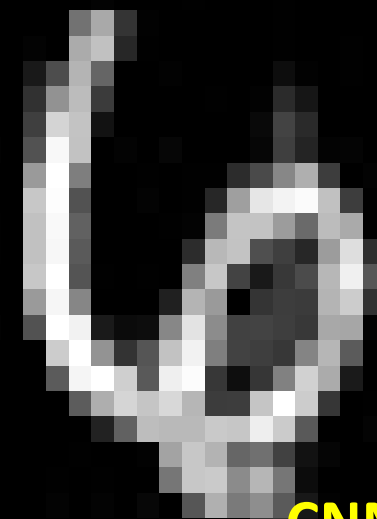
CNN:9
HNN:9



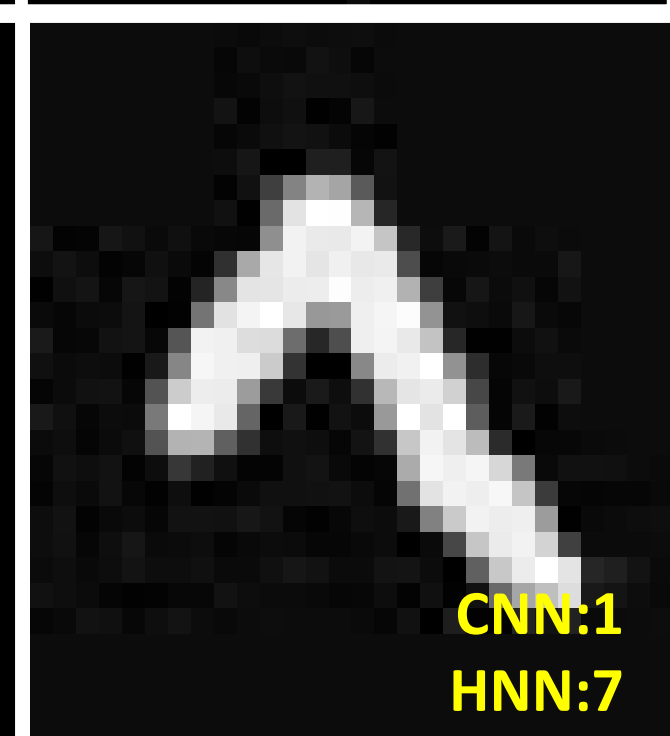
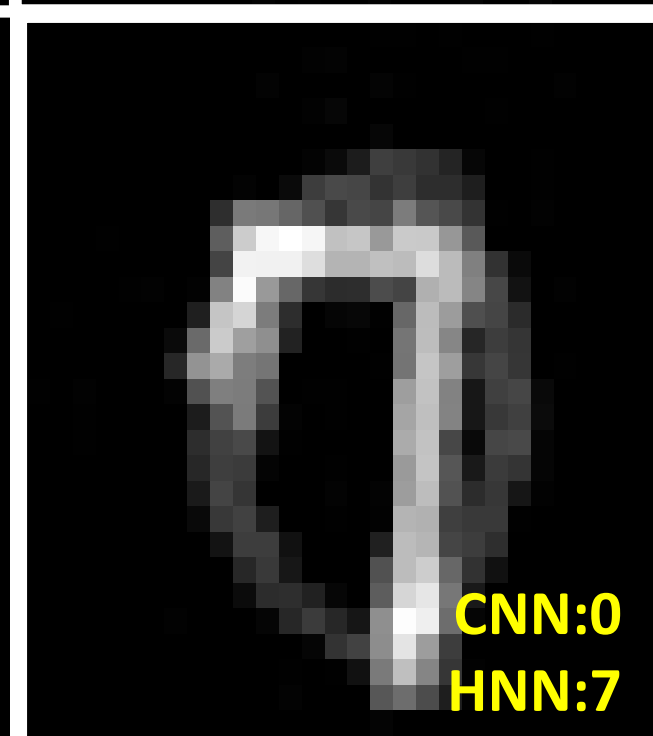
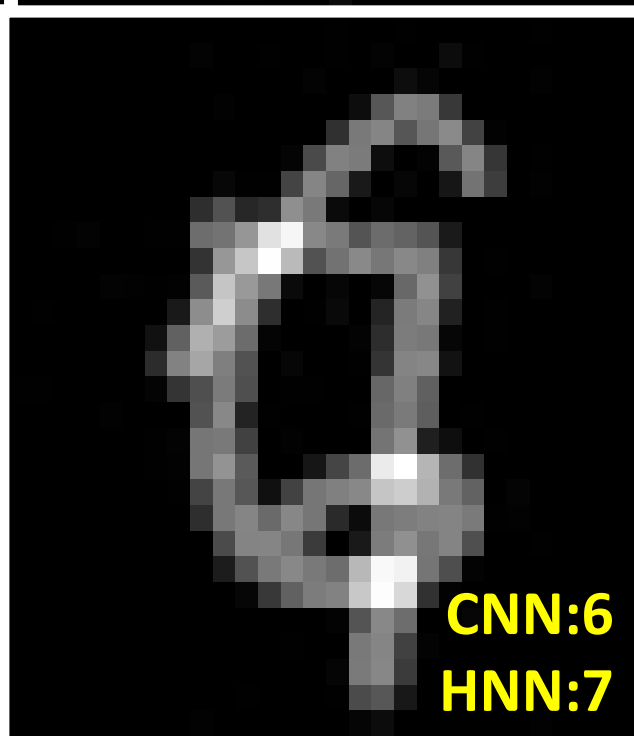
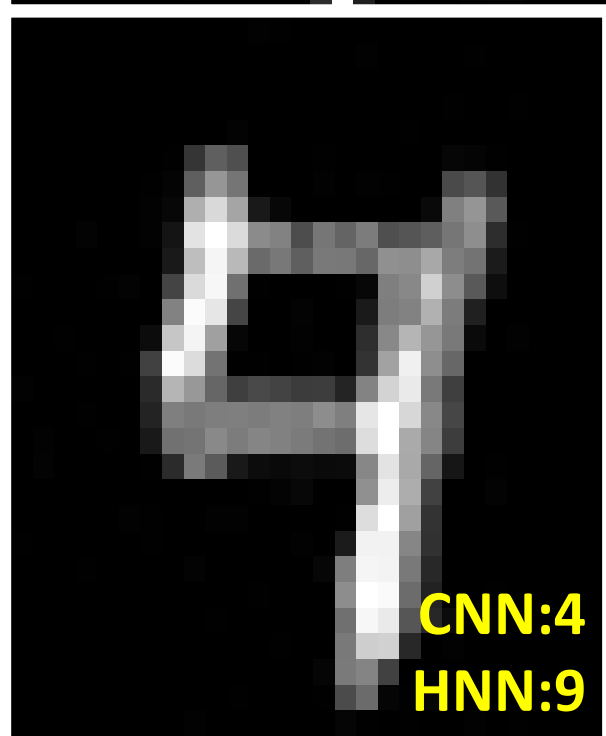
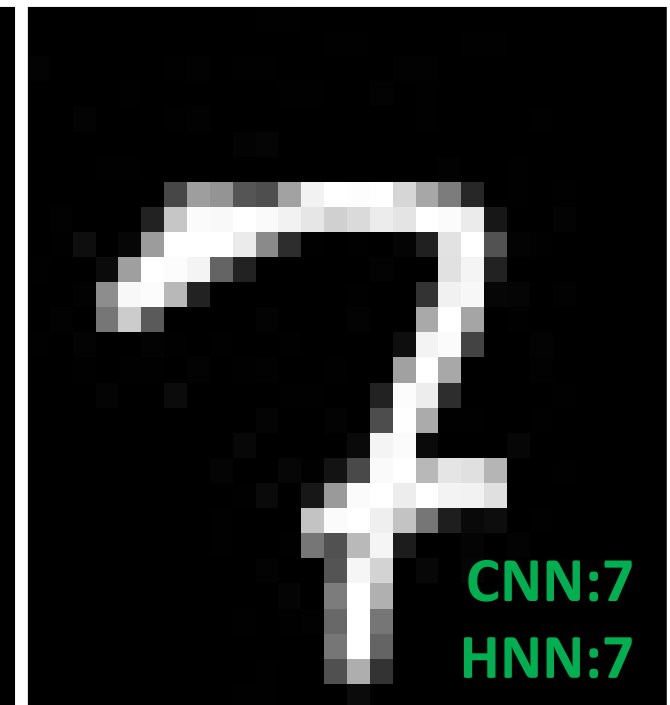
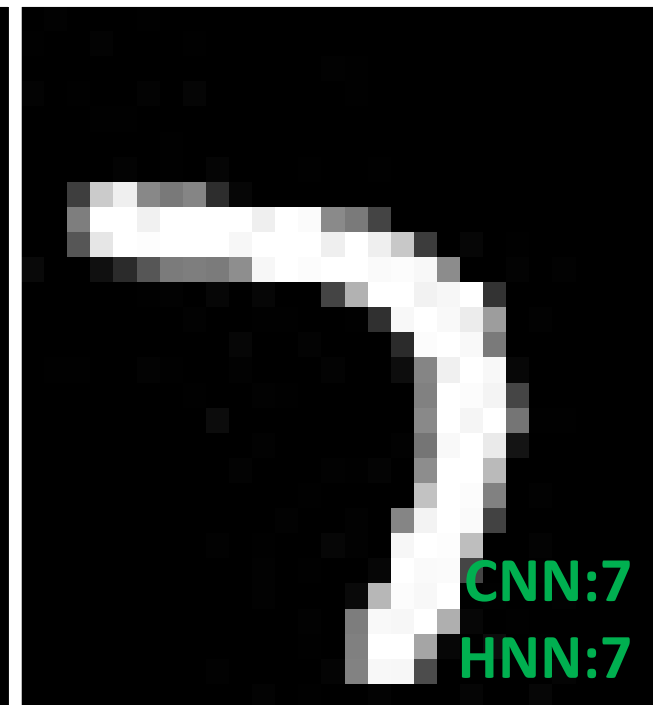
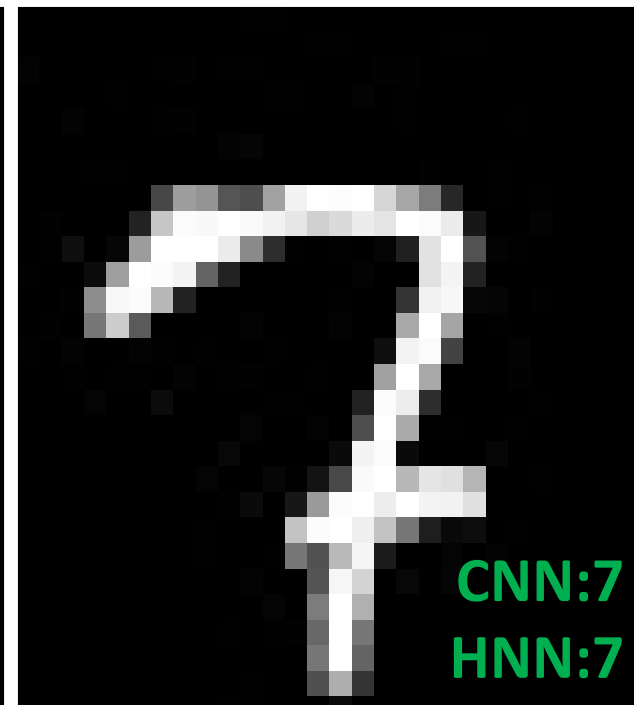
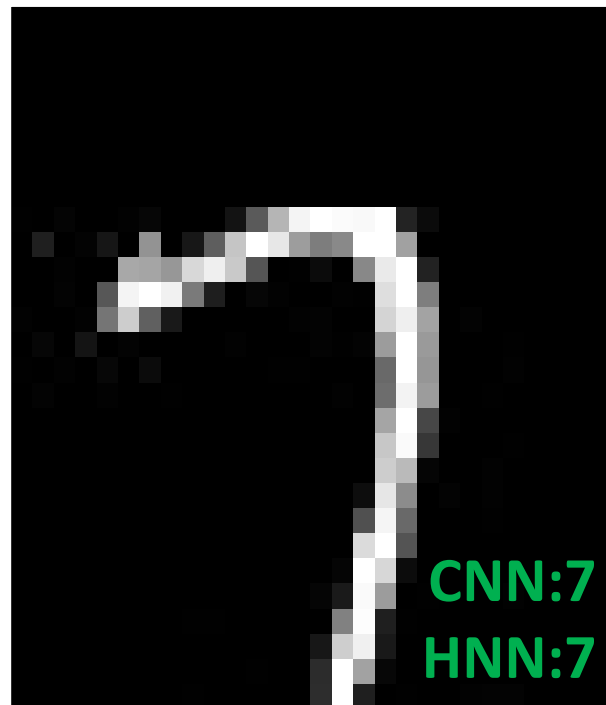
CNN:4
HNN:6

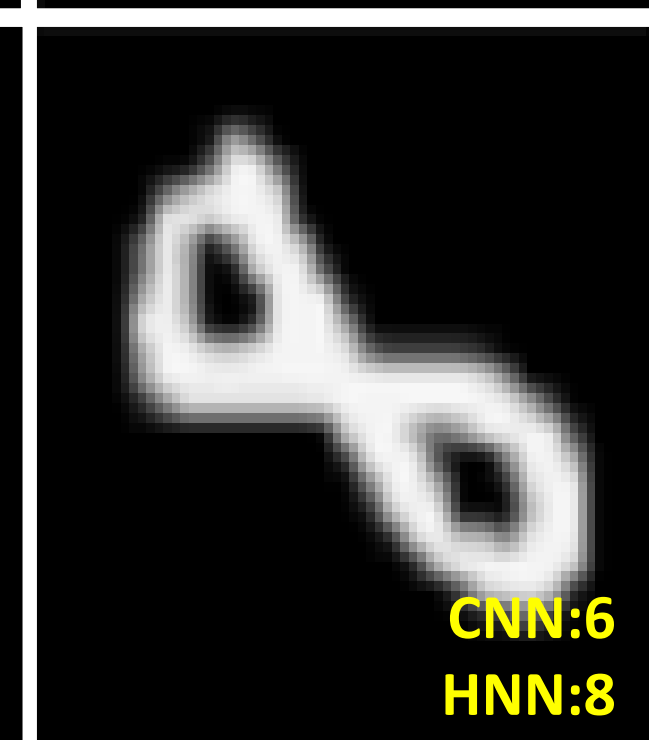
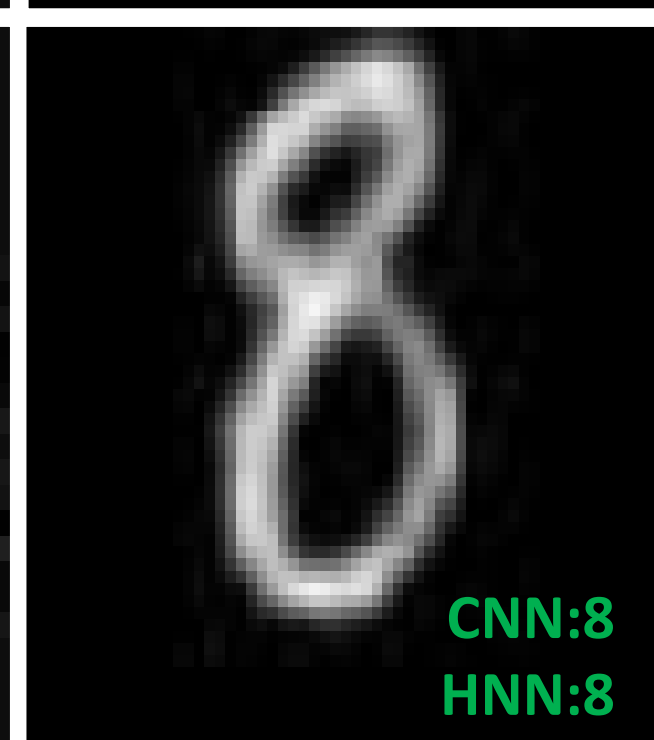
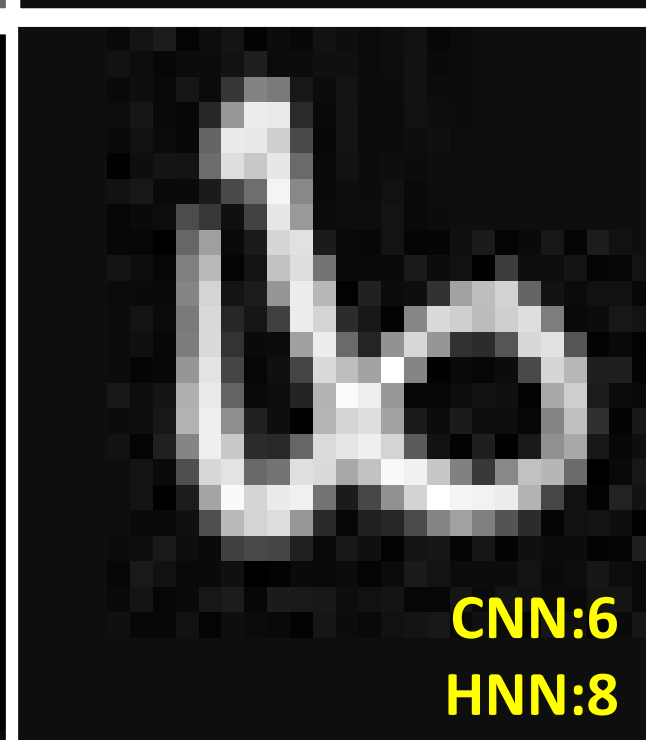
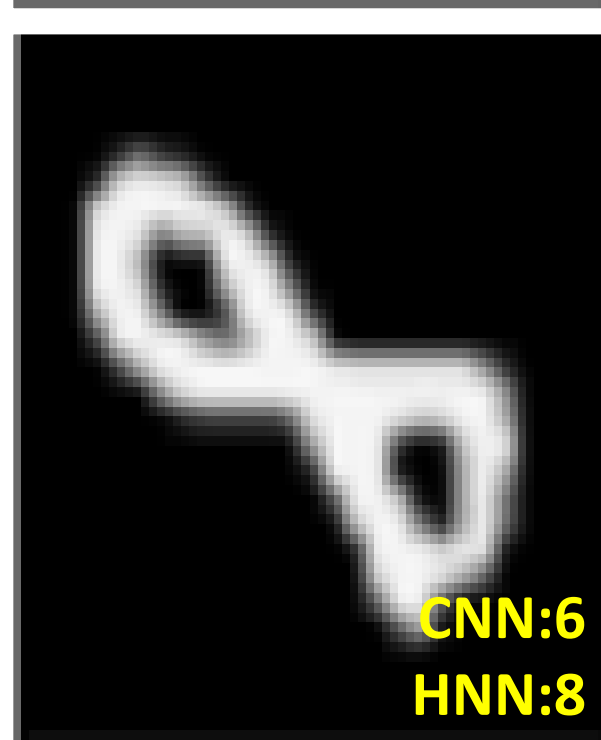
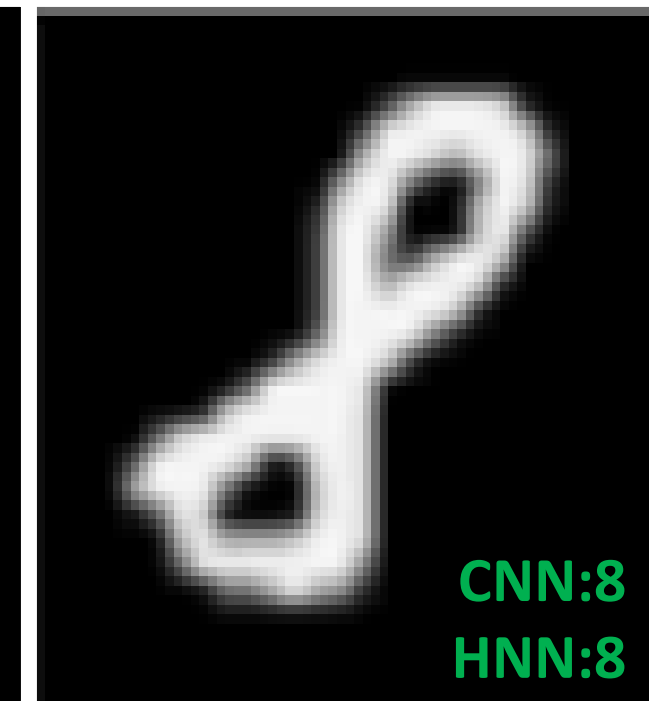
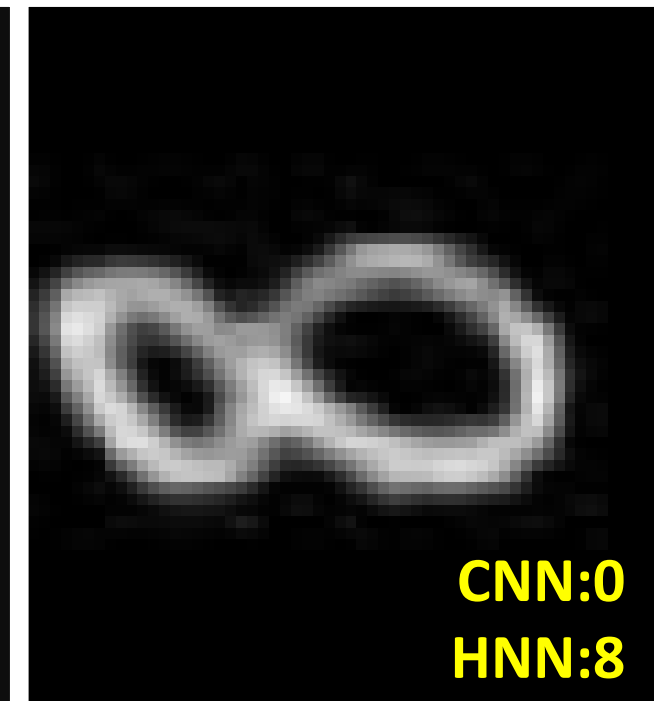
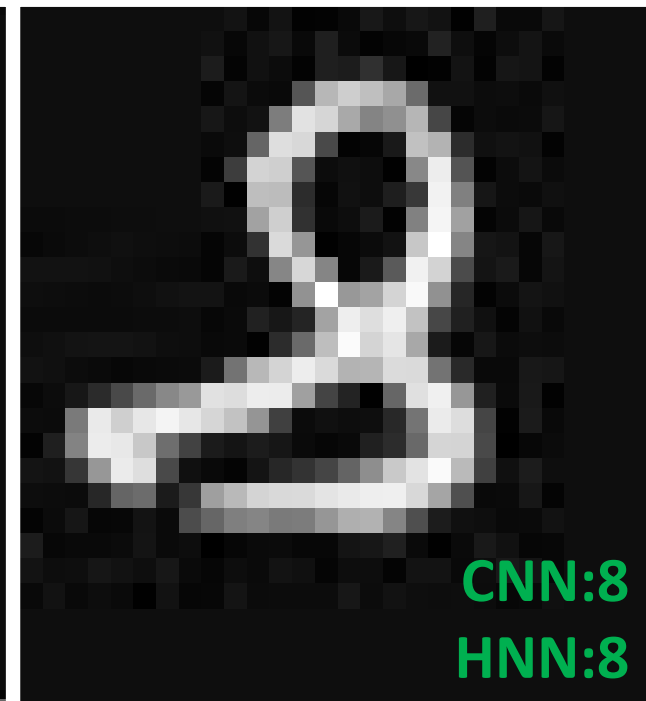
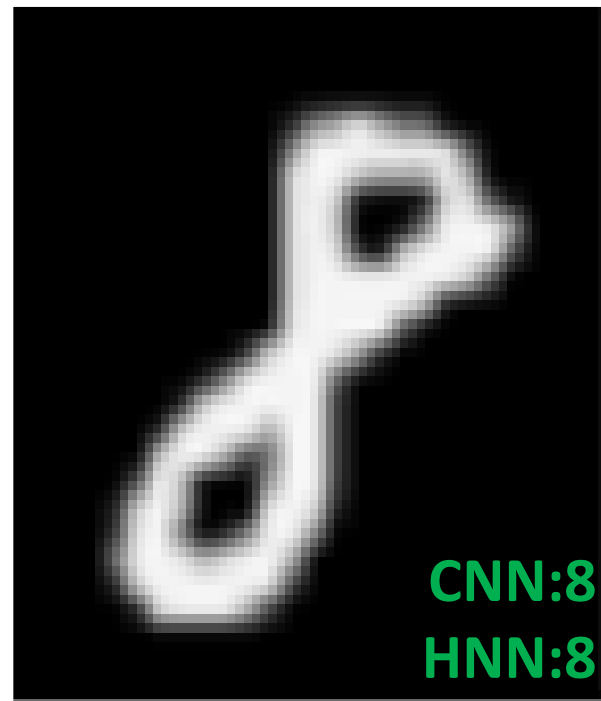


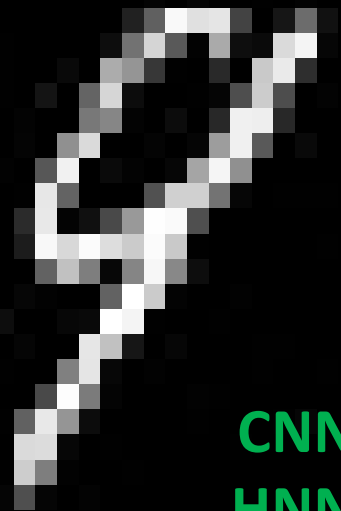
CNN:6
HNN:6



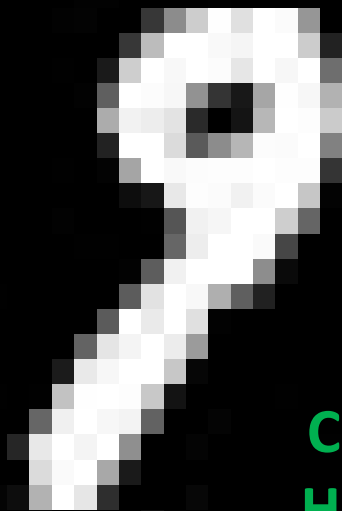
CNN:0
HNN:6







CNN:9
HNN:9



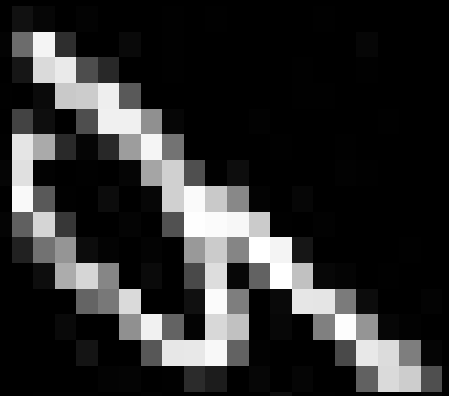
CNN:9
HNN:9



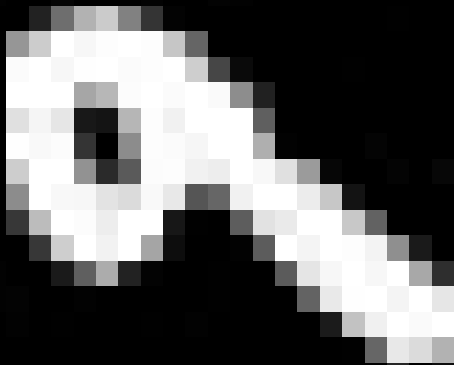
CNN:9
HNN:9



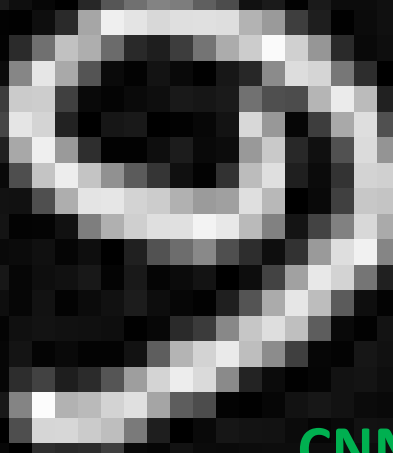
CNN:4
HNN:9



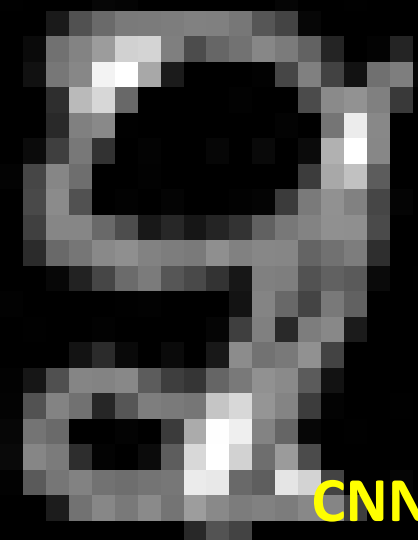
CNN:0
HNN:9



CNN:0
HNN:9



CNN:9
HNN:9



CNN:2
HNN:9

Observations

- Longer training duration compared to CNN
 - Further investigation and optimization required

GitHub

- <https://github.com/sureshsugumar/HierNet>

References

- **Concept inspiration –**

Sara Sabour, Nicholas Frosst, Geoffrey E Hinton. Dynamic Routing Between Capsules. [arXiv:1710.09829](https://arxiv.org/abs/1710.09829), 2017.

- **MNIST Video generation –**

<http://blog.otoro.net/2016/04/01/generating-large-images-from-latent-vectors/>

- **Demo inspiration –**

[https://github.com/lexfridman/mit-deep-learning/blob/master/tutorial deep learning basics/deep learning basics.ipynb](https://github.com/lexfridman/mit-deep-learning/blob/master/tutorial%20deep%20learning%20basics/deep%20learning%20basics.ipynb)

Project Demo

YouTube Video

<https://youtu.be/TL81fI0X6vg>