A Neural Model of Human Image Categorization

Eric Hunsberger, Peter Blouw, James Bergstra, Chris Eliasmith

Centre for Theoretical Neuroscience
University of Waterloo, Canada

Thursday, August 1, 2013

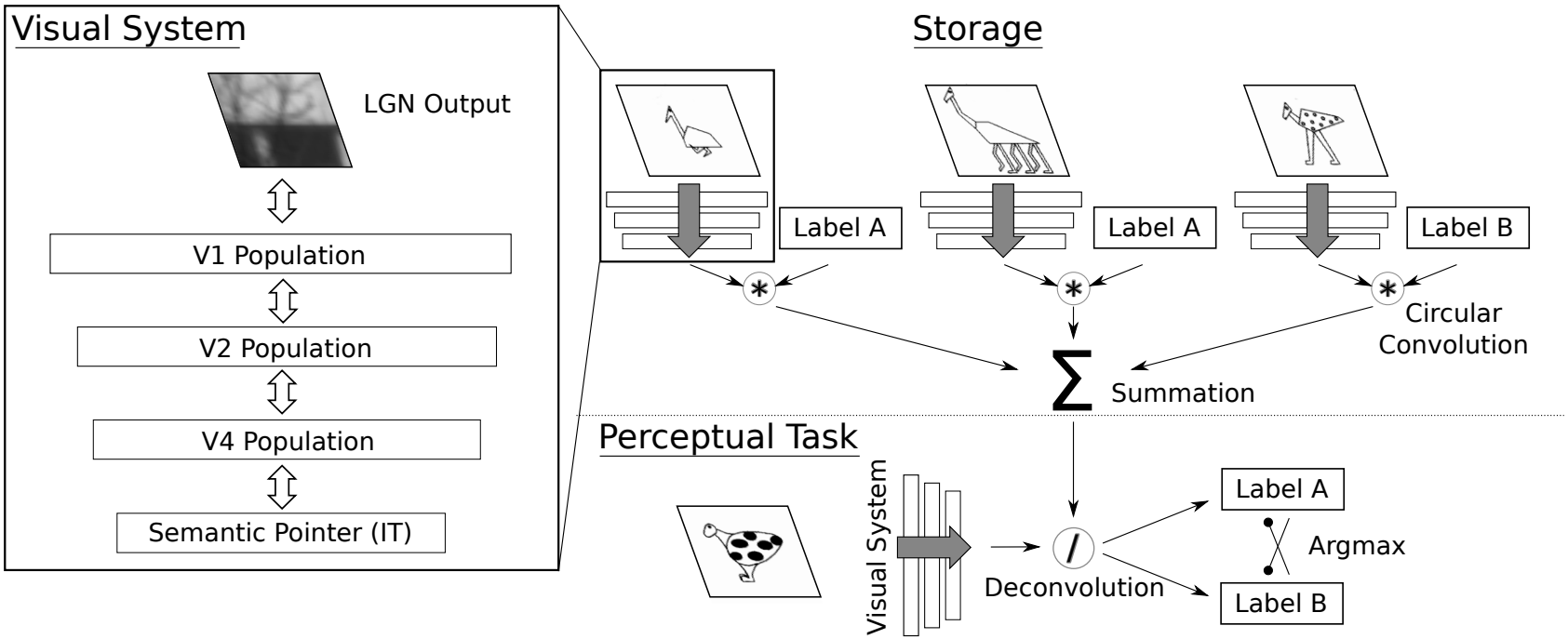
Goals

- Create a neural model of human image classification
- Make the model realistic
  - Spiking neurons
  - Multiple brain areas
  - Realistic parameters
- Reproduce results from two visual categorization experiments
  - Posner and Keele (1968): prototype theory
  - Regehr and Brooks (1993): exemplar theory

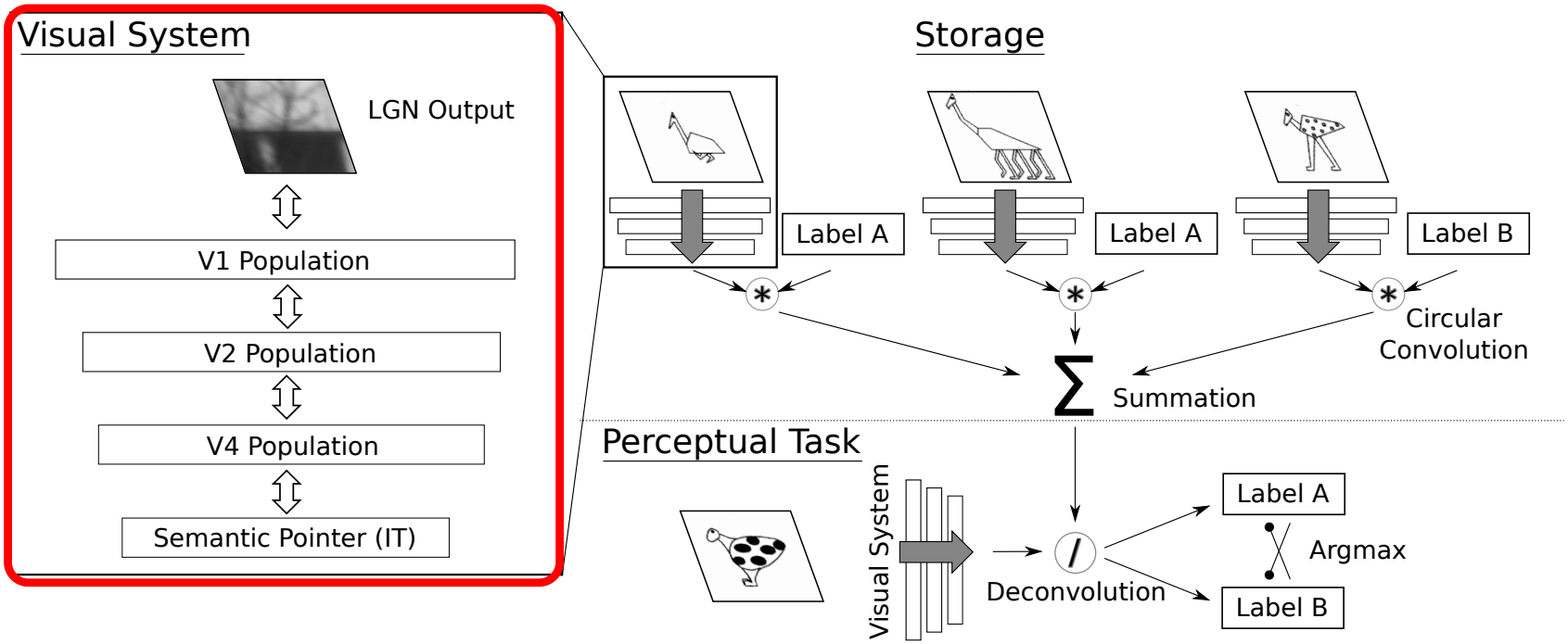
Model

- Starts at the perceptual level
- All implemented in spiking neurons
  - ~ 25000 simulated neurons
- Realistic parameters
- No free parameters
- Unchanged across multiple tasks

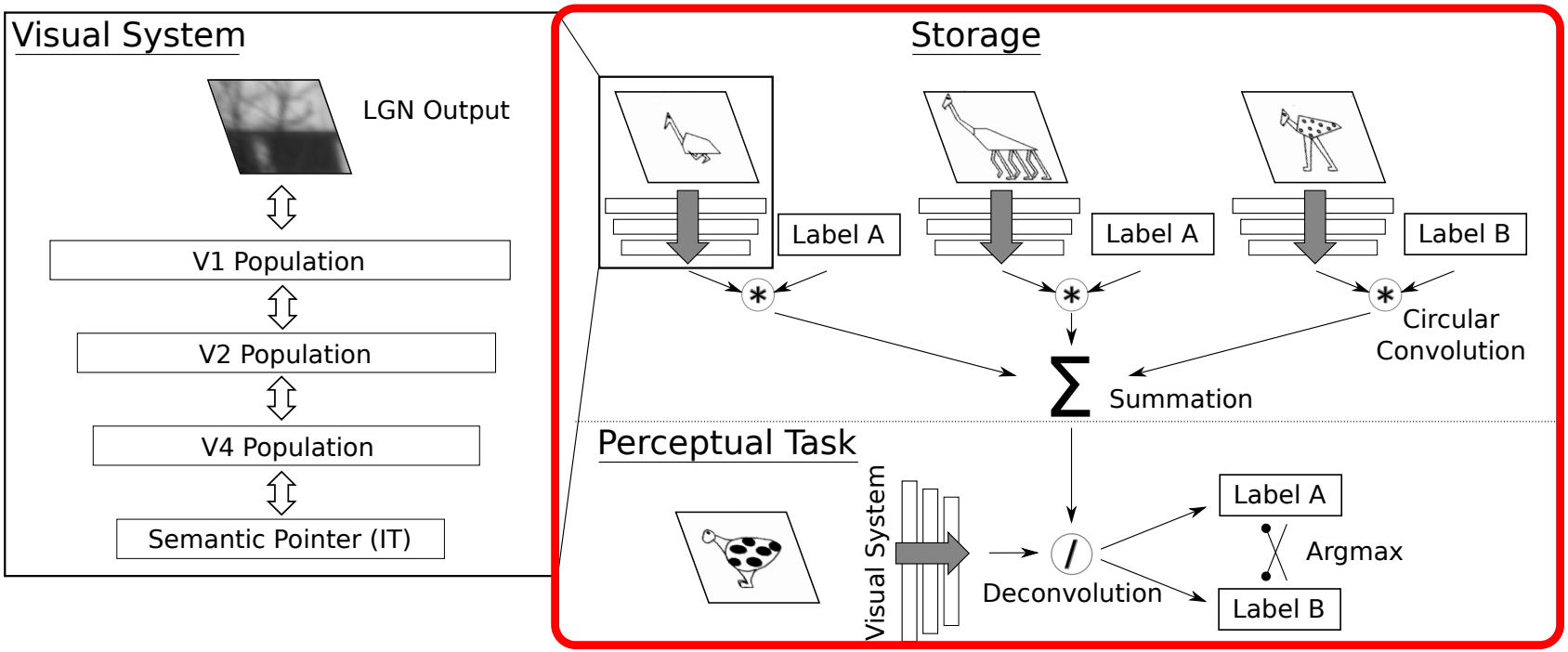
Model


Visual System
LGN Output
V1 Population
V2 Population
V4 Population
Semantic Pointer (IT)
Storage
Label A
Label A
Label B
Circular
Convolution
Summation
Perceptual Task
Visual System
Deconvolution
Label A
Argmax
Label B

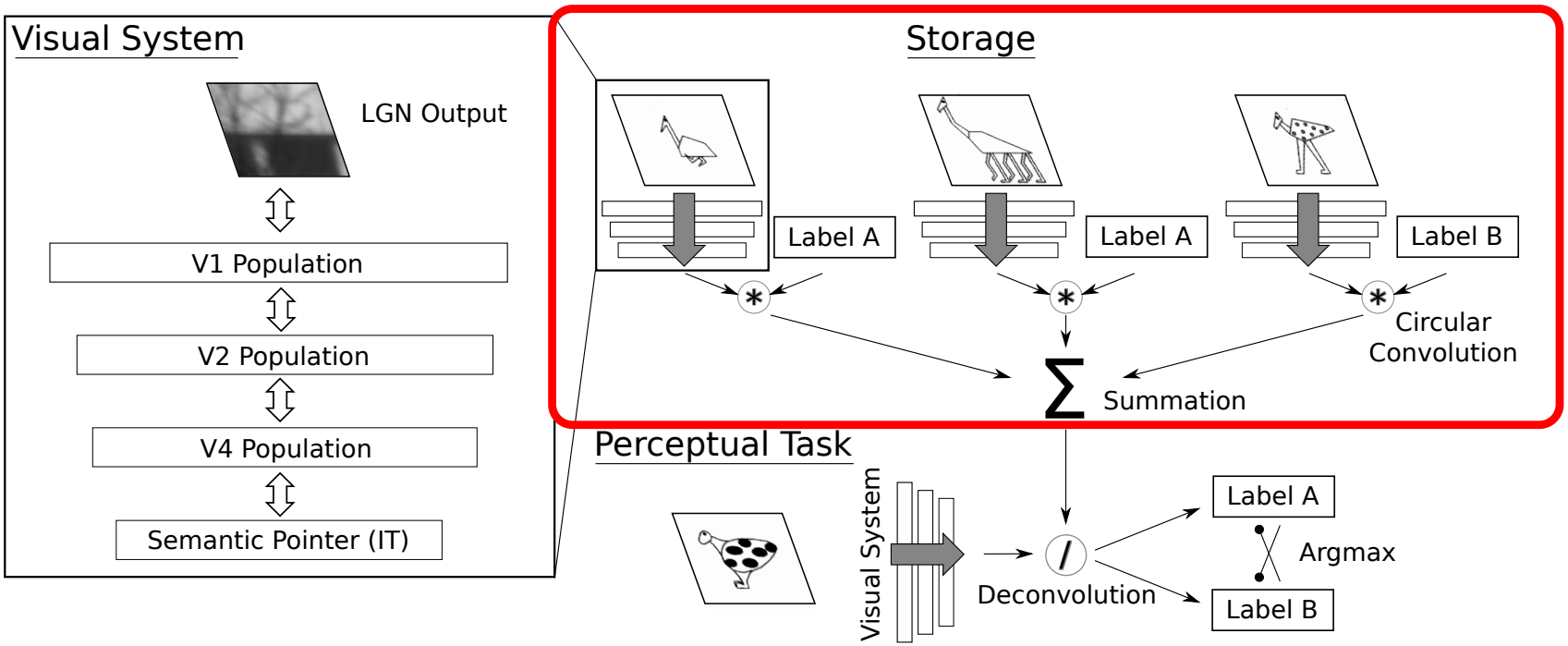
Model


Visual System
LGN Output
V1 Population
V2 Population
V4 Population
Semantic Pointer (IT)
Storage
Label A
Label A
Label B
Circular
Convolution
Summation
Perceptual Task
Visual System
Deconvolution
Label A
Argmax
Label B

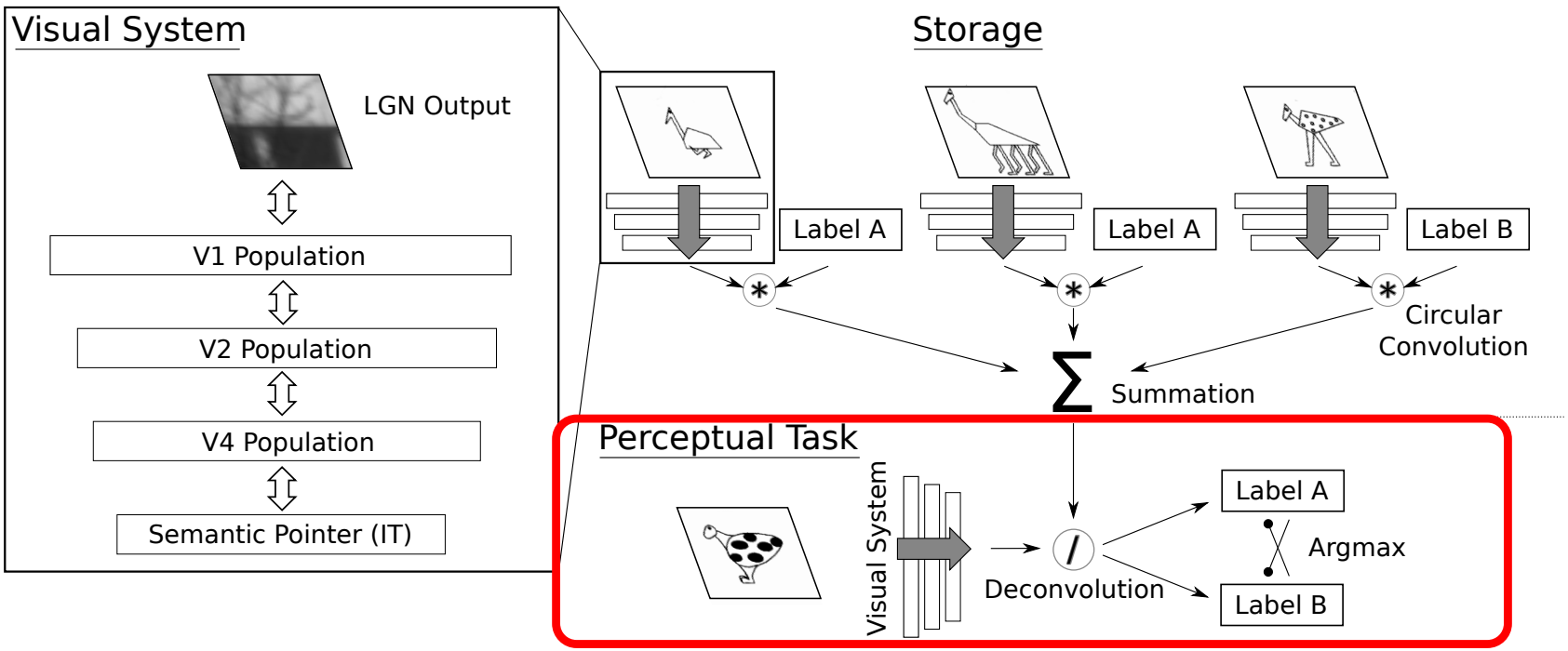
Model


Visual System
LGN Output
V1 Population
V2 Population
V4 Population
Semantic Pointer (IT)
Storage
Label A
Label A
Label B
Circular
Convolution
Summation
Perceptual Task
Visual System
Deconvolution
Label A
Argmax
Label B

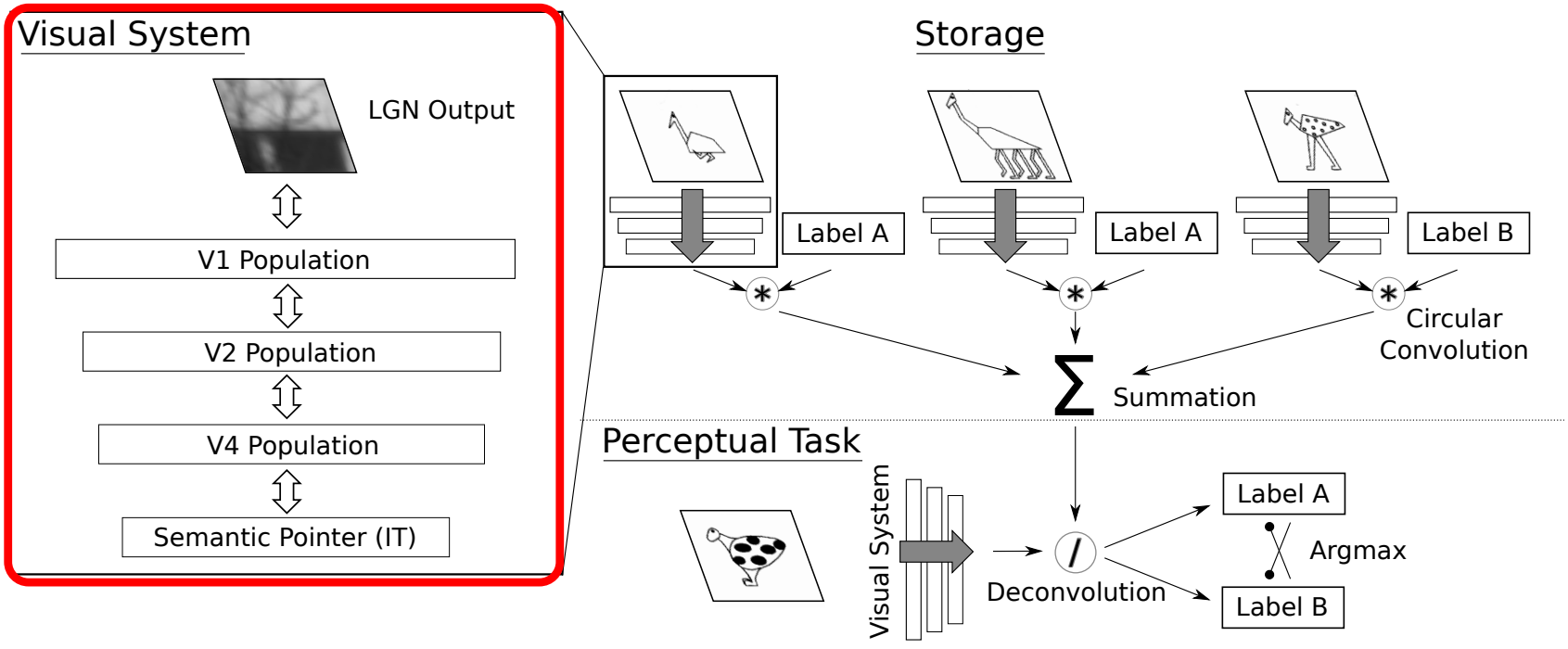
Model


Visual System
LGN Output
V1 Population
V2 Population
V4 Population
Semantic Pointer (IT)
Storage
Label A
Label A
Label B
Circular Convolution
Summation
Perceptual Task
Visual System
Deconvolution
Label A
Argmax
Label B

Model


Visual System
LGN Output
V1 Population
V2 Population
V4 Population
Semantic Pointer (IT)
Storage
Label A
Label A
Label B
Circular
Convolution
Summation
Perceptual Task
Visual System
Deconvolution
Label A
Argmax
Label B

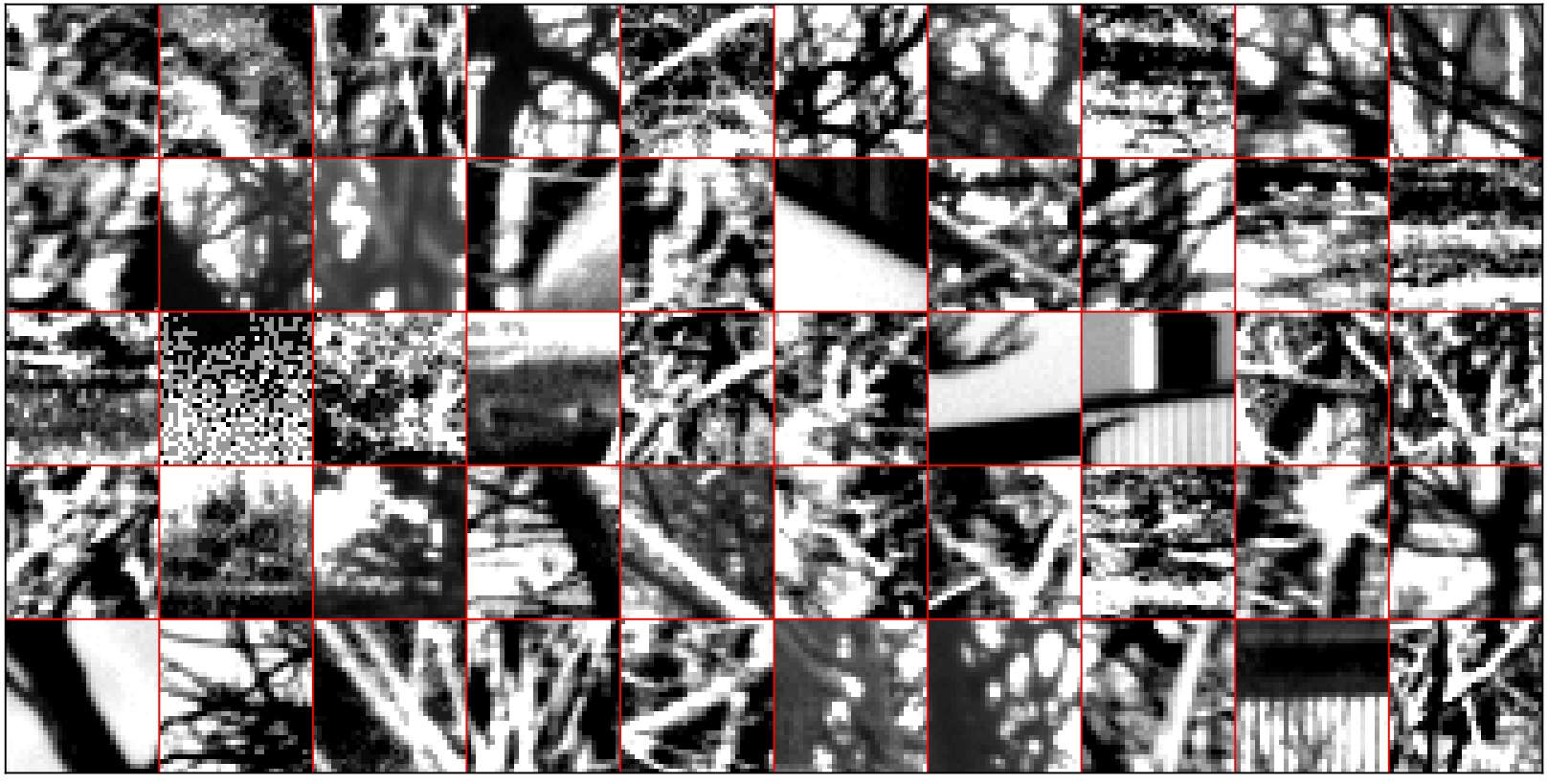
Visual System



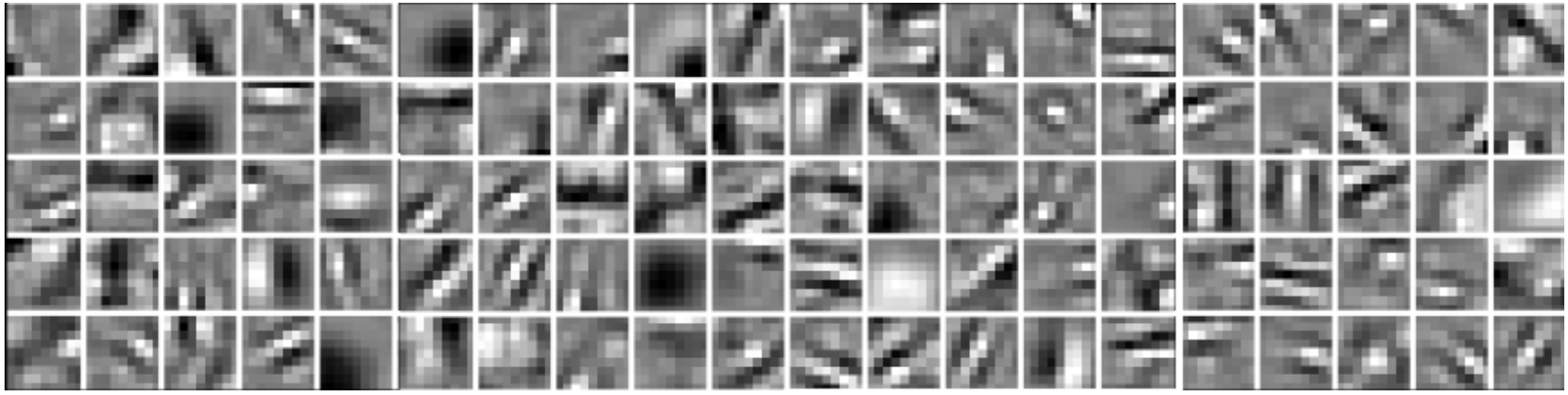
- Compress 900 pixel image into 225-vector
- Deep autoencoder with four layers
  - 2500, 900, 400, and 225 neurons respectively
  - Each layer trained on output of previous layer

Visual System Training

Trained on unwhitened image patches from the Van Hateren natural image database

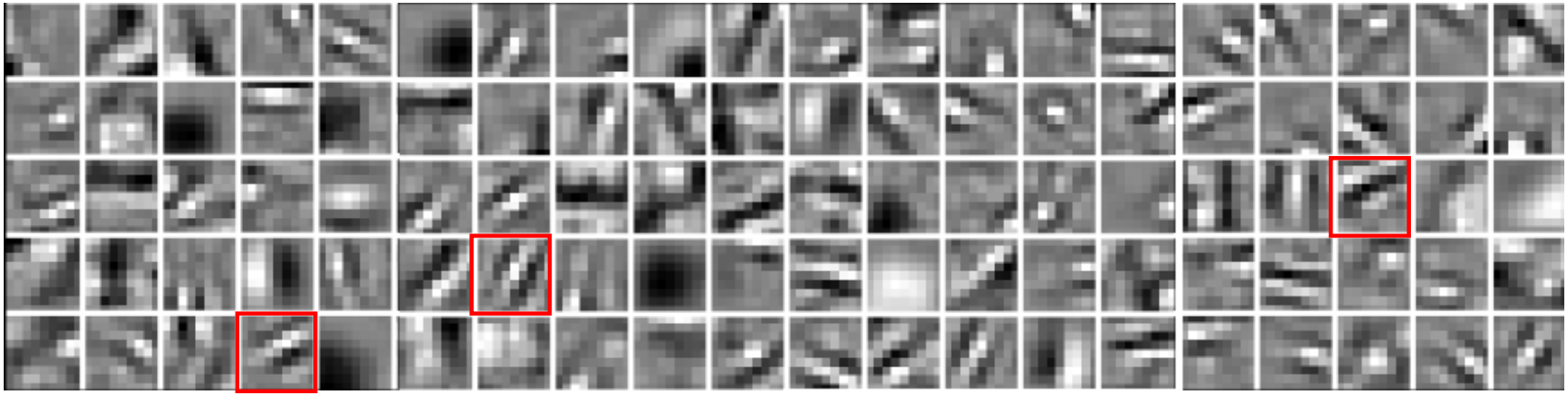


Visual System Results



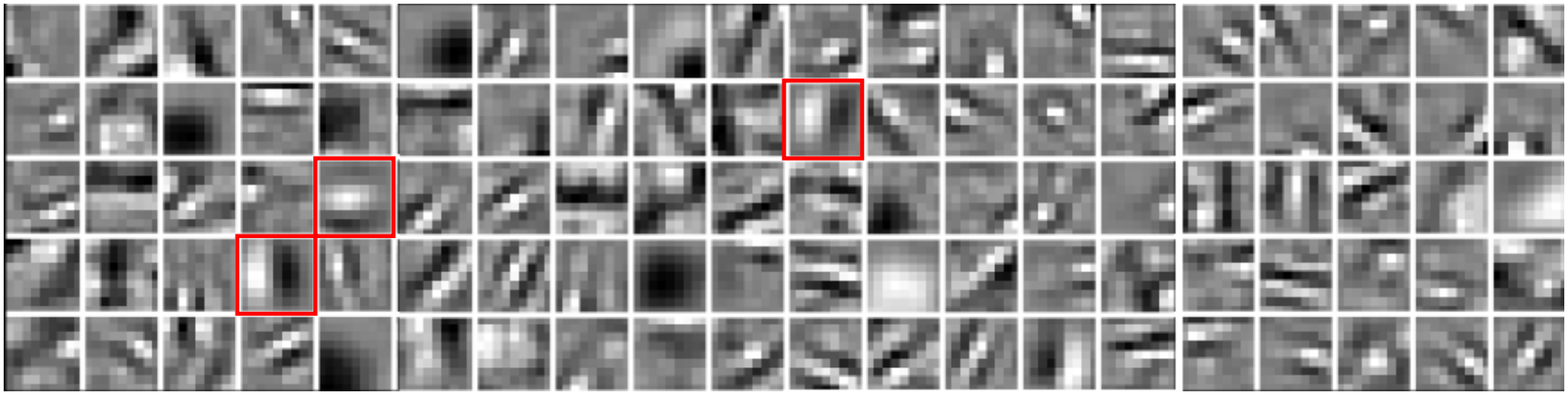
- Detect luminance edges at a variety of orientations and frequencies
- Some filters also detect general lightness or darkness

Visual System Results



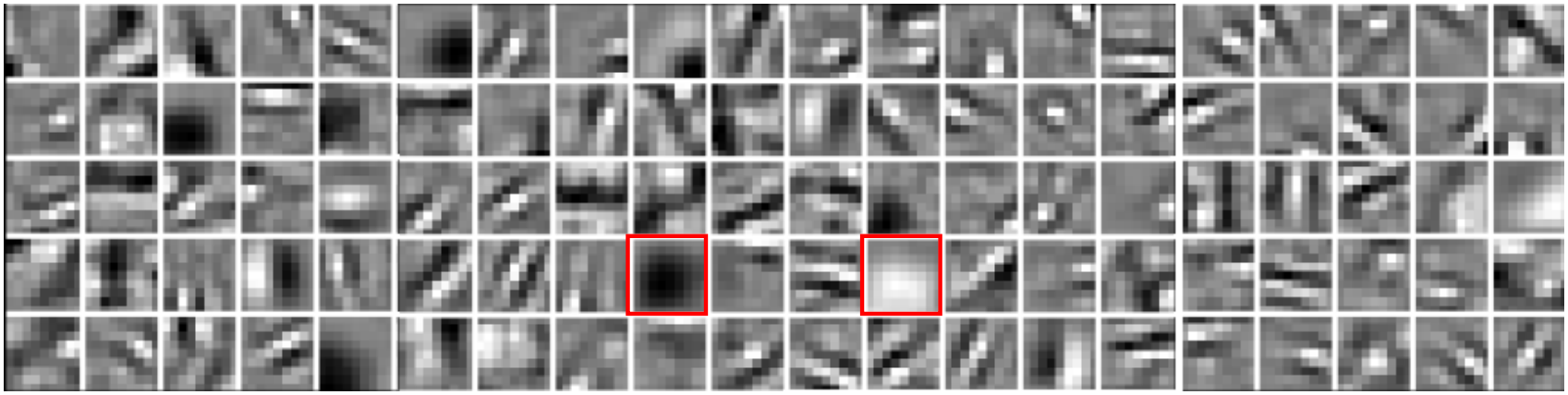
- Detect luminance edges at a variety of orientations and frequencies
- Some filters also detect general lightness or darkness

Visual System Results



- Detect luminance edges at a variety of orientations and frequencies
- Some filters also detect general lightness or darkness

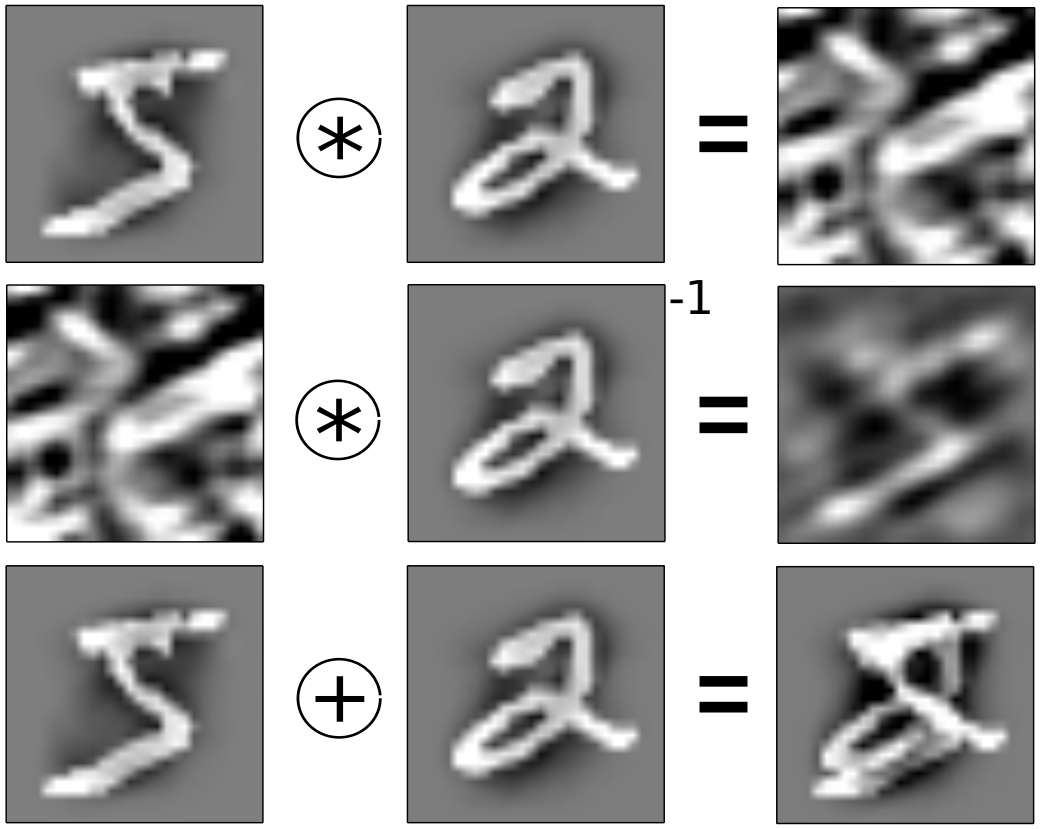
Visual System Results



- Detect luminance edges at a variety of orientations and frequencies
- Some filters also detect general lightness or darkness

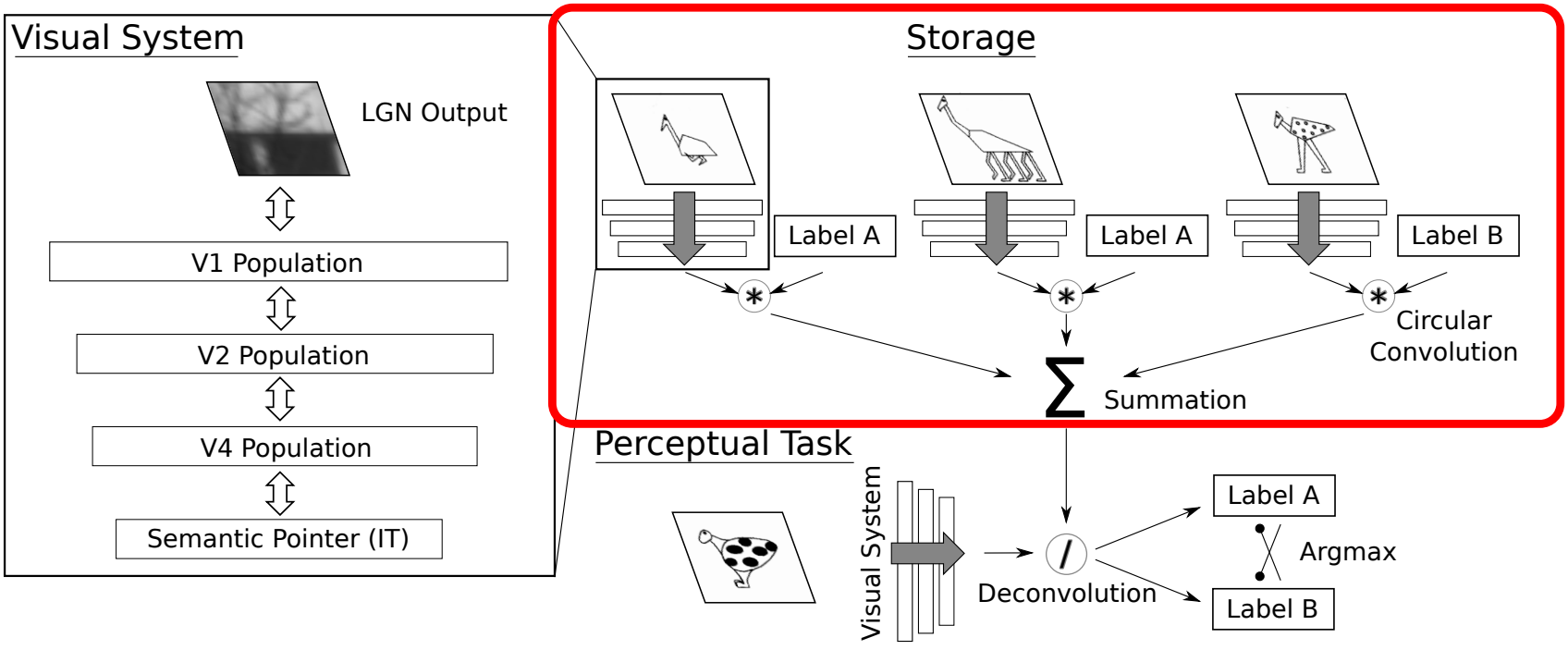
Classifier

- Based on Tony Plate's theory of Holographic Reduced Representations (HRRs)



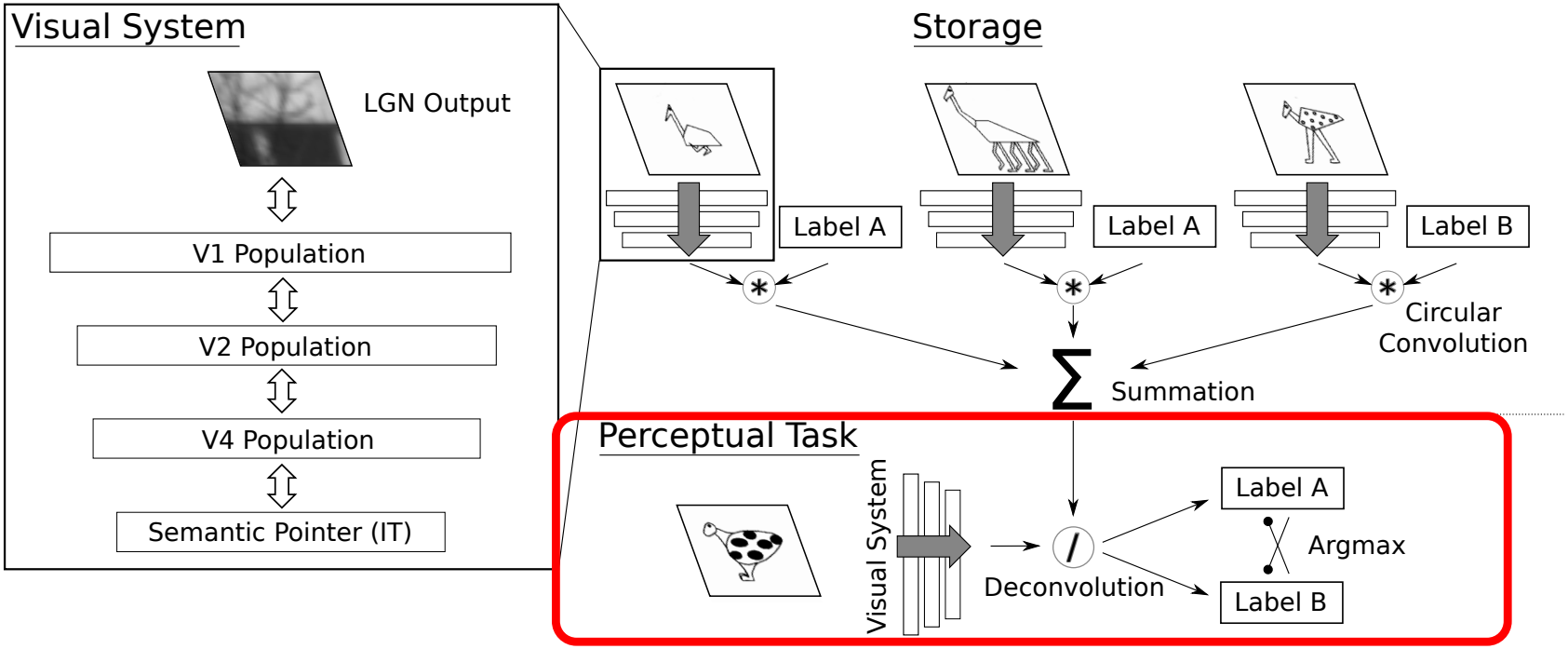
- Implemented in spiking neurons using the Neural Engineering Framework

Classifier: Training



$$M = \sum_{i=1}^{N} (P_i * L_i)$$

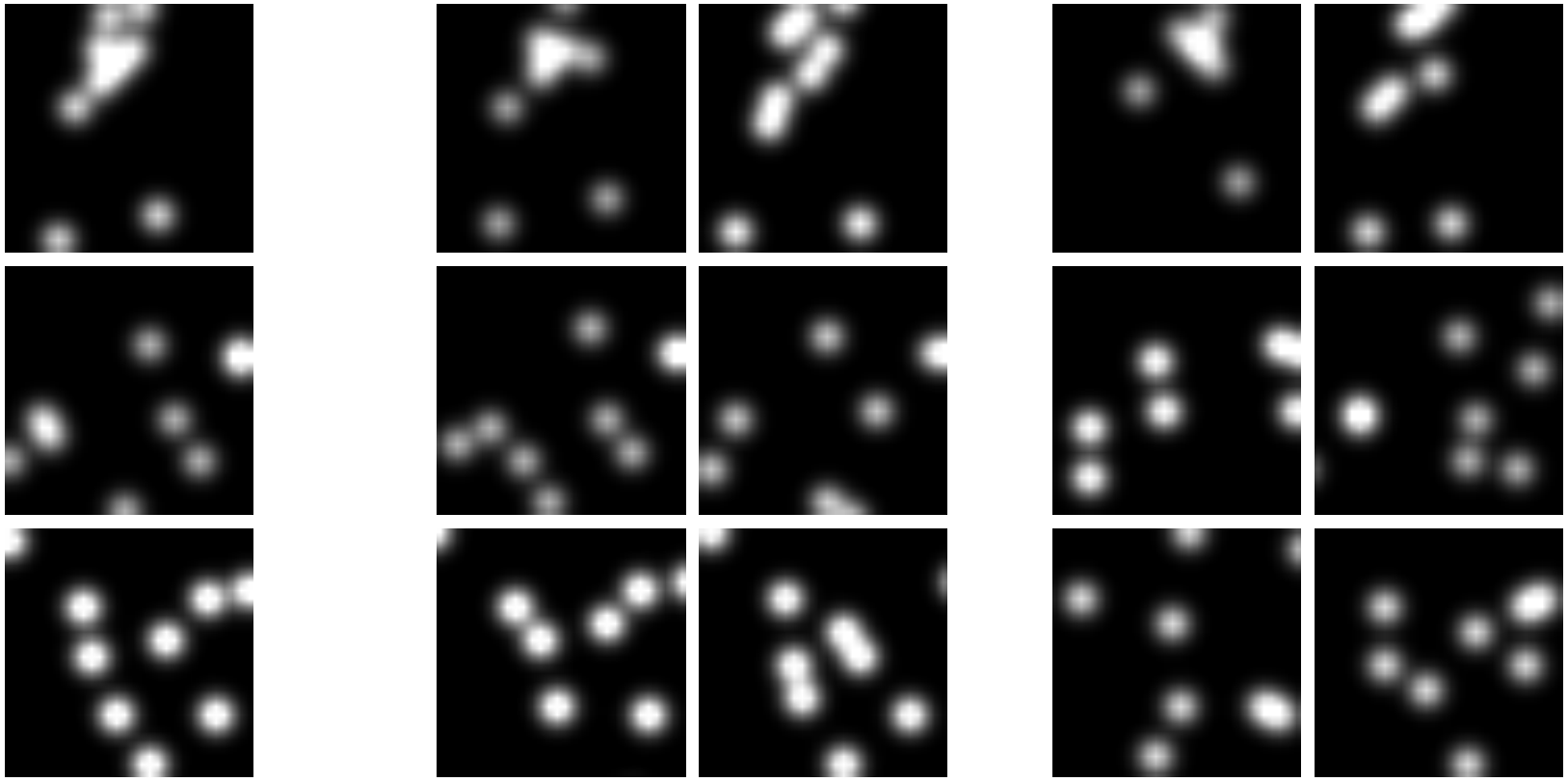
Classifier: Testing



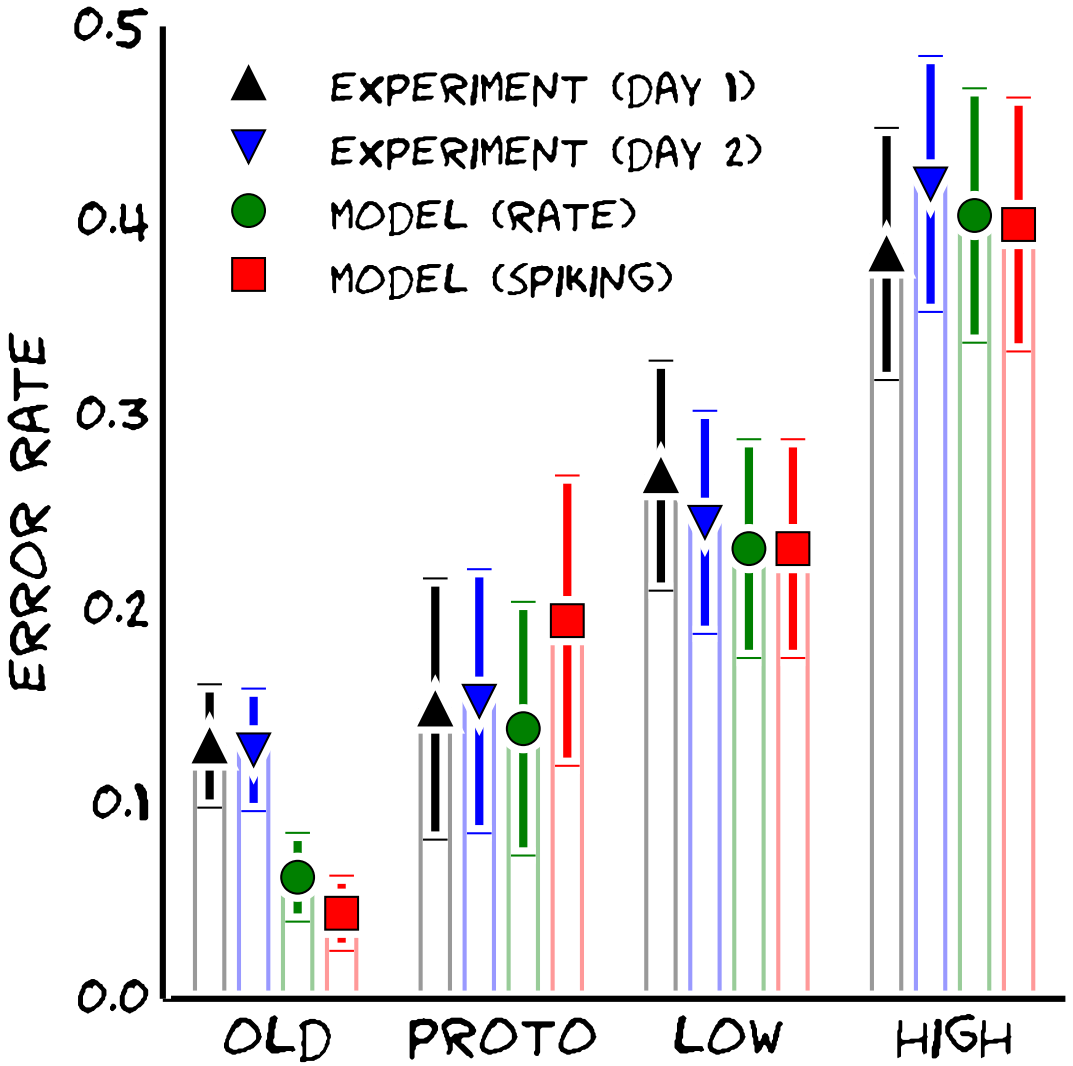
$$c = \text{argmax}_c \left[\left(P^{-1} * M \right) \cdot L_c \right]$$

Experiment 1: Posner & Keele

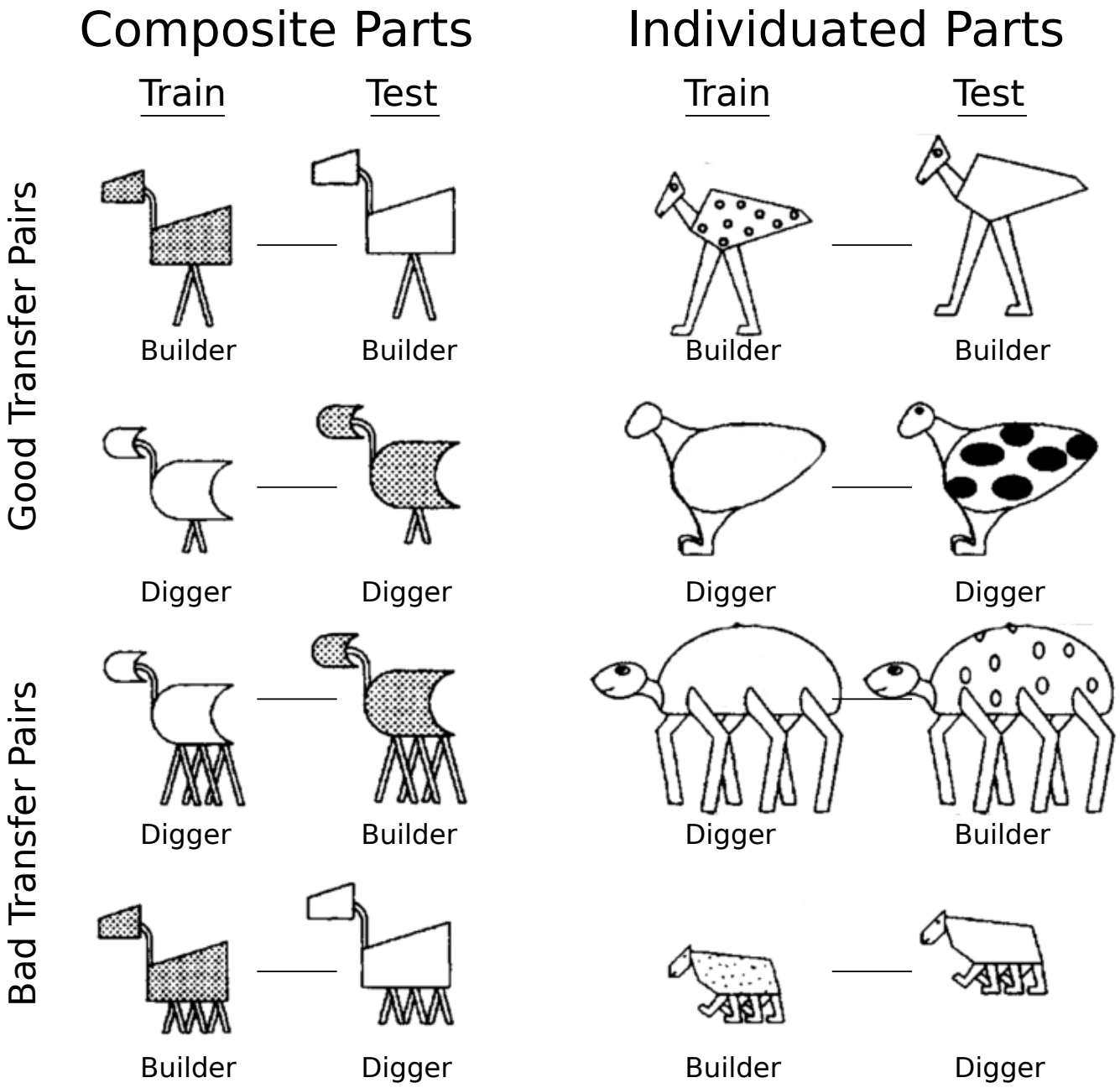
Prototype | Low Distortion | High Distortion



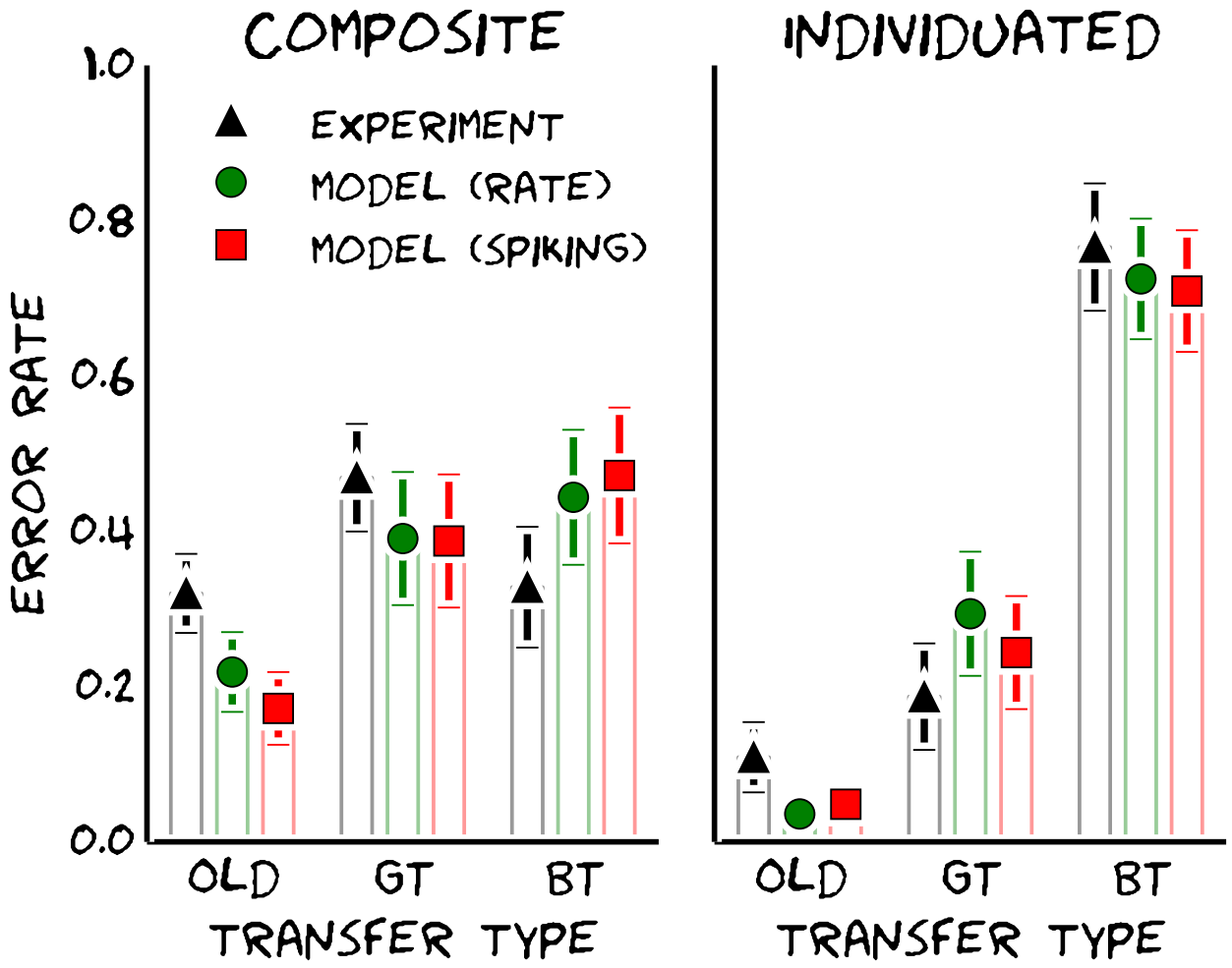
Experiment 1 Results



Experiment 2: Regehr & Brooks



Experiment 2 Results



Conclusions

Our model:

- Reproduces human performance on two distinct tasks
- Does not change at all between tasks
- Provides a basic explanation of how humans perform image categorization

Questions

Visual System Training

Each layer is an autoencoder with sparsity constraint:

$$O = \overbrace{\frac{1}{K}\sum_{i,k}\left(x_i^{(k)} - y_i^{(k)}\right)^2}^{\text{reconstruction error}} + \lambda \overbrace{\sum_{j}\left|q_j - \rho\right|}^{\text{sparsity error}}$$

$x_i^{(k)}$ visual node i, example k

$y_i^{(k)}$ reconstruction of visual node i, example k

q_j average activation of hidden node j

λ sparsity importance parameter

ρ sparsity level parameter

Visual System Training