A COMPUTATIONAL MODEL FOR THE BIOLOGICAL UNDERPINNINGS OF INFANT VISION AND FACE RECOGNITION

Laura A. Mueller
Department of Mathematics
University of Wisconsin, Madison WI 53706

Liya Wang
Department of Biomedical Engineering
University of Wisconsin, Madison WI 53706

Amir H. Assadi *
Department of Mathematics
University of Wisconsin, Madison WI 53706
assadi@math.wisc.edu

Abstract
A scientists' challenge is to understand how an infant's brain learns to efficiently process visual information. An infant's range of focus is short for several months, and faces enter into this range more often than any other scene. Within a few months of birth, the brain can differentiate faces from other objects and an infant can recognize a known face from a stranger's. The repetition of face stimulus in conjunction with the relatively short time it takes for recognition to occur suggests there may be more regularity among facial stimuli, or perhaps a more sensitive information processing mechanism biased towards facial stimuli.

Keyword
Face Recognition
Infant Vision
Principal Component Analysis
Backpropagation Neural Network

Vision has been a popular subject of study, but most research has focused on the fully developed visual systems of adult humans. To further understand the brain's role in perceptual processes such as face recognition, one needs focus on the visual system's development from infancy to adulthood. Infants experience the world through all senses, some more developed than others. The visual learning of a mother's face or other caretakers is accomplished by feedback from a combination of touch, taste, smell, and sound. During early developmental stages, there are communication pathways between the visual and other sensory areas of the cortex, showing how the biological network is self-organizing. For theoretical study, a model with feedback, as if from other senses, provides comparable "behavioral" output.

---

* Laura Mueller was partially supported through NSF-KDI-LIS and other grants to Amir Assadi, by the NSF-VIGRE to UW-Madison Mathematics Department, and by the Office of Provost and Vice Chancellor for Academic Affairs, UW-Madison for the Symmetry Project. Amir Assadi and Liya Wang are partially supported through an NSF-KDI-LIS grant.

The authors acknowledge helpful and inspiring discussion with Dr. Michael Struck of Dept of Ophthalmology, UW-Madison.

In this project, we investigate the learning process of face recognition for an infant's brain, using simulation by an artificial neural network (ANN) as a model. The biological hypotheses for this model are based on plasticity and well-established behavioral research on the role of response to low frequencies in early stages [1]. Learning in the brain occurs through creation or elimination of synapse connections or changes in synapse connection strengths in a neural pathway. If a neuron repeatedly contributes to the firing of another neuron, the strength between those two neurons increases. The Hebbian learning rule reflects this biology in ANN's by strengthening the connection weights between connected neurons. After training the network on a set of "typical faces" and also on a few non-facial stimuli, the ANN with tested with a previously unseen face as well as new objects. Higher success rates in tests for recognizing faces compared to other objects by the ANN suggests regularity amongst low frequency components of faces is essential for the brain to learn more efficiently to recognize a face.

The model for infant vision consists of a Multilayer Perceptron, based on a Backpropagation Algorithm, used to classify images as a face or non-face [Fig 3]. Images are filtered prior to being fed to the neural network (NN). In accordance to behavioral research, early infant vision responds to low frequencies only. A low-pass filter is first used to remove high frequency components in the images [Fig 1,2]. Then Generalized Hebbian Learning is applied for principal components extraction, using the statistical technique Principal Component Analysis (PCA). PCA consists of solving the correlation matrix of a matrix of images for eigenvalues and eigenvectors. The number of eigenvalues carrying at least 90% of the information indicates the number of eigenfaces to be used as inputs for the NN. The Back Propagation algorithm involves the feedforward step, where each input is multiplied by a weight matrix and the hyperbolic tangent function is taken of the resulting matrix. Backpropagation involves finding the error between the expected output and actual output, altering the weight matrices, and feeding the network forward again. This cycle (termed an epoch) continues until the error is no greater than 10% [Fig 4].

All the training and testing samples images represent 5% of the entire data set after passing the same low-pass filter. During the training process, the network performed with 100% accuracy for 30 training samples [Fig 4]), of which 16 were face images and 14 were non-face images. During testing, the network performed with 100% accuracy for the test samples, of which 4 were face images and 4 were non-face images.

Equipped with no previous knowledge of facial structure, the NN learned to distinguish faces from non-faces accurately in a similar manner that an infant's brain builds the neural connections after birth. Preprocessing images to remove high frequency eliminated unnecessary information from the images, helping the NN to efficiently learn the face versus non-face distinction. Successful results suggest this artificial neural network serves as a biologically viable model to explain the learning process an infant brain undergoes for face recognition.

Bibliography

[1]Banks, M. "The Role of Low Frequency in Infant Vision", PNAS, 1981.
[2]Barlow, Horace. The Exploitation of Regularities in the Environment by the Brain, Brain and Behavioral Sciences. May, 2001.
[3]Barlow, Horace and Tripathy, Srimant P. Correspondence Noise and Signal Pooling in the Detection of Coherent Visual Motion, The Journal of Neuroscience. October 15, 1997. 17(20)7954-7966.
[4]Crick, Watson. The Astonishing Hypothesis: The Scientific Search for the Soul. NY, NY: Touchstone, 1994.

[5]Dan, Yang, et. al. Efficient Coding of Natural Scenes in the Lateral Geniculate Nucleus: Experimental Test of a Computational Theory, The Journal of Neuroscience. May 15, 1996. 16(10)3351-3362.
[6]Palmer, Stephen. Vision Science: Photons to Phenomenology. Cambridge, Mass: MIT Press, 1999.
[7]Smith, Murray. Neural Networks for Statistical Modeling. Boston, MA: International Thomson Computer Press, 1996.

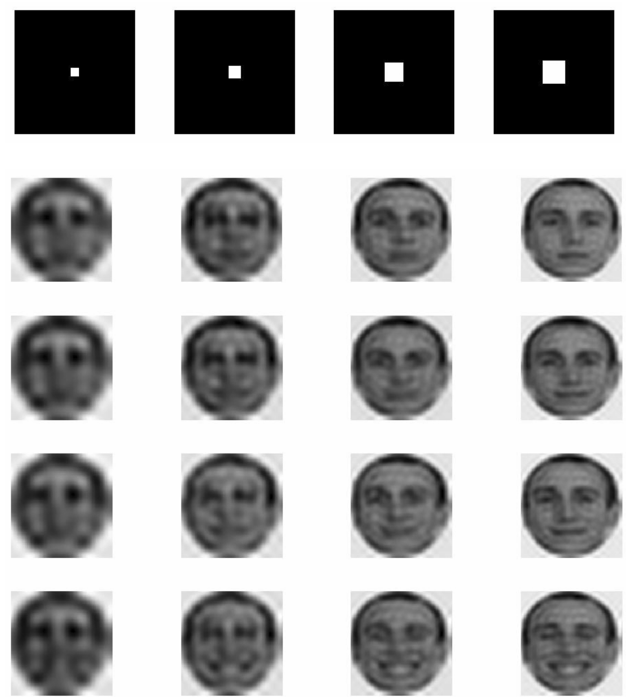
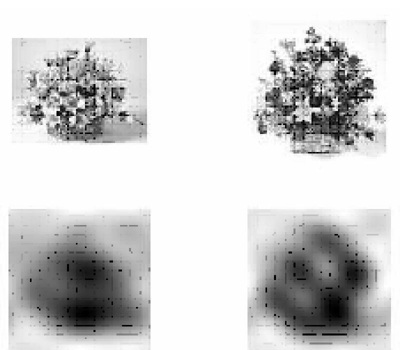


Figure 1: Low-pass frequency filters displayed on the top row. Corresponding filtered faces displayed below the filters.



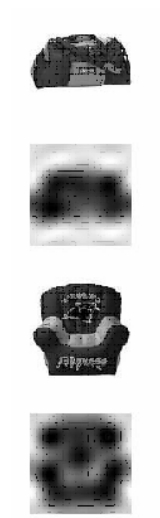


Figure 2: Examples of filtered non-face images.

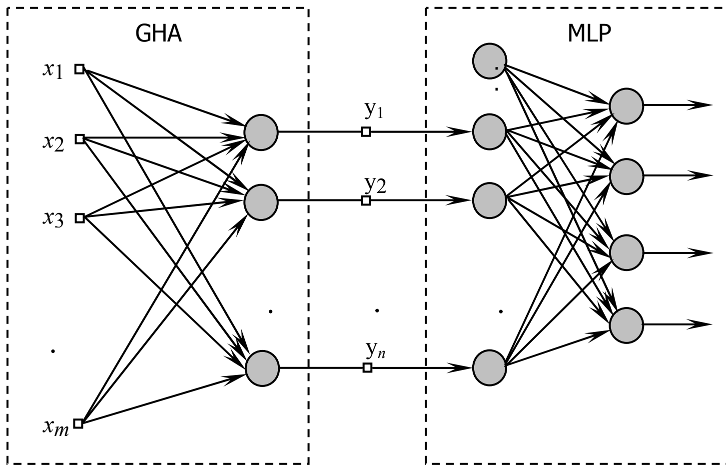


Figure 3: Neural Network Design

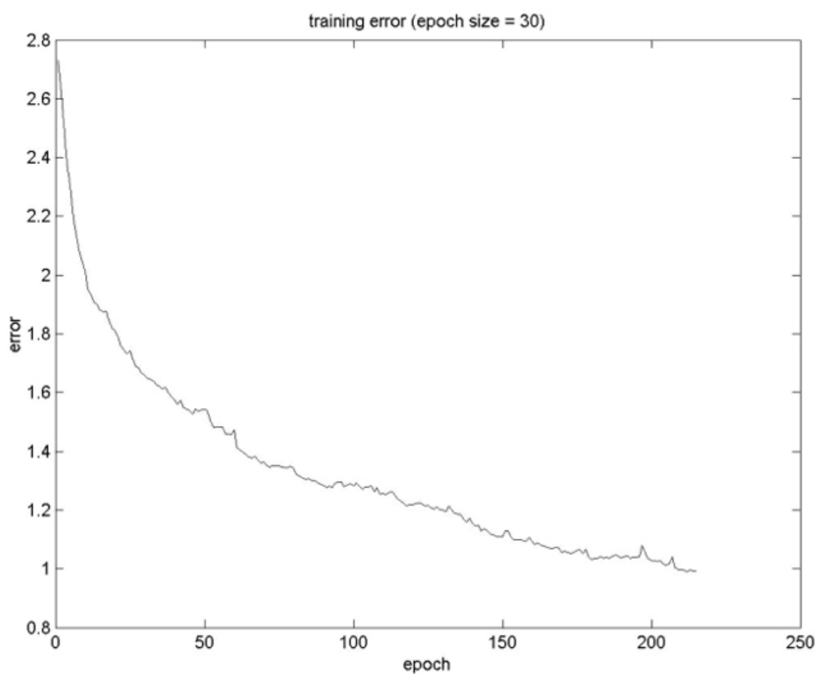


Figure 4: Training error for the neural network for about 200 epochs. Error decreases with each epoch until all the training samples are correctly identified.