Celebrated Backpropagation Applications

Sonar Target Recognition
(Gorman and Sejnowski, 1988)

- **2-layer network** distinguishes between reflected sonar signals of rocks vs. metal cylinders at bottom of Chesapeake Bay
- 60 input units, 2 output units
- input patterns based on **Fourier transform of raw time signals**
- tried varying numbers of hidden units (0, 3, 12, 24)
- best performance with 12 hidden units (close to 100% accuracy on training set)
- 85-90% classification accuracy for signals not in training set

IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, VOL. 36, NO. 7, JULY 1988

Learned Classification of Sonar Targets Using a Massively Parallel Network

R. PAUL GORMAN AND TERRENCE J. SEJNOWSKI

Gorman and Sejnowski, 1988

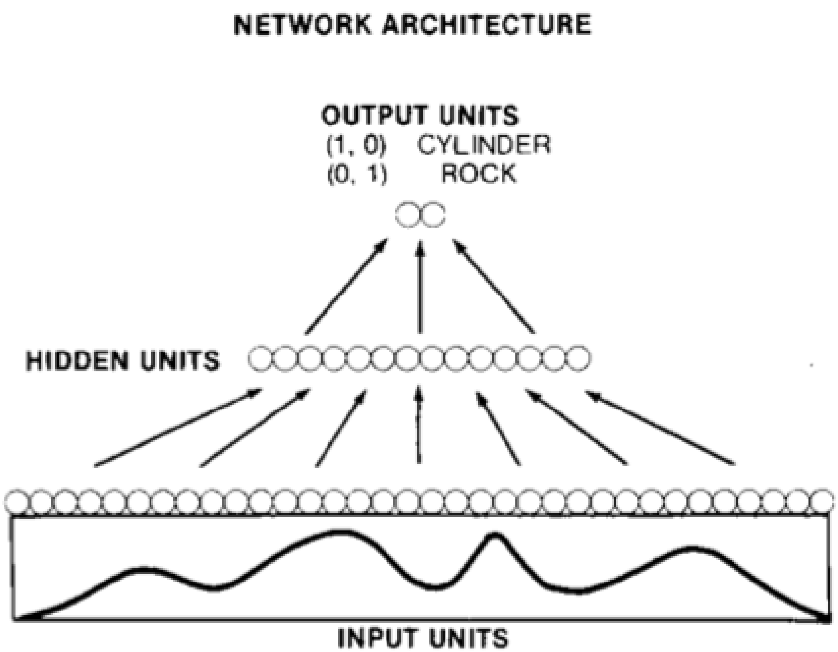


Fig. 2. Architecture of the network. The bottom layer consists of 60 processing units with their inputs ''clamped'' to the amplitude of the preprocessed sonar return. The hidden layer has modifiable weights on both the input and output connections, which allows the network to extract high-order features from the input waveform.

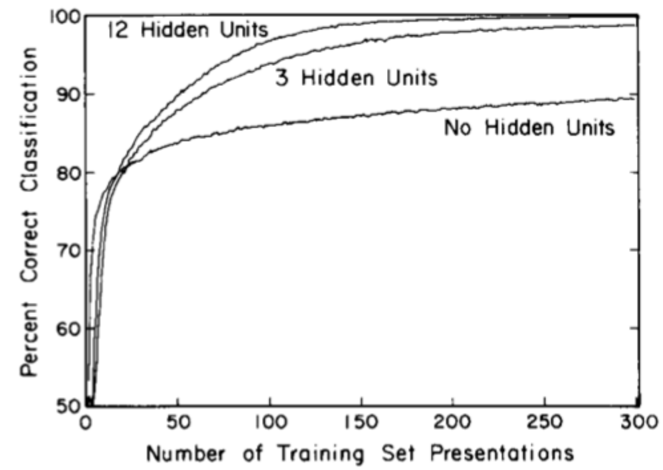


Fig. 4. Network learning curves for experiments with randomly chosen training sets. Each curve represents an average of 130 learning trials for a network with the specified number of hidden units.

Follow On:
Dror, Zagaeski, and Moss, 1995

Pergamon

Neural Networks, Vol. 8, No. 1, pp. 149–160, 1995

0893-6080/95 $9.50 + .00

0893-6080(94)00057-3

CONTRIBUTED ARTICLE

Three-Dimensional Target Recognition via Sonar: A Neural Network Model

ITIEL E. DROR,[1] MARK ZAGAESKI,[2] AND CYNTHIA F. MOSS[2]

[1]Miami University and [2]Harvard University

(*Received* 16 *March* 1993; *revised and accepted* 10 *May* 1994)

Echolocation à la Bat

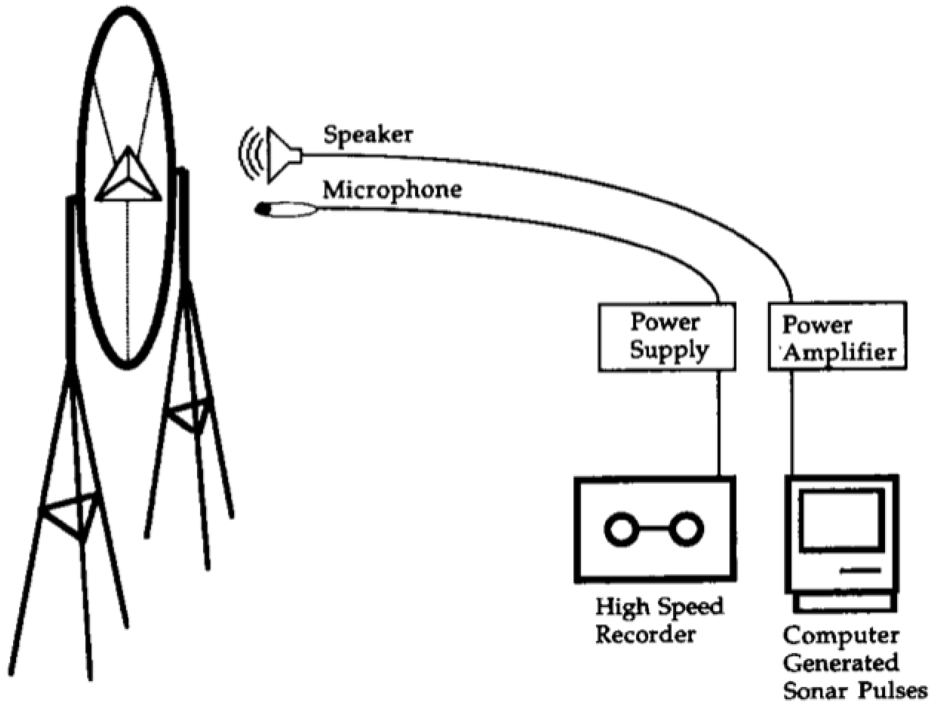


FIGURE 2. Schematic of apparatus used to record echoes from geometric shapes. The shape was suspended by fine thread in the center of a plastic hoop and positioned 34 cm from an ultrasound speaker and microphone. The speaker broadcast 1 ms FM sound sweeping from 100–25 kHz in two harmonics. Echoes from the target at a variety of orientations were recorded on an analog tape and digitized off-line.

Use of Time Bins

classification (2 output units)

6-neuron hidden layer

240 input units
(12 time bins
x 20 frequencies)

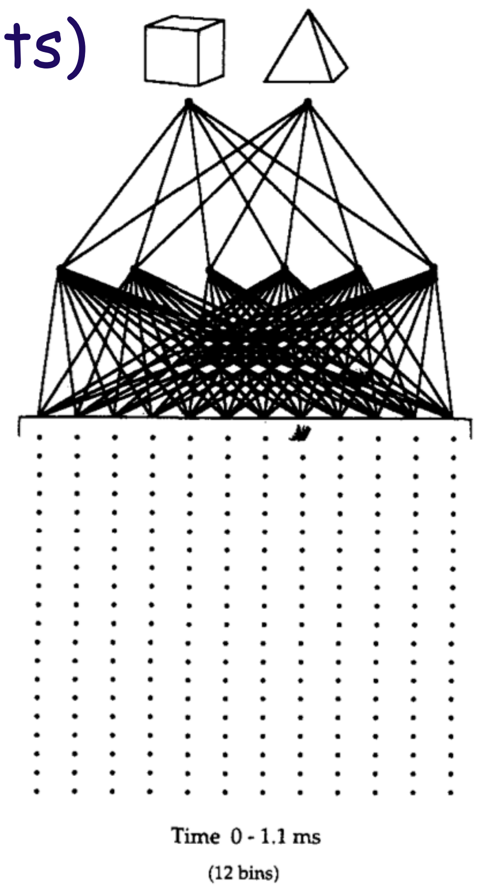


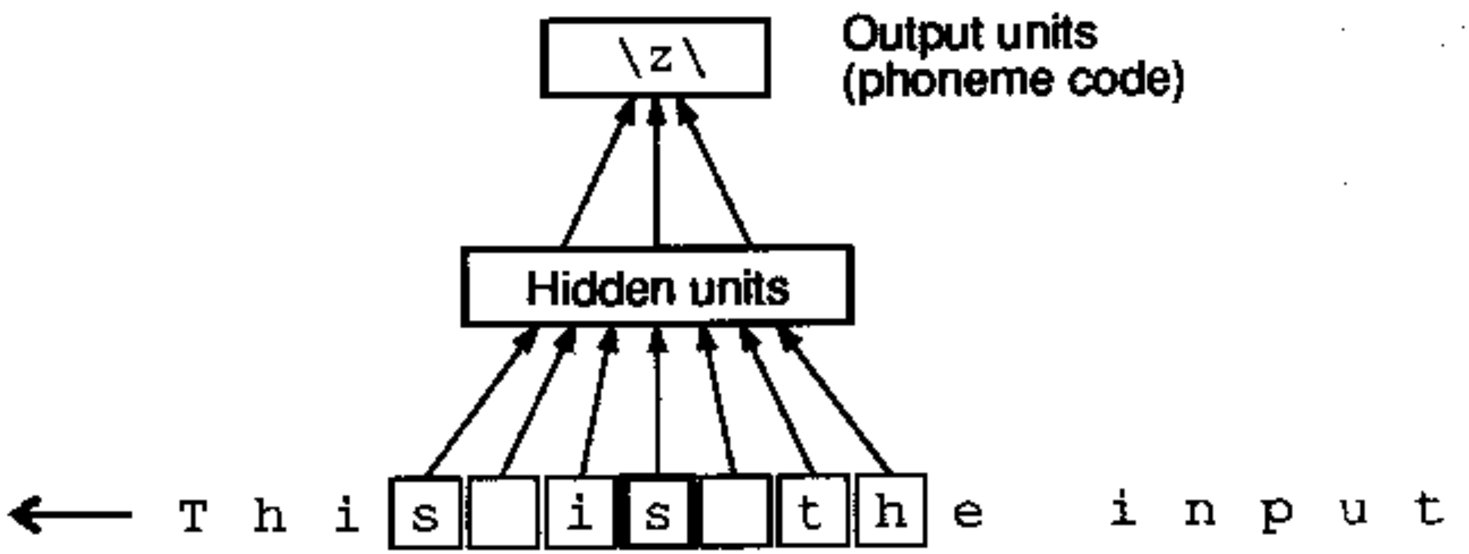
FIGURE 3. Sonar recognition of three-dimensional shapes. The architecture of the network: 240 input units, six hidden units, and two output units. The input layer in this example has been arranged to illustrate the layout of the spectrogram representation.

NETtalk
(Sejnowski and Rosenberg, 1986)

- Sejnowski, T. J. and Rosenberg, C. R. (1986) NETtalk: a parallel network that learns to read aloud, *Cognitive Science*, 14, 179-211.
- The authors taught a neural network to "read" using backpropagation.
- A stream of words were given to the network, along with the phoneme pronunciation of each in symbolic form.
- A speech generation device was used to convert the phonemes to sound.
- The network learned the phoneme pronunciations, thus was able to "speak" the words from a stream of words.

NETtalk

- 203-80-26 Multi-layer network
- Input is rolling sequence of 7 characters
- 203 = 7 x 29 different characters
- Output is the phoneme (if any) for the middle letter in the sequence



NETtalk

- 80 hidden units
- trained on 1024 words using a side-by-side English/phoneme source
- intelligible speech after 10 training epochs; 95% accuracy on training corpus after 50 epochs
- some hidden units developed meaningful responses (e.g., vowels vs. consonants)
- generalization: 78% accuracy on continuation of training text
- damaging the network produced graceful degradation, with rapid recovery on retraining
- DECtalk performed better, but used hand-coded linguistic rules developed over a decade

Sejnowski & Rosenberg, 1987

Complex Systems **1** (1987) 145–168

Parallel Networks that Learn to Pronounce English Text

Terrence J. Sejnowski
Department of Biophysics, The Johns Hopkins University, Baltimore, MD 21218, USA

Charles R. Rosenberg
Cognitive Science Laboratory, Princeton University, Princeton, NJ 08542, USA

GregTalk (2000)

- Similar results were reproduced in the spring 2000 offering of CS 152 by two Pomona College students:
  Greg Fishbein and Greg Schueler.

GloveTalk (Fels & Hinton, 1992)

IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 3, NO. 6, NOVEMBER 1992

Glove-Talk: A Neural Network Interface Between a Data-Glove and a Speech Synthesizer

S. Sidney Fels and Geoffrey E. Hinton

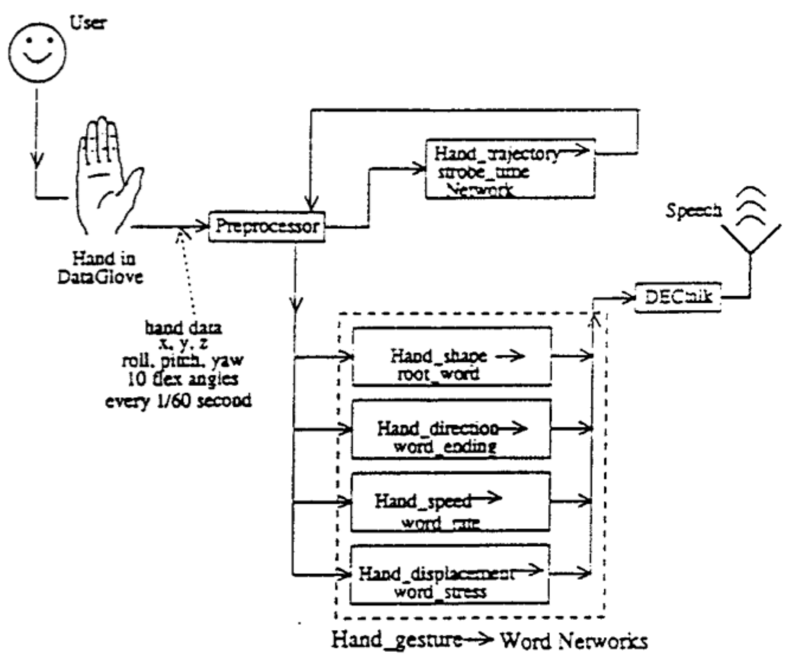


Fig. 1. Glove-Talk System.

TABLE I
EXAMPLES OF GLOVE-TALK LANGUAGE

root word	hand shape
come	
go	
I	
you	
short	

GloveTalk Network

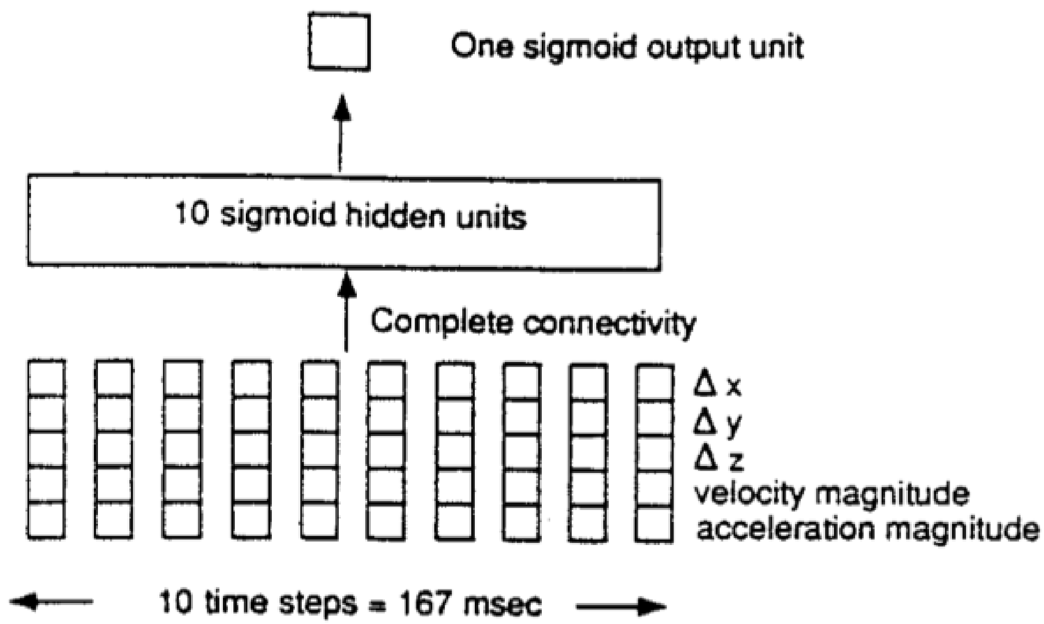


Fig. 2. The Strobe Network.

Performance: 1% incorrect, 5% no recognition

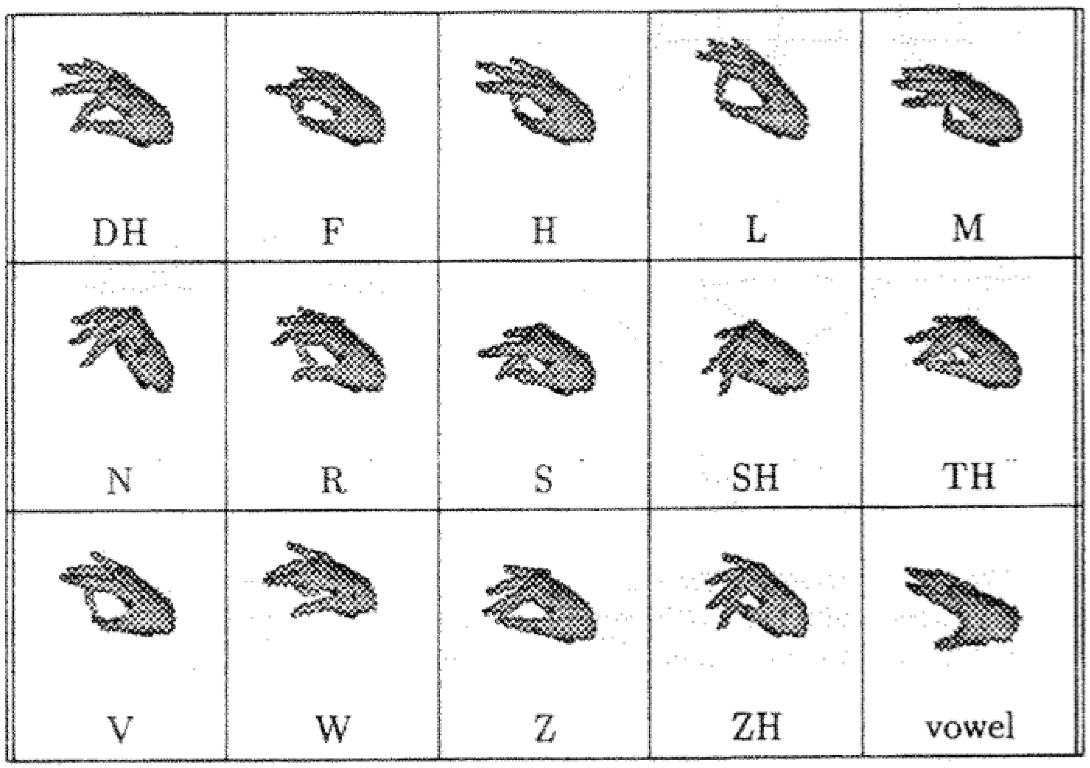
GloveTalk II (Fels & Hinton, 1995)

CHI '95 Proceedings

www.acm.org/sigchi/chi95/Electronic/documnts/papers/ssf_bdy.htm

TABLE I

STATIC GESTURE-TO-CONSONANT MAPPING FOR ALL PHONEMES. NOTE, EACH GESTURE CORRESPONDS TO A STATIC NONSTOP CONSONANT PHONEME GENERATED BY THE TEXT-TO-SPEECH SYNTHESIZER



GloveTalk II

Demo: http://www.youtube.com/watch?v=Mb7K2IpZn6E

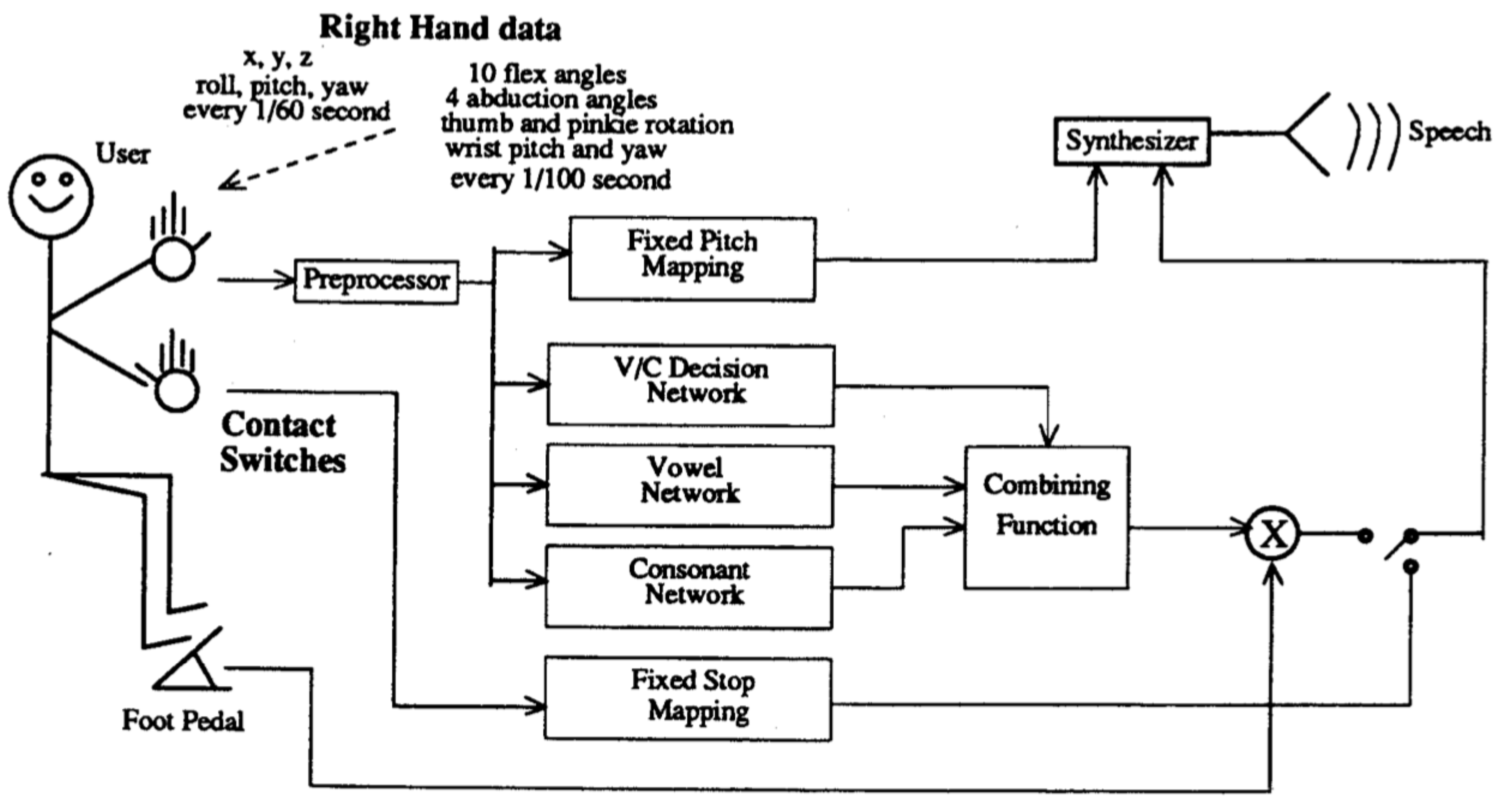
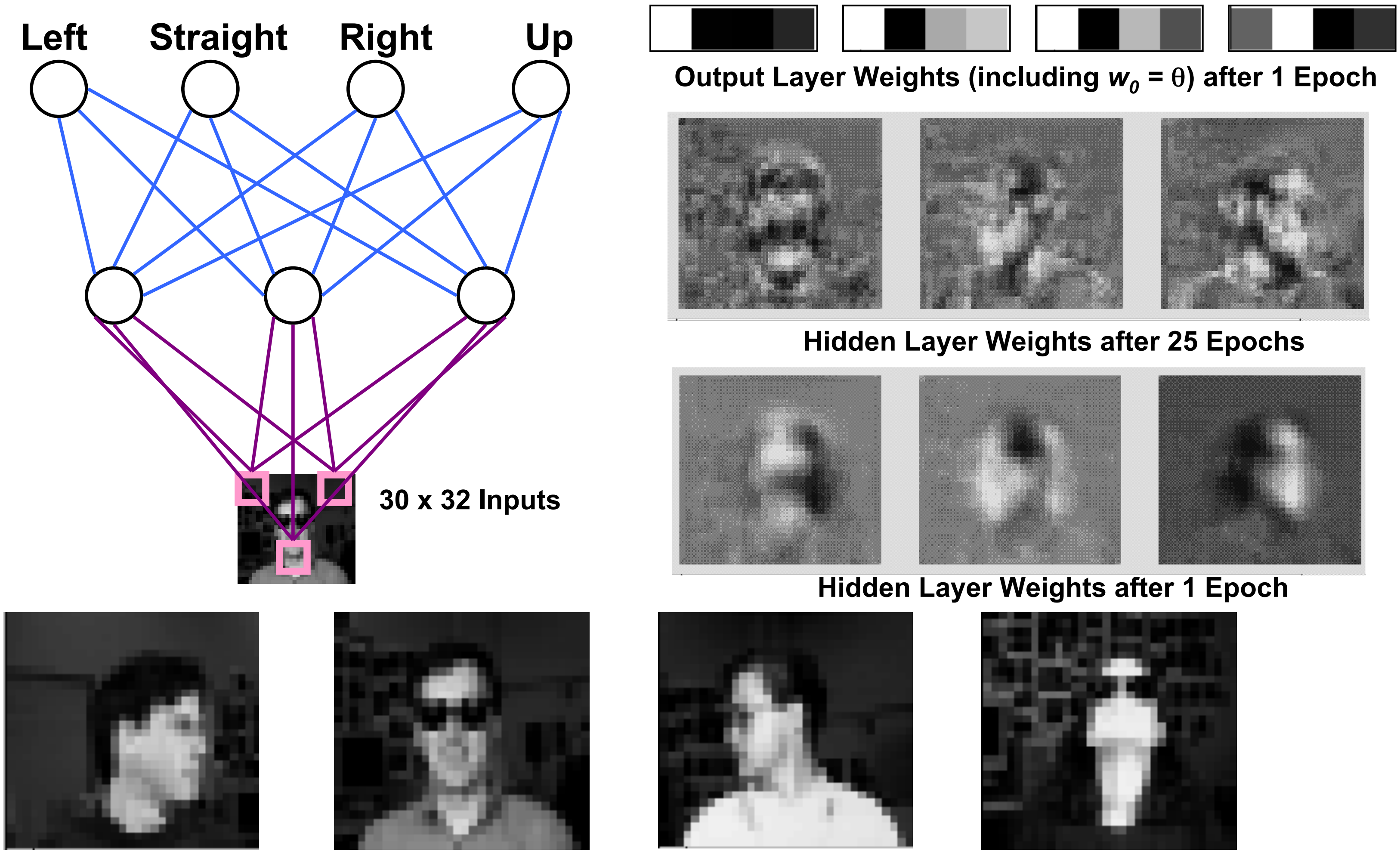


Fig. 2. Block diagram of Glove-TalkII: input from the user is measured by the Cyberglove, polhemus, ContactGlove, and foot pedal, then mapped using neural networks and fixed functions to formant parameters which drive the parallel formant synthesizer [12].

Glove-Talk II Summary

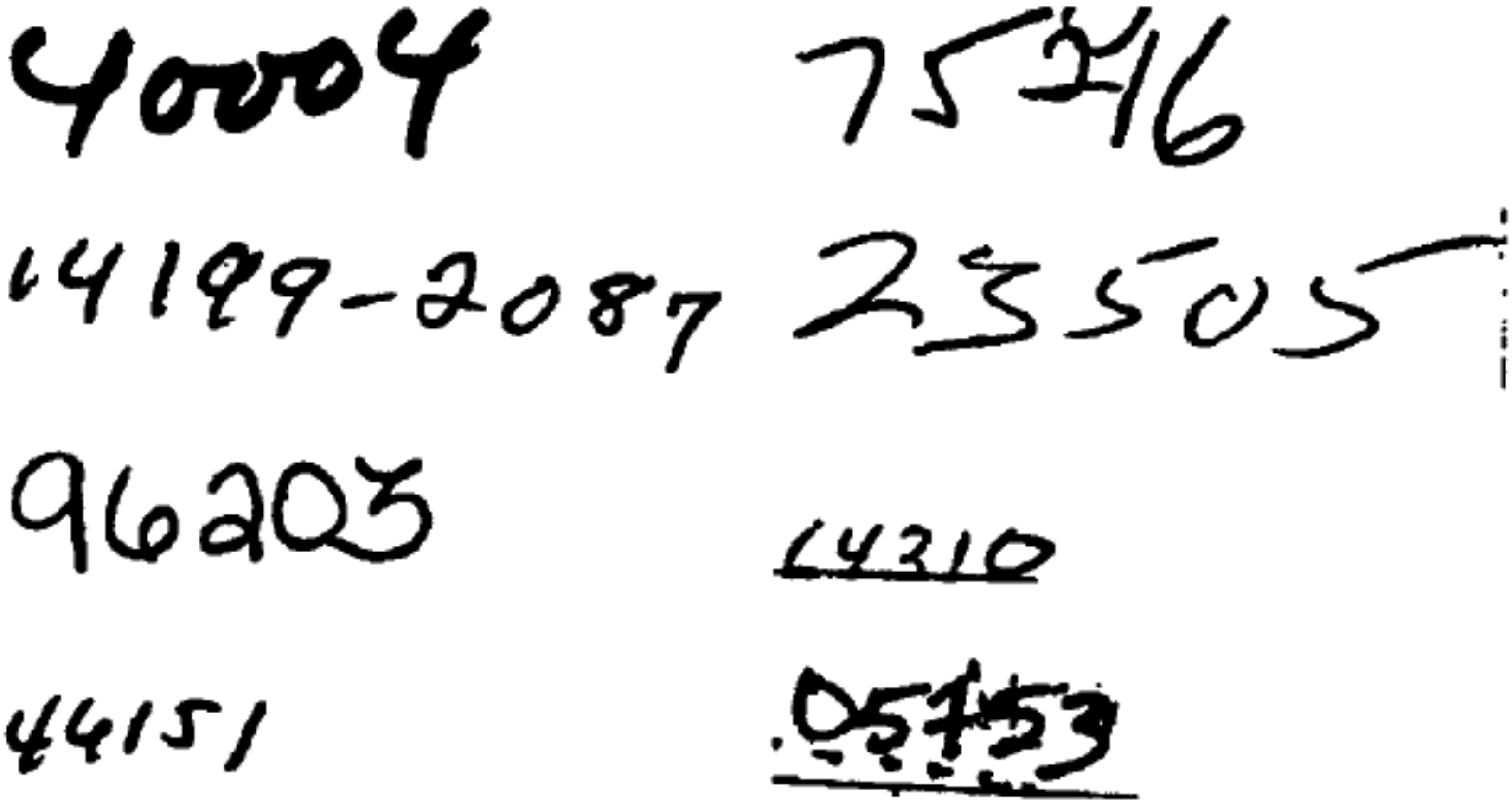
- One subject was trained to use Glove-TalkII. After 100 hours of practice he is able to speak intelligibly. The subject passed through 8 distinct stages while he learned to speak. His speech is fairly slow (1.5~to~3 times slower than normal speech) and somewhat robotic. It sounds similar to speech produced with a text-to-speech synthesizer but has a more natural intonation contour which greatly improves the intelligibility and naturalness of the speech. Reading novel passages intelligibly usually requires several attempts, especially with polysyllabic words. Intelligible spontaneous speech is possible but difficult.

Face&Pose Recognition (Mitchell, 1997)



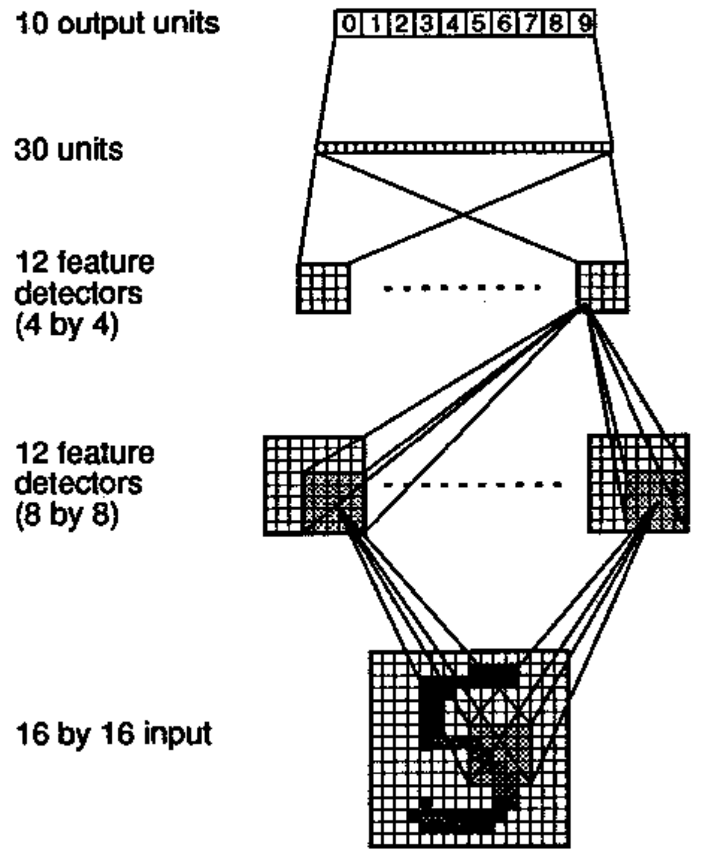
- 90% Accurate Learning Head Pose and in Recognizing 1-of-20 Faces
- http://www.cs.cmu.edu/~tom/faces.html

Zipcode Recognition (Yann LeCun, 1990)



Normalize Digits First

Network Structure



Feature Detectors

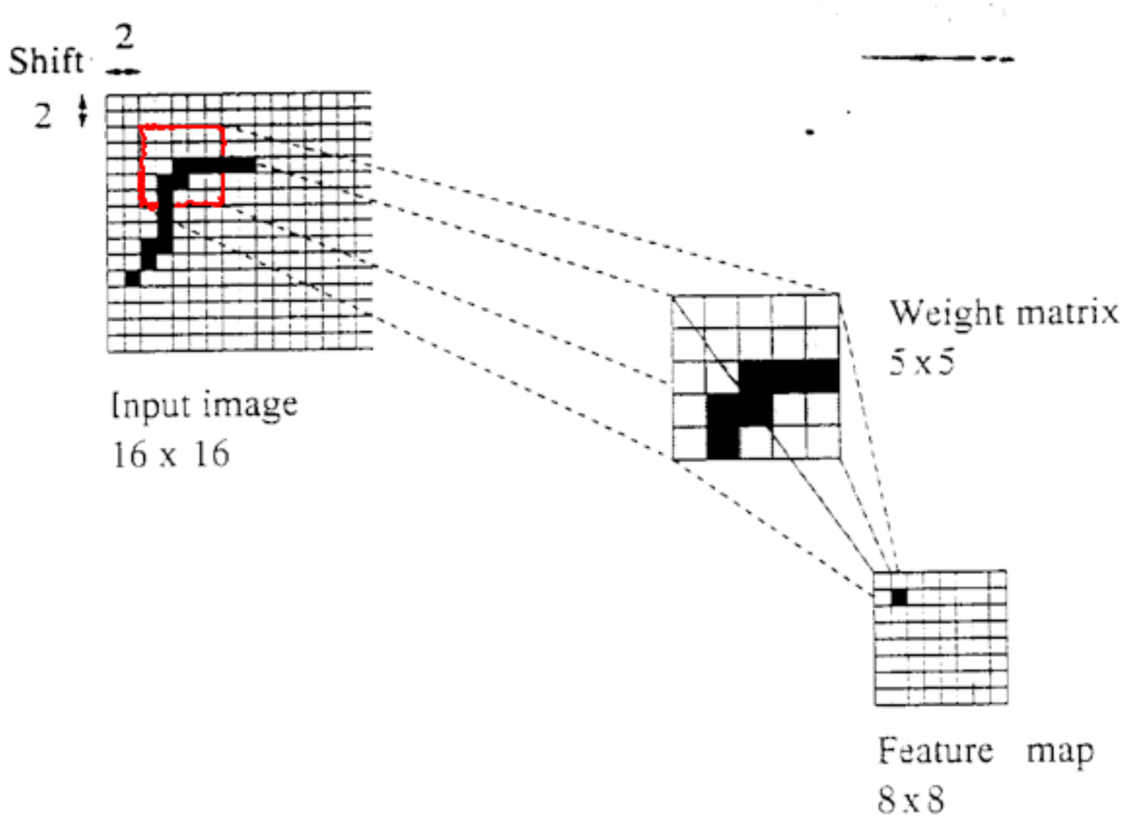


Figure 5.3: A feature map.

Sub-sampling Map

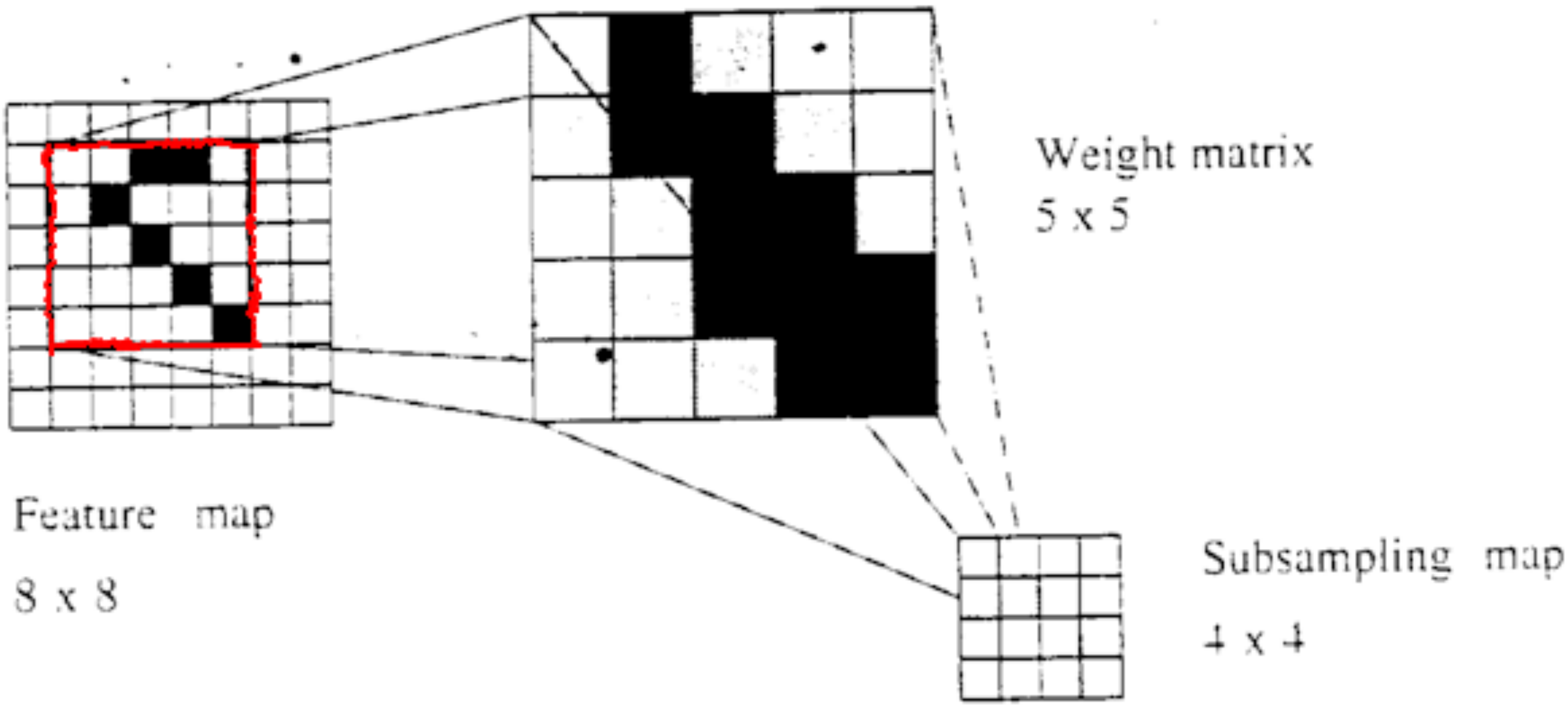


Figure 5.4: A sub-sampling map.

Architecture

- **Input layer**: 16x16 = 256 neurons with input values in range [-1, 1].
- **Hidden layer H1**: consists of 12 feature maps H1.1, … , H1.12.
- Feature map:
  - 8x8 neurons.
  - Each neuron in the feature map has the *same* incoming weights, but is connected to a square at a unique position in the input image. This square is called a *template*.

Architecture

- Hidden layer H2: consists of 12 sub-sampling maps H2.1, … , H2.12.

- Sub-sampling map:
  - Consists of 4x4 neurons.
  - Each neuron of the sub-sampling map is connected to a 5x5 square of H1.j, for each j in 8 of the 12 feature maps.
  - All neurons of the sub-sampling map share the same 25 weights.

Architecture

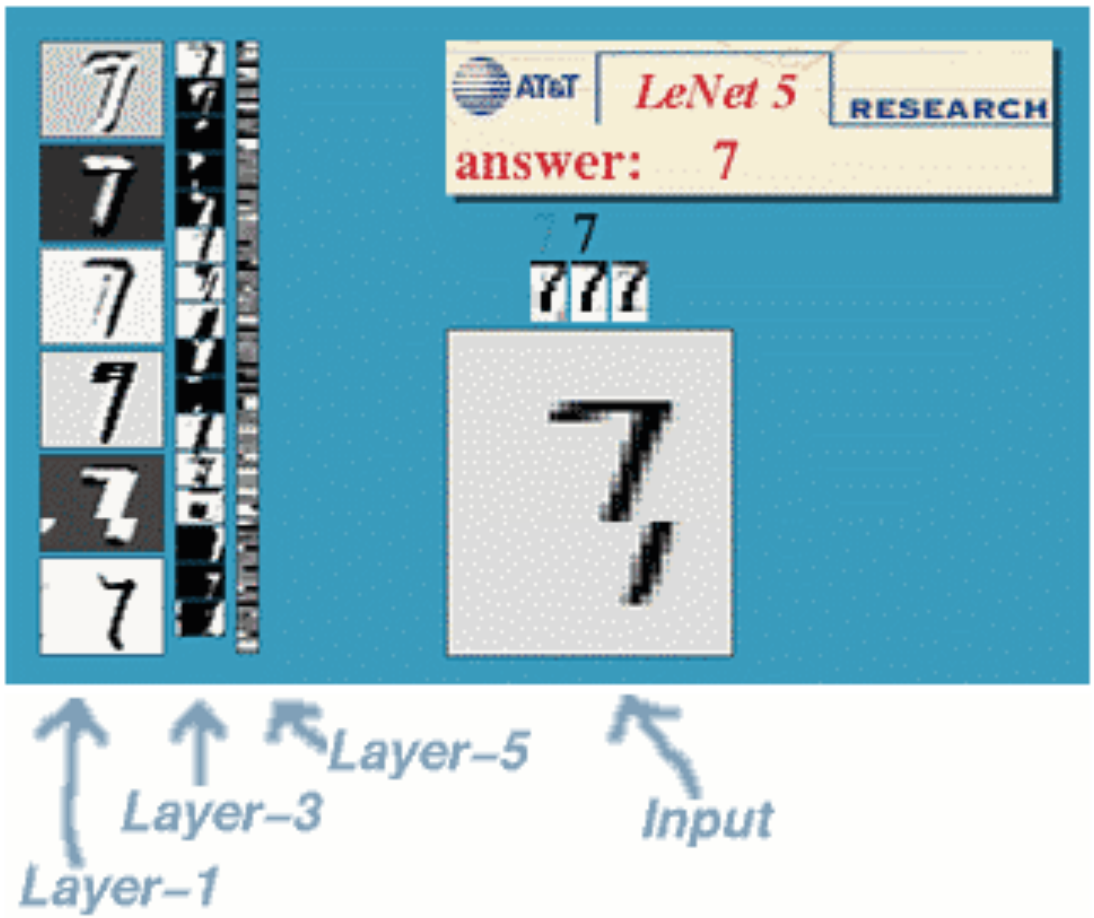
- Hidden layer H3:
  - Consists of 30 neurons.
  - H3 is completely connected to the sub-sampling layer (H2).
- Output layer: consists of 10 neurons, numbered 0, … , 9 and the neuron with the highest activation value is chosen. The digit recognized is equal to the cell number.

Atypical Data Recognized

Further Details and Results

- ~10,000 digits from the U.S. mail were used to train and test system
- ZIP codes on envelopes were initially located and segmented by a separate system (difficult task in itself): preprocessing of input data was crucial to success
- ***weight sharing*** used to constrain number of free parameters
- 1256 units + 30060 weights + 1000 biases, but only 9760 free parameters
- used an accelerated version of backprop (pseudo-Newton rule)
- trained on 7300 digits, tested on 2000
- error rate of ~1% on training set, ~5% on test set
- if marginal cases were rejected (two or more outputs approximately the same), error reduced to ~1% with 12% rejected
- used "optimal brain damage" technique to prune unnecessary weights
- after removing weights and retraining, only ~1/4 as many free parameters as before, but better performance: 99% classification accuracy with 9% rejection rate
- problem-specific knowledge was designed into the network architecture

http://yann.lecun.com/exdb/lenet/



ALVINN (Pomerleau, 1996)

- network controlled steering of a car on a winding road
- network inputs: 30 x 32 pixel image from a video camera, 8 x 32 gray scale image from a range finder
- 29 hidden units
- 45 output units arranged in a line corresponding to steering angle
- achieved speeds of up to 70 mph for 90 minutes on highways outside of Pittsburgh

Tabulation of Some Large Backpropagation Examples

application	#weights	#samples	error	ref.
text -> speech	25000	5000	0.20	Sejnowski
sonar target rec	1105	192	0.15	Gorman
car control	>36000	1200	car drives on winding road	Pomerleau
back-gammon	>11000	3000	computer champion	Tesauro
sex rec from faces	>36000	90	0.09	Golomb
char rec	9900	5000	0.055	Sato
remote sensing	1800	50	0.05-0.10	Kamata
signature verif.	480	280	0.05	Sabourin

T.J. Sejnowski and C.R. Rosenberg, *NETtalk: a parallel network that learns to read aloud*, The John Hopkins University Electrical Eng. and Comp. Science, 1986.

P. Gorman and T.J. Sejnowski, *Learned Classification of Sonar Targets Using Massively Parallel Network*, IEEE Transactions on ASSP, vol. 36, no. 7, July 1988.

D. Pomerleau, ALVINN: *Ann Autonomous Land Vehicle in a Neural Network*, in: David S. Touretzky, Advances in Neural Information Processing Systems I, 1989

G. Tesauro, *Neurogammon wins computer olympiad*, Neural Computation, vol. 1, pp 312-323, 1990

B.A. Golomb, D.T. Lawrence, T.J. Sejnowski, Sexnet: *A neural network identifies sex from human faces*, Adv. in Neural Inf. Proc. Sys. I, 1989

A. Sato, K. Yamada, J. Tsukumo, and T. Temma, *Neural network models for incremental learning*, ICNN, Helsinki, 1991.

S.-I. Kamata, R.O. Eason, A. Perez, and E. Kawaguchi, *A Neural Network Classifier for LANDSAT Image Data*, Proc. 11th ICPR, The Hague, Vol 2, 573-576, 1992

R. Sabourin and J-P. Drouhard, *Off-Line Signature Verification Using Directional PDF and Neural Networks*, Proc. 11th ICPR, The Hague, Vol 2, 321-325, 1992

Other Applications

Multi-channel Piezoelectric Quartz Crystal Sensor for Mixed Organic Vapours

Tamkang Journal of Science and Engineering, Vol. 5, No. 4, pp. 209-217 (2002)

Center of Mass Estimation for Use in a Clinical Environment

*Proceedings of the **25" Annual** International Conference of the IEEE EMBS Cancun, **Mexico September 17-2 1,2003***

Predicting Product Quality With Backpropagation: A Thermoplastic Injection Molding Case Study

International Journal of Advanced Manufacturing Technology

The application of a non-linear back-propagation neural network to study the mass balance of Grosse Aletschgletscher, Switzerland

Journal of Glaciology, Vol. 51, No. 173, 2005