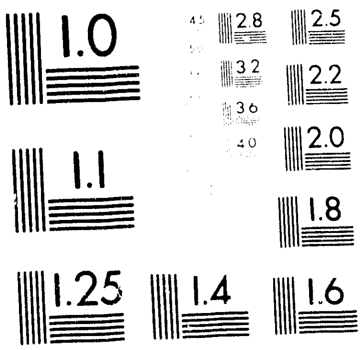

1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

Conf-931101--2

HIGH QUALITY GARBAGE: A NEURAL NETWORK PLASTIC SORTER IN HARDWARE AND SOFTWARE

Suzanne L. Stanton *, M. Kathleen Alam, Gregory A. Hebner

Sandia National Laboratories, Albuquerque, NM 87185
Operated for U.S. Department of Energy under contract no. DE-AC04-76DP 00789

Abstract

In order to produce pure polymer streams from post-consumer waste plastics, a quick, accurate and relatively inexpensive method of sorting needs to be implemented. This technology has been demonstrated by using near-infrared spectroscopy reflectance data and neural network classification techniques. Backpropagation neural network routines have been developed to run real-time sortings in the lab, using a laboratory-grade spectrometer. In addition, a new reflectance spectrometer has been developed which is fast enough for commercial use.

Initial training and test sets taken with the laboratory instrument show that a network is capable of learning 100% when classifying 5 groups of plastic (HDPE and LDPE combined), and up to 100% when classifying 6 groups. Initial data sets from the new instrument have classified plastics into all seven groups with varying degrees of success.

One of the initial networks has been implemented in hardware, for high speed computations, and thus rapid classification. Two neural accelerator systems have been evaluated, one based on the Intel 80170NX chip, and another on the AT&T ANNA chip.

Background

Plastics must be sorted by polymer type prior to reprocessing to ensure that the recycled plastic resin has virtually the same properties as the original resin. Presently many manufacturers are voluntarily placing codes on their containers to indicate that they *may* be recycled. Each code refers to a different type of polymer: 1- Polyethylene terephthalate (PET), 2- High Density Polyethylene (HDPE), 3- Polyvinyl Chloride (PVC), 4- Low Density Polyethylene (LDPE), 5- Polypropylene (PP), 6- Polystyrene (PS), and 7- Other (combinations, layers, and anything not a 1-6).[1,2]

Present automated plastic separating techniques commonly used are hydrocyclones and electromagnetic scanning. Hydrocyclones do not always produce a pure polymer stream, and electromagnetic scanning technologies separate only specific types of polymers, such as PET or PVC. In the U.S., the major method for sorting postconsumer plastics is hand sorting. This procedure is time consuming and subject to failure if no codes are present on the plastics.

In this new system, waste plastics are processed by focusing light onto the surface of each plastic container which generates a specific reflectance pattern. These patterns are then input into a neural network program which classifies the plastic.

Near Infrared Spectroscopy Data Collection Methods

In the laboratory, a Fourier transform near infrared spectrometer, model PCM 4000, from Laser Precision (now KVB Analect) with 4.0 cm^{-1} resolution was used to obtain the data. This instrument has CaF_2 beamsplitters and a thermoelectrically cooled PbSe detector. An Axiom diffuse/specular reflectance attachment set at 15^o was found to be the best for collection of the reflectance data from the plastic samples. For each training spectrum, an average of five scans in the range of 1000 - 2500 nm, ratioed to the background of a standard ceramic disk (Coors Ceramic Company, Golden, CO), and plotted as Log(1/Reflectance), was used. This procedure took about ten seconds per training spectrum. Typical responses for each plastic type can be seen in Figure 1.

Preprocessing

Preprocessing and the selection of input data for the network are the key elements to successful classification. Various methods were tested, including autoscaling (Equation 1), maximum scaling (Equation 2) and area scaling (Equation 3). Autoscaling (subtracting out the mean, and dividing by the variance for each waveform) gave the best

All papers must include the following statement:
This work performed at Sandia National Laboratories is supported by the U.S. Department of Energy under contract DE-AC04-76DP00789.

results in terms of convergence of the networks and percent correct for testing. In each case, the inputs were range scaled according to the range of each input element in the training set, y_j being the input element (1 - j) and x'_j being the corresponding preprocessed data point. (Equation 4.)

(1) $x' = (x - avg(x))/ std(x)$

(2) $x' = (x - avg(x))/ max(x)$

(3) $x' = (x - avg(x))/ sum(x)$

(4) $y_j = (x'_j - min(y_j))/ (max(y_j) - min(y_j))$

Software Results

Standard backpropagation neural networks with sigmoidal transfer functions (Equation 5) were chosen for their relative ease of use and portability to hardware.[3] The output of the neuron is z, the summed input is y, and the gain of the sigmoid is ϕ .

(5) $z = 1 / (1 + \exp(-y\phi))$

With the hardware limitations in mind, the number of network inputs and nodes in each of the hidden layers was minimized. A backpropagation network was trained with two hidden layers, learning coefficients of 0.9 for hidden layer one, and 0.6 for hidden layer two, and a convergence criteria of 0.005 rms output error. Networks with only one hidden layer were tested, but did not generalize as well, due to the dimensionality of the input space, and the way in which the different groups were clustered. [4]

Table 1 shows the results from a few of the best performing software solutions. The networks which had only 5 outputs classified the HDPE(2) and LDPE(4) together. (These polymers differ only in the amount of branching, and are thus difficult to distinguish with only near infrared data.) All of these software solutions were run on a PC in an interactive system with the spectrometer.

Table 1. Software Emulation Network Performances

Networks In - Out	Training	Testing
11-6	97.57%	98.41%
11-5	99.6%	100%
17-6	100%	98.18%
17-5	100%	100%
33-6	98.99%	100%
33-5	99.6%	100%

Hardware Results

The results from the 11 input - 6 output network were considered "worst case" of the main converged networks, due to the small number of inputs, and maximum number of outputs, and highest feedforward training error. This network was implemented on the Intel 80170NX development system, and the AT&T ANNA chip system, in order to compare the implementation and accuracy issues. The specifications for each chip can be seen in Table 2.

Table 2 Neural Network Chip Specification Comparison

	AT&T ANNA	INTEL 80170NX
Max Speed (cps)	5 billion	2 billion
Max # wts (on chip)	4096	10,240
Wts (bits)	6	6
I/O Type	Digital	Analog
11-6 network Results		
Training	85.22%	94.33%
Testing	93.65%	87.30%

Intel 80170NX Chip

The DynaMind (NeuroDynamX, Boulder, CO) software interface package was used to run the 11-6 network on the Intel 80170NX development system (iNNTS) (Intel, Santa Clara, CA). The number of inputs were limited to 128 with a maximum of 64 neurons in the network. Using two hidden layers was not difficult, only the total number of neurons was limited. The weights were limited to a range of +/- 2.5 for hardware reasons. The sigmoid function on the chip had a slope of five, versus the slope of one used in the original software solution. These two factors made it necessary to retrain the network using the available software package to match the network to the chip. A learning rate of 0.05 was used to converge to 0.03 rms error in the software simulation. This gave a feedforward accuracy of 95.34% on the training data and 84.13% on the test set. (See Table 3.)

Table 3 Feedforward results for 80170NX

	Training Set	Test Set
Simulation	95.34	84.13
Chip in Loop	93.93	87.30
Chip in Loop (w/CIL training)	94.94	84.13

An important note in the success of training this network for the hardware was the scaling of the data. The outputs had to be scaled to +/- 0.8 instead of +/- 1, which is the entire range of the sigmoid function. This put the desired neuron outputs more in the linear range and thus made training easier.

The hardware produced an initial feedforward training accuracy of 93.93%, with 87.3% for the test set. Chip-in-loop training then proceeded at a learning rate of 0.001 for approximately 2000 epochs, converging to 0.05 rms error. This gave 94.94% accuracy on the training set and 84.13% on the test set. The breakdown of the errors by plastic type shows that the main error comes from the differentiation between the 2s and the 4s. (See Table 4.)

Table 4 Testing Accuracies by Plastic Type

Plastic	Training set			Test set		
type	Sim	CIL	CIL-T	Sim	CIL	CIL-T
1	100	100	100 %	100	100	100
2	81	69.6	86	100	80	100
3	100	97	95	100	100	100
4	93.1	96.6	93.1	0	60	0
5	100	100	100	100	92.3	100
6	97.4	100	96	100	100	100

AT&T ANNA Chip

The AT&T ANNA chip (AT&T Bell Laboratories, Holmdel, NJ) was optimally designed for locally connected, weight-sharing or time-delay networks, but can also be used for fully connected and recurrent nets. [5] The prototype application for this chip was an optical character recognition system, with a large, sparsely connected network. [6]

The 11-6 network was tested on this hardware with appropriate modifications made by retraining the network within the original software [7]routines. Unfortunately, this did not include using an accurate model of the neuron squashing function, which is a digitized output. The weights were limited to +/- 2.5 for the chip constraints. The weights had to be modified to accommodate the bias saturation levels. The software simulation gave a feedforward accuracy of 94.94% and 92.06% for training and testing sets, respectively.

The modified networks were tested using the Lisp interface to the ANNA chip on a VME board. Feedforward results can be seen in Table 5. In this case, the testing data gave better results (93.65%) than the training data (85.5%), partially due to the fact that fewer of the samples were ambiguous . This ambiguity was caused by the relatively few quantization levels on the inputs, meaning that some samples with different classifications had identical inputs when they were digitized upon application on this chip. This did not appear in the software training, since the digitized transfer functions were not used.

Table 5 Testing Accuracies by Plastic Type ANNA chip

Plastic	Training set			Test set		
type	CIL		# Ambig.	CIL		#Ambig.
1	97.4 %	(76/78)	2	100	(11/11)	0
2	62.0	(49/79)	6	100	(10/10)	0
3	96.0	(96/100)	4	100	(10/10)	0
4	70.0	(61/87)	6	70	(7/10)	2
5	100	(74/74)	0	100	(13/13)	0
6	85.5	(65/76)	7	92	(8/9)	1
Total	85.5%		5.04%	93.65%		4.76%

Custom Instrument Results

The new instrument uses broadband light with a circular variable filter and an InAs detector. This configuration obtains 13 scans per second, versus the 0.5 scans per second with the laboratory instrument. The trade-off came in the resolution of this instrument: approximately 20 nm, compared to the 0.4 - 3.5 nm resolution on the earlier instrument. The data as gathered from the new instrument is in the range of 1300 - 2300 nm. (See Figure 2.) With a small set of initial data, networks were trained to classify up to 100% for seven categories of plastic. Larger sets of data have now been taken, and training is underway for the new networks.

Conclusions

The initial neural network software simulation accuracies are acceptable for implementation into a real-time system, and it is believed that the speed of the software will be adequate as well. The limiting factor of the speed proved to lie in the instrument.

Both chips performed well, considering the less than ideal situations for the network which was implemented. If either hardware solution is chosen for implementation, more accurate methods for initial training will be used.

Acknowledgments

Special thanks to Susan Tucker, Senior Member of the Technical Staff, Sandia National Laboratories, who did the ANNA chip implementation. Also, thanks to Edi Säckinger at AT&T Bell Labs, Holmdel, NJ, for loaning us the ANNA chip and giving us the necessary support.

---

[1]Angela Kirkman, Charles H. Kline, "Recycling Plastics Today," *Chemtech*, October 1991, pp 606-614.

[2]K.R. Kreisher, K.F. Lindsay, P. Mapleston, J.K. Rogers, G.R. Smoluk, R.A. Toensmeier, R.V. Wilder, and A.S. Wood, "Waste Solutions," (additional insert), R.D. Leaversuch (Ed.), *Modern Plastics*, April 1990.

[3]D.E. Rumelhart and J.L. McClelland, Eds., *Parallel Distributed Processing*. Cambridge, MA: M.I.T. Press, 1986.

[4]Bernard Widrow, Michael A. Lehr, "30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation," *Proceedings of the IEEE*, Vol. 78 No. 9, September 1990.

[5]Bernhard E. Boser, Eduard Säckinger, "An Analog Chip for Time Delay and Feature Extraction Neural Networks," March 25, 1991,

[6]Eduard Säckinger, Bernhard Boser, Lawrence D. Jackel, "A Neurocomputer Board Based on the ANNA Neural Network Chip," Advances in Neural Information Processing Systems -4-, Edited by John E. Moody, et al. Morgan Kaufman Publisher, Inc., San Mateo, CA 1992

[7]Eduard Säckinger, Bernhard Boser, Jane Bromley, Yann LeCun, Larry D. Jackel, "Application of the ANNA Neural Network Chip to High-Speed Character Recognition," IEEE Transactions on Neural Networks, Vol 3 No. 3, May 1992.

Figure 1

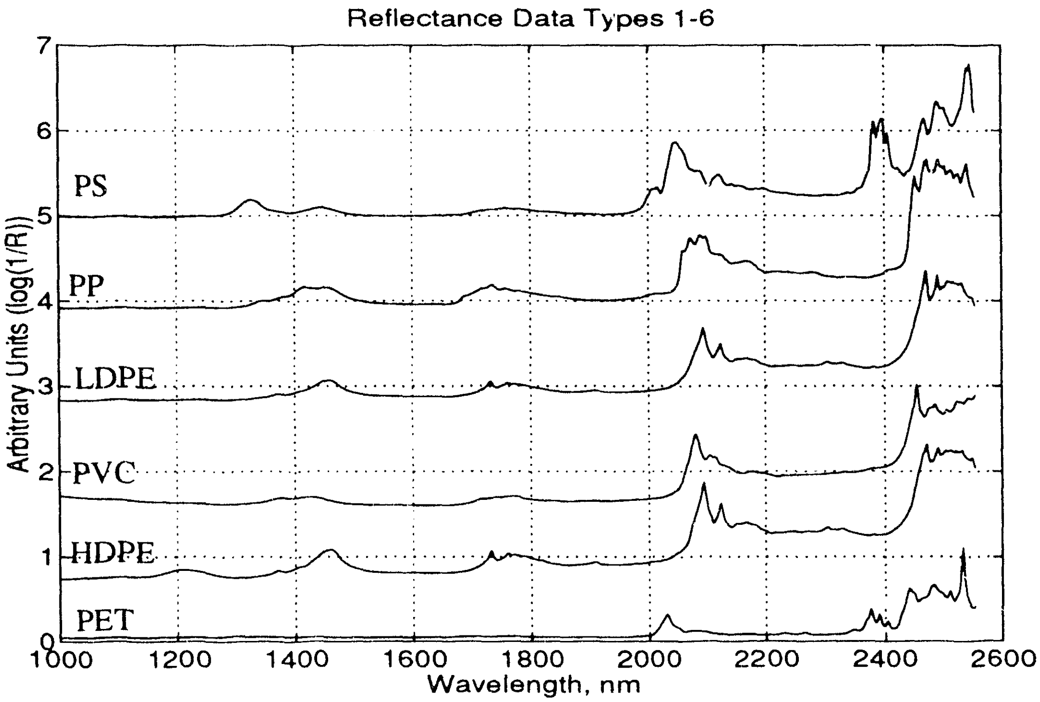
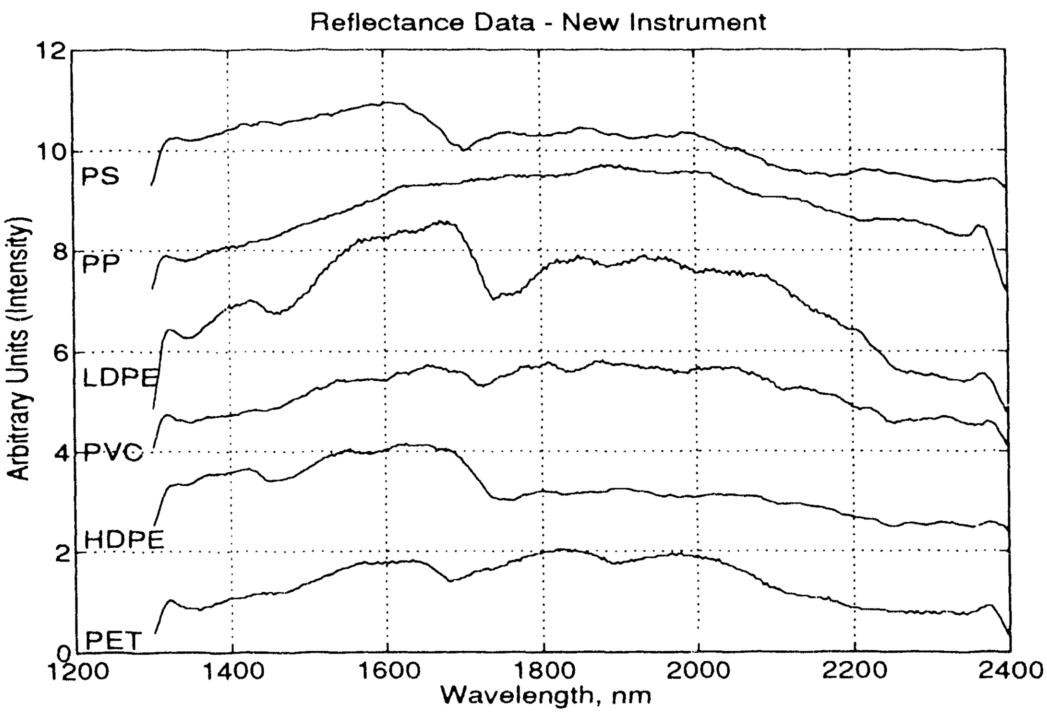


Figure 2



DATE
FILMED
12 / 6 / 93
END