The Use of Neural Networks in an Image Analysis System to Distinguish Between Laser Prints and Their Photocopies

J. Tchan, R. C. Thompson and A. Manning

Materials Research Group, School of Printing and Publishing, The London College of Printing, The London Institute, London, England

An image analyzer that employed neural networks was used to classify prints according to whether they were laser printed or photocopies of the laser prints. The prints analyzed were monochrome images of squares, the text character 'a' and a circle. Each image was reproduced in a range of tones. The image analysis system produced raw image data from the prints and a pre-processing program was used to extract features from this raw image data. Neural networks employed the features to find classification models for the three different sets of images. In the analysis a classification rate of 100% was achieved for the squares, 95% for the letter 'a' and 93% for the circle.

Journal of Imaging Science and Technology 44: 132–144 (2000)

Introduction

A research program aimed at developing an image analysis system that can identify the production source of a print has been initiated. This approach to print identification has hitherto not been investigated. It is envisaged that the system could find applications in the print security and anti-counterfeiting fields.

The production process of a print sample can be identified by visual inspection of edge defects and mottle characteristics of the print under a microscope[1] by an expert observer. However, to be of practical use, the expert observer must not only be able to distinguish between a variety of printing processes, but also be able to distinguish between prints from different makes of printing machine that employ the same printing process. The latter is not easily achievable by an expert observer.

It is proposed to produce a system that can replace the expert observer and be able to identify the production source of a print where expert observers cannot manage this task. Also an important aspect of this research is to see whether the system can be made adaptive so that it performs the identification task in a cognitive or artificially intelligent way. This means that the system will be able to make automatic assessments on prints of different images. This study demonstrates the feasibility of producing such a system by demonstrating its ability to distinguish between prints from two printing sources.

This investigation describes the development of an adaptive artificially intelligent system that can automatically differentiate between digital laser prints and their optical photocopies over a range of tones as well as a solid print. To achieve this it uses artificial intelligence (AI) techniques in the form of neural networks. The advantage of using neural networks is that their theoretical principles are simple to comprehend. This arises from the fact that a neural network can be viewed as a system that learns to find the correct solution to a problem through training.

The role of the neural network is to find empirical models from a combination of physical or geometrical features from a print to produce an adaptive system. Some of the features employed in this investigation have been previously studied using image analysis[2–10] however these variables have only been studied individually and have not been used together to produce an adaptive system. In this project a computer program that can process raw CCD camera data to create many different features pertaining to the print under investigation has been developed. These variables are linked to a neural network and can be used in any combination to produce different network models in order to find the correct solution to the print identification problem. In addition the program is modular so that variables can be changed at any time to allow other print identification problems to be investigated.

Artificial Neural Networks (ANNs)

ANNs are algorithms that can perform classification tasks by producing empirical models from data.[11–14] Here ANNs were used to classify prints according to whether they were produced by a laser printer or as photocopied reproductions. Before ANNs can be applied a process known as feature extraction is performed on the prints. Each feature is a measurement of a physical or geometrical property of the print. In this investigation, the features were obtained using an image analysis system that incorporated a computer program known as a pre-pro-

Original manuscript received April 2, 1999

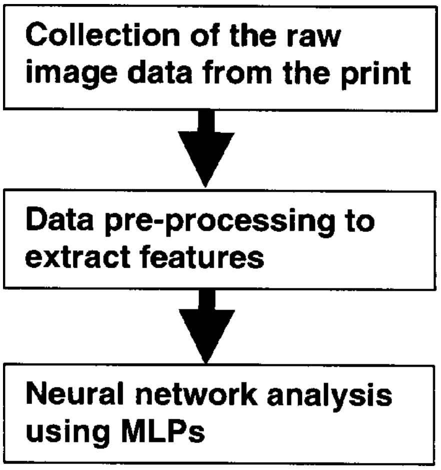


Figure 1. The main components of the experimental procedure.

cessor. Figure 1 illustrates the procedure adopted. Firstly, the image analysis system captured raw image data from the print in the form of a 735 × 582 pixel array with a 0-255 tonal range. Secondly, features were extracted from the raw image data by the pre-processing program. Thirdly, the features were used by the neural networks to produce empirical classification models. The **Multilayered Perceptron** section of this article deals briefly with a class of neural networks known as multilayered perceptrons (MLPs) which have been used in this investigation. The **Multilayered Perceptron** and **Data Pre-processing** sections cover the importance of data pre-processing.

Multilayered Perceptron (MLP)

Neural networks solve classification problems by producing an empirical model between a set of input training vectors V_{Tk}, and a corresponding set of output target vectors O_{Tk}, (k = 1,2,3,4....t where t is the number of training vectors). The empirical model is a mapping between the input vectors and the output target vectors. Therefore a neural network can only produce a solution provided that there is a relationship of the form $y = f(x)$ between the input and output target vectors .

A neural network with an MLP architecture can find both linear and non-linear mappings between the input and the output vectors. The structure of an MLP is shown in Fig. 2. The circular nodes are computational units known as neurons and the square nodes are inputs. In Fig. 2, an input vector is applied to the MLP. The input vector contains a set of measurements known as features (X_1, X_2, X_3, X_4...., X_f) that are used to produce a corresponding output state or vector O_{Nk}. The output state of the network O_{Nk} is dependent on the values of the inputs features and the weights, W. For a given input vector, when any of the weights are adjusted, the output state of the network can be altered.

A neuron calculates its output response by multiplying each of its input values by a different weight and adding the resultant values before normalization using transfer functions. The output neuron uses the responses from the neurons in the hidden layer as its input values, as in Eq. 1 where F_{out} is a normalizing transfer function. Regarding the neurons in the hidden layer the features are used as input values (Eq. 2) and the transfer function is denoted as F_{hid}. The role of the transfer function is to normalize the neuron output to either [0,1] or [–1,1]. MLPs commonly use logsig and tansig functions, Fig. 3, that have [0,1] and [–1,1] as limits respectively.

$$O_{Nk} = F_{out}\left(\sum_{i=1 \text{ to } n} W_{oi} V_{Tk}(Y_i)\right) \tag{1}$$

$$V_{Tk}(Y_i) = F_{hid}\left(\sum_{j=1 \text{ to } f} W_{h,i,j} V_{Tk}(X_f)\right) \tag{2}$$

The objective is to adjust the weights so as to reduce the difference between the corresponding target vectors O_{Tk} and the actual network outputs O_{Nk}. The difference between the two sets of vectors is commonly expressed as a sum squared error E_n (Eq. 3) that is reduced in steps using an iterative learning algorithm. This process is known as network training or network learning. A type of iterative algorithm that can achieve this task is a back propagation algorithm. It is known as back propagation because in the weight adjustment process, the weights are reduced from the output towards the input of the network. Two commonly used forms are standard back propagation and the Levenberg–Marquart (LM) algorithms. They perform the same task of adjusting the weight values to reduce E_n, but have different computational requirements.

$$E_n = \frac{1}{2}\sum_{k=t} (O_{Tk} - O_{Nk})^2 \tag{3}$$

If an LM and a standard back propagation algorithm are implemented on identical networks the network using the LM algorithm will be much faster[16] to train. However, the LM network will require more computer memory than the standard back propagation network. Because the amount of memory and data processing speed are also determined by the size of the network, it is only advantageous to use the LM method for small networks.

The error reduction process for the print classification problem is illustrated in the flow chart Fig. 4. The objective was to differentiate between laser and photocopied prints of simple monochrome images that varied in tone, using a neural network model. The image analysis system produced experimental data in the form of a series of feature vectors, V_{Tk}, for each print. The output target vectors, O_{Tk}, were assigned the value 1 for the laser print and 0 for the photocopier print. The MLP network and a back propagation algorithm were applied to the set of input vectors until E_n (Eq. 1) was reduced to a pre-determined level.

One of the properties of an MLP network is that it can produce many different solutions. This property is illustrated by the sum squared error, E_n, landscape in Fig. 5 for two weights W_1 and W_2. The position of the ball represents an E_n value and is related to the number of back propagation loops the program has performed. The objective of back propagation is to move the value of E_n towards a minimum. However, there are many mimima that E_n can move towards. The position where the network finishes and therefore the form of the solution can be changed by:

1. Changing the starting position of E_n. This is adjusted by changing the initial weight values of the network.
2. Altering a parameter known as the step size or the learning rate of the network. This controls the amount by which the weights are adjusted in each back propagation loop.
3. Changing the sum-squared error goal for the network.

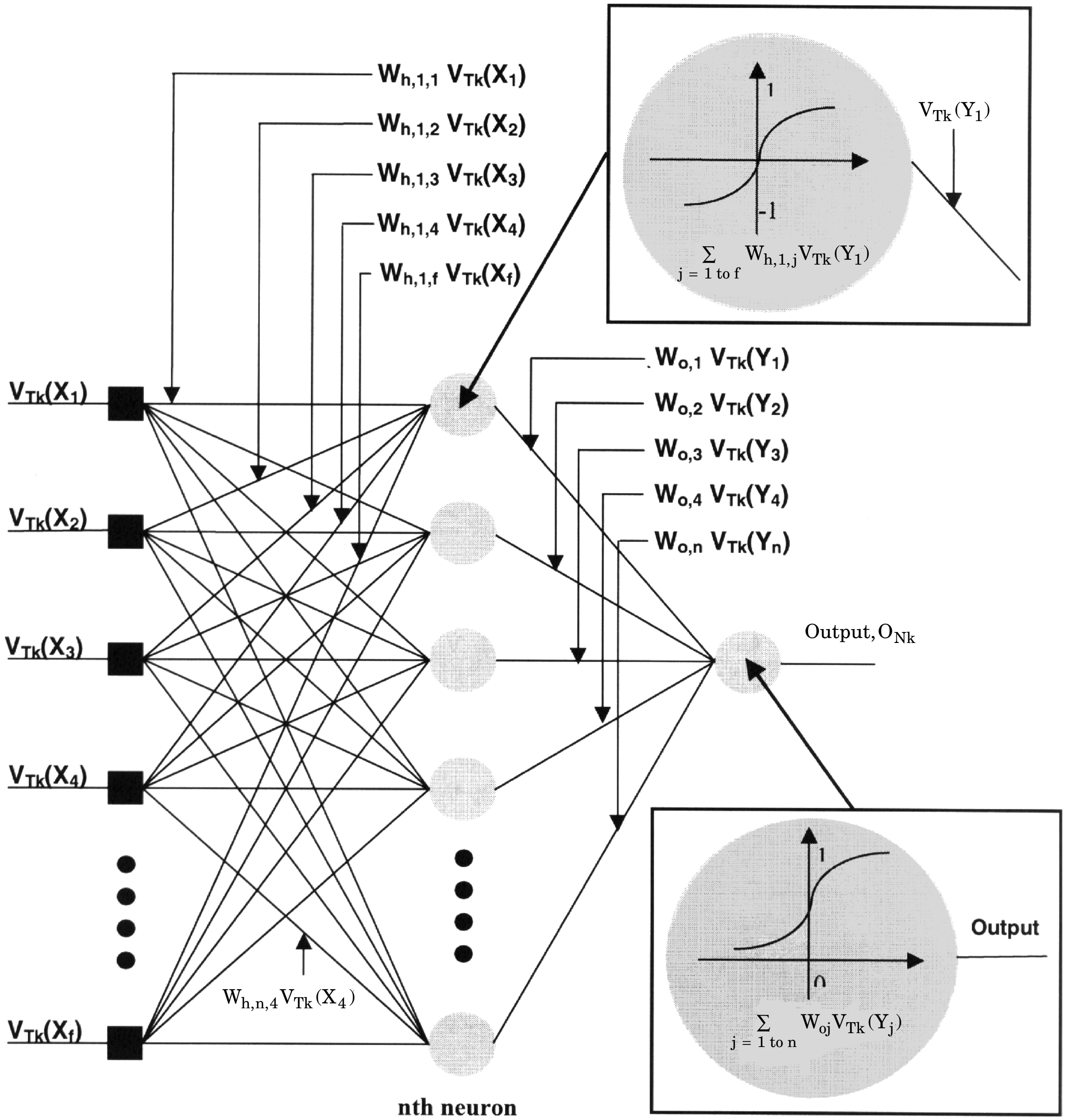


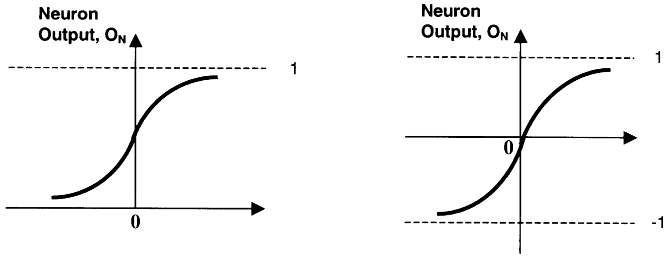
Figure 2. The MLP neural network architecture used to implement back propagation.

Another way of obtaining different solutions is to change the energy landscape. This can be achieved in two ways. Firstly, by changing the input features or their number which increases the dimensionality of the problem. Secondly, by using a different number of neurons in the hidden layer. Increasing the number of neurons in the hidden layer enables more complex classification problems to be modeled. This is illustrated in Fig. 6 for a 2-D problem.

The MLP neural network method described above is known as supervised learning. This is because a training set of input vectors, $V_{Tk,}$ with known output vectors, O_{Tk}, is used to train the network to produce the correct responses O_{Nk}. Once a network is trained, it is used to predict the output from input vectors that it has not previously encountered. The assessment of the accuracy of a model produced by a network is known as validation and testing. A procedure used for validation and testing is described below.

A validation set of data is employed to check the classification accuracy of the models produced by the training set. This can be achieved by training a network to a specific E_n level using the training set and then storing the network weights. The validation input vectors are then applied to the network model and the network output for this set of data is compared with the actual output. If the accuracy needs to be improved the network can be trained further to a lower E_n value and the network is reassessed using the same validation set. If the results are acceptable, training is stopped and the network weights are stored. The network model is then tested again using other sets of data. These further sets of data are known as the test sets.

The explanation of neural networks given so far emphasises their heuristical and empirical nature. They can produce many different solutions for a particular problem. Different solutions can be found by adjusting the learning rate, the sum squared error goal, the number of hidden neurons in the network or the number



$$O_{Nk} = \text{Logsig}\left(\sum_{i=1 \text{ to } n} W_{oi} V_{Tk}(Y_i)\right) \qquad O_{Nk} = \text{Tansig}\left(\sum_{i=1 \text{ to } n} W_{oi} V_{Tk}(Y)\right)$$

Figure 3. Typical transfer functions used in the MLP. The graphs refer to the output neuron.

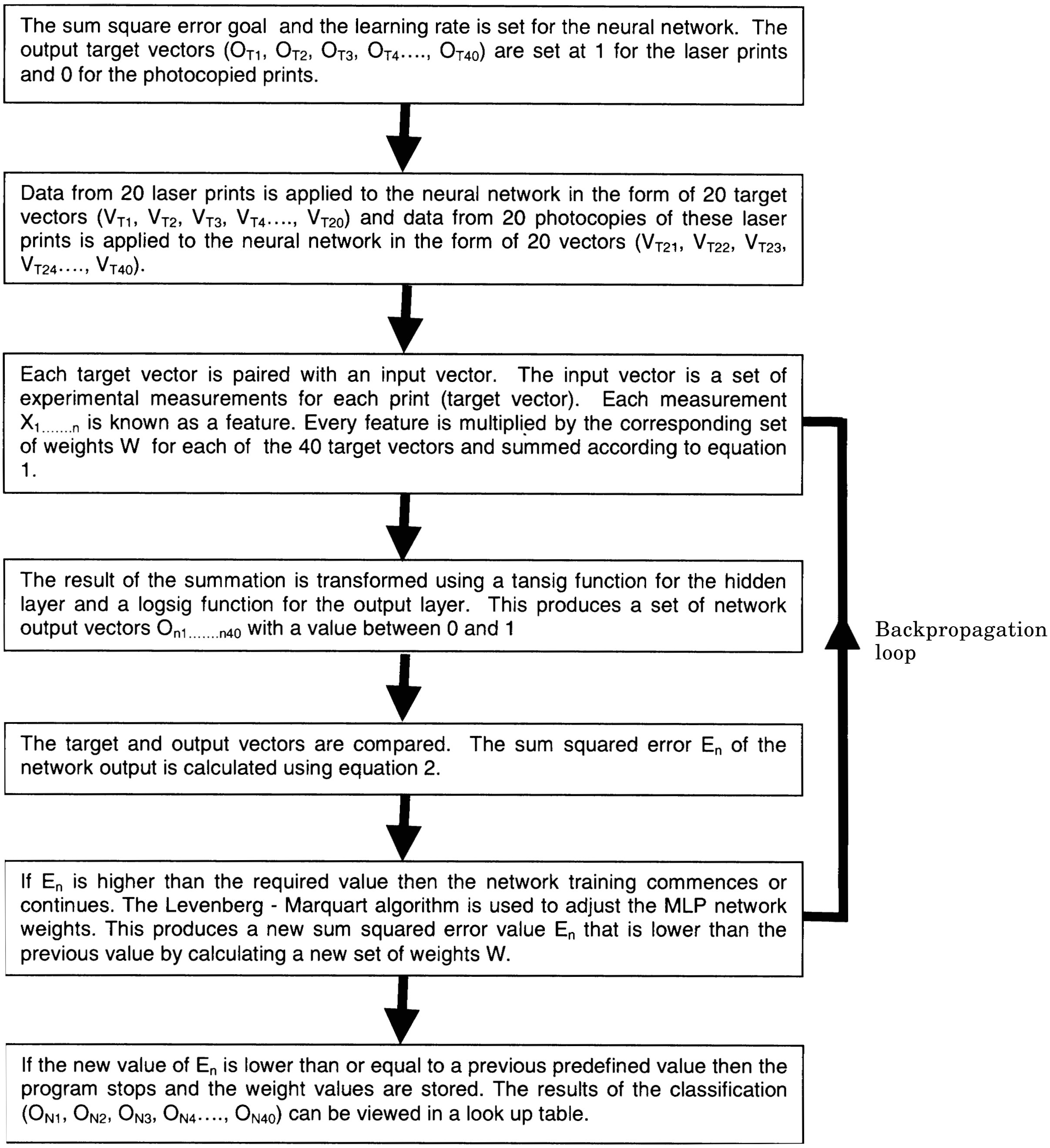


Figure 4. The stages in the application of the MLP neural network algorithm.

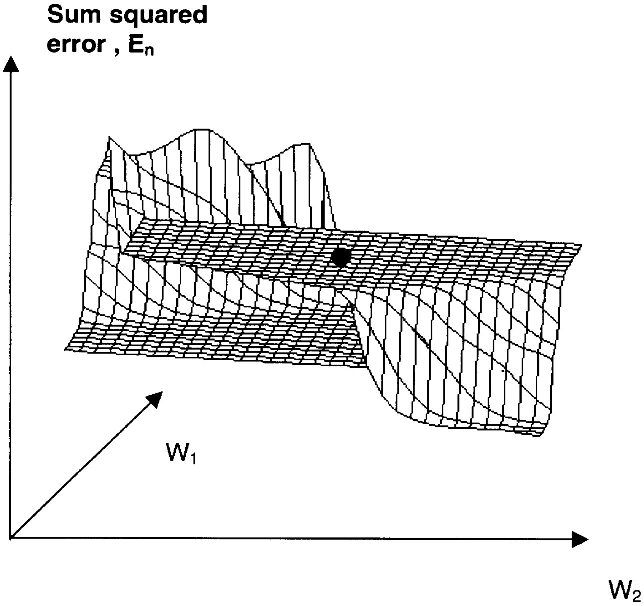


Figure 5. A representation of 2-D weight landscape produced by an MLP.

and nature of the inputs used in the network. The input parameters have a very important affect on the precision of the model that a network can produce and the input data quality can affect the likelihood of a network finding a satisfactory result to a problem. As mentioned earlier the input to a network can also be referred to as a feature. If a network result is unsatisfactory, it can be improved by adding other features (inputs) to the network. This increases in the dimensionality of the problem can sometimes improve the performance of the network.

However, using more features will increase the quantity of data required by the network to produce an accurate model. This is because a network can be considered as a method of fitting a curve between classes of data points. Increasing the dimensionality of the network will increase the number of data points needed for accurate interpolation. Therefore, it is important to use the highest quality data possible when applying neural networks. This minimises the number of dimensions required and maximises the accuracy of interpolation between data points. The process of selecting data is known as feature extraction and is a major part of neural network analysis. The **Data Pre-processing** sec-

Figure 6. Two dimensional example illustrating how the number of neurons in the hidden layer of a neural network equals the number of linear decision boundaries that can be produced.

Figure 7. A selection of images used in the investigation.

tion develops the concept of feature extraction in terms of the prints used in this investigation. The next section describes the prints used in the classification trials.

Print Samples

The following images were chosen: 10 mm squares, the letter ‘a’ in a 36 point bold Arial font and a 10 mm diameter circle. Illustrations of these images are shown in Fig. 7.

The first series comprised squares of 15 differing tones and 8 solids. The squares were drawn using Word 7 and produced using a Hewlett Packard 4M plus laser printer. These images were subsequently photocopied using a Sharp SF-2022, an Oce Bookcopier and a Canon 6030 photocopier.

The second series comprised 20 text characters of the letter ‘a’ of 19 differing tones and a solid. The letters were created using Adobe Photoshop™ and printed from a Hewlett Packard 4M plus laser printer. These images were subsequently photocopied using the following photocopiers: a Sharp SF-2022, a Xerox 5343 and an Infotec 5402.

The third series comprised circles of 14 differing tones plus a solid. The circles were drawn using Word 7 and printed from a Hewlett Packard 4M plus and a Kyocera FS-1700 laser printer. These images were subsequently reproduced using a Sharp SF-2022 photocopier.

Three sets of classification trials were carried out. In the first classification trial, the system was used to differentiate between images of the square that were produced by the Hewlett Packard laser printer and three different photocopiers. In the second classification trial, the squares were replaced by images of the text character ‘a’ and a similar procedure followed. The third classification trial used images of circles. This trial demonstrated that the system can classify prints from more than one laser printer source.

Data Pre-processing

In the **Multilayered Perceptron** section it was stated that the quality of the input data determines the accuracy of results obtained from a neural network. This can be illustrated by considering the print classification problem in this investigation.

The image analysis system was used to produce magnified images of the print samples, an example of which is given in Fig. 8(a). The image data takes the form of an 8 bit 735 × 582 array. It is possible to use each pixel in the CCD raw image data array as a feature for an input vector. This would mean that a neural network with 735 × 582 inputs would be required. This number of inputs is excessively high because a large volume of data would be required for accurate interpolation in 735 × 582 dimensions. Pre-processing reduces the dimensionality of the problem by extracting features. Features can be either physical or geometric characteristics of the image extracted from the raw image data that are applied to the neural network to produce the classification models.

For example, information about the tonal density and noise in an image can be extracted from the raw image data components. It is possible that these two features when applied to neural networks can classify the images more accurately than by using the raw image components, because the volume of data required for accurate interpolation in 2 dimensions is considerably less than that required for 735 × 582 dimensions. If the network using the tonal density and noise features cannot classify the laser prints and photocopies accurately, other features can be extracted from the image and applied to the neural network. The next section describes the pre-processing program that was used to extract features from the CCD images of the print samples.

Pre-processing Program

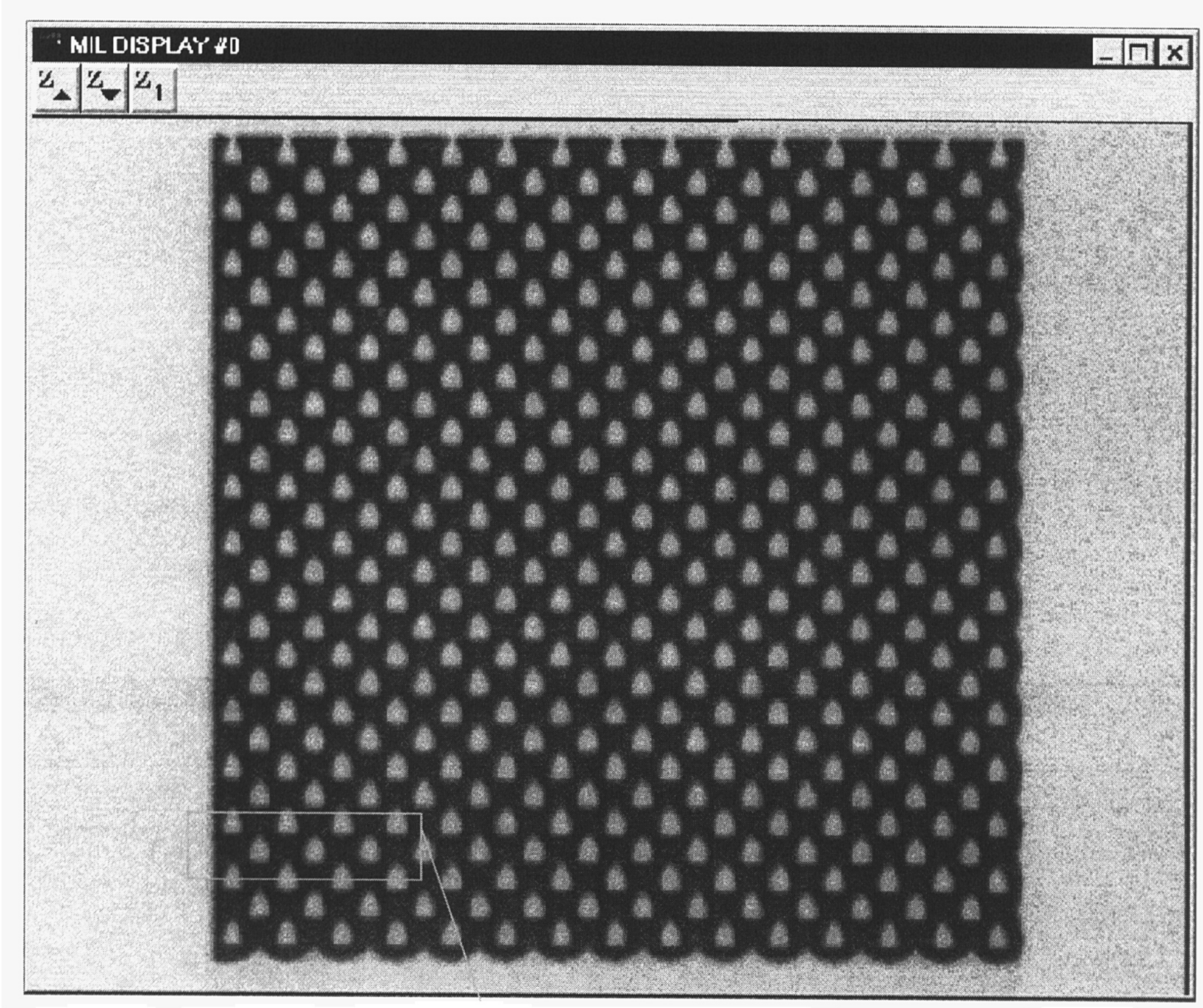
A pre-processing algorithm was developed that could produce eight features derived from the CCD image of a 10 mm halftone square, see Fig. 8(a). These features were used in different combinations to produce a range of classification models. The algorithm calculates values for these features by measuring in a variety ways, the noise, the edge and tonal characteristics of the print. The measurement was achieved by scanning the image matrix, line by line, and locating the image regions of the print using a thresholding technique. Image areas were detected by measuring the tonal gradient between neighbouring pixels, when the gradient exceeded a predefined positive value an image region was recorded. Conversely when the gradient exceeded a predefined negative value then a background region had been detected. Figures 8(b) and 9 illustrate the thresholding process for part of a scanned line.

Feature 1 is a noise measurement obtained by comparing the lengths of neighboring image regions $I_{m,y}$ and $I_{m+2,y}$ for the entire image array. After every comparison, T is increased by 1 and when $I_{m,y}$ and $I_{m+2,y}$ are equal, a counter N is increased by 1. Each time y is incremented at the end of a scanned line, m is reset to 1. Equation 4 is used to calculate a noise factor for feature 1.

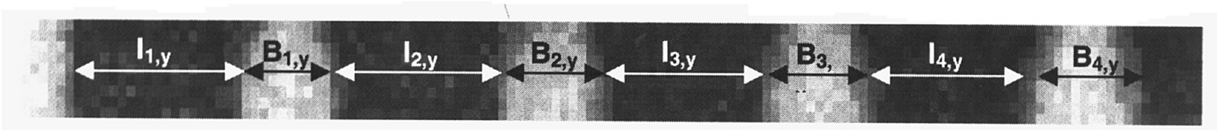
$$\text{Feature1(noise factor)} = 1 - \frac{N}{T} \tag{4}$$

Feature 2 is the modal halftone frequency of the print, where the halftone frequency is $I_{n,y} + B_{n,y}$ (n = 1,2,3,4.........). This is illustrated in Figs. 8(b) and 9. Feature 2 is calculated by determining the halftone frequency size distribution for the entire image and then locating the peak halftone value in the distribution. This is illustrated in Fig. 10.

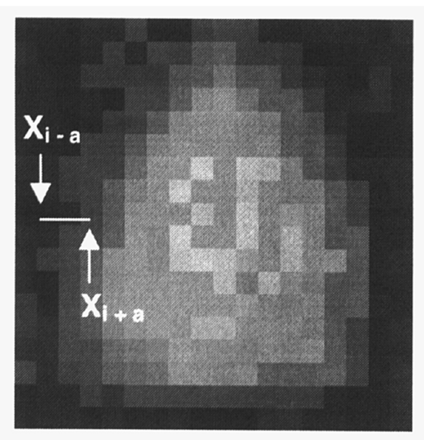
Feature 3 is a width calculation of the size distribution of the halftone in an image, (Fig. 10). For this feature, if there are more than five counts for a given length of cycle in a halftone image, a counter in the algorithm is incremented by one. In Fig. 10 for example, the number of halftone frequency values with more than 5 counts is 8. The noisier an image is, the higher the counter value will be.



a



b



c

Figure 8. Illustrations that demonstrate how features are extracted from the CCD image using the preprocessing algorithm. (a) A magnified image of a 10mm square used in the experimental work. (b) An expanded view of a part of Fig. 8a used to explain how noise measuremetns are made. (c) An expanded view of a part of Fig. 8a showing how edge sharpness calculations are made.

Figure 9. Part of a scanned line showing successive image and background areas.

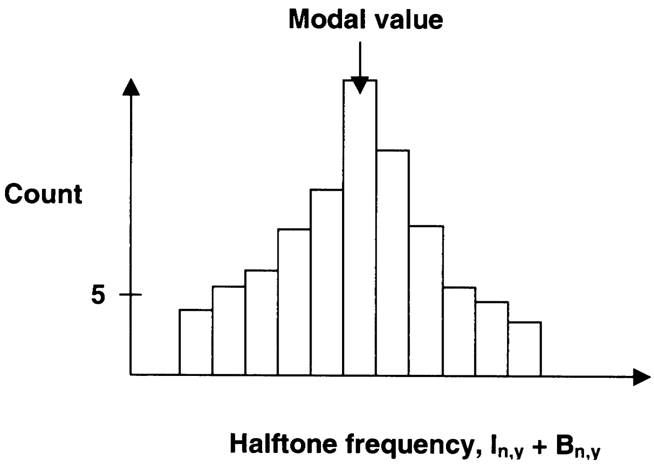


Figure 10. An illustration showing how the modal frequency factor, feature 2, is calculated.

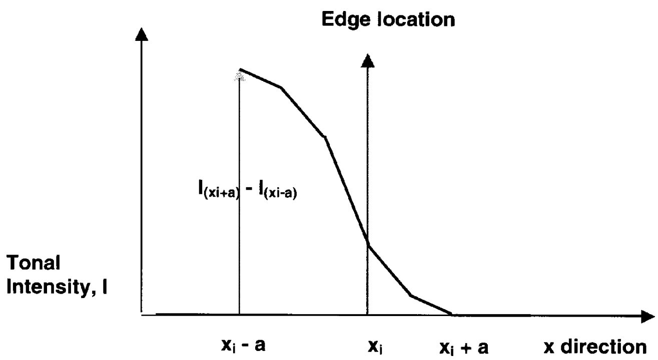


Figure 11. Diagram illustrating how the tonal gradient between the values of $x_i + a$ and $x_i - a$, at the location of the edges was calculated.

Feature 4 is obtained by calculating the average value of the descending gradients, using Eq. 5, in the x direction at the position where an image region has just been detected. The method of determining the gradient value using Eq. 5 is illustrated in Figs. 8(c) and 11.

$$\text{Feature 4 (gradient factor)} = \frac{\sum_i \frac{I_{(x_i+a)} - I_{(x_i-a)}}{(x_i+a)-(x_i-a)}}{N} \quad (5)$$

Feature 5 is the same calculation as feature 4 but performed on ascending gradients.

Feature 6 is a measure of the print contrast, found by integrating the tonal values of all the image regions illustrated by the shaded areas in Fig. 9.

Feature 7 is determined from recording the tonal value of every pixel throughout the image regions. The tonal distribution for all the image regions is then calculated. Feature 7 is the tonal value in the distribution with the largest pixel population. This is illustrated in Fig. 12.

Feature 8 is the pixel population value for the peak in the distribution shown in Fig. 12. This peak value corresponds to the tonal value calculated for feature 7.

Experimental Hardware and Software

The CCD camera which captured the image data was connected to a computer that operated the image analysis software.[15] The software comprised the Matrox MIL frame grabber application and an image data pre-processing program connected to a neural network. The preprocessing algorithm was written in Visual Basic version 5 and the neural networks implemented using Matlab.[16] The complete image analysis system is shown schematically in Fig. 13.

The arrangement of the experimental apparatus is illustrated in Fig. 14. The first stage of the image analysis system employed a JVC TK-S350 monochrome CCD camera fitted with a Computar M6Z zoom lens. The camera was mounted on a Polaroid MP4 Land camera stand. An annular TLE 22W/29 lamp surrounding the camera provided the illumination. A magnification that gave the image analysis system the same resolution as the human eye at a viewing distance of 300 mm from an image was selected. This meant that the print samples for the investigation did not exceed 10 × 13 mm in dimensions.

Classification Trials

A preliminary investigation was undertaken to establish whether the system was capable of producing accurate classification models. This involved the selection of different combinations of the eight features.

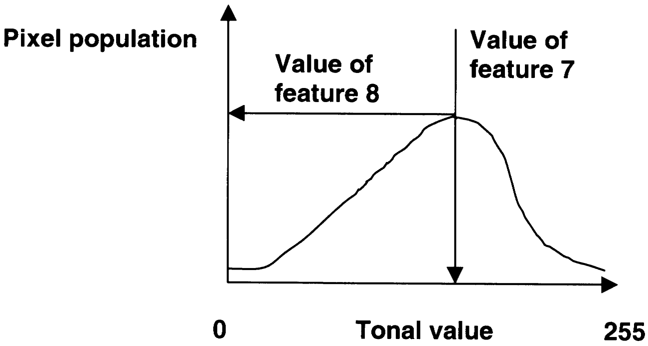


Figure 12. Diagram showing how features 7 and 8 are determined.

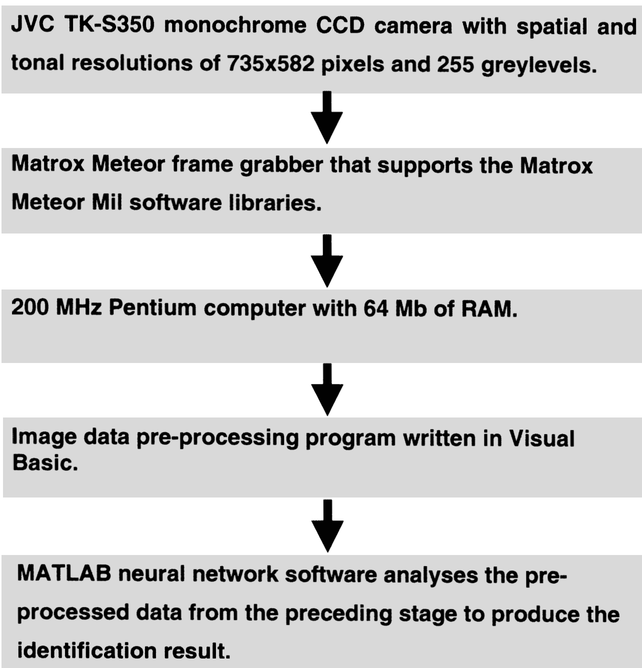


Figure 13. An overview of the image analysis system.

The aim of the classification trials was to show that the system was capable of differentiating between laser prints and their photocopies for the following images:

i 10 mm squares
ii the letter 'a' in a 36 pt bold Arial font
iii 1 mm diameter circles

In the first classification trial, a series of 23 squares (15 differing tones and 8 solids) was produced using a Hewlett Packard 4M plus laser printer. These images were subsequently photocopied using a Sharp SF-2022 photocopier. These 46 images were pre-processed individually by the image analysis system; this formed the training set for the neural network. Each laser print was assigned an output target value 1 and each photocopy an output target value 0. A further series of the 23 squares was printed using the Hewlett Packard 4M plus laser printer and reproduced using an Oce Bookcopier and Canon 6030 photocopiers.

The first step was to produce a classification model for the training set described in Fig. 15 by applying different combinations of the eight features to the MLP network. If this is achieved then all the training data is correctly classified. The next step is to see whether the data in the validation set can be correctly classified. If there are classification errors in the validation set, the network can be trained further to a lower E_n value, the number of hidden neurons changed, the network retrained or different features applied.

Figure 14. The arrangement of the CCD camera and illumination system.

The results from two network models produced for the prints of the squares are shown in Tables I and II. The results of Table I were produced using features 1, 4, and 6; the results of Table II were produced using features 2, 3, and 7. The data from the training sets for these two models are also plotted in Figs. 16(a) and 17(a). Figures 16(b) and 17(b) show the corresponding network structures, the transfer functions and the network parameters used to produce the results of Tables I and II. The hidden layer and output neuron used the tansig function and logsig function respectively.

The results shown in Tables I and II were the most accurate produced in the classification trials. These 'optimum' models were produced by adjusting the network parameters to 'tune the network'. The data presented here (including the results for the text character and circle) was the result of running the neural network program approximately five hundred times. The LM algorithm converged much faster than standard back propagation. The LM method needed approximately 10 s while the back propagation method required about 1 h for each run. Because the accuracy of the two methods was found to be comparable it was decided to use LM method in this investigation.

In the next classification trial images of the text character 'a' replaced the squares and the same network was used to classify the laser prints from their photocopies.

Training set

23 images of squares of 15 differing tones and 8 solids produced using a Hewlett Packard 4M laser printer, Target output = 1

Sharp SF-2022 photocopier reproductions of the 23 images of the squares of 15 differing tones and 8 solids produced from the Hewlett Packard 4M printer, Target output = 0

Validation set

23 images of squares of 15 differing tones and 8 solids produced using a Hewlett Packard 4M laser printer (not the same set of images as the training set), Target output = 1

Oce Bookcopier photocopier reproductions of the 23 images of the squares of 15 differing tones and 8 solids produced from the Hewlett Packard 4M printer, Target output = 0

Test set

23 images of squares of 15 differing tones and 8 solids produced using a Hewlett Packard 4M laser printer (not the same set of images as the training set but the same set as the validation set), Target output = 1

Canon 6063 photocopier reproductions of the 23 images of the squares of 15 differing tones and 8 solids produced from the Hewlett Packard 4M printer, Target output = 0

Figure 15. Block diagram showing the training, validation and test sets of data used in the classification trials for the images of the squares.

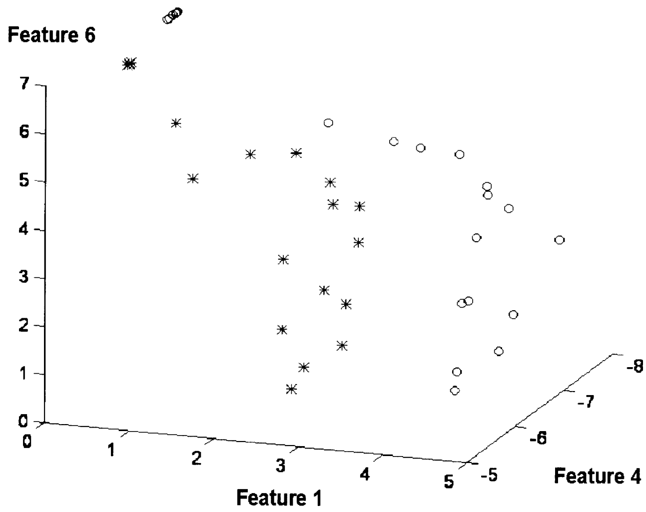
Two different photocopiers were used for the validation and test sets; a Xerox 5343 and an Infotec 5402 MF replaced the Oce Bookcopier and Canon 6030 respectively. The first set of results for this trial, shown in Table III, used features 1, 4 and 6. The same network architecture and parameters that were used to classify the squares were employed. Figure 18 shows this network along with its transfer functions and network parameters. From these results, it can be seen that for the text character ‘a’, there are nine errors, compared to zero errors for the ‘squares’. However with the following parameters: features 1, 2, 3, 7 and 8, 3 hidden neurons, E_n set to 0.1 and the learning rate set to 0.1, the number of errors was reduced to two. The network architecture and results for this test are shown in Fig. 19 and Table IV.

Figure 20 shows the neural network, the transfer functions and the network parameters used to obtain the results in Table V for the images of circles. A difference between this trial and the previous trials is that a change of laser printer had been made instead of a change in photocopier for the validation set.

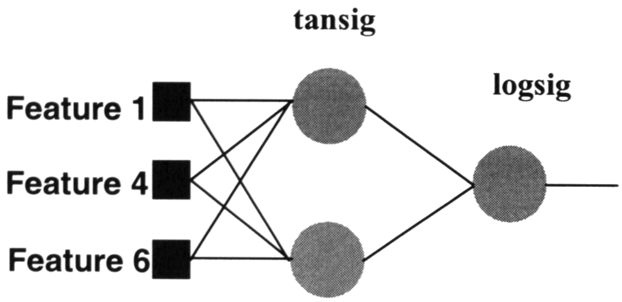
The results demonstrate that the system was able to correctly classify the majority of the prints. The squares were classified with the greatest accuracy and the results took the shortest time to produce, approximately 30 min. Classification of the circles and the ‘a’s both took approximately 8 h. Possible reasons why the system performed the classification task on the squares better than the two other prints are firstly, the number of prints examined and secondly the area of the print. 23 squares were investigated as opposed to 20 ‘a’s and 15 circles. The size of the print determines the volume of data about that print. From Fig. 7 it can be seen that the squares present the greatest area with the ‘a’ covering the smallest.

Conclusions

It has been demonstrated that the system can differentiate between laser prints and their photocopies over a range of tones. It also showed that an MLP neural network employing the LM algorithm was able to perform the classification task efficiently. The classification ac-



(a)



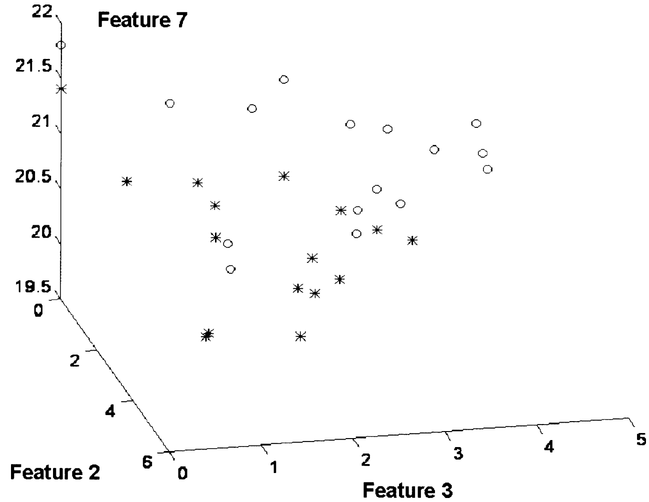
(b)

Figure 16. (a) A plot of features 1, 4, and 6 for the prints of the squares in the training set. Laser = *o*, Photocopy = *; (b) the neural network used to produce the data for the squares shown in Table I.

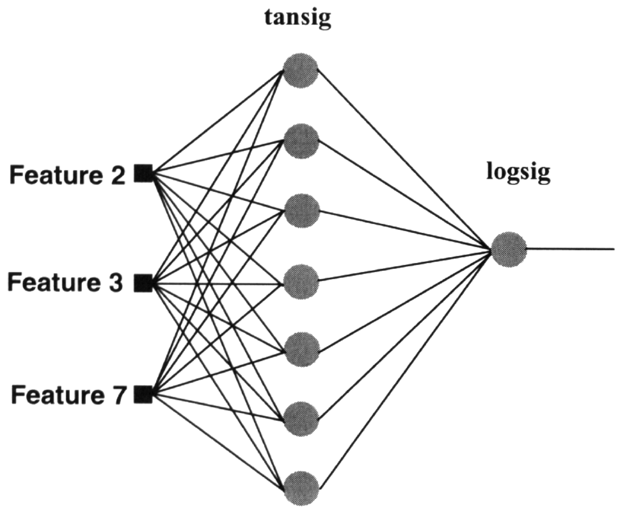
curacy varied and depended both on the image employed and the combination of features used.

Pre-processing the data using the present system required approximately fifteen minutes for each print, this time excludes the time taken for running the neural network. In order to improve both the accuracy and the processing speed of the system both the hardware and software of the system need to be improved. This should permit the solution of the following two problems: (1) identifying whether a print sample is a laser print or an optical photocopy using only a single neural network model and (2) identifying the make of laser printer or copier that produces the print sample. To solve these problems, a different set of experimental features is required which involves analysis of toner satellites on non-image areas of prints and print mottle.

To achieve the above a new system with a faster processor (400Mhz cpu) is under development. A more sensitive digital camera, the Hamamatsu C4742-95 has replaced the JVC TK-S350 analogue camera. The Hamamatsu camera has the advantages of Peltier cooling, which also increases its signal to noise ratio, and gives higher tonal and spatial resolutions. Lenses that are capable of greater magnification will enable smaller print samples to be analysed and a wider variety of



(a)



(b)

Figure 17. (a) A plot of features 2, 3, and 7 for the prints in the training set. Laser = o, Photocopy = *; (b) the neural network used to produce the data for the squares shown in Table II.

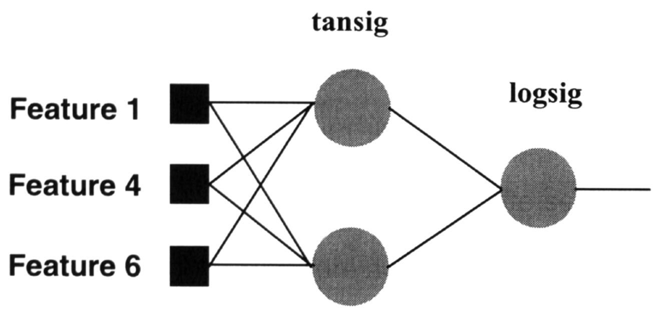


Figure 18. The neural network architecture used with the Levenberg–Marquart algorithm to produce the results for the letter ‘a’ as shown in Table III.

Figure 19. The neural network architecture used to produce the results for the letter ‘a’ shown in Table IV.

Figure 20. The neural network architecture used to produce the results shown in table 5 for the images of a circle.

TABLE I. The Output Results for the Training, Validation and Test Sets for the Square Using the Network Structure Shown in Fig. 15(b).

Neural network results for the square using features 1,4,6					
Training Set		Validation Set		Test Set	
HP 4M laser printer (set1)	Sharp-2020 photocopier	HP 4M laser printer (set2)	Oce Bookcopier photocopier	HP 4M laser printer (set2)	Canon 6030 photo-copier
1.0000	0.0053	0.9998	0.0028	0.9998	0.0062
1.0000	0.0062	1.0000	0.0025	1.0000	0.0051
1.0000	0.0266	1.0000	0.0026	1.0000	0.0480
0.9887	0.0032	0.9998	0.0025	0.9998	0.0267
1.0000	0.0082	1.0000	0.0164	1.0000	0.0052
1.0000	0.0077	0.9998	0.0047	0.9998	0.0102
1.0000	0.0153	1.0000	0.0026	1.0000	0.0041
0.9955	0.0029	0.9354	0.0025	0.9354	0.0032
1.0000	0.0127	1.0000	0.0025	1.0000	0.0028
1.0000	0.0078	1.0000	0.0025	1.0000	0.0034
1.0000	0.0194	1.0000	0.0025	1.0000	0.0027
1.0000	0.0025	1.0000	0.0025	1.0000	0.0038
1.0000	0.0335	1.0000	0.0025	1.0000	0.0026
1.0000	0.0053	1.0000	0.0025	1.0000	0.0025
0.9835	0.0036	0.9115	0.0025	0.9115	0.0026
0.9878	0.0139	1.0000	0.0035	1.0000	0.0048
0.9870	0.0148	1.0000	0.0037	1.0000	0.0054
0.9935	0.0138	1.0000	0.0040	1.0000	0.0057
0.9999	0.0165	1.0000	0.0037	1.0000	0.0064
0.9998	0.0139	1.0000	0.0042	1.0000	0.0079
0.9998	0.0121	1.0000	0.0047	1.0000	0.0061
0.9999	0.0137	1.0000	0.0259	1.0000	0.0064
0.9991	0.0149	1.0000	0.0042	1.0000	0.0070

TABLE II. The Output Results for The Training, Validation and Test Sets for the Square Using the Network Structure Shown In Fig. 16 (b). The Classification Errors Are Highlighted In Bold.

Neural network results for the square using features 2, 3, and 7					
Training Set		Validation Set		Test Set	
HP 4M laser printer (set1)	Sharp-2020 photocopier	HP 4M laser printer (set2)	Oce Bookcopier photocopier	HP 4M laser printer (set2)	Canon 6030 photo-copier
0.9978	0.0012	**0.2661**	0.0002	**0.2661**	0.0000
0.9860	0.0010	0.9918	0.0002	0.9918	0.0000
0.9981	0.0023	0.9859	0.0001	0.9859	**0.7348**
0.9976	0.0001	0.9982	0.0000	0.9982	0.3285
0.9983	0.0017	0.9980	**0.9992**	0.9980	0.0053
0.9919	0.0000	0.9873	0.0003	0.9873	0.0000
1.0000	0.0022	1.0000	0.0002	1.0000	0.0010
0.9999	0.0000	1.0000	0.0000	1.0000	0.0007
1.0000	0.0004	**0.3604**	0.0000	**0.3604**	0.0004
1.0000	0.0004	1.0000	0.0000	1.0000	0.0004
1.0000	0.0042	1.0000	0.0000	1.0000	0.0004
0.9833	0.0004	**0.0026**	0.0004	**0.0026**	0.0004
1.0000	0.0004	1.0000	0.0000	1.0000	0.0004
1.0000	0.0004	1.0000	0.0000	1.0000	0.0004
0.9838	0.0004	1.0000	0.0004	1.0000	0.0004
0.9746	0.0092	0.9366	0.0001	0.9366	0.0064
0.9721	0.0084	0.9386	0.0001	0.9386	0.0105
0.9740	0.0077	0.9371	0.0001	0.9371	0.0166
0.9748	0.0075	0.9382	0.0000	0.9382	0.0272
0.9738	0.0074	0.9376	0.0001	0.9376	0.0467
0.9751	0.0070	0.9384	0.0001	0.9384	0.0164
0.9764	0.0073	0.9397	0.0001	0.9397	0.0269
0.9749	0.0087	0.9381	0.0001	0.9381	0.0523

TABLE III. The Classification Results for the Laser and Photocopied Print Samples of the Letter 'a'. The Neural Network Structure Shown in Fig. 18 was used to Obtain These Results. The Classification Errors are Highlighted in Bold.

Neural network results for the letter 'a' using features 1, 4, and 6					
Training Set		Validation Set		Test Set	
HP 4M laser printer (set1)	Sharp-2020 photocopier	HP 4M laser printer (set2)	Oce Bookcopier photocopier	HP 4M laser printer (set2)	Canon 6030 photo-copier
0.9861	0.0819	0.9859	0.0034	0.9991	**0.6776**
0.9338	0.0723	0.6160	0.0392	0.5097	0.0000
0.9999	0.0064	0.9996	0.0071	0.9997	**0.9998**
1.0000	0.0000	1.0000	0.0618	1.0000	**0.9998**
0.9998	0.0000	1.0000	0.0010	1.0000	0.3231
1.0000	0.0000	1.0000	0.0055	0.9999	0.0000
1.0000	0.0000	1.0000	0.0001	0.9999	0.0000
1.0000	0.0000	1.0000	0.0001	0.9999	0.0013
1.0000	0.0000	1.0000	**1.0000**	1.0000	0.0000
1.0000	0.0000	1.0000	0.0307	1.0000	**0.9755**
1.0000	0.0000	1.0000	0.0045	1.0000	0.0321
1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
1.0000	0.0000	1.0000	0.0000	0.9999	0.0000
1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
1.0000	0.0000	0.9999	0.0000	1.0000	0.0000
1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
1.0000	0.0000	0.9997	0.0000	0.9998	0.0000
0.9954	0.0000	**0.0002**	0.0000	**0.0056**	0.0000
0.9998	0.0000	**0.0000**	0.0000	1.0000	0.0000
0.9999	0.0000	0.9999	0.0000	**0.0000**	0.0000

TABLE IV. The Classification Results for the Laser and Photocopied Print Samples of the Letter 'a'. The Neural Network Structure Shown in Fig. 19 Was Used to Obtain These Results. The Classification Errors are Highlighted in Bold.

Neural network results for the letter 'a' using features 1,2,3,7,8					
Training Set		Validation Set		Test Set	
HP 4M laser printer	Sharp SF-2022 photocopier	HP 4M laser printer	Xerox 5343 photocopier	HP 4M laser printer	Infotec 5402 MF photo-copier
0.9987	0.0443	0.9269	0.1607	0.9977	0.0207
0.8124	0.0848	0.6122	0.0246	**0.2606**	0.0099
0.9893	0.0414	0.9512	0.2458	0.7552	0.0073
0.9993	0.0419	0.9993	0.3016	0.9988	0.0052
0.9646	0.1428	0.9964	0.0368	0.8209	0.0037
0.9939	0.0902	0.9963	0.0860	0.9935	0.0038
0.9971	0.0134	0.9988	0.1028	0.9975	0.0037
0.9984	0.0243	0.9976	0.0583	0.9986	0.0057
0.9995	0.0180	0.9995	**0.9860**	0.9995	0.0063
0.9996	0.0121	0.9996	0.1259	0.9996	0.0075
0.9996	0.0050	0.9996	0.3711	0.9996	0.0034
0.9997	0.0084	0.9994	0.0375	0.9996	0.0059
0.9998	0.0099	0.9996	0.0073	0.9998	0.0079
0.9996	0.0073	0.9995	0.0171	0.9997	0.0059
0.9999	0.0097	0.9995	0.0103	0.9998	0.0124
0.9997	0.0048	0.9999	0.0320	0.9996	0.0082
0.9998	0.0037	0.9996	0.0087	0.9999	0.0073
0.9995	0.0083	0.9988	0.0066	0.9995	0.0110
0.9996	0.0054	0.9980	0.0065	0.9995	0.0046
0.9997	0.0049	0.9996	0.3214	0.9997	0.0046

TABLE V. The Classification Results for the Image of a Circle Produced by Two Different Laser Printers and Their Photocopies. The Neural Network Structure, Using the Transfer Functions and Network Parameters Shown in Fig. 20 Was Used to Obtain These Results. The Classification Errors are Highlighted in Bold.

Neural network results for the circle using features 1, 2, 3, 4, and 7			
Training Set		Validation Set	
Hewlett Packard 4M Laser Printer	Sharp SF2022 photocopier	Kyocera FS-1700 Laser printer	Sharp SF-2022 photocopier
0.9872	0.0108	**0.0129**	0.0106
0.9874	0.0106	0.9869	0.0106
0.9874	0.0106	0.9833	0.0106
0.9868	0.0106	**0.0125**	0.0106
0.9874	0.0106	0.9866	0.0106
0.9874	0.0106	0.9870	0.0106
0.9874	0.0106	0.9874	0.0107
0.9874	0.0106	0.9864	0.0474
0.9874	0.0107	0.8832	0.0820
0.9874	0.0122	0.9836	0.1041
0.9874	0.0900	0.9332	0.2000
0.9874	0.0496	0.9862	0.1115
0.9874	0.1050	0.9873	0.4226
0.9137	0.1372	0.9604	0.0800
0.9503	0.1627	0.9798	0.1863

features to be extracted. The application of C or machine code for the pre-processing program is also under consideration.

Acknowledgements. The authors would like to thank Colin Parsons for his help with the installation of the computer hardware.

References

1. C. J. Baron and E. J. Pritchard, *Print Recognition*, PIRA, Surrey, England, 1971.
2. G. E. Kowalczyk, *Tappi J.* **81**, 181 (1998).
3. H. Saarelma and P. Oittinen, *Graphic Arts in Finland* **26**, 8 (1998).
4. C. Antoine, *Proc. Pan-pacific and international printing and graphic arts conference*, Pulp and Paper Technical Association of Canada, Montreal, Canada, 1998, p. 49.
5. Y. Kipman, *Proc. Print and Image Quality Conference,* San Diego, California, USA, Information Management Institute Inc., Windham, MA (1996).
6. P. Spencer and D. Schwartzwald, *NIP12: International Conference on Digital Printing Technologies,* IS&T, Springfield, VA, 156 (1996).
7. T. Safadi, *Professional Printer* **38** 10 (1994).
8. M. Duhamel, M. Auvray and F. Chatton, *Proc. 7th International Congress on Non-Impact Printing Technologies,* IS&T, Springfield, VA, 1991, p. 939.
9. J. Edinger, *J. Imaging Sci.* **31**, 177 (1987).
10. Y. Tanaka and T. Abe, *J. Imaging Technol.* **13**, 202 (1987).
11. R. Beale, *Neural Computing: An Introduction*, Institute of Physics Publishing Ltd, Bristol, England,1994, pp.1-105.
12. S. Haykin, *Neural Networks,* Macmillan College Publishing, Ontario, Canada, 1994, pp.1-229.
13. L. Tarassenko, *A Guide to Neural Computing Applications*, Arnold, London, England,1998, pp.1-18.
14. C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, England, 1995, pp.1-163.
15. J. L. Automation, *Image Analysis Principle and Practice*, J. L. Automation, Joyce Loebl Ltd, Gateshead, England, 1989, pp.3-57.
16. H. Demuth and M. Beale, *Neural Network Toolbox*, For use with Matlab, User's Guide, The MathsWork Inc, Natick, Mass., USA,1994.