NEURAL NETWORKS IN IMAGE PROCESSING: A REVIEW OF CURRENT APPLICATIONS

SHENG-HUNG CHUNG¹ and RAJASVARAN LOGESWARAN²
Faculty of Engineering, Multimedia University, 63100 Cyberjaya, Malaysia. (shchung@mmu.edu.my¹ and loges@mmu.edu.my²)

ABSTRACT
This paper describes current applications of neural networks in image processing. Artificial neural networks (ANNs) are methods of computation and information processing modelled by the brain. Many recent attempts to improve the flexibility and effectiveness of ANNs have focused on the implementation level. In this article, we look into ANNs used in the different stages of image processing, specifically in the pre-processing, data reduction, segmentation, object recognition and image understanding phases. The focus is on current and future ANNs, including feed-forward networks, Kohonen feature maps, Hopfield networks, goal-seeking neuron (GSN) and cellular neural network (CNN). New types of ANNs are fast increasing. Through this survey of introducing the findings, implementations and recent advances of ANNs in image processing, it is hoped that this paper will serve as a summary, or base to accelerate further development and use of ANNs in the field of image processing, and improving the accuracy and speed of image processing tasks in the future.

Oktem and Joony (2004) implemented the detection of malignant tumours at an early stage which is an important stage in diagnosis of cancer regions in mammograms by using back-propagation network and a self-organizing map. There is a significant and growing usage of ANNs in image processing to solve complex problems automatically. ANNs give professionals, as well as non-specialists, the tools to train, visualize and validate neural network models to solve interesting problems in image processing.

Neural networks for image processing have been largely derived from the use of conventional techniques, for example: a Laplacian operator with thresholding (Cheng et al., 1993; Kendall et al., 2002) and moment-preserving techniques (Cheng, 1993). The term “image processing” refers to enhancing and manipulating an image, such as by adjusting its size, resolution, or colour palette. Image processing consists of five different tasks, namely pre-processing, data reduction, segmentation, object recognition and image understanding. Optimisation is used as an auxiliary tool that is available in all steps of the image processing chain (Figure 1).

This paper attempts to provide a brief survey of ANN techniques developed for the mentioned tasks, provide a foundation for the reader to work on further development of ANN techniques in other image processing applications.

INTRODUCTION
There has been much interest recently in developing neural network models and applications to solve complicated information processing problems. Several recent applications that use artificial neural networks (ANN) in image processing have obtained results in solving real world problems such as gender recognition (David, 2001) and gender classification (Sun, 2000) from facial images using neural network and Genetic Algorithm (GA). Jakubowska et al. (2004) presented the classification of the thermal images in order to discriminate between healthy and pathological cases during breast cancer screenings.

PRE-PROCESSING
In pre-processing, image analysis techniques are used to enhance and extract desired characteristics of the image by applying algorithms such as filtering and edge detection. Image pre-processing is necessary to facilitate the remaining image processing tasks for identifying the portion of an image to be studied. Neural networks are
used in terms of algorithm application, with the numerical parameters of the image, such as the gray scale intensity and contrast taken in as the input. Chang et al. (2004) utilized a statistical pattern recognition algorithm to identify defects of images of real life objects, based on their numerical representations. The same concept was applied by Bishop (1996), where statistical pattern recognition was used in the pre-processing of intensity images.

Pre-processing operations are categorised into image reconstruction, image restoration, and image enhancement (Petersen, 2002). Applications of several ANNs used in pre-processing, mainly Hopfield ANNs, ADALINE, regression feed-forward, Cellular Neural Network (CNN), random neural network, and Fuzzy Cellular Neural Network (FCNN), and Arbitration Neural Network are discussed below.

**Image reconstruction**
Images are frequently reconstructed from data obtained from a number of sensor measurements. Wang et al. (1997) trained a Hopfield ANN for reconstruction of 2D images from pixel data obtained from various projections. The ADALINE network is trained to perform an electrical impedance tomography (EIT) reconstruction, a 2D image based on 1D measurement on the circumference of the image Peterson et al. (2002). Srinivasan et al. (1993) trained a Hopfield ANN to perform the inverse Radon transform for reconstruction of computerised tomography (CT) images.

The Hopfield ANN consists of “summation” layers to avoid having to interconnect all units. In (Srinivasas's study, 1993), a regression feed-forward network is trained to learn the mapping E(y|x), with x the vector of input variables and y the desired output vector to reconstruct images from electron holograms.

**Image restoration**
Image restoration removes aberrations introduced by the sensors, including noise. In general, one wants to restore an image that is distorted by the (physical) measurement system. Such distortions include noise, motion blur, out-of-focus blur, distortion and low resolution. Image restoration employs information about the nature of the distortions introduced by the system. In the studies by Celebi et al. (1997), a 3D CNN is applied for restoration of noisy and blurry images. Contrary to larger solutions offered using a discrete-time Hopfield network, only eight cells are used for each pixel, reducing the network size. The presented image restoration scheme does not have a convergence problem which is inherent in discrete-time Hopfield network solutions.

**Image enhancement**
Image enhancement is used in the accentuation of certain desired features, which may facilitate later processing steps such as segmentation or object recognition. The most well known enhancement is edge detection. In (Egmont-Petersen et al., 2002), a statistical approach to multi-slice edge detection was proposed. Chandrasekran (1992) reported a novel feed-forward architecture which was used to classify an input window in determining the existence of an edge as shown in Figure 2 (c). Edge detection significantly reduces the amount of data by filtering out useless information, while preserving the important structural properties in an image.

In recent studies by Gelenbe et al. (2001),

---

**Figure 1.** Image Processing Chain
the random neural network (RNN) is trained using zero-order interpolation to perform enhanced image enlargement. The RNN network is trained to produce, as its output, a difference image \( T_{diff,i} \), which minimises the square of the sum of pixel errors using. Results in (Gelenbe et al., 2001) have shown that the network learning time has been successfully reduced in the implementation.

Another ANN, Fuzzy Cellular Neural Network (FCNN) is used in application of gray-scale mathematical morphology and fuzzy inference edge detection. The FCNN structures are based on the uncertainties in the human cognitive processes and in modeling neural systems, and provide an interface between the human expert and the classical CNN. A result from Yang et al. (1996) using a FCNN is shown in Figure 3 below:

Ramalho and Curtis. (1995) implemented edge detection using a representative number of the various edge detection schemes. The different edge detection techniques are utilized in parallel for generalized edge detection of different image types. This allows the attainment of several edge maps containing different features. From the edge maps, edges are arbitrated with a Back-propagation neural network (BPN) (Ramalho and Curtis, 1995), performing the merging of the different maps. To overcome the increased computing time, implementation is performed on a multi-transputer array, using the inherent parallelism of the BPN.

The result of the edge detection is shown in Figure 4 where the neural network arbitration is shown to be effective in improving the achieved edge map.

A recently introduced type of CNN, namely Hysteretic Cellular Neural Network (HCNN), may be used for edge detection of simple binary images. HCNN extracts the contour of a single black figure on a white background as shown in Figure 5. A comprehensive theoretical analysis of this HCNN’s stability and equilibrium points is given in (Matei, 1998; Slavova, 1996).
DATA REDUCTION

In this article, we look at two popular applications of ANNs in data reduction, image compression and feature extraction. Applications using BPN (Ramonha, 1995), goal-seeking neuron (GSN) (Matei, 1998; Slavova 1996), Boolean neural network (Jin’no, 1996) and CNN (Carvalho, 1991) are discussed in this section. The BPN (Ramonha, 1995) model applied for image compression is set up as a multilayer correspondence network with supervised learning. It allows the calculation of a functional relationship between input and output.

Each input pattern is associated with its corresponding output pattern (Lanzarini, 1997). The back-propagation algorithm (BPA) has been a very popular training algorithm for multilayer perceptrons. The disadvantage of BPA is that the number of the training steps for a given training set may be very high. The number of steps depends on the initial values of the node input weights \( W \) that may not be determined for complex problems (Lanzarini, 1997).

The GSN is a deterministic neuron designed to overcome many of the problems found in other Boolean models, maximizing the efficiency of the storage of values in its memory, storing new information without disrupting previously absorbed information, and employing single-pass learning (Carvalho et al., 1997). Feature extraction using GSN is focused more on the appropriate encoding of the raw data through a feature

**Figure 5.** (a) Original Binary input image with white background (b) Rescaled output

The result of edge detection capabilities of the CNN is shown in Figure 6. It is shown in (Jin’no and Tanaka, 1996) that the neural network clearly provides superior edge detection and only takes 10% of the processing time of conventional techniques.

An idea of cellular neural filtering has been initiated by the idea of CNN (Carvalho et al., 1991). Non-linear cellular neural filters (NCNF) include multi-valued non-linear filters (MVF) and cellular neural Boolean filters (CNBF). MVF is successfully used for noise reduction (Wang and Wahl, 1997; Stassopoulo et al., 1996). Gaussian, uniform and other kinds of noises may be removed from the image, but at the same time MVF ensures maximal preservation of the image boundaries. An example of MVF application to noise reduction is shown in Figure 7.

**Figure 6.** Result of edge detection
extraction process realized by means of a Boolean network. In many image classification tasks, an important focus is on the feature extraction process. In the Boolean network architecture, neurons are grouped in a set of blocks, where each block is a set lattice structure and is associated with the sub-area of the input image. By using an unsupervised learning algorithm, the block extracts its own feature set. It has been proven that the development of integrated Boolean network architecture, encompassing both feature extraction and classification networks, is possible (Adler and Guordo, 1994).

Another particular area of image filtering and image enhancement is colour processing. Enormous enhancements have been made by many researchers in this field (Adler and Guordo, 1994). The CNN model has been rapidly used to cover a wide range of image filtering applications which are typically characterized by their spatial dynamics. Simulation strategies based on the CNN multilayer model is an ideal architecture for colour image processing. This is because each pixel's colour can be handled as an RGB (red, green blue) triplet whose combinations yield a secondary colour (Meyer and Heindl, 1998).

Each layer of CNN cells represents each primary colour component, carrying out the processing independently to produce the results. The purpose of the colour model is to facilitate the specification of colours. Basically, a colour model is a specification of a three-dimensional coordinate system and a subspace within that system, where each colour is represented by a point. The RGB model is an additive primary system that describes a colour in terms of the percentage of red, green, and blue in the colour (Lauterbach, 2001). Using the RGB model has the advantage that each primary colour can be represented by a CNN layer, as shown in Figure 8.

SEGMENTATION
Segmentation is the process of dividing a scene into a number of individual objects or sections, differentiating them from each other and the image background. Segmentation used in classification tasks refers to assignment of image objects to one or more possible groups. Decisions are made by evaluating features either statistically or structurally based on relationships.

Gan (1992) implemented fuzzy neural network to improve the resolution of multi-resolution images and the segmentation of images. According to Gan (1992), the back-propagation neural network is used to obtain an optimized membership function where the algorithm is implemented for both types of application. The advantage of using fuzzy neural networks compared with conventional neural networks is to reduce the number of elements in each neural network layer. Thus computation time can be reduced. Recently, the moment-preserving technique that is most widely used in clustering may be thought of as an information transformation method that groups the pixels of an image into classes (Cheng, 1993). The difference between the moments of an image and those of a processed version of the image may be
used as a goodness criterion for an image operation.

The concept of moment preserving techniques is to group the gray values of the pixels image into several classes and replace all the gray values in each class with an identical gray value. The variables in the moment-preserving equations can be determined interactively by a recurrent neural network and a connectionist neural network, working cooperatively. Both of the networks are designed in such a way that the sum of square errors between the moments of the input image and those of the output version is minimized. This approach is suitable for clustering problems when an appropriate distance function or similarity measure is defined. Other popular segmentation techniques implemented using neural networks are given below:

**Texture classification**

Restriction is added to the texture classification network to identify that the same set of weights between the input sub-image and the hidden layer units is the same for each unit. Texture classification is convenient using an ANN because a model of the texture has been 'learnt' by the Hopfield ANN (Aizenberg, 1996) by training with sample textures.

**GSN Self-Organising Architecture**

GSN uses single-layer clusters instead of pyramids and works by creating new clusters to store unfamiliar patterns and grouping similar patterns in the same cluster (Lee et al., 1996). GSN requires only two training phases: a learning phase and a matching phase. By using an unsupervised learning algorithm, this architecture self-organises input patterns into clusters of neurons, associating each cluster with one class. Before a new pattern can be learned, the matching phase is invoked to decide if the pattern is similar to any already existing clusters. Otherwise, a new cluster is created and taught the input pattern. In classifying an unknown pattern, the matching phase is invoked to calculate the degree of similarity between the input pattern and each cluster, in order to assign the unknown input to an appropriate class.

**OBJECT RECOGNITION**

Object recognition consists of locating the positions and possibly orientations and scales of instances of objects in an image. Several ANNs have been trained to locate individual objects based directly on pixel data. In computer vision, the object recognition system typically involves some sort of a sensor, the use of a model database in which all the object "models" representation are saved, and a decision-making ability. The digitized image is processed so as to represent it in the same way as the models are represented in the database.

**Neural networks**

An interesting approach that performs object recognition, which is invariant to 2D translation, in-plane rotation and scale, is the neurally inspired Pyramidal neural network (PNN) (Kim, 2001). PNN works similar to a multilayer feed-forward network with reducing numbers of processing elements (nodes) in each layer, for example, 5x5 inputs to 4x4 first hidden layer to 3x3 to 2x2 and finally, all the connection weights are fixed to 1x1 output, and thresholding in each layer is set to the minimum value of the layer's elements to identify an object in a test image.

Most of the neural networks approaches studies were focused on examining pixel intensities of image blocks for object detection. An adaptive on-line neural network which is incrementally grown and based on a shape-based approach has been used for single shape and clustered shape object detection in (Carpenter et al., 1998). Self-organizing map (SOM) is used in pattern recognition, where it could be applied to dimensionality reduction in images, achieving recognition rates of 75.7% (Kyoung, 2003).

Koch et al. (1990) proposed a selective attention neural network for recognising multiple objects in any position or orientation in a image; selective attention controls in their research focused on the scene and at specific times, through neural networks implementation. The object recognition module (ORM) is implemented using a recurrent neural network where the ORM takes the output of the selective attention module, builds up an internal representation of
the object and uses it to make a decision on the object's identity.

The images used in the object module can also consist of multiple objects and possibly partially occluded objects. By using selective attention controls, we may avoid processing an entire pixel image, where the number of connections needed for a 256 x 256 medium resolution image can easily go beyond the conventional neural network simulator.

**Object recognition module**

Figure 9 below shows the approach of Koch and Zhou (1990). The object recognition module receives the current field of focus or receptive field, and the change in position from the last field of focus. Based on these two inputs the recognition needs to build an internal model of the object and determine the object's identity by using the internal model. A recurrent neural network allows the module to have a memory of the input's history and its current internal model. Part of this feedback includes the recognition module's current decision about the object's identity.

The selective attention module determines where to look next and outputs the relative position of the new receptive field with respect to the last receptive field's position. The receptive field extractor moves the current focus of attention and produces a new receptive field. It determines this new receptive field from the selective attention module outputs and the current input scene, and supplies the new receptive field into the recognition module.

Many approaches have been introduced for 3-D Object recognition (3DOR) (Sahambi and Khorasani, 2003). To achieve 3DOR, the pose of the objects is saved in the database and the ANN algorithm is not only to recognize the object correctly but also to identify its pose as viewed by camera. Figure 10 below shows the training views of one of the objects stored in the database taken from Sahambi and Khorasani, (2003).

A Modular neural network classifier has been applied to the problem of automatic target recognition using forward-looking infrared (FLIR) imagery (Wang et al., 1998). The classifier consists of several independently trained neural networks. Each neural network makes a decision based on local features extracted from a specific portion of a target image. The classification decisions of the individual networks are combined to determine the final classification.

The experiment in Wang et al. (1998) shows that decomposition of the input features results

---

**Figure 9.** A Selective Attention Neural Network for Object Recognition

**Figure 10.** Training views of one of the objects. The pose angle was sampled at every 50°
in performance superior to a fully connected network in terms of both network complexity and probability of classification. The two types of modular networks applied in the experiment are a mixture of experts modular network (MEMN), which is the decomposition of data, and committee of networks modular network (CNMN), for the decomposition of features.

Another interesting approach by Philips (1998) is by using feed forward networks and a few relevant techniques such as principal component analysis (Sanger, 1989), projection pursuit, factorial analysis, dynamic link architectures and entropy-based wavelet encoding of images for face identification. Face recognition is more difficult than character recognition as the size of the gallery (individuals) is large and only one training example per person is available. Therefore, the above techniques pursue automatic selection of features that distinguish between classes and objects.

**IMAGE UNDERSTANDING**

It is well known that most approaches to image understanding require a number of distinct steps (Barnard, 2001). Image understanding couples techniques from segmentation or object recognition with knowledge of the expected image content. Image understanding requires a combination of a number of distinct steps: noise removal, segmentation, edge detection, and object recognition.

Barnard and Casasent (2001) employ conventional methods of image processing which can be used in conjunction with neural-net classifiers. This allows us to emphasise the relevant choices in the domain of image processing and pattern recognition. As images are typically subjected to a number of transformations, such as scaling, translation and rotation, the image understanding system must be able to detect objects independently of such transformations, which means that the classifier must be invariant to these transformations. There exist three classes of techniques for invariant neural network recognition (Barnard and Casasent, 2001).

Firstly, the structure of the network can be designed in such a way that its output is always invariant to certain transformations. Then an input image can directly be presented to the net without pre-processing, and the net will be able to recognize the detected objects despite arbitrary transformations. A wide variety of rotated views of a set of objects is presented to a supervised neural network, with all views of the same object given the same class label. If the ANN is able to learn to discriminate these objects properly, and the number of examples shown is large enough, the ANN will also generalize correctly other transformations than those shown.

One disadvantage of this approach is that it is not clear how an ANN is trained to recognize new objects invariantly, or how it is able to use this training to also recognize new objects invariantly. Thus, with present techniques, it would be necessary to retrain the ANN on all the transformations for every new object to be recognized. Another problem with this approach is that the demands placed on the classification system might be very severe if the dimensionality of the feature space is high. The main improvement ANN classifiers offer over most conventional classifiers is that the ANN classifiers allow more general decision boundaries in the feature space.

The majority of conventional classifiers use decision boundaries such as hyperplanes or hyperquadrics whereas ANN classifiers use piecewise-linear or piecewise-quadratic decision boundaries (Barnard, 2001). In a similar case, Stassopoulou et al. (1996) mapped the trained ANN onto a Bayesian belief network after training had been performed.

**Neural Tree Network**

Zhen and Fulcher. (1996) used a new pattern classification method called Neural Tree Networks (NTN) for face perspective understanding. The NTN consists of neural networks connected in a tree architecture. The use of a neural network at each tree node as in the NTN, results in better classification performance compared with conventional classifiers. The results show that the NTN compares favourably with both neural networks and decision trees. However, a single neural network at each node is not good enough
for complex pattern recognition as the decision function for complex patterns is always non-
continuous and non-smooth. Zhen and Fulcher (1996) implemented a neural network group-
based node tree for complex image understanding with non-continuous and non-smooth decision functions.

Image understanding is an interpretation process of the input signal by a knowledge set. Omori et al. (2001) used a set of images, each having some deformation and some objects in the images occluded others. The knowledge about the shape of the targets with no deformation is represented as the coefficients of the recognizing network, the Neo-Cognitron (Kim, 2001). The result was a sequence of each recognized target, indicating successful interpretation of the input image.

CONCLUSION
This paper summarises some of the advances made using ANNs to perform image processing tasks efficiently. Computation of the training time and resources of a neural network is important and should be taken into account when implementing ANNs solutions. Neural solutions are truly interesting when existing algorithms fail or when ANNs may reduce the amount of computation time and effort considerably, mainly through inherent parallelism and intelligence. Various ANNs are currently used for enhancing image processing. The CNN shows promise and further research is required on enhancements to the cellular network structure and training methods in order to improve the training time. It is believed that neural networks in image processing would also be very useful for medical practitioners to diagnose diseases more efficiently.

REFERENCES


*Neural Networks Technologies Ltd*, Israel.


