NOVEL NEURAL NETWORK METHODS FOR DESCRIBING ATTRIBUTES CONTAINED WITHIN LESIONS IMAGES

S.G.ANDREWS

QinetiQ
KISystems, Consulting
The Alba Centre, Alba Campus
LIVINGSTON
EH54 7EG
sandrews@QinetiQ.com
stuart@embedded-software.org.uk

Abstract – Several novel methods based on intelligent recognition techniques are presented. This is an extension of our earlier work that utilises a number of these techniques, which included production rules, genetic algorithms and associative memories. The first additions involve the use of an implied grammar that is derived from a dermatology heuristic that describes lesion morphology characteristics; this is also known by its mnemonic 'ABCDE'. The technique is combined with both our novel neural adaptive architecture and an image transform, the HALOGRAPH. This combination extracts the requisite components from the lesion image defined by the grammar, from which subsequently a diagnosis is generated. Another aspect of the lesion morphology is also considered: texture. The technique applied here makes use of Laws basis function to describe its structure. The texture landscape is defined as an energy profile. One of the limitations of our previous system involved the creation of volumes of information during the analysis phase, and to limit this effect we now apply PCA. The inclusion of this technique provides a reduction in the information content while retaining the essential detail, and it also acts as a method of feature recognition. A further side effect of this approach is that it can be used as an encoder to assist the simplification process of the internal definition of the evolving neural architecture.

Keywords: Production rules, Evolving Neural architectures, Image Recognition, and Genetic Algorithms, Grammars.

1. INTRODUCTION

The aim of the research of which this paper is a subset was to overcome some of the limitations that standard unsupervised neural architecture behaviour exhibits, and to develop an intelligent assistant that could be used both by general practitioners, and specialist dermatologists without having to resort to lengthy training.

Neural networks are used with increasing frequency in the context of medicine. Inspiration for artificial neural networks derives from their biological analogy the neurone. The earliest example of the artificial neural networks was the 'Perceptron', [Rosenblatt, 1958]. More complex neural architectures such as the Back-propagation network (BNN) [Rumelhart et al, 1986; Werbos, 1974], have been used in the detection of cancer cells [Moallemi, 1991], and with skin cancer recognition [Stoecker et al, 1997; Tomatis et al; 1997; Burroni et al, 1997]; A different architecture, namely the radial basis function network (RBF) [Burroni et al, 1997] was used in conjunction with spectroscopics for detection of cervical precancer [Tummer et al, 1998]. Other topics such as breast cancer screening have used neural networks [Bridgett et al, 1995]. A method of learning medical image shapes, and also in the context of image compression was applied by [Panagiotidis et al, 1996] using a probabilistic neural network (PNN). The topic of cranial pressure and monitoring of neuro-surgical patients was investigated [Swiercz et al, 1995, 1998] using a recurrent neural architecture (RNN) [Tomatis et al, 1977]. More recently neural networks have been used in capturing and defining MRI images [Reyes-Aldasori et al, 2000].

2. METHOD

In this paper we present a number of extensions and improvements to our novel neural network [Andrews et al, 1998, 1999, 2000]. The network is used to recognise skin cancer images. The architecture uses intelligent techniques, namely production rules [Nilsson, 1980], and the genetic algorithm (GA)

Proceedings 17th European Simulation Multiconference
(c) SCS Europe BVBA, 2003 ISBN 3-936150-25-7

[Holland, 1975; Harp, et al, 1991; Goldberg, 1989]**.** The first of these is a method for describing the internal structure of the neural architecture. One of the main behavioural features is its ability to remember feature patterns encountered. The concept that we make use of is that of association. We make reference to the concept of *associativity*. This concept has been extensively researched by [Anderson, 1970; Kohonen, 1974, 1988]**.** Training is handled by use of the memory concept, which allows any pattern previously encountered to be recalled. Any pattern anomalies are rapidly assimilated into the memory, leading to the modification of the neural definition. The behaviour of the novel neural network architecture makes use of an associative memory (AM) to control the assembly of the lesion image analysis. Each time a feature (pixel) is encountered a reference is made to the AM as to whether there had been a prior encounter with this feature, if not a neuron is assigned. Any further encounter with the same feature causes the neural strength to be incremented. A further fact that is elicited is the geographic distribution. This is assigned both neurons and associated strength. This process continues until all features are processed. Genetic algorithms (GA) [Holland, 1975; Harp, et al, 1991; Goldberg, 1989] are used in place of Hebbian learning [Hebb, 1949], and the feature histogram is used to influence the behaviour of the GA in its production of the weight definition.

Two further methods are used, which define characteristics contained within the lesion image. The first of these uses an implied grammar [Friedmann et al, 1985]**.** The second involves the lesion surface texture in terms of an energy map [Laws, 1979]**.** The first of these methods operates in conjunction with the implied grammar and an image transform, the HALOGRAPH [Andrews et al, 1998]. Other artificial intelligence techniques involved include information reduction and encoding (PCA) [Jolliffe, 1986; Sanger, 1989] and these operate with the novel neural network [Andrews et al, 1989, 2000].

The PCA is a powerful data analysis tool that is used in the context of multivariate analysis [Amari, 1977]. The technique used in Sanger's GHA is a generalisation in terms of the standard approach to PCA. One of the limitations of PCA which is that it makes various assumptions (e.g Gaussian Distribution) which as a result causes it not to characterise all the trends within the data. The transform simplifies the extraction and definition of the image details while simultaneously adding rigour to the diagnosis process thereby avoiding ambiguity.

The topic of surface texture is an important diagnostic tool [Stoecker et al, 1992, 1997]**.** Laws basis functions are used [Laws, 1979] to describe the details. These are used to differentiate the various micro-features of the surface texture that are present within the image. Having elicited the energy profile using this technique we then simplify the resulting product using a combination of PCA [Jolliffe, 1986] and neural network [Sanger, 1989]**.** The resultant definition is in effect a précis of the energy mapping. Using this approach allows for both recall and reconstruction of the précis. A further side effect of using PCA is the ability to be able to generalise. For instance, another image can be reconstructed using an already available PC definition [Cichocki et al, 1994, 1996]. A more comprehensive definition of the energy profile is constructed using the novel neural architecture [Andrews et al, 1998, 2000]**.**

The reason for using Laws is that it enables the characterisation of micro-features that are contained within the lesion image. It also provides a method of extracting a scaled variance profile with regard to each type of micro-feature.

3. APPLICATION

The following application illustrates how the novel neural architecture [Andrews et al, 1998, 2000] acts in combination with associated neural architectures [Oja, 1982; Sanger, 1989] to describe facets of lesion morphology. The first of these processes was explained earlier. This defines the lesion surface texture in terms of an energy profile**.** Using the energy profile [Laws, 1979] allows several levels of definition to be described. The resulting details are extracted and encoded using PCAGHA [Sanger, 1989]**.**

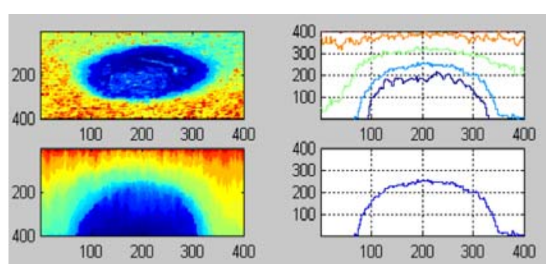

Figure 1 Lesion images & HALOGRAPH transformation of image.

This method allows for rapid recall of the original information. Modifications of the behaviour of the architecture [Andrews et al, 1998, 2000] are effected using the same process [Sanger, 1989] in which the detail [Nilsson, 1980] is simplified (see fig 1**).** This is applied at each level of the definition decomposition.

The above definition illustrates how the implied grammar (mnemonic, 'ABCDE') and describes the lesion morphology in terms of its geometric profile, boundary condition, and pigmentation distribution. Using standard imaging techniques this decomposition becomes extremely complex.

Figure 2 shows how the grammar unfolds the aspects of the lesion morphology and then uses the results to assemble a diagnosis. The grammar forms a partnership with the novel image transform [Andrews et al, 1998] and three other neural architectures [Andrews et al, 1998; Nilsson, 1980; Sanger, 1989].

R1: if $\sigma 1$ & $\sigma 2$	then 'A'
R2: if $\beta\chi 1$ & $\beta\chi 2$	then 'B'
R3: if $\chi\delta 1$ & $\chi\delta 2$, ..., $\chi\delta n$	then 'C'

Figure 2 Example rules describing initial level of the implied grammar.

The grammar acts in a supervisory role to extract and define salient aspects of each lesion image in relation to each aspect of the heuristic. The neural networks [Oja, 1982; Sanger, 1989] are used to encode firstly the lesions boundary contour, and secondly the pigmentation distribution.

4. RESULTS

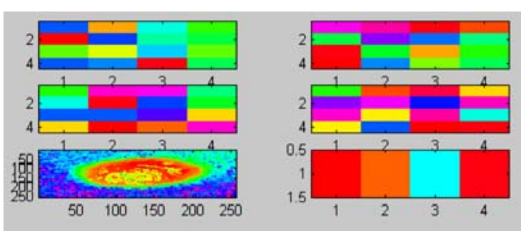

Figure 3 PCA definition of a lesion image

The above diagram, figure 3 shows how by applying PCA techniques based on Sangers GHA neural architecture [Sanger, 1989] the complexity of the information content can be significantly reduced. Without using the above approach the level of description is large. We mentioned earlier that the neural assembler decomposes the lesion image into a hierarchy that forms a binary-tree. Each level of the hierarchy has its own definitions. The initial level contains significant detail. Each subsequent level contains less and less detail (*see figure 5 & figure 6).*

The maximum level of decomposition is five after which there is little or no relevant information present. To illustrate the process involved we shall transform figure 1 into symbolic form.

In figure 4 we have four masks having 16 coefficients, which subsequently reduces to four PC's. By comparison with the table shown in figure 7, a similar table for images in figure 5 if expanded to their full form would occupy several pages. Should the full version be required it is still available.

R1 :If $\omega 1$ & $\omega 2$ & $\omega 3$ &, …, & $\omega 16$	Then mask 1
R2: if $\omega 1$ & $\omega 2$ & $\omega 3$ &, …., & $\omega 16$	Then mask 2
R3: if $\omega 1$ & $\omega 2$ & $\omega 3$ &, ….,& $\omega 16$	Then mask 3
R4: if $\omega 1$ & $\omega 2$ & $\omega 3$ &, …., & w16	Then mask 3
R5: if $\pi\chi 1$ & $\pi\chi 2$ & $\pi\chi 3$ & $\pi\chi 4$	Then PC coeff 1

Figure 4 Rules that defines the PC's

The other table describes some of the rules that describe the expanded form without the inclusion of PCA.

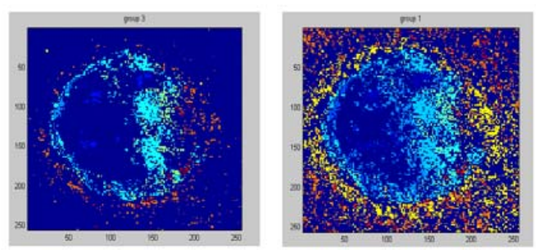

Figure 5 internal definition of a lesion image created by neural network

R1: If a0	Then b
R2: If b0 & b1 & b2 & b3 & b4	Then c
R4: If c1 & c2 & c3 & c4	then d1
R5: If a1> k1 & a1 <k2	Then g1
R6: If a2 >k3 & a1 <k4	then g2
R7: if a1>k5 & a1 <k6	then g3
R8: if a3>k76 & a1<k8	then g4

Figure 6 Extended rule definition

The graphs in figure 7, illustrate another aspect of each level of the hierarchy, the feature profile, and the details that influence the genetic algorithm (GA) in the computation of the connective mesh.

The above detail shown in figure 6 pertains to the symbolic definition of the features contained within the lesion. The first of which is a single definition, the others increase in complexity. The last four rules pertain to the hierarchical decomposition which defines the level at which the feature resides in terms of the decomposition.

5 CONCLUSIONS

In our previous paper [Andrews et al, 2000] we showed how a novel neural architecture could be used to emulate the process of edge detection.

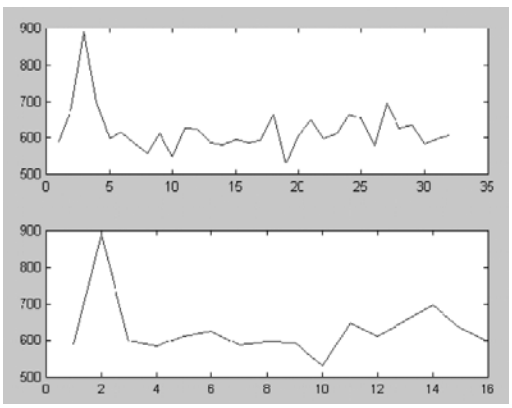

Figure 7 Graph profile of feature signature

In this paper we have extended the repertoire and tackled a few of the system's previous limitations. We have enhanced its capability by the inclusion of PCA that is used to simplify but also to encode the lesion image, and by the inclusion of a dermatology heuristic that describes the attributes of the lesion morphology [Friedman et al, 1985]**.** The heuristic, in the form of an implied grammar, is linked with the image transform, the HALOGRAPH [Andrews et al, 1998]. This combination adds rigour to the diagnostic process, and as a result there is less likelihood of ambiguity.

REFERENCES

ANDERSON J.A, (1970),'Two Models of Memory Organisation', Mathematical Bioscience, 137-160.

ANDREWS S.G, HARRISON P V. (1998),' The Fractal Dimension a Clue to Chaotic Growth an Approach to Skin Cancer Diagnosis'. Intl. Conf EADV, Nice, 1998.

ANDREWS S.G, ZEIN N.F, HONARY.B, HARRISON P, (1999), 'Analysis and Diagnosis of Dermatological Images with an Adaptive Associative Novel Neural Architecture', EANN99, Warsaw Sept 13-19.

ANDREWS S.G, ZIEN N F, (2000), 'Evolving Neural Architectures for Encoding and Recognition in Skin Cancer Lesion Images.' ICAI2000, Las Vegas.

AMARI S, (1977). ' Neural theory of association and concept formation'. Biological Cybernetics, 26: pp 175-185, 1977.

BRIDGETT N.A, BRANDT J, and HARRIS C.J. (1995): 'A Neurofuzzy Route to Breast Cancer Diagnosis'. 4th International Conference Artificial Neural Nets, 448-453.

BROOMHEAD D.S, LOWE D, (1988). ' Multivariant Functional interpolation and adaptive networks'. Complex Systems, 2, 321-355.

BURRONI M, DELL'EVA G, PUDDU P, ATZORI F, BONO R, FERRANTI G, MACCHINI, PALLOTTA S, (1997): ' Early diagnosis of melanoma using artificial neural networks.' Melanoma Research. Vol. 7, Supplement 1.

CICHOCKI A, KASPRZAK W, SKARBEK W, (1996) . ' Adaptive learning algorithm for principal component analysis with partial data'. In: Cybernetics and Systems '96 volume 2. Austrian Society or Cybernetic Studies. Vienna, Austria, 1996, pp 1014-1019.

CICHOCKI A, and UNBEHAUEN R, (1994). 'Neural networks for optimisation and signal processing'. John Wiley & Sons. New York, 1994.

GOLDBERG D.E. (1989): ' Genetic Algorithms in search, optimisation, and machine learning'. Addison-Wesley, 63.

GOUGH M.P. (1997): ' Associative List Memory '. Neural Networks, Vol. 10. No 6, 1117-1131.

FRIEDMAN R, RIGEL D, KOPF A, (1985). 'Early detection of malignant melanoma: The role the physician examination and self-examination of the skin'. CA-35 (3). 130-150
HARP S.A, SALMAD T. (1991), 'Genetic Synthesis of Neural Network Architecture', (in Handbook of

Genetic Algorithm L. Davis Ed.) Van Norstrand Reinhold N.Y., 202-222.

HOLLAND J.H. (1975): ‘ Adaptation in natural and artificial systems’. Ann Arbor: University of Michigan Press.

KENERVA P, (1984). 'Self-Propagating Search: A unified theory of Memory'. PhD Thesis, Stanford University, CA. USA.

JOLLIFFE I.T, (1986). 'Principal Components Analysis', New York, Springer-Verlag.

KOHONEN T, (1974). ‘An adaptive Associative Memory Principle’, IEEE Trans. on Computers C-23, 444-445.

LAWS, K I, (1979). ‘Texture Energy Measures’. DARPA Image Understanding Workshop (Los Angeles): 47-51.

MOALLEMI C. (1991): ‘Classifying Cells for Cancer Diagnosis using Neural Networks’. IEEE Expert, 8-12.

NILSSON N J, (1980). ‘Principals of Artificial Intelligence’. Palo Alto, Tioga.

OJA E, (1982), Principal components, minor components, and linear neural networks, Neural Networks, 5, 1992, 927-935

PANAGIOTIDIS N.G, KALOGERAS D, KOLLIAS S, D, STAFYLOPATIS A, (1996): ‘Neural Network-Assisted Effective Lossy Compression of Medical Images’. Proceedings of IEEE, Vol. 84. Vol. 10, 1474-1487.

REYES-ALDASORA C C, ALGORRI GUZMAN M, (2000). ‘ A Combined algorithm for image segmentation using neural networks and 3D surface reconstruction using dynamic meshes’. V IBERO-AMERICAN SYMPOSIUM ON PATTERN RECOGNITION, LISBON, Portugal, September 11-13.

ROSENBLATT F, (1958). ‘The Perceptron: A probabilistic model for information storage and organisation in the brain’. Psychological Review **65**, 386-408.

RUMLEHART D.E, HINTON G.E, WILLIAMS R.J, (1986). ‘ Learning internal representation by error propagation’. In Parallel Distributing Processing, 1 (D.E RumelHart, and J.L McCelland Eds.) MA, MIT Press.

SPECHT D.F, (1990). ‘ Probabilistic Neural Networks and Polynomial Adaline as Complementary Techniques for Classification’. IEEE Transaction on Neural Networks. Vol. 1, No. 1, 111-121.

SANGER T. D, (1989). ‘Optimal unsupervised learning in a single-layer feed-forward neural network’, Neural Networks, 2, 459-473.

STOECKER W.V. MOSS R.H, MADSEN R.W, LEE H, C, ERCAL F, and UMBAUGH S.E. , (June 1997).‘ Which Computer-detected features of pigmented lesions best predict a diagnosis of malignant melanoma.’ Melanoma Research. Vol. 7, Supplement 1.

STOECKER W.V. MOSS R.H, (1992). ‘Digital imaging in dermatology’. Compt Med imaging graphics, 145-150.

SWIERCZ M, LEWKO J, and MARIAK Z, (1995), ‘ Application of neural networks to the prediction of interacranial pressure’. Biocybern, Biomed. Eng., 15, 93-104.

SWIERCZ M, MARIAK Z, LEWKO J, CHOJNACKI K, and KOZLOWSKI PIEKARSKI P, (1998).‘Neural Network for detecting emergency states in neurological patients’. Medical & Biological Engineering & Computing, Incorporating CELLULAR ENGINEERING, Vol.36, No6. 717-722.

TOMATIS S, BARTOLI C, and BONO A, CASCINELLI N, CLEMENTE C, FARINA B, RADAELLI S, TRAGNI G, and MARCHESIN R, (1977). ' Advances in computer-diagnosis of Melanoma: the telespectrophotometric approach'. Melanoma Research, Vol. 7, Suppl. 1.

TUMER K, RAMANUJAM N, GOSHT J, RICHARDS-KORTUM R, (1998). ‘Ensembles of Radial Basis Function Networks for Spectroscopic Detection of Cervical Precancer’. IEEE Transaction on Biomedical Engineering, Vol. 45, No 8. 953-961.

WERBOS P.J, (1974). ' Beyond Regression: New Tools for Prediction and Analysis in Behavioural Science'. PhD Thesis. Harvard University.

Biography

Stuart Andrews has over 35 years computing expertise covering many different domains. This expertise was gained both in industry and academia. He has also run his own business specialising in VLSI design tools research. Over the last two years he was a Technology manager assisting SME's solving many different types of problems related to embedded systems. He is now an Independent consultant. His main interests cover real-time systems, DSP as applied to communications systems, imaging systems, intelligent systems as applied to medical procedures. Currently his main topic of research is intelligent recognition & categorisation techniques in the form of hybrid neural networks.