

# Landmark Detection using a Domain Independent Technique in Cephalograms

by  
S John<sup>1</sup>, V Ciesielski<sup>2</sup>, A Innes<sup>1</sup> and J Mamutil<sup>3</sup>

1. RMIT University, School of Aerospace, Mechanical & Manufacturing Engineering, Bundoora East Campus, Victoria. 3083, Australia.
2. 1. RMIT University, School Computer Science & IT, La Trobe Street, City Campus, Victoria. 3001, Australia.
3. Braces Pty Ltd, 404 Windsor Road, NSW 2153, Australia.

## Abstract

*The main purpose of this paper is about the relative merits of automated feature detection protocols on cephalometric-related digital radiology images. While the domain dependent techniques such as handcrafted masks are shown to be more accurate than pixel-based domain independent methods, the process for determining these shapes was unwieldy with respect to time taken and subjectivity. This paper discusses the use of a Pulse-Coupled Neural Network (PCNN) technique as a segmentation technique for parts of a radiology image. The results of this investigation showed a significant improvement in detection performance when compared to the handcrafted shapes when applied to several cephalometric landmarks. This encouraging outcome augurs well for a near-automated feature detection protocol for such medical-related radiology images.*

**Keywords:** Neural Networks, Segmentation, Cephalogram, Feature Detection, Landmarks.

## Introduction

Whilst there are many segmentation techniques used in image processing, a promising technique that was discussed by Innes (2003, 2006) for segmenting regions of interest in complex images is the Pulse-Coupled Neural Network (PCNN). The PCNN is a relatively new edge detection and segmentation method that has produced promising results in segmenting regions of interest in medical images such as in reports by Keller *et al* (1999) and Wolfer *et al* (1999). The difficulty of segmenting regions of interest in typical Cephalograms is that a number of areas located within the bony tissue of the head are subject to both noise and low contrast. While fairly accurate segmentation can be achieved for a few images using segmentation algorithms, a lot of algorithms tend to fail when applied to a larger suite of images unless the parameters are constantly adjusted. Because the ultimate goal of the problem in this study is to develop an automated methodology, a process whereby continually altering parameters in an ad hoc manner to achieve the best possible outcome is obviated.

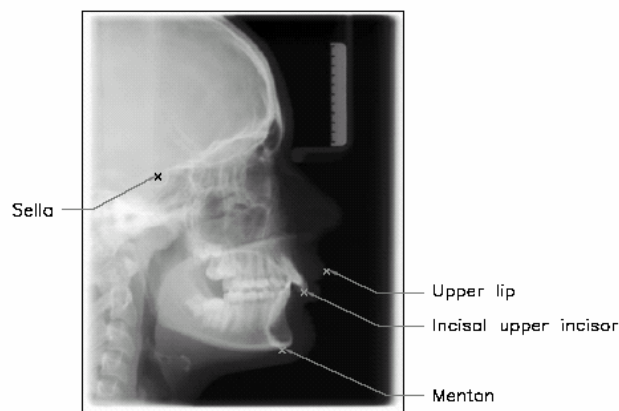


Figure 1 A Cephalogram showing the landmarks (marked X) used in this study

In a comprehensive cephalometric analysis, between 50 and 60 landmark points need to be located. In this paper, we address the 4 landmarks shown in Figure 1. We treat the finding of these landmark points as 4 independent object detection problems. The work presented in this paper is in the context of an existing object detection system in which a window of pixels of size *square\_size* is input to into a program which outputs the likelihood that the window is centred on the object of interest. This program is evaluated at each pixel position in the image and the position with the highest output is returned as the predicted position of the object. The program is obtained by a genetic programming process. The program uses a number of pixel level features, the means and standard deviations of a number of regions (we refer to them as `shapes') in the window. In previous work we have hand-crafted these shapes for each object of interest. Full details of the approach are given in Ciesielski *et al* (2003) and Innes (2006). In this paper, we describe an automated approach which uses pulse coupled neural networks. Examples of the kinds of shapes used by the object detection programs are the binary images in the far right column of Figure 2 and Table 2. The white pixels comprise one shape, the black pixels another. The means and standard deviations of these shapes would be the inputs to the object detection program.

### Pulse-Coupled Neural Network Segmentation

The aim of this section is to determine whether the outputs from the Pulse Coupled Neural Network (PCNN) are able to produce a segmented output that highlights regions of interest that will be useful for landmark detection. It is anticipated that segmenting the output will provide useful shapes for landmark detection programs by assisting with discriminating the landmark from background. The PCNN algorithm used in these experiments is based on the code from Lindblad and Kinser (1998). The PCNN will focus on highlighting regions of interest on both soft and bony tissue. The segmented regions will later be used to assist with locating the position of the landmark. To determine whether pre-processing by a PCNN could be useful in landmark detection four landmarks are selected. Two easy landmarks (the menton and upper lip landmarks), one of medium difficulty (incisal upper incisor landmark) and a hard landmark (sella landmark) are selected. The sella landmark is located in an area of bony tissue that is shown on the X-ray as subtle changes in greyscale. The other types of landmarks are located on the edge of bone/soft-tissue and soft-tissue. The different landmark types used in Innes (2006) that regularly achieved 100% detection performance have been omitted from this investigation. By using the PCNN it is anticipated that a set of parameters for each landmark will produce a binary output that have shapes relevant for landmark identification.

### Segmentation Results

The results shown in Figure 2 are the binarised output from the PCNN applied to examples of each type of landmark. A set of parameters were empirically determined for each type of landmark prior to segmentation. The parameter values remained constant throughout segmentation of the training data.

To determine the likelihood that the PCNN output could be used to assist with locating landmarks, the segmented outputs were manually classified into three categories, i.e. Definitive, Partially defined and Failure. The classes qualify the output and establish the validity of the PCNN parameters. Table 1 is a summary of segmentation results from the PCNN applied to the training set of 83 examples. The results indicate the PCNN method was able to accurately segment regions located on the edge of bone/soft-tissue or soft-tissue, however, the technique was less successful for highlighting the semi-circular region that encompasses the sella landmark. The reduced segmentation success rate for the sella landmark is because of the low contrast between the region of interest (semi-circular region) and background. The images shown in Figure 2 are a sample of results from the training data. We expect accurate detection of the Menton and Upper Lip points, moderately accurate detection of the Incisor and not particularly accurate detection of the Sella point.

	<b>Menton</b>	<b>Upper lip</b>	<b>Incisal</b>	<b>Sella</b>
<b>Definitive (%)</b>	<b>88.0</b>	<b>97.6</b>	<b>85.5</b>	<b>32.5</b>
<b>Partially defined (%)</b>	<b>12.0</b>	<b>2.4</b>	<b>12.0</b>	<b>33.7</b>
<b>Failure (%)</b>	<b>0</b>	<b>0</b>	<b>2.4</b>	<b>33.7</b>

Table 1: Summary of segmentation results for the four types of landmarks as shown in Figure 2.

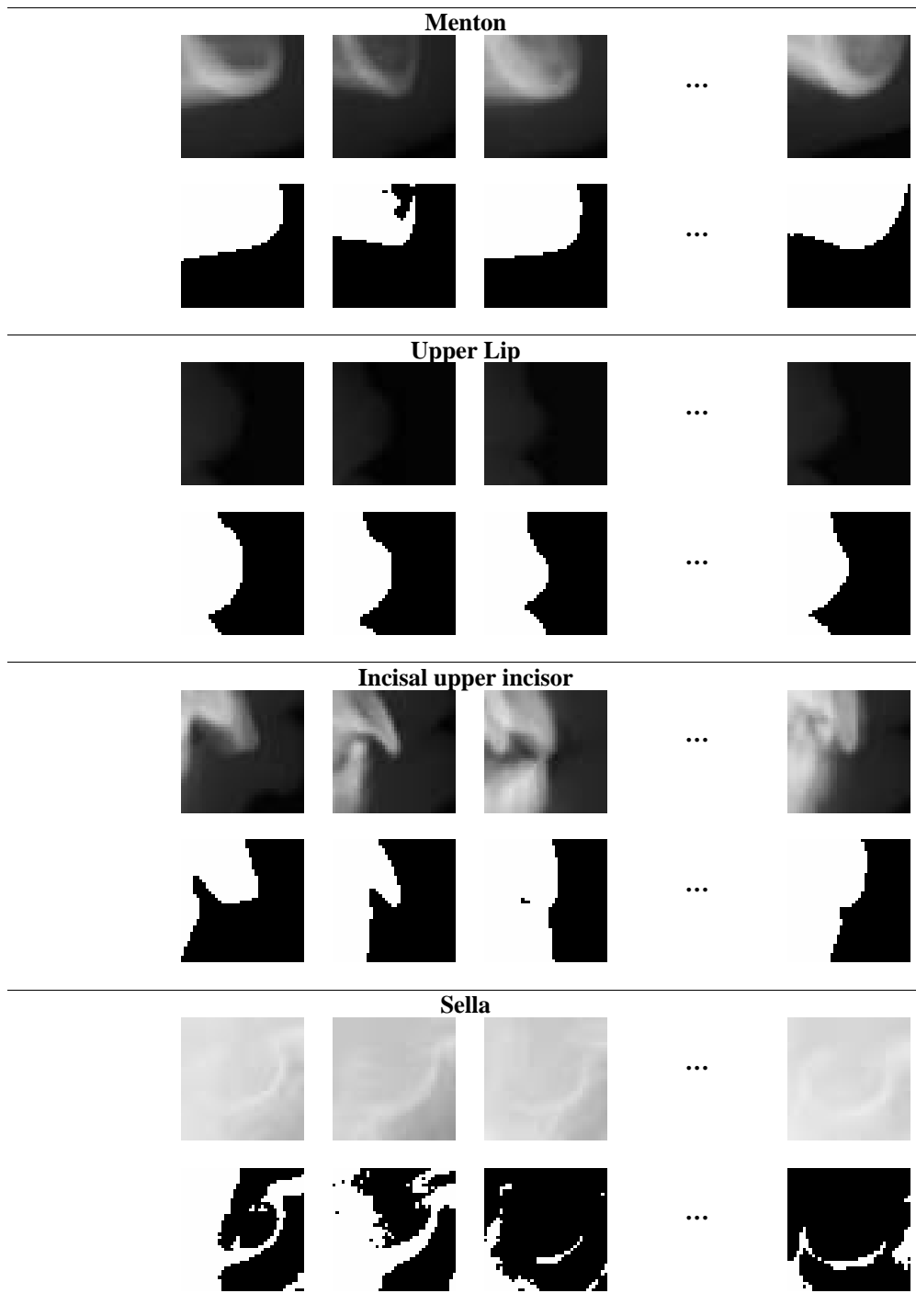


Figure 2: Segmentation results using the PCNN for the menton, upper lip, incisal upper incisor and sella cut-outs of size 40 x 40 pixels. Below each image cut-out is the corresponding output of the PCNN. The images have been scaled by 130% to enhance clarity.

**Algorithm for Determining the PCNN Derived Shapes**

In this section we describe our method for computing the shapes to be used by the detection program from the training examples. Because one segmentation was created from each cut-out, the method is limited to extracting two shapes from each cut-out, the black area and the white area, shown in Figure 2. The PCNN derived shapes are created as follows:

1. An image cut-out,  $I_k$ , of dimension *square\_size* (the same size as the input window to the detection program) is extracted from a training image. The image cut-out is centred on the known position of the landmark. The binary output,  $O_k$ , is created by applying the PCNN to the image cut-out.
2. Step one is iteratively applied to each image in the training set.
3. The binary output from the PCNN,  $O_k$  is used to create a template, by computing the average at each pixel position (i,j). The template is calculated using Equation 1.

$$Template(i, j) = \frac{1}{n} \sum_{k=1}^n O_k(i, j) \quad (1)$$

where  $O_k$  is the segmented image and n is the number of images in the training set. An example template is shown in the top row of Figure 3.

4. Two shapes are extracted by applying the thresholding rule below to each template. Shapes A and B corresponds to white and black pixels in the bottom row of Figure 3. The term given to describe these shapes will be PCNN derived shapes.

**If        Template(i, j) < threshold**  
                  **Shape A**  
**Else**  
                  **Shape B**

The procedure for segmenting a Template into two distinct shapes.

Grey pixels in the template, as shown in the top row of Figure 3 indicate that the PCNN has either not produced an ideal segmentation or there is biological variability that has been captured whilst averaging the binary outputs. Grey pixels are defined as pixels having grey level intensities between the interval [1, 254]. Because the PCNN method performs well at extracting soft and bony tissue, the grey pixels for the menton, upper lip, and incisal upper incisor landmarks are caused by biological variability. However, segmentation results for the sella cut-out were less successful and achieved an ideal segmentation of only 32.5%. This produces a lower contrast template compared to the other three landmarks. For example if a pixel in the template has a greyscale intensity of 127, then approximately half the images at pixel(i,j) were segmented black and an the other half white.

A value of 70 was chosen to threshold the template into two distinct regions, i.e. shape A and shape B. This value was empirically determined for the sella template and was subsequently applied for thresholding the templates for the other landmarks shown in Figure 3. If the process for segmenting regions of interest in areas that are subject to subtle changes of greyscale were improved, this may improve the PCNN derived shapes.

This section demonstrated that shapes could be automatically generated using a PCNN. The PCNN derived shapes have found regions of interest that we believe will be useful for locating landmarks.

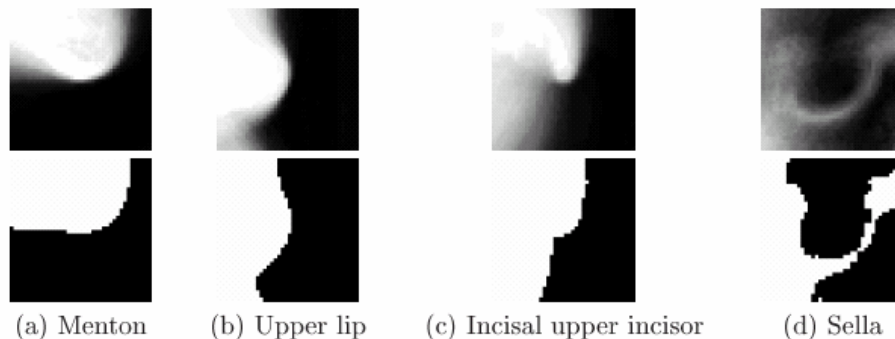


Figure 3 Templates computed using the outputs from a PCNN

## Learning from PCNN derived shapes using Genetic Programming

Previous work by Innes (2006) has used means and standard deviations of hand crafted shapes as the inputs to the object detection programs. However, constructing a set of handcrafted features is difficult and time consuming. Therefore, it is preferred to determine if the detection performance of programs using PCNN derived shapes are comparable to programs that used handcrafted shapes.

### Methodology Used

To do this, we will investigate four feature sets that use PCNN derived shapes. Each feature set is a progression from manually created handcrafted shapes to a method that can generate shapes automatically. The feature sets are defined in Table 2.

The first feature set uses handcrafted shapes as reported in Innes (2006) and is used as a benchmark for measuring detection performance for the following three feature sets. The second feature set substitutes the PCNN derived shapes for some of the handcrafted shapes from the previous feature set. This allows the determination of whether PCNN derived shapes can improve detection performance using the same number of terminals. The aim of the third feature set is to measure the detection performance for programs that use only PCNN derived shapes. The fourth feature set determines if additional square shapes combined with PCNN derived shapes can improve detection performance. The fourth feature set is automatically generated. The features are calculated from the mean and standard deviations of grey level pixel values for each shape. To determine the effectiveness of each feature set for evolving programs, four landmarks of varying detection difficulty have been selected. They are the menton, upper lip, incisal upper incisor and sella landmarks. A schematic of the methodology is shown in Figure 4 below.

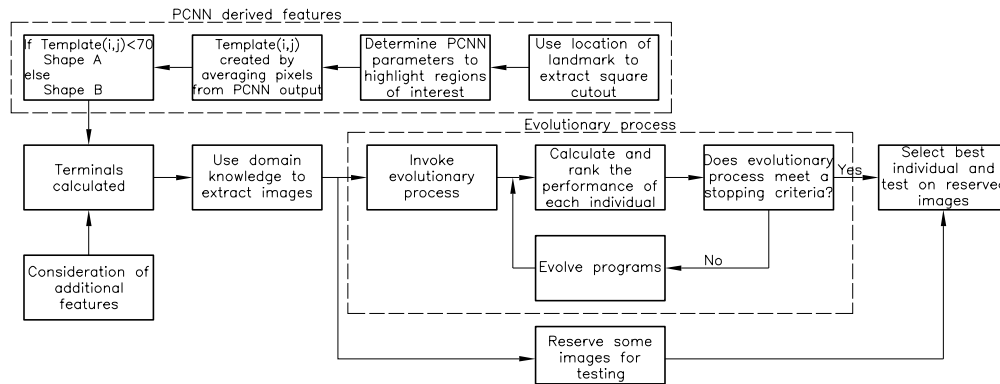


Figure 4: Diagram depicting an approach for extracting PCNN and additional features from the extracted images along with the methodology for evolving and evaluating detection programs for the task of locating a landmark.

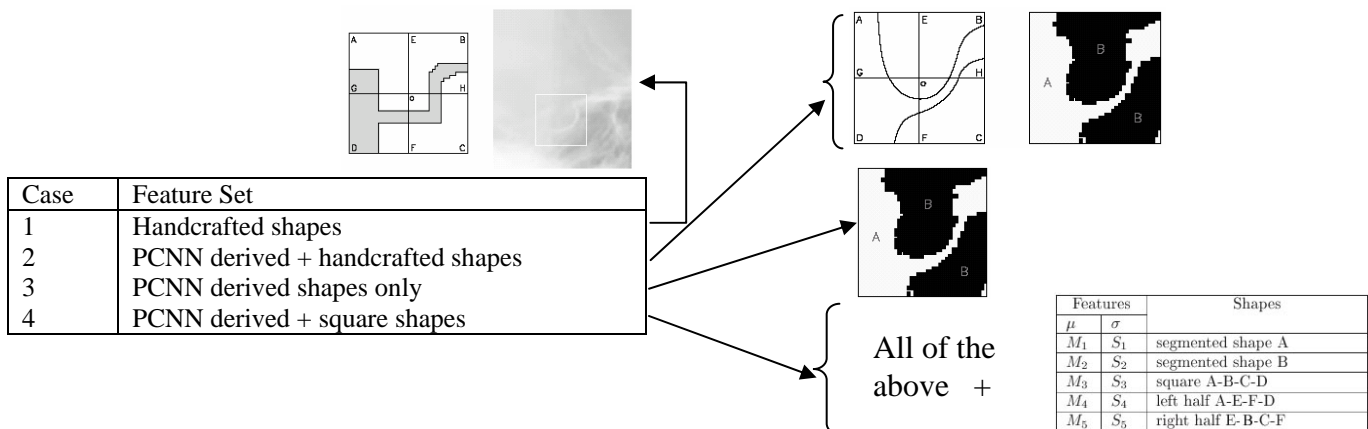


Table 2: Definition of four feature sets that are made available during the genetic search using the sella point as an illustrated example. M and S in the sub-tables represent the mean and standard deviation respectively.

### Feature Set Comparison

This section compares the detection performance of programs that were evolved using the different feature sets shown in Table 2 for comparing the four feature sets, a one-way ANOVA procedure described by Smith (1995) can be used to measure the differences in the mean detection rate. Let the null hypothesis be that the mean detection rate of programs that have been evolved from the four features will be the same, i.e.  $H_0: \mu_{\text{case1}} = \mu_{\text{case2}} = \mu_{\text{case3}} = \mu_{\text{case4}}$ .

The p-values for comparing the detection rate of programs when applied to training images are less than an alpha value of 0.05 for the upper lip, incisal upper incisor and sella landmarks. Therefore, the null hypothesis, that the mean detection rate of programs that were evolved from the four feature set sets is the same, is rejected. This indicates that at least two of the means are significantly different. Since the alternative hypothesis is supported, a Tukey's pairwise comparison (Smith (1995)) is conducted to determine discrepancies between the different feature sets. The results of the pairwise comparison are described extensively in Innes (2006).

A full comparison of all the 4 cases and the 4 landmarks used in this study is graphed in Figure 5. An analysis of the detection rates shown in Figure 5 indicates that on average, the feature sets using the PCNN derived shapes produced better performing programs compared to the feature sets using the handcrafted shapes with the exception of programs for the incisal upper incisor. The results for the PCNN derived shapes only when applied to the incisal upper incisor indicates that other shapes were required to discriminate the object from background when comparing the results from the other feature sets.

A further analysis of the pairwise comparison revealed that there was no significant difference between programs that were evolved from features using PCNN and handcrafted shapes or PCNN and square shapes. The results of programs using the PCNN and square shapes produced detection programs that were either equivalent to or exceeded the detection performance of the handcrafted shapes used by Innes (2006). This is a desirable result since the method to create the PCNN and square shapes is a process where only a small amount of prior knowledge is required about a landmark.

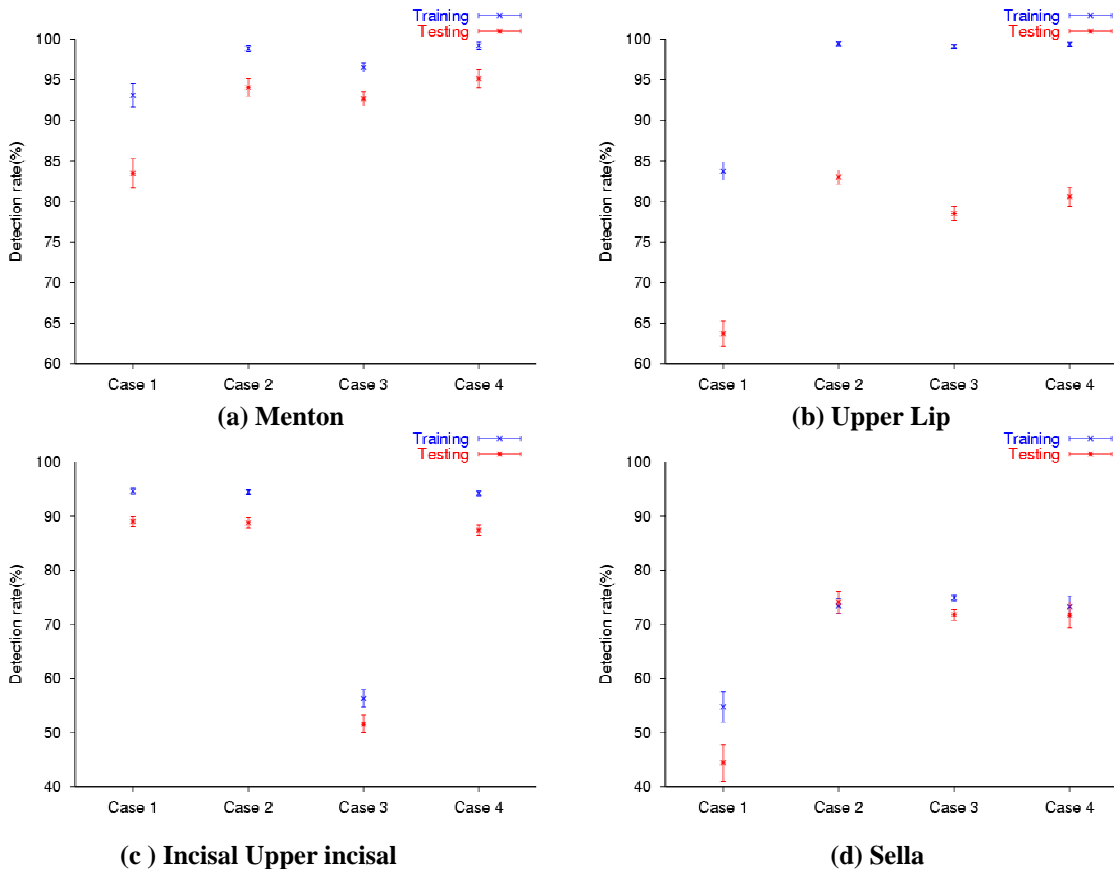


Figure 5: A summary of programs' detection performance that are evolved using variations of handcrafted, PCNN derived and square shapes. The error bars represent 95% confidence intervals.

## Conclusion

The main aim of the work presented in this paper was to determine whether an existing manual procedure for determining a partitioning of an input window into shapes for an object detection system could be improved by an automated procedure. The automated procedure involved segmentation with a pulse coupled neural network, computation of an average template and then thresholding the template to obtain the shapes. An investigation of the detection programs that used the PCNN derived shapes demonstrated a significant improvement in detection performance when compared to the handcrafted shapes when tested on the menton, upper lip and sella landmarks. There was no improvement for the incisal upper incisor landmark. Overall, the results suggest that the automated shape generation procedure will give detection performance that is not worse than the hand crafting procedure.

Ideally, an approach for landmark detection which is entirely automated is desired with respect to the creation of useful features to be used within a program for the identification of landmarks within an image. Even for a domain expert, it is not a trivial task to construct shapes for accurate object detection. This is a compelling argument for the utility of an automated approach to the problem. The study presented here goes some way to achieving such automation.

The threshold used to create the PCNN shapes was constant for all landmarks and determined by the sella landmark. Because segmentation of the sella region is a difficult problem, this may have affected the thresholding of the menton and incisal upper incisor. Choosing a larger threshold for the incisal upper incisor may have produced a shape more indicative of the landmarks shape and improved object detection accuracy.

## References

1. A Innes, V Ciesielski, J Mamutil, S John (2003) "Finding Templates for Cephalometric landmarks using Pulse-coupled Neural Networks and genetic Algorithms", Proceedings of the International Conference on Image Science, Systems and Technology (CISST'03), Las Vegas, Nevada, USA, CSREA press Volume 2, pg 456-462, ISBN 1 892512 47 5.
2. A Innes (2006), 'Feature Detection using Genetic Programming in Cephalometrics'. Ph.D thesis (In preparation), School of Aerospace, Mechanical & Manufacturing Engineering, RMIT University, Melbourne, Australia.
3. Vic Ciesielski, Andrew Innes, John Mamutil, and Sabu John (2003). Landmark detection for cephalometric radiology images using genetic programming. *International Journal of Knowledge Based Intelligent Engineering Systems*, 7(3):164-171, ISSN: 1327-2314.
4. P E Keller & D McKinnon (1999), 'Pulse-coupled Neural Networks for Medical Image Analysis' Proceedings of SPIE – Applications of Science and Computational Intelligence II, Editors – K L Priddy, P E Keller, D B Fogel and J C Bezdek, Volume 3722, pages 444-451, 1999.
5. J Wolfer, SH Lee, J Sandelski, R Summerscales, J Soble, J Roberge (1999), ' Endocardial Border Detection in Contrast enhanced echocardiographic cineloops using a Pulse-coupled Neural Network', *Computers in Cardiology*, 1999, pages 185-188, 26-29 September 1999.
6. T Lindblad and K Kinser (1998), 'Image processing using Pulse-coupled Neural Networks'. Springer-Verlag, London.
7. P Smith (1995), 'Applied Linear Models-Part II, RMIT University Lecture notes in Statistics and Operations Research, Preliminary version edition.