Middle-East Journal of Scientific Research 24 (S2): 169-175, 2016
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2016.24.S2.141

Feature Extraction and Classification for Segmentation of Overlapping Cervical Cells by Multiple Level Set Functions Optimization

P. Shanmugapriya and N. Priscilla Vilma Manorathi

Department of ECE, PG Scholar, Dept of ECE, Saranathan College of Engineering, Trichy, India

Abstract: Feature Extraction and classification is an important task in medical applications. In medical examination of specimens the presence of blood reduces the accuracy and it leads to false results. In this work an algorithm for extraction of features is given by moment invariants. Using Legendre moment invariants seven moments are calculated and normalization is done using the central moments. The classification of the specimen to identify whether it is normal or abnormal is done with the help of Artificial Neural Networks. Pap smear test is usually done in the lower part of the cervix of humans.To find any unusual developments the specimen is viewed under the microscope to find any abnormal changes which leads to cancer. Proper examining of cell images reduces the deaths of cervical cancer. Feature extraction is by using Legendre Moment Invariants where moment's upto seven are calculated. With the help of Artifical Neural Networks classification is done. Thetype of learning used is Supervised Learning.Segmentation of overlapping cells is done with the help of level set functions taking into account the unary and binary constraints. Instead of segmenting the cell as a whole, individual cells within the clumps are identified and segmented. In this manner the specimen examination and classification in medical diagnosis for the treatment of diseases is improved.

Key words: Cell segmentation · Neural networks · Moment invariants · Feature extraction · Classification

INTRODUCTION

To analyze any microscopic images segmentation is essential and it remains one of the problems which are facing challenges.In the cervical regions proper detection of cells is essential.Complex nature of cell structures due to tightly packed cells in cervix regions, results in difficulty in segmenting cells, in these regions and it is one of the problems faced in medical analyzes.The specimen must be clearly identified so that any abnormal changes are detected in early stages [1].

The proper classification of images is most needed factor for screening and cancer diagnosis to provide accurate results of cervical cancer. By the process of segmentation a cytologist is able to get more information of any diseased part.This becomes one of the major advantage to improve the specimen classification. In this manner the better treatment can be provided thereby reducing thedeaths of victims. By proper examination of cell structures any abnormalities can be easily detected and the number of deaths due to cervical cancer can be reduced [1].

A method based on graph cuts in [2], which deals with the segmentation of overlapping nuclei and the boundaries of cell clumps are focused. This method delineate the boundary of a clump of cells and individual contours of overlapping nuclei of normal and abnormal cervical cells images. Instead of providing accurate boundaries for each overlapping cell, the method generates a contour of whole clump of overlapping cells. It segments only entire clumps instead of individual cells which is an issue in limiting the information for classification of a cytologic specimen.

A technique is proposed in [3] in which a sliding band filter is used to segment overlapping cell nuclei and cytoplasm on Drosophila MelanogasterKC167 dataset. This technique aims at the detection of cells based on intensity thresholding method and estimation of shapes on multivibrate medical images. This filter designed using this approach is able to detect overall convex shapes and performs well for cell detection. Due to high variability of cell shapes and frequent cluster overlap between the cells are the difficulties faced in this method. Cervical cancer is tumor which is caused bygrowth of cervical tissue

Corresponding Author: P. Shanmugapriya, Department of ECE, PG Scholar, Dept of ECE, Saranathan College of Engineering, Trichy, India.

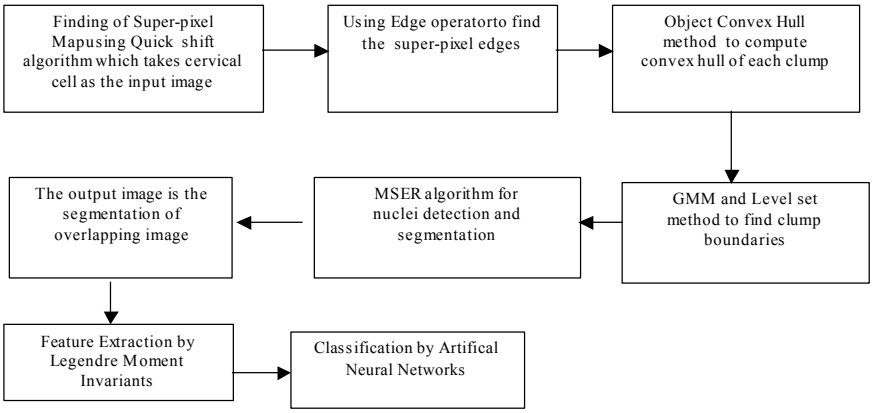

Fig. 1: Illustration of Feature Extraction and Classification of overlapping cell segmentation

cellsand leads to abnormal changes which leads to division of cells and leads to death of cells. If the tumor is malignant, its cell flow through the blood stream and other regions which are normal also gets affected. These infected cells are then distinguished as normal and abnormal using classification by Artifical neural networks.

Proposed Methodology: The proposed methodology deals with steps such as, segmentation of cell which is termed as segmentation of scene objects, after the scene method level set is used, followed by Feature Extraction and classification given in Fig 1.

Initial Scene Segmentation: The method to segment the overlapping cells deals with the following stages: first stage consists of the scene segmentation which is used to segment the cell clumps. Region of interest and background are the two regions which are involved. Particular region of interest are extracted by setting threshold value. By this process the region of interest are separated from background. The methods involved in this segmentation are summarized below.

Quick-Shift Algorithm: The input to this method is the cervical cell. This method finds the parzen density estimate by using the equation,

$$p(x) = \frac{1}{N}\sum_{i=1}^{N} k(x - x_i)\ , x \in R^d \quad (1)$$

where $k(x)$ is Gaussian window an N is the data points x_1, ... ,$x_N \in R^d$ around each point. In this algorithm each point get connected to the point which is nearest and thereby forming a link which looks like a tree. By setting the threshold segmentation is obtained which involves in breaking the links under the condition dx > threshold (τ). Super-pixel map is obtained by this method.

Canny edge Detection: Using super-pixel map as input from the previous step the super-pixel edges which are prominent are obtained by canny edge detection. The smoothed image is obtained by convolving the input image with a Gaussian function. The intensity gradients in horizontal and vertical directions are obtained. The prominent edges are obtained by double threshold method [3].

Object Convex Hull: By the method of the connected components analysis convex hull image is obtained. Object convex hull method is preferred. This method takes as input the super-pixel edge map obtained using the canny edge detector. The hull image is returned as output of each connected component [4].

GMM AND LEVEL SET for Computation of clump boundaries: Using GMM (Gaussian Mixture Model) the boundaries of clumps are obtained. Level set functions like kernel function, signed distance function, Heaviside step function are involved in computation of clump boundaries.

Maximally Stable Extremal Regions Algorithm: The nuclei regions are segmented and blobs are detected by making use of Maximally Stable Extremal Regions algorithm, which uses as input the previously obtained cell clumps. The thresholds are set which are used to find blobs [5] by this method.

Level Set Method: The initial segmentation is used for further processing which uses level set method [6]. Thus, overlapping cell segmentation is performed with the initial scene segmentation followed by the detailed segmentation which involves several level set functions [7]. The energy is represented by the following equation and is given by,

$$\varepsilon\left(\{\emptyset_i\}_{i=1}^{|N|}\right) = \sum_{i=1}^{N} \varepsilon_{u(\emptyset_i)} + \sum_{i=1}^{N} \sum_{j \in N(i)} \varepsilon_{b(\emptyset_i \emptyset_j)} \quad (2)$$

The energy calculation by Mean Curvature is given below,

$$E_{image}(z, \emptyset) = \sqrt{\left\{\left(\log_{10} \frac{p_{in}(z,\emptyset)}{p_{out}(z,\emptyset)}\right)^2\right\} - \left\{\left(\log_{10} \frac{p_{in}(z,\emptyset)}{p_{out}(z,\emptyset)}\right)\right\}^2} \quad (3)$$

where ε representsthemeanfunction.
The Gradient Flowis given as specified below,

$$\nabla_{\emptyset} E_{image} = -\frac{\varepsilon_e(\emptyset)}{E_{image}} . [\{B.G\} - \{B\}.\{G\}] \quad (4)$$

B and G is given by,

$$B = \log_{10} \frac{p_{in}(z,\emptyset)}{p_{out}(z,\emptyset)} \quad (5)$$

$$G = \left[\left(\frac{1}{A_{in}} + \frac{1}{A_{out}}\right) - K(z - l(x))\left(\frac{1}{A_{in} p_{in}(z,\emptyset)} + \frac{1}{A_{out} p_{out}(z,\emptyset)}\right)\right] \quad (6)$$

Here A_{in} is given by, $\int_{\Omega} H_{\varepsilon}(\emptyset)\, dx$ the PDE describes the evolution of the curve which optimally maximizes the "distance" between the distributions which correspond to the exterior and interior of segmenting curve. $H_{\epsilon}(\phi)$ represents the Heaviside step function.

Feature Extraction: From completed cell segmentation as input features are extracted by making use of moment invariants technique. There are many types of moment invariants like Legendre, Geometric, Zernike and Complex moment Invariants. Among which Legendre moment Invariants is choosen as its performance is better when compared with others. It is used in applications like pattern recognition. The normalization is done by making use of complex and geometric moment invariants [8].

For two dimensional image of size M * M the moments are given by,

$$m_{pq} = \sum_{x=0}^{x=M-1} \sum_{y=0}^{M-1} (x)^p . (y)^q f(x,y)$$
$$p, q = 0, 1, 2 \ldots \ldots \quad (7)$$

By translating an amount the moments are given by,

$$\mu_{pq} = \sum_{x} \sum_{y} (x + a)^p . (y + b)^q . f(x,y) \quad (8)$$

By substituting $x = \bar{x}$ and $y = \overline{y}$ the central moments are given by,

$$\overline{x} = \frac{m_{10}}{m_{00}} \text{ and } \overline{y} = \frac{m_{01}}{m_{00}} \quad (9)$$

$$\mu_{pq} = \sum_x \sum_y (x - \overline{x})^p . (y - \overline{y})^q . f(x,y) \quad (10)$$

Scaling normalization when applied the equation for central moments is given by,

$$_{pq} = {}^{\mu_{pq}}\!/_{\mu^{\gamma}_{pq}} \quad \text{and}$$

$$\gamma = |(p + q)/2| + 1 \quad (11)$$

Seven moments are obtained as order three which are invariants to scale, position and orientation is given by set of equations namely,

$$M_1 = (_{20} - _{02}) \quad (12)$$

$$M_2 = (_{20} - _{02}) + 4^2_{11} \quad (13)$$

$$M_3 = (_{20} - 3_{12})^2 + 3(_{21} + _{03})^2 \quad (14)$$

$$M_4 = (_{20} + _{12})^2 + (_{21} + _{03})^2 \quad (15)$$

$$M_5 = (_{20} - 3_{12})(_{30} + _{12}) |(_{30} + _{12})^2 - 3(_{21} + _{03})^2| + (3_{21} - _{03})(_{21} + _{03}) |3(_{30} + _{12})^2 - (_{21} + _{03})^2| \quad (16)$$

$$M_6 = (_{20} - _{02})|(_{30} + _{12})^2 - (_{30} + _{12})^2| + 4\,_{21}(_{30} - _{12})(_{21} - _{03}) \quad (17)$$

$$M_7 = (3_{21} - _{03})(_{30} - _{12})|(_{30} + _{12})^2 - 3(_{21} + _{03})^2| - (_{30} - 3_{12})(_{21} + _{03})|3(_{30} + _{12})^2 - (_{21} + _{03})^2| \quad (18)$$

Calculation of moment invariants of an image, the output is a feature vector, a column vector containing M_1 M_2--------$M_7$and this is the feature vector which is obtained as the column vector and is given by,

M = [M_1 $M_2$$M_3$ M_4 $M_5$$M_6$ M_7]

Classification by Neural Networks: The complete cell segmentation which is obtained is fed as input to the feature extraction where moments are calculated and centralized moments are computed and seventh invariants moments are calculated which results seven features

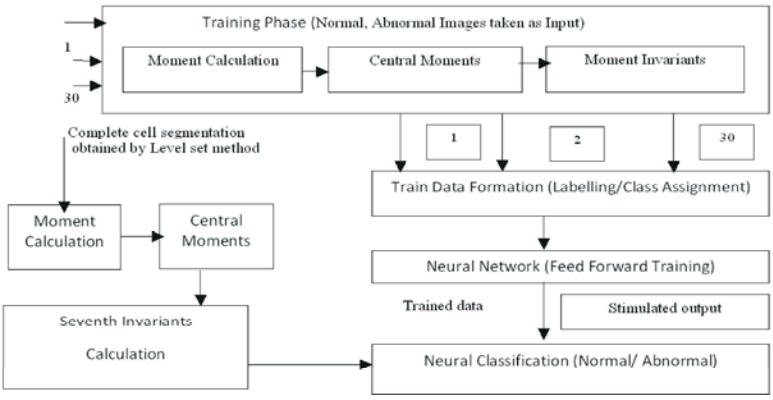

Fig. 2: Illustration of Feature Extraction and Classification

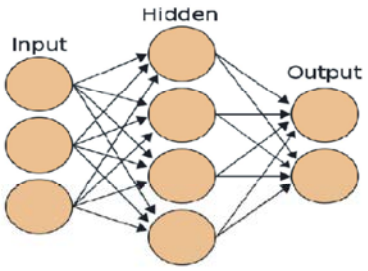

Fig. 3: Neural Network Classifier

vectors are obtained. These feature vectors are fed as input to the neural classifier is given in Fig 2. In the training phase thirty images are taken, ten of which are abnormal and normal. For each of the image taken the moment calculation is done using Legendre moment invariants. From the moments obtained the centralized moments are calculated. From the centralized moments the seventh invariants moments are calculated. Seven features are obtained from each images.These features are fed as input to the train data formation where labeling and class assignment is done to classify the images as normal and abnormal class [9].The resulting output is fed as to the neural network where multilayer feed forward network is used. This network consists of three layers namely input layer, hidden layer and output layer given in Fig. 3. The learning method used is supervised learning where each input is matched to respective target or output pattern. The input to the feed forward network is data and the output is stimulated output, which represents the hidden layer of the neural classifier.In the feed forward network the network structure is created, the data is assigned according to the network and stimulated output is obtained. The method used is Levenberg method. Gradient method is not used as it deals with data analysis. Here features are analyzed so Levenberg method is used.

The inputs are fed to the classifier, one input from the features calculated from moment invariants and the another input is from the stimulated output coming from feed forward network. These are fed to the classifier, where the input data will be checked with the hidden layer data getting the output layer data. The resulting outputs from classifier will be either normal or abnormal.

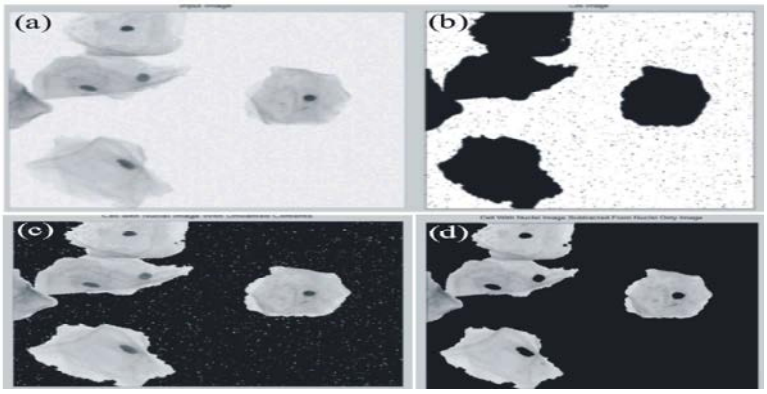

Fig. 4: (a) Input cell image (b) cell image (c) Cell image with unwanted contents (d) Cell image subtracted from nuclei only image.

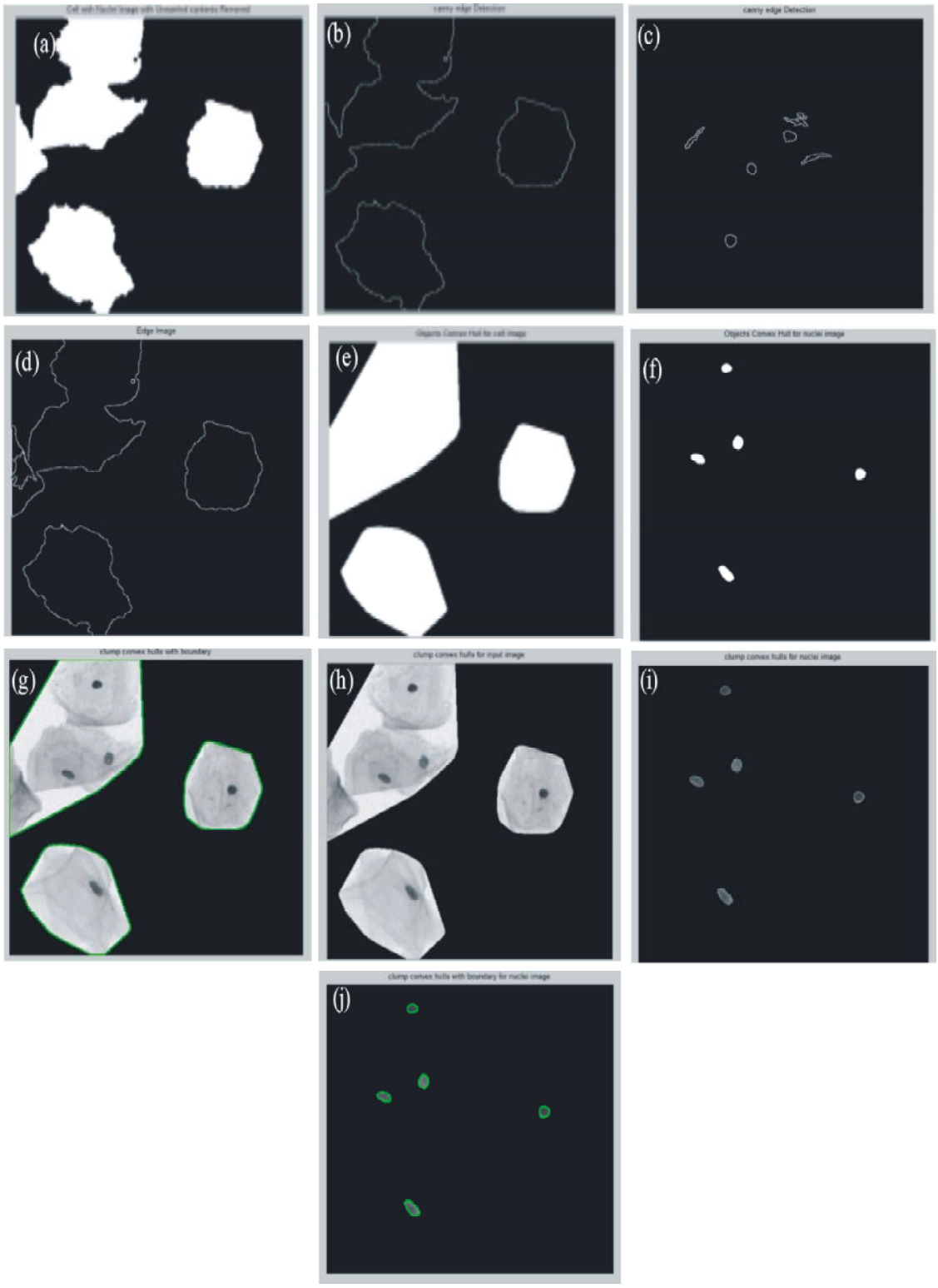

Fig. 5: (a) Cell with nuclei with unwanted contents removed (b) Cell image (c) nuclei image (d) cell image (e) Convex hull image (f) Convex hull nuclei image (g) Object convex hull image (h) Convex hull boundary image (i) Convex hull nuclei image (j) Convex hull boundary for nuclei image.

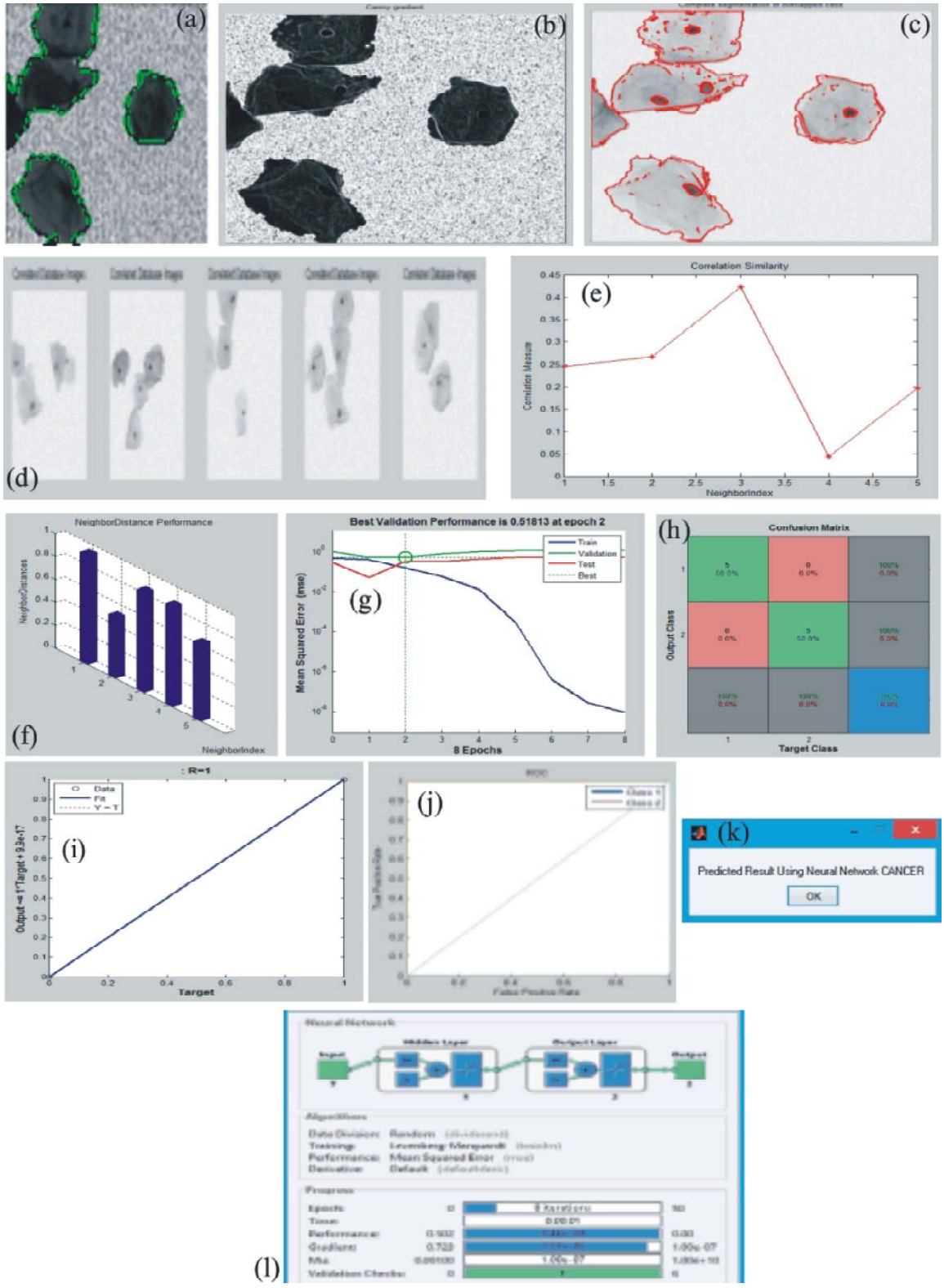

Fig. 6: (a) Clump boundaries (b) Canny gradient (c) segmented output (d) Matched convergence images (e) Correlation boundary (f) Neighbour distance performance (g) Validation performance (h) Confusion matrix (i) Target performance (j) False/True positive rate (k) Predicted Result (l) Neural network.

RESULTS AND DISCUSSION

The cell segmentation is performed using cervical image of dimension 512x512 given in Fig. 4. In the training phase ten images are collected which are normal and abnormal. The image types are EDF and synthetic.Using object convex hull method hull image is obtained given in Fig. 5. Levenberg method is preferred in neural networks as it results in stimulated output. The segmented output is used for, further processing like feature extraction.Classification which results in distinguishing the cervical image whether it is normal, abnormalgiven in Fig. 6.

CONCLUSION

The clumps of cervical cell is segmented and features are extracted and classified using neural network classifier.Using Legendre moment invariants seven invariants moments are computed. Classification whether it is normal or abnormal is performed by using Neural network classifier which uses multilayer feed forward network The learning method used is supervised learning.The performance metric parameters based on Mean Squared Error (MSE) is 0.51813.

REFERENCES

1. Plissiti, M.E., C. Nikou and A. Charchanti, 2011. Automated detection of cell nuclei in Pap smear images using morphological reconstruction and clustering, IEEE Trans. Inf. Technol. Biomed., 15(2): 233-241.
2. Zhang, L., H. Kong, C. Chin, S. Liu, Z. Chen, T. Wang and S. Chen, 2014. Segmentation of cytoplasm and nuclei of abnormal cells in cervical cytology using global and local graph cuts. Computerized Medical Imaging and Graphics, 38: 369-380.
3. Shrivakshan, G.T., and Dr C. Chandrasekar, 2012. A Comparison of various Edge Detection Techniques used in Image Processing, IJCSI International Journal of Computer Science, 9(5), No 1, September 2012.
4. Dillencourt Michael, B., Hannan Samet and Markku Tamminen, 1992. A general approach to connected-component labeling for arbitrary image representations. Journal of the ACM (JACM) 39(2): 253-280.
5. Matas, J., O. Chum, M. Urban and T. Pajdla, 2002. Robust wide baseline stereo from maximally stable extremal regions, in Proc. BMVC, Cardiff, U.K., pp: 384-396.
6. Li C. Xu, C. Gui and M.D. Fox, 2010. Distance regularized level set evolution and its application to image segmentation, IEEE Trans. Image Process., 19(12): 3243-3254.
7. Gao, X., B. Wang, D. Tao and X. Li, 2011. A relay level set method for automatic image segmentation, IEEE Trans. Syst., Man, Cybern. B, Cybern., 41(2): 518-525.
8. Vaidya, Y.M., B.S.V. Swati and Narmada Kantipudi, 2015. Moment Invariants Based Feature Techniques for Segmentation of Retinal images using Supervised Method, 2015 International Conference on Industrial Instrumentation and Control (ICIC) May 28-30,2015
9. Abid Sarwar, Mehbob Ali, Jyotsna Suri, Vinod Sharma, 2015. Performance Evaluation of Machine Learning Techniques for Screening of Cervical Cancer", 2015 IEEE.