Automatic Detection for Healthy and Unhealthy Kidneys on Abdominal CT Images using Machine Learning Algorithm

Israt Jahan Tulin
Lecturer of CSE department
BGMEA University of Fashion and Technology
Uttara, Dhaka-1230

ABSTRACT
In this paper, we have proposed a machine learning (Support Vector Machine) approach for detecting healthy and unhealthy kidneys in CT (Computed Tomography) images. At first, kidney region have been segmented from the abdomen area using region growing algorithm. After successful segmentation, the kidney region is extracted and it is given to Support Vector Machine algorithm for the final detection of which kidney is healthy and unhealthy. Our proposed approach consists of training process and testing process. In training process we train our algorithm with the CT images of healthy kidney and unhealthy kidney. In testing process our algorithm detect healthy and unhealthy kidneys from the input images with an accuracy of 73.3%. The proposed algorithm has been implemented in MATLAB and experiment result tested on 70 images downloaded from internet.

General Terms
Terms are kidney segmentation, detection, region growing method and machine learning algorithm.

Keywords
Keywords are CT images, kidney, segmentation, detection and algorithm.

1. INTRODUCTION
Computed Tomography (CT) of an abdomen is an imaging method that uses x-rays and computer technology to produce images of the belly area. For diagnosis of abdominal organs pathologies, different imaging techniques such as Magnetic Resonance Imaging (MRI), Positron Emission Tomography, Computed Tomography (CT) and ultrasound scan are a standard instrument in medical field. Due to good spatial resolution and high Signal-to-Noise ratio Computed Tomography (CT) images are chosen to segment kidney region and detect healthy kidney images.

Image segmentation is the process of partitioning a digital image into multiple segments and the goal is to simplify or change the representation of an image into something that is more meaningful and easier to analyze.

Worldwide research indicates that one out of 10 adults had kidney problems and by 2015 it is estimated that about 36 premature deaths due to kidney disease will happen [1]. Kidney function is very important and impairment can be life threatening. So it is necessary to diagnosis the kidney disorders and diseases in the early stages. Abdominal CT images widely used for the diagnoses kidney diseases.

Hence our main objective is to segment kidney region from CT (Computed Tomography) images and detect the healthy and unhealthy kidneys from the segmented kidney regions by applying machine learning algorithm.

2. LITERATURE REVIEW
Till now, many researches have been focused for kidney segmentation. Pohle and Toennies developed a region-growing algorithm that automatically learns its homogeneity criterion from the characteristics of the region that is segmented [2-4]. This approach is less sensitive to the seed point location. Based on the pixel value distribution of the organ and the mesh operation [5], [6], Kim et al. and Yoo et al. proposed similar approaches for kidney segmentation. To identify and extract organs from normal CT images [7], Kobashi and Shapiro explained a knowledge-based procedure. The detection result was rated 85% grade A from testing of 75 images from three patients. Through tissue texture analysis and distribution of directional maximums [8], Mavromatis et al. performed medical image segmentation. Wang et al. proposed a constrained optimization approach. In this process deformable contour can be computed as extra constraints within the contour energy minimization frame work [9]. Tsagaan and Shimizu described a deformable model approach for automatic kidney segmentation [10], [11]. They used a deformable model represented by the grey level appearance of kidney and its statistical information of the shape. They tested about 33 abdominal CT images. The degree of correspondence between automatic segmentation and manual positioning was 86.9%.

3. KIDNEY SEGMENTATION
An abdominal CT image is very complicated. It contains kidney, liver, spleen, spine, fat, and pathologies. To segment the kidney region from the abdominal CT images, at first we have applied an abdomen contour detection algorithm to delineate the abdominal cavity. After successfully determining the abdominal boundary, we have segmented kidney region by applying region growing segmentation algorithm.

3.1 Region Growing Process
Region growing algorithm is a pixel-based image segmentation method. It involves the selection of initial seed points. This algorithm works with a set of seed points. Each seed points correspond to an individual region. The seed points compare to their neighboring pixels based on a similarity criterion. These neighboring pixels are then computed either by 4-connectivity or 8-connectivity. The neighboring pixels in 4-connectivity are connected horizontally and vertically, whether the neighboring pixels in 8-connectivity are connected horizontally, vertically, and diagonally. The simplest similarity criteria are used generally. It can be obtained by calculating the difference between the

intensity value of the image pixel and the corresponding region mean. If the difference is less than a specified threshold, then the pixel belongs to that region and is subsequently labeled. Otherwise the pixel is not labeled and skipped.

3.2 Region Growing Algorithm

Step 1: Every seed point corresponds to a single region and by applying the seed selection algorithm, these seed points are computed.

Step 2: For a given seed point,

- Seed point is assigned as the region label.
- Initialize the region mean and it is equal to the pixel intensity at the seed point.
- Neighbours of the seed point are computed and stored in the neighbour matrix. This matrix stores the neighboring pixels addresses, to be checked.

Step 3: For every pixel stored in the neighbour matrix,

- The pixel which is not labeled and fulfill the similar criteria,
  - I. The pixel in the corresponding region is labeled.
  - II. The new region mean of the corresponding region is computed.
  - III. The neighboring of pixels are computed and stored in the neighbour matrix.
- Otherwise, skip the pixel and choose the next neighboring pixel.
- Repeat step 3, until all the listed neighboring pixels are not checked.
- Leave the region pixels unlabelled, if the region size is very small.

Step 4: The next seed point is chosen and repeat step 2 for next region.

Step 5: If all the pixels are not labeled, the seed points are computed again using only the unlabelled pixels.

Step 6: Repeat the process until all pixels of the image are labeled into their corresponding regions. In this way, the image is divided into various segmented regions.

3.3 Region Modification

After applying the region growing method to the abdomen area we have obtained different gray scale images of kidney region. But it may remain some trivial and irregular objects or holes that may scatter inside the region. So at first we have applied ROI (Region Growing Interest) process to find a portion of an image that we want to filter. After that we have applied a series of image processing skills such as pixel filling, erosion, lebeling and dialation to improve the accuracy of the segmented area under various conditions arising from clinical practices such as the timing and rate of contrast media injection, image intensity variations, and blood flow rate etc.

3.4 Experiment Result

3.4.1 Kidney Segmentation

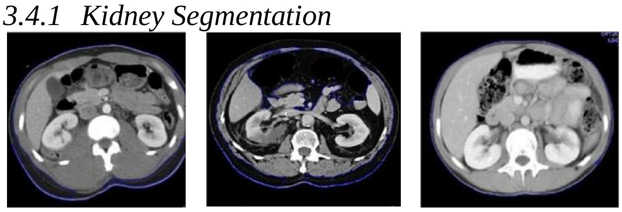

Fig.1: Detection of abdominal boundary

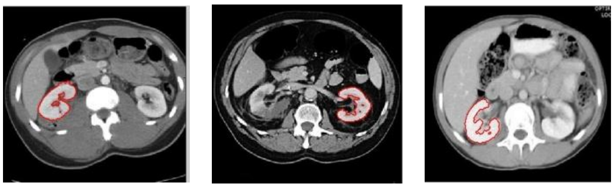

Fig.2: Kidney extraction

Fig.3: Kidney Segmentation

3.4.2 Region Modification

Fig.4: Binary Images

Fig.5: Gray Scale Images

4. HOG AND SVM

4.1 HOG

Histogram of Oriented Gradient (HOG) is a feature descriptor. It is a representation of an image or an image patch that simplifies the image by extracting useful information and throwing away extraneous information. This procedure counts occurrences of gradient orientation in localized portions of an image. Gradients of an image are very useful because the magnitude of gradients is large around the edges and corners. These edges and corners contain lot more information about object shape than flat regions. We use HOG descriptor of cell size 4 × 4 to encode the patterns (HOG feature vectors). After that the HOG feature vectors are fed to SVM.

4.2 SVM

In our work, Support Vector Machine (SVM) is used to perform classification [12], [13]. SVM is trained for two sets of models: one for Healthy Kidney CT images and another for Unhealthy Kidney CT images. The main idea behind training of SVMs is to find the separating hyperplane optimally so that the classification error is minimized for the given test samples. Assume a set of M training samples of two separable classes are represented by (x_1,y_1), (x_2,y_2), …, (x_M,y_M); where $x \in R^N$ is an N-dimensional space and class label is denoted by y, $(y_i \in \{-1,+1\})$. A SVM attain the optimal hyperplane which linearly classifies (separates) the larger portion of the training data points while maximizing the distance from the hyperplane. Twice of this distance is called the margin. The hyperplane discriminant function is described by the following equation:

$$f(x) = \sum_{i=1}^{M} yi, ai. k(x, xi) + b$$

Where, the membership of x is determined by the sign of f(x) and kernel function is denoted by k(·,·). Finding all the nonzero ai is the equivalent of constructing an optimal hyperplane. A vector xi is said to be supported vector (SV) of the hyperplane, if it comply to a nonzero ai. SVMs provide a compact classifier as the number of training data points which are maintained as the support vectors is generally very small.

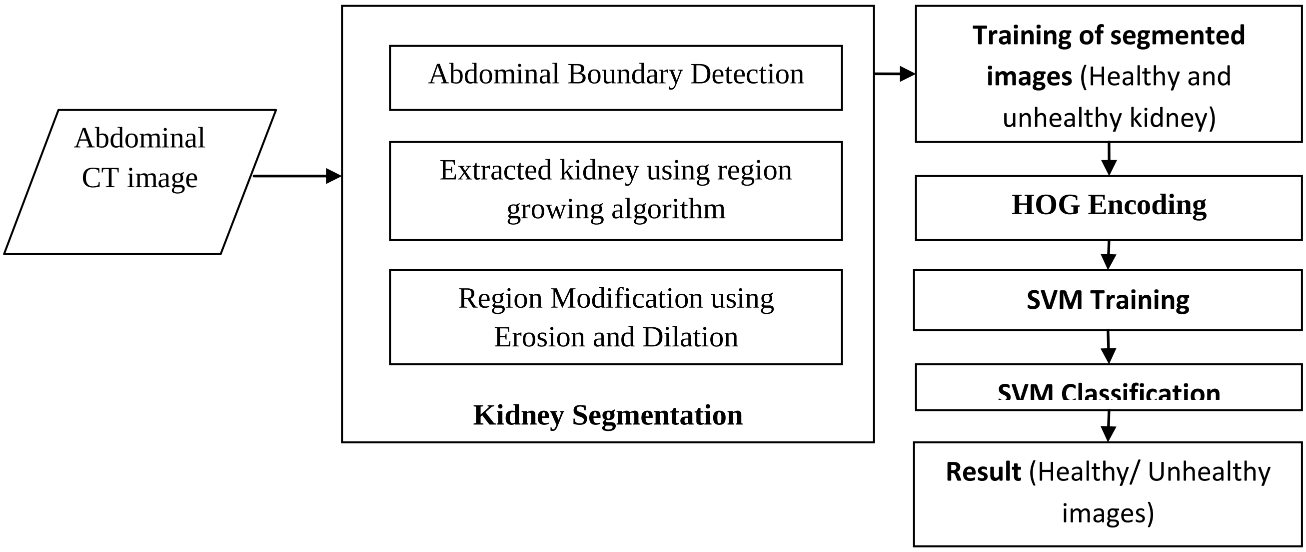

Fig 3: Overall methodology of our proposed approach

5. RESULT DISCUSSION

We have implemented our algorithm in MATLAB (version 8.4.0 (64bit)). All the experiment were performed on a computer bearing the configuration as follows:

CPU: Intel Core i3-2350M 2.30 GHz

RAM: 8 GB DDR3 1333MHz

Operating System: Windows 10 64bit.

There are 70 kidney CT images in our dataset. 40 images are used to train our system where 20 images are of healthy kidney CT images and another 20 images are of unhealthy kidney CT images.. Test sample contains 30 images where 15 images are of healthy kidney CT images and 15 images are of unhealthy kidney CT images. Each feature of healthy and unhealthy kidney CT images is trained and tested separately. So 2 sets of SVM models are trained and each set contains 3 models for 3 features. For simplicity, watermark, latent image and micro-printing models are going to be denoted using f1, f2 and f3, respectively. The output of these models is either 0 if it is unhealthy or 1 for healthy. Finally, we combine these results using the following equation.

$$\text{Result} = \left(\frac{f1}{3} + \frac{f2}{3} + \frac{f3}{3}\right) \times 100$$

If the result value is more than 50, then the image is detected as healthy. That means 2 out of 3 features have to recognize as healthy to testify the image as healthy. Our method produces 73.3% recognition accuracy so far. This is because of images are collected from different patients individually with the help of digital camera as the doctors don't give permission to access their confidential databases. Confusion matrices for Healthy and Unhealthy kidney images are given in Table I.

Table I: Confusion Matrix

		Predicted Class	
Actual Class		**Healthy Kidney**	**Unhealthy Kidney**
	Healthy Kidney	11	4
	Unhealthy Kidney	5	10

6. CONCLUSION

In this paper, an image-based methodology has been proposed to segment the kidney region and identify healthy and unhealthy kidney from the segmented CT images. We used region growing method to segment images and SVM classifier to detect healthy and unhealthy images after extracting three security features (watermark, latent image and micro-text) from the acquired segmented CT images of the kidneys. Here we have considered two types of images (Healthy and Unhealthy). With limited testing we have got 73.3% recognition accuracy. In the future work, this proposed machine learning algorithm will be focused on improving the system performance for both gray scale and color images.

7. ACKNOWLEDGMENTS

At first author offers gratitude to Almighty Allah who gave the ability to complete the whole thesis work without any breaking.

Author was partially supported by the Department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka, Bangladesh.

Author would also like to express gratitude to all the respected colleagues & all staffs of BUFT (BGMEA University of Fashion and Technology) who helped in different ways for this research project.

Lastly author wants to express hearty complement to the parents & family members for their constantly encouragement to come at this stage.

8. REFERENCES

[1] E. Supriyanto, Hafizah W.M, Wui Y.J, Arooj A. Automatic Non Invasive Kidney Volume Measurement Based On Ultrasound Image Proceedings of 15th International WSEAS GSCC Multiconferences held in Corfu Island , Greece, July 14- 16,2011 ISBN: 978-1-61804-019-0 Page 387-392

[2] R. PohleandK.D. Toennies,“Anewapproachformodel-basedadaptive region growing in medical image analysis,” in the 9th Int. Conf. Computer Analysis of Images and Patterns, vol. 2124, Warsaw, Poland, 2001, pp. 238–246.

[3] R. Pohle and K. D. Toennies, “Segmentation of medical images using adaptive region growing,” in Proc. Int. Soc. Opt. Eng. (SPIE), vol. 4322, 2001, pp. 1337–1346.

[4] R. Pohle and K. D. Toennies, “Self-learning model-based segmentation of medical images,” Image Process. Commun., vol. 7, pp. 97–113, 2001.

[5] S.H. Kim,S.W. Yoo,S.J. Kim,J.C. Kim,andJ.W. Park,“Segmentation ofkidneywithoutusingcontrastmediumonabdominalCTimage,”in5th Int. Conf. Signal Processing, vol. 2, Beijing, China, 2000, pp. 1147– 1152.

[6] S. W. Yoo, J. S. Cho, S. M. Noh, K. S. Shin, and J. W. Park, “Organ segmentation by comparing of gray value portion on abdominal CT image,” in 5th Int. Conf. on Signal Processing, vol. 2, Beijing, China, 2000, pp. 1201–1208.

[7] M. Kobashi and L. G. Shapiro, “Knowledge-based organ identification from CT images,” Pattern Recogn., vol. 28, pp. 475–491, 1995.

[8] S.Mavromatis, J.M.Bo, and J.Sequeira, “Medical image segmentation using texture directional features,” in IEEE 23rd Annu. Int. Conf. Eng. Med. Biol. Soc., vol. 3, Istandbul, Turkey, 2001.

[9] X. Wang, L. He, and W. G. Wee, “Deformable contour method: A constrained optimization approach,” Int. J. Comput. Vision, vol. 59, no. 1, pp. 87–108, 2004.

[10] B. Tsagaan, A. Shimizu, H. Kobatake, K. Miyakawa, and Y. Hanzawa, “Segmentation of kidney by using a deformable model,” in Int. Conf. Image Processing, vol. 3, Thessaloniki, Greece, 2001, pp. 1059–1062.

[11] B. Tsagaan,A. Shimizu,H. Kobatake,andK. Miyakawa,“Anautomated segmentation method of kidney using statistical information,” in Proc. Medical Image Computing and Computer Assisted Intervention, vol. 1, 2002, pp. 556–563.

[12] V. Vapnik, The Nature of Statistical Learning Theory, Springer, 1995.

[13] C. Cortes and V. Vapnik. Support vector networks. Machine Learning, 20, 1995.