Adrenal Tumor Classification on T1 and T2-weighted Abdominal MR Images

Mucahid BARSTUGAN[1], Rahime CEYLAN[1],
Electric and Electronics Engineering Department[1],
Engineering and Natural Sciences Faculty[1]
Konya Technical University[1]
Konya, Turkey
(mbarstugan, rceylan)@ktun.edu.tr

Semih ASOGLU[2], Hakan CEBECI[2], Mustafa KOPLAY[2]
Radiology Department, Medicine Faculty[2]
Selcuk University[2]
Konya, Turkey
dr.semihasoglu@gmail.com

***Abstract*—Adrenal tumors occur on adrenal glands and can be malignant. In this study, adrenal tumors on T1 and T2-weighted MR images were classified by the SVM algorithm. Before the classification stage, different feature extraction algorithms and filtering methods were used for preprocessing. The classification results that were obtained by four different methods were evaluated on five different evaluation metrics as sensitivity, specificity, accuracy, precision, and F-score. The best classification performance was obtained with Method 2 on T1-weighted MR (Magnetic Resonance) images where the sensitivity, specificity, accuracy, precision, and F-score metrics were obtained as 99.17%, 90%, 98.4%, 99.17%, and 99.13%, respectively.**

***Keywords*— *adrenal gland, adrenal tumor, MR image, classification, feature extraction, image filtering*.**

I. Introduction

Adrenal glands were firstly defined by Eustachius in 1563. Cortex and medulla of adrenal glands excrete cortisol and adrenalin hormone, respectively. If there is not enough cortisol excretion to the blood, inaccetence, asthenia and atonia situations occur. Adrenalin hormone expedites the heartbeat. Computed

MR, CT (Computed Tomography) and Ultrasound techniques are used to detect the pathological and physiological situation of the adrenal tumors. Usage of the anatomical and functional imaging systems has important role on diagnosis of the anomaly on adrenal glands.

Saiprasad et al. [1] selected ROI (Region of Interest) to segment the adrenal glands, which have an adenoma, on CT images. Adrenal glands were automatically segmented on the selected ROI, and the mean classification accuracy was obtained as 87%. Saiprasad et al. [2] used Random Forest algorithm to detect the adrenal anomaly, automatically. In their proposed method, CT images were classified into three different classes by pixel-based classification as adrenal, left adrenal and background. Then, the adrenal anomaly was detected by histogram analysis. The mean sensitivity and specificity values were found as 80% and 90% on 20 images. Koyuncu and Ceylan [3] classified 32 dynamic CT images that have 24 benign and 8 malign tumors. The feature extraction process was implemented on images by GLCM and secondary statistical features methods. The extracted features were classified by the proposed Bounded PSO-ANN method. The mean classification performance was achieved as 78.95%. Li et al. [4] extracted the features of benign and malign tumors by GLCM on CT images. Then, the extracted features were classified by spatial Bayesian modeling and the classification accuracy was obtained as 80%. Koyuncu and Ceylan [5] classified dynamic CT images that consisted of 90 benign and 24 malign tumors. They used different feature extraction algorithms and classifiers. They achieved at classification performance of 80.7%. Foti et al. [6] studied adrenal tumor classification on CT images. They used 80 adenoma images and 47 non-adenoma images. The statistical analysis was done on classification, and the classification accuracy was found as 97.6%. Romeo et al. [7] classified adrenal tumors on MR images. The feature extraction process was implemented on manually segmented adrenal tumors. The extracted features were classified by J48 algorithm. They used three types of tumor as lipid-rich adenoma, lipid-poor adenoma, and non-adenoma. The dataset consisted of 60 MR images, and the classification accuracy was obtained as 80% after the classification process.

As seen in literature studies, CT images are mostly used on adrenal tumor classification. In this study, adrenal tumors were classified into benign/malign on 122 images. Machine learning methods were used during the classification process. This paper has five parts. The second part analyzes the images as statistically and visually and explains the segmentation method and the algorithms used. The third part presents the results. The fourth part concludes withthe obtained findings.

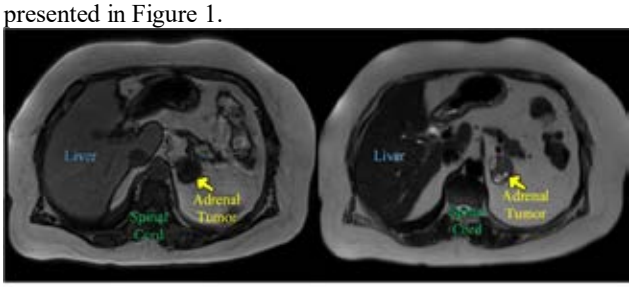
II. Material and Method

The data consists of 122 T1 and T2 phase MR abdominal images from the Radiology Department of the Medicine Faculty at Selcuk University. MR images were procured by SIEMENS AREA 1.5 2013 and obtained in DICOM format. The images were transformed into JPEG format, and have dimensions of 962x1160, resolution of 96 DPI, and depth of 24 bits. Three experienced radiologists in the Radiology Department detected the adrenal tumors on the MR images.

The experts detected nine different types of as adenoma, fatty adenoma, cyst, bloody cyst, lipoma, myelolipoma,

angiomyolipoma, adenomyolipoma, lymphangioma on 112 images in benign class. 10 images, which have the malign tumor, were defined as malignant.

The used dataset includes T1 and T2-weighted MR images. Four different methods in this study were implemented on adrenal tumors, which were manually segmented from T1 and T2-weighted MR images. The sample T1 and T2 images of the same MR scan are presented in Figure 1.



a) T1-weighted b) T2-weighted

Fig. 1. T1 and T2 images of the same MR scan

Figure 1 shows the differences between T1 and T2-weighted MR images. One of these differences is the grey-level differences of the organs and adrenal tumor. Another difference is that there is not fatty area inside adrenal tumor on T1 image. However, there is the fatty area inside adrenal tumor on T2 image. This situation shows that T1 and T2-weighted MR images may contain different information about the adrenal tumor.

A. The methods used

Four different methods were proposed to examine the effects of filtering and feature extraction methods on the classification process. The pipelines for the used methods are presented in Figure 2.

The most comprehensive method is Method 4, so only Method 4 is explained in detail.

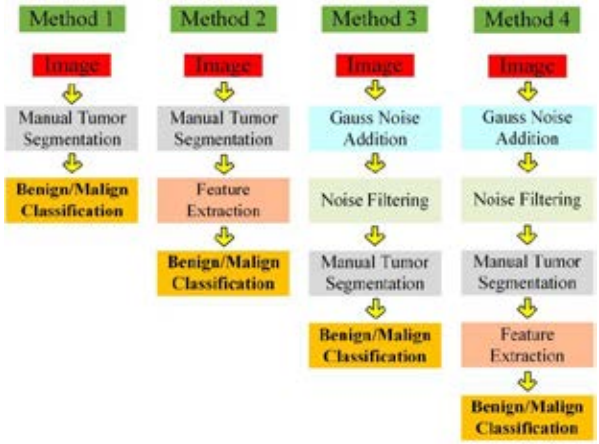


Fig. 2. The methods used

Method 4

In Method 4, Gaussian noise is added to images by taking σ=10, 15, 20, 25, 50, 75 and 100. Then, 2D Order Statistics [8], Anisotropic Diffusion Filter [9], Non-Local Mean Filter [10], and Wiener2 [11] techniques are used to remove the noise. Tumors were segmented when σ=25 value, which gave the best filtering result. The feature extraction process is implemented on the tumors, which are segmented from the results of four different filtering techniques, by GLSZM (Gray-Level Size Zone Matrix) [12], GLCM (Gray-Level Co-Occurrence Matrix) [13-15], GLRLM (Gray-Level Run Length Matrix) [16, 17], LDP (Local Directional Patterns) [18], SFTA (Segmentation-based Fractal Texture Analysis) [19], 2D-FFT (Fast Fourier Transform) [20], 2D-DWT (Discrete Wavelet Transform) [21], and 2D-DCT (Discrete Continuous Transform) [22] algorithms. The extracted features are classified by SVM [23].

B. Noise addition-filtering process

Noise addition process was implemented to remove the noises on the images. σ value was taken as 10, 15, 20, 25, 50, 75, and 100. The noise was removed by 2D Order Statistics Filtering, Anisotropic Diffusion Filter, Non-Local Mean Filter, and Wiener2 Filter methods. PSNR (Peak Signal to Noise Ratio) [24] and NERR (Noise Energy Reduction Ratio) [25] values of each method were obtained for different σ values.

III. Experimental Results

In this section, the results are presented as noise filtering results and classification results in two different parts. During the classification process, 2-folds, 5-folds, and 10-folds cross-validation methods were implemented. Five different evaluation metrics were used to examine the classification performances of the methods used. These metrics are sensitivity (SEN), specificity (SPE), accuracy (ACC), precision (PRE) and F-score [26].

A. Noise filtering results

PSNR and NERR metrics (Equation 1&2) were used to evaluate the noise filtering performances of the techniques used. PSNR and NERR results obtained for σ=10, 15, 20, 25, 50, 75, 100 by using 2D Order Statistics, Anisotropic Diffusion Filter, Non-Local Mean Filter, and Wiener2 Filter are presented in Table I. Table I shows the mean results of 122 images.

$$PSNR = 10log_{10}\left(\frac{MAX_I^2}{\frac{1}{mn}\sum_{i=0}^{m}\sum_{j=0}^{n}[I(i,j)-K(i,j)]^2}\right) \quad (1)$$

$$NERR = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}|D_n(u_i,v_j)|^2 - \sum_{i=1}^{N}\sum_{j=1}^{M}|D_f(u_i,v_j)|^2}{\sum_{i=1}^{N}\sum_{j=1}^{M}|D_n(u_i,v_j)|^2} \quad (2)$$

As seen in Table I, the highest result for NERR metric was found as 0.98 for σ=25, 75, and 100. In this study, after noise filtering process for σ=75 and 100, the image still included noise. Therefore, σ was taken as 25 in this study. Then, the noises were filtered and manual segmentation process was implemented.

TABLE I. THE MEAN PSNR AND NERR VALUES OF THE FILTERS

σ	Metric	Filter Type			
		Wiener2	2D-Order Statistics	Anisotropic Diffusion	Non-local Mean
σ=10 PSNR_in= 32.9 dB	PSNR_out (dB)	33.83	32.63	33.76	32.12
	NERR	0.87	0.28	0.91	0.38
σ=15 PSNR_in= 29.38 dB	PSNR_out (dB)	30.42	29.73	30.43	29.59
	NERR	0.91	0.42	0.96	0.49
σ=20 PSNR_in= 26.88 dB	PSNR_out (dB)	27.97	27.47	27.99	27.50
	NERR	0.92	0.60	0.97	0.71
σ=25 PSNR_in= 24.98 dB	PSNR_out (dB)	26.05	25.65	26.09	25.77
	NERR	0.94	0.68	**0.98**	0.81
σ=50 PSNR_in= 18.92 dB	PSNR_out (dB)	20.08	19.77	20.08	20.04
	NERR	0.96	0.79	0.97	0.95
σ=75 PSNR_in= 15.4 dB	PSNR_out (dB)	16.57	16.27	16.26	16.57
	NERR	0.97	0.81	0.76	**0.98**
σ=100 PSNR_in= 12.9 dB	PSNR_out (dB)	14.08	13.78	13.35	14.09
	NERR	0.97	0.81	0.41	**0.98**

B. Classification results

In this study, adrenal tumors were classified as benign/malign. SVM algorithm was used for the classification process. In Method 1, the images were classified without any preprocessing. In Method 2, the feature extraction process was implemented by eight different algorithms and the extracted features were classified. In Method 3, the noise filtering process was implemented by four different techniques. Then, the classification process was implemented. In Method 4, eight different feature extraction algorithms were implemented on four different noise filtering results. Then, the extracted features were classified. 2-fold, 5-fold and 10-fold cross-validation processes were realized during all classification processes. These methods were implemented on T1 and T2-weighted MR images and the obtained classification results were examined. When all of the results are examined, it is seen that Method 2 achieved the best classification performance.

So, a different experimental study was realized by Method 2. The best feature extraction method was obtained. Semi-automatic segmentation results of adrenal tumors, which were obtained in our previous study [27], were classified by Method 2 as seen in Figure 3. Table II presents the classification results of all of the proposed methods and semi-automatic segmentation results. As seen in Table III, the best classification performance was obtained as 98.4% on T1-weighted MR images by Method 2 with 10-fold cross-validation. GLCM showed the best performance during feature extraction on T1-weighted MR images. The highest classification performance on T2-weighted MR images was obtained as 97.57% by Method 2 with 10-fold cross-validation. The best result was obtained when the GLCM algorithm was used as a feature extraction algorithm on T2-weighted MR images. The highest classification results were obtained by Method 2 on T1 and T2-weighted MR images. Therefore, semi-automatic segmentation results were classified by Method 2. The classification accuracy of semi-automatic segmentation results was obtained as 94.23%. The classification performance on semi-automatic segmentation results was obtained lower than manual segmentation results around 4.2%. Also, the classification results of semi-automatic segmentation results show that malign tumors were not classified with high performance. Although the classification accuracy is over 90%, specificity value was obtained as 50%. The reason is that our previous study segmented the tumors with a segmentation performance around 60%. The classification performance became lower because of an information loss during segmentation of malign tumors.

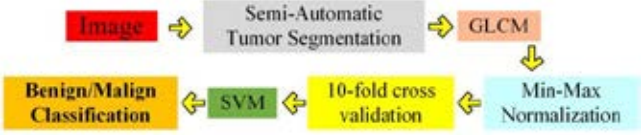


Fig. 3. The classification process of the semi-automatic segmentation results of adrenal tumors

IV. DISCUSSION AND CONCLUSION

This study classified adrenal tumors as benign/malign. The literature studies show that the highest classification performance on adrenal tumor classification was obtained in this study. There is not any standardized database on adrenal tumors in the literature. Therefore, the dataset used in this study and other studies is different from each other. The studies on adrenal tumors mostly use CT images. Table III shows the adrenal tumor classification results in the literature and compares with this study. As seen in Table III, there are only two studies on MR images. The highest classification performance on CT images was obtained as 97.6% in [6]. However, [6] obtained the results with statistical methods and did not use machine learning methods.

Future studies will try to manually segment and classify adrenal tumors on 3D images.

TABLE II CLASSIFICATION RESULTS OF THE MANUAL AND SEMI-AUTOMATIC SEGMENTATION RESULTS

Method	Image	Filter	FEM	CV	SEN (%)	SPE (%)	ACC (%)	PRE (%)	F-score (%)
Method 1	T1-weighted Manual Segmentation	x	x	5-fold	96.44±3.7	40±22.4	91.83±4.9	94.74±1.9	95.57±2.7
Method 2		x	GLCM	10-fold	99.17±2.6	90±31.6	98.4±3.3	99.17±2.6	99.13±1.8
Method 3		Anisotropic Diffusion Filter	x	5-fold	96.4±3.8	50	92.6±3.5	95.57±0.2	95.96±2
Method 4		All Filters	GLCM	5-fold	99.13±1.9	80±27.4	97.57±2.2	98.3±2.3	98.69±1.2
Method 1	T2-weighted Manual Segmentation	x	x	5-fold	96.44±3.8	50±35.4	92.67±3.3	95.72±2.9	96.01±1.8
Method 2		x	GLCM	10-fold	99.17±2.6	80±42.2	97.57±3.9	98.3±3.5	98.7±2.1
Method 3		Wiener Filter & 2D Order Filter	x	10-fold	94.7±8.7	60±51.6	91.86±7.7	96.67±4.3	95.37±4.7
Method 4		All Filters	GLCM	5-fold	99.13±1.9	80±27.4	97.53±2.3	98.26±2.4	98.67±1.2
Method 2	Semi-automatic Segmentation	x	GLCM	10-fold	98.26±3.7	50±52.7	94.23±4	95.83±4.4	96.92±2

FEM: Feature Extraction Method

TABLE III. LITERATURE COMPARISON

Study	Number of Total Tumors	Method	Data	SEN(%)	SPE(%)	ACC(%)	AUC(%)
[2]	20 (10 benign, 10 malign)	Random Forest	CT	80	90	----	----
[3]	32 (24 benign, 8 malign)	PSO-YSA	Dynamic CT	----	----	78.95	84.29
[4]	230 (121 benign, 109 malign)	Bayesian Modelling of GLCM	CT	----	----	80	----
[5]	114 (90 benign, 24 malign)	Bounded PSO-YSA	Dynamic CT	75	85.22	80.7	78.61
[6]	129 (82 benign, 47 malign)	Statistical Methods	CT	94.5	97.6	97.6	----
[7]	60 (40 benign, 20 malign)	J48	MR	78.67	79.67	80	79.47
This study	122 (112 benign, 10 malign)	GLCM – SVM (Method 2)	MR	99.17	90	98.4	----

REFERENCES

[1] G. Saiprasad, N. J. Saenz, C.-I. Chang, and E. L. Siegel, "Prototype Decision Support System for Evaluation of Adrenal Glands Incorporated into Routine CT Workflow," presented at the Radiological Society of North America Scientific Assembly and Annual Meeting, 2010.

[2] G. Saiprasad, C.-I. Chang, N. Safdar, N. Saenz, and E. Siegel, "Adrenal gland abnormality detection using random forest classification," *Journal of digital imaging,* vol. 26, no. 5, pp. 891-897, 2013.

[3] H. Koyuncu and R. Ceylan, "Classification of adrenal lesions by bounded PSO-NN," in *Signal Processing and Communications Applications Conference (SIU), 2017 25th*, 2017, pp. 1-4: IEEE.

[4] X. Li, M. Guindani, C. Ng, and B. Hobbs, "Classification of adrenal lesions through spatial Bayesian modeling of GLCM," in *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*, 2017, pp. 147-151: IEEE.

[5] H. Koyuncu, R. Ceylan, S. Asoglu, H. Cebeci, and M. Koplay, "An extensive study for binary characterisation of adrenal tumours," *Medical & biological engineering & computing,* pp. 1-14, 2018.

[6] G. Foti, G. Malleo, N. Faccioli, A. Guerriero, L. Furlani, and G. Carbognin, "Characterization of adrenal lesions using MDCT wash-out parameters: diagnostic accuracy of several combinations of intermediate and delayed phases," *La radiologia medica,* pp. 1-8, 2018.

[7] V. Romeo *et al.*, "Characterization of Adrenal Lesions on Unenhanced MRI Using Texture Analysis: A Machine‐Learning Approach," *Journal of Magnetic Resonance Imaging,* 2018.

[8] T. Huang, G. Yang, and G. Tang, "A fast two-dimensional median filtering algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing,* vol. 27, no. 1, pp. 13-18, 1979.

[9] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on pattern analysis and machine intelligence,* vol. 12, no. 7, pp. 629-639, 1990.

[10] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, vol. 2, pp. 60-65: IEEE.

[11] J. S. Lim, "Two-dimensional signal and image processing," *Englewood Cliffs, NJ, Prentice Hall, 1990, 710 p.,* 1990.

[12] G. Thibault *et al.*, "Shape and texture indexes application to cell nuclei classification," *International Journal of Pattern Recognition and Artificial Intelligence,* vol. 27, no. 01, p. 1357002, 2013.

[13] D. A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization," *Canadian Journal of remote sensing,* vol. 28, no. 1, pp. 45-62, 2002.

[14] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics,* vol. 3, no. 6, pp. 610-621, 1973.

[15] L.-K. Soh and C. Tsatsoulis, "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices," *CSE Journal Articles,* p. 47, 1999.

[16] A. Chu, C. M. Sehgal, and J. F. Greenleaf, "Use of gray value distribution of run lengths for texture analysis," *Pattern Recognition Letters,* vol. 11, no. 6, pp. 415-419, 1990.

[17] A. S. M. Sohail, P. Bhattacharya, S. P. Mudur, and S. Krishnamurthy, "Local relative GLRLM-based texture feature extraction for classifying ultrasound medical images," in *Electrical and Computer Engineering (CCECE), 2011 24th Canadian Conference on*, 2011, pp. 001092-001095: IEEE.

[18] T. Chakraborti, B. McCane, S. Mills, and U. Pal, "LOOP Descriptor: Local Optimal-Oriented Pattern," *IEEE Signal Processing Letters,* vol. 25, no. 5, pp. 635-639, 2018.

[19] A. F. Costa, G. Humpire-Mamani, and A. J. M. Traina, "An efficient algorithm for fractal analysis of textures," in *Graphics, Patterns and Images (SIBGRAPI), 2012 25th SIBGRAPI Conference on*, 2012, pp. 39-46: IEEE.

[20] R. C. Singleton, "Algorithms: Algorithm 338: algol procedures for the fast Fourier transform," *Communications of the ACM,* vol. 11, no. 11, pp. 773-776, 1968.

[21] M. M. Eltoukhy, I. Faye, and B. B. Samir, "A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation," *Computers in biology and medicine,* vol. 42, no. 1, pp. 123-128, 2012.

[22] A. K. Jain, *Fundamentals of digital image processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.

[23] S. R. Kulkarni and G. Harman, "Statistical learning theory: a tutorial," *Wiley Interdisciplinary Reviews: Computational Statistics,* vol. 3, no. 6, pp. 543-556, 2011.

[24] H. Koyuncu and R. Ceylan, "Elimination of white Gaussian noise in arterial phase CT images to bring adrenal tumours into the forefront," *Computerized Medical Imaging and Graphics,* vol. 65, pp. 46-57, 2018.

[25] E. Michel-González, M. H. Cho, and S. Y. Lee, "Geometric nonlinear diffusion filter and its application to X-ray imaging," *Biomedical engineering online,* vol. 10, no. 1, p. 47, 2011.

[26] S. Ruuska, W. Hämäläinen, S. Kajava, M. Mughal, P. Matilainen, and J. Mononen, "Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle," *Behavioural processes,* vol. 148, pp. 56-62, 2018.

[27] M. Barstuğan, R. Ceylan, S. Asoglu, H. Cebeci, and M. Koplay, "Adrenal tumor segmentation method for MR images," *Computer methods and programs in biomedicine,* vol. 164, pp. 87-100, 2018.