Research Article

Supervised Classification of Lymph Nodes based on ADC Maps Construction from Whole Body Diffusion Weighted MRI

Radhia Ferjaoui[1]*, Mohamed Ali Cherni[2], Nour El Houda Kraiem[3] and Tarek Kraiem[1,4]

[1]Department of Biophysics and Medical Technologies, Research Laboratory of Biophysics and Medical Technologies (LRBTM), University of Tunis El-Manar, ISTMT, 1006, Tunis, Tunisia; [2]Department of Signal and Image Processing, LR13 ES03 SIME Laboratory, University of Tunis, ENSIT, Montfleury 1008, Tunis, Tunisia; [3]Department of Radiology, Hospital Aziza Othmana of Tunis, 1006, Tunis, Tunisia; [4]Department of Medicine, University of Tunis El Manar, Faculty of Medicine of Tunis, 1007, Tunis, Tunisia

Corresponding author:
Radhia Ferjaoui,
Department of Bophysics and Medical Technologies,
Research Laboratory of biophysics and Medical Technologies (LRBTM),
University of Tunis El Manar, ISTMT, 1006, Tunis, Tunisia,
Tel: +21694441475;
E-mail: radhiaferjeoui@gmail.com

Abstract

Background and Aim: The aim of this study was to evaluate and analyze the different Apparent Diffusion Coefficient (ADC) values of components of heterogeneous lymph nodes by using the K-means technique, compared with a whole-lesion mean ADC value alone in discriminating benign and malignant pathologies. **Methods:** In this paper, we propose a new method based on functional information to recognize the malignancy of lymph nodes in DW MRI images. Twenty patients with a total of 102 lesions were included in this work, and the regions of interest (ROIs) were automatically extracted using the segmentation process based on the Chan-Vese algorithm. The functional information is obtained through the reconstruction of ADC maps with two diffusion factors: b-values at 0 and at 600 s.mm^2 for each ROI. Then the classification by K-means into solid and non-solid parts was done and the feature means ADC values were calculated for each cluster separately. And the distinguishing between cancerous lesions and benignant was done by using the K-nearest neighbors classifier (K-NN). **Results:** The results showed that the mean ADC values (in 10-3.mm^2.s^{-1}) of necrotic part 1.03±0.03 were significantly higher than those measured in the solid parts 0.84±0.02. The optimal ADC threshold value for differentiating benign from malignant lymph nodes was determined using the analysis of the receiver operating characteristic(ROC) at 1.12 * 10-3 .mm^2.s^{-1} with the sensitivity (SE), specificity (SP) and area under the curve (AUC) being 94.12%, 89.19% and 0.972%, respectively. And the mean ADC values of benign and malignant lymph nodes were 2.1 * 10-3 .mm^2.s^{-1} and 0.80±0.27 * 10-3 mm^2.s^{-1} respectively. The ADC values obtained for benign lymph nodes were higher ADC values than those in malignant lesions. **Conclusion:** Thus, the proposed computer-aided diagnosis is a helpful tool for automatic lymph nodes classification into clusters and it can successfully distinguish solid from non-solid parts in lymph nodes from the Whole body. It can also help users in predicting lesions pathologies (malignant or benign) based on the computer-aided diagnosis (CAD) system based on the K-NN classifier with accuracy higher than 93.43% and F1_measure and Geometric-mean values reach respectively 96%, 86.84%, when used ROIs placed in the solid partitions.

Keywords: Lymph nodes; DW MRI; Computer-aided diagnosis system; Segmentation; Feature extraction; Classification; K-nearest neighbors; Support vector machine

Introduction

Cancer is a major public health problem worldwide and is the second leading cause of death in the world. In particular, the Lymphoma, a form of cancer that affects the lymphocyte cells (B cell and T cell), is one of the most frequent cancers on the planet these days. It represents 3.2% of all cancers and about 225,000 deaths are estimated worldwide (2.7% of all deaths). [1] Typically, determination of the extent of cancer is very important for appropriate treatment planning and prognosis determination. [2,3] Hence, imaging modality (Computed Tomography (CT), Positron Emission Tomography (PET), PET/CT, Magnetic Resonance Imaging (MRI)...) plays a vital role in the initial staging and evaluation of response to therapy. However, both of these modalities (PET/CT and FDG-PET/CT) have several disadvantages such as the significant amount of radiation exposure which result in the development of secondary malignancies [4,5] and the relatively high cost they have. [6] Also,

How to Cite this Article: Ferjaoui R, et al. Supervised Classification of Lymph Nodes based on ADC Maps Construction from Whole Body Diffusion Weighted MRI. Ann Med Health Sci Res. 2020;10: 980-988.

not many institutions have these modalities. Recently has emerged an important imaging modality, Whole Body Diffusion Weighted Imaging MRI (WB-DWI-MRI) [7] that does not use any ionizing radiation and can be used to obtain physiological information about lymphoma for the whole human body. For this reason, DWI is a potentially useful technique for the staging of Lymphomas. [8,9] And it provides functional information in that it allows quantification of the random motion of water molecules by measuring Apparent Diffusion Coefficient (ADC) value. [10] For quantitative measurement, the region-of-interest (ROI) was manually extracted by doctors on one or more representative slices, but the position was determined by visual identification. This usually depends on the user experience and these methods are time-consuming. Also, usually only the main apparent diffusion coefficient (ADC) was obtained for whole-lesions. Zhang et al. on 2013 [11] concluded in their study that a variation in the positions of ROIs affects the study results. Therefore, a careful selection of ROIs positions within suspect lesions is needed for measuring ADC values in order to achieve a successful diagnosis with the lowest times. The main contributions in this research work are: (1) we propose an automated technique using computer-aided approaches for detection and location estimation for all tumors inside the whole body to determine the best course of treatment early on. (2) We apply the proposed segmentation method to extract lesion from the whole body DW image (with b value=600 s.mm²). (3) And for quantitative analysis, we construct a parametric image named ADC map by using two images acquired with two different b values (b0, b600). (4) Then for each constructed parametric ADC map, the K-means clustering technique is used to differentiate between the components of lesions and detect the solid parts. (5) ADC features were extracted and calculated from each cluster separately and used for input vector to the K-nearest neighbor's classifier to choose the best component with high accuracy for distinguishing benign from malignant lesions. Finally, abundant experimental results show better performance (Accuracy, sensitivity, and specificity) of our research compared with the state-of-the-art image segmentation methods in the four regions in the whole body. And, found encouraging results to use the recently and non-ionizing Diffusion-weighted MRI modality for the staging of lymphoma.

The paper is organized as follows: Section 2 is about the experimental setup of the proposed method. Then we present and discuss our results in section 3. Finally, conclusions and future enhancements are given in section 4.

Research Methodology

In this section, we start by describing the clinical DW WB MRI image database. Next, we present the experimental setup of the proposed method. The proposed system showing in Figure 1 is done using several steps wish are given in the subsequent subsections: firstly, automatically extract ROIs within lymph nodes and affected organs, then calculate the ADC values and construct the parametric ADC maps of whole-tumors. After that, the clustering phase based on the K-means algorithm was performed, and the mean ADC values were compared with the measured average ADC value of each component. Finally, the classification step was drawn to give the decision for distinguishing malignant from benign lymph nodes, by using the K-NN classifier compared with support vector machine (SVM).

Study population

Twenty participants with enlarged lymph nodes in one or/ and more regions (head/neck, chest, abdominal, pelvis) were included in this study. Patients aged 19 to 81 years old, where the mean age is 48.16 years, were diagnosed. For 10 males and 10 females a total of 204 images (102 with b=0 s.mm², and 102 for b=600 s.mm²) with malignant and benign lymph nodes were collected. The diagnosis was performed in the period from 27 June 2015 till February 2018. Conventional MRI and DWI studies were used for lymph nodes evaluation using 1.5Tesla MR scans.

WB-DW-MR protocol

For all patients the imaging protocol consisted of using the Whole Body DW-MRI exams which are conducted in daily practice conditions. All DW-MR imaging was performed with a 1.5T MR scanner (Signa HDxt, General electric, healthcare) in Tunis, with a four-element phased array surface coil in all body regions for signal reception. In all patients, diffusion weighted images of 4 regions [Figure 2]. Head/neck, chest, abdomen, and the pelvis region, were obtained using a sequence EPI

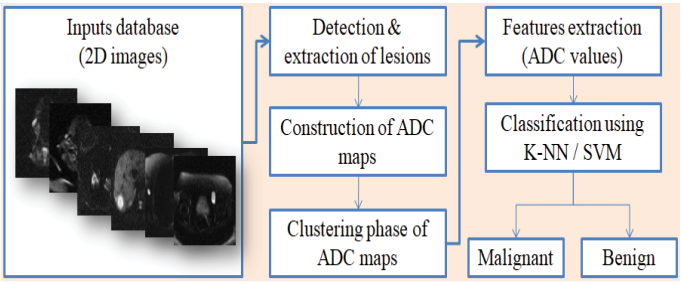


Figure 1: Proposed block diagram of the proposed method.

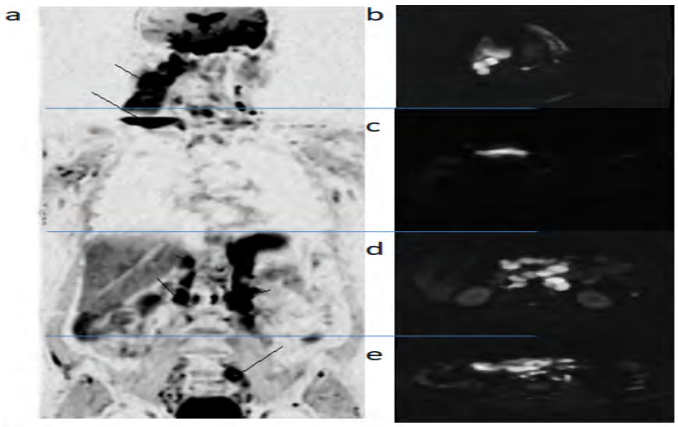


Figure 2: Coronal maximum intensity projection gray-scale inverted whole-Body (64-year old man) shows extensive supra-diaphragmatic and infra-diaphragmatic lymph node involvement (arrows) (a). Axial DW images acquired with b-value of 600 s (mm) 2 showing lymphomatous nodes in the head/neck, Chest, Abdominal, and the pelvic regions respectively (b), (c), (d) and (e).

(single-shot echo-planer imaging) with repetition time (TR) and echo delay time (TE) [Table 1]. All 2D images were obtained in the axial, coronal, and sagittal planes. And they were encoded with .dcm extension (Digital Image and Communications in Medicine: DICOM).

Segmentation phase

As shown in Figure 3, lymphoma segmentation is based on filtering the DW slice, the choice of which depends on the ROI containing the tumor that is biggest in diameter. The filtering is done using a median filter to remove noise caused by operator performance, equipment, and the environment. The first step in segmentation is the creation of the initial mask. To do this, we apply the morphological transformation by using many operators defined in mathematical morphology as: dilation and erosion, opening and closing. [12] Then the Chan-Vese algorithm [13] based on technique of curve evolution is used to extract the lesion in the diffusion weighted image (b600). The basic idea starts by simplifying the energy-based model proposed by The Mumford-Shah [14] describing with the energy functional equation (1) to obtain equation (2) formulated as below.

$$E^{MS}(v,C) = \int_{\Omega} |I - v|^2 dxdy + v \times \int_{\Omega/C} |\nabla v|^2 \mathrm{dxdy} + \mu|C| \quad (1)$$

Where ν is a nearly piecewise smooth approximation to the image I, and the curve C is a set of edges between disjoint regions within I. ν and μ are fixed scale parameters, Ω/C is the domain excluding the curve C (edges between disjoint regions within I), and the length of the curve C is induquetd by |C|.

$$E(c1,c2,C) = \mu * length(C) + v * Area(inside(C)) + \lambda_1 \int_{inside\ (C)} |I(x,y) - c1|^2 dxdy + \lambda_2 \int_{outside\ (C)} |I(x,y) - c2|^2 dxdy \quad (2)$$

Construction of ADC maps

After automatic extraction of the ROIs, the ADC values are calculated pixel by pixel using an implicit monoexponentiel model according to simple equation (Eq. 3) for each extracted lymph node: [15]

Table 1: Technical parameters of DWI at 1.5T and study population.

Parameters	Values and Units
TR (Repetition Time)	7050 (msec)
TE (Echo Time)	84.8 (msec)
TI (Inversion Time)	0 (msec)
Number of signal averaged	8
Slice thickness	8 (mm)
Dw images	102 images
MRI images	102 images
Matrix size	256 × 256 (pixel)
Population	20 patient
Gender	10 Male
	10 Female
Ages	Median age 47 years
	Range 16-81 years
Total lesions	102 lesions
head/neck region	4 lesions
Chest region	35 lesions
Abdomen region	34 lesions
Pelvis region	29 lesions

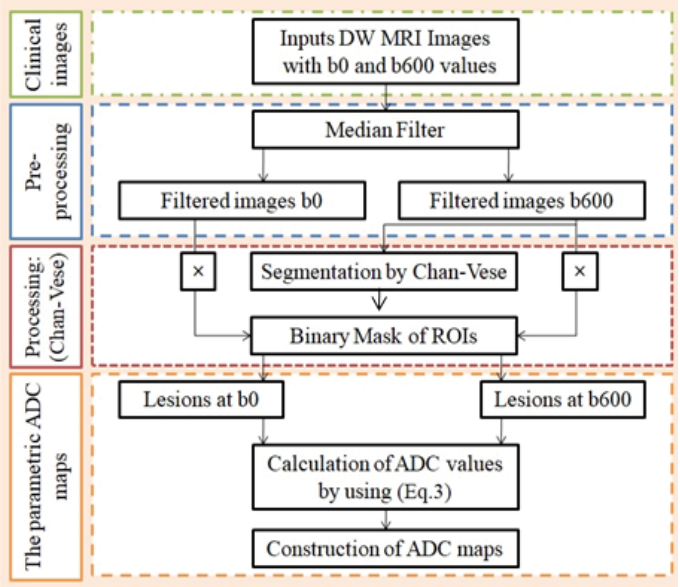


Figure 3: Proposed method to calculate ADC values and construction of the parametric ADC maps.

$$ADC = \frac{\ln(\frac{S_0}{S_i})}{b_i - b_0}\ (mm^2.s^{-1}) \quad (3)$$

With S_0 and S_i is the signal intensity of the same selected region of interest (ROI) on the image obtained with the b_0 and b_i values, respectively. b is the diffusion factor (the motion-probing gradient factors (in s.mm²)). In this study we used b_0=0 and b_i=600. The diagram in Figure 3 shows the different steps of proposed method to calculate and construct the ADC map for each lesion. The calculated ADC value from the ROIs of DW image b_0 and b600 were then averaged to a mean ADC values. Hence, if the ROIs are with homogenous signal intensity, the ADC of the whole-lesion for each ROI is calculated. On the other hand, if the extracted ROI is heterogeneous, a clustering step is needed to distinguish the lesion components into solid and non-solid. And the mean ADC value was calculated for each component separately.

Clustering phase

The clustering step aims to group's pixels of the segmented lesion such a way that pixels in the same cluster have a higher degree of similarity to each other than to those in other clusters. The proposed method is based on the unsupervised learning algorithm, the K-Means clustering to partition the heterogeneous ROIs from the parametric ADC maps into k cluster. The procedure is simple and easy in classifying the ROIs across a number of clusters k that is fixed from the beginning. Secondly, each pixel (xi) was associated with the nearest centroid (ck) by calculating the Euclidean distance (d) using the distance formula (equation 4, 5) between the pixel and the centroid. If the pixel has the shortest distance among all then it is moving to a particular cluster (solid, non-solid, or background cluster). Finally, this process is repeated until all pixel compares to cluster centroids and the cluster centers are stable. [16] Figure 4 shows the diagram of the proposed method for the clustering of various components of the segmented lymph nodes and the

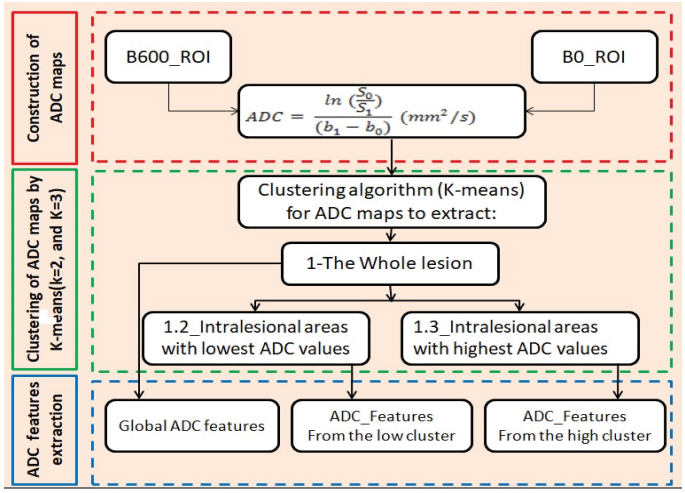


Figure 4: Proposed method for clustering of various components of the ADC maps.

choice of the cluster number.

$$KM(X, C) = \sum_{i=1}^{n} \min \|x_i - c_j\|^2 \ (4)$$

$$d(p, q) = d(q, p) = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2} \ (5)$$

The ADC values of the whole homogenous ROIs were calculated from the parametric ADC maps. But if the signal intensity of the ADC maps or the segmented ROIs are heterogeneous, then we set the class number at 2 in the K-means algorithm to classify those segmented ROIs into the background and grayscale lesions. The mean ADC value of whole lesions was calculated. On the other hand, the number of classes K is fixed to 3, since there are three classes in each input: the background and the two classes included in the grayscale lymph node: the solid part, which it is represented by the high-intensity pixels, and the non-solid part, which is represented by the low-signal intensity in diffusion-weighted image (b600). The mean ADC values were calculated for each cluster separately and compared with each other.

ADC features extraction

ADC measurements were performed by using the result of the clustering step. Therefore, the mean ADC values (ADC mean) were calculated over each cluster. This resulted in 102 features for each cluster of the entire database. All of this was used for input data to The K-nearest neighbor's classifier (K-NN) to differentiate between benign and malignant lesions. The description of this step is given in the classification phase.

Classification phase

K-nearest neighbors algorithm (K-NN) was used in our work, to classify the input datasets 'D' into two-class classification: Cancer lesion-present and cancer lesion-absent named respectively malignant and benign. The idea of this easy, non-parametric and supervised approach is to maximize the intra-class variance and to minimize the inter-class variance. Hence, predictions are made for a new feature vector (x) by searching through the entire training datasets (D) for their K nearest neighbors according to Euclidean distance (d) and summarizing the output variable for those K cases. Then in classification x is assigned to the class to which the majority of those k nearest neighbors belong to.

The algorithm for the k-nearest neighbors is given below as Figure 5.

Evaluating CAD system

Statistical analysis was used in this study. A box-and-whisker plot was used to represent the scatter plot of mean ADC values of benign and malignant lymph nodes for the three cluster models. An unpaired bilateral Student t-test was performed for all groups to determine differences in mean benign and malignant lymph nodes. After the correlation of the histopathology findings by a 20-year-experience radiologist, and the results of the K-NN technique for each ROI, the optimal ADC threshold value for differentiating benign from malignant lymph nodes was determined using the analysis of the receiver operating characteristic (ROC). Also, we used five evaluation parameters, which are used to determine the performance of the proposed method in the following categories: The sensitivity (SE), Specificity (SP), Accuracy (AC), Precision (PR), and the Recall (RC). The above mentioned parameters are computed using Eqs. (6), (7), (8), (9) and (10) respectively.

$$Sensitiity = \frac{TP}{TP + FN} * 100 \ (6)$$

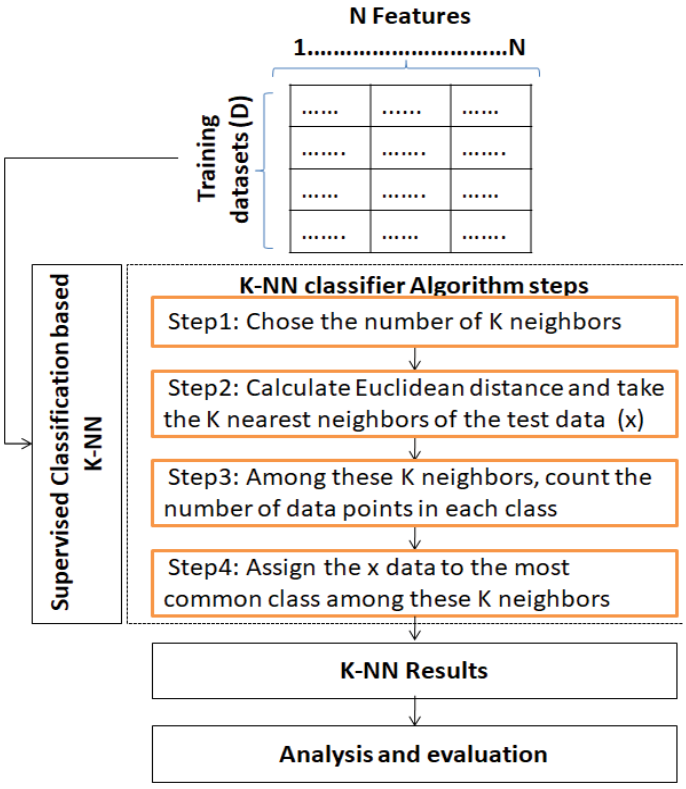


Figure 5: The detailed diagram of the proposed classification methodology based on the K-NN classifier.

$$specificity = \frac{TN}{TN + FP} * 100 \ (7)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} * 100 \ (8)$$

$$Precision = \frac{TP}{TP + FP} * 100 \ (9)$$

$$Recall = \frac{TP}{TP + FN} \ (10)$$

Where,

- TP is the True Positive, which represents the count of pixels correctly segmented with the center of the lesion is inside the manually segmented lymph node by the expert (a radiologist with 20 years of experience in DW-MR imagine).
- The True negative (TN) is the count of pixels correctly segmented as none manually segmented regions.
- FP and FN are the False positive and False negative respectively which are the number of pixels falsely identified as manually segmented ROI and non-manually segmented regions.
- Precision (PR) is the correct segmentation that refers to the percentage of True Positive (TP).
- The sensitivity (SE) is the number of the TP divided by the total number of elements that belong to the positive cluster.

A cut-off value was determined to predict lesion malignancy using the maximum value of the Youden index, calculated for each point of the ROC curve (SE + SP-1). All of these parameters and the AUC (area under the curve) were calculated using 95% confidence intervals. Differences in the diagnostic performance among clustered ADC values were analyzed by comparing the ROC curves according to the method described by DeLong et al. [17] We also used the Bland-Altman plot to verify the level of agreement between the result obtained when we use solid parts and the whole-lesions average ADC value alone. The performance of classifier is done by measuring the accuracy which represent the probability that a diagnostic test is correctly performed, F1 _ measure Eqs. (11) and Geometric mean Eqs. (12).

$$F1_{measure} = \frac{2TP}{2TP + FP + FN} \ (11)$$

$$\text{Geometric\_mean} = \sqrt{\text{Sensitivity} * \text{specificity}} \ (12)$$

Results and Discussion

Experiments on clinical data composed by 102 tumors demonstrated that the proposed method shows encouraging results not only in one region of the body, but in the whole body for these four regions: the head/neck, the chest, the abdomen, and the pelvis region. And this is the first work use the imaging processing technique to Automatic segment, detect, clustering and classify the Lymphoma lesion in The Whole Body Diffusion Weighted MRI technique, the non-ionizing modality wish has a low cost compared with other modality (FDG-PET/CT, PET/CT). The segmentation results of the proposed and existing methods are presented below. An example of segmentation and detection of lymphoma lesion in the four regions is shown in Figure 6. The results presented in Table 2 show that the proposed method [18] is more significant in terms of segmentation performance in the abdomen region than the other regions. In this region, the specificity is 99.99% and the accuracy is equal

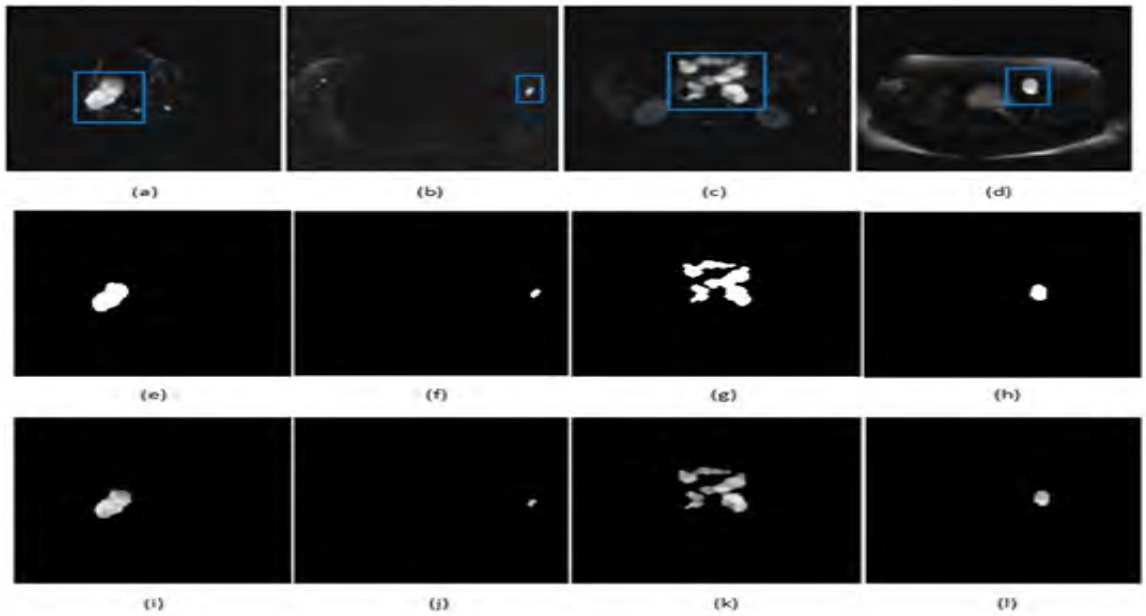


Figure 6: Results of the proposed Methods based on a Chan-Vese algorithm for the four regions in the WB with a heterogeneous lesion in the box. The axial DW MR image showing enlarged lymphadenopathy (a). Axillaries Lymph nodes is showing in the Chest region (b). Intestinal LN present in the abdominal region (c). And the inguinal LN is on the pelvis region (d). (e), (f), (g), and (h) are their binary masks, respectively. (i), (j), (k), and (l) represent respectively the segmented lesions on axial images at b=600s.mm².

to 99.97%. In the head/neck region, the specificity and accuracy are respectively 99.99% and 99.96% by using the proposed segmentation algorithm.

To evaluate the performance of the proposed method, we also used a receiver operating characteristic (ROC) curve analysis and the area under the curve (AUC) was measured for each region to indicate the successful results obtained by this method. The results of the ROC curve and AUC are presented in Figure 7 and Table 3 respectively. From these results, we can conclude that the AUC value for the majority of the 4 regions are more than 0.97 when we use the proposed method and the highest AUC value (98%) is obtained for the abdomen regions. This method is compared with Kamdi et al. method. [19] It is better than the AUC values obtained by Kamdi et al. which range between 0.5 and 0.94 for the head neck regions and chest regions, respectively.

An example is showing in Figure 8 explain the results of the segmentation method and indicates the construction of the parametric ADC map in the chest regions. Figure 9 showed the results of the clustering step based on the K-means cluster to extract the components of the whole lesion, where mean ADC features are measured.

The box-and-whisker showing in Figure 10 represents the distribution of mean ADC values of solid portions compared with those obtained in necrotic parts of lymph nodes across all 102 ROI's among the different clusters of K-means. The horizontal line in each box is a median (50th percentile) of the

Table 2: Quantitative evaluation of performances measured using the proposed method.

	Head/neck	Chest	Abdomen	Pelvis	All regions
SE (%)	88.54	96.19	94.76	92.52	93
SP (%)	99.99	99.95	99.92	99.99	99.96
AC (%)	99.96	99.95	99.88	99.97	99.94
PR	0.94	0.90	0.86	0.91	0.90
RC	0.88	0.96	0.94	0.92	0.93

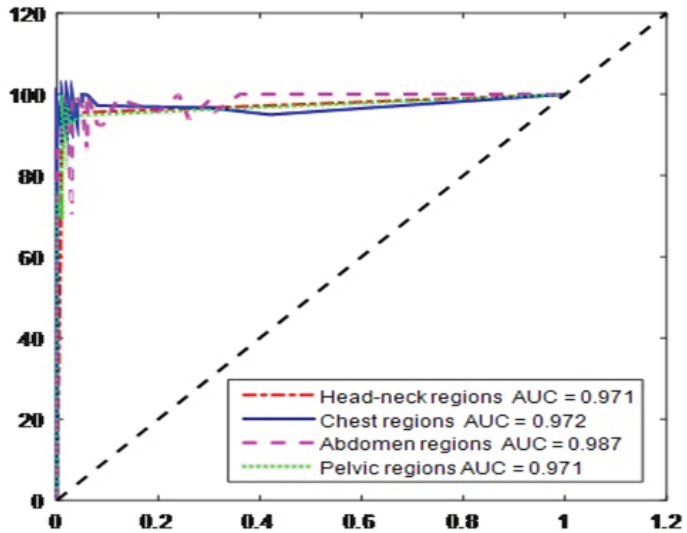


Figure 7: Results of the Receiver Operating Characteristic (ROC) curve of the detected lesions in the four regions.

Table 3: The area under the ROC curve (AUC) values achieved with the proposed method and Kamdi et al. method for the different region.

	Regions			
Methods	Head/neck	Chest	Abdomen	Pelvis
S. Kamdi et al. [19]	0.5	0.94	0.89	0.91
Proposed method	0.97	0.97	0.98	0.97

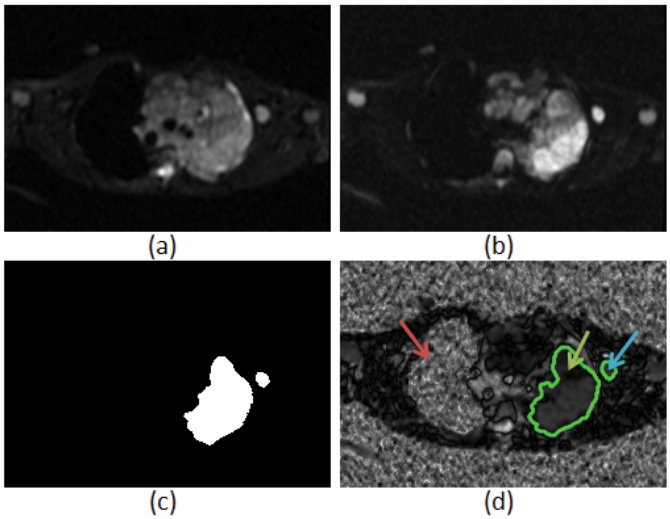


Figure 8: Results of the segmentation method and construction of the ADC maps for MRI images of cervical lymph nodes in a 29-year-old patient. Lymph nodes show with a heterogeneous signal intensity on axial images obtained at b=0s. (mm) 2 (a) and at b=600s.(mm) 2 (b). The binary mask of the segmented region by using Chan-Vese (c). Constructed ADC maps by using Eq.3. (d) The red arrow, and both of the blue and green arrows represent respectively healthy and lymphomatous regions.

measured values. The top and bottom of the boxes represent respectively 25the and 75th percentiles. Whiskers indicate the range between the smallest and largest data point observed. For the heterogeneous ROIs, the cluster characterized by high signal intensity at b value of 0 s. mm² and reduced signals at b 600 s.mm², and which represents the non-solid parts as shown bright on ADC maps and the ADC values ranging from 0.3 ×10−3 $mm^2.s^{-1}$ to 1.87×10^{-3} $mm^2.s^{-1}$ with mean equal to 1.03 ± 0.3 $\times10^{-3}$ $mm^2.s^{-1}$. These ADC values were significantly higher than those measured in the solid parts [Figure 10]. The other cluster of the solid portions is represented by low signal intensities at b=0 s.mm², bright on images at b=600 s. mm² and hypointense on ADC maps. In this cluster the mean ADC value was 0.84 ± 0.2×10^{-3} $mm^2.s^{-1}$ and the ADC value was between 0.24 and 1.29 $\times10^{-3}$ $mm^2.s^{-1}$.

MRI images of cervical lymph nodes in a 29-year-old patient. Lymph nodes show with a heterogeneous signal intensity on axial images obtained at b=0s.(mm) 2 (a) and at b=600s. (mm) 2 (b). The binary mask of the segmented region by using Chan-Vese (c). Constructed ADC maps by using Eq.3 .(d) The red arrow, and both of the blue and green arrows represent respectively healthy and lymphomatous regions.

When we use two clusters for the K-means clustering, the ADC values of the whole lesions were 0.1 ×10−3 $mm^2.s^{-1}$ to 2.9 $\times10^{-3}$ $mm^2.s^{-1}$. The homogenous regions are characterized by high

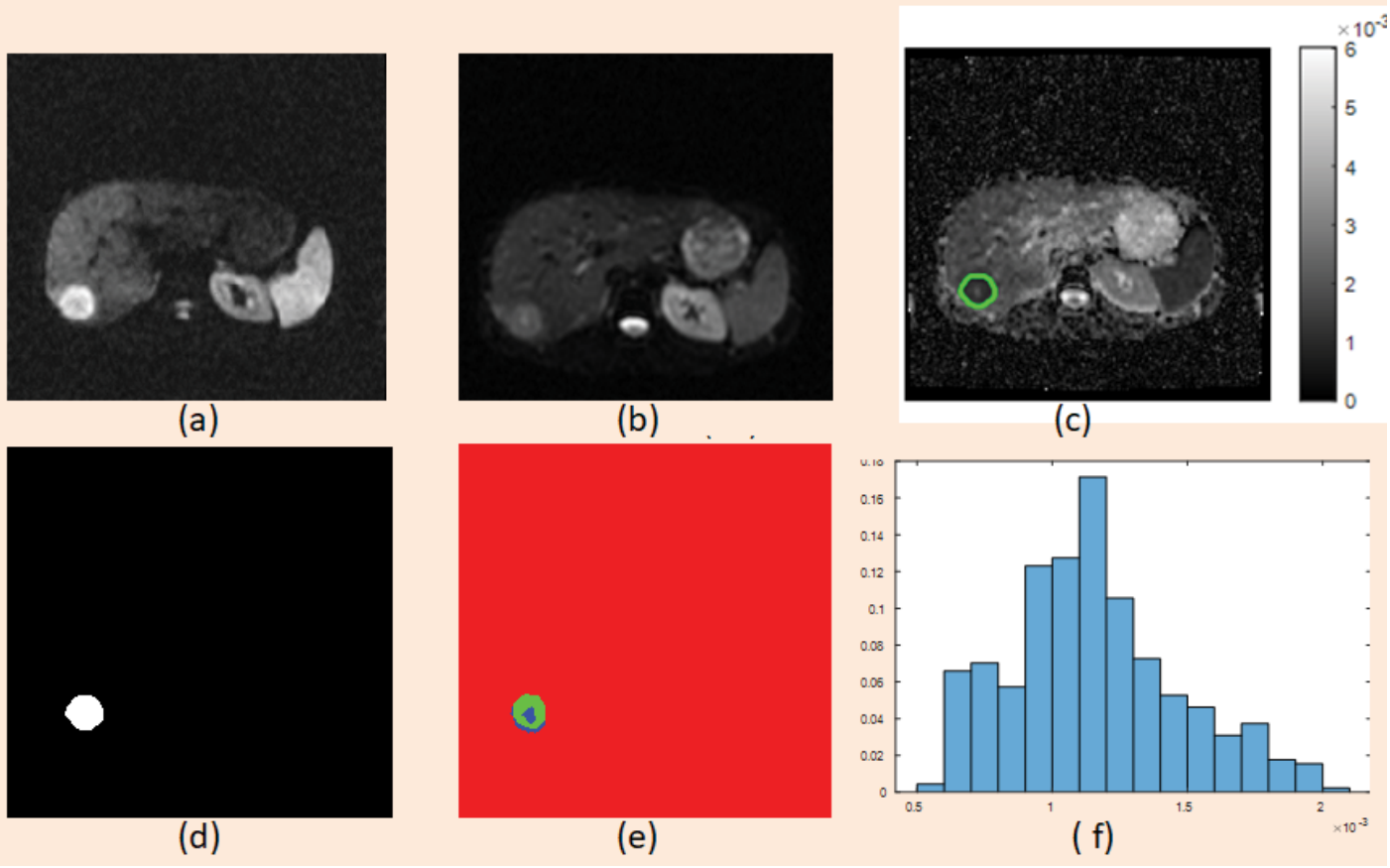


Figure 9: Malignant lymphoma in the hepatic region. Segmentation and clustering phase based on K-means algorithm to extract the solid-portions cluster of the whole-lesions.

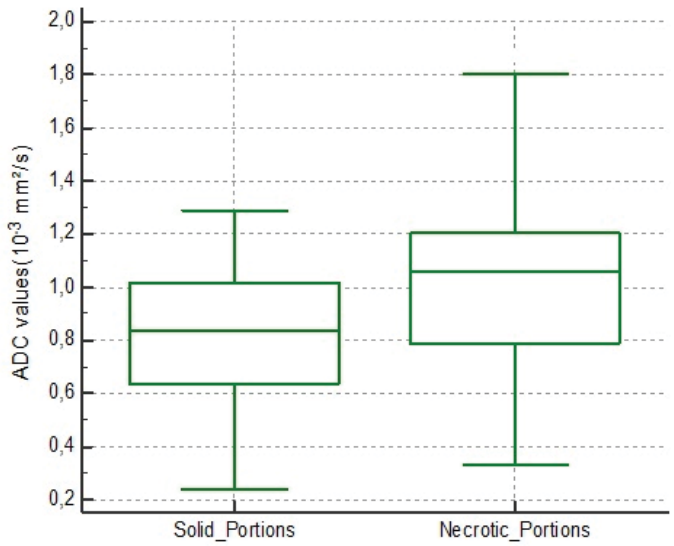


Figure 10: A box-and-whisker was used to represent the scatter plot of the average ADC values of the solid and non-solid portions of the segmented lesions.

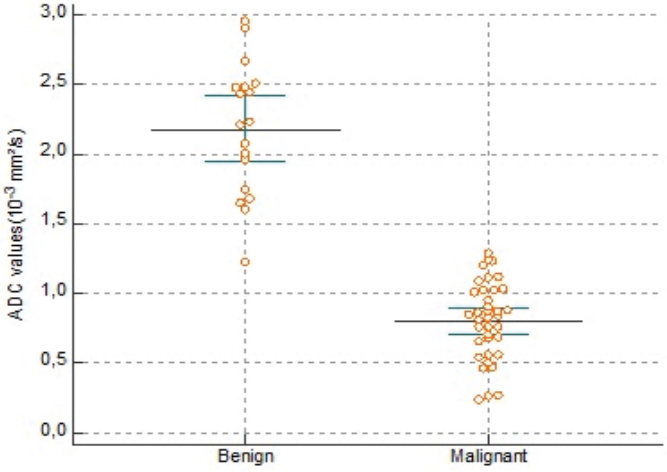


Figure 11: Dot plot illustrated the distribution of apparent diffusion coefficients of benign and malignant lymph nodes in patients with lymphoma. Black and blue line in the dot plot represented mean and standard deviation respectively.

signal intensities at both of b0 and b600 and hypo-intense on ADC maps.

The K-nearest neighbor's classifier was used in this study to classify the ROIs based on their ADC values, within benign and malignant lesions because identified metastatic nodes are very important for the staging and therapy planning. Figure 11 show that the benign lymph nodes had higher ADC values than those of malignant ROIs. However, the mean ADC value of malignant ROIs was $0.80 \pm 0.27 \times 10^{-3}$ mm².s^{-1} and the ADC values ranged between 0.71 and 0.89×10^{-3} mm².s^{-1}.

The decrease of the ADC values for malignant lesions [Figure 11] is indicated by the enlarged nuclei and hypercellularity. Therefore, it could consequently reduce the diffusion space of water protons in the extracellular and intracellular dimensions.

Our results demonstrated that the ADC values could distinguish malignant from benign ROIs.

Figure 12 presents a comparison of ROC curves of the solid-parts cluster and whole-lesion cluster. We showed that through a clustered analysis of the ADC values within lesions into different partitions and especially in solid portions, we could better differentiate malignant from benign extracted ROIs than through analysis of whole-lesions mean ADC values. This is because in whole lesions there exists other components of lymph nodes such as necrotic parts increase the ADC values. The ROC curves were done to determine the threshold ADC value which equals 1.12 × 10^{-3} $mm^2.s^{-1}$ for differentiating the two groups, benign and malignant, with SE, SP, and the AUC are 94.12%, 89.19 %, and 0.972, respectively.

Our work showed that the average ADC values in malignant lymph nodes were significantly lower than of benignant nodes by Whole Body DW applying b values of 0 and 600 s.mm^2. Compared with recent studies [Table 4] about differentiation of pathologies lymph nodes by DW imaging, there were some similarities and differences. Abdel Razek et al. [20] have indicated that the average ADC value of benign equal 1.64 ± 0.16 ×10^{-3} $mm^2.s^{-1}$ was significantly higher than that 1.09 ± 0.11 ×10^{-3} $mm^2.s^{-1}$ of malignant lymph nodes. Similarly, Li et al. [21] have showed that the mean ADC value (10^{-3} $mm^2.s^{-1}$) of benign (1.02 ± 0.23) were significant higher than that of malignant lymph nodes (0.67±0.10). The mean ADC value of malignant lymph nodes 0.78±0.09 ×10^{-3} $mm^2.s^{-1}$ was also lower than that of benignant lymph nodes 1.24±0.16 ×10^{-3} $mm^2.s^{-1}$ in the Holzapfel et al. [22] study. However, Cintra MB et al. [23] had reported that measuring average ADC values could not allow distinguishing between benign and malignant lymph nodes.

Leave-one-out cross-validation is used to estimate the performance and the accuracy of the K-NN classifier compared with the support vector machine (SVM) model. The results obtained are showing in Figure 13. These training results demonstrate that K-NN is the best learning model with a learning accuracy reaching 93.43% where the missed classification value is 6.57% for all the training data. And with F1 measure and

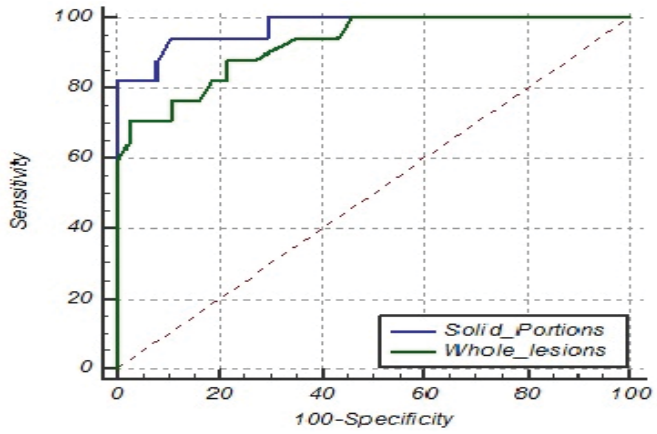
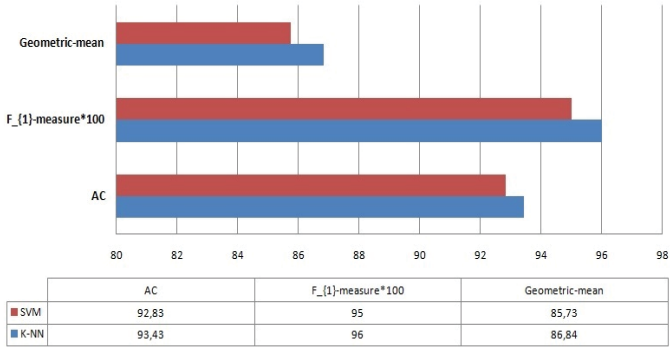


Figure 12: Receiver operating characteristic curves obtained through the proposed method based on K-means analysis of the solid-portions cluster and the whole-lesions.

Table 4: The Mean ADC values achieved with the proposed method and literature studies for malignant and benign lymph nodes.

Methods	Mean ADC values ± STD	
	Malignant	Benign
Proposed Method (2020)	0.8 ± 0.27	2.1 ± 0.17
Abdel Razek et al. [20]	1.09 ± 0.11	1.64 ± 0.16
Li et al. [21]	0.67 ± 0.10	1.02 ± 0.23
Holzapfel et al. [22]	0.78 ± 0.09	1.24 ± 0.16



	AC	F_{1}-measure*100	Geometric-mean
SVM	92,83	95	85,73
K-NN	93,43	96	86,84

Figure 13: Evaluation performance of the proposed method based on K-NN and SVM classifiers.

Geometric mean values higher than these obtained with SVM, which reaches respectively 96%, 86.84%, and 95%, 85.73% and used ROIs placed in the solid partitions. Therefore, we can conclude from these obtained results the superiority of the proposed CAD system.

Our study has several limitations. First, the sample of regions of interest with the lesions was small. Secondly, the quality of the DWI images was not adequate for the processing steps induced by the susceptibility artifacts. And only functional parameter values are used for the classification. Our future work will be on developing more features and especially the anatomical information obtained not only from the diffusion image, with b600 value. But, we will be extending our work to extracting more morphological and anatomical parameters from the fusion image obtained from the b600 and b0 images. And, we need to increase the database, especially regarding the samples of suspicious tissues and lymphoma in all regions to confirm our results.

Conclusion

Experiments on clinical application data composed of 102 tumors demonstrated that this new proposed method presents good results for the four processes: segmentation, clustering, feature extraction, and classification phase. 18 The segmentation using the Chan-Vese algorithm to automatically extract lymph nodes in the four regions of the Whole body gives encouraging results with high sensitivity (SE) and specificity (SP). Applying the K-means technique to ADC maps in the clustering phase not only facilitated the quantification of heterogeneous ROIs by calculating the mean ADC values for each component of lymph nodes but also enabled us to compare these values with the mean whole-lesion ADC values. Moreover, the analysis of these partitions provided a better characterization of ROIs and was of additional benefit in distinguishing benignancy versus

malignancy lymph nodes compared with whole-lesion ADC values alone. The use of cut-off value at 1.12×10^{-3} $mm^2.s^{-1}$ to distinguish benign from malignant segmented ROIs gave motivating results with the SE, SP, and AUC were 94.12%, 89.19 %, and 0.972, respectively when we used ROI' placed in the solid partitions. Finally, these results of the developed proposed CAD system confirm the superiority of the proposed method with accuracy reaching 93.43% which is higher than 92.83% obtained by using the SVM classifier.

Acknowledgments

The authors would like to thank the department of radiology of the Tunisian National Hospital Aziza Othmana-Tunis, and the Tunisian National Hospital for Children Bechir Hamza-Tunis for their support and for providing us the MRI-exams used in this work.

Competing Interests

Authors declare that they have no conflict of interest.

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136:E359-E386.
2. Armitage JO. Staging non-Hodgkin lymphoma. CA-Cancer J Clin. 2005;55:368-376.
3. Connors JM. State-of-the-art therapeutics: Hodgkin's lymphoma. J Clin Oncol. 2005;23:6400-6408.
4. Tubiana M, Aurengo A, Averbeck D, Masse R. Recent reports on the effect of low doses of ionizing radiation and its dose–effect relationship. Radiat Environ Bioph. 2006;44:245.
5. Huang B, Law MW, Khong PL. Whole-body PET/CT scanning: estimation of radiation dose and cancer risk. J Radiol. 2009;251:166-174.
6. Plathow C, Walz M, Lichy MP, Aschoff P, Pfannenberg C, Bock H, et al. Cost considerations for whole-body MRI and PET/CT as part of oncologic staging. J Radiol. 2008;48:384-396.
7. van Ufford HM, Kwee TC, Beek FJ, van Leeuwen MS, Takahara T, Fijnheer R, et al. Newly diagnosed lymphoma: initial results with whole-body T1-weighted, STIR, and diffusion-weighted MRI compared with 18F-FDG PET/CT. Am J Roentgenol. 2011;196:662-669.
8. Lin C, Luciani A, Itti E, El-Gnaoui T, Vignaud A, Beaussart P, et al. Whole-body diffusion-weighted magnetic resonance imaging with apparent diffusion coefficient mapping for staging patients with diffuse large B-cell lymphoma. Eur radiol. 2010;20:2027-2038.
9. Yap E, Law ZK, Abdullah NM, Wahid SF. Consolidation radiotherapy for advanced-stage aggressive B-cell non-Hodgkin lymphoma: A systematic review and meta-analysis. Excli J. 2017;16:1233.
10. Le Bihan D. Apparent diffusion coefficient and beyond: what diffusion MR imaging can tell us about tissue structure. 2013
11. Zhang Y, Chen J, Shen J, Zhong J, Ye R, Liang B. Apparent diffusion coefficient values of necrotic and solid portion of lymph nodes: differential diagnostic value in cervical lymphadenopathy. Clin Radiol. 2013;68:224-231.
12. Mendonca AM, Campilho A. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. IEEE Trans Med Imaging. 2006;25:1200-1213.
13. Chan TF, Vese LA. Active contours without edges. Ieee T Image Process. 2001;10:266-277.
14. Mumford DB, Shah J. Optimal approximations by piecewise smooth functions and associated variational problems. Commun Pur Appl Math. 1989.
15. Kwee TC, Basu S, Torigian DA, Nievelstein RA, Alavi A. Evolving importance of diffusion-weighted magnetic resonance imaging in lymphoma. PET Clin. 2012;7:73-82.
16. Nazeer KA, Sebastian MP. Improving the Accuracy and Efficiency of the k-means Clustering Algorithm. Proc World Congr Eng 2009;1:1-3
17. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988:837-845.
18. Ferjaoui R, Cherni MA, Kraiem NE, Kraiem T. Lymphoma Lesions Detection from Whole Body Diffusion-Weighted Magnetic Resonance Images. In 2018 5th International Conference on Control, Decision and Information Technologies (CoDIT) 2018;364-369.
19. Kamdi S, Krishna RK. Image segmentation and region growing algorithm. Int J comput (IJCTEE). 2012;2:103-107.
20. Razek AA, Soliman NY, Elkhamary S, Alsharaway MK, Tawfik A. Role of diffusion-weighted MR imaging in cervical lymphadenopathy. Eur Radiol. 2006;16:1468-1477.
21. Lee MC, Chuang KS, Chen MK, Liu CK, Lee KW, Tsai HY, et al. Fuzzy C-means clustering of magnetic resonance imaging on apparent diffusion coefficient maps for predicting nodal metastasis in head and neck cancer. Brit J Radiol. 2016;89:20150059.
22. Holzapfel K, Duetsch S, Fauser C, Eiber M, Rummeny EJ, Gaa J. Value of diffusion-weighted MR imaging in the differentiation between benign and malignant cervical lymph nodes. Eur J Radiol. 2009;72:381-387.
23. Cintra MB, Ricz H, Mafee MF, Santos AC. Magnetic resonance imaging: dynamic contrast enhancement and diffusion-weighted imaging to identify malignant cervical lymph nodes. Radiologia brasileira. 2018;51:71-75.