Computerized Medical Imaging and Graphics 73 (2019) 11–18

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

A decision support tool for early detection of knee OsteoArthritis using X-ray imaging and machine learning: Data from the OsteoArthritis Initiative

Abdelbasset Brahim[a,*], Rachid Jennane[a], Rabia Riad[a], Thomas Janvier[a], Laila Khedher[a], Hechmi Toumi[a,b], Eric Lespessailles[a,b]

[a] University of Orléans, I3MTO Laboratory, EA 4708, 45067 Orléans, France
[b] Hospital of Orléans, Rheumatology Department, 45032 Orléans, France

ARTICLE INFO

Article history:
Received 27 October 2017
Received in revised form 14 March 2018
Accepted 23 January 2019

Keywords:
Intensity normalization
Computer aided diagnosis system
Classification
OsteoArthritis

ABSTRACT

This paper presents a fully developed computer aided diagnosis (CAD) system for early knee OsteoArthritis (OA) detection using knee X-ray imaging and machine learning algorithms. The X-ray images are first preprocessed in the Fourier domain using a circular Fourier filter. Then, a novel normalization method based on predictive modeling using multivariate linear regression (MLR) is applied to the data in order to reduce the variability between OA and healthy subjects. At the feature selection/extraction stage, an independent component analysis (ICA) approach is used in order to reduce the dimensionality. Finally, Naive Bayes and random forest classifiers are used for the classification task. This novel image-based approach is applied on 1024 knee X-ray images from the public database OsteoArthritis Initiative (OAI). The results show that the proposed system has a good predictive classification rate for OA detection (82.98% for accuracy, 87.15% for sensitivity and up to 80.65% for specificity).

1. Introduction

OsteoArthritis (OA) is a progressive joint disease causing pain, swelling, stiffness and loss of function (Goldring and Goldring, 2006; Sellam et al., 2009). OA is a chronic disease, whose origin is not well understood. It is considered the most common form of arthritis and is characterized by cartilage degradation and bone changes (Goldring and Goldring, 2010). At the early stage of the disease it is challenging to diagnose OA as it has been suggested that the first changes occur in the subchondral bone (Radin et al., 1970) before the occurrence of joint space narrowing and osteophytes (Wang et al., 2013). Hence, early detection, diagnosis and intervention are strongly desired to overcome this highly disabling disease since early-stage treatment could prevent the breakdown of cartilage and bone. Currently, the diagnosis of OA is based upon patient-reported symptoms and X-ray imaging. Conventional radiographic images remain the "gold standard" for the imaging diagnosis of knee OA (Hayashi et al., 2017), despite their known limitations in detecting early disease and subtle changes over time (Buckland-Wright, 2006; Guermazi et al., 2013). Moreover, this imaging technique is cheap, fast, easy to use and still provides valuable information about changes occurring in human bones (Buckland-Wright, 2006). These radiographs can depict joint space loss, as well as subchondral bony sclerosis and cyst formation (Duryea et al., 2003).

Knee OA detection consists of classifying a knee radiograph into healthy or OA (Kellgren and Lawrence, 1957). The clinical grading of X-ray images is currently performed by experienced clinicians using several methods, such as the Kellgren–Lawrence (KL) score (Kellgren and Lawrence, 1957). This approach is based on the classification of individual joints into one of five grades, with 0 representing normal and 4 being severe OA (Kellgren and Lawrence, 1957). Individual grading is time consuming, can be subjective and prone to error, since the symptoms involved in OA classification are continuous. Hence, the assessment of OA by the human reader may differ from one clinician to another. To make a valid OA diagnosis, a considerable amount of knowledge and experience is required. Consequently, there is a growing interest in developing a low-cost and non invasive image-based computer aided diagnosis (CAD) system for the early detection of knee OA (Kraiger et al., 2012). Such a bone-based decision support system, when fully developed, would provide an objective and observer-independent quantification of OA. This diagnosis tool could be used for clinical studies

* Corresponding author.
E-mail addresses: brahim.abdelbasset@yahoo.fr, abdelbasset.brahim@univ-orleans.fr (A. Brahim).

https://doi.org/10.1016/j.compmedimag.2019.01.007
0895-6111/

such as the evaluation of medication, intra-articular injections, or surgical interventions that have an effect on the progression of OA. Several approaches for the detection and analysis of OA using different knee images have been proposed in the literature (Wu et al., 2014; Shamir et al., 2009a,b; Woloszynski et al., 2010; Janvier et al., 2017a,b; Antony et al., 2016; Stachowiak et al., 2016; Kotti et al., 2017). Subtle differences in textures and intensity variations within an X-ray knee image can be detected by computer-aided image analysis methods, without clinical bias. These approaches have been employed to analyze texture and intensity changes in the bone of the knee, as reviewed in the following paragraph.

A system that analyses the textural information across the entire joint was proposed in Shamir et al. (2009), and was validated on 193 plain radiographs. The classification accuracy for discrimination between KL grade 0 (non-OA) and KL grade 2 (minimal OA) was 80.4%. In Woloszynski et al. (2010), the texture analysis was performed on the trabecular bone (TB). The authors developed an approach for TB texture classification in knee radiographs using a signature dissimilarity measure (SDM). Further analysis revealed that the SDM method is suitable for TB texture classification in knee OA detection and that it can be of interest for the texture classification of medical images in general. This computer-aided method was validated on 137 X-ray images (68 CC and 69 OA) and achieved an accuracy rate of 78.8%. In Antony et al. (2016), the authors proposed a new automatic method which is able to quantify the severity of knee OA from X-ray imaging using deep convolutional neural networks (CNN). In addition, an automatic method to extract the knee joint images was used. Their proposed segmentation method was based on using a linear support vector machine (SVM) with the Sobel horizontal image gradients as features to detect the knee joints, as regions of interest. Their results were validated on a dataset of X-ray images from the OsteoArthritis Initiative (OAI) and showed a sizable improvement. The accuracy rate achieved was 77.6% to classify non-OA (KL = 0) from minimal OA (KL = 2). In Kotti et al. (2017), an automatic detection of knee OA is presented. The authors developed a computer system that is based on the body kinetics. They validated their proposed system on a dataset of 94 subjects (47 CC and 47 OA) and the system achieved a mean accuracy of $72.61\% \pm 4.24\%$.

The early detection of knee OA is still challenging today as the above-mentioned methods do not satisfy the clinical constraints of being highly efficient on a large amount of data, since they were validated on limited and private databases. Moreover, almost all these methods suffer either from the problem of computational complexity, such as the computation of Zernike features in Shamir et al. (2009), the complexity of the model as in Antony et al. (2016) or its sensitivity to image size when using a global approach based on similarity patterns (Woloszynski et al., 2010).

In this paper, a novel image-based CAD system (Fig. 1) is proposed that uses computational methods from machine learning to early detect knee OA. The system comprises five stages. First, three regions of interest (ROI) in the medial, middle and lateral sides of size 128×128 pixels are extracted from the knee radiograph. These ROIs are then preprocessed in the Fourier domain using circular filtering in order to stationarize the X-ray images. To keep the essential information of the data, a quantization of the gray level is next performed. In the following stage, a novel method of intensity normalization, which is based on predictive modeling using multivariate linear regression (MLR) is applied to the data in order to reduce the variability between the subjects. Lastly, in order to reduce the dimensionality of the normalized sub-images, independent component analysis (ICA) is used. Thus, each ROI is represented by few components and the proposed system will not be affected by the small sample size problem. The resulting features will be used for the classification task using the Naive Bayes and the random forest classifiers.

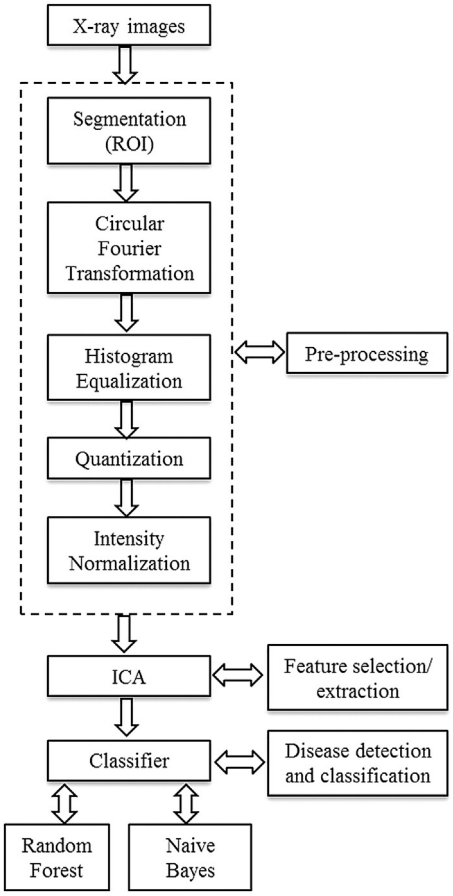

Fig. 1. The different modules of the proposed CAD system.

Our approach provides a different insight from previous methods (Wu et al., 2014; Shamir et al., 2009a,b; Woloszynski et al., 2010; Janvier et al., 2017a,b; Antony et al., 2016; Stachowiak et al., 2016; Kotti et al., 2017), namely: (i) a novel preprocessing step, based on predictive modeling using MLR is employed to reduce inter-subject variability; (ii) there are no descriptors, and no image transforms or texture analysis are performed to detect the OA disease. Thus, in this proposed approach, only the pixel intensities are used directly as features to train the classifiers; (iii) except for the segmentation process, which is semi-automatic to extract the ROIs, the proposed bone-based decision support system for the early detection of OA is fully automatic; and (iv) finally, the proposed aided diagnosis method is validated on a large multi-centric dataset in which different methods of radiographic acquisition were used such as radiographic imaging and computed radiography (Initiative, 2013).

The paper is organized as follows: Section 2 presents the X-ray image database used in this work and details the different modules of the proposed CAD system. The experimental results and discussion are given in Section 3 while some conclusions are drawn in Section 4.

2. Materials and methods

2.1. Osteoarthritis Initiative (OAI) Database

The radiographic knee images used in this study were collected from the public dataset OsteoArthritis Initiative (OAI) (Lester, 2008; Eckstein et al., 2012), which is available for public access at "http://www.oai.ucsf.edu/". The OAI is a multi-centric, longitudinal, prospective observational study of knee OA. The main goal of

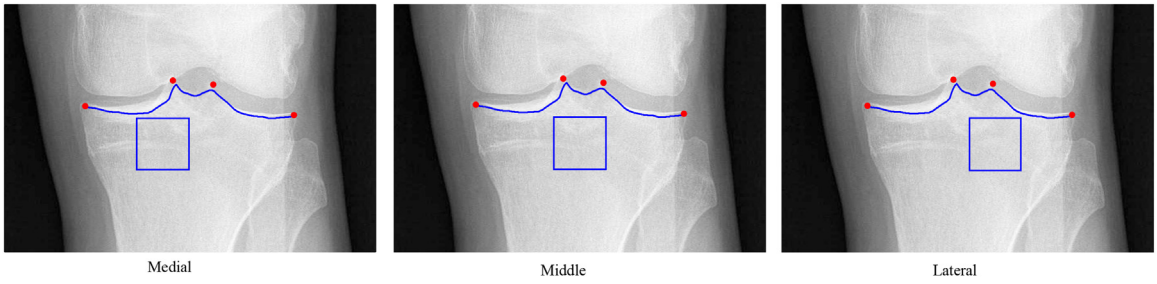

Fig. 2. The Trabecular bone ROIs extracted for each side of the knee. The red points are the manually defined anatomical markers, the blue line is the estimated tibial edge and the blue boxed areas show the ROIs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the OAI is to develop a public domain research resource to facilitate the scientific evaluation of biomarkers for knee OA as potential surrogate endpoints for disease onset and progression (Galván-Tejada et al., 2014). The OAI database contains bilateral fixed flexion knee radiographs of 4796 men and women followed since 2008. Their ages are between 45 and 79.

The present study focuses on the early detection of knee OA, which is the most challenging. Therefore, only the radiographs with KL grade = 0 (non-OA, considered as CC) and KL grade = 2 (minimal OA, considered as OA) were taken into consideration. Thus, 1024 knee radiographs were used in this study and to avoid any statistical bias of imbalanced datasets, the same number of knees was considered for both groups: 514 knees from CC subjects and 514 knees from OA subjects.

2.2. Segmentation

In order to extract the tibial trabecular bone ROIs, a semi-automatic segmentation method was used (Janvier et al., 2017a). This approach is based on marking the tibial spines and the lateral and medial extremities of the tibia. The algorithm then approximates the tibial subchondral bone plate as the brightest path going through these anatomical markers. Vertical adjustment was carried out in such a way that the upper boundaries of the medial and lateral compartments were placed immediately below the lowest point on the tibial edge, while the horizontal adjustment of the ROI position was defined as the center between knee extremities. Finally, three ROIs are extracted on each knee for the medial, middle and lateral sides with a fixed size of 128 × 128 pixels as depicted in Fig. 2. X-ray images were read from the database along with the information whether they represent the left or right knee. Based on this information, the ROIs in the right knees were mirrored horizontally to be comparable to the ROIs in the left knees.

2.3. Pre-processing

The OAI database needs to be preprocessed previous to the analysis and interpretation of the results. The Pre-processing step is relevant to extract or enhance significant data characteristics and to prepare the database for the application of data analysis approaches. Thus, data pre-processing is a critical step in order to obtain meaningful results. Several computational methods were used, as described in the following subsections.

2.3.1. Circular Fourier Filtering (CFF) of the data

After the segmentation process, each sub-image, i.e. ROI, was filtered in the Fourier domain in order to keep only the essential information related to the TB structure and avoid the information related to the acquisition process. In other words, this step aims to stationarize the data (Houam et al., 2014). Due to the complexity of the TB structure, the mean periodograms performed on the lines of the ROIs show two distinct regions separated by a cut-off frequency f_c, as depicted in Fig. 3. The physical meaning of these two frequency regions is that the low frequencies ($|f| < f_c$) correspond to the region which is not affected by osteoarthritis and belong to inter-trabecular spacing, whereas the high frequencies ($|f| > f_c$) correspond to the area concerned by osteoarthritis, i.e. the TB area. Thus, in order to characterize only the TB area in the X-ray sub-images satisfactorily, a filtering process in the Fourier domain was used for each image. Hence, based on Fig. 3, the first 60 rays of the Fourier spectrum were removed before performing the inverse Fourier transform and only the high frequency part of each ROI was considered. Fig. 4 illustrates the resulting images for each class after the filtering process using CFF.

2.3.2. Histogram equalization and quantization

In order to enhance the tibial trabecular bone patterns and to keep the relevant imaging information, a quantization process over fewer gray levels was performed. However, the contrast of the source images is low. Thus, a nonlinear approach, based on histogram equalization (HE) was applied to the filtered sub-images in order to improve the contrast by increasing the dispersion of the highest frequency and decreasing that of other frequencies. Following HE, the number of gray levels was reduced to 16, which provides better and more easily exploitable images that are suitable for TB analysis (Houam et al., 2014). The HE and quantization process are shown for a given subject from each class in Fig. 4.

2.3.3. Intensity normalization using multivariate linear regression (MLR)

In general, MLR is based on the following statistical model (Heij et al., 2004) using a matrix notation:

$$Y = X\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1}$$

where $\mathbf{Y}$ is a vector of the response variables, $\mathbf{X}$ is a vector of the regressor values, $\boldsymbol{\beta}$ is the vector of the regression coefficient and $\boldsymbol{\epsilon}$ is the vector of the error terms. In this paper, MLR is applied for intensity normalization in order to reduce the inter-subject variability and artifacts in ROIs. Intensity normalization is a relevant process in any image analysis with more than one subject to ensure comparability within and across the images of the different subjects (Gonzalez and Woods, 2002; Brahim et al., 2015a). Several assumptions have to be made in Eq. (1) to perform the normalization task of each ROI:

- Let the number of observations N be the number of pixels for each ROI.
- Let $\bar{\mathbf{y}} \in \mathbf{Y}$, an $N \times 1$ vector of observations on the dependent variable, that is an ROI from a template that will be defined later (Eq. (3)).
- Let $\mathbf{x} \in \mathbf{X}$, an $N \times 1$ vector of regressors where the N observations are on one independent variable, that is the ROI of each image

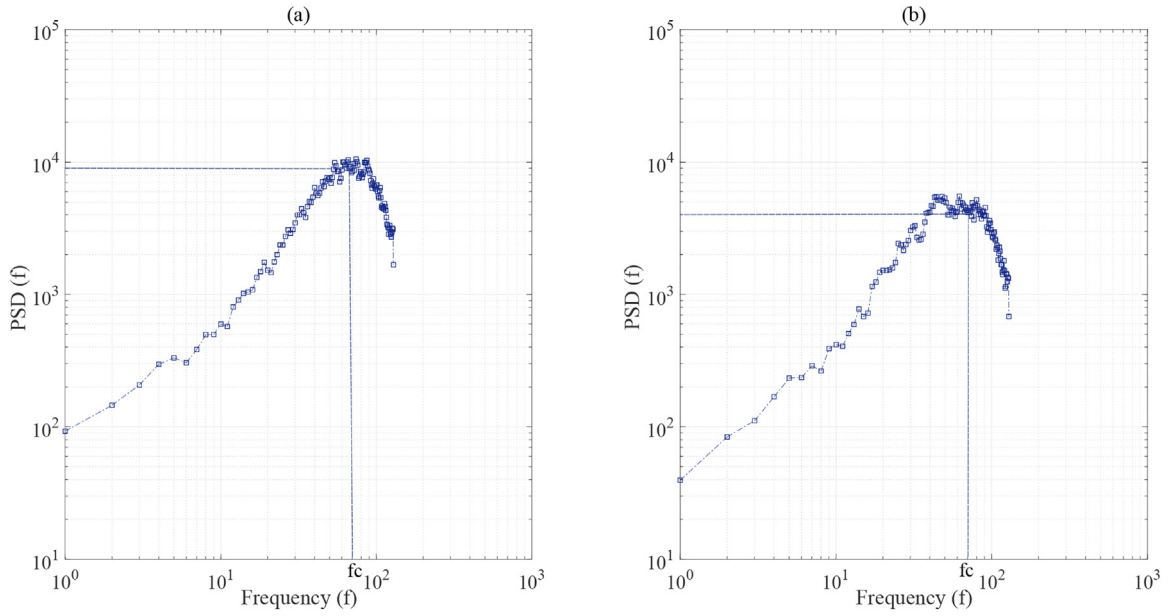

Fig. 3. Two representative mean periodograms computed over the lines of two ROIs. For a CC subject (a), for an OA patient (b).

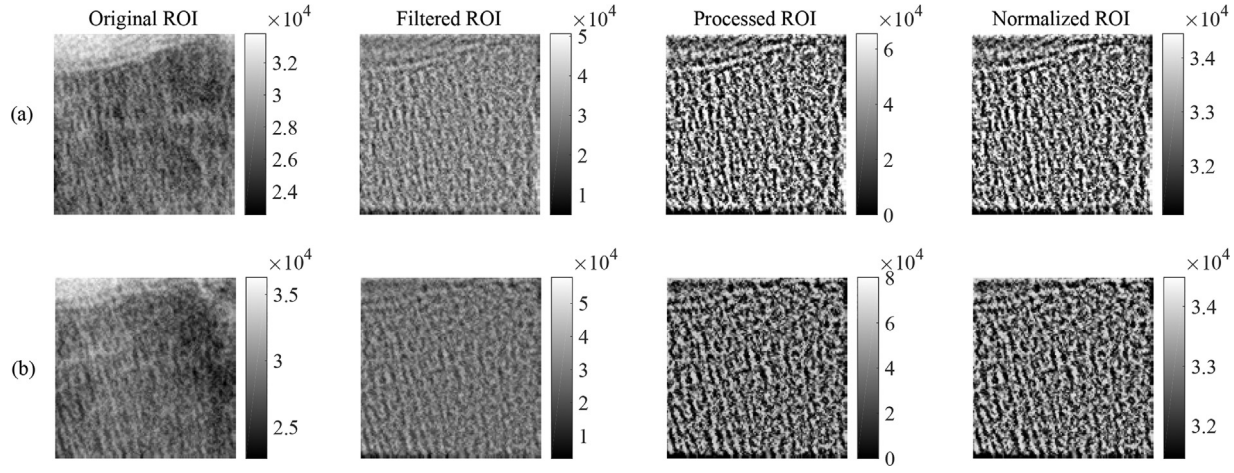

Fig. 4. A representative ROI of the tibial trabecular bone, Filtered ROI using CFF, Processed ROI using HE and quantization and Normalized ROI using MLR: (a) is from a CC subject and (b) is from an OA patient.

subject $\mathbf{X}$ ($\mathbf{x}$ is either the medial, middle or lateral side of each knee).
- Let $\boldsymbol{\beta}$ be a 1×1 vector of an unknown parameter that has to be estimated and $\boldsymbol{\epsilon}$ be an $N \times 1$ vector of errors.

Therefore, by estimating $\boldsymbol{\beta}$ and $\boldsymbol{\epsilon}$ values, a linear intensity transformation is used to transform the intensity distribution of $\mathbf{x}$ to $\hat{\mathbf{x}}$ using the following expression:

$$\hat{\mathbf{x}} = \mathbf{x}a + b \tag{2}$$

where $\hat{\mathbf{x}}$ is the normalized image, $a = \boldsymbol{\beta}$ and $b = \bar{\boldsymbol{\epsilon}}$ are the intensity normalization parameters, which represent the scale and offset of the intensity transformation (Brahim et al., 2015b). Further details about the computation of these parameters are given in A.

Finally, the intensity normalization, which is based on a linear transformation of the intensity values for the different ROIs using predictive modeling based on MLR is summarized as follows:

- First, the template $\bar{\mathbf{y}}$ is computed as:

$$\bar{\mathbf{y}} = \frac{1}{N_{cc}} \sum_{i \in X_{cc}} \mathbf{x}_i(x, y) \tag{3}$$

where X_{cc} denotes the subset of CC ROIs, N_{cc} is the number of CC ROIs, $\mathbf{x}_i(x, y)$ is the ith ROI.
- Second, the intensity normalization parameters $\boldsymbol{\beta}$ and $\boldsymbol{\epsilon}$ of the MLR model are estimated using Eqs. (A.5) and (A.6).
- Finally, the normalized ROIs are computed by linearly transforming the pixel intensities of each ROI using the model in Eq. (2).

A normalized ROI using MLR for each class is depicted in Fig. 4 (Last column).

2.4. Feature selection/extraction and classification stages of the proposed CAD system

In order to extract the most significant features from the ROIs, Independent Component Analysis (ICA) (Hyvärinen and Oja, 2000; Li and Yin, 2013; Khedher et al., 2017) based feature extraction was

used. This approach is employed as a feature extraction strategy to avoid the small sample size problem (Khedher et al., 2017). In this paper, the number of samples is 1024 images, and the number of features is (128 × 128) pixels. Thus, a reduction of the input vector is desirable. Using ICA, the first 10 discriminant components were retained and were used for classification.

The classification performances of the proposed CAD were tested using random forest and Naive Bayes classifiers (Lee et al., 2010; Martínez-Murcia et al., 2013) and the Leave-One-Out (LOO) cross validation strategy (Brahim et al., 2015b). Two different classifiers are used to test the effectiveness of the proposed normalization method in order to keep the research finding more objective. Several statistics were computed: accuracy (Acc), sensitivity (Sen) and specificity (Spe) (Brahim et al., 2015a; Imani et al., 2015).

3. Results and discussion

3.1. Qualitative and statistical analysis

A data set of 1024 knee radiographs was used to assess the performance of the proposed system developed for the early detection of knee OA. The visual analysis of OA and CC bone images, as illustrated in Fig. 4, shows no visual difference, making the discrimination task very challenging. In order to reduce the variability of pixel intensity between the different subjects, as illustrated in Fig. 5.(a), a novel normalization method based on the MLR was applied on the different ROIs. To assess the effectiveness of the proposed normalization in a quantitative analysis, different statistical metrics that quantify the difference between two sets of samples, such as the Kullback-Leibler divergence (KLD) (Kullback, 1959) and the Jeffreys divergence (JD) (Jeffreys, 1946) were used. In our case, the set of samples to be evaluated with these measures is the histogram of the normalized ROI, referred to the histogram of a reference target ROI (typically, the mean ROI of the CC subjects).

When these divergence measures have lower values, this means less discrepancy between these two histogram distributions. The difference between the probability distribution of each ROI ($\mathbf{x}$) denoted by Q and the probability distribution of the template ($\bar{\mathbf{y}}$) denoted by P was computed for all subjects before and after intensity normalization using KLD and JD divergences, which are defined as (Kullback, 1959; Jeffreys, 1946):

$$KLD(P\|Q) = \sum_{i=1}^{n} \ln\left(\frac{P(i)}{Q(i)}\right)(P(i)) \tag{4}$$

$$JD\ (P\|Q) = \sum_{i=1}^{n} P(i)\ \ln\left(\frac{P(i)}{Q(i)}\right)\ d(i) + \sum_{i=1}^{n} Q(i)\ \ln\left(\frac{Q(i)}{P(i)}\right)\ d(i) \tag{5}$$

where n is a fixed number of bins, which can be chosen arbitrarily.

The computation of the inter-subject distance is quantitatively calculated between these two distributions, both before (Raw data) and after normalization (Normalized data) for all subjects and for the different sites of the tibia in Table 1. As can be seen, the lowest KLD and JD values and the lowest errors were obtained (in terms of standard deviation) using the data normalized by MLR when compared to the raw data. This table shows that the inter-subject variability in terms of intensity values are quantitatively mitigated after the normalization process by the proposed approach.

Moreover, after the normalization process, the separation between the two classes is increased, as shown in Fig. 5(b). Hence, the classification task between CC and OA groups after the preprocessing step should be improved when compared to raw data. In the following section, the classification performances of the proposed CAD system will be evaluated.

Table 1
Mean Kullback-Leibler and Jeffreys divergences and standard deviation for the ROIs before and after the intensity normalization in the medial, middle and lateral sites of the tibia.

ROI	Dataset	Class	KLD	JD
Medial	Raw data	CC	1.421 ± 0.805	1.811 ± 0.820
		OA	2.215 ± 0.656	2.206 ± 0.673
		CC+OA	2.086 ± 0.689	2.096 ± 0.689
	Normalized data	CC	**0.029 ± 0.085**	**0.037 ± 0.100**
		OA	**0.099 ± 0.067**	**0.087 ± 0.109**
		CC+OA	**0.121 ± 0.175**	**0.109 ± 0.191**
Middle	Raw data	CC	1.431 ± 0.788	1.758 ± 0.750
		OA	2.240 ± 0.653	2.208 ± 0.671
		CC+OA	2.103 ± 0.673	2.091 ± 0.676
	Normalized data	CC	**0.037 ± 0.071**	**0.034 ± 0.112**
		OA	**0.065 ± 0.070**	**0.063 ± 0.079**
		CC+OA	**0.115 ± 0.081**	**0.053 ± 0.070**
Lateral	Raw data	CC	1.479 ± 0.827	1.807 ± 0.786
		OA	2.263 ± 0.671	2.251 ± 0.685
		CC+OA	2.130 ± 0.701	2.132 ± 0.698
	Normalized data	CC	**0.016 ± 0.071**	**0.016 ± 0.098**
		OA	**0.061 ± 0.048**	**0.053 ± 0.068**
		CC+OA	**0.083 ± 0.043**	**0.034 ± 0.054**

The bold values are for the proposed approach.

3.2. Quantitative classification performance of OA

The proposed CAD system was evaluated on its ability to distinguish OA patients from CC subjects by computing the classification rates(Acc, Sen and Spe measures) and the values of Positive and Negative Likelihood (PL and NL) (McGee, 2002). These statistical metrics give an assessment of the supervised learning system's ability to reproduce a medical diagnosis carried out by experts. The use of assessment parameters that are sample prevalence dependent can be inappropriate, however. Other statistical values such as PL and NL can be more reliable as they are prevalence independent, i.e. they do not depend on the ratio of the classes. The PL indicates by how much the probability of disease increases if the test is positive, while the NL indicates by how much the probability decreases if the test is negative.

As can be seen in Table 2, the classification results obtained by the proposed system with two different classifiers outperform those obtained using the raw data. The proposed model of pixel-wise classification for the medial ROI achieves a gain of 11.38% for accuracy, 21.51% for specificity and 2.72% for sensitivity using the random forest classifier. For the middle and lateral ROIs, it achieves a gain of 9.63% and 10.12% for accuracy, 19.27% and 21.02% for specificity, 1.01% and 0.79% for sensitivity. There is also a remarkable improvement in the classification rates when the Naive Bayes classifier is used. Our findings also demonstrate that the different ROIs could be a discriminating region for the early detection of OA. Furthermore, there is a slight increase in the classification rates when using ROIs located immediately under the medial compartment.

In addition, the likelihood ratios are one of the best ways to measure and express the diagnosis accuracy. Clinicians can use them for a quick comparison between different diagnostic strategies and thus refine their clinical judgment. In clinical medicine, values of PL greater than 5 or NL values less than 0.2 can be applied to the pre-test probability of a patient having the disease tested in order to estimate a post-test probability of the disease state existing (McGee, 2002). These statistical values are computed in

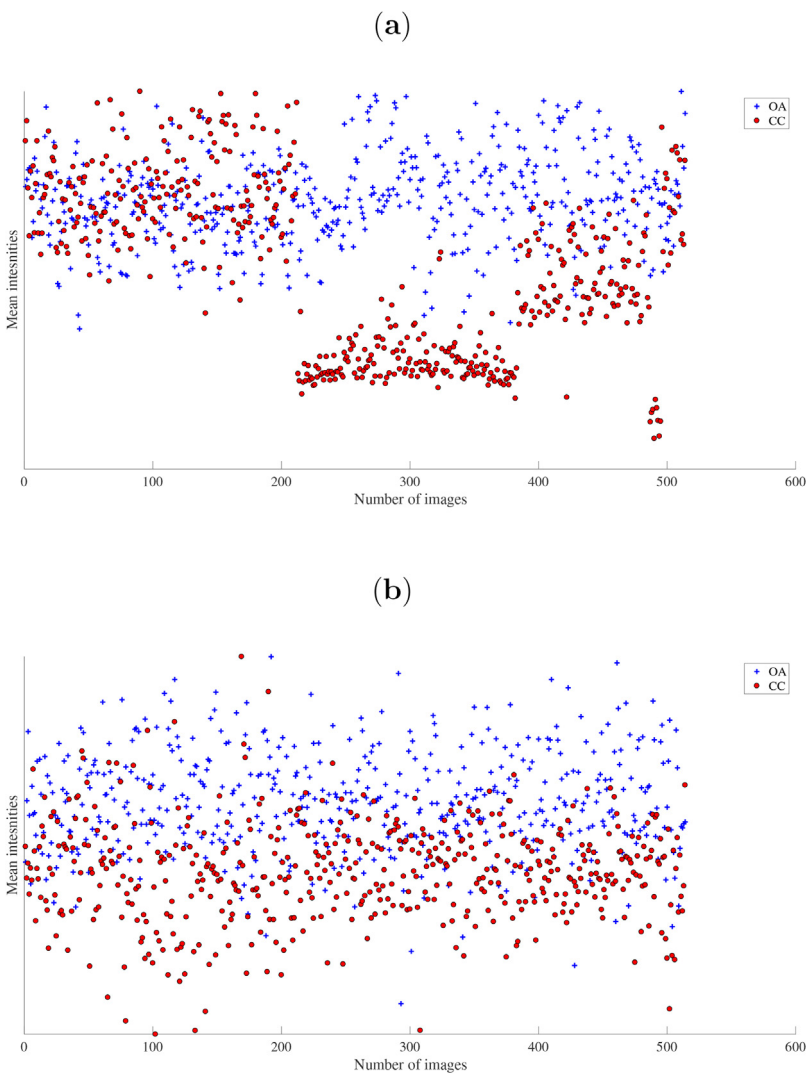

Fig. 5. Distribution of the mean intensities of CC and OA subjects in the medial region before (a) and after normalization (b). The *x*-axis represents the number of images for each class. The *y*-axis indicates the mean intensity.

Table 2
The classification rates, PL and NL achieved with the raw data and the proposed CAD system in the different ROIs using random forest and naive bayes classifiers.

Classifier	Random Forest									
Statistical measures	Acc (%)		Sen (%)		Spe (%)		PL		NL	
Data	Raw	Normalized	Raw	Normalized	Raw	Normalized	Raw	Normalized	Raw	Normalized
Medial	71.60	**82.98**	84.05	**87.15**	59.14	**80.65**	2.06	**6.50**	0.27	**0.16**
Middle	72.76	**82.39**	85.68	**86.77**	58.75	**78.02**	2.10	**7.24**	0.23	**0.18**
Lateral	72.37	**82.49**	85.40	**86.19**	58.56	**79.57**	2.08	**5.21**	0.24	**0.17**
Classifier	Naive Bayes									
Statistical measures	Acc (%)		Sen (%)		Spe (%)		PL		NL	
Data	Raw	Normalized	Raw	Normalized	Raw	Normalized	Raw	Normalized	Raw	Normalized
Medial	70.62	**82.49**	86.38	**87.93**	53.31	**79.57**	1.88	**4.33**	0.23	**0.15**
Middle	70.33	**82.20**	85.41	**87.55**	53.11	**78.99**	1.87	**4.16**	0.23	**0.16**
Lateral	69.75	**82.10**	85.21	**86.38**	53.11	**78.02**	1.84	**3.93**	0.26	**0.17**

The bold values are for the proposed approach.

Table 2. The PL values of the proposed aided diagnosis method are greater than the value considered to be a good class discriminant, and the NL values are lower than 0.2, also considered as a good value. The higher PL values and the lower NL values are provided by the processed sub-images, using the proposed system when compared to the raw data for the different classifiers. Thus, the proposed CAD system can make a significant contribution to OA diagnosis.

The proposed aided diagnosis method was compared with several approaches for OA diagnosis from the literature, as shown

Table 3
The classification rates of the proposed system compared to existing studies in the literature.

Approaches	ROIs	Number of X-ray images	Acc (%)	Sen (%)	Spe (%)
The porposed CAD system	Medial, Middle and Lateral Knee sides	1024: (514 CC+514 OA)	**82.98 ± 2.12**	**87.15 ± 4.25**	**80.65 ± 1.42**
Shamir et al. (2009a)	Joint Areas	193: (154 CC + 39 OA)	80.4	81.4	79.3
Woloszynski et al. (2010)	Trabecular Bone (different parts)	137: (68 CC + 69 OA)	78.8	76.8	80.9
Antony et al. (2016)	Knee Joints	The whole OAI database	77.2	NA	NA
Kotti et al. (2017)	Body Kinetics	–	72.61 ± 4.24	NA	NA

The bold values are for the proposed approach.

in Table 3. In Shamir et al. (2009), the discrimination accuracy between KL grade 0 (non-OA) and KL grade 2 (minimal OA) was 80.4%; the sensitivity and specificity were 81.4% and 79.3% respectively. In Woloszynski et al. (2010), the SDM system achieved an accuracy of 78.8%, a sensitivity of 76.8% and a specificity of 80.9%. In Antony et al. (2016) and Kotti et al. (2017), their proposed computer-aided methods achieved accuracy values up to 77.2% and 72.61%, respectively. Thus, the classification rates of the proposed CAD are higher than those obtained in previous studies (Shamir et al., 2009a; Woloszynski et al., 2010; Antony et al., 2016; Kotti et al., 2017) for the early detection of knee OA. Furthermore, our CAD was validated on a large dataset. The statistical significance of the proposed CAD system results was measured using a non-parametric test, i.e. the Mann–Whitney U test. The p-value was 0.0108 under the medial compartment, which is considered statistically significant ($p < 0.05$). In addition, for medical applications, the slightly higher sensitivity than specificity of the proposed system could be advantageous. This is because the cost of treating OA at later stages is usually higher than the cost of misdiagnosing a non-progressive knee. The results obtained reveal the potential of the proposed system as a decision support system for the early detection of knee OA based on TB analysis. Moreover, this CAD system avoids the subjectivity and importance of operator expertise involved in manual operations, and is able to make predictions on unseen data.

4. Conclusion

This study has introduced an automated computer-aided diagnosis approach for the detection of OA, which makes use of a combination of normalization based on predictive modeling using MLR and feature extraction using ICA. The normalization using MLR allowed us not only to reduce the intersubject variability, but also to increase the separation between CC and OA groups. Further analysis revealed that the proposed system can provide high-classification performance in distinguishing between healthy and osteoarthritic patients in different knee sides. In performance terms, it outperforms several methods proposed in the literature. We hypothesize that such an analysis approach could serve for the early detection of OA. However, the generalization capabilities of the proposed method could be overestimated and should be handled with caution. First, the X-ray devices were not calibrated and neither the variations in the acquisition parameters, nor the impact of the soft tissue in the X-ray absorption and diffusion were taken into account. Second, age and gender were not taken into account. The multicentric data used underline the robustness of our approach since we were able to detect early knee OA whatever the center or the various acquisition devices used.

Acknowledgments

The authors would like to thank the OAI study participants and clinical staff as well as the coordinating center at UCSF. The authors wish to thank the ANR for financial support under the project ANR-12-TECS-0016-01.

Appendix A. Computation of the intensity normalization parameters

In linear algebra terms, and according to Eq. (1) and the assumptions above, the minimization of the residual sums of squares (RSS) is the criterion used to obtain an estimation of the parameter $\boldsymbol{\beta}$, which can be defined as:

$$\begin{aligned}\sum_{i=1}^{N}\epsilon_i^2 &= \epsilon^T\epsilon \\ &= (\bar{\mathbf{y}}-\mathbf{x}\boldsymbol{\beta})^T(\bar{\mathbf{y}}-\mathbf{x}\boldsymbol{\beta}) \\ &= \bar{\mathbf{y}}^T\bar{\mathbf{y}}-\bar{\mathbf{y}}^T\mathbf{x}\boldsymbol{\beta}-\boldsymbol{\beta}^T\mathbf{x}^T\bar{\mathbf{y}}+\boldsymbol{\beta}^T\mathbf{x}^T\mathbf{x}\boldsymbol{\beta} \\ &= \bar{\mathbf{y}}^T\bar{\mathbf{y}}-2\boldsymbol{\beta}^T\mathbf{x}^T\bar{\mathbf{y}}+\boldsymbol{\beta}^T\mathbf{x}^T\mathbf{x}\boldsymbol{\beta}\end{aligned} \tag{A.1}$$

where $\epsilon^T, \bar{\mathbf{y}}^T, \boldsymbol{\beta}^T, \mathbf{x}^T$ are the transpose of the vector of error ϵ, the transpose of the ROI $\bar{\mathbf{y}}$ of a template, the transpose of the unknown parameter $\boldsymbol{\beta}$ and the transpose of the ROI of each subject $\mathbf{x}$, respectively.

In order to estimate $\boldsymbol{\beta}$ that minimizes the RSS, the derivative of Eq. (A.1) is calculated with respect to $\boldsymbol{\beta}$. The equation obtained is:

$$\frac{\partial\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}}{\partial\beta} = -2\mathbf{x}^T\bar{\mathbf{y}}+2\mathbf{x}^T\mathbf{x}\beta = 0 \tag{A.2}$$

From Eq. (A.2), the "normal equation", associated with the RSS problem is:

$$(\mathbf{x}^T\mathbf{x})\beta = \mathbf{x}^T\bar{\mathbf{y}} \tag{A.3}$$

After the multiplication of both sides of Eq. (A.3) by the inverse $(\mathbf{x}^T\mathbf{x})^{-1}$, the equation becomes:

$$(\mathbf{x}^T\mathbf{x})^{-1}(\mathbf{x}^T\mathbf{x})\beta = (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\bar{\mathbf{y}} \tag{A.4}$$

By definition, $(\mathbf{x}^T\mathbf{x})^{-1}(\mathbf{x}^T\mathbf{x}) = \mathbf{I}_N$, where $\mathbf{I}_N$ is the identity matrix and N is the number of pixels. As a result, the least square solution, $\boldsymbol{\beta}$ is:

$$\beta = (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\bar{\mathbf{y}} \tag{A.5}$$

Then, according to Eq. (1) and using the above assumptions, the residuals are computed as:

$$\boldsymbol{\epsilon} = \bar{\mathbf{y}}-\mathbf{x}\beta = \bar{\mathbf{y}}-\mathbf{x}(\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\bar{\mathbf{y}} = (\mathbf{I}_N-\mathbf{x}(\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T)\bar{\mathbf{y}} \tag{A.6}$$

References

Antony, J., McGuinness, K., O'Connor, N.E., Moran, K., 2016. Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks. 2016 23rd International Conference on Pattern Recognition (ICPR), 1195–1200.

Brahim, A., Górriz, J., Ramírez, J., Khedher, L., Salas-Gonzalez, D., 2015a. Comparison between different intensity normalization methods in 123 I-ioflupane imaging for the Automatic Detection of Parkinsonism. PLOS ONE 10 (7), 1–20.

Brahim, A., Górriz, J., Ramírez, J., Khedher, L., 2015b. Intensity normalization of DaTSCAN SPECT imaging using a model-based clustering approach. Appl. Soft Comput. 37, 234–244.

Buckland-Wright, C., 2006. Which radiographic techniques should we use for research and clinical practice? Best Pract. Res. Clin. Rheumatol. 20 (1), 39–55.

Duryea, J., Zaim, S., Genant, H.K., 2003. New radiographic-based surrogate outcome measures for osteoarthritis of the knee. Osteoarthr. Cartil. 11 (2), 102–110.

Eckstein, F., Wirth, W., Nevitt, M.C., 2012. Recent advances in osteoarthritis imaging – the Osteoarthritis Initiative. Nat. Rev. Rheumatol. 8, 622–630.
Galván-Tejada, J.I., Trevino, V., Celaya-Padilla, J.M., Tamez-Pena, J.G., 2014. Knee osteoarthritis pain prediction from X-ray imaging: Data from osteoarthritis initiative. 2014 International Conference on Electronics, Communications and Computers (CONIELECOMP), 194–199.
Goldring, S.R., Goldring, M.B., 2006. Clinical aspects, pathology and pathophysiology of osteoarthritis. J. Musculoskelet. Neuronal Interact. 6 (4), 376–378.
Goldring, M.B., Goldring, S.R., 2010. Articular cartilage and subchondral bone in the pathogenesis of osteoarthritis. Ann. N. Y. Acad. Sci. 1192, 230–237.
Gonzalez, R.C., Woods, R.E., 2002. Digital Image Processing, Second Edition. Tom Robbins.
Guermazi, A., Hayashi, D., Roemer, F.W., Felson, D.T., 2013. Osteoarthritis: a review of strengths and weaknesses of different imaging options. Rheum. Dis. Clin. N. Am. 39 (3), 567–591.
Hayashi, D., Roemer, F.W., Jarraya, M., Guermazi, A., 2017. Imaging in osteoarthritis. Radiol. Clin. N. Am. 55 (5), 1085–1102.
Heij, C., de Boer, P., Hans Franses, P., Kloek, T., van Dijk, H.K., 2004. Econometric Methods with Applications in Business and Economics.
Houam, L., Hafiane, A., Boukrouche, A., Lespessailles, E., Jennane, R., 2014. One dimensional local binary pattern for bone texture characterization. Pattern Anal. Appl. 17, 179–193.
Hyvärinen, A., Oja, E., 2000. Independent component analysis: algorithms and applications. Neural Netw. 13 (4–5), 411–430.
Imani, E., Pourreza, H.-R., Banaee, T., 2015. Fully automated diabetic retinopathy screening using morphological component analysis. Comput. Med. Imaging Graphics 43 (Supplement C), 78–88.
Initiative, O., 2013. Procedure Manual for Radiographic Examinations of the Knee, Hand, Pelvis and Lower Limbs.
Janvier, T., Jennane, R., Valery, A., Harrar, K., Delplanque, M., Lelong, C., Loeuille, D., Toumi, H., Lespessailles, E., 2017a. Subchondral tibial bone texture analysis predicts knee osteoarthritis progression: data from the Osteoarthritis Initiative Tibial bone texture & knee OA progression. Osteoarthr. Cartil. 25 (2), 259–266.
Janvier, T., Jennane, R., Toumi, H., Lespessailles, E., 2017b. Subchondral tibial bone texture predicts the incidence of radiographic knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthr. Cartil., 1–8.
Jeffreys, H., 1946. An invariant form for the prior probability in estimation problems. Proc. R. Soc. Lond. Ser. A 186, 453–461.
Kellgren, J.H., Lawrence, J.S., 1957. Radiologic assessment of osteoarthritis. Ann. Rheum. Dis. 16, 494–501.
Khedher, L., Illán, I.A., Górriz, J.M., Ramírez, J., Brahim, A., Meyer-Baese, A., 2017. Independent component analysis-support vector machine-based computer-aided diagnosis system for Alzheimer's with visual support. Int. J. Neural Syst. 27 (3), 1650050–1650118.
Kotti, M., Duffell, L.D., Faisal, A.A., McGregor, A.H., 2017. Detecting knee osteoarthritis and its discriminating parameters using random forests. Med. Eng. Phys. 43, 19–29.
Kraiger, M., Martirosian, P., Opriessnig, P., Eibofner, F., Rempp, H., Hofer, M., Schick, F., Stollberger, R., 2012. A fully automated trabecular bone structural analysis tool based on T2*-weighted magnetic resonance imaging. Comput. Med. Imaging Graphics 36 (2), 85–94.
Kullback, S., 1959. Information Theory and Statistics, Dover Books on Mathematics. John Wiley & Sons, New York.
Lee, S., Kouzani, A., Hu, E., 2010. Random forest based lung nodule classification aided by clustering. Comput. Med. Imaging Graphics 34 (7), 535–542.
Lester, G., 2008. Clinical research in OA-the NIH Osteoarthritis Initiative. Musculoskelet. Neuronal Interact. 8 (4), 313–314.
Li, C.-F., Yin, J.-Y., 2013. Variational Bayesian independent component analysis-support vector machine for remote sensing classification. Comput. Electr. Eng. 39 (3), 717–726.
Martínez-Murcia, F.J., Górriz, J.M., Ramírez, J., Puntonet, C., Illán, I.A., for the Alzheimer's Disease Neuroimaging Initiative, 2013. Functional activity maps based on significance measures and independent component analysis. Comput. Methods Prog. Biomed. 111 (1), 255–268.
McGee, S., 2002. Simplifying likelihood ratios. J. Gen. Intern. Med. 17 (8), 646–649.
Radin, E.L., Paul, I.L., Tolkoff, M.J., 1970. Subchondral bone changes in patients with early degenerative joint disease. Arthritis Rheum. 13 (4), 400–405.
Sellam, J., Herrero-Beaumont, G., Berenbaum, F., 2009. Osteoarthritis: Pathogenesis, Clinical Aspects and Diagnosis, EULAR Compendium on Rheumatic Diseases. BMJ Publishing Group LTD., Italy, pp. 444–463.
Shamir, L., Ling, S.M., Scott, W.W., Bos, A., Orlov, N., Macura, T., Eckley, D.M., Ferrucci, L., Goldberg, I.G., 2009a. Knee X-ray image analysis method for automated detection of osteoarthritis. IEEE Trans. Biomed. Eng. 56 (2), 407–415.
Shamir, L., Ling, S.M., Scott, M., Hochbergk, L., Ferrucci, I.G., Goldberg, 2009b. Early detection of radiographic knee osteoarthritis using computer-aided analysis. Osteoarthr. Cartil. 17, 1307–1312.
Stachowiak, G.W., Wolskin, M., Woloszynski, T., Podsiadlo, P., 2016. Detection and prediction of osteoarthritis in knee and hand joints based on the X-ray image analysis. Biosurf. Biotribol., 162–172.
Wang, T., Wen, C.Y., Yan, C.H., Lu, W.W., Chiu, K.Y., 2013. Spatial and temporal changes of subchondral bone proceed to microscopic articular cartilage degeneration in guinea pigs with spontaneous osteoarthritis. Osteoarthr. Cartil. 21 (4), 574–581.
Woloszynski, T., Podsiadlo, P., Stachowiak, G.W., Kurzynski, M., 2010. A signature dissimilarity measure for trabecular bone texture in knee radiographs. Med. Phys. 37 (5), 2030–2042.
Wu, Y., Yang, R., Jia, S., Li, Z., Zhou, Z., Lou, T., 2014. Computer-aided diagnosis of early knee osteoarthritis based on MRI T2 mapping. Biomed. Mater. Eng. 24 (6), 3379–3388.