Artif Intell Rev (2017) 47:531–559
DOI 10.1007/s10462-016-9492-8

# Image descriptors in radiology images: a systematic review

**Mariana A. Nogueira[1] · Pedro Henriques Abreu[1] · Pedro Martins[1] · Penousal Machado[1] · Hugo Duarte[2] · João Santos[2]**

Published online: 28 June 2016


**Abstract** Clinical decisions are sometimes based on a variety of patient's information such as: age, weight or information extracted from image exams, among others. Depending on the nature of the disease or anatomy, clinicians can base their decisions on different image exams like mammographies, positron emission tomography scans or magnetic resonance images. However, the analysis of those exams is far from a trivial task. Over the years, the use of image descriptors—computational algorithms that present a summarized description of image regions—became an important tool to assist the clinician in such tasks. This paper presents an overview of the use of image descriptors in healthcare contexts, attending to different image exams. In the making of this review, we analyzed over 70 studies related to the application of image descriptors of different natures—e.g., intensity, texture, shape—in medical image analysis. Four imaging modalities are featured: mammography, PET, CT and MRI. Pathologies typically covered by these modalities are addressed: breast masses and microcalcifications in mammograms, head and neck cancer and Alzheimer's disease in the

✉ Pedro Henriques Abreu
pha@dei.uc.pt

Mariana A. Nogueira
mn@student.dei.uc.pt

Pedro Martins
pjmm@dei.uc.pt

Penousal Machado
machado@dei.uc.pt

Hugo Duarte
hugo.duarte@ipoporto.min-saude.pt

João Santos
joao.santos@ipoporto.min-saude.pt

1 Department of Informatics Engineering, Faculty of Sciences and Technology, Centre for Informatics and Systems, University of Coimbra, Pólo II, Pinhal de Marrocos, 3030-290 Coimbra, Portugal

2 IPO-Porto Research Centre (CI-IPOP), Rua Dr. António Bernardino de Almeida, 4200-072 Porto, Portugal

case of PET images, lung nodules regarding CTs and multiple sclerosis and brain tumors in the MRI section.

**Keywords** Image descriptors · Medical images · Computer vision · Healthcare contexts

## 1 Introduction

The design of algorithms aimed at providing a summarized description of image regions (or of the whole image) is a prolific and central research topic in the fields of computer vision and image analysis. Such algorithms, commonly known as *descriptors*, have been successfully used in a wide range of applications such as matching (Tola et al. 2010), object (class) recognition (Mikolajczyk et al. 2005), and image retrieval (Jégou et al. 2011), just to name a few distinct and prominent ones.

By performing a research in Thomson Reuters website enclosing 6 different datasets and using 'image descriptor' as the query (encompassing the content of the title, abstract and keywords list), it is readily seen that the interest in this topic has grown substantially since 2000 (Fig. 1) reaching in 2013 more than 1500 publications. The literature has been prolific in introducing new types of descriptors as well as refinements over older ones, but this growth is largely explained by the increasing number of application domains in which descriptors have been playing an important role. For example, descriptors have been widely and successfully used in medical image analysis. They are an effective tool to assist in the interpretation of different medical image modalities, contributing to improve the diagnosis and monitoring of a large number of diseases.

In this paper, we focus on descriptors in medical image analysis. We are primarily interested in analyzing the specificities of certain problems and how these affect the choice of descriptors. Note that the literature reporting the use of descriptors in medical image analysis has became vast to be comprehensively reviewed herein. As such, we focus on presenting a list of relevant and heterogeneous works that illustrate the diversity of the field.

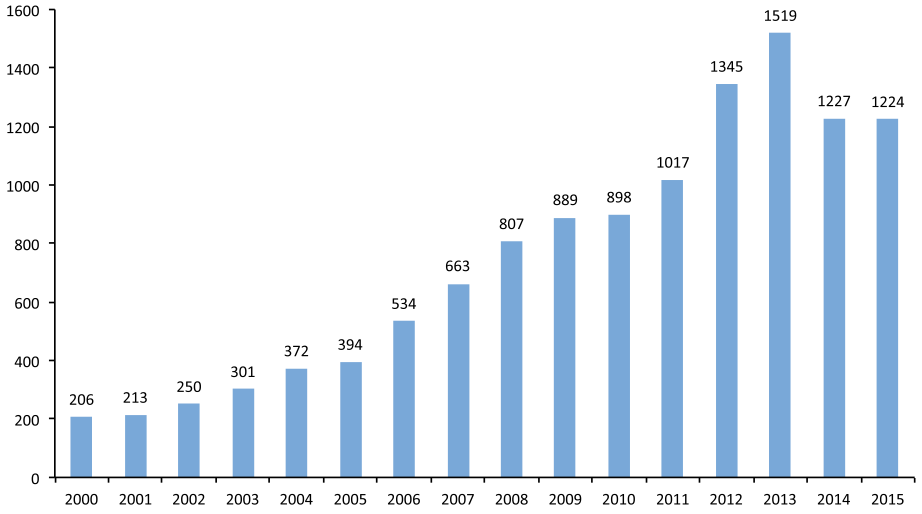


**Fig. 1** Distribution per year of indexed papers related to works with image descriptors over the period 2000–2015

The remainder of this paper is organized as follows. In Sect. 2, an introduction to descriptors is made by providing basic definitions and outlining different techniques to construct descriptors. In Sect. 3, a review of the use of local descriptors in medical image analysis is made. Finally, in Sect. 4, remarks are made about the state of the art and future research directions in image description under a medical image analysis context.

## 2 Image descriptors: preliminaries

In short, a descriptor is an algorithm that outputs a vector whose components describe the content of a contiguous part of an image. The simplest form of a descriptor is the one that outputs the pixel intensities of an image region (Mikolajczyk and Schmid 2005). However, this straightforward description has two major drawbacks: (1) it is not computationally efficient due to the relatively high dimensionality of the vector; (2) the descriptor is not robust to image distortions. While such descriptor makes sense in some scenarios, state-of-the-art-descriptors represent more sophisticated solutions in which computational efficiency, robustness and invariance properties are not overlooked.

Different techniques have been used to design efficient and effective descriptors. One of the most common approaches is to use histograms to represent certain characteristics related to appearance or shape (Mikolajczyk and Schmid 2005). Popular descriptors such as scale invariant feature transform (SIFT) (Lowe 1999, 2004), local binary patterns (LBP) (Ojala et al. 2002), or histogram of oriented gradients (HOG) (Dalal and Triggs 2005) make use of this distribution-based strategy. In some cases, the histogram depicts differential information, like in SIFT, where a 3D histogram of gradient locations and orientations is built in order to describe the edge distribution. Another common strategy relies on the spatial-frequency analysis of the image regions. Examples from this category are descriptors constructed from the Gabor filter banks or the wavelet transform. Moment-based techniques are also frequent in the design of descriptors. In this case, descriptors are designed to take advantage of invariance properties of moments with respect to geometric transformations. For example, there are a number of shape descriptors based on Zernike moments (Teague 1980), which are invariant to rotation and robust to noise.

This diversity of techniques allows the description of elements such as intensity, texture and shape, i.e., elements from which we can discriminate different features in the image (in Table 1, we group some of the most popular descriptors according to the image property that they are most often associated with). This is particularly useful in a healthcare context. Descriptors can be a powerful tool to assist health professionals in their clinical decision-making, namely in tasks such as the classification of lesions as well as in the detection and quantification of abnormalities or anatomic structures.

Before starting the review itself, we briefly explain the concepts of the image descriptors that the reader will most often come across later on in the literature review. Moreover, as a considerable share of these works combine classification methods with the image descriptors, we also present a brief summary of some prominent classifiers.

### 2.1 Image descriptors

#### *2.1.1 Gray-level histogram*

The gray-level histogram is a vector containing the absolute frequency of each gray level in a 2D or 3D segmented patch.

**Table 1** Grouping of popular descriptors according to the image property that they are most often associated with

| Intensity | Texture | Shape |
|---|---|---|
| Intensity vector | Co-occurrence matrix | Zernike moments |
| Histogram | Run-length matrix | SIFT |
| | Wavelets (and curvelets) | HOG |
| | Gabor filter banks | |
| | LBP | |

### *2.1.2 Gray-level co-occurrence matrix (GLCM)*

Haralick et al. (1973) the GLCM is a matrix containing in the element $(i, j)$ the number of times a pixel of intensity $i$ is at distance $d$ from a pixel of intensity $j$, in a pre-defined direction. There is a 3D version of this matrix, which allows for the description of volumetric texture features (Tsai et al. 2007).

### *2.1.3 Gray-level run-length matrix (GLRL)*

Galloway (1975) a gray-level run is a sequence of consecutive pixels or voxels with the same gray-level, in a certain direction. The GLRL is a matrix that contains, in the element $(i, j)$, the number of $j$-length runs of pixels (or voxels) with the gray-level $i$, in a pre-defined direction.

### *2.1.4 2D wavelets*

2D discrete wavelet decomposition consists of two successive 1D wavelet decompositions, one in the horizontal direction and the other in the vertical direction of the patch matrix. A single level decomposition results in 4 matrices with half the resolution of the initial image matrix – the approximation matrix and the vertical, horizontal and vertical details matrix.

Decompositions to further levels are obtained by successively applying the same mechanism to the previous level approximation matrix. Beforehand, a wavelet family must be chosen – it defines the morphology of the low-pass and high-pass filters.

### *2.1.5 Gabor filter banks*

Gabor filters are linear filters which consist of Gaussian kernels modulated by a sinusoidal plane wave, and are very popular in edge detection. They can either be found in 2D (Porat and Zeevi 1989) or 3D versions (Feng and Reed 2007).

### *2.1.6 LBP*

Ojala et al. (2002) the patch is divided into blocks and, for each pixel of the block, a binary string of the size of a pre-defined neighborhood will be produced. Each of the neighbors is responsible for a bit: 0 if the neighbor's gray level is inferior to the central pixel's, and 1 otherwise. This binary string is usually converted to its decimal form. Then, for each block, a histogram is produced, with the count of the number of times each decimal number is produced by pixels in the block. This descriptor is intensity- and scale-invariant. In addition,

it can also be rotation invariant, if we only consider the so-called uniform patterns. The LBP descriptor has also been extended to deal with volumetric data (Fehr 2007).

#### *2.1.7 Zernike moments*

Teague (1980) the Zernike polynomials are a set of complex polynomials orthogonal to each other in the unit disk. Each polynomial is characterized by its order ($n$) and its number of repetitions of the azimuthal angle ($m$). The Zernike moment of an image, for a given $n - m$ combination, consists of its projection onto the respective Zernike basis function. The magnitudes of the moments are rotation-invariant. Also, since the polynomials are orthogonal to each other, there is no redundancy or overlap of information between moments. There is also a 3D version of these moments (Canterakis 1999).

#### *2.1.8 SIFT*

Lowe (1999, 2004) SIFT is a rotation—and scale-invariant local descriptor in the form of a 3D histogram of gradient locations and orientations. More precisely, it is a 128-dimensional vector representing the gradient information of $4 \times 4$ locations around a pixel in an orientation histogram with 8 bins. There is also an extension of the SIFT descriptor to 3D (Scovanner et al. 2007).

#### *2.1.9 HOG*

Dalal and Triggs (2005) like SIFT, this descriptor is based on the idea that local shape of an object can be rather well characterized by the distribution of local intensity gradients. It is put in practice by dividing the image in smaller cells of pixels, accumulating for each cell a 1D-histogram of gradient directions over its pixels. A 3D version of this descriptor can be found in Kläser et al. (2008).

### 2.2 Classification methods

Classifiers are a class of algorithms which can be trained by examples in order to learn to classify objects based on a set of features. Herein, the theory behind some state-of-the-art classifiers is approached.

#### *2.2.1 Bayes classifier*

The minimum-error Bayes classifier assigns a test sample to the class with the highest *a posteriori* probability, the latter being computed based on a model built on the training set.

#### *2.2.2 k Nearest Neighbors (kNN)*

In order to classify a test sample, its distance (various metrics can be used) to every training sample is computed (not a very good practice for large datasets). The assigned class will be the most occurring in the k nearest training samples.

#### 2.2.3 Random forests

RF are ensemble learning methods that output a final decision based on the decisions of multiple decision trees, addressing the eventual problem of overfitting obtained with individual decision trees.

#### 2.2.4 Support vector machines (SVM)

SVMs aim at class discrimination through determining the hyperplane which maximizes class separation margin. The most simple SVMs are the linear ones, with linear hyperplanes. However, in the (most frequent) scenario of the classes not being linearly separable, non-linear kernel functions are used to map data to a higher-dimensional space where they become linearly separable.

#### 2.2.5 Artificial neural networks

A generic neural network is composed of one or more layers of "neurons". Feature vectors are fed to the first layer. Each entry of the vector is assigned a weight by each neuron of this layer, which outputs a value based on the response of an activation function to the weighted sum of all entries (plus an optional bias). These outputs will be used as inputs to the following layer, and so on, and the outputs of the last layer are the network outputs. Several standard ANN-architectures have been proposed. The multilayer perceptron, typically trained with a backpropagation algorithm, is probably the most often adopted one.

## 3 Literature review

In this section, we provide a review of the use of image descriptors in healthcare contexts. Our main goal is to give the reader an introductory document to this subject. The review focuses on four different types of image exams: mammograms, PET and CT scans, and MR images. We decided to narrow down our choice to these image exams, as they are illustrative of the diversity and relevance of works based on the use of image descriptors. We also opted not to include descriptor-based medical image registration techniques, since there are already several surveys covering this topic (e.g., Sotiras et al. 2013; Zitová and Flusser 2003 or Oliveira and Tavares 2014).

Most of the works described herein are classification systems aimed at supporting clinical decisions. As such, classifiers play also an important role in the performance of the system. In our analysis, we focus on the selection of the descriptors to construct the feature vectors. Therefore, we are primarily interested in large-scale studies that explore different types of image descriptors.

### 3.1 Mammographic images

Mammographic images are generally used for the analysis of lesions related to breast cancer, which usually correspond to masses and/or calcifications. These lesions stand out as more intense regions due to their higher density (and thus higher x-ray absorption), when compared to normal breast tissue.

Most breast cancer computer-aided-diagnosis (CAD) studies using image descriptors of the lesions aim for:

– the detection of lesions regardless of the type;
– the detection of masses or calcifications; or
– the discrimination between benign and malignant forms of lesions.

The selection of the appropriate image descriptors for each purpose is the focus of this section.

The first descriptors to be explored were statistical features computed directly from the gray levels of the regions of interest (ROIs) (Christoyianni et al. 2000; Huo et al. 1998; Ramos-Pollan et al. 2012) or from the gray-level histograms of the patches (Sheshadri and Kandaswamy 2007). Instead of building simple *intensity descriptors*, some authors proposed the extraction of texture information. For example, fractal features were often used in early studies (Priebe et al. 1994; Marchette et al. 1994; Guliato et al. 2008; Cabral and Rangayyan 2012). Second-order statistics computed from the GLCM (a set of 14 of these features is commonly referred to as Haralick features Haralick et al. 1973) (Dhawan 1996; Wang et al. 2009; Yu and Guan 2000; Dua et al. 2009) and third-order statistics computed from the GLRL (Galloway 1975; Sahiner et al. 2001; Kim and Park 1999) were also, and still are, widely explored. Later, multi-resolution texture features, such as the ones extracted from wavelets (Rashed et al. 2007; Ferreira and Borges 2003; Dhawan 1996; Soltanian-Zadeh and Rafiee-Rad 2004) and curvelets (Candès et al. 2006) or Gabor-based also started to be used (Eltoukhy et al. 2010; Wang et al. 2009; Buciu and Gacsadi 2011). More recent descriptors such as LBP (Ojala et al. 2002) have also been adopted for breast cancer CAD applications (Oliver et al. 2012). Shape parameters such as compactness, eccentricity, area, radius, perimeter, concavity and symmetry were the first shape descriptors to be explored (Bagui et al. 2003). Then, Hu's invariant moments (Hu 1962) were used in some studies, but Zernike moments (Teague 1980) appeared as a more advantageous alternative in several aspects (e.g., non-redundant information Constantinidis et al. 2001; Belkasim et al. 1991). More recent appearance/shape descriptors such as HOG (Dalal and Triggs 2005) have also been adapted to breast cancer CAD systems (Moura and Guevara-López 2013).

As an exhaustive review of breast cancer CAD would be too long, we focus on a set of representative works, which highlight the importance of descriptor selection. We selected three studies, whose main features are summarized in Table 2.

Moura and Guevara-López (2013) collect a set of 11 descriptors that had been previously used in breast cancer diagnosis studies, as well as a novel descriptor, and:

1. Compute the performances of standalone descriptors, standalone clinical data and combinations of both, with the main hypothesis that combining image descriptors with clinical data will be advantageous over using either of them isolated, and also to assess whether the highest performing combinations are based on the best standalone descriptors;
2. Identify a breast lesion regardless of the type, identify masses or calcifications, expecting that different descriptors will be more suitable for different lesions; and finally
3. They perform the experiments using two different mammographic images databases, which can help us understand whether the highest performing solutions for each scenario are different for distinct databases or there are solutions with consistent performances regardless of the database.

The 11 mentioned descriptors were the following: two intensity-based descriptors—intensity statistics computed directly from the gray-levels of the patches (IS) and from the gray-level histogram; six descriptors usually associated with texture features—GLCM, GLRL and gray-level-difference matrices, Gabor filter banks (Gab), wavelets and curvelets; and three descriptors usually associated with shape features—Zernike moments (Zer), the Hu set of invariant moments and HOG. Within the last group, a novel descriptor was proposed,

**Table 2** Main features of the studies reviewed in Sect. 3.1

| Authors | Objectives | Dataset(s) | Descriptors Families |
|---|---|---|---|
| Moura and Guevara-López (2013) | A comparison of the suitability of a very comprehensive set of image descriptors for detecting lesions (e.g. masses or calcifications); improving performance by combining other clinical information with the image descriptors | 913 segmentations of benign findings and 849 of malignant findings, extracted from the digital database for screening mammography (DDSM) (Heath et al. 2000)<br>187 benign findings and 175 malignant findings from the BCDR-F01 dataset of the breast cancer digital repository (BCDR) (Guevara-López et al. 2015)<br>Both databases are public repositories | Gray-level statistics computed from the patch<br>Gray-level histogram<br>Invariant moments<br>Zernike moments<br>Haralick features<br>GLDM<br>GLRL<br>Gabor filter banks<br>Wavelets<br>Curvelets<br>HOG<br>HGD (novel) |
| Tahmasbi et al. (2011) | Extraction of shape and margin properties of breast lesions using Zernike moments, for discriminating between benign and malignant lesions | 209 normal breasts, 67 ROIs with benign lesions and 54 ROIs with malignant lesions, extracted from digital mammograms from the mammographic image analysis society (MIAS) database (Suckling et al. 1994) | Zernike moments |
| Sharma and Khanna (2015) | Zernike moments for malignancy classification of mammographic patches | 534 malignant and 266 benign samples from IRMA database. 407 benign and 857 malignant findings from DDSM database | Zernike moments |

**Table 3** AUC values for the standalone clinical data experiments

| Dataset | All lesions | Masses | Calcifications |
|---|---|---|---|
| DDSM | 0.853 | 0.867 | 0.807 |
| BCDR | 0.712 | 0.829 | 0.725 |

**Table 4** Highest performing standalone descriptors for each experiment

| Dataset | All lesions | Masses | Calcifications |
|---|---|---|---|
| DDSM | GLRL (0.743) | GLRL (0.733) | Wavelet (0.733) |
| | HGD | HGD | Har |
| BCDR | HGD (0.825) | HGD (0.860) | Har (0.793) |
| | HOG | HOG | Gab, HGD, Wav, Curv |

The AUC value is only presented for the descriptor with the best performance

histograms of gradient divergence (HGD), which is especially designed for round-shaped objects, such as masses, and aims at describing the regularity of their shape.

The mammographic images used in the study belong to the digital database for screening mammography (DDSM) (Heath et al. 2000) and the breast cancer digital repository (BCDR) (Guevara-López et al. 2015). Both databases provide clinical information alongside the images. A total of 1762 segmentations were extracted from the DDSM, 913 of them corresponding to benign lesions and 849 to malignant lesions. Regarding clinical information, images from the DDSM are complemented with patient age, density of the breast, subtlety of the lesion and with a set of observations of the radiologists about the lesions. These lesions are classified as masses and/or calcifications. In the case of masses, a characterization of the shape and margin is given; in the case of calcifications, their type and distribution are descibed. Regarding the BCDR dataset, a total of 362 segmentations were extracted, 187 corresponding to benign findings and 175 to malignant ones. Each segmentation is complemented with the following clinical information: age, breast density and the types of abnormalities observed in the segmentation (masses, calcifications and/or four other types of lesions).

For each database, three subsets were explored: the whole dataset (with all types of lesions), a subset exclusively with masses and a subset exclusively with calcifications; for each of these subsets, and for each of the 12 descriptors, the following three scenarios were considered: feeding only the descriptor to the classifier, feeding the descriptor combined with the clinical information to the classifier and feeding only the clinical information to the classifier. This was done for several classifiers, such as SVM, Random Forests, Logistic Model Trees, kNN and Naive Bayes, and the highest area under the curve (AUC) of the receiver operating characteristic (ROC) among all classifiers was selected as the performance measure for each case.

The main results are presented in Tables 3–5.

In Table 3 the results for standalone clinical data are presented. It was observed that standalone clinical data achieves consistently higher average performances in images of the DDSM database than in images from the BCDR database. This indicates that the clinical data associated with images from the DDSM database is more discriminatory than that associated with images from the BCDR database.

In Table 4, the name and AUC value of the highest performing standalone descriptor is presented for each lesion type and for each database. The names of other descriptors

**Table 5** Highest performing standalone descriptors for each experiment

| Dataset | All lesions | Masses | Calcifications |
|---|---|---|---|
| DDSM | IS (0.868) | IS, Zer (0.89)<br>HGD | Gab (0.803) |
| BCDR | HGD (0.817) | HGD (0.894) | Gab<br>Har (0.815) |

The AUC value is only presented for the descriptor with the best performance

with performances similar to the highest performing one are also indicated. Comparing with Table 3, we can see that standalone clinical data always outperforms all standalone descriptors in DDSM database, and is always outperformed by some standalone descriptors in the BCDR database. It is clear that the best performances achieved by standalone image descriptors in the BCDR database are consistently higher than those of DDSM database. The highest performing descriptors in the all-lesions subset are generally the same as those of the only-masses subset. We observe that the performances of the descriptors depend on the used database (GLRL and HOG work better in the DDSM and BCDR databases, respectively), but we can also find descriptors that show consistently good performances transversely to the used database—the proposed HGD descriptor is always on the top two for all the four cases. Moving on to the calcifications subset, texture descriptors clearly dominate the best performing group, with Wavelet and Haralick descriptors in the top group for both databases.

In Table 5 the name and AUC value of the highest performing descriptor-clinical-data combination is presented for each lesion type and for each database. The names of other combinations with performances similar to the highest performing one are also indicated. Except for the only-calcifications subset of the DDSM database and for the all-lesions subset of the BCDR database, there were always some combinations which outperformed both standalone clinical data and all standalone descriptors, not necessarily based on the highest performing standalone descriptors. In the DDSM database, intensity statistics-based combinations were the highest performing ones in all-lesions and only-masses subsets, with Zernike and HGD descriptors also in the top group for the only-masses subset. Regarding the BCDR database, HGD-based combinations significantly outperformed all others for the all-lesions and only-masses subsets. Regarding calcifications, again combinations based on texture descriptors, namely Haralick and Gab combinations, outperformed others.

By analyzing Tables 4 and 5, we can conclude that, in general, the group of highest performing standalone descriptors and combinations is similar in the all-lesions and only-masses subset. These combinations generally consisting of intensity, texture and shape descriptors—GLRL, HGD, HOG, IS and Zer achieved the best performances. Particular attention should be drawn to the proposed shape descriptor HGD, which consistently was in the highest performing group of descriptors, either isolated or combined with clinical data. In the case of calcifications, the groups of the best standalone descriptors and combinations are almost exclusively based on texture—wavelets, Haralick features, Gabor filter banks, and curvelets achieved the best performances—as it could be expected given the calcifications' texture characteristics.

In summary, a few useful conclusions can be drawn from Moura and Guevara-López's study:

1. In most scenarios, combining descriptors with other clinical information significantly outperforms standalone clinical data and all standalone descriptors, with the highest per-

forming combination not being necessarily based on the highest performing standalone descriptor.
2. The suitability of a descriptor depends on the type of lesion: when all lesions are mixed and when only masses are present, texture and shape descriptors (e.g. GLRL, HGD and HOG) lead to the highest performances, whereas regarding calcifications, texture descriptors (e.g. wavelets, Haralick features, Gabor filter banks and curvelets) clearly outperform others—this is intuitive given the characteristics of the lesions.

Sharma and Khanna (2015) extract Zernike moment magnitudes from mammographic patches for malignancy classification. The images were extracted from the Image Retrieval in Medical Applications (IRMA) database—mammographic patches extracted from images of the DDSM database (Heath et al. 2000) (534 malignant and 266 benign samples),—and from the DDSM itself (407 benign and 857 malignant samples).

For training, ROIs of abnormal regions were segmented manually. For testing, they were segmented manually from automatically identified dense regions, based on k-means clustering.

Zernike magnitudes were computed and fed to k-NN and radial basis function (RBF)-SVM classifiers. With SVM, performance increased up to order 20, and decreased for higher order moments, on both databases (sensitivity = 0.99 and specificity = 0.99 for the IRMA database and sensitivity = 0.96 and specificity = 0.96 for the DDSM database, with order 20). kNN generally performed poorer than SVM, and needed much higher moment order to achieve similar performances (at order 35, sensitivity = 0.97 and specificity = 0.92 for IRMA database and sensitivity = 0.94 and specificity = 0.93 for the DDSM database).

Compared to GLCM and discrete cosine transform (DCT)-based features, the Zernike-based descriptor outperformed both (sensitivity values of 0.90, 0.78 and 0.99 and specificity values of 0.93, 0.78 and 0.99 for GLCM, DCT and Zernike-based features respectively, using SVM and the IRMA database).

The main conclusions to draw from the work by Sharma and Khanna are:

1. Zernike descriptors can outperform well-known descriptors as GLCM and DCT in malignancy classification of mammographic patches, achieving very high performances.
2. An optimizing classifier should be selected—in this case SVM was clearly advantageous over k-NN, achieving higher performances with lower dimensionality feature vectors.

Tahmasbi et al. (2011) also focus on detecting a particular type of breast lesion: malignant masses. Given a dataset with samples of normal tissue, benign masses and malignant masses, the authors aim to detect only the malignant masses. For this purpose, they must find a descriptor that is discriminative between benign and malignant masses. Given the fact that the shape and margin of malignant masses are usually more eccentric and irregular than those of benign masses, they use shape and margin features based on Zernike moments.

The mammographic images used in this study belong to the Mammographic Image Analysis Society (MIAS Suckling et al. 1994) digital mammogram database. Manual segmentation was performed by two radiologists (209 normal breasts, 67 ROIs containing benign lesions and 54 ROIs with malignant lesions) and radial averaging was performed to calculate the final boundary of the mass.

Images were processed in different ways in order to compute shape and margin features: to enhance shape properties, boundaries were binarized and masses were “filled”. To make margins more visible, histogram equalization was performed, increasing the contrast of the ROI.

Zernike moment magnitudes were then computed on both types of processed images as features of mass shape and margin, and fed, individually and mixed in different proportions, to a multilayer perceptron neural network.

It was observed that the best performance belonged to the systems that used only shape features, particularly those based on low-order Zernike moments (achieving a false positive rate (FPR) of 11.13 %, a false negative rate (FNR) of 0.0 %, an Accuracy of 92.8 % and an AUC of 0.975) , and that, as the proportion of margin features increased, the performance decreased.

Furthermore, the only-shape low-order Zernike-moment-based feature vector outperformed several margin and shape features of previous studies that used the MIAS database, namely spiculation index and other spiculation metrics, compactness, index of convexity, Fourier factor, fractal dimension, Fourier transform of radial distance, fractional concavity and fuzziness of mass margins (Guliato et al. 2008; Cabral and Rangayyan 2012; Mu et al. 2008; Rojas-Domínguez and Nandi 2009; Tahmasbi et al. 2010).

The main conclusions to draw from the work by Tahmasbi et al. are:

1. Tuning parameters of a descriptor (if they exist), in this case the order of the Zernike moments, can be determinant in the performance, as features based on low-order Zernike moments outperformed those based on high-order ones.
2. Processing images to enhance certain properties can also be determinant, as in this case the shape-enhancing processing led to higher descriptor performances than the margin-enhancing processing.

### 3.2 PET images

In PET images, functional anomalies are addressed through the analysis of the uptake of radioactive tracers (regions of higher uptake stand out as more intense regions).

We decided to restrain to studies on FDG-PET images, in which the tracer is a radioactively-labeled glucose analogue (FDG), and thus useful for the analysis of pathologies characterized by regions of abnormally high (e.g. tumors or infections/inflammations) or low metabolism [e.g. Alzheimer's disease (AD)]. That choice was supported by two (correlated) reasons: (1) most CAD studies of PET images are on FDG-PET images; (2) FDG-PET allows for analysis of very prominent pathologies these days, as tumors and dementia.

We selected three studies: one related to tumor detection, another related to prediction of tumor treatment outcome and another related to AD CAD. Those studies are representative not only pathology-wise, but also for featuring typical image descriptors for analysis and addressing common problems. The main characteristics of the three studies are summarized in Table 6.

Intensity features such as intensity-volume histograms (Naqa et al. 2009; Yu et al. 2009), texture features such as those based on GLCM (Naqa et al. 2009; Yu et al. 2009) and shape features as solidity and eccentricity have been used in in FDG-PET tumor CAD Yu et al. (2009).

Regarding AD CAD, most approaches rely solely on voxel intensity for classification (Silveira and Marques 2010; Fan et al. 2008; Davatzikos et al. 2006). However, other types of features have been explored: intensity-based features as first order statistics (Ramírez et al. 2013), texture features as those based on GLCM (Ramírez et al. 2013), wavelets (Bicacro et al. 2012), gradient magnitude and orientation (Bicacro et al. 2012) and LBP (Morgado 2012), and shape features as coefficients of spherical harmonics (Gerardin et al. 2009).

Wu et al. (2012) design a system for the detection of primary tumor and metastasis of nasopharyngeal carcinoma (NPC). The lesions were identified by radiologists on 25 sets

**Table 6** Main features of the studies reviewed in Sect. 3.2

| Authors | Objective | Dataset | Descriptors families |
|---|---|---|---|
| Wu et al. (2012) | Automatic detection and classification of nasopharyngeal carcinoma, combining image descriptors with a priori clinical knowledge | 25 PET/CT examinations of patients suffering from NPC from the PET/CT Unit in the University of Hong Kong | Intensity<br>Texture (second-order moments)<br>Shape (area, eccentricity, compactness) |
| Naqa et al. (2009) | Predictive model of treatment outcome based on intensity, texture and shape features of pre-treatment PET images | 14 images of patients with cervix tumor (7 with persistent disease after treatment) and 9 images of patients with head and neck tumors (4 of them died after treatment) | Intensity (SUV-based)<br>GLCM<br>Shape (compactness, eccentricity, extent, Euler's number). |
| Morgado (2012) | Comparison of feature extraction and feature selection techniques for automated diagnosis of MCI and AD using PET images | 59 PET images of each class (healthy patients, MCI patients, AD patients) retrieved from the ADNI database (Jack et al. 2008) | Intensity<br>Local variance<br>LBP |

of PET/CT images of patients suffering from NPC. Then, the lesions were segmented out and labeled as benign or malignant, to posteriorly assess the accuracy in segmentation and classification of the algorithm. In the proposed method, the suspicious regions (i.e., regions of high FDG uptake) were automatically segmented from PET images using a region growing algorithm and then registered to the corresponding CT images using a rigid registration algorithm (Viola and Jones 2001). A few texture (second-order moments) and shape features (area, eccentricity, compactness) were computed from the segmented ROIs.

A common problem in these studies is the false positive rate, due to the fact that some anatomical locations naturally exhibit high metabolism and are mistaken for tumors. A set of features was computed aiming to help the classifier to differentiate some of the regions taken as suspicious from anatomic regions that naturally show high uptakes. The average intensity of CT values was used to differentiate tumor regions from regions of physiological marrow uptake in bones and brown fat uptake in fatty tissues. The FDG intensity difference between the regional peak and its surroundings was used to differentiate regions of true bone metastasis from regions of normal bone marrow uptake. Anatomical location information was also used, such as the likelihood of a segmented candidate to be a part of the primary tumor or its nodal metastasis differs according to its anatomic location (by definition the primary tumor arises from the nasopharynx and a pattern of nodes spreads in the neck). The symmetry of the segmented candidate about the medial plane was also considered (symmetric organs as tonsils, salivary glands and thyroid, naturally show high FDG uptake). Different combinations of all features were input to an RBF-SVM classifier. The algorithm is evaluated in two fronts: segmentation performance and classification performance.

To evaluate the segmentation performance, an identification was considered positive if a lesion volume produced by the algorithm overlapped at least 80 % of the lesion volume drawn by radiologists. The system identified all the lesions but five asymmetrical and unusually hot tonsils and larynx were mistakenly classified as lesions.

Regarding classification, using fivefold cross-validation on 20 of the 25 sets of imaging data, the image feature combination of relative position, average intensity, area, eccentricity and symmetry has the higher true positive rate (TPR), 99.3 %, and the lowest false positive rate (FPR), 4.8 %, with the relative position being the most important feature, since taking it from the feature vector clearly worsened performance more than taking the other features. Applying this model to the remaining 5 sets, a TPR of 95.1 %, an FPR of 7.0 % and an accuracy of 93.3 % were obtained.

The work by Wu et al. shows that combining image descriptors as area and eccentricity with other clinical information as average CT values, symmetry measures to the medial plane, and anatomical location can significantly improve the performance of tumor detection on PET, helping to differentiate physiological from pathological uptakes.

Besides tumor detection, image descriptors have also been used in some studies for treatment response assessment and prediction. Naqa et al. (2009) use logistic regression to construct a predictive model of treatment outcomes of cervix and head and neck cancer types, based on intensity, texture and shape features computed from pre-treatment images.

The FDG-PET images were acquired at the Mallinckrodt Institute of Radiology at Washington University School of Medicine—14 of cervix tumor patients (half of them with persistent disease) and 9 from head and neck tumor patients (four of whom did not survive). The outcomes were measured in terms of disease persistence for the cervix tumor patients and in terms of survival regarding head and neck tumor patients.

The gross tumor volume (GTV) was determined by region growing [down-limited by 40 % of maximum standardized uptake value (SUV)] in cervix cancer scenarios and manually by radiologists in head ad neck cancer situations, according to the normal practice in the institute.

Intensity, texture and shape features were extracted from each GTV segmentation. Regarding intensity features, cumulative SUV volume histograms were obtained and, from these, metrics as the fractions of volume with SUV under 10 and 90 % and above 10 and 90 % of the maximum SUV, the differences between them, and common statistics such as mean, standard deviation, maximum and minimum were computed. Regarding texture and shape features, GLCM-based features such as energy, contrast, homogeneity and entropy, and shape features like eccentricity, Euler's number, solidity and extent were computed.

The Spearman coefficient and the AUC value were used to assess the discriminatory power of the features. Regarding the cervix tumor patients, the difference between the fractional volumes above 90 and 10 % of the maximum SUV and energy achieved the highest scores. Regarding patients with head and neck tumors, the fractional volume above 90 % of maximum SUV and the shape extent were the most discriminative features.

For both tumor types, logistic regression analysis was used to obtain predictive models of the outcomes based on the top two discriminative features. When compared with the real outcomes, the cervix tumor model obtained a Spearman coefficient of 0.49 and an AUC value of 0.76, whereas the head and neck tumor treatment outcome predictive model obtained a Spearman coefficient of 0.89 and an AUC value of 1.0.

Hence, attractive results were obtained with simple two-feature logistic regression-based predictive models of treatment outcomes, particularly in the head and neck cancer. Texture, shape and SUV-volume features turned out to be more important in discrimination than the usual SUV statistics.

Morgado (2012) computes four different descriptors in PET images and compares their performances in the diagnosis of mild cognitive impairment (MCI), a syndrome that is proved to be related with the pre-clinical stage of Alzheimer's disease (AD), and AD itself. The four descriptors are Local Variance (LVAR), standard LBP (2D-LBP), a proposed three-dimensional generalization of LBP (3D-LBP), and voxel intensity (VI) for several resolutions (using a Gaussian scale-space pyramid) and were computed from PET images from 59 subjects of each class (MCI, AD and Cognitive Normal (CN)), extracted from the ADNI database (Jack et al. 2008).

A usual problem in dementia scenarios is the decision of computing features from a ROI or from the entire brain image. Both have their advantages and disadvantages: in the first case, one gets a lower dimensionality feature vector, but the segmentation work is required. In the second case, the segmentation work is not necessary, but one gets high dimensionality descriptors.The authors decided to compute the features from the entire brain image. Thus, dimensionality reduction is required. For that purpose, five feature selection algorithms were compared: squared point biserial correlation coefficient (PBCC), mutual information maximization (MIM), minimal redundancy maximal relevance (mRMR) and two medically driven selection alternatives, built over Eye Tracking data recorded while an expert was examining each subject's PET image— time-independent eye track driven selection (TI-ETDS) and time-dependent eye track driven selection (TD-ETDS). Regarding classification, linear and RBF-kernel SVMs were used.

Preliminary tests showed that:

1. The performances of TD-ETDS and TI-ETDS selection algorithms were similar or even better than those of the other selection algorithms in the AD versus CN task, but significantly worse in the other tasks (even worse than that of a random selector).
2. The performance accuracy computing VI features on the level 1 of the scale-space was similar to that when level 0 was used, but represented lower computational load, so level 1 was chosen.

3. The usage of an RBF kernel did not significantly improve the systems' performances, so the linear kernel was used.

The best feature-type–selection-algorithm combination was 3D-LBP and PBCC for the AD versus CN task (91.4 %), VI and MIM for the MCI versus CN task (79.4 %) and LVAR and MIM for the MCI versus AD task (73.4 %). However, other combinations reached similar performances. From the results, one may observe that the most accurate classification task was AD versus CN, with a clear drop in performance in the tasks that involve MCI. Over all the classification tasks, MIM was most frequently the best selection algorithm, followed by PBCC with similar results and mRMR with generally poorer results. Regarding feature types, it is not so easy to choose one that consistently had the best performance over all tasks.

In conclusion, all 2D-LBP, 3D-LBP, LVAR and VI performed well in the diagnosis of MCI and AD, with the PBCC and MIM feature selection algorithms leading to high performances after dimensionality reduction.

### 3.3 CT images

CT is mostly used for the analysis of bone and lung pathologies. Bone regions stand out as very intense regions when compared to the remaining tissues due to the higher density and thus higher x-ray absorption. On the contrary, lungs stand out as especially dark regions due to their particularly low density (high air percentage). CT is also used for analysis of pathologies of other tissues, but MRI is usually preferred over CT due to the high soft tissue contrast the former provides.

Most studies using CT images for CAD are related to lung diseases, particularly lung nodules, although other pathologies such as liver pathology or polyps in CT colonography are often addressed. Lung nodules are clearly seen in CT images, as they are dense regions and contrast with the dark lung background. In fact, the high CT resolution allows for the identification of very small nodules, which sometimes is not possible using MRI. However, when one's CAD system is sensitive to very small nodules, it gets more susceptible to false positives, such as noise or small artifacts, which can be considered nodules. For this reason, there are several studies on false positive reduction strategies.

The main features of the studies reviewed in this subsection are summarized in Table 7.

Intensity, texture and shape descriptors have all been used for lung cancer CAD: intensity descriptors are usually based on statistics computed from gray-level histograms (Depeursinge et al. 2007). Texture features are usually based on GLCM (McNitt-Gray et al. 1999), but others such as fractal, wavelet (Depeursinge et al. 2007) and SIFT features (Kato et al. 2009) have also been used. Regarding shape features, shape index, curvedness (Murphy et al. 2009) and diameter/size (Ko and Betke 2001; Brown et al. 2001) have been used.

Depeursinge et al. (2007) developed a system for discriminating among 5 different classes of lung tissue—healthy, emphysema, ground glass nodules, fibrosis and micronodules—based on texture features.

The ROIs were extracted from radiologist-annotated high-resolution CT images of an internal database: 77 samples of healthy tissue, 72 of emphysema, 155 of micronodules, 64 of fibrosis and 113 of ground glass nodules.

Three feature vectors were computed from the ROIs and compared: one based on wavelet statistics, another based on the gray-level histogram and percentage of air pixels, and their combination.

Regarding the wavelet descriptor, third order B-splines family was used as wavelet basis; the mean and the standard deviation of the detail, approximation and composite coefficients of each resolution level were concatenated to produce the feature vector. Grey-level his-

**Table 7** Main features of the studies reviewed in Sect. 3.3

| Authors | Objectives | Datasets | Descriptors families |
|---|---|---|---|
| Depeursinge et al. (2007) | Texture descriptors for classification of 4 different pathologies of lung tissue | 77 samples of healthy tissue, 72 of emphysema, 155 of miconodules, 64 of fibrosis and 113 of ground glass nodules from an internal database | Gray-level histogram<br>Wavelet |
| Boroczky et al. (2005) | Finding the optimal feature subset, out of a pool of texture, shape and intensity features, for false positive reduction of a previous lung nodule detection system | 52 true positives and 443 false positives output by a previous CAD system | Gray-level statistics computed from the patch<br>Gray-level histogram<br>Gradient matrix<br>Shape (sphericity, elongated and flat shape, compactness) |
| Dettori and Semler (2007) | Comparison of texture descriptors for classification of tissues of different organs in CT scans | 2 healthy chest and abdomen CT studies from Northwestern Memorial Hospital | Wavelet<br>Ridgelet<br>Curvelet<br>GLCM<br>GLRL |

tograms were also computed: pixel values are expressed in Hounsfield Units (H.U.), thus a density measure. As different tissues have different densities, pixel values are useful for discrimination of different tissues. The second descriptor consisted of the gray-level histogram plus the percentage of air in the lungs (pixel values under 1000 H.U.). Before classification, feature vectors were mapped to the [0,1] interval. A k-nearest neighbor classifier (based on the Euclidean distance) was used and the leave-one out approach was used for validation.

The following experiments were performed: 5 classification tasks of each tissue type versus the remainder, and a multiclass task. With the exception of the healthy tissue versus all task, where the histogram + air % descriptor outperformed the other two (98 % accuracy), the combination feature vector systematically outperformed the isolated ones, with accuracies in the range 95–100 %. In the multiclass task, 92.2 % of the healthy samples were correctly classified, all emphysemas, 86.7 % of the ground glass lesions, 92.9 % of the micronodules and 93.8 % of fibrosis.

In summary, a combination of gray-level histogram statistics and wavelet features can achieve high performance in the detection and discrimination between lung pathologies.

Boroczky et al. (2005) address false positive reduction of a previous lung nodule CAD system, with a feature selection process based on genetic algorithms (GA) coupled with SVM, among intensity, shape and texture features. The CT images were provided by a private clinic in New York. Out of 495 Volumes of Interest (VOIs) classified as nodules by a previous CAD system (Wiemker et al. 2002), radiologists labeled 52 of them as being true nodules and 443 as false nodules.

From each VOI, 23 features were extracted: statistics computed directly from the VOI and from the gray-level histogram, difference between the mean values of the gray-levels within the nodule and in its vicinity, statistics computed from the gradient matrix, shape features such as spheric shape, flat shape and elongated shape, sphericity and compactness. Then, a CHC GA (Eshelman 1991) was used for finding the optimal subset size and the optimal subset of features: a hierarchical fitness function was used, with the first priority assigned to the sensitivity, the second to the specificity and the third to the number of features in the subset. A first GA run was performed to determine the optimal subset size, based on the occurrences of chromosomes representing each subset size. Out of 8000 trials, chromosomes representing subsets of 10 features were clearly the most frequent. Then, subset size was fixed to 10, and a second GA run was performed to determine the optimal 10-feature group.

The fittest chromosome corresponded to the following subset: gray level minimum, compactness, flat shape, elongated shape, sphericity, contrast, gradient maximum, gradient standard deviation, gradient skewness, and gradient small value ratio. Using this feature subset, all the true positives were retained (sensitivity of 100 %), and a 50 % reduction of false positives was achieved.

For validation purposes, the performance was determined for optimal subsets of adjacent sizes (6–14), and the optimal subset of size 10 remained the one with the highest performance. Furthermore, using a Tomek downscaling of factor 3, false positive reduction went up to 56.41 %—higher factors caused the sensitivity to fall from 100 %. Thus, a false positive reduction of more then 50 % was achieved, with the selection of an optimal feature subset consisting of intensity, texture and shape features.

Image descriptors can also be very discriminative between organic tissues in CT images, as observed in Dettori and Semler (2007), where Dettori et al. compare the performance of wavelet, ridgelet, and curvelet texture descriptors, as well as two standard texture descriptors—GLCM and GLRL—in the discrimination between tissues of five different organs.

**Table 8** Average performances over all organs of the highest performing descriptors within the wavelet, ridgelet and curvelet groups and of the GLCM and GLRL descriptors

| Descriptor | TPR (%) | TNR (%) | Prec. (%) | Acc. (%) |
|---|---|---|---|---|
| Haar wavelet | 74.4 | 93.7 | 74.4 | 89.9 |
| Ridgelet | 83.8 | 96.0 | 85.0 | 93.6 |
| Curvelet | 94.6 | 98.7 | 94.7 | 97.9 |
| Co-occurrence | 89.1 | 97.3 | 89.7 | 95.8 |
| Run-length | 84.3 | 96.1 | 84.7 | 93.9 |

Using an active contour mapping algorithm (Xu et al. 2005), pure tissues of five classes (363 backbone, 446 heart, 506 liver, 411 kidney and 364 spleen) were segmented from two healthy chest and abdomen CT studies.

For each segmented candidate, wavelet (Haar, Daubechies and Coiflet), ridgelet and curvelet transforms were applied and statistics (e.g. mean, standard deviation, Haralick features, energy, entropy) calculated over the resulting coefficients were combined and used as texture features. As mentioned before, standard texture descriptors—GLCM and GLRL—were also computed.

The different feature vectors were fed to a decision tree classifier based on the cross-validation classification and regression tree approach, to classify the segmented candidate as backbone, heart, liver, kidney or spleen tissue.

The performances of the descriptors with the higher average performance over all organs within the wavelet, ridgelet and curvelet groups and of the GLCM and GLRL descriptors are presented in Table 8. Within the wavelet group, a Haar feature vector merging the mean, standard deviation and 9 Haralick features computed and averaged over the details of two resolution levels, outperformed all other Haar feature vectors and those based on Daubechies and Coiflet wavelets, for most organs. Within the ridgelet group, a feature vector consisting of the entropy values of all 32 radial directions of each of two resolution levels clearly outperformed all other feature vectors. Regarding the curvelet group, a feature vector combining mean, standard deviation, entropy and energy measures computed from each of 18 wedges (of two levels of resolution) significantly outperformed all other curvelet-based feature vectors.

Observing the results in Table 8, one can see that the best curvelet-based feature vector clearly outperforms the best wavelet and ridgelet-based feature vectors. Regarding the comparison with standard texture descriptors GLCM and GLRL, one can conclude that the the best wavelet and ridgelet feature vectors are outperformed by GLCM and GLRL-based ones. On the other hand, the best curvelet feature vector outperforms both.

In conclusion, curvelet-based image features can be very powerful in the discrimination of textures of different tissues on CT images, outperforming in many scenarios other standard texture descriptors as GLCM and GLRL and wavelet and ridgelet-based features.

### 3.4 MRI images

MR images are based on the different relaxation times of tissues after being subjected to an electromagnetic stimulus, and provide excellent soft tissue contrast. Moreover, contrast can be further enhanced through the injection of a contrast enhancement agent. For this reason, MRI is widely applied in the diagnosis and treatment of neurological, cardiovascular, musculoskeletal, liver and gastrointestinal diseases.

Most CAD studies using MR images are related to brain pathologies as tumors, lesions related to white matter hyperintensities (WMH) and dementia, although other pathologies as breast or prostate cancer are also frequently object of study (Madabhushi et al. 2005; Meinel

et al. 2007). WMH have shown to be associated with several prominent pathologies, such as multiple sclerosis, vascular disease and dementia. In Table 9 we summarize the main features of the studies we review in this subsection.

Concerning brain tumor CAD, intensity, texture and shape features have been used. In the intensity group, first-order statistics are the most common (Zacharaki et al. 2009). In the texture group, GLCM (Herlidou-Même et al. 2003), GLRL (Herlidou-Même et al. 2003) and gradient matrix-based features (Herlidou-Même et al. 2003) are the most frequent, although fractal features (Iftekharuddin et al. 2009), Gabor features (Zacharaki et al. 2009) and more recently descriptors such as LBP have also been used. Regarding the shape group, circularity, irregularity and surface-to-volume ratio (Zacharaki et al. 2009) have been used, but also more recent descriptors such as HOG. As a representative study, we chose (Reddy et al. 2012), where Reddy et al. use intensity, texture (LBP) and shape (HOG) descriptors for brain tumor classification.

Unay et al. (2007) suggest an LBP-based texture analysis to achieve robustness to bias-fields and misalignments (namely rotation).

Until recently, analysis of brain MR images was mostly exclusively based on intensity features. Since soft tissue contrast is high, intensity features are naturally discriminative. However, if our aim is brain lesion segmentation for accurate tumor volume measurement, intensity-based analysis may not be enough: bias-fields (intensity inhomogeneity caused by equipment interferences during acquisition) and inter- and intra- patient misalignment significantly degrade the performance of automatic segmentation techniques. Moreover, normal tissues may also be enhanced with contrast agent, resulting in the segmentation of a larger region than the actual lesion; on the other hand, the presence of noise or non-uniformity of the distribution of contrast agent in the lesion may result in an incomplete extraction. These are some of the reasons why some studies as the one by Unay et al. are turning to other types of descriptors as potentially more robust complements or alternatives to intensity-based analysis.

Dual MR scans of 549 subjects were used. In order to test robustness of LBP to bias fields and rotation degradations, the original MR images were degraded using a set of simulated bias fields (with larger or smaller intensity and spatial variations) and rotated by several angles using three different interpolation methods—nearest neighbor, bilinear and bicubic.

Then, standard LBP, rotation invariant LBP and uniform and rotation invariant LBP were computed from the original and degraded images. The Bhattacharyya distance between each descriptor computed in the original and the degraded image was used as the dissimilarity measure.

It was observed that dissimilarity increased with the increase of intensity and spatial variations of bias fields. In the case of rotation degradation, dissimilarity was higher for large rotation angles and for lower-complexity interpolation methods. Despite these increases, all dissimilarity values fall below 0.04 % in the case of bias field degradation and bellow 4 % in the case of rotation degradation. Regarding the different variants of LBP, introducing rotation invariance and uniformity in LBP always increased performance.

In summary, the uniform and rotation invariant LBP variant is quite robust to bias-fields and rotation degradation, which makes it a promising complement to the usual intensity-based analysis.

Reddy et al. (2012) also believe that it is advantageous to add texture features to the usually used intensity features on MRI, based on the fact that normal brain tissues differ also in structure from lesions.

They propose a confidence guided segmentation method and compare its performance with that of other known segmentation methods. For that purpose, 19 groups of multiparametric

**Table 9** Main features of the studies reviewed in this Sect. 3.4

| Authors | Objectives | Datasets | Descriptors families |
|---|---|---|---|
| Unay et al. (2007) | To demonstrate robustness of LBP texture descriptors to bias field and rotation degradations | Dual (T2 and PD) MR scans from 549 subjects from Leiden University Medical Center | LBP |
| Reddy et al. (2012) | To prove that incorporating the information of a confidence surface in segmentation methods, built over the output scores of a brain tumor classifier, significantly improves their performance | 19 groups of MRI images with brain tumor | Intensity<br>HOG<br>LBP |
| Theocharakis et al. (2009) | Discrimination between two white matter hyperintensities-related lesions, multiple sclerosis (MS) and cerebral microangiopathy (CM) | 47 CM and 31 MS ROIs of MR images of an internal database | Gray-level histogram<br>GLCM<br>GLRL |

MRI images from 11 subjects were used. Each group consisted of a pre-contrast T1 weighted (T1pre), a post-contrast T1 weighted (T1post), a T2 weighted (T2) and a fluid attenuated inversion recovery (FLAIR) MRI image. To start, a mask for the enhanced region is generated with the difference between T1pre and T1post images.

Then, mean intensity (MI), LBP and HOG features are computed for each pixel within the enhanced region mask from each of T1pre, T1post, T2 and FLAIR images, and concatenated to form a single feature vector.

After that, the feature vectors of each pixel are input into two different classifiers, SVM and AdaBoost (Hill et al. 2001), for tumor pixel classification. A confidence surface is then constructed based on the classification output scores. The authors propose to use the generated confidence surface to guide the segmentation process: two classical segmentation methods, level set and region growing, are slightly modified to incorporate the confidence surface information in the segmentation process.

Regarding classification results, ROC curves were plotted and it was observed that AdaBoost outperformed SVM, with larger AUC values. Checking the AdaBoost weights for the different features, the MI features from T1pre and T1post had the larger weights, indicating that these features still play the most important role in tumor detection. On the other hand, the HOG features from T1pre and T1post images and the LBP feature from T1pre had larger weights than the MI feature from T2 and FLAIR images, suggesting that these texture features are also useful for discrimination.

For assessing segmentation accuracy, the average Dice similarity score (DSS) was computed. Using the original level set method, a DSS of $0.3 \pm 0.27$ was obtained, whereas for original region growing DSS value was $0.29 \pm 0.22$. Using the confidence guided versions, DSS significantly improved for both methods, with confidence guided region growing segmentation outperforming confidence guided level set segmentation: DSS was $0.68 \pm 0.13$ for level set and $0.69 \pm 0.14$ for region growing.

Thus, it can be concluded that intensity features are still probably the most important for brain tumor diagnosis, but LBP and HOG features also contribute for discrimination. Moreover, incorporating confidence guiding in the segmentation methods significantly improved their performances.

WMH have been associated with prominent pathologies as MS, vascular disease and dementia. For this reason, several studies focus on the analysis of these lesions, particularly on accurate segmentation. Segmentation is generally based on intensity features, although some studies have already experimented other types of descriptors. Another common objective of WMH-related studies is the discrimination between different WMH-related lesions, for instance between MS and vascular diseases. Intensity, texture and shape descriptors have been used for this purpose: intensity descriptors are usually based on statistics computed from gray-level histograms; textural features are usually based on GLCM (Loizou et al. 2011; Yu et al. 1999; Mathias et al. 1999; Zhang et al. 2008), gradient matrix (Yu et al. 1999), GLRL (Yu et al. 1999; Zhang et al. 2008), but others such as fractal (Loizou et al. 2011), Fourier (Loizou et al. 2011), shape (Loizou et al. 2011), or wavelet-based (Zhang et al. 2008) features have also been used.

Theocharakis et al. (2009) developed a system for discriminating between multiple sclerosis lesions (MS) and cerebral microangiopathies (CM), based on texture features. Four experts consensually specified 47 CM and 31 MS rectangular ROIs. From each ROI, 23 textural features were extracted: 4 features from the ROI's histogram, 14 from the co-occurrence matrices (the Haralick features) and 5 from the run-length matrices. All features were normalized to zero mean and unit variance.

For feature selection and classification, four methods were compared: minimum distance, linear discriminant analysis, logistic regression and probabilistic neural network (PNN). With a leave-one-out approach, the best classification accuracy belonged to the PNN classifier (88.46 %), using the mean value, sum of variance and run-length nonuniformity features. However, a cross-validation scheme with $\frac{2}{3}$ train – $\frac{1}{3}$ test dataset partition led to an average accuracy of 72.96 % (over 10 random partitions), with different features in the top 3 for each of the repetitions. The most frequent features on the top 3 were mean, contrast, sum of average and sum of variance (4 times the mean and 3 times the others). A total of 15 of the 23 features were at least once in the top 3 (mean, contrast, sum of average, sum of variance, angular second moment, correlation, inverse different moment, IMC2, gray-level nonuniformity, run-length nonunniformity, skewness, sum of entropy, entropy, difference variance and run percentage). A Mann-Whitney U-test was performed to assess significant difference between both classes for each of these 15 features, and all of them showed significant difference except for skewness and gray-level nonuniformity.

One important conclusion to draw from the study of Theocharakis et al. is that a combination of features computed from the gray-level histogram and the GLCM and GLRL matrices can lead to relatively high performances in the discrimination between MS and CM.

### 3.5 Further notes

For the previous review, we selected four prominent imaging modalities and, for each one, we selected a few studies that together would give the reader an idea of (1) which are the main pathologies object of CAD studies within each imaging modality, (2) which are the main image descriptors used in the analysis of such pathologies, (3) which are the main challenges associated with each scenario, and some ways of addressing them (e.g. dimensionality reduction in AD CAD, false positive reduction in lung nodule CAD). Nonetheless, in order to provide the reader with a richer insight on the state-of-the-art techniques, we believe to be of interest to briefly address prominent related medical challenges, where the performances of different algorithms over the same database can be compared. In Table 10 we list some challenges related to this review and some of the descriptors and classifiers that were used by the different teams. Moreover, we also provide the official homepages of the challenges, for further details on the proposed strategies, scores or even for dataset download.

For further information to the reader, we list in Table 11 some reference data repositories, where data related to the different imaging modalities and pathologies discussed herein (among others) can be found.

Finally, if the reader is interested in a deep comparison of the performance of the different descriptors in several different contexts/datasets, both medical and non-medical, there are some studies focusing on such task, such as the work by Nanni et al. (2013, 2014, 2015).

## 4 Discussion and conclusion

Computer-aided diagnosis has benefited from the research in image descriptors. Algorithms aimed at providing a summarized description of image regions have been an important and powerful to their daily practice. In our study, four different modalities of medical exams were included based on their practical use in healthcare contexts trying to help researchers dealing with different kinds of studies. Concerning mammograms, image descriptors were used for CAD of lesions associated with breast cancer, such as masses or calcifications. Studies usually focus on one of three tasks: (1) detecting abnormal tissue, i.e., lesions; (2) detecting

**Table 10** Descriptors and classifiers adopted in some related medical challenges

| Challenge | Descriptors | Classifiers | Homepage |
|---|---|---|---|
| CADDementia | Volumes, thicknesses and shapes | RF | http://caddementia.grand-challenge.org/ |
| (MICCAI 2014) | LBP | SVM | |
| | Gaussian filter banks | | |
| Longitudinal | Atlas-based prior models | RF | http://iacl.ece.jhu.edu/MSChallenge |
| Multiple | Voxel intensity at different smoothing scales | CNN | |
| Sclerosis lesion | Histogram | | |
| Segmentation | | | |
| Challenge | | | |
| (ISBI 2015) | | | |
| Multimodal | Voxel intensity | RF | http://braintumorsegmentation.org/ |
| Brain Tumor | Intensity statistics | SVM | |
| Image | Atlas-based prior models | CNN | |
| Segmentation | Gaussian filter banks | | |
| Challenge | Histogram | | |
| (MICCAI-BRATS 2015) | Gradient magnitude | | |
| | Laplacian | | |
| | Haralick features | | |

*CNN* convolutional neural networks

**Table 11** Some reference databases for different modalities and pathologies

| Modalities | Database |
|---|---|
| Mammographies | DDSM (Heath et al. 2000) |
| | MIAS (Suckling et al. 1994) |
| | BCDR (Guevara-López et al. 2015) |
| For several collections of Mammographies, Lung CT and PET/CT, Neuro MRI | The cancer imaging archive (Clark et al. 2013) |
| Lung CT | Japanese Society of Radiological Technology (Shiraishi et al. 2000) |
| BrainWeb Normal/MS (simulated data) | BrainWeb (Cocosco et al. 1997) |
| Challenges | Portal hosting several medical challenges: http://grand-challenge.org/allchallenges/ |

particular lesion types as masses or calcifications; (3) detecting malignant forms of lesions. The most suited image descriptors differ among tasks. For the first task, texture and shape descriptors (e.g GLRL and HOG) proved to be good choices. As for calcifications, texture descriptors were the most suitable option. Some examples of such descriptors are GLRL, curvelets, wavelets, GLCM and Gabor filters. Finally, for discriminating between benign and malignant masses, shape descriptors such as Zernike moments can be used, since shapes of malignant masses are characterized by accentuated irregularity and eccentricity. With regard

to PET images, most CAD applications are based on FDG-PET images. These are especially useful to address anomalies related to abnormally high or low glucose metabolism, as tumors or dementia, respectively. In the case of tumors, studies are usually focused on detection, diagnosis or treatment response evaluation or prediction. Intensity/SUV features, texture descriptors such as GLCM and shape descriptors such as compactness and eccentricity proved to be suitable for tumor detection. A usual problem regarding tumor studies is the false positive detection rate: some anatomical regions naturally show high uptake, and often systems cannot differentiate such regions from tumors. We observed that using CT information, as well as information related to anatomical location and symmetry measures to the medial plane can be useful to perform such differentiation. As for dementia scenarios, intensity features and texture features based on LBP can yield high diagnosis performances. It is usual to find studies on dementia that use descriptors computed from the entire brain image, instead of computing them from a segmented ROI. In that case, dimensionality reduction approaches should be considered, for example feature selection algorithms such as MIM or PBCC. In the CT context, most CAD applications are aimed at lung nodule detection or diagnosis. These applications are often affected by high false positive rates, and a considerable volume of work is directed towards its reduction. Works based on texture descriptors (e.g gray-level-histogram, gradient matrix or wavelets) and shape descriptors (e.g sphericity and compactness) proved to yield good detection performances with relatively low false positive rates. In addition, texture descriptors such as curvelets, GLCM and GLRL computed from CT images can also be very effective in the analysis of other types of tissues (e.g. besides lung tissue). Finally, from the MRI context, most CAD applications using image descriptors are related to detection, segmentation and diagnosis of brain pathologies such as brain tumors and lesions associated with white matter hyperintensity (e.g MS, dementia, or vascular pathologies). The well-known bias fields and inter- and intra-patient misalignments affecting MR images are sometimes impeding factors for accurate automated segmentation, when the latter is based solely on intensity values. LBP-based texture analysis can be more robust to bias field and rotation degradations. Regarding diagnosis, intensity features play an important role. However, texture and shape descriptors such as LBP and HOG proved to contribute for class discrimination in tumor diagnosis. Texture descriptors as the gray-level histogram, GLCM and GLRL have proven to be effective in the discrimination of lesions related to white matter hyperintensities.

Most of the works covered in this survey are in its essence feature-based classification systems. Over the last few years, classification systems based on deep learning (e.g., using deep convolutional neural networks) have been proposed as a quite successful alternative to traditional feature-based systems. Deep learning techniques, which do not require hand-crafted features, have shown to be quite effective in tasks such as object and speech recognition or natural language processing. But the revolution brought by deep learning has not made feature-based systems obsolete. There are cases where these traditional systems are more adequate, either because they are a more efficient solution or because they do not provide a black-box approach. In particular, the use of image descriptors in a healthcare is still a very active and promising research field. As such, it is expected that the design of novel image descriptors will happen. These contributions can either belong to the classical descriptor families or somehow help to define new families. Exploring the advantages of combining descriptor information with other clinical data seems to be a trend in the field. While some works have already shown the benefits of such approach, there is still room for improvement.

## References


Bagui SC, Bagui S, Pal K, Pal NR (2003) Breast cancer detection using rank nearest neighbor classification rules. Pattern Recognit 36(1):25–34

Belkasim S, Shridhar M, Ahmadi M (1991) Pattern recognition with moment invariants: a comparative study and new results. Pattern Recognit 24(12):1117–1138

Bicacro E, Silveira M, Marques JS (2012) Alternative feature extraction methods in 3D brain image-based diagnosis of Alzheimer's disease. In: Proceedings of the 19th IEEE international conference on image processing (ICIP). pp 1237–1240

Boroczky L, Luyin Z, Lee K (2005) Feature subset selection for improving the performance of false positive reduction in lung nodule cad. In: Proceedings of the 18th IEEE symposium on computer-based medical systems. pp 85–90

Brown MS, McNitt-Gray MF, Goldin JG, Suh RD, Sayre JW, Aberle DR (2001) Patient-specific models for lung nodule detection and surveillance in CT images. IEEE Trans Med Imaging 20(12):1242–1250

Buciu I, Gacsadi A (2011) Directional features for automatic tumor classification of mammogram images. Biomed Signal Process Control 6(4):370–378

Cabral TM, Rangayyan RM (2012) Analysis of breast masses in mammograms. Synth Lectures Biomed Eng 7(3):1–118

Candès E, Demanet L, Donoho D, Ying L (2006) Fast discrete curvelet transforms. Multiscale Model Simul 5(3):861–899

Canterakis N (1999) 3D Zernike moments and Zernike affine invariants for 3D image analysis. In: Proceedings of the 11th scandinavian conference on image analysis, pp 85–93

Christoyianni I, Dermatas E, Kokkinakis G (2000) Fast detection of masses in computer-aided mammography. IEEE Signal Process Mag 17(1):54–64

Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F (2013) The cancer imaging archive (TCIA): maintaining and operating a public information repository. J Digit Imaging 25(6):1045–1057

Cocosco CA, Kollokian V, Kwan RK, Evans AC (1997) BrainWeb: online interface to a 3D MRI simulated brain database. Neuroimage 5(4):1045–1057

Constantinidis AS, Fairhurst MC, Rahman AFR (2001) A new multi-expert decision combination algorithm and its application to the detection of circumscribed masses in digital mammograms. Pattern Recognit 34(8):1527–1537

Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: Proceedings of the 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). pp 886–893

Davatzikos AC, Fan AY, Wu AX, Shen AD, Resnick BSM (2006) Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. Neurobiol Aging 29(4):514–523

Depeursinge A, Sage D, Hidki A, Platon A, Poletti P, Unser M, Muller H (2007) Lung tissue classification using wavelet frames. In: Proceedings of the 29th annual international conference of the IEEE engineering in medicine and biology society. pp 6259–6262

Dettori L, Semler L (2007) A comparison of wavelet, ridgelet, and curvelet-based texture classification algorithms in computed tomography. Comput Biol Med 37(4):486–498

Dhawan A (1996) Analysis of mammographic microcalcifications using gray-level image structure features. IEEE Trans Med Imaging 15(3):246–259

Dua S, Singh H, Thompson HW (2009) Associative classification of mammograms using weighted rules. Expert Syst Appl 36(5):9250–9259

Eltoukhy MM, Faye I, Samir BB (2010) A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram. Comput Biol Med 40(4):384–391

Eshelman LJ (1991) The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. Found Genet Algorithms 1:265–283

Fan Y, Resnick SM, Wu X, Davatzikos C (2008) Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study. Neuroimage 41(2):277–285

Fehr J (2007) Rotational invariant uniform local binary patterns for full 3d volume texture analysis. In: Proceedings of FinSig, 6 pp

Feng M, Reed TR (2007) Motion estimation in the 3-D Gabor domain. IEEE Trans Image Process 16(8):2038–2047

Ferreira CBR, Borges DL (2003) Analysis of mammogram classification using a wavelet transform decomposition. Pattern Recognit Lett 24(7):973–982

Galloway MM (1975) Texture analysis using gray level run lengths. Comput Graph Image Process 4(2):172–179

Gerardin E, Chetelat G, Chupin M, Cuingnet R, Desgranges B, Kim H, Niethammer M, Dubois B, Lehericy S, Garnero L, Eustache F, Colliot O (2009) Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. Neuroimage 47(4):1476–1486

Guevara-López MA, Posada NG, Moura DC, Pollán RR, José M (2015) BCDR: a breast cancer digital repository. In: Proceedings of the 15th international conference on experimental mechanics. pp 1–5

Guliato D, de Carvalho JD, Rangayyan RM, Santiago SA (2008) Feature extraction from a signature based on the turning angle function for the classification of breast tumors. J Digit Imaging 21:129–144

Haralick RM, Shanmuga K, Dinstein I (1973) Textural features for image classification. IEEE Trans Syst Man Cybern 3(6):610–621

Heath M, Bowyer K, Kopans D, Moore R, Jr, PK (2000) The digital database for screening mammography. In: Proceedings of the 5th international workshop on digital mammography. pp 212–218

Herlidou-Même S, Constans J, Carsin B, Olivie D, Eliat P, Nadal-Desbarats L, Gondry C, Rumeur EL, Idy-Peretti I, de Certaines J (2003) MRI texture analysis on texture test objects, normal brain and intracranial tumors. Magn Reson Imaging 21(9):989–993

Hill DLG, Batchelor PG, Holden M, Hawkes DJ (2001) Medical image registration. Phys Med Biol 46(3):R1–R45

Hu MK (1962) Visual-pattern recognition by moment invariants. IRE Trans Inf Theory 8(2):179–187

Huo Z, Giger ML, Vyborny CJ, Wolverton DE, Schmidt RA, Doi K (1998) Automated computerized classification of malignant and benign masses on digitized mammograms. Acad Radiol 5(3):155–168

Iftekharuddin KM, Zheng J, Islam MA, Ogg RJ (2009) Fractal-based brain tumor detection in multimodal MRI. Appl Math Comput 207(1):23–41

Jack CR, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, Borowski B, Britson PJ, Whitwell JL, Ward C, Dale AM, Felmlee JP, Gunter JL, Hill DL, Killiany R, Schuff N, Fox-Bosetti S, Lin C, Studholme C, DeCarli CS, Krueger G, Ward HA, Metzger GJ, Scott KT, Mallozzi R, Blezek D, Levy J, Debbins JP, Fleisher AS, Albert M, Green R, Bartzokis G, Glover G, Mugler J, Weiner MW (2008) The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J Magnet Reson Imaging 27(4):685–691

Jégou H, Perronnin F, Douze M, Sánchez J, Pérez P, Schmid C (2011) Aggregating local image descriptors into compact codes. IEEE Trans Pattern Anal Mach Intell 34(9):1704–1716

Kato N, Fukui M, Isozaki T (2009) Bag-of-features approach for improvement of lung tissue classification in diffuse lung disease. In: Proceedings of the SPIE 7260, medical imaging 2009: computer-aided diagnosis, vol 7260. pp 1–10

Kim JK, Park H (1999) Statistical textural features for detection of microcalcifications in digitized mammograms. IEEE Trans Med Imaging 18(3):231–238

Kläser A, Marszalek M, Schmid C (2008) A spatio-temporal descriptor based on 3D-gradients. In: Proceedings of the British machine vision 2008 (BMVC'08), 10 pp

Ko JP, Betke M (2001) Chest CT: automated nodule detection and assessment of change over time-preliminary experience. Radiology 218(1):267–273

Loizou CP, Kyriacou EC, Seimenis I, Pantziaris M, Christodoulos C, Pattichis CS (2011) Brain white matter lesions classification in multiple sclerosis subjects for the prognosis of future disability. In: Iliadis L, Maglogiannis I, Papadopoulos H (eds) Artificial intelligence applications and innovations, IFIP advances in information and communication technology, vol 364. Springer, Berlin, pp 400–409

Lowe DG (1999) Object recognition from local scale-invariant features. In: Proceedings of the 7th IEEE international conference on computer vision (ICCV'99). pp 1150–1157

Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60:91–110

Madabhushi A, Feldman MD, Metaxas DN, Tomaszeweski J, Chute D (2005) Automated detection of prostatic adenocarcinoma from high-resolution ex vivo MRI. IEEE Trans Med Imaging 24(12):1611–1625

Marchette D, Priebe CE, Julin E, Rogers G, Solka JL (1994) The filtered kernel estimator. In: Proceedings of the 26th symposium on the interface. p 25

Mathias J, Tofts P, Losseff NA (1999) Texture analysis of spinal cord pathology in multiple sclerosis. Magn Reson Med 42(5):929–935

McNitt-Gray M, Wyckoff N, Sayre J, Goldin J, Aberle D (1999) The effects of co-occurrence matrix based texture parameters on the classification of solitary pulmonary nodules imaged on computed tomography. Comput Med Imaging Graph 23(6):339–348

Meinel LA, Stolpen AH, Berbaum KS, Fajardo LL, Reinhardt JM (2007) Breast MRI lesion classification: improved performance of human readers with a backpropagation neural network computer-aided diagnosis (CAD) system. J Magn Reson Imaging 25(1):89–95

Mikolajczyk K, Leibe B, Schiele B (2005) Local features for object class recognition. In: Proceedings of the 10th IEEE international conference on computer vison (ICCV'05). pp 1792–1799

Mikolajczyk K, Schmid C (2005) A performance evaluation of local descriptors. IEEE Trans Pattern Anal Mach Intell 27(10):1615–1630

Morgado PM (2012) Automated diagnosis of Alzheimer's disease using PET images a study of alternative procedures for feature extraction and selection electrical and computer engineering. Master's thesis, MSc thesis at Electrical and Computer Engineering Dep., Higher technical institute, Technical University of Lisbon

Moura DC, Guevara-López MA (2013) An evaluation of image descriptors combined with clinical data for breast cancer diagnosis. Int J Comput Assist Radiol Surg 8(4):561–574

Mu T, Nandi AK, Rangayyan RM (2008) Classification of breast masses using selected shape, edge-sharpness, and texture features with linear and kernel-based classifiers. J Digit Imaging 21(2):153–169

Murphy K, van Ginneken B, Schilham A, de Hoop B, Gietema H, Prokop M (2009) A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification. Med Image Anal 13(5):757–770

Nanni L, Brahnam S, Ghidoni S, Menegatti E (2014) Region-based approaches and descriptors extracted from the co-occurrence matrix. Int J Latest Res Sci Technol 3(6):192–200

Nanni L, Brahnam S, Ghidoni S, Menegatti E (2015) Improving the descriptors extracted from the co-occurrence matrix using preprocessing approaches. Expert Syst Appl 42(22):8989–9000

Nanni L, Brahnam S, Ghidoni S, Menegatti E, Barrier T (2013) Different approaches for extracting information from the co-occurrence matrix. PLoS One 8(12):1–9

Naqa IE, Grigsby P, Apte A, Kidd E, Donnelly E, Khullar D, Chaudhari S, Yang D, Schmitt M, Laforest R, Thorstad W, Deasy J (2009) Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. Pattern Recognit 42(6):1162–1171

Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987

Oliveira FP, Tavares JM (2014) Medical image registration: a review. Comput Methods Biomech Biomed Eng 17(2):73–93

Oliver A, Torrent A, Llado X, Tortajada M, Tortajada L, Sentís M, Freixenet J, Zwiggelaar R (2012) Automatic microcalcification and cluster detection for digital and digitised mammograms. Knowledge-Based Syst 28:68–75

Porat M, Zeevi YY (1989) Localized texture processing in vision: analysis and synthesis in the gaborian space. IEEE Trans Biomed Eng 36(1):115–129

Priebe C, Julin E, Rogers G, Healy D, Lu J, Solka JL, Marchette D (1994) Incorporating segmentation boundaries into the calculation of fractal dimension features. In: Proceedings of the 26th symposium on the interface. pp 52–56

Ramírez J, Gorriz J, Salas-Gonzalez D, Romero A, Lopez M, Alvarez I, Gomez-Río M (2013) Computer-aided diagnosis of Alzheimer's type dementia combining support vector machines and discriminant set of features. Inf Sci 237:59–72

Ramos-Pollan R, Guevara-Lopez MA, Suarez-Ortega C, Díaz-Herrero G, Franco-Valiente JM, del Solar MR, de Posada NG, Vaz MAP, Loureiro J, Ramos I (2012) Discovering mammography-based machine learning classifiers for breast cancer diagnosis. J Med Syst 36(4):2259–2269

Rashed EA, Ismail IA, Zaki SI (2007) Multiresolution mammogram analysis in multilevel decomposition. Pattern Recognit Lett 28(2):286–292

Reddy KK, Solmaz B, Yan P, Avgeropoulos NG, Rippe DJ, Shah M (2012) Confidence guided enhancing brain tumor segmentation in multi-parametric mri. In: Proceedings of the international symposium on biomedical imaging. pp 366–369

Rojas-Domínguez A, Nandi AK (2009) Development of tolerant features for characterization of masses in mammograms. Comput Biol Med 39(8):678–688

Sahiner B, Chan H, Petrick N, Helvie MA, Hadjiiski LM (2001) Improvement of mammographic mass characterization using spiculation measures and morphological features. Med Phys 28(7):1455–1465

Scovanner P, Ali S, Shah M (2007) A 3-dimensional sift descriptor and its application to action recognition. In: Proceedings of the 15th ACM international conference on multimedia, MM '07. pp 357–360

Sharma S, Khanna P (2015) Computer-aided diagnosis of malignant mammograms using Zernike moments and SVM. J Digit Imaging 28(1):77–90

Sheshadri H, Kandaswamy A (2007) Experimental investigation on breast tissue classification based on statistical feature extraction of mammograms. Comput Med Imaging Graph 31(1):46–48

Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu K, Matsui M, Fujita H, Kedera Y, Doi K (2000) Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. Am J Roentgenol 174(12):71–74

Silveira M, Marques J (2010) Boosting alzheimer disease diagnosis using pet images. In: Proceedings of the 20th international conference on pattern recognition (ICPR). pp 2556–2559

Soltanian-Zadeh H, Rafiee-Rad F, Pourabdollah-Nejad SD (2004) Comparison of multiwavelet, wavelet, haralick, and shape features for microcalcification classification in mammograms. Pattern Recognit 37(10):1973–1986

Sotiras A, Davatzikos C, Paragios N (2013) Deformable medical image registration: a survey. IEEE Trans Med Imaging 32(7):1153–1190

Suckling J, Parker J, Dance DR, Astley SM, Hutt I, Boggis CRM, Ricketts I, Stamatakis E, Cerneaz N, Kok SL, Taylor P, Betal D, Savage J (1994) The mammographic image analysis society digital mammogram database. In: Proceedings of the international workshop on digital mammography. pp 211–221

Tahmasbi A, Saki F, Shokouhi SB (2010) Mass diagnosis in mammography images using novel ftrd features. In: Proceedings of the 17th Iranian conference of biomedical engineering (ICBME). pp 1–5

Tahmasbi A, Saki F, Shokouhi SB (2011) Classification of benign and malignant masses based on Zernike moments. Comput Biol Med 41(8):726–735

Teague MR (1980) Image-analysis via the general-theory of moments. J Opt Soc Am 70(8):920–930

Theocharakis P, Glotsos D, Kalatzis I, Kostopoulos S, Georgiadis P, Sifaki K, Tsakouridou K, Malamas M, Delibasis G, Cavouras D, Nikiforidis G (2009) Pattern recognition system for the discrimination of multiple sclerosis from cerebral microangiopathy lesions based on texture analysis of magnetic resonance images. Magn Reson Imaging 27(3):417–422

Tola E, Lepetit V, Fua P (2010) Daisy: an efficient dense descriptor applied to wide baseline stereo. IEEE Trans Pattern Anal Mach Intell 32(5):815–830

Tsai F, Chang CK, Rau JY, Lin TH, Liu GR (2007) 3D computation of gray level co-occurrence in hyperspectral image cubes. In: International workshop on energy minimization methods in computer vision and pattern recognition. Springer, Berlin, pp 429–440

Unay D, Ekin A, Cetin M, Jasinschi R, Ercil A (2007) Robustness of local binary patterns in brain mr image analysis. In: Proceedings of the IEEE engineering in medicine and biology society annual conference, vol 2007. pp 2098–2101

Viola P, Jones M (2001) Rapid object detection using a boosted cascade of simple features. In: Proceedings of the conference on computer vision and pattern recognition. pp 511–518

Wang D, Shi L, Heng PA (2009) Automatic detection of breast cancers in mammograms using structured support vector machines. Neurocomputing 72(13–15):3296–3302

Wiemker R, Rogalla P, Zwartkruis A, Blaffert T (2002) Computer-aided lung nodule detection on high-resolution ct data. In: Proceedings of the SPIE 4684, medical imaging 2002: image processing. pp 677–688

Wu B, Khong P, Chan T (2012) Automatic detection and classification of nasopharyngeal carcinoma on PET/CT with support vector machine. Int J Comput Assist Radiol Surg 7(4):635–646

Xu D, Lee J, Raicu D, Furst J, Channin D (2005) Texture classification of normal tissues in computed tomography. In: Proceedings of the 2005 annual meeting of the society for computer applications in radiology. Orlando, Florida

Yu H, Caldwell C, Mah K, Mozeg D (2009) Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning. IEEE Trans Med Imaging 28(3):374–383

Yu O, Mauss Y, Zollner G, Namer I, Chambron J (1999) Distinct patterns of active and non-active plaques using texture analysis on brain NMR images in multiple sclerosis patients: preliminary results. Magn Reson Imaging 17(9):1261–1267

Yu S, Guan L (2000) A CAD system for the automatic detection of clustered microcalcifications in digitized mammogram films. IEEE Trans Med Imaging 19(2):115–126

Zacharaki EI, Wang S, Chawla S, Yoo DS, Wolf R, Melhem ER, Davatzikos C (2009) Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn Reson Med 62(6):1609–1618

Zhang J, Tong L, Wang L, Li N (2008) Texture analysis of multiple sclerosis: a comparative study. Magn Reson Imaging 26(8):1160–1166

Zitová B, Flusser J (2003) Image registration methods: a survey. Image Vis Comput 21(11):977–1000