Rotation Invariant Features Based on Three Dimensional Gaussian Markov Random Fields for Volumetric Texture Classification

Yasseen Almakady[a,**], Sasan Mahmoodi[a], Joy Conway[b,c], Michael Bennett[c]

[a]*Faculty of Engineering and Physical Sciences, Electronics and Computer Science, University of Southampton, Highfield Campus, Southampton, SO17 1BJ, United Kingdom.*
[b]*Faculty of Health Sciences, University of Southampton, Highfield Campus, Southampton, SO17 1BJ, United Kingdom.*
[c]*Southampton NIHR Respiratory and Critical Care Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust, Southampton SO16 6YD, United Kingdom .*

ABSTRACT

This paper proposes a set of rotation invariant features based on three dimensional Gaussian Markov Random Fields (3D-GMRF) for volumetric texture image classification. In the method proposed here, the mathematical notion of spherical harmonics is employed to produce a set of features which are used to construct the rotation invariant descriptor. Our proposed method is evaluated and compared with other method in the literature for datasets containing synthetic textures as well as medical images. The results of our experiments demonstrate excellent classification performance for our proposed method compared with state-of-the-art methods. Furthermore, our method is evaluated using a clinical dataset and show good performance in discriminating between healthy individuals and COPD patients. Our method also performs well in classifying lung nodules in the LIDC-IDRI dataset. Our results indicate that our 3D-GMRF-based method enjoys more superior performance compared with other methods in the literature.

1. Introduction

The presence of texture in our natural world demonstrates its importance on the characterization of objects or regions. Textures therefore exhibit rich information about our surroundings. In computer vision, textures provide essential visual cues about objects appearance and properties. They play a key role in many applications, including remote sensing, medical image analysis, object recognition, and automated inspection. Due to the significance of textures in computer vision discipline, various methods have been developed for texture analysis. Such techniques can be broadly categorized into statistical, structural, spectral, and model-based methods. They generally aim to extract features for texture description in order to classify, segment and synthesize textures (Tuceryan and Jain, 1993). Most of these methods are initially developed to characterize textures existing in two-dimensional images. However, recent advances in three-dimensional (3D) imaging technology, particularly in medical imaging, and the increasing availability of high-resolution images, require new methods for 3D texture analysis since 2D texture analysis methods ignore some valuable information contained in 3D textures (Kovalev et al., 2001). Consequently, many of the methods that are originally proposed for 2D texture images have been extended to characterize textures in 3D images. Nevertheless, the extension of these methods into 3D is not trivial and faces some challenges, such as issues regarding voxels sampling, computation time, and rotation invariance. Markov Random Fields (MRFs) have been introduced about two decades ago as a tool for solving visual perception problems (Wang et al., 2013). One important type of MRFs is the Gaussian Markov Random Fields (GMRF) which is among texture analysis methods that has been proved to demonstrate excellent performance in characterizing textures in 2D images, and are then employed for texture classification (Chellappa and Chatterjee, 1985; Dharmagunawardhana et al., 2016) and segmentation (Dharmagunawardhana et al., 2014; Manjunath and Chellappa, 1991; Mahmoodi and Gunn, 2011; Xia et al., 2006). GMRF model parameters can capture the essential structures of texture, and it has been a popular choice for modelling textures

**Corresponding author:
e-mail: yham1n15@soton.ac.uk (Yasseen Almakady)

(Tuceryan and Jain, 1993). GMRF extension to characterize textures in volumetric images has not received much attention, hence, exploring this would be beneficial in many applications, especially in medical image analysis applications in which rich texture information is available in the form of volumetric images.

The main challenge in texture analysis is to deal with potential changes such as rotation, scales, and translation in texture images. In the natural world, textures may be observed at various scales and orientations. An optimal texture descriptor, therefore, should be invariant to transformations such as scaling and rotation to enable us to develop efficient methods for texture analysis. Rotation invariance is an essential property of any efficient texture analysis method because many computer vision problems presume images are captured at random orientations. Various methods have been developed to achieve rotation invariance for 2D textures. However, extending these methods for 3D texture is far from being a simple task. Optimal rotation invariance requires equidistance sampling, which is a straightforward task on the circular neighborhood of a 2D texture. However, such a task becomes complex on a sphere representing a neighborhood in a 3D texture. Another issue relates to the order of sampled points, which is easily manageable in the circle but difficult to deal with on the surface of a sphere due to the dimensionality. These issues establish the need for developing new methods for 3D texture analysis, enjoying a rotation invariance property.

Existing methods of 3D texture analysis handle the rotation invariance problem in a few ways. In the method proposed by Fehr and Burkhardt (2008), the spherical harmonics function are employed to achieve the rotation invariance by computing the minimization of the full correlation over all angles between grey-values of all points on a spherical surface, with a given radius and the volume representation. These points are weighted in an arbitrary but fixed order to provide a fast method for a correlation across all angles. Another method based on spherical harmonics is proposed by Banerjee et al. (2012) to describe 3D regions; however, unlike the method of Fehr and Burkhardt (2008), the rotation invariant descriptors are constructed by computing a set of histograms for each voxel on a sphere equivalent to a set of variables that represent frequency components of local binary patterns (LBP) codes in a spherical harmonics domain. By measuring the distance between two histograms, the similarity between two regions can then be estimated. Based on a region growing algorithm and uniform patterns, an extension to the 3D LBP method is proposed in (Paulhac et al., 2008) by searching for uniform regions located on the surface of a sphere. A novel method is introduced by Kazhdan et al. (2003) to construct a rotation invariant descriptor based on the spherical harmonic representation of 3D shapes. This method depends on the amount of energy measured within each frequency for each sphere in concentric spheres of different radii. The function defined in a spherical coordinate system is decomposed into its harmonics, and these harmonics are then summarized within each frequency. The L_2-norm is computed for each frequency component to produce 2D descriptors indexed by radius and frequency. Although this method suffers from infor-

mation loss due to the possibility of having the same harmonic reorientations for two different models, it proves to display better results in comparison with other descriptors. Despite the fact that this method describes the 3D shapes rather than volumetric texture; the idea of using spherical harmonics to achieve rotation invariance is adopted by Banerjee et al. (2012) to construct a rotation invariant texture descriptor. Additionally, the method presented by Cid et al. (2017) attempts to classify 3D solid textures, based on so-called 'steerability', a key property of the Riesz wavelets. the steerability property indicates that the local response of a randomly rotated image can be obtained analytically from a linear combination of the responses of all components of the filter bank without the requirement of a rotated filter. Depeursinge et al. (2018) investigate the local rotation invariance (LRI) and directional sensitivity (DS) of radiomics features stating that optimal radiomics image operators must combine LRI with DS. To achieve this, invariant texture operators based on spherical harmonic wavelet (SHW) are proposed to combine LRI with DS which are then employed for 3D texture classification. The significance of the local directional information is also exploited in (El khadiri et al., 2018) to develop the local directional ternary pattern (LDTP). In this method, features are extracted from the local ternary pattern (LTP) and combined with the directional features extracted from local directional patterns (LDP) to construct a new descriptor for texture classification. Han et al. (2015) calculate Haralick features of 3D texture extracted from Lung CT images and utilize them for pulmonary nodules classification.

In addition to these methods, the deep convolutional neural networks (CNN) traditionally achieved rotation invariance by learning through rotation-based data augmentation. Such a method becomes costly for 3D images due to an intractable number of rotation arrangements in a 3D space. Therefore, rotation invariance has recently been exploited in CNN for various problems, including texture analysis. Winkels and Cohen (2019) introduce 3D CNNs with group convolutions (3D G-CNNs) by exploiting the prior knowledge found in images such as orientation to perform rotation invariance. Worrall et al. (2017) investigate the issue of CNNs being not rotation invariant and present the Harmonic Networks (H-Nets) to achieve rotation invariance. This is carried out by replacing the CNN filter with the circular harmonic that represents all rotated versions of the filter. Andrearczyk and Depeursinge (2018) evaluate the employment of equivariance and invariance to rotations to classify 3D textures and conclude that it is beneficial to classify 3D textures by including built-in equivariance and invariance in CNNs, but this involves considerable computational complexity.

The choice of feature extraction from 3D images using 3D approaches rather than 2D approaches has been investigated in (Almakady et al., 2018; Paulhac et al., 2008; Yan et al., 2016; Griffiths and Boehm, 2019). In medical image analysis, the development of 3D approaches is curial since most medical imaging modalities produce volumetric images. However, it is suggested to use CNN for 2D slices of a 3D volumetric texture to provide more training data. Such an approach could tackle the issue of a small number of training data (Xie et al.,

2019). Although 3D CNN is numerically expensive in terms of time and memory, it achieves better results compared with 2D CNN because it captures features in 3D space (Yan et al., 2016). In the method presented in (Xie et al., 2019), semi-supervised adversarial classification (SSAC) model is proposed using labelled and unlabelled data which is employed for benign and malignant lung nodule classification in chest CT scans. Unlabelled data in this method is used to minimize the requirement for data annotation by high-skilled professionals that are not always available. Hussein et al. (2017) propose 3D CNN multi-task learning for lung nodules characterization. This method exploits the volumetric information provided by 3D images of the CT scan. Such information is lost when 2D slices are used. The method is then applied for lung nodules classification. The problem of nodule classification is also investigated in (Shen et al., 2017) where a method based on multi-crop convolutional neural networks (MC-CNN) is introduced for lung nodule malignancy classification. This method utilizes a multi-crop pooling strategy that captures nodule salient information and uses it for lung nodule classification.

Despite the fact that CNNs have made remarkable improvements in image-related tasks such as image classification and segmentation (Xie et al., 2019; Litjens et al., 2017; Xie et al., 2018), there are substantial remaining challenges which could limit the applications of deep learning-based methods in the medical field. This is because CNNs require a large-scale annotated dataset for training to achieve better performance which usually does not exist in the medical domain (Kumar et al., 2016). In medical imaging, it is challenging to obtain large amount of data due to ethical issues associated with patient privacy (Winkels and Cohen, 2018). Furthermore, the lack of sufficient annotated data complicates these challenges as these datasets require accurate annotations by high-skilled experts which is a time consuming task (Tajbakhsh et al., 2016). Besides, the training stage in CNNs requires extensive computational and memory resources. Although utilizing transfer learning overcomes the issue associated with small training dataset, its performance could be influenced by different medical applications (Tajbakhsh et al., 2016). Such complications can be avoided by methods based on handcrafted features to become beneficial for medical image applications in the absence of extensive training data and powerful hardware.

In this paper, we propose a rotation invariant descriptor for volumetric texture classification by extending the descriptor proposed in (Almakady et al., 2018). Unlike deep learning-based methods, our proposed method does not require large training data or powerful hardware, which makes it suitable for medical image applications where datasets are usually small. The main contribution of this paper is to achieve the rotation invariance by employing spherical harmonics to remove the effect of rotation for 3D-GMRF parameters. Unlike the descriptor previously proposed in (Almakady et al., 2018), the size of the descriptor in the method proposed here does not increase if more voxels are sampled. This is because the concatenated histograms are computed in terms of spherical harmonic frequencies rather than 3D-GMRF estimated parameters which leads to a stable computation time. Our proposed method is suitable for
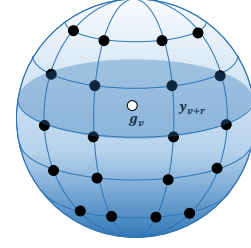
Fig. 1. Graphical representation of voxels y_{v+r} sampled over a sphere surface where g_v is the central voxel.

tasks that deal with texture features of volumetric images with any rotated texture patterns. Since volumetric images are commonly produced and used in medical imaging applications, our method can be used in disease detection by classifying discriminative texture features extracted from 3D medical images. In some diseases, the pathological changes can be characterized by alterations affecting the texture of tissue, hence, capturing these textural alterations under various rotation can help to detect the disease. In such cases, our proposed method with the rotation invariance property can be exploited to recognise texture patterns that characterize the diseases under any random rotation. The rest of this paper is organized as follows. A brief background about 3D-GMRF is given in Section 2. Section 3 introduces our proposed method. Section 4 discusses the results and performance measurements with comparisons to other methods. The evaluations of the methods on COPD detection and lung nodule classification are presented in section 5, and finally, the paper is concluded in section 6.

2. Background

2.1. Three-dimensional Gaussian Markov Random Field Model (3D-GMRF)

GMRF has seen various applications in image processing and analysis including volumetric texture analysis. The 3D-GMRF model parameters are utilized as features which have been used for volumetric texture classification (Almakady et al., 2018) and segmentation (Almakady et al., 2019).

In the method presented in (Almakady et al., 2018), the 2D-GMRF based method proposed in (Dharmagunawardhana et al., 2016) is extended to 3D-GMRF for volumetric texture classification. In this method, the 3D-GMRF model is generated at each voxel, where the estimated parameters of the model are employed as texture features in addition to the mean of a processed image region. Let $\Omega = \{v = (i, j, k) \mid 1 \leq i \leq H, 1 \leq j \leq W, 1 \leq k \leq D\}$ denote the set of points indexed by (i, j, k) on a $H \times W \times D$ 3D lattice corresponding to voxels in three-dimensional image volume. The local conditional probability density function of the intensity value g_v at location v is defined by:

$$p(g_v|y_{v+r}, r \in V_v) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2\sigma^2}\left(g_v - \lambda - \sum_{r \in V_v} \alpha_r(y_{v+r} - \lambda)\right)^2\right\}$$

$$(1)$$

where V_v is a neighborhood of the voxel at location v, α_r is a set of interaction coefficients that represent the influence on a voxel by intensity values y_{v+r} located at a relative position r in the neighborhood (Dharmagunawardhana et al., 2016; Petrou and Sevilla, 2006). The neighborhood scheme adopted here is voxels sampled over a sphere surface with radius R, so that $V_v \in \{\theta, \phi \mid 0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi\}$ and $|V_v|$ is equivalent to the number of voxels. Fig 1 shows the sampled voxels over a sphere surface equivalent to the neighbors y_{v+r}, where the center of the sphere corresponds to the gray scale of the central voxel g_v. The model parameters in equation (1) are estimated by using the maximum likelihood estimation (MLE) and is found by taking the partial derivative of a log-likelihood function with respect to α_r, σ^2 and setting it to zero. The partial derivative of the log-likelihood function calculated based on (1), with respect to λ leads to a non-linear equation. Therefore, in this paper, λ being the mean intensity of the neighborhood patch, is computed separately for each neighborhood patch and is subtracted from the intensity of the patch to avoid dealing with non-linear equations and the issues associated with them. Parameters α_r and σ^2 are then calculated from patches whose mean is subtracted beforehand. Such a technique would then lead us to a set of linear equations for each voxel,

$$\alpha_v = \left(\sum_{v \in \Omega_v} \mathbf{y}_v \mathbf{y}_v^T + c^2 I \right)^{-1} \left(\sum_{v \in \Omega_v} \mathbf{y}_v g_v \right) \tag{2}$$

$$\sigma_v^2 = \frac{1}{|\Omega_v|} \sum_{v \in \Omega_v} (g_v - \alpha_v \mathbf{y}_v)^2 \tag{3}$$

where $\alpha_v = row[\alpha_r]$ and $\mathbf{y}_v = col[y_{v+r}]$ for $r \in V_v$, I is the identity matrix with the same size of the observation matrix $\mathbf{y}_v \mathbf{y}_v^T$, the superscript T is the transpose operation, Ω_v is the neighborhood patch known as estimation cube and c is a constant number to control the strength of regularization (Friedman et al., 2001; Bjorkstrom, 2001) . The constant c is a regularization parameter to ensure that the observation matrix is non-singular and invertible. The value of c is empirically chosen to maximize the classification accuracy. Parameters estimation is conducted locally at each voxel resulting in a feature vector $\mathbf{f}_v = \{\alpha_v, \sigma_v^2, \lambda_v\}$ for each voxel location v. A descriptor is therefore constructed by computing a histogram for each of calculated parameters $\{\alpha_r, \sigma^2, \lambda\}$ over the whole volume. The calculated histograms are then concatenated to form a descriptor for the entire volume. However, this descriptor is not rotation invariant since rotating the texture results in a rotation of neighbors y_{v+r} leading to some changes in relevant parameters α_r . Hence, the same process of computing the histograms of calculated parameters leads to different descriptors in comparison with the descriptors computed before the rotation.

3. Rotation Invariant Descriptor

If a descriptor does not change under any arbitrary rotation, it is known as a rotation invariant descriptor. However, as previously stated, constructing such a descriptor is not a trivial task in a spherical neighborhood. To achieve rotation invariance,

Table 1. Number of classes in each category.

Category name	Classes
interpolated	30
Geometric	25
Fourier	15
Mixed texture	25
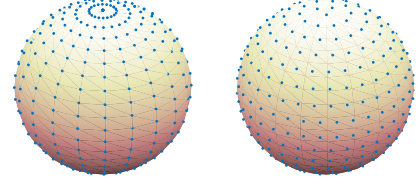
Fig. 2. Illustration of different sampling systems: On the left, points are sampled on the surface of the sphere to result in unequally spaced points. On the right, the points are uniformly distributed on the surface of the sphere.

we decompose model parametres as a linear combination of orthogonal spherical harmonics in this paper to eliminate the effect of rotating textures that particularly results in changing the interaction parameters α_r. Before applying spherical harmonics, it is important to retain the locations of voxels under any random rotation. Such a procedure can be achieved by distributing equidistant points on the neighborhood sphere surrounding the central voxel g_v.

3.1. Equidistant Spherical Sampling

A significant step towards achieving rotation invariance in 2D texture methods proposed by Ojala et al. (2002), is the distribution of P equally spaced points indexed by p on the circumference of a circle with radius R. The coordinates of such points can be simply given by $(-Rsin(2\pi p/P), Rcos(2\pi p/P))$. Following the same notion, the uniform distribution of points on a sphere is essential to achieve rotation invariance for their indices. The solution proposed by Deserno (2004) is adopted in this paper where equidistant points are regularly placed on the surface of a sphere. The circles of latitude are defined at a constant interval d_θ, while points on those circles are also sampled at constant intervals d_ϕ such that $d_\theta \simeq d_\phi$ and,

$$d_\theta d_\phi = 4\pi r^2/N \tag{4}$$

where r is the radius of a sphere, N is the number of points required to be uniformly distributed on the surface of a sphere. Fig 2 (right) shows the arrangement of points in an equidistant sampling scheme, resulting in uniformly distributed points on the surface of the sphere. In contrast, the popular sampling scheme samples points on ϕ and θ with equal distance; however, the conversion to the Cartesian coordinate system results in points bunched around the poles ($\theta = \pi, \theta = 0$) where it is sparse around ($\theta = \pi/2$), this sampling scheme then leads to the incorrect distribution of points (See e.g. Fig. 2 (left)).
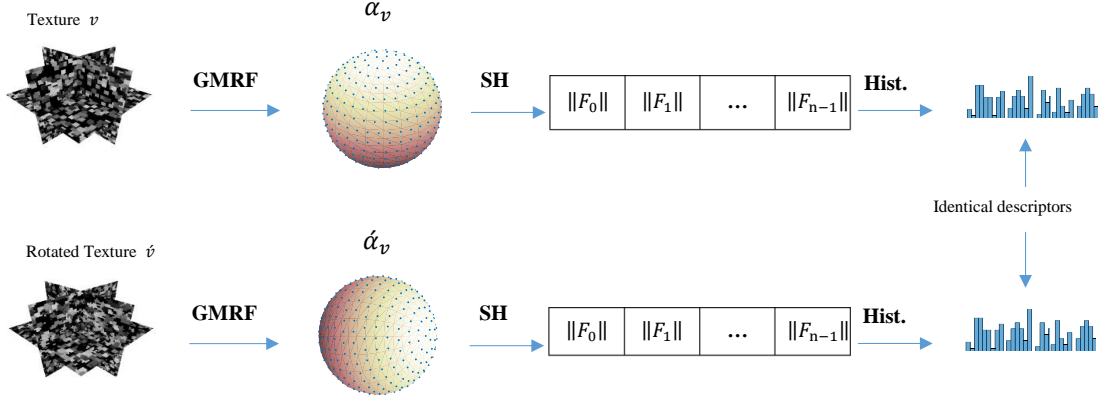
Fig. 3. Example of constructing rotation invariant descriptors for normal and rotated samples.

3.2. Spherical Harmonics

Spherical harmonics (SH) are a set of orthogonal functions with which any function in a spherical coordinate system can be linearly decomposed. They are analogous to Fourier kernels in Fourier transformation. The transformation with spherical harmonics results in complex valued coefficients (Green, 2003). For given sampled points on a unit sphere, the real $Y_{R\ell}^m(\theta, \phi)$ and imaginary $Y_{I\ell}^m(\theta, \phi)$ parts of spherical harmonic functions with degree ℓ and order m are defined by:

$$Y_{R\ell}^m = \begin{cases} \sqrt{2}K_\ell^m cos(m\phi)P_\ell^m(cos\theta) & m \neq 0 \\ K_\ell^0 P_\ell^0(cos\theta) & m = 0 \end{cases} \quad (5)$$

$$Y_{I\ell}^m = \begin{cases} \sqrt{2}K_\ell^m sin(|m\phi|)P_\ell^m(cos\theta) & m \neq 0 \\ K_\ell^0 P_\ell^0(cos\theta) & m = 0 \end{cases} \quad (6)$$

where $\ell \in \mathbb{R}^+$, $-\ell \leq m \leq \ell$, P_ℓ^m is the associated Legendre polynomial and K_ℓ^m is a scaling factor to normalize the function and is given by:

$$K_\ell^m = \sqrt{\frac{(2\ell + 1)}{4\pi} \frac{(\ell - |m|)!}{(\ell + |m|)!}} \quad (7)$$

The projection of functions in spherical coordinates into real and imaginary spherical harmonics are coefficients $C_{R\ell}^m$ and $C_{I\ell}^m$ with degree ℓ and order m. These coefficients are simply computed by integrating the product of spherical function f and the spherical harmonics function as follows:

$$C_{R\ell}^m = \oint f(\theta, \phi) Y_{R\ell}^m(\theta, \phi) \, d\theta d\phi \quad (8)$$

$$C_{I\ell}^m = \oint f(\theta, \phi) Y_{I\ell}^m(\theta, \phi) \, d\theta d\phi \quad (9)$$

The discrete approximation of function f by using these coefficients $C_{R\ell}^m$ and $C_{I\ell}^m$ for n number of bands is written as follows:

$$\tilde{f}(\theta, \phi) = \sum_{\ell=0}^{n-1} \sum_{m=-\ell}^{\ell} \left((C_{R\ell}^m + jC_{I\ell}^m)(Y_{R\ell}^m(\theta, \phi) + jY_{I\ell}^m(\theta, \phi)) \right) \quad (10)$$

Where $j = \sqrt{-1}$, ℓ is a positive integer value in the range [0, $n - 1$], and $C_{R\ell}^m$, $C_{I\ell}^m$ are calculated in (8) and (9) respectively. An important property of spherical harmonics is the rotation invariance. According to such a rotational invariance property, the projection of function f is identical to the rotation of projected function $\tilde{f}$ (Kazhdan et al., 2003). Hence, any rotation of function f does not change its L_2-norm (Banerjee et al., 2012). Based on this key property, the rotation invariant descriptor is proposed by computing the amount of energy contained within each spherical harmonic frequency, which are constant over any random rotation of the spherical function.

3.3. Construction of 3D-GMRF Rotation Invariant Descriptor

The 3D-GMRF model parameters α_v at voxel v measure the influence of voxels y_{v+r}, on the surface of a spherical neighborhood, in the central voxel g_v (see Fig. 1). As previously stated, the rotation of neighbors y_{v+r} results in the changes of parameters α_r. To remove the effect of rotation, a function f is defined at voxel location v to represent the estimated parameters α_v (see equation 2) that are sampled on the neighborhood sphere. Let $(\theta, \phi) \in \Omega_s$ where Ω_s represents the uniformly distributed points on the surface of a sphere and let the function f represents the GMRF model parameters α_v. The 3D-GMRF rotation invariant descriptor is therefore defined by:
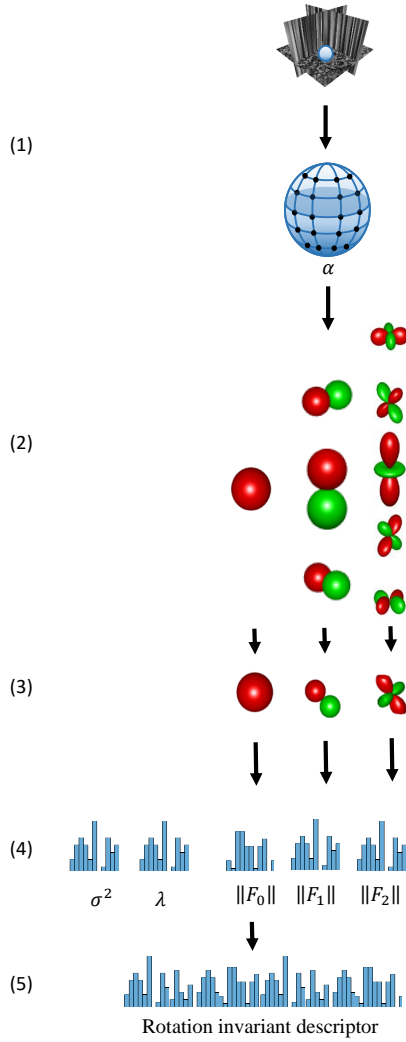
$$GMRF_{P,R}^{3Dri} = \{H(\|F_0\|), H(\|F_1\|) \dots, H(\|F_{n-1}\|), H(\sigma^2), H(\lambda)\} \quad (11)$$

Where P is the number of points sampled on the surface of neighborhood sphere with radius R, $H(.)$ is the histogram of the descriptor's elements over the entire volume and $\|F_\ell\|$ is the L_2-norm of the frequency component (Kazhdan et al., 2003) of the function f computed by:

$$\|F_\ell\| = \sqrt{\sum_{\theta \in \Omega_s} \sum_{\phi \in \Omega_s} |F_\ell(\theta, \phi)|^2} \quad (12)$$
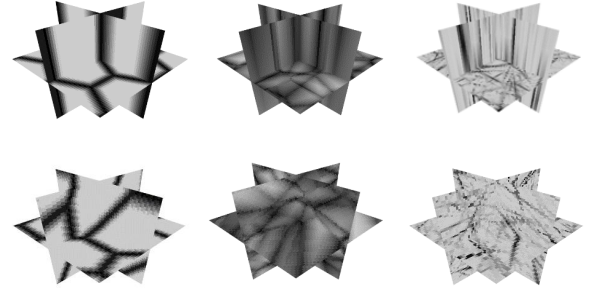
where $F_\ell(\theta, \phi)$ is calculated as:

Fig. 5. Example of volumetric textures from the solid texture database synthesized by the interpolation method. The top row represents a normal subset and samples in the bottom row are rotated versions of the normal subset.

variant descriptor representing the entire volumetric image at any random rotation. In the example illustrated in Fig. 3, two volumetric texture patches with different orientation produce identical descriptors. The construction process of our proposed method is illustrated with more details in Fig. 4.

4. Volumetric Texture Classification

4.1. Dataset

There are few synthetic volumetric texture datasets available for methods evaluation. The RFAI database is one of the available databases for synthetic 3D texture images (Paulhac et al., 2009), which is created to evaluate texture classification and segmentation methods. This database is constructed using two-dimensional dataset such as Brodatz and fractal textures. Four different methods are employed to generate 3D textures in this database. The first method is by interpolating two or more 2D texture images to construct a volumetric texture (See e.g. Fig. 5). In the second method geometric shapes such as spheres and cubes, are used to generate volumetric textures (See e.g. Fig. 6). Fourier transformation is employed to synthesize volumetric textures in the third method (See e.g. Fig. 7), whereas the fourth method is a mixture of the three previous methods to generate volumetric textures. Table 1 lists the number of classes contained within each category. These methods produce four different categories in which each category contains five types of textures according to the type of distortion: normal, rotated, Gaussian blur, Gaussian noise, and subsampling. Each of these types contain several classes and each class is composed of 10 volumetric images with a size of $64 \times 64 \times 64$. The database also offers a dataset for segmentation problems. This diversity of synthetic textures provides an appropriate way to evaluate texture analysis methods under different conditions. All datasets with their subsets are considered to evaluate our proposed rotation invariant descriptor. A simple program is developed to easily construct the volumetric texture with size $64 \times 64 \times 64$.

Fig. 4. Construction of the rotation invariant descriptor $GMRF_{P,R}^{3Dri}$. **(1) For each voxel in the volume image, discrete points are parametrized on the sphere and parameters α_r are estimated to represent a function in the spherical coordinate system. (2) The function is decomposed into its harmonic frequencies. (3) Within each frequency, the harmonics are accumulated, and a L_2-norm is computed for each frequency component to result in a set of variables for each voxel. (4) A histogram is computed for each variable, including variance σ^2 and mean λ. (5) Histograms are concatenated to construct a rotation invariant descriptor for the whole volume.**

$$F_\ell(\theta, \phi) = \sum_{m=-\ell}^{\ell} \left(\left(C_{R\ell}^m + jC_{I\ell}^m \right) \left(Y_{R\ell}(\theta, \phi) + jY_{I\ell}^m(\theta, \phi) \right) \right) \quad , (\theta, \phi) \in \Omega_s$$

(13)

Each voxel in the volumetric image is composed of a set of elements representing the L_2-norm of frequency components $\|F_\ell\|$, in addition to other model parameters σ_v^2 and λ_v since variance and mean are not affected by texture rotation. A histogram is computed for each of these elements to produce a rotation in-

Table 2. Classification accuracies [%] of our method and some other methods using RFAI datasets.

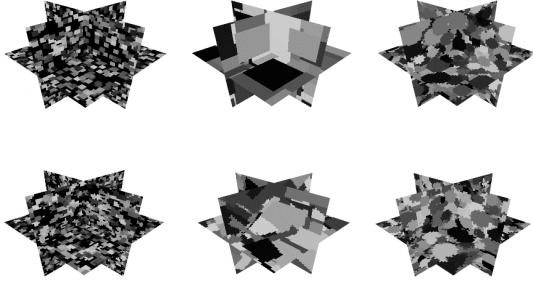| | | Synthetic Texture Dataset | | | | | | | | | | | | | | |
	Descriptor	Fourier				Geometric				Interpolated				Mixed texture			
		rotate	noise	smooth	subsampling	rotate	noise	smooth	subsampling	rotate	noise	smooth	subsampling	rotate	noise	smooth	subsampling
A	$GMRF_{42,1}^{3Dri}$	**97.3**	100	72.0	**47.0**	**98.0**	100	82.8	28.4	99.6	93.0	78.59	49.16	**98.8**	99.6	92.0	26.8
B	$GMRF_{42,1}^{3D}$	59.33	-	-	-	70.4	-	-	-	43.81	-	-	-	70.0	-	-	-
C	3D Riesz	97.0	100	**87.0**	45.0	86.0	96.0	42.0	21.0	-	-	-	-	-	-	-	-
D	$NI/RD/CI - LBP^{riu3g}$	-	-	-	-	-	-	-	-	95.3	-	-	-	-	-	-	-

Fig. 6. Example of volumetric textures synthesized by the geometrical method. The top row represents a normal dataset and samples in the bottom row are rotated versions of the normal dataset.
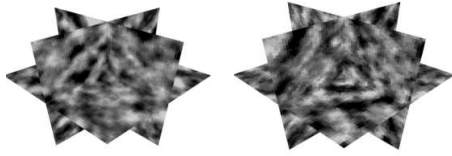

Fig. 7. Example of volumetric textures synthesized by the Fourier method. The sample of the left represents a normal dataset and the sample on the right is the rotated version of the normal dataset.

4.2. Metrics and Parameter Selection

The rotation invariant descriptor proposed here are examined by using different datasets, metrics and parameter settings. The similarity between descriptors is estimated by using L_1-norm as a distance metric and the k-nearest-neighbors (kNN) with ($k=1$) is exploited to classify the textures. The leave-one-out strategy is used to measure classification accuracy. The best number of histogram bins is selected to produce the highest rate of accuracy. However, the number of bins should not be selected in such a way that it leads to an increase in the descriptor size. Various numbers of points P sampled on the surface of the neighborhood sphere are selected, $P = \{32, 42, 60, 100, 200\}$ to assess the performance of the descriptor, while the radius R of the sphere is assigned to $\{1, 2\}$. The normal (unrotated textures) subset of each dataset is used as a training set, and each of the other subsets including the randomly rotated textures subset is used as a test set.
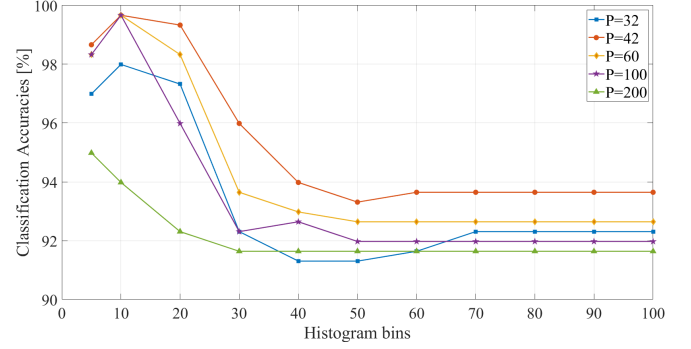
Fig. 8. Classification accuracies [%] achieved for various sampling points (P) and $R = 1$ by employing different numbers of histogram bins.

4.3. Results and Discussions

In this section, our proposed descriptor $GMRF_{P,R}^{3Dri}$ is evaluated and compared with other methods using RFAI database. The performance of our method using different settings is also evaluated using a selected dataset from RFAI.

4.3.1. Performance Evaluation and Comparison to Other Methods

The aim of the first experiment here is to ensure that the $GMRF_{P,R}^{3Dri}$ is truly rotation invariant. Our proposed method here along with the rotation variant method $GMRF_{P,R}^{3D}$ proposed in (Almakady et al., 2018) are both evaluated on the rotated subset of the four datasets. For each dataset, we train the normal subset and test the methods on the rotated subset. Since the rotation variant descriptor $GMRF_{P,R}^{3D}$ is constructed by concatenating the histograms of GMRF parameters, any texture rotation leads to the change of parameters order and therefore results in unsatisfactory classification performance. The correct classification accuracies of both methods presented in Table 2 (A-B) show that our proposed rotation invariant descriptor successfully classifies the rotated set while the other method poorly performs when the rotated subset is used. Further investigation is carried out by employing more rotation angles to examine the rotation invariance property of the descriptor proposed here. The classification performance of $GMRF_{P,R}^{3Dri}$ is evaluated under various rotation angles and is compared with the performance of the rotation variant method $GMRF_{P,R}^{3D}$. The normal (unrotated) subset from the volumetric texture dataset is selected as a training set, then the two methods are tested on the textures rotated with angles ranging from 0° to 340° around axes x, y and z. The results of this experiment displayed in Table 3 also indicate that despite the fact that our rotation invariant descriptor is

Table 3. Classification accuracies achieved with various rotation angles

		Classification accuracy [%] for each angle																	
Method	Axis	0°	20°	40°	60°	80°	100°	120°	140°	160°	180°	200°	220°	240°	260°	280°	300°	320°	340°
$GMRF^{3Dri}_{42,1}$	x	100	99.6	96.4	94.4	94.0	89.6	90.4	88.8	87.2	94.4	88.0	86.4	90.0	94.4	95.2	94.8	95.6	99.2
	y	100	99.2	96.4	97.6	96.0	93.6	90.8	88.8	87.6	96.0	90.4	90.0	93.2	96.0	96.8	97.6	98.4	99.2
	z	100	97.6	95.2	94.8	96.0	95.2	92.0	89.2	88.0	92.8	88.4	86.0	87.6	89.2	94.0	95.2	95.2	98.0
$GMRF^{3D}_{42,1}$	x	100	69.6	48.4	42.8	45.6	46.4	40.8	39.2	43.2	67.2	48.0	34.8	41.6	48.0	48.0	47.2	49.6	72.4
	y	100	76.0	59.6	59.6	62.8	56.0	46.8	47.6	51.2	62.8	51.2	48.4	52.8	60.8	59.2	55.6	61.2	76.4
	z	100	65.6	51.2	44.4	44.4	50.8	44.4	38.4	45.6	63.2	40.4	36.8	36.0	43.6	45.2	44.8	47.2	63.2

only trained on the normal (unrotated) samples; it successfully classifies the samples under various rotations. The rotation invariant descriptor proposed here achieves higher performance than the rotation variant method through all rotation angles, except for the 0 where no rotation has been applied. The rotation variant method $GMRF^{3D}_{P,R}$, on the other hand, performs lower in classifying the rotated samples as the descriptors constructed by this method are not rotation invariant and therefore the classification is highly affected by the rotation. The aforementioned results indicate that $GMRF^{3Dri}_{P,R}$ proposed here is invariant to rotation.

To demonstrate the classification performance of $GMRF^{3Dri}_{P,R}$ descriptor, a comparison with two recent texture-based methods is made by using the same synthetic dataset. The first method is an LBP-based method proposed in (Citraro et al., 2017) to combine three descriptors $NI/RD/CI-LBP^{riu3g}$ extracted from 3D volumetric textures. These are intensity, radial difference, and contrast descriptors respectively. This method is evaluated on the rotated subset of the interpolated dataset in the RFAI database consisting of 30 classes, each comprising 10 volumetric images (Paulhac et al., 2009), (see Table 1 for further description regarding this synthetic dataset). The second method proposed in (Cid et al., 2017) based on the 3D Riesz wavelet, consists of three approaches for local image orientation estimation, based on multidimensional Riesz, uni-directional Riesz, and first-order Riesz components. Our method is evaluated using all datasets of the RFAI database and compared with the reported results of the each method involved in this comparison. All subsets, normal, rotate, noise, smooth and subsampling are considered in this experiment in which the normal subset of each dataset is used as a training set, and the testing is performed on the remaining subsets of the corresponding dataset.

The classification accuracies presented in Table 2 (A,C,D) demonstrate the excellent performance of our proposed descriptor $GMRF^{3Dri}_{P,R}$, outperforming the other two descriptors. For Geometric dataset in particular, our proposed descriptor outperforms the 3D Riesz wavelet-based method (Cid et al., 2017) with a significant margin. The descriptor proposed here also comparatively performs well with noise, smooth and subsampling subsets data. Comparing with Riesz wavelet-based method, our proposed descriptor is favorably competitive for Fourier dataset and performs better for Geometric dataset.

The proposed method here is further evaluated using different sampling rates and histogram bins values. The classification accuracies presented in Table 4 demonstrate the classification performance of $GMRF^{3Dri}_{P,R}$ at different sampling rates using the interpolated dataset of the RFAI database. It is observed that our descriptor achieves the best results with sampling rate

Table 4. Classification accuracies [%] for various sampling rates (P), R = 1 and two histogram bins values using interpolated dataset of the RFAI database.

P	Classification Accuracies [%]	
	Histogram bins =5	Histogram bins =10
32	96.99	97.99
42	98.66	99.66
60	98.32	99.66
100	98.32	99.66
200	94.98	93.97

$P = \{42, 60, 100\}$ with 10 number of histogram bins. With too few or too many sampling points such as $P = \{32, 200\}$ the performance is slightly lowered. This is because a high sampling rate on a sphere with radius $R = 1$ tends to capture noise, while a low sampling rate leads to the ignorance of fine details and therefore reduces the discriminative power of the descriptor. With only $P = \{42\}$, the descriptor can capture significant details that are enough to characterize the local textures with less sensitivity to noise. With $P = \{60\}$ sampling points, our descriptor also achieves a satisfactory result; however, it requires more computation time when compared with $P = \{42\}$.

Fig. 8 depicts the classification accuracies obtained by setting a different number of histogram bins using the interpolated dataset. It is evident that the accuracy is improved when using a histogram with 10 bins but starts gradually dropping until it reaches a histogram with 40 bins. The accuracy then remains steady at low values with very slight changes. Therefore, the best appropriate value of histogram bins is 10 leading to a high classification performance with a low-dimensional descriptor.

The scales can be controlled by altering the value of radius R, as higher values of R can capture larger structures. Classification accuracies displayed in Fig.9 indicate that $R = \{1\}$ performs better than $R = \{2\}$ on the Interpolated dataset except for higher sampling rate $P = \{200\}$. As previously explained, this high number of sampling points tends to capture noise in $R = \{1\}$. The relatively low classification accuracy associated with $R = \{2\}$ compared with the accuracy for $R = \{1\}$ is potentially due to the importance of adjacent neighbors of the central voxel in texture structures. This is because for larger radii, for instance $R = \{2\}$, the adjacent neighbors of the central voxel are ignored and faraway neighbors, which contribute less, are considered. Nevertheless, higher scales i.e. larger R, can characterize macro structures or dominant patterns that are located at a large scale (Ojala et al., 2002). The size of the descriptor is based on the number of frequency components resulting from the accumulation of the spherical harmonics in addition to the variance σ^2 and mean λ. In other words, given the number of
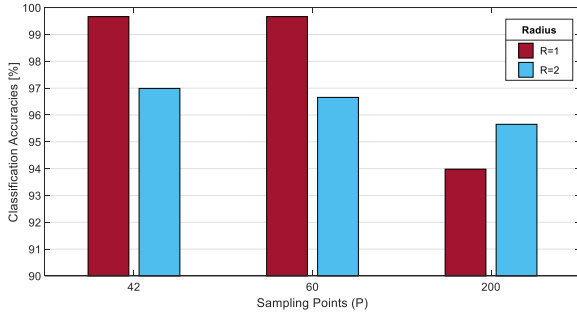
Fig. 9. Classification accuracies [%] of different scales $R = \{1, 2\}$ for different sampling rate $P = \{42, 60, 200\}$ when the number of histogram bins is set to 10.

Table 5. Comparison of our method with other handcrafted feature based methods on the COPD dataset for various values of R.

	Method	Classification accuracy [%]		
		R=1	R=2	R=3
A	$GMRF_{42,1}^{3Dri}$	**90.63**	81.25	84.37
B	$GMRF_{42,1}^{3D}$	81.25	75.00	78.12
C	LPH	75.0	56.25	75
D	LBP	78.12	71.88	75
E	3D GLCM	75.00	-	-
F	Gabor filters	78.12	-	-
G	SIFT	75	-	-
H	Intensity features method	68.75	-	-
I	Density-based method	71.88	-	-

spherical harmonic bands n, the size of descriptor is calculated by:

$$|GMRF_{P,R}^{3Dri}| = (n + 2) * b \qquad (14)$$

where b is the number of histogram bins, the additional terms refer to the variance σ^2 and mean λ. It is obvious that the number of sampling points P does not affect the size of the descriptor. This is because the parameters α_v of 3D-GMRF model are decomposed into a fixed number of frequency components of spherical harmonics proportional to the number of bands n. This overcomes the problem of the growth of sampling points, resulting in a reduction in descriptor sizes and therefore less computational time is required for matching histograms.

The classification performance of the descriptor is reported here through the Cumulative Match Characteristic (CMC) curve. The CMC represents the classification accuracies achieved within the Rank-k where k is the kth nearest neighbor. As can be seen from Fig.10, the perfect classification among all datasets can always be achieved within Rank-5 and in particular within Rank-2 for Fourier dataset.

5. Application to Medical Image Classification

5.1. Application to COPD Detection

In this experiment, the descriptor proposed here is employed for the detection of chronic obstructive pulmonary disease (COPD). Clinically, COPD refers to a group of progressive lung diseases defined as a common, preventable, and treatable disease characterized by persistent respiratory symptoms and airflow limitation that is due to airway and/or alveolar abnormalities (Vogelmeier et al., 2017). This disease is characterized by pathological changes affecting the texture of a healthy lung in the CT scan images of lungs. The changes in lung textures can be detected to discriminate COPD patients from healthy individuals by classifying the features extracted from lung images.

A clinical dataset of full-lung HRCT volumetric images composed of 32 subjects is exploited in this experiment. These are 19 healthy subjects and 13 subjects diagnosed with COPD. The rotation invariant descriptor is constructed for each subject by applying our method on extracted randomly rotated volumes of interest (VOIs) from the HRCT images of lungs (See e.g.

Fig. 11). We conducted three experiments investigating our proposed method performance on the COPD dataset. The first experiment given in Section 5.1.1 presents the comparison with the texture based methods. Section 5.1.2 compares our method with the intensity based methods. The comparison with deep learning based methods is finally presented in Section 5.1.3.

5.1.1. Comparison to texture based methods

This experiment aims at comparing our method with texture feature based methods using the COPD dataset. Five different texture-based methods are considered in this experiment in addition to our previously proposed method $GMRF_{P,R}^{3D}$ to examine the rotation invariance property. These methods are; the local parameter histogram (LPH) method based on 2D-GMRF (Dharmagunawardhana et al., 2016), LBP (Ojala et al., 2002), three-dimensional gray-level co-occurrence matrix (3D GLCM) (Haralick et al., 1973; Han et al., 2015), features extracted by a bank of Gabor filters (Manjunath and Ma, 1996) and scale invariant feature transform (SIFT) (Lowe, 2004) . In 3D GLCM, a set of texture features is derived for each VOI from the 3D GLCM by using 13 angular directions and 4 distances. For LPH and LBP, we set $R = 1$ and $P = 8$, while Gabor filters are computed using six orientations and four scales as reported in (Manjunath and Ma, 1996). SIFT is a rotation and scale invariant method for local feature extraction, which has been used for various tasks including texture classification (Yang and Newsam, 2008; Xu et al., 2012). In this experiment, a number of keypoints are detected for each slice in VOIs, and then the 128 dimensional descriptors are constructed for these keypoints. To perform the classification task, a normalized global descriptor is obtained by computing the distribution of the local descriptors after being labelled using the k-mean clustering algorithm (Yang and Newsam, 2008). The above methods are applied to each slice in the same extracted VOIs, and then the best result achieved among all slices in VOIs is selected to ensure a fair comparison. The classification accuracies presented in Table 5 (A-G) demonstrate the discriminatory power of the descriptor proposed here to distinguish between healthy individuals and COPD patients compared with all other methods. Moreover, the outcomes of the comparison between the GMRF-based methods (A and C in Table 5), in addition to the performance of Gabor filters and SIFT, support the assumption that information contained in 3D HRCT images is better
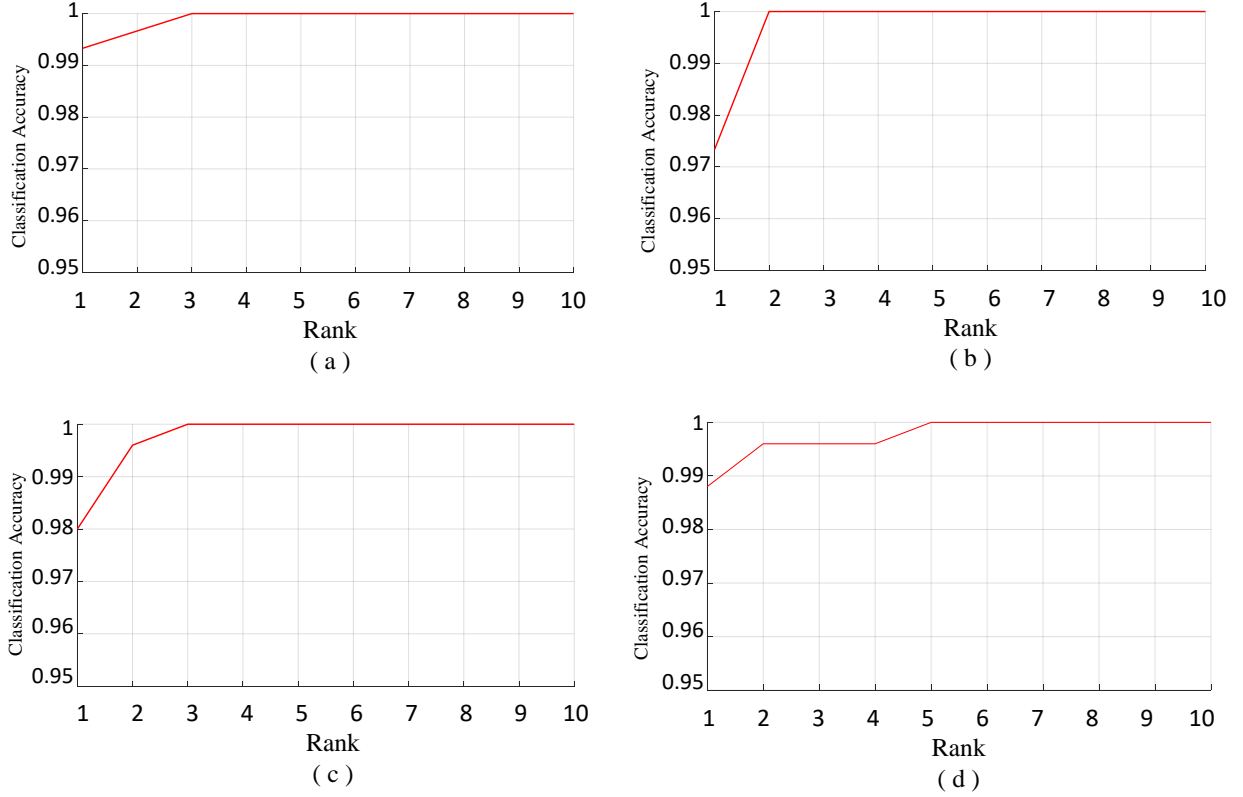
Fig. 10. CMC curves demonstrating the classification performance of our proposed descriptor on different datasets : (a) Interpolated dataset (b) Geometric dataset, (c) Fourier dataset and (d) Mixed texture dataset.

described by 3D descriptors, therefore, when extracting texture features from 3D images, the 3D model can better describe a 3D texture. In contrast, the 2D based methods process each slice in the volume separately which would lead to the negligence of valuable information which could be found by including the third dimension. Moreover, since the method that shown in Table 5 (B) is variant to the rotation, it is not able to capture different rotated textures and then performs less than the method presented in Table 5 (A) on the rotated VOIs. Regarding the Gabor filters-based method shown in Table 5 (F), besides the low classification performance, a disadvantage associated with the method of Gabor filter bank, is that it requires filtering the bank of filters with textures rotated under a maximum number of angles to extract features for different possible rotated patterns. Such a process of filtering would be numerically expressive and time consuming and also interactable for all possible rotation angles. In contrast, our method extracts the rotated features using the rotation invariant property. As can be seen from Table 5 (A-D), increasing the neighborhood size (Radius R) is not always beneficial because this incorporates irrelevant voxels that could have less influence on the central voxel. Larger neighborhood sizes also could lead to the negligence of micro textures which could be found in a small neighborhood size.

5.1.2. Comparison to intensity based methods

The importance of texture information extracted from CT images has already been proved to be effective in COPD de-

tection (Park et al., 2008; Uppaluri et al., 1997, 1999; Sluimer et al., 2003; Sorensen et al., 2010). Additional experiments are conducted here to measure the performance of our method when texture features are excluded in COPD images analysis. The aim of this experiment is to demonstrate the discriminative power of texture features for COPD diagnosis by analysing 3D lung HRCT scans. The same VOIs extracted previously from the clinical dataset are used in the following experiments.

A. Comparison with features excluding texture information

In this experiment, our proposed method is tested by excluding the texture features α_v and σ_v^2 and keeping only the intensity feature represented by λ_v . This results in the feature vector $f_v = \{\lambda_v\}$ which leads to the final descriptor i.e. $GMRF_{P,R}^{3Dri} = H(\lambda)$. This descriptor is constructed for each of the VOIs and then the same settings and process described in Section 4.2 are followed for $R = \{1\}$. The classification accuracy obtained by using only intensity features is 68.75% as presented in Table 5 (H). This accuracy is much less than the accuracy with full texture model 90.63% presented in Table 5 (A). This clearly demonstrates the importance of the texture features forming our descriptor to distinguish between healthy individuals and COPD patients.

B. Comparison with density-based method

In this experiment, our method is compared with a density-based method on COPD detection to emphasize the importance of texture compared with intensity. The presence of COPD (i.e. emphysema) can be detected by quantifying the low attenuation area (LAA) in CT images of lungs attributable to the abnormalities and destruction of the lung tissues (Muer et al., 1988; Mascalchi et al., 2017) . Computing the histogram of LAA in the CT images of the lung can lead to the detection of the disease (Sorensen et al., 2010). There are different measurements derived from the histogram of LAA such as mean density, relative area (RA) and the percentile of LAA distribution (Mascalchi et al., 2017) . We select the percentile of LAA distribution as a measurement derived from the histogram of LAA to compute the lung density based only on the intensity values of the VOIs. The percentile of LAA distribution is defined as the distribution of all values in Hounsfield units (HU) under a specific threshold (Mascalchi et al., 2017). The 15th percentile has been used practically as an optimal threshold (Stoel and Stolk, 2004; Lynch, 2014), therefore it is chosen in this experiment where a threshold of -950 HU is selected for density quantification (Lynch, 2014). The classification is performed according to the percentage of voxels under -950 HU in the VOIs. If the 15th percentile of a given VOI is -950 HU or less, then it indicates the presence of emphysema, otherwise the subject is healthy. The classification accuracy of this experiment is 71.88% as presented in Table 5 (I). Such a result supports the outcomes of previous research indicating the advantages of texture analysis over the intensity analysis for COPD detection.

5.1.3. Comparison with deep-learning-based methods

Deep learning approaches such as CNN have demonstrated a magnificent performance in many fields of machine learning, including image classification. However, the size of training data plays a vital role in the performance of the CNN models. Therefore, in this experiment we aim to compare the performance of our method with CNN based methods when the size of the training data is small . Such a scenario is prevalent in medical domain where only a small size of training data is available for many reasons discussed earlier. We employ two pre-trained CNN models; Residual Network architecture with 50 layers (ResNet-50) (He et al., 2016) and AlexNet (Krizhevsky et al., 2012) trained on the ImageNet database and have rich feature representations for various image classes. We adapt those networks to our classification problem by replacing the last fully connected layer and the final classification layer, with new layers to have a two classe output to match the number of classes in the COPD dataset. The size of the images is also adjusted to fit into the input layer of the networks. The 2D slices are extracted from each VOIs across the third dimension and then randomly split into training and validation data sets where 90% of the data is used for training. During the training, images are randomly augmented to increase the size of the training set. The adjusted networks are then fine-tuned for 20 epochs with a learning rate set to 0.0001 and employed for COPD classification. The results presented in Table 6 are the mean and standard deviation of classification accuracies of each network running ten times independently. It can be noticed that our method performs better in comparison to the pre-trained ResNet-50 and Alex-Net.

Table 6. Comparison of our method with deep learning-based methods on the COPD dataset.

	Method	Classification accuracy [%]
1	$GMRF_{42,1}^{3Dri}$	**90.63**
2	ResNet-50	87.50±0.09
3	AlexNet	85.94±.05
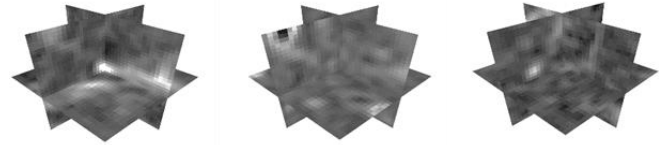
Fig. 11. VOIs sample taken from HRCT images of lungs in the clinical dataset.

However, it is important to mention that the relativly low performance of the deep learning methods is most likely due to the insufficient dataset involved in this expreiment. Nevertheless, our proposed method here is not affected by the small size of the dataset. These results suggest that our proposed method can be used in medical applications where the datasets usually have small sizes.

5.2. Application to Lung Nodule Classification

The dataset used in this experiment is the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI), which is publicly accessible for medical imaging research (Armato III et al., 2011). The database contains 1018 clinical chest CT scans, each of which is associated with an XML file that involves detailed information about each case, such as nodules locations and annotations. Each nodule in the (LIDC-IDRI) database is rated by a 5-point scale of its level of malignancy by up to four experienced thoracic radiologists. We extract the nodules with their associated centres according to the nodule collection report (A. P. Reeves, 2011). We compute the median malignancy level (MML) of each nodule to annotate its level of malignancy. Each nodule with MML> 3 is labelled as malignant whereas a nodule with MML< 3 is labelled as benign. A nodule with MML=3 is regarded as uncertain and is excluded from this experiment. Therefore, after removing some nodules with ambiguous information and nodules with missing slices, we end up with 884 benign and 514 malignant nodules included in this experiment. A cube with a size of $64 \times 64 \times 64$ centred at each extracted nodule is considered to capture the majority of the nodule, and then we place an adaptive sphere inside the cube to include only the nodule and exclude the surrounding tissue based on the information provided by the report (A. P. Reeves, 2011), and then we apply our method on extracted patches of the nodules. Figure 12 depicts some examples of benign (left) and malignant (right) lung nodules in chest CT scans visualized in the 3D space. Results are reported as the mean, standard deviation, sensitivity/recall and specificity of 10-fold cross-validation. Our proposed method is compared with two CNNs based methods (Shen et al., 2017) and (Hussein et al., 2017) . Since their datasets are different in the number of

Table 7. Comparison of our method with deep learning-based methods on the (LIDC-IDRI) database.

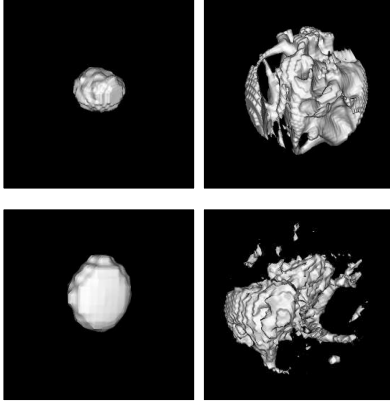| No.	Method	Number of Nodules		Results [%]		
		B	M	Accuracy	Sensitivity/Recall	Specificity
A	$GMRF^{3Dri}_{42,1}$	884	514	91.45±0.10	83.48±0.02	**96.07±0.01**
B	$GMRF^{3Dri}_{42,1}$	528	297	**92.63±1.51**	87.27±0.03	95.66±0.01
C	Multi-crop CNN	528	297	87.14	77.00	93.00
D	$GMRF^{3Dri}_{42,1}$	635	509	91.57±1.43	**87.86±0.03**	94.55±0.01
E	3D CNN	635	509	91.26	-	-

Fig. 12. 3D visualization of benign (left) and malignant (right) lung nodules in chest CT scans.

nodules, our dataset is adjusted to match the same number of nodules reported in those methods for a fair comparison with methods. The reduction of the dataset is performed through a random selection of nodules from the dataset repeated ten times to ensure an unbiased selection of data. Table 7 presents the classification performance of the methods involved in this experiment. In row A of this table, our proposed method uses all extracted nodules (884 benign and 514 malignant) while a part of the dataset is employed in rows B and D of Table 7. It can be seen from Table 7 that our proposed method performs well in classifying the lung nodules in comparison to the other methods.

6. Conclusion

In this paper, a 3D-GMRF-based rotation invariant method is proposed for volumetric texture classification by exploiting the rotation invariance property of the spherical harmonics. The function that defines the estimated parameters α_r on the surface of the neighborhood sphere is decomposed into its spherical harmonics with multiple bands. The harmonics within each band are then accumulated and an L_2-norm is computed for frequency components. A histogram of each component is then computed over the volume in addition to the histograms of σ^2 and λ. These histograms are concatenated to construct the rotation invariant descriptor. The classification performance of the descriptor is examined by classifying the rotated subset of

various datasets in the RFAI synthetic database for volumetric textures. Our descriptor demonstrates good classification performance by achieving high classification accuracy with various sampling rates and a reduced number of histogram bins. To demonstrate the classification performance of our descriptor, we present various experiments to compare our method with the state-of-the-art methods on synthetic and medical datasets of COPD and LIDC-IDRI. The results presented here demonstrate the excellent classification performance achieved by our proposed descriptor in comparison with the methods investigated here. Our descriptor also demonstrates good performance in comparison with a 2D-GMRF based method to discriminate COPD patients from healthy individuals in a clinical dataset to indicate the importance of extracting features from the structure of 3D volumetric data. To examine the significance of the texture features, additional experiments are discussed here by considering only the intensity features. The results show the superior performance of the texture features for COPD detection.

References

A. P. Reeves, A. M. Biancardi, s., 2011. The lung image database consortium (lidc) nodule size report.

Almakady, Y., Mahmoodi, S., Conway, J., Bennett, M., 2018. Volumetric Texture Analysis Based on Three-Dimensional Gaussian Markov Random Fields for COPD Detection, in: Annual Conference on Medical Image Understanding and Analysis (MIUA), Springer. pp. 153–164.

Almakady, Y., Mahmoodi, S.M., Bennett, M., 2019. Gaussian markov random fields-based features for volumetric texture segmentation, in: 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), IEEE. pp. 212–215.

Andrearczyk, V., Depeursinge, A., 2018. Rotational 3d texture classification using group equivariant cnns. arXiv preprint arXiv:1810.06889 .

Armato III, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A., et al., 2011. The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. Medical physics 38, 915–931.

Banerjee, J., Moelker, A., Niessen, W.J., van Walsum, T., 2012. 3D LBP-based rotationally invariant region description, in: Asian Conference on Computer Vision, Springer. pp. 26–37.

Bjorkstrom, A., 2001. Ridge regression and inverse problems. Stockholm University, Department of Mathematics .

Chellappa, R., Chatterjee, S., 1985. Classification of textures using Gaussian Markov random fields. IEEE Transactions on Acoustics, Speech, and Signal Processing 33, 959–963.

Cid, Y.D., Mller, H., Platon, A., Poletti, P.A., Depeursinge, A., 2017. 3D solid texture classification using locally-oriented wavelet transforms. IEEE Transactions on Image Processing 26, 1899–1910.

Citraro, L., Mahmoodi, S., Darekar, A., Vollmer, B., 2017. Extended three-dimensional rotation invariant local binary patterns. Image and Vision Computing 62, 8–18.

Depeursinge, A., Fageot, J., Andrearczyk, V., Ward, J.P., Unser, M., 2018. Rotation invariance and directional sensitivity: Spherical harmonics versus radiomics features, in: International Workshop on Machine Learning in Medical Imaging, Springer. pp. 107–115.

Deserno, M., 2004. How to generate equidistributed points on the surface of a sphere. P.-If Polymerforshung (Ed.) .

Dharmagunawardhana, C., Mahmoodi, S., Bennett, M., Niranjan, M., 2014. Gaussian Markov random field based improved texture descriptor for image segmentation. Image and Vision Computing 32, 884–895.

Dharmagunawardhana, C., Mahmoodi, S., Bennett, M., Niranjan, M., 2016. Rotation invariant texture descriptors based on Gaussian Markov random fields for classification. Pattern Recognition Letters 69, 15–21.

El khadiri, I., Chahi, A., Ruichek, Y., Touahni, R., et al., 2018. Local directional ternary pattern: A new texture descriptor for texture classification. Computer vision and image understanding 169, 14–27.

Fehr, J., Burkhardt, H., 2008. 3D rotation invariant local binary patterns, in: 19th International Conference on Pattern Recognition . ICPR, IEEE. pp. 1–4.

Friedman, J., Hastie, T., Tibshirani, R., 2001. The elements of statistical learning. volume 1. Springer series in statistics New York.

Green, R., 2003. Spherical harmonic lighting: The gritty details, in: Archives of the Game Developers Conference, p. 4.

Griffiths, D., Boehm, J., 2019. A review on deep learning techniques for 3d sensed data classification. Remote Sensing 11, 1499.

Han, F., Wang, H., Zhang, G., Han, H., Song, B., Li, L., Moore, W., Lu, H., Zhao, H., Liang, Z., 2015. Texture feature analysis for computer-aided diagnosis on pulmonary nodules. Journal of digital imaging 28, 99–115.

Haralick, R.M., Shanmugam, K., Dinstein, I.H., 1973. Textural features for image classification. IEEE Transactions on systems, man, and cybernetics , 610–621.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Hussein, S., Cao, K., Song, Q., Bagci, U., 2017. Risk stratification of lung nodules using 3d cnn-based multi-task learning, in: International conference on information processing in medical imaging, Springer. pp. 249–260.

Kazhdan, M., Funkhouser, T., Rusinkiewicz, S., 2003. Rotation invariant spherical harmonic representation of 3D shape descriptors, in: Symposium on geometry processing, pp. 156–164.

Kovalev, V.A., Kruggel, F., Gertz, H.J., von Cramon, D.Y., 2001. Three-dimensional texture analysis of MRI brain datasets. IEEE transactions on medical imaging 20, 424–433.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, pp. 1097–1105.

Kumar, A., Kim, J., Lyndon, D., Fulham, M., Feng, D., 2016. An ensemble of fine-tuned convolutional neural networks for medical image classification. IEEE journal of biomedical and health informatics 21, 31–40.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. Medical image analysis 42, 60–88.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints.

International journal of computer vision 60, 91–110.

Lynch, D.A., 2014. Progress in imaging COPD, 2004-2014. Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation 1, 73.

Mahmoodi, S., Gunn, S., 2011. Snake based unsupervised texture segmentation using Gaussian Markov random field models, in: Image Processing (ICIP), 2011 18th IEEE International Conference on, IEEE. pp. 3353–3356.

Manjunath, B., Chellappa, R., 1991. Unsupervised texture segmentation using Markov random field models. IEEE Transactions on Pattern Analysis and Machine Intelligence , 478–482.

Manjunath, B.S., Ma, W.Y., 1996. Texture features for browsing and retrieval of image data. IEEE Transactions on pattern analysis and machine intelligence 18, 837–842.

Mascalchi, M., Camiciottoli, G., Diciotti, S., 2017. Lung densitometry: why, how and when. Journal of thoracic disease 9, 3319.

Muer, N.L., Staples, C.A., Miller, R.R., Abboud, R.T., 1988. density mask: an objective method to quantitate emphysema using computed tomography. Chest 94, 782–787.

Ojala, T., Pietikainen, M., Maenpaa, T., 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on pattern analysis and machine intelligence 24, 971–987.

Park, Y.S., Seo, J.B., Kim, N., Chae, E.J., Oh, Y.M., Do Lee, S., Lee, Y., Kang, S.H., 2008. Texture-based quantification of pulmonary emphysema on high-resolution computed tomography: comparison with density-based quantification and correlation with pulmonary function test. Investigative radiology 43, 395–402.

Paulhac, L., Makris, P., Ramel, J.Y., 2008. Comparison between 2D and 3D local binary pattern methods for characterisation of three-dimensional textures, in: International Conference Image Analysis and Recognition, Springer. pp. 670–679.

Paulhac, L., Makris, P., Ramel, J.Y., 2009. A Solid Texture Database for Segmentation and Classification Experiments, in: VISAPP (2), pp. 135–141.

Petrou, M., Sevilla, P.G., 2006. Image processing: dealing with texture. volume 1. Wiley Online Library.

Shen, W., Zhou, M., Yang, F., Yu, D., Dong, D., Yang, C., Zang, Y., Tian, J., 2017. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. Pattern Recognition 61, 663–673.

Sluimer, I.C., van Waes, P.F., Viergever, M.A., van Ginneken, B., 2003. Computeraided diagnosis in high resolution CT of the lungs. Medical physics 30, 3081–3090.

Sorensen, L., Shaker, S.B., De Bruijne, M., 2010. Quantitative analysis of pulmonary emphysema using local binary patterns. IEEE transactions on medical imaging 29, 559–569.

Stoel, B.C., Stolk, J., 2004. Optimization and standardization of lung densitometry in the assessment of pulmonary emphysema. Investigative radiology 39, 681–688.

Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B., Liang, J., 2016. Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE transactions on medical imaging 35, 1299–1312.

Tuceryan, M., Jain, A.K., 1993. Texture analysis. World Scientific. pp. 235–276.

Uppaluri, R., Hoffman, E.A., Sonka, M., Hunninghake, G.W., McLennan, G., 1999. Interstitial lung disease: a quantitative study using the adaptive multiple feature method. American journal of respiratory and critical care medicine 159, 519–525.

Uppaluri, R., Mitsa, T., Sonka, M., Hoffman, E.A., McLennan, G., 1997. Quantification of pulmonary emphysema from lung computed tomography images. American journal of respiratory and critical care medicine 156, 248–254.

Vogelmeier, C.F., Criner, G.J., Martinez, F.J., Anzueto, A., Barnes, P.J., Bourbeau, J., Celli, B.R., Chen, R., Decramer, M., Fabbri, L.M., 2017. Global strategy for the diagnosis, management and prevention of chronic obstructive lung disease 2017 report. Respirology 22, 575–601.

Wang, C., Komodakis, N., Paragios, N., 2013. Markov random field modeling, inference & learning in computer vision & image understanding: A survey. Computer Vision and Image Understanding 117, 1610–1627.

Winkels, M., Cohen, T.S., 2018. 3d g-cnns for pulmonary nodule detection. arXiv preprint arXiv:1804.04656 .

Winkels, M., Cohen, T.S., 2019. Pulmonary nodule detection in ct scans with equivariant cnns. Medical image analysis 55, 15–26.

Worrall, D.E., Garbin, S.J., Turmukhambetov, D., Brostow, G.J., 2017. Harmonic networks: Deep translation and rotation equivariance, in: Proceed-

ings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5028–5037.

Xia, Y., Feng, D., Zhao, R., 2006. Adaptive segmentation of textured images by using the coupled Markov random field model. IEEE Transactions on Image Processing 15, 3559–3566.

Xie, Y., Xia, Y., Zhang, J., Song, Y., Feng, D., Fulham, M., Cai, W., 2018. Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest ct. IEEE transactions on medical imaging 38, 991–1004.

Xie, Y., Zhang, J., Xia, Y., 2019. Semi-supervised adversarial model for benign–malignant lung nodule classification on chest ct. Medical image analysis 57, 237–248.

Xu, Y., Huang, S., Ji, H., FermüLler, C., 2012. Scale-space texture description on sift-like textons. Computer Vision and Image Understanding 116, 999–1013.

Yan, X., Pang, J., Qi, H., Zhu, Y., Bai, C., Geng, X., Liu, M., Terzopoulos, D., Ding, X., 2016. Classification of lung nodule malignancy risk on computed tomography images using convolutional neural network: A comparison between 2d and 3d strategies, in: Asian Conference on Computer Vision, Springer. pp. 91–101.

Yang, Y., Newsam, S., 2008. Comparing sift descriptors and gabor texture features for classification of remote sensed imagery, in: 2008 15th IEEE international conference on image processing, IEEE. pp. 1852–1855.