**Article 25fa pilot End User Agreement**

This publication is distributed under the terms of Article 25fa of the Dutch Copyright Act (Auteurswet) with explicit consent by the author. Dutch law entitles the maker of a short scientific work funded either wholly or partially by Dutch public funds to make that work publicly available for no consideration following a reasonable period of time after the work was first published, provided that clear reference is made to the source of the first publication of the work.

This publication is distributed under The Association of Universities in the Netherlands (VSNU) ‘Article 25fa implementation’ pilot project. In this pilot research outputs of researchers employed by Dutch Universities that comply with the legal requirements of Article 25fa of the Dutch Copyright Act are distributed online and free of cost or other barriers in institutional repositories. Research outputs are distributed six months after their first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and/or copyrights owner(s) of this work. Any use of the publication other than authorised under this licence or copyright law is prohibited.

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please contact the Library through email: copyright@ubn.ru.nl, or send a letter to:

University Library  
Radboud University  
Copyright Information Point  
PO Box 9100  
6500 HA Nijmegen

You will be contacted as soon as possible.
Bag-of-Frequencies: A Descriptor of Pulmonary Nodules in Computed Tomography Images

Francesco Ciompi*, Colin Jacobs, Ernst Th. Scholten, Mathilde M. W. Wille, Pim A. de Jong, Mathias Prokop, and Bram van Ginneken

Abstract—We present a novel descriptor for the characterization of pulmonary nodules in computed tomography (CT) images. The descriptor encodes information on nodule morphology and has scale-invariant and rotation-invariant properties. Information on nodule morphology is captured by sampling intensity profiles along circular patterns on spherical surfaces centered on the nodule, in a multi-scale fashion. Each intensity profile is interpreted as a periodic signal, where the Fourier transform is applied, obtaining a spectrum. A library of spectra is created and labeled via unsupervised clustering, obtaining a Bag-of-Frequencies, which is used to assign each spectra a label. The descriptor is obtained as the histogram of labels along all the spheres. Additional contributions are a technique to estimate the nodule size, based on the sampling strategy, as well as a technique to choose the most informative plane to cut a 2-D view of the nodule in the 3-D image. We evaluate the descriptor on several nodule morphology classification problems, namely discrimination of nodules versus vascular structures and characterization of spiculation. We validate the descriptor on data from European screening trials NELSON and DLCST and we compare it with state-of-the-art approaches for 3-D shape description in medical imaging and computer vision, namely SPHARM and 3-D SIFT, outperforming them in all the considered experiments.

Index Terms—Chest computed tomography (CT), computer-aided detection, frequency analysis, nodule characterization, pulmonary nodules, three-dimensional (3-D) descriptor.

I. INTRODUCTION

Lung cancer is the most common cancer worldwide, accounting for 1.3 million of deaths annually. A mere 15% of all diagnosed lung cancers are detected at an early stage, resulting in a five-year survival rate of only 16% [1]. A recent study [2], showed that screening using low-dose computed tomography (CT) significantly reduces lung cancer and overall mortality. As a result, interest in lung cancer screening programs for the early detection of pulmonary nodules has substantially increased. In this context, the LungRADS guidelines [3] have been recently presented, with the aim of defining a standardized procedure for patient-tailored follow-up. In the guideline, the two elements necessary to establish the kind of management are the nodule type and the nodule growth-rate.

The appearance of a solid pulmonary nodule on an axial view of a CT scan is shown in Fig. 1(a). A characteristic for nodules in CT images is the appearance as bright tissues with a rounded-like 3-D shape. However, bright intensity is also characteristic of vascular structures such as vessels or bifurcations [Fig. 1(c) and (d)].

Fig. 1. Examples of intensity profiles sampled at three different distances \(r_1, r_2, r_3\) from a voxel of interest centered on a solid pulmonary nodule (a), a spiculated nodule (b) and vascular structures [(c), (d)]. In (e), the average and standard deviation of spatial frequencies spectra computed on the four profiles at distance \(r_1\) are depicted; in (f), frequencies of the intensity profiles sampled in (c) (top) and (d) (bottom) for \(r_2\) and \(r_3\) profiles are represented (best viewed in color).
Prior to characterization, lung nodule detection is required. Given the large amount of scans to read in a lung screening scenario, systems for computer-aided detection (CAD) have been proposed [5]–[8]. As a consequence, the use of CAD as both first and second reader has demonstrated to improve the sensitivity of nodule detection and the speed of scan reading. Nevertheless, CAD systems suffer from the presence of false positives (FPs), which are often identified in vascular structures such as vessels and bifurcations. In [7], it was shown that the main causes of FPs in the proposed CAD system for the detection of solid nodules are vascular structures. In this scenario, it becomes clear that the analysis of morphology is useful both to assess the malignancy of nodules and to reduce the number of false positives in nodule detection. We name this kind of tasks as nodule characterization problem.

In order to characterize a pulmonary nodule based on its morphology, methods relying on prior 3-D nodule segmentation have been proposed. In [9], segmented nodules are characterized as solid, part-solid and nonsolid through the analysis of their composition in a supervised learning framework. In [10], smoothness and irregularity of a nodule shape are characterized using image and surface features. When prior nodule segmentation is required, available methods for automatic or semi-automatic segmentation such as the ones presented in [11]–[13] may be used. Although highly accurate segmentations can be obtained for solid nodules, fully automatic procedures may fail in the presence of nonsolid or semi-solid nodules. For these reasons, it is desirable to make nodule characterization independent of accurate nodule segmentation as much as possible.

Approaches that do not rely on the shape of the segmented nodule have been also presented. In [14], unsupervised clustering is applied to a set of nodule patches, to create a dictionary used to label nodule voxels at testing time. The method showed high correlation with histology when characterizing the nodule as aggressive or indolent. In [15], nodule characterization is performed using the 2-D SIFT descriptor [16] and Linear Discriminant Analysis. The SIFT descriptor has appealing properties such as scale-invariance and rotation-invariance, which have made it popular in the computer vision community. In medical imaging, the 3-D formulation of SIFT (3-D SIFT) is a more suitable descriptor to characterize objects in 3-D images. Although 3-D SIFT has been used in quite some applications in medical imaging [17]–[20], it has been only applied to few problems using CT data, solely restricted to surveillance applications [21].

Given a voxel centered on a region of interest in a CT scan, information on local morphology in the lung may be obtained by analyzing the intensity of the image in a sufficiently large neighborhood. Examples are depicted in Fig. 1, where using three circular sampling patterns with radius $r_1 < r_2 < r_3$ around a voxel of interest, characteristic intensity profiles at the three scales are obtained in the presence of nodules [(a), (b)] or vascular structures [(c), (d)]. If we consider the distance $r_1$, the corresponding signal has high intensity and subtle variations along all the cases, while discriminative profiles with more variation in intensity are obtained at larger distances $(r_2, r_3)$. The interesting behavior of intensity profiles is reflected on the frequency domain. In Fig. 1(e), the mean spectrum computed using the four profiles at distance $r_1$ is depicted, along with the standard deviation per frequency. It is worth noting that relatively small variations are observed around each frequency, meaning that the behavior in regions close to the voxel of interest is similar regardless the kind of structure. In (f), spectra for spatial frequency $f > 0$ cycles/mm for $r_2$ and $r_3$ profiles in (c) and (d) are also depicted. In these cases, it is interesting to observe that spectra at low frequency $(1 < f < 5$ cycles/mm) have a similar behavior, even though they are computed at different distances from the centre. A larger variability is observed at higher frequency components, due to the presence of noise and ripple on the main signal.

In these examples, the information provided by each intensity profile is extracted from a 2-D domain, obtained by intersecting the 3-D scan with a plane. During their reading routine, radiologists are used to analyzing the nodule morphology by combining the view of orthogonal planes, obtaining axial, coronal and sagittal views. The message underlying this procedure is that a combination of 2-D views provides enough information to characterize nodule morphology.

1) Our Contribution: In this paper, we investigate the use of frequency-based features to describe nodule morphology in CT scans. Our hypothesis is that a multi-scale 3-D analysis of intensity profiles can represent the basis for extracting information to compactly encode nodule morphology. We introduce a novel descriptor for lung nodule characterization in chest CT, which we call Bag-of-Frequencies (BoF). The descriptor is based on the analysis of the image intensity in a 3-D spherical neighborhood of a voxel of interest, without the need for prior nodule segmentation. Based on this idea, we extend the use of axial, coronal and sagittal planes to a generic number of N planes, not limited to orthogonal directions. For each plane, circular intensity profiles are sampled at several scales. At each scale, the set of N circles approximate a spherical surface. Given the circularity of the sampling pattern, each profile can be interpreted as a periodic signal, containing characteristic spatial frequencies. The spectrum of each profile is computed via Fourier analysis and used as basic information for the descriptor. A dictionary of spectra is collected and labelled using unsupervised clustering. The descriptor is finally obtained as the histogram of labels of spectra in the sampled sphere. We formulate the descriptor in a scale-invariant fashion, based on the estimation of the nodule size. For this purpose, we also propose a method to derive the nodule size from the sampling pattern. Finally, we propose a technique to automatically compute planes to extract 2-D views from the 3-D scan which maximize the information on nodule morphology for visual inspection. To the best of our knowledge, this is the first time a similar approach is proposed to design a descriptor for nodule characterization in chest CT. The approach closest to ours is SPHARM [22], [23], which given a binary 3-D segmentation, approximates the object using 3-D spherical harmonics. We evaluate the application of the BoF descriptor on two problems: 1) the discrimination of nodules versus vessels and bifurcations and 2), the characterization of pulmonary nodule as spiculated. In the evaluation, we use data from the Dutch-Belgian Randomized Lung Cancer Screening Trial (NELSON) [24] and the Danish Lung Cancer Screening Trial (DLCST) [25].
Fig. 2. Schematic representation of the used sampling pattern. Set of circles is considered around a pulmonary nodule, obtained as the intersection between spheres centered on the nodule and planes (a). For each circle, points are sampled (b), to generate an intensity profile. The icosahedron geometry for a semi-sphere is depicted in (c), where the vector $\mathbf{b}$ is used to associate a circle with radius $r_F$ with the barycentre of a triangle. In (d)-(f), examples of spherical surfaces approximated by circles generated using levels $L = 1, 2, 3$ of the icosahedron are depicted.

II. METHOD

The goal of the proposed method is the design of a compact descriptor of local lung morphology for a given voxel of interest $v_{xyz} = [x, y, z]$ in a 3-D CT image. The descriptor is computed based on the following steps. 1) Sampling the image intensity in a neighborhood $N_v$ of $v_{xyz}$ in the image $I(x, y, z)$, based on a specific sampling pattern. Sequences of samples are used to define intensity profiles around $v_{xyz}$. 2) Frequency analysis of intensity profiles, to obtain a spectral signature for each profile. 3) Quantization of the feature space defined by collecting several spectral signatures, to assign each spectrum a characteristic label. The descriptor is finally obtained by computing statistics of labeled spectra in $N_v$. In the following sections, each part of the method is described in details. For the sake of clarity, the BoF descriptor is first formulated assuming that the diameter $d$ of the nodule of interest is known. In Section II-E, a technique to estimate the nodule diameter is introduced. The general procedure to compute the descriptor is schematized in Algorithm 1. Additional procedures of sampling and nodule size estimation are outlined in Algorithms 2 and 3.¹

A. Sampling

The construction of the sampling pattern is schematized in Fig. 2. Given the voxel of interest $v_{xyz}$, a set of points contained into the volume of a sphere of radius $R$ centered on $v_{xyz}$ is sampled. The sampling pattern inside the volume is defined using a set of $G$ concentric spherical surfaces with different radii $r_1, \ldots, r_G$, in a multi-scale fashion. Each spherical surface is approximated by a set of concentric circles [see Fig. 2(a)]. In order to equally distribute the circles to cover the spherical surface, the geometrical properties of the icosahedron are exploited [26]. Using the icosahedron, a semi-spherical surface is obtained by $L$ iterative subdivision of four initial ($L = 0$) triangles into four sub-triangles each (step 2.8) [see Fig. 2(c)].

¹In the text, we refer to a generic step B of algorithm A as step A.B.
Given the nodule diameter $d$, two sets of spheres $\Gamma_{\text{int}}$ and $\Gamma_{\text{ext}}$, respectively internal and external to the nodule are defined [see Fig. 2(a)]. The total set of spheres is $\Gamma = \Gamma_{\text{int}} \cup \Gamma_{\text{ext}}$. Consequently, we define $G = \Gamma_{\text{int}} - |\Gamma_{\text{int}}|$ and $G_{\text{ext}} = |\Gamma_{\text{ext}}|$ (step 1.2). The distance $\Delta r$ between subsequent spheres is defined as $\Delta r = d/2G_{\text{int}}$ (step 2.1) and, consequently, $R = G\Delta r$. This procedure makes the sampling strategy scale-invariant, since the distance between spheres and, consequently, the whole sampling pattern is adapted based on the nodule diameter. It is worth noting that the distance between spheres can be arbitrarily defined.

For each circle, the image intensity values $s_m = I(\odot r_{\phi_k}, \theta_j, \psi_l)$ taken at a set of angles $\Omega = \{\omega_m\}$ ($m = 1, \ldots, M$) with constant angular distance $\Delta \omega$ along the circle are sampled [steps 2.10–2.12; Fig. 2(b)]. We define the signal obtained from the ordered sequence of samples $s_1, \ldots, s_M$ along the circle as intensity profile. For computation and visualization purposes, the profiles are organized as rows of a matrix $\mathcal{M}_{\text{xyz}}$ (step 2.11). An example is depicted in Fig. 3. As a consequence, $G$ groups of signals, one for each sphere, are visible in the matrix. In terms of implementation, the multi-scale sampling procedure is applied by resampling the volume of interest at each scale using trilinear interpolation.

### B. Frequency Analysis

Given the set of profiles sampled at different locations of the CT scan (see Figs. 1 and 3), one can state that 1) different regions of interest show different characteristic patterns; 2) given the circularity of the sampling strategy, each profile can be treated as a periodic signal. As a consequence, frequency analysis can be applied to each intensity profiles, treated as an independent signal, in order to extract information related to the morphology captured through sampling. It is known that periodic signals have a discrete Fourier transform, also called spectrum. For this purpose, the Fourier transform is applied to each intensity profile as

$$\Psi_{\gamma,k,j}[f] = \sum_{m=0}^{M-1} I(\odot r_{\phi_k}, \theta_j)[m]e^{-i2\pi f m R}.$$  

The absolute value of $\Psi$ represents a spectral signature, which encodes information on the morphology in the CT scan around the voxel of interest. Examples of spectra are depicted in Fig. 1(e) and (f) and in Fig. 3(a). For computational efficiency, we apply the fast Fourier transform (FFT) algorithm (step 1.8).

Given the symmetry of the absolute value of the spectrum, we only take the first half of values of each $\Psi_{\gamma,k,j}[f]$, with $f = 0, \ldots, M/2$. It is worth noting that the FFT is separately applied to each intensity profile. As a consequence, given the input signal matrix $\mathcal{M}_{\text{xyz}} \in \mathbb{R}^{G \times N \times M}$, the output of the FFT is a matrix $\mathcal{M}_f \in \mathbb{R}^{G \times N \times M/2}$. In Fig. 3, an example of matrix $\mathcal{M}_{\text{xyz}}$ and the corresponding $\mathcal{M}_f$ is depicted. The information encoded in each spectrum can be also interpreted as the amount of variations in the corresponding 2-D view of the object in the volume of interest. Based on this assumption, a technique for the selection of the most informative views of a pulmonary nodule is proposed in Appendix A.

### C. Feature Space Quantization

As observed in Fig. 1(c) and (d), similar intensity profiles and, consequently, similar spectral signatures in $\mathcal{M}_f$ are shared by different regions of interest at different distances from $v_{\text{xyz}}$ in $\mathcal{N}_Y$. Based on this observation, we assume that the information on nodule morphology is contained in the occurrence of each signature in $\mathcal{N}_Y$. In order to obtain a compact representation of the occurrence of each spectrum in $\mathcal{N}_Y$, we perform a quantization of the feature space defined by a large collection of spectra, where each frequency represents a dimension of the space. The quantization of the feature space has two advantages. 1) It allows to identify a number of representative spectra, interpreted as a sequence of frequency components, which can be used to label each input signal. In this way, each spectrum at both training and testing time can be identified using a label instead of a series of frequency components. 2) Given the large amount of spectra extracted from the 3-D sampling pattern, the quantization procedure allows to reduce the variability of the data and the dimensionality of the problem, which makes the classification task much easier. For this purpose, we apply K-Means clustering algorithm to the set of collected spectra. Prior to clustering, data are normalized to have zero mean and unit variance, and the Euclidean distance is used to define clusters. As a result, K centroids are obtained, which represent a codebook of spectral signatures. The use of a codebook is in line with the Bag-of-Visual-Words framework [27]. For this reason, we name our approach as Bag-of-Frequencies.
D. Bag-of-Frequencies Descriptor

As stated before, we assume that the information on nodule morphology is contained in the occurrence of each signature in $N_{y}$. Once the codebook is computed, we use it to assign each new spectrum the label $y \in \{y_1, \ldots, y_K\}$ corresponding to the closest centroid [see Fig. 3(a)]. In this way, a vector $V_y \in \mathbb{N}_1^{G} \times N$ is obtained from each matrix $M_y$ (step 1.9). The BoF descriptor is finally computed as the normalized histogram of labels in $V_y$:

$$BoF = \text{hist}(V_y, nBins - K)$$

(step 1.10). The normalization $\sum_{k=1}^{K} BoF(k) = 1$ is applied in order to make descriptors computed using different configuration parameters comparable. Based on the aforementioned assumptions, the K elements length BoF descriptor compactly encodes statistics of spectra at multiple scales. As a consequence, BoF encodes information on the morphology in $N_{y}$. It is worth noting that using K-Means clustering algorithm for feature space quantization allows to compute a descriptor whose length can be easily tuned with the parameter $K$, which allows to calibrate the complexity of the problem in terms of memory usage and computational cost.

E. Nodule Size Estimation

The BoF descriptor has been formulated based on the assumption that the nodule diameter $d$ is known. In practical terms, this assumption is valid in most of the reading scenarios in radiology, since available reading software often contains tools to measure nodule size. However, for the sake of completeness, we propose an automatic procedure to estimate the nodule size based on the matrix $M_{x,y}$. The technique is outlined in Algorithm 3, and it is based on the fact that 1) the region in the inner part of the nodule generally presents high and fairly homogeneous intensity; 2) the region outside the nodule has instead a lower intensity in average. Those two regions can be clearly distinguished in Fig. 3(a). In order to detect the regions belonging to the nodule, a threshold $T_{HU}$ is applied to $M_{x,y}$, obtaining a binary matrix $B$ (step 3.4). Afterwards, the percentage of positive area in $B$ for each radius $r_y$ is compared against a threshold $P_{HU}$ (step 3.5). The result of the comparison is saved in a vector $\beta$ of binary values, where $\beta(1)$ is forced to be equal to 1 (step 3.1). The presence of nodule border is estimated as the position where the first transition $1 \rightarrow 0$ is detected (step 3.12), and converted into nodule size in mm multiplying by $\Delta_r$ (step 3.13). In the rest of the paper, we refer to this method as approximate Nodule Size (aNS).

For this purpose, we designed classification problems where BoF was used as feature vector for nodule characterization. In these experiments, we focused on problems where difference in morphology is the key feature to capture to discriminate candidates, namely spiculated nodules and vascular structures. Second, we experimentally evaluated the properties of the descriptor in terms of quality of nodule size estimation for the aNS technique, robustness with respect to the central voxel of the nodule, and rotation-invariance property. Finally, we compared the performance of the BoF descriptor and of the nodule size estimator versus state-of-the-art methods. For comparison purposes, we considered 3-D SIFT and SPHARM as state-of-the-art descriptors that encode information on morphology. Furthermore, a descriptor based on statistics of spectra as well as a set of features used in a CAD system were also considered in the comparison. Additional details on comparisons are provided in Section III-C.

A. Nodule Characterization

We considered two classification problems where nodules can be characterized based on their morphology. The first problem consists in automatically discriminating nodules from vascular structures, namely vessels. The second problem consists in assessing the spiculation property of a nodule by discriminating spiculated nodules versus nonspiculated nodules, namely spiculation. In both cases, the nodule characterization problem is formulated as a supervised machine learning problem, where a set of labeled training data is provided and a classifier is trained in a supervised fashion. The aim of these experiments is twofold. First, they allow to assess the effectiveness of BoF as a descriptor of pulmonary nodule morphology. Second, the presented problems can be also interpreted in terms of CAD usage in lung cancer screening scenarios. The vessels problem can be in fact seen as a false positives reduction approach, used to filter out vascular structures detected as nodules in CAD. The spiculation problem can be seen as part of the automatic assessment of nodule malignancy probability, where the presence of spiculation plays an important role [4]. Details on the datasets and on experimental settings are provided in the next section.

1) Nodules Versus Vessels: The first experiment aimed at evaluating the performance of BoF when used to discriminate nodules against vascular structures in CT images. In order to collect a representative dataset for this problem, we randomly selected 1000 scans from the NELSON [24] dataset and 1000 scans from the DLCST [25] dataset. Successively, we applied a recently published CAD system for automatic lung nodule detection [5]. The CAD analysis consists of two steps, namely 1) candidate detection and 2) false positives reduction. In 1), voxels are selected based on double thresholding of image intensity followed by morphological filtering and connected component analysis. From the set of obtained regions, candidates with a diameter smaller than 5 mm were discarded (see [5] for details). In 2), a set of 128 features are extracted and each candidate is classified using a cascade of two classifiers, namely Linear Discriminant Classifier, followed by a combination of 10 GentleBoost classifiers. As output of the CAD, we obtained 673 locations classified as pulmonary nodules. The provided

III. EXPERIMENTS

We evaluated three aspects of the BoF descriptor. First, we evaluated its descriptiveness of lung nodule morphology. For this purpose, we designed classification problems where BoF was used as feature vector for nodule characterization. In these experiments, we focused on problems where difference in morphology is the key feature to capture to discriminate candidates, namely spiculated nodules and vascular structures. Second, we experimentally evaluated the properties of the descriptor in terms of quality of nodule size estimation for the aNS technique, robustness with respect to the central voxel of the nodule, and rotation-invariance property. Finally, we compared the performance of the BoF descriptor and of the nodule size estimator versus state-of-the-art methods. For comparison purposes, we considered 3-D SIFT and SPHARM as state-of-the-art descriptors that encode information on morphology. Furthermore, a descriptor based on statistics of spectra as well as a set of features used in a CAD system were also considered in the comparison. Additional details on comparisons are provided in Section III-C.

Authorized licensed use limited to: Radboud University Nijmegen. Downloaded on February 21,2022 at 10:28:15 UTC from IEEE Xplore. Restrictions apply.
locations correspond to the center of mass of the detected objects. An experienced thoracic radiologist reviewed the findings and labeled 224 nodules, classifying them as (a) perifissural nodules, (b) part-solid nodules, (c) solid nodules, (d) nonsolid nodules. Among the remaining 449 findings, 124 were labeled as vascular structures. In this way, the CAD system in [5] is simply used as a black box to provide points that are further processed and used as input for our experiments. We defined the classification problem using the classes nodule, vessel. Examples of nodules and vessels used in this problem are depicted in Fig. 4(a)–(c) and Fig. 4(d)–(f), respectively.

2) Spiculated Nodules Characterization: The second experiment aimed to evaluate the capability of BoF to detect spiculation in nodules. For this purpose, we took 800 nodules from the DLCST [25] dataset and asked an experienced thoracic radiologist to score for presence of spiculation. As a result, we obtained 51 spiculated nodules. From the remaining 749 nodules, we randomly selected 204 samples, to obtain a ratio of abnormal: normal samples of 1:4. Examples of spiculated nodules used in this problem are depicted in Fig. 4(g)–(i). As for vessels, we defined the spiculation problem as a binary classification problem, where the two classes are spiculated, nonspiculated.

3) Experimental Results: Supervised learning was applied to each voxel of interest in vessels and spiculation, where BoF was computed as described in Section II. The BoF descriptor has a set of free parameters that can be tuned based on the specific problem. In Table I, the set of free parameters is listed along with the values used in our experiments. We define the reported values as the default values for the descriptor, assigned based on a trade-off between reasonable minimal measures used to capture nodule morphology and computational cost of the algorithm. In presence of vascular structures, the aNS technique might underestimate the size of the region of interest. For this reason, we empirically defined a minimal diameter of $d_{\text{min}} = 5$ mm. This value is in line with the size of available data for the vessels problem, since the CAD system in [5], discards every candidate with a diameter $d < 5$ mm by design. In order to create the codebook, a subset of 100 locations for each class was randomly selected from the dataset. In this way, the K-means algorithm had $200NG = 256\,000$ spectra available to compute the centroids. We used a Random Forests classifier [28] for supervised learning purposes, with a number of trees empirically set to $N_{\text{tree}} = 100$. In order to define training and testing sets, for each dataset a ten-fold cross-validation procedure was applied, avoiding intersection of data from the same patient within folds, and the area under the ROC curve ($A_x$) was used as measure for performance evaluation.

The values of $A_x$ as a function of $K$ are depicted in Fig. 5. It is worth noting that after an initial steep improvement, the
Table II

<table>
<thead>
<tr>
<th>Dataset</th>
<th>BoF</th>
<th>MS</th>
<th>3-D SIFT</th>
<th>SPHARM</th>
<th>CAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>vessels</td>
<td>0.867/0.855</td>
<td>0.748/0.743</td>
<td>0.805/0.832</td>
<td>0.693/0.510</td>
<td>0.849</td>
</tr>
<tr>
<td>spiculation</td>
<td>0.907/0.906</td>
<td>0.818/0.822</td>
<td>0.456/0.561</td>
<td>0.630/0.685</td>
<td>-</td>
</tr>
<tr>
<td>nFeat</td>
<td>100</td>
<td>640</td>
<td>640</td>
<td>507</td>
<td>129</td>
</tr>
</tbody>
</table>

Fig. 6. ROC curves for nodule characterization problems vessels and spiculation. Curves using BoF, 3-D SIFT, SPHARM, and MS as descriptors are depicted, and each method is applied based on the nodule size estimated using both aNS and the method in [11]. In (a), the ROC curve for CAD features is also depicted.

performance remains approximately constant, or even slightly worse, for \( K \geq 100 \). For this reason, we set the default parameter to \( K = 100 \). The classification performance for \( K = 100 \) are reported in Table II. In Fig. 6, the ROC curves obtained for \( K = 100 \) are depicted. It can be observed that the value of \( A_z \) obtained for spiculation is higher than the one for vessels. This might be surprising, since from the point of view of morphology the characterization of vascular structures is a much easier problem. However, it is worth noting that the dataset used in vessels is the output of a CAD system, where a previous false-positives reduction step was applied. Therefore, the considered vascular structures are really similar to nodules, making the problem more difficult. Visual examples of BoF are depicted in Fig. 7, where examples of nodules, vessels and spiculated nodules are considered. Three views of each object are depicted, along with the corresponding matrix \( M_{zpp} \). The labels of each spectrum are shown in colors and placed on circular sectors of rings that correspond to the radius of the sphere they belong to.

B. Bag-of-Frequencies Descriptor Properties

We evaluated the properties of BoF in terms of 1) performance of the aNS technique, 2) robustness with respect to the central point and to the orientation of the sampling pattern, and 3) importance of scale invariance.

1) Nodule Size Estimation Error: In order to evaluate the performance of the aNS technique, we randomly took 400 nodules from the NELSON dataset, for which manual measurements were provided. We computed the error in terms of absolute value of the difference between manual and automatic diameters, obtaining \( err = 0.94 \pm 1.16 \) mm. We experimentally observed that despite the good accuracy in nodule size estimation, the general trend for aNS is to underestimate the nodule size.

2) Robustness to Translation: In order to evaluate the robustness of BoF with respect to the central voxel of the nodule, we compared nodule classification results when the center of mass of the nodule is perturbed with Gaussian noise. For this purpose, we added a vector of random displacements \( \Delta_{zpp} = \{ \delta_x, \delta_y, \delta_z \} \) to \( v_{zpp} \), where \( \delta_i \in \mathcal{N}(0, \sigma) \). The results obtained by varying the value of \( \sigma \) between 0.4 and 2.0 mm are reported in Table III. We can observe that BoF is fairly robust with respect to the position used as center of mass of the nodule, since the performance are practically unchanged for a perturbation of \( v_{zpp} \) up to 1.2 mm. It is worth noting that the mean error for nodule size estimation is within this range of values. A significant reduction in performance is observed starting from a perturbation of 2 mm. Since the perturbation is applied independently to each dimension in 3-D, this means that the real distance from the center could reach the value of \( \sqrt{12} \approx 3.46 \) mm, which is almost comparable with the minimum size of a significant nodule (\( \approx 4 \) mm). One could observe that a perturbation of the center of mass makes the descriptor less aware of the exact size of the nodule, when estimated using aNS. In this sense, the aforementioned experiment also gives insights on the robustness of the method to variations of estimated nodules size.

3) Robustness to Rotation: We also investigated the robustness of BoF with respect to rotation of the nodule. In terms of 3-D sampling, it is not possible to provide an analytical proof of rotation-invariance for the descriptor. This is due to the fact that 1) we use an approximation of the spherical surface and 2) in practical terms, CT scans are not perfectly isotropic, making the sampling pattern profile in the \( z \) axis different from the ones in the \( x \) and \( y \) axis. However, we can provide an empirical evaluation of the robustness, based on experiments. For this purpose, we compared the descriptors computed centered on a fixed voxel of interest \( v_{zpp} \), varying the rotation of the scan using different angles. We chose as example a vascular junction, which is expected to exhibit the higher amount of variation along several directions in 3-D. In particular, we applied rotations of \( \pi/2 \) with respect to each dimension, in order to take into account for the maximum variation of ratio in voxel size. As a result, three different orientations of the image are considered. We indicated the three rotations with the name of axial, sagittal and coronal, since the corresponding view can be considered as the reference plan in the three cases. An example of BoF computed for the three rotations is depicted in Fig. 8. We can observe that: 1) the same spectral signatures are present in the three descriptors; 2) the proportion of spectral signatures is comparable for the three cases; 3) variations, when present, are relatively small (\( \approx 0.001 \)), with a maximum difference of value \( \approx 0.02 \). Finally, in order to evaluate the impact of these small variations in terms of CAD usage, we repeated the experiments presented in Sections III-A1 and III-A2 when the three rotations are applied. Performance results are depicted in Table IV. Since the obtained results do not show significant differences, it is possible to conclude that BoF provides a comparable degree of descriptiveness.
Fig. 7. Examples of nodules, vessels, and spiculated nodules used in the experiments. Each example is depicted using 1) its view in the axial, coronal, and sagittal plane, 2) the matrix $M_{xyz}$, and 3) the set of labels assigned by applying the codebook to the set of spectra obtained from $M_{xyz}$. Labels from $y = 1$ to $y = K$ are represented with different colors, where each color represents a spectral signature in the codebook. Labels are depicted in circular pattern, where each ring corresponds to a sphere and each angular sector corresponds to a circle, i.e., an intensity profile, associated to a specific pair of angles $(\phi_1, \phi_2)$, which is assigned a label $y$.

TABLE III

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>INDICATES THE DISTANCE IN mm FROM THE CENTER OF MASS OF THE NODULE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$ [mm]</td>
<td>0.0</td>
</tr>
<tr>
<td>vessels ($A_z$)</td>
<td>0.867</td>
</tr>
<tr>
<td>spiculation ($A_z$)</td>
<td>0.907</td>
</tr>
</tbody>
</table>

Fig. 8. Comparison of BoF descriptors ($K = 106$) computed at the same location $r_{xyz}$ using three different orientations of the image, namely axial, coronal, and sagittal.

TABLE IV

| COMPARISON OF PERFORMANCE FOR THE EXPERIMENTS VESSELS AND SPICULATION WHEN THE IMAGE $I(x, y, z)$ IS ROTATED. CASES AXIAL, SAGITTAL, AND CORONAL INDICATE A ROTATION OF THE IMAGE OF $\pi/2$ RADIANTS AROUND THE $z$, $y$, AND $x$ AXIS, RESPECTIVELY |
|---------------------------------|-------|-------|-------|-------|
| experiment                     | axial | sagittal | coronal |
| vessels ($A_z$)                | 0.867 | 0.864 | 0.871 |
| spiculation ($A_z$)            | 0.907 | 0.903 | 0.911 |

regardless the orientation of the image, which empirically corroborates its rotation-invariance property.

4) Scale-Invariance: Finally, we are interested in evaluating the importance of nodule size estimation, used to achieve scale-invariance properties of the descriptor. For this purpose, we considered the vessels experiment and we set all the parameters as detailed in Table I, but with the scaleInvariant option disabled (see Algorithm 2 for details). We applied the same evaluation procedure and we obtained $A_z = 0.834$ that, compared versus the value $A_z = 0.867$ obtained with scaleInvariant enabled, represents an empirical proof of the effectiveness of such feature.

C. Comparison

In this section, we present the comparison of BoF with state-of-the-art methods. First, we compare BoF versus existing descriptors for 3-D analysis of nodule morphology. For this purpose, we considered 3-D SIFT [16] and SPHARM [22]. Furthermore, a descriptor obtained by simply joining the averaged spectra computed per radius in the sampling pattern, namely Mean Spectrum (MS), is also considered. Beside that, we are also interested in comparing BoF with ad-hoc features used in a CAD system for lung nodules detection [5]. Finally, we investigate the effect of the method used to estimate the nodule size on classification performance. For this purpose, we compared the performance of BoF when $\alpha NS$ is applied against the performance obtained when the method in [11] for automatic lung nodule segmentation is applied. All the comparisons were made using the same experimental settings described in Section III-A3 in terms of classification and cross-validation.

1) Descriptors: For 3-D SIFT, we used a publicly available implementation\(^2\) adopting the default parameters as they are set in the available code. Since the available code of 3-D SIFT does not implement scale-invariance, in order to make a fair comparison with BoF, we applied the descriptor to the volume of interest with size $2G_0$, successively resampled to have a fixed

\(^2\)Available online: http://crcv.ucf.edu/source/3D
size of 15 mm. As a result, a 640 elements length descriptor for each voxel of interest was obtained.

For SPHARM, we used the publicly available SPHARM-PDM library. Since SPHARM requires a 3-D binary map of the segmented nodule, for the sake of comparison we segmented the region in the volume of interest applying the same default threshold value $T_{HU}$ as for BoF, and the volume of interest was resampled as done for SIFT. By using the default parameters for SPHARM, three coefficients for 169 spherical harmonics are obtained. In our experiments, we joined all the coefficients in a unique 507 elements length feature vector, used as descriptor for the 3-D object.

In order to corroborate the effectiveness of the information contained in spectra, in the comparison we also considered a descriptor based on simple statistics of spectra in $N_c$. For this purpose, we chose the most simple statistics one can compute, i.e., the mean spectrum (MS). In order to take into account for the spectral components of each sphere in $N_c$, we averaged the spectra belonging to each sphere with radius $r_k$ independently. The final descriptor is then the concatenation of $G$ mean spectra. If we use the default parameters, the length of the MS descriptor is $G \cdot M/2 = 640$ elements, which is the same length of 3-D SIFT. The main advantage of this approach is that the creation of the codebook via unsupervised clustering is no longer necessary.

The final descriptor we are interested in for comparison purposes is the set of ad-hoc features that have been used in a recently presented CAD system for nodule detection [5]. The set of features in [5], which we refer to as CAD, consists of a combination of specifically tuned descriptors, consisting of 129 values, where the posterior probability of a previous classification task is also used (see [5] for details). We repeated the experiments described in Section III-A1 and III-A2 for 3-D SIFT, SPHARM, and MS, and the experiment in Section III-A1 for CAD as well. The performance in terms of ROC curve are depicted and compared in Fig. 6. The comparison in terms of area under the ROC curve $A_\beta$ is reported in Table II, where the number of features for each descriptor is also indicated.

The BoF descriptor outperforms all the other approaches considered in the comparison. SIFT showed good results for vessels but poor performance for spiculation. Since SIFT relies on features based on local gradients, differences in morphology can be easily encoded in presence of vascular structures. On the one other hand, it is less effective to describe slight variations in morphology of spiculated nodules, mostly observed at the border of the nodule. The MS descriptor is based on simple statistics of the spectra obtained through sampling. Despite the simplicity of the descriptor, it reported better results than SPHARM in all the experiments and it outperformed SIFT in spiculation. On the one hand, this means that the information embedded in spectra is descriptive of morphology. However, despite the larger number of features for MS, the performance is lower than the one obtained using BoF. This fact corroborates the usefulness of feature space quantization through the creation of the codebook. Moreover, using the codebook allows BoF to outperform the other descriptors using a smaller feature vector (see Table II). BoF slightly outperforms CAD in vessels problem. This corroborate the capability of BoF to encode relevant morphological features that are useful to solve a specific problem, though with a formulation of general validity for the descriptor. Finally, it is worth noting that BoF, as MS and SIFT, does not require a prior segmentation of the nodule, which is instead necessary for SPHARM and CAD.

2) Nodule Size Estimation: We investigate the behavior of the considered descriptors when the nodule size is provided either by the $aNS$ technique or by an existing algorithm for automatic lung nodule segmentation. For this purpose, we applied the method presented in [11], where morphological processing is used and special attention is given to the separation between nodules and attached vascular structures (see [11] for details). As a result, the method provides the measure of the volume of the segmented nodule. We compute the approximated diameter as the one of a sphere with volume equivalent to the one given by the algorithm. It is worth noting that the computed measure might not correspond to the exact nodule diameter. However, the BoF descriptor only needs a rough estimation of the nodule size. For SIFT descriptor, in order to fulfill the scale-invariance condition, similarly to what done for BoF, the size of the volume of interest was defined as $2G(d/G_{\text{min}}) + 1$. The results of the comparison are detailed in Table II, beside the results obtained using $aNS$.

We can observe that the effects produced by the choice of the method used for nodule size estimation are negligible in most of the case. Slight differences can be in fact observed in the comparison (see Table II). The only remarkable differences are observed when 3-D SIFT is used in the spiculation problem and when SPHARM is used in vessels. In spiculation, the difference is maybe due to the effect of general underestimation of the nodule size produced by our approach. As a consequence, the subtle variations that characterize the border of a spiculated nodule, compared to a nonspiculated one, may receive a lower weight in SIFT, giving more weight to the internal part of the nodule, which is common, in terms of morphology, to the two classes. In vessels, the difference might be due to the limited volume used to obtain the binary segmentation of the vessel, which may not give enough information to make SPHARM perform well.

When BoF is considered, using $aNS$ allows to obtain better results compared to the initialization using the method in [11]. This might be due to the fact that $aNS$ slightly underestimates the nodule size in average, dealing to a smaller $\Delta r$. As a consequence, closer spheres are defined in the sampling pattern, allowing a finer sampling strategy. This suggests that either smaller values of $\Delta r$ or a larger number of spheres could be defined as default parameters of the descriptor. This aspect will be further investigated in the future.

IV. DISCUSSION

The standard architecture of a CAD system for pulmonary nodules detection consists of two parts, 1) the candidate selection stage and 2) the false-positive reduction stage. In 1), each voxel of the CT scan is assigned a label as possible candidate to belong to a nodule, based on a defined criterion. In 2), groups of
selected voxel are classified as nodules using a classifier trained in a supervised fashion. Based on the obtained experimental results, BoF can be effectively used in both stage 1) and 2) of a CAD system. If computed voxel-wise, BoF may serve for candidate selection, classifying nodules versus parenchyma and other structures; in this case, the vocabulary may consist of an extremely heterogeneous set of frequencies that may appear in a CT scan. If applied after candidate selection, BoF may serve for false positive reduction: in this case, a vocabulary may be created ad-hoc, collecting only frequencies that are expected to be present in the final problem. In this scenario, we experimentally observed that BoF outperforms the results that could be obtained using a descriptor designed ad-hoc for pulmonary nodule analysis in a CAD system [9], [5]. In order to further improve the performance, the combination of BoF with other descriptors such as the ones in [9], [5] is a meaningful approach and it is planned as future research.

In case BoF is used for candidate selection, computational time and memory requirements might be an issue. To deal with these problems, the parameters of the descriptor may be finely tuned for a specific problem. In particular, we observed that good performance can be already obtained with a relatively small value of $K$ [$K < 100$, see Fig. 5(a)]. Furthermore, the number of samples, the level of the icosahedron and the number of spheres can be adapted to fit the speed and memory requirement for the CAD system. The application-specific fine tuning of parameters as well as the evaluation of performance using different kind of classifiers is out of scope for the current paper and will be investigated in the future.

The results obtained in the spiculation problem indicate that BoF is an effective descriptor for presence of spiculation. In this sense, it could be used in CAD systems to assess the malignancy probability of a nodule, based for example on a recently presented model [4]. Additional research could be carried out to investigate the potential of BoF to characterize other types of nodules, such as solid, nonsolid, and part-solid. These categories are taken into account when it comes to deciding the management for a patient, as detailed in the recently presented LungRADS guidelines [3].

Unsupervised clustering is a necessary step in the BoF formulation, in order to define the vocabulary of signatures for the descriptor. Furthermore, using the codebook allows to tune the length of the descriptor, based on the parameter $K$. It is worth noting that the creation of the codebook needs to be performed only once, a task that is in line with the pipeline of supervised learning approaches. Moreover, the same codebook can be reused to label spectra in several problems. In this sense, we could also think of creating a unique codebook from a big set of data collected from several heterogeneous sources. As future development of the proposed descriptor, the use of alternative techniques for feature space partitioning and quantization such as Gaussian mixture models, principal component analysis as well as random partitions will be investigated and compared with the performance obtained using K-Means clustering. An alternative way of encoding the morphologic properties would be to combine histograms computed per radius. However, this approach has the drawback of increasing the length of the descriptor ($GN$ times larger). In order to deal with this drawback, techniques of feature selection or feature reduction may be applied.

In cases where the creation of the codebook is not possible or represents a limitation for the final application, one may consider using the MS descriptor, which carries information on basic statistics of spectra. The MS descriptor has shown lower performance than the BoF descriptor, since it does not take advantage of the space quantization step. In future research, statistics more complex than the mean value per frequency could be used to take into account for the distribution of spectra.

As additional research, we plan to study the applicability of BoF to other characterization problems in different 3-D medical image modalities, not limited to chest CT. Furthermore, the applicability of BoF to general computer vision problems, such as 3-D video characterization or key-point detection represents a straightforward step in our research.

V. CONCLUSION

We have presented bag-of-frequencies, a novel descriptor for the characterization of pulmonary nodules in CT images. The descriptor is solely based on morphological information extracted by direct sampling of image intensities and unsupervised clustering of features, without the need for explicit definition of higher level features. We have shown the properties of the descriptor in terms of nodule morphology characterization, and empirically proven its properties of scale-invariance and rotation-invariance. The validation was performed on a large set of data extracted from the two main European lung cancer trials, NELSON and DLCST. Experimental results demonstrated that BoF outperforms well-known methods for 3-D shape description such as 3-D SIFT and SPHARM. The potential application of BoF to CAD systems for false-positives reduction and nodule characterization has also been shown, with the aim of providing additional and more precise tools for computing nodule malignancy probability.

APPENDIX A

OPTIMAL CROSS-SECTION PLANE

As an additional contribution, we describe a technique to obtain the optimal 2-D view of the nodule as the intersection between the 3-D image and a plane. We define the optimal 2-D view as the one that allows to visualize the most characteristic morphological feature of the nodule. The selection of the optimal plane is based on the analysis of intensity profiles obtained during the sampling procedure. Based on the icosahedron geometry, we generate a set of circles, each of them laying on a different plane. Since the information about nodule morphology captured by each circle is different, we assume that the plane, i.e., the cut in the 3-D image that contains more information about morphology is the one where the intensity profiles show the larger amount of variations. Variations in a signal in the spatial domain imply presence of high frequency components in the frequency domain. For this reason, we consider the high-frequency content measure ($hfcm$) [29] as an indicator of the amount of variations per profile. Since the same sampling pattern is replicated at different scales along the sphere,
we choose the optimal plane as the one that maximizes the averaged $h_{f \text{cm}}$ along all the circles at multiple scales belonging to the same plane defined by the pair of parameters $(\phi_k, \theta_j)$ in spherical coordinates, namely multi-scale $h_{f \text{cm}}$ ($h_{f \text{cm}}^k$):

$$h_{f \text{cm}}^k \cdot j = \frac{1}{G} \sum_{g=1}^{G} \sum_{f=0}^{K} \Psi_{g,k,j}[f] .$$

We define the parameters of the optimal plane as $(\phi_k, \theta_j) = \arg \max_{k,j} (h_{f \text{cm}}^k)$ With this approach, a list of planes ranked from the most to the less informative can be obtained. As a result, the set of the first $n$ optimal 2-D views can be showed to radiologists. Visual examples of the three most informative planes, compared with standard axial–coronal–sagittal views are depicted in Fig. 9 for examples of spiculated nodules. Given these visual results, some observations can be made. First, in all the views provided by $h_{f \text{cm}}$, the variability in nodule morphology is observed. Second, in cases where spiculation is not a marked characteristic of the nodule [(c), (e)], other interesting morphological features such as presence of vessels or pleural wall are highlighted. Finally, even though we have limited the visualization to three planes, it is worth considering that a much larger number of available 2-D views can be obtained and ranked based on the $h_{f \text{cm}}$ measure, which may represent a useful tool for radiologists.

REFERENCES


