



OPEN

A dynamic lesion model for differentiation of malignant and benign pathologies

Weiguo Cao¹, Zhengrong Liang^{1,2}✉, Yongfeng Gao¹, Marc J. Pomeroy^{1,2}, Fangfang Han³, Almas Abbasi¹ & Perry J. Pickhardt⁴

Malignant lesions have a high tendency to invade their surrounding environment compared to benign ones. This paper proposes a dynamic lesion model and explores the 2nd order derivatives at each image voxel, which reflect the rate of change of image intensity, as a quantitative measure of the tendency. The 2nd order derivatives at each image voxel are usually represented by the Hessian matrix, but it is difficult to quantify a matrix field (or image) through the lesion space as a measure of the tendency. We conjecture that the three eigenvalues contain important information of the Hessian matrix and are chosen as the surrogate representation of the Hessian matrix. By treating the three eigenvalues as a vector, called Hessian vector, which is defined in a local coordinate formed by three orthogonal Hessian eigenvectors and further adapting the gray level occurrence computing method to extract the vector texture descriptors (or measures) from the Hessian vector, a quantitative presentation for the dynamic lesion model is completed. The vector texture descriptors were applied to differentiate malignant from benign lesions from two pathologically proven datasets: colon polyps and lung nodules. The classification results not only outperform four state-of-the-art methods but also three radiologist experts.

Malignant lesions have a high tendency to invade their surrounding environment compared to benign ones, thus the lesion growth rate has been widely used as a figure of merit (FOM) for computer-aided diagnosis (CADx) of the lesion, for example, of colon polyps^{1,2} and lung nodules^{3,4}. However, this FOM requires at least two measurements over a time period of weeks or even months apart, incurring additional cost, patient stress and risk of delayed treatment.

On the other hand, numerous image features have been extracted from the lesion volume, followed by sophisticated feature selection and classification operations for the same task of CADx⁵⁻⁸. A typical example of a CADx pipeline takes a few steps of (1) localizing the lesion in a medical image [this step can be computerized, called computer-aided detection (CADe)]; (2) segmenting the lesion volume; (3) extracting features from the lesion volume, and (4) classifying the features for lesion diagnosis. By expanding the pipeline to include not only medical images but also genetic data and more, many successful applications have been reported in the field of radiomics⁹⁻¹¹.

Recently, great efforts have been devoted to construct various convolutional neural network (CNN) architectures to deeply learn the features directly from the input medical images and simultaneously classify the learnt features for the same ultimate objective, i.e. CADx of lesions¹²⁻¹⁵. The efforts have been further devoted to expanding the CNN architectures to learn not only the medical images but also the textures and more^{16,17}.

To our knowledge, both the above research endowers for sophisticated feature extraction and classification operations and the recent CNN-based deep machine learning architectures have not explicitly considered the clinical observations that malignant lesions have high tendency to invade their surrounding environment compared to benign ones¹⁸⁻²³. This paper proposes a dynamic lesion model and explores the feasibility of explicitly considering the high invading tendency for the task of CADx of lesions.

The proposed dynamic lesion model is based on a conjecture that the high tendency is related to the rate of image intensity changing at each image voxel inside the lesion volume. The rate is mathematically described by the 2nd order derivative operation, usually represented by the Hessian matrix at each image voxel²⁴. Thus, the

¹Department of Radiology, State University of New York at Stony Brook, Stony Brook, NY, USA. ²Department of Biomedical Engineering, State University of New York at Stony Brook, Stony Brook, NY, USA. ³School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, Guangdong, People's Republic of China. ⁴Department of Radiology, School of Medicine, University of Wisconsin, Madison, WI, USA. ✉email: jerome.liang@sunysb.edu

dynamic lesion model is represented by a matrix field within the lesion space. Extracting quantitative descriptors from the matrix field or matrix image to measure the tendency is mathematically challenging. To relieve this mathematical challenge, we have another conjecture below.

Based on the recent report that the eigenvalues and the eigenvectors of the Hessian matrix are mathematically related to each other²⁵, we conjecture that the three eigenvalues contain most, if not all, the information of the Hessian matrix and, therefore, the three eigenvalues are chosen as the surrogate representation of the Hessian matrix. By treating the three eigenvalues as a vector, called Hessian vector hereafter, described by its two angular variables and one magnitude variable at each image voxel, we can obtain a vector image through the lesion volume. By adapting the well-known gray level co-occurrence (GLCM) computing method^{7,26} to extract texture measures (or descriptors) from the vector image, followed by classification of the extracted vector texture descriptors, a quantitative description for the dynamic lesion model is completed.

Experimental results from pathologically proven lesion datasets demonstrate the superiority of the dynamic lesion model over state-of-the-art classification methods^{26–29}. Particularly, the demonstration of outperformance of the presented dynamic lesion model over three experienced radiologists indicate the great potential of model-based, task-driven, artificial intelligence (AI)-enabled CADx, which not only learns the experience that experts have learned, but also knows the task that experts are heading to. A similar demonstration of the great potential is the task-driven, AI-enabled AlphaGo³⁰, which can outperform experts because it not only learns how to play the game from the experts, but also knows the task of occupying the largest space when taking each step further, while experts may not be able to achieve that task at every step.

The remainder of this paper is organized as follows. “Methods” describes the dynamic lesion model and the model representation by Hessian matrix and its eigenvalue vector at each image voxel, followed by the presentation of GLCM computation for extraction of the vector texture descriptors from the Hessian vector image. “Experiments and results” presents the experimental outcomes from two pathologically proven datasets, colon polyps and lung nodules, for the task of differentiating malignant from benign lesions with comparison to four existing state-of-the-art lesion classification CADx methods and three radiologist experts. Lastly, discussions and conclusions are drawn in “Conclusions and discussions”.

Methods

As introduced above, this paper proposes a novel dynamic lesion model to perform the CADx task. Inspired by the clinical observation of different invasion properties for different lesion types, the model is deliberately designed to express the invading tendency using Hessian Matrix representation. In this section, we will describe the dynamic model in detail.

Dynamic lesion model and hessian matrix representation. Our proposed dynamic lesion model is based on the observations that a malignant lesion has a high tendency to invade its surrounding environment compared to a benign one^{20,21}. We hypothesize that the tendency for lesions to invade their surrounding environment is related to the rate of change of the image intensity values at each image voxel. The proposed method uses a Hessian matrix representation of the 2nd order derivative to model these changes in intensity.

The rate of the change is mathematically expressed by the 2nd order derivative operator of a scalar function or intensity image $I = I(x, y, z)$ in three-dimensional (3D) space and could be represented by the Hessian matrix as follows²⁴:

$$\mathbf{H}(x, y, z) = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{xy} & I_{yy} & I_{yz} \\ I_{xz} & I_{yz} & I_{zz} \end{bmatrix}, \quad (1)$$

where I_{xx} , I_{xy} , I_{xz} , I_{yy} , I_{yz} , and I_{zz} are the 2nd order derivatives of $I(x, y, z)$. To compute these partial derivatives of the intensity image, we use the well-established Deriche filters with parameter $\alpha = 1$ as a default in the article³¹. The lesion model is then represented by a matrix at each image voxel through the lesion space, i.e. a matrix field or matrix image. Constructing image feature measures or descriptors from a matrix field for the task of CADx of lesions is difficult. To mitigate this difficulty, we adapt the traditional analysis of decomposing the Hessian matrix at each image voxel location.

Hessian eigenvalues as the surrogate representation of hessian matrix. Traditionally, Hessian matrix is frequently utilized to describe local image geometries, such as edges and corners in image processing^{32,33}. Because of the 2nd order derivative operation on a scalar function, Hessian matrix reflects the rate of the scalar function changing, which describes not only the local geometric (shape) properties but also the local moving tendency. While Hessian matrix contains very useful local geometric and dynamic information of a lesion in a medical image, it is very challenging to directly utilize it to construct image feature descriptors for the task of CADx of lesions. To combat this challenge, we adapt the traditional approach of decomposing the Hessian matrix as follows²⁴:

$$\begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{xy} & I_{yy} & I_{yz} \\ I_{xz} & I_{yz} & I_{zz} \end{bmatrix} (\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3) = (\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3) \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \quad (2)$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are three eigenvalues, \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 are their corresponding eigenvectors, respectively, which are all orthogonal with each other and form three new basis in the 3D Euclidean space.

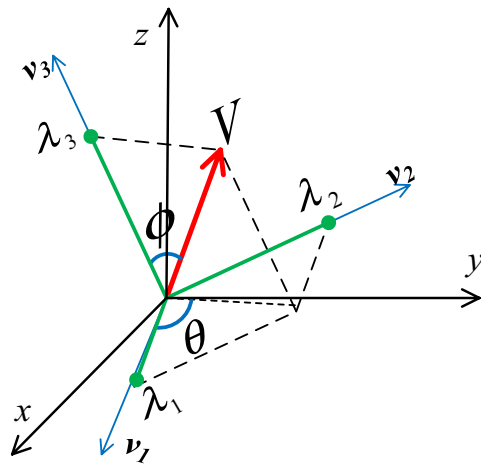


Figure 1. Diagram showing the azimuth and polar angles θ and ϑ respectively from the Hessian vector $V(x, y, z) = (\lambda_1, \lambda_2, \lambda_3)$.

Based on the recent report that the eigenvalues and the eigenvectors are mathematically related to each other²⁵, we can choose either the three eigenvalues or the three eigenvectors to construct image feature descriptors. Since the three eigenvalues contain fewer variables than the three eigenvectors and have advantages in computation, the eigenvalues would be a better choice.

Furthermore, the three eigenvalues can be viewed as the projections of a vector on the three orthogonal eigenvectors, respectively, and the vector contains not only the strength of motion tendency (i.e. the vector magnitude), but also the direction of motion tendency (the vector orientation). Thus, we have another conjecture that the vector, made from the three eigenvalues and called Hessian vector hereafter, is a surrogate representation of the Hessian matrix.

Based on the two conjectures above, our dynamic lesion model is then represented by the Hessian vector. For the CADx task, we will turn to how to extract quantitative measures about the lesion tendency from the vector field.

Vector-based co-occurrence matrix or vector texture image. From the decomposition of Hessian matrix, one image voxel in volumetric data could produce three eigenvalues which could be formed as a local vector (or Hessian vector), i.e. $(\lambda_1, \lambda_2, \lambda_3)$. This 3D vector image could be represented at each image voxel by:

$$V(x, y, z) = (\lambda_1(x, y, z), \lambda_2(x, y, z), \lambda_3(x, y, z)), \quad (3)$$

$$\lambda_m < \lambda_i(x, y, z) < \lambda_M, \quad (4)$$

where (x, y, z) is the coordinate in the original image $I(x, y, z)$, and the values λ_m and λ_M are the minimum and maximum eigenvalues among all the image voxels. Thus, each image voxel inside the Hessian vector field is specified by its three values of $(\lambda_1, \lambda_2, \lambda_3)$ of Eq. (3) within the range of Eq. (4). Figure 1 illustrates that $V(x, y, z)$ should be a Hessian vector in a new local 3D Euclidean space, defined by its three orthogonal eigenvectors, i.e. v_1, v_2 , and v_3 .

As a vector, $V(x, y, z)$ could be expressed by the spherical coordinates as:

$$V(x, y, z) = (\lambda_1, \lambda_2, \lambda_3) = |\lambda|(\cos\theta\cos\vartheta, \sin\theta\cos\vartheta, \sin\vartheta), \quad (5)$$

where $|\lambda|$ is the magnitude, θ and ϑ are two angles representing the vector direction as shown in Fig. 1. For simplicity in presentation, $|\lambda| = \lambda$ will be assumed hereafter. The Hessian vector of Eq. (5) is defined by the three variables as follows:

$$\lambda = \sqrt{(\lambda_1)^2 + (\lambda_2)^2 + (\lambda_3)^2}, \quad (6)$$

$$\theta = \begin{cases} \cos^{-1}\left(\frac{\lambda_1}{\sqrt{(\lambda_1)^2 + (\lambda_2)^2}}\right) & \lambda_2 \geq 0 \\ \pi + \cos^{-1}\left(\frac{\lambda_1}{\sqrt{(\lambda_1)^2 + (\lambda_2)^2}}\right) & \lambda_2 < 0, \end{cases} \quad (7)$$

$$\vartheta = \cos^{-1}\left(\frac{\lambda_3}{\sqrt{(\lambda_1)^2 + (\lambda_2)^2 + (\lambda_3)^2}}\right). \quad (8)$$

Since λ , θ and ϑ are three geometric measures, they are independent with any coordinates even if they are expressed by λ_1, λ_2 , and λ_3 which are three functions expressed by (x, y, z) . So λ , θ , and ϑ are three very stable local metrics, which are more suitable to describe local properties in lesions.

Given the Hessian vector images of Eqs. (5)–(8) from the acquired intensity image, $I = I(x, y, z)$, our task is to find or develop a computing method to extract quantitative measures from these vector images.

The magnitude (λ), azimuth angle (θ) and polar angle (ϑ) form another equivalent Hessian vector of Eq. (3) which produces the vector image, represented at each image voxel by:

$$V(x, y, z) = (\lambda, \theta, \vartheta), \tag{9}$$

$$\lambda_m < \lambda < \lambda_M, \quad 0 \leq \theta \leq 2\pi, \quad 0 \leq \vartheta \leq \pi, \tag{10}$$

where the minimum value λ_m and the maximum value λ_M are determined by Eq. (4), given the acquired image $I(x, y, z)$. Thus, each image voxel inside the Hessian vector field is specified by its three values of $(\lambda, \theta, \vartheta)$ of Eq. (9) within the ranges of Eq. (10), respectively.

Inspired by the well-known GLCM texture descriptor for CADx of lesions^{7,26}, our task here is to find an adequate number of gray levels for each of the three ranges of Eq. (10). Then, GLCM measures can be computed from the vector image of Eq. (9).

Digitalization of the three variables in their ranges of Eq. (10), respectively. Let Q^a and Q^b denote the numbers of gray levels of θ in the range $[0, 2\pi]$ and ϕ in the range $[0, \pi]$, respectively. The digitalization of the angular variables θ and ϑ in the vector image domain is given by:

$$\Theta = \begin{cases} \lfloor (Q^a \cdot \theta) / 2\pi \rfloor & \theta \neq 2\pi \\ Q^a - 1 & \theta = 2\pi, \end{cases} \tag{11}$$

$$\Phi = \begin{cases} \lfloor (Q^b \cdot \phi) / \pi \rfloor & \phi \neq \pi \\ Q^b - 1 & \phi = \pi, \end{cases} \tag{12}$$

where $\lfloor X \rfloor$ is defined as the infimum of X ²⁸.

Due to the scale difference between the angular variables (θ, ϑ) and the magnitude variable (λ), we introduce a scale mapping parameter τ , where τ is an integer. Let Q^c denote the number of gray levels of ρ in the range $[\rho_m, \rho_M]$ and $\Delta\rho = \rho_M - \rho_m$. The digitalization of the vector magnitude ρ is given by:

$$\rho = \sqrt[\tau]{\lambda}, \tag{13a}$$

$$\rho = \begin{cases} \lfloor ((\rho - \rho_m) \cdot Q^c) / \Delta\rho \rfloor & \rho \neq \rho_M \\ Q^c - 1 & \rho = \rho_M. \end{cases} \tag{13b}$$

Thus, we obtain the gray level vector image from the acquired intensity image, $I = I(x, y, z)$, as follows:

$$T = (\rho, \Theta, \Phi). \tag{14}$$

Our next task is to compute the GLCM-alike measures from this gray level vector image.

Computation of vector-based co-occurrence matrix. Given the above gray level vector image of Eq. (14), we present in this section a method which utilizes derivative-based vector to define a new type of co-occurrence matrix (CM), called vector-based CM or VCM and expressed as:

$$VCM_{V_1, V_2}(\Delta l, \Delta m, \Delta n) = \sum_{n=1}^N \sum_{m=1}^M \sum_{l=1}^L \begin{cases} 1 & T(l, m, n) = V_1 \ \&\& \ T(l + \Delta l, m + \Delta m, n + \Delta n) = V_2 \\ 0 & \text{otherwise} \end{cases}, \tag{15}$$

where $T(l, m, n)$ represents the digitalized gray level vector images in 3D space, (L,M,N) indicate the 3D volume size, V_1 and V_2 are a vector pair in $T(l, m, n)$, and $(\Delta l, \Delta m, \Delta n)$ is the offset from $T(l, m, n)$ to $T(l + \Delta l, m + \Delta m, n + \Delta n)$, indicating a direction. Figure 2 illustrates the calculation of Eq. (15).

How to find all offsets around one concerned voxel is one of several key steps in the calculation of VCM. It is well known that the nearest neighbors have the closest relevance to the concerned voxel in topology and statistics. Therefore, the 26 nearest neighbor voxels around the concerned voxel are chosen as our candidates to obtain the offsets, i.e. $(\Delta l, \Delta m, \Delta n)$, in Eq. (15). These offsets can be viewed as the data sampling in the 3D vector image domain along the 26 directions defined by the 26 nearest neighbors³⁴.

By the VCM definition of Eq. (15), the two VCMs with one offset and its inverse offset (or inverse direction) are symmetrical so that half of the 26 nearest neighbor voxels are redundant and should be discarded. Therefore, 13 offsets are retained while their inverse directions are removed. The retained 13 offsets for the VCM calculation are denoted as: (0, 0, 1), (0, 1, 0), (1, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 0), (-1, 1, 0), (0, 1, -1), (1, 0, -1), (1, 1, 1), (-1, 1, 1), (1, 1, -1) and (-1, 1, -1). Thus, using the VCM definition of Eq. (15), we will compute 13 VCMs with 13 offsets through a lesion volume. Each of these 13 VCMs is called vector texture pattern or image hereafter.

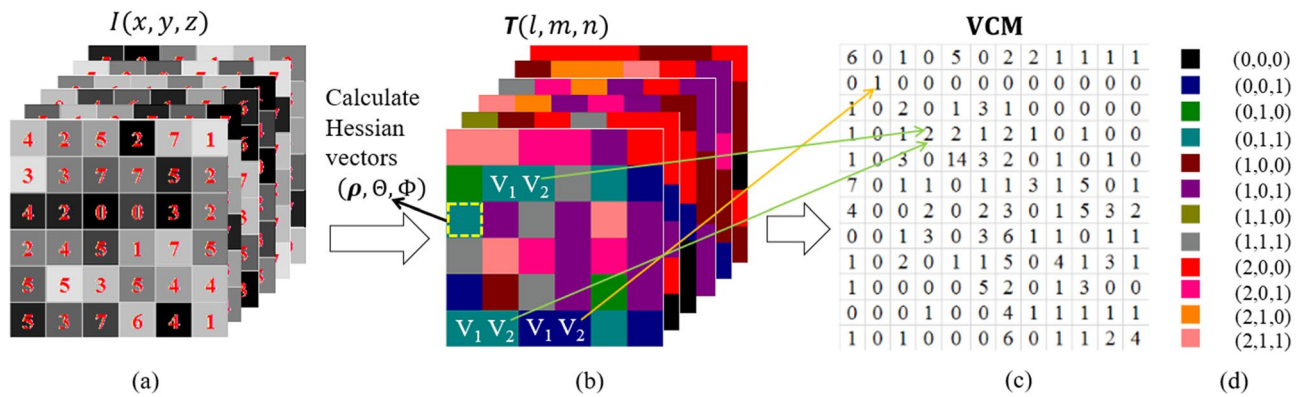


Figure 2. An illustration of vector-based CM calculation. (a) Illustration of one $6 \times 6 \times 6$ volumetric data with 8 Gy levels as an example for illustration purpose. (b) Visualization of Hessian vector-based 3D image where every element is a Hessian vector and $Q^c = 3$, $Q^a = Q^b = 2$ are chosen as an example. (c) The VCM of (b) with offset (or direction) (0, 1, 0) where the total Hessian vector number is equal to $3 \times 2 \times 2 = 12$. (d) The color maps used by 12 vectors in (b).

Extraction of vector texture measures from VCMs. By adapting the 28 texture measure definitions from each offset³⁵, we will extract a total of 364 ($= 13 \times 28$) vector texture measures per lesion, i.e. 28 measures per direction over 13 offsets. Since these 364 texture measures are derived from a 3D lesion and the redundancy from the 13 reverse offsets is not included, they will be called vector texture features (VTFs) hereafter. Classifying these VTFs and evaluating the classification outcomes are presented in the following section.

Classification of the vector texture features. This study proposed a dynamic lesion model to differentiate malignant from benign pathologies. The evaluation of the lesion model is based on two pathology proven datasets, i.e. polyp and nodule datasets. The polyp data acquisition is approved by the Institutional Review Board (IRB) from University of Wisconsin. The nodule data acquisition is approved by the IRB of Stony Brook University. The use of the datasets for the evaluation is approved by the IRB of Stony Brook University. All methods in this study were performed in accordance with the relevant guidelines and regulations.

Once the VTFs are computed for each lesion's gray level vector image of Eq. (14), which is the 2nd order derivative representation of the original image of $I(x, y, z)$, many classifiers can be used to classify these vector texture features, as presented in the introduction section above. To focus on the demonstration of the impact of the proposed lesion model with the associated VTFs extraction strategy for the task of CADx of lesions, we adapt the R package "random forest (RF)" to perform the classification^{36,37}, as it has shown effectiveness in our previous classification experiments^{7,34}. Exploring new machine learning classification methods on the VTFs is beyond the scope of this work.

Due to the limited sample size of pathologically proven medical image datasets, the RF classifier (and all machine learning algorithms) are susceptible to bias from the input data of limited sample size, depending on how it is divided into training and testing data subsets. To reduce this bias effect, 100 randomized groups of lesions are generated, and each group is divided into training and testing sets. The division is made such that there is an equal number of each class (or lesion label) in both the training and testing sets.

We employ a feature selection algorithm from the R package that first reads in the entire list of the vector texture features and generates an importance value for each variable in the feature set based on the GINI index³⁶. We use a forward step feature selection (FSFS) method^{7,38} to perform lesion classification, where the features are added iteratively to that algorithm in decreasing importance order, such that the best performing features are added first to optimize performance. This procedure is repeated for each of the 100 randomized groups of testing and training datasets. The results are evaluated for each group and for each number of features based on the area under the curve (AUC) of receiver operating characteristics curve (ROC) with average over the 100 groups. Results of the above 100 repeated two-fold cross validation experiments are shown in the following section.

Experiments and results

In this section, we first describe the clinical datasets, which will be used to test the proposed lesion model, and evaluate two parameters associated with the digitalization of gray levels and the scale mapping of the magnitude of the Hessian vector to obtain some preliminary results for further experiments. Then lesion classification is performed on the extracted vector texture features or VTFs. At the end, we compare our classification results with four existing CADx methods and three radiologist experts.

Polyp and nodule datasets. *Polyp dataset.* A total of 59 patients including 51% males and 49% females, who were scheduled for clinical colonography examination, were recruited to this study under informed consent after approval by the IRB. Their ages range from 45.9 to 91.6 years old (mean age of 66.5 years old). The patients were scanned by a routine clinical non-contrast CT colonography (CTC) scanning protocol covering the entire abdomen volume prior to the clinical colonography examination. Tube voltage was 120 kVp and dose was deter-

	Training			Testing/prediction		
	Benign	Malignant	Observation group	Benign	Malignant	Observation group
Polyps	15	16	100	16	16	100
Lung nodules	10	24	100	10	24	100

Table 1. The training and testing scheme for the polyp and lung nodule dataset.

	AUC	Acc	Sen	Spe
Polyps	0.986 ± 0.012	0.959 ± 0.026	0.939 ± 0.055	0.980 ± 0.034
Lung nodules	0.861 ± 0.061	0.841 ± 0.052	0.863 ± 0.087	0.786 ± 0.103

Table 2. The classification results of the polyp and lung nodule dataset.

mined by automatic exposure control. A total of 63 polyp masses were found and resected by the clinical examination. The pathological reports indicate 31 benign and 32 malignant polyps. The size of the polyp masses ranges from 3 to 8 cm (mean of 4.2 cm). In the group of benign, four sub-categories are recorded from the pathology reports, they are Serrated Adenoma (3 cases), Tubular Adenoma (2 cases), tubulovillous adenoma (21 cases), Villous Adenoma (5 cases). In the category of malignant, it is Adenocarcinoma only (32 cases). Each abdominal CTC image volume consists of more than 400 image slices, each image slice has an array size of 512×512 , and each image element or voxel is nearly cubic with edge size of 1 mm. The contour of each polyp image slice inside the CTC abdominal image volume was manually delineated by radiological experts on a slice-by-slice manner using a semi-automated segmentation algorithm.

Pulmonary nodule dataset. A total of 66 patients including 52% males and 48% females, who were scheduled for CT-guided lung nodule needle biopsy, were engaged in this study under informed consent after approval by the IRB. The average age of the patients is 69.5 years old, ranging from 33 to 91 years old. A total of 68 lung nodules with 20 benign and 48 malignant were biopsied under the CT scanning with 120 kVp tube voltage and automatic exposure control. The diameter of these nodules ranges from 0.91 to 13.08 cm (mean size of 3.15 cm). Each CT scan covers a portion of the patient entire chest volume, resulting in 100–200 image slices of 512×512 array size, and each image voxel is nearly cubic with edge size of 1 mm. The border of each nodule image slice inside the volumetric patient CT scan was also drawn by experts on a slice-by-slice manner using a semi-automated segmentation algorithm.

Experimental results from polyp dataset. To calculate the VCMs of polyps, the nearest 26 neighbors around the concerned voxel were used to determine the 13 independent offsets or directions which were described in “Vector-based Co-occurrence Matrix or Vector Texture Image”. According to the digitalization scheme of Eqs. (11)–(13b), the total gray levels are equal to $Q^a \times Q^b \times Q^c$. To study the number of total gray levels (i.e. Q^a, Q^b, Q^c) as a parameter in calculating the VCM, fifteen different total gray levels were considered in our experiments, and the parameter varies in the range of {16, 24, 32, 40, 48, 56, 64, 72, 80, 88, 96, 104, 112, 120, 128}. After integer factorization, these 15 total gray levels will generate 235 different combinations of Q^a, Q^b , and Q^c as listed in the Appendix of this paper. These combinations will produce 235 feature sets which will be fed to the RF classifier as training and testing datasets. The parameters which produced the best outcome were then chosen as the total gray levels for the digitalization scheme. The parameters of (2, 2, 32) and (2, 8, 6) for (Q^a, Q^b, Q^c) were chosen in the experiments for classification of polyp and nodule datasets respectively.

The above trial-and-error empirical method for parameter determination was also applied to determine the parameter τ for the scale mapping of the vector magnitude in Eq. (13a). We tested this parameter in the range of 1, 2, ..., 6. The range of 2, 3, 4, 5 is reasonable choice. $\tau = 5$ was chosen in the experiments for classification of our polyps and $\tau = 2$ is selected for our pulmonary nodules.

Once the two parameters were determined, the following steps are taken to generate the final classification results. For the 63 polyps of limited sample size, including 31 benign and 32 malignant ones, we designed three testing schemes, aiming to perform a thorough investigation. We performed the two-fold cross validation as described in the method section above. To obtain robust results, 100 observation groups were randomly selected from the dataset. For each observation group, we selected 31 polyps as the training set and the remaining 32 polyps as the testing set. The experimental set up is presented in the first row of Table 1. After classification, we evaluate the results with four widely used metrics, i.e. AUC, Accuracy (Acc), Sensitivity (Sen), and Specificity (Spe). The classification results, in terms of average and standard deviation from the 100 observation groups, are shown in the first row of Table 2.

Experimental results from nodule dataset. VCM and VTF calculations of the pulmonary nodules share the same steps as the VCM and VTF calculations of the colon polyps, i.e. going through the 13 independent offsets or directions and performing the digitalization scheme to obtain the 235 different combinations of Q^a, Q^b , and Q^c for the corresponding 235 features. By inputting these 235 features to the RF package and going

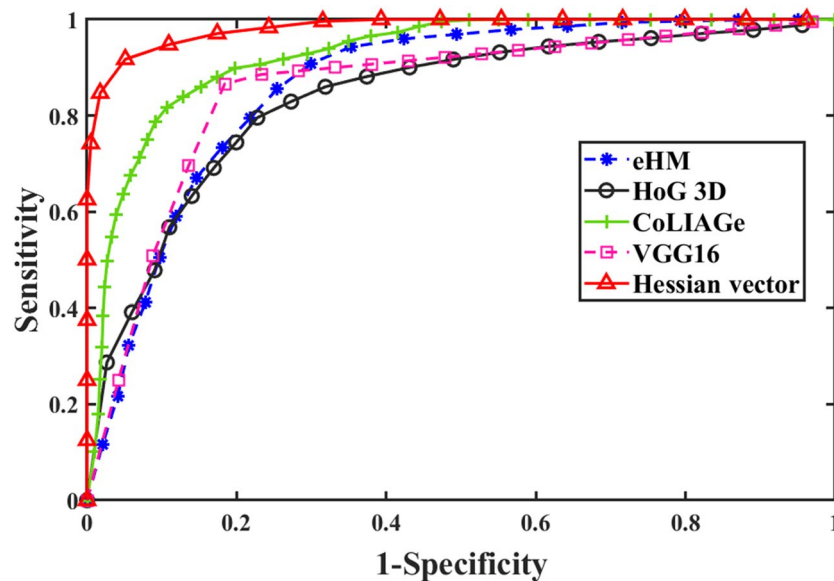


Figure 3. The ROC curves presented for each comparative method and our proposed lesion model over the 63 polyps with twofold cross validation.

through the GINI-based feature sorting and the FSFS-based feature selection^{7,38}, the two parameters (one is the total number of gray levels and the other is the scale mapping of the vector magnitude) will be determined. Once the two parameters were determined, the following procedure is taken to generate the final classification results, similar to the procedure in generating the final classification results of the polyps.

In this classification, the 68 nodules, including 20 benign and 48 malignant, were stochastically divided into 100 groups of training and testing datasets with twofold cross validation as illustrated in the second row of Table 1. The final classification results, in terms of the average and standard deviation over the 100 groups, are shown in the second row of Table 2. Compared to the polyp data results of the first row in Table 1, it is very interesting to note the gap of 8% between the sensitivity and specificity in the very unbalanced nodule dataset (of 20/48) vs. the gap of 4% in the nearly balanced polyp dataset (of 31/32). This may reflect the robustness of the lesion model for unbalanced datasets.

Comparison to other CADx methods. For reference on how well our proposed dynamic lesion model performs, we compare our results to some well-known texture extraction and classification methods and a state-of-the-art CNN deep learning architecture on both the polyp and the nodule datasets as follows.

- Haralick texture method²⁶ with extended measures (eHM)^{7,35}. The Haralick method is widely cited for the GLCM texture descriptor in the field of CADx.
- HoG3D²⁷—this method counts the occurrences of gradient orientation (i.e. only the angular variables of a vector) in some cropped portions of the original intensity image and generates some histograms which are joined to form gradient features.
- CoLIAGe (co-occurrence of local anisotropic gradient orientation)²⁸—this method also employs the gradient angles to extract the entropy of every local patch (or a local voxel group) to form texture features by two joint histograms.
- VGG16²⁹—it is a widely cited CNN deep learning method. It was implemented by the following architecture: total of 20 salient slices were extracted from each lesion volume and inputted to the VGG16 pipeline for lesion feature extraction and classification.

Figure 3 shows the ROC curves of the Hessian vector representation of our lesion model with comparison to the above four CADx methods over the 63 polyps. These ROC curves demonstrate that VTF based polyp classification is the best one among the four state-of-art methods.

In addition to the visual assessment via ROC plots, we further computed the four evaluation metrics. Table 3 shows the comparison results among these four different methods and our proposed lesion model over the 63 colon polyps, where the same RF classifier was used for eHM, HoG3D and CoLIAGe. Our proposed lesion model performed the best. We can see quite clearly that when compared to the eHM, the *Hessian vector* derived texture features outperformed the Haralick texture features with an AUC of 0.982 against 0.876. Compared against the gradient-based features of HoG3D and CoLIAGe, our proposed vector-based texture features improved the performance substantially over the HoG3D features (AUC = 0.804) and the CoLIAGe texture features (AUC = 0.923). The gain of our lesion model is also substantially higher over the VGG-16 outcome (AUC = 0.833).

	AUC	Acc	Sen	Spe
eHM	0.876	0.807	0.858	0.757
HoG3D	0.804	0.713	0.726	0.700
CoLIAGe	0.923	0.836	0.839	0.833
VGG16	0.833	0.740	0.709	0.771
Hessian vector	0.986	0.959	0.939	0.98

Table 3. AUC, accuracy, sensitivity, and specificity values over the 63 colon polyps from the comparative methods and our proposed lesion model.

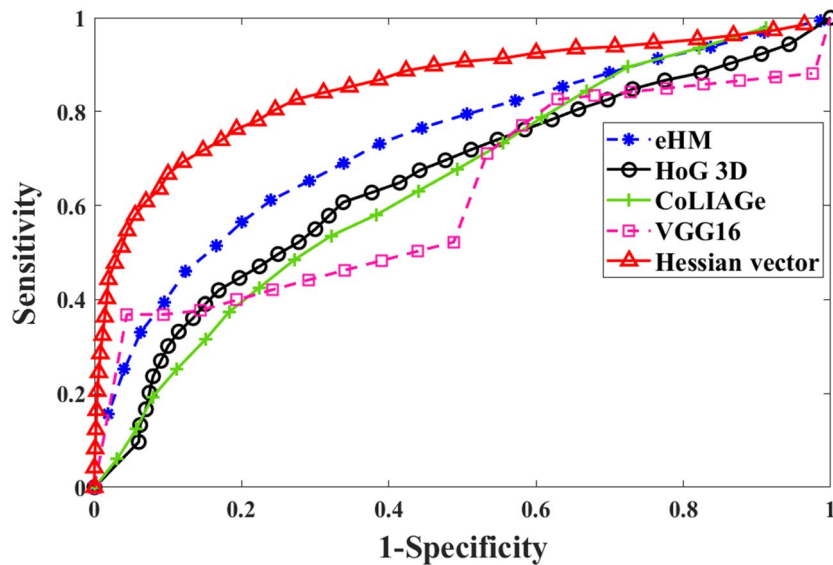


Figure 4. The ROC curves presented for each comparative method and our proposed lesion model over the 68 lung nodules with twofold cross validation.

	AUC	Acc	Sen	Spe
eHM	0.792	0.805	0.881	0.631
HoG3D	0.671	0.706	0.496	0.939
CoLIAGe	0.654	0.756	0.930	0.358
VGG16	0.669	0.687	0.673	0.700
Hessian vector	0.861	0.841	0.863	0.786

Table 4. AUC, accuracy, sensitivity, and specificity values over the 68 lung nodules from the comparative methods and our proposed lesion model.

Figure 4 shows the ROC curves of the Hessian vector representation of our lesion model with comparison to the above four CADx methods over the 68 nodules. The ROC curves in this figure illustrate that VTF based lung nodule classification is much better than eHM, HoG 3D, CoLIAGe and VGG16.

Table 4 shows the comparison results among different methods over the 68 lung nodules, where the same RF classifier was used, except for the deep learning VGG16 which performs feature analysis and classification simultaneously. Our proposed lesion model outperformed all the comparison methods, rendering a similar outcome as that from the 63 polyp dataset. When compared to the eHM, our Hessian vector derived texture features outperformed the Haralick texture features with an AUC of 0.861 against 0.792. Compared against the gradient based features, our proposed vector texture features improved the performance substantially over the HoG3D features (AUC=0.671) and the CoLIAGe texture features (AUC=0.654). The gain is also substantially higher over the VGG-16 outcome (AUC = 0.669).

Dataset/method	eHM	HoG3D	CoLIAGe	VGG16
63 colon polyps	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$
68 lung nodules	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$

Table 5. P-values comparing proposed methods to comparison methods using Wilcoxin ranked sum test over 63 polyps and 68 lung nodules.

In addition to the comparisons by the visualization with the ROC plots and the quantitative measures with the four metrics of AUC, Acc, Sen and Spe, we further performed experiments to show the statistical significance among all the CADx methods. Using the Wilcoxin ranked t-test, we obtained the quantitative measures for significant difference between the results of our lesion model or Hessian vector representation and the four reference methods. Table 5 shows the P-values between our Hessian vector model and the four reference methods, indicating that the classification performance of our lesion model is statistically significantly better than the performances of the four reference methods for both the polyp and nodule datasets (p-value $\ll 0.05$).

Comparison to human observers. Inspired by the competitive game of AlphaGo vs. Expert³⁰, we invited three radiologists to score on the above 59 patients' CTC images, each includes the entire abdomen volume. A total of 63 polyps are embedded inside the 59 CTC images. The human experts are experienced radiologists with 15–35 years of clinical practice in their field. They had all the visualization tools in current clinical setting to exam each polyp, and achieved AUC values of 0.869, 0.926, and 0.960, respectively³⁹. Our dynamic lesion model, called AlphaPolyp here, outperformed all three radiologists by AUC score of 0.986. By this comparison study, we have gained the insights below about why both AlphaGo and AlphaPolyp can outperform the human experts.

For the AlphaGo, it not only learns how to play the game, but also knows the task of occupying an area as large as possible. Given a current situation for next move, AlphaGo can make the next move to occupy the largest possible nearby area while human expert may not be able to achieve that task every time. For our AlphaPolyp design, the task-driven AI-enabled algorithm is built on the clinical observations (i.e. the human experts' learning process) and is trained for the task of matching the pathological outcome of the true lesions. The AlphaPolyp will take all the learnt observations into consideration to predict the pathological outcome while the human experts may not be able to consider all the observations to predict the pathological outcome.

Conclusions and discussions

A dynamic lesion model was proposed to consider the clinical observation that malignant lesions have a high tendency to invade their surrounding environment compared to benign ones. To mathematically describe the lesion model, one conjecture was made that the lesion growing or invading tendency is related to the rate of image intensity changing at each image voxel, and the rate is represented by the 2nd order derivatives at that voxel, leading to a field of Hessian matrix across the image array space. To circumvent the difficulty of extracting quantitative measures from a matrix field for the task of CADx of lesions, another conjecture was made that the three eigenvalues of the Hessian matrix can be a surrogate representation of the Hessian matrix, and further be treated as a vector in the orthogonal system of the eigenvectors, i.e. Hessian vector.

Inspired by the well-known co-occurrence texture measures for the CADx task, we extracted the co-occurrence vector texture features, i.e. VTFs, from the Hessian vector field or image. Classifying the VTFs for the CADx task on two pathologically proven lesion datasets of polyps and nodules provided striking results, outperforming both the state-of-the-art CADx methods and radiologist experts.

While the VTF was demonstrated as a good feature descriptor of the dynamic lesion model, other vector feature descriptors can be explored for the CADx task. This is one of our future research topics.

While the Hessian vector was demonstrated as a good surrogate of the Hessian matrix, use of the Hessian matrix for generation of matrix feature descriptors for the CADx task is desired. This is another topic of our future research endeavors.

More importantly, by the insights on why AlphaGo can outperform the human experts, we believe task-driven AI-enabled CADx systems would be a direction of our future research endeavors^{40–42}.

It can be seen that the evaluation study of the VTF descriptor for the dynamic lesion model is limited by the small sample size of both the polyp and the nodule datasets. We have been devoting significant effort to continually acquire more pathologically proven lesion datasets, although the data acquisition is highly costly.

Appendix

In our experiments, we choose 15 kinds of total gray levels to compute different VCMs, i.e. {16, 24, 32, 40, 48, 56, 64, 72, 80, 88, 96, 104, 112, 120, 128}. According to various of integer factorizations, we generated 235 different combinations of Q^a , Q^b , and Q^c which are listed in Table 6.

Gray level	The set of (Q ^a , Q ^b , Q ^c)
16	(2, 2, 4), (2, 4, 2), (4, 2, 2)
24	(2, 2, 6), (2, 3, 4), (2, 4, 3), (2, 6, 2), (3, 2, 4), (3, 4, 2), (4, 2, 3), (4, 3, 2), (6, 2, 2)
32	(2, 2, 8), (2, 4, 4), (2, 8, 2), (4, 2, 4), (4, 4, 2), (8, 2, 2)
40	(2, 2, 10), (2, 4, 5), (2, 5, 4), (2, 10, 2), (4, 2, 5), (4, 5, 2), (5, 2, 4), (5, 4, 2), (10, 2, 2)
48	(2, 2, 12), (2, 3, 8), (2, 4, 6), (2, 6, 4), (2, 8, 3), (2, 12, 2), (3, 2, 8), (3, 4, 4), (3, 8, 2), (4, 2, 6), (4, 3, 4), (4, 4, 3), (4, 6, 2), (6, 2, 4), (6, 4, 2), (8, 2, 3), (8, 3, 2), (12, 2, 2)
56	(2, 2, 14), (2, 4, 7), (2, 7, 4), (2, 14, 2), (4, 2, 7), (4, 7, 2), (7, 2, 4), (7, 4, 2), (14, 2, 2)
64	(2, 2, 16), (2, 4, 8), (2, 8, 4), (2, 16, 2), (4, 2, 8), (4, 4, 4), (4, 8, 2), (8, 2, 4), (8, 4, 2), (16, 2, 2)
72	(2, 2, 18), (2, 3, 12), (2, 4, 9), (2, 6, 6), (2, 9, 4), (2, 12, 3), (2, 18, 2), (3, 2, 12), (3, 3, 8), (3, 4, 6), (3, 6, 4), (3, 8, 3), (3, 12, 2), (4, 2, 9), (4, 3, 6), (4, 6, 3), (4, 9, 2), (6, 2, 6), (6, 3, 4), (6, 4, 3), (6, 6, 2), (8, 3, 3), (9, 2, 4), (9, 4, 2), (12, 2, 3), (12, 3, 2), (18, 2, 2)
80	(2, 2, 20), (2, 4, 10), (2, 5, 8), (2, 8, 5), (2, 10, 4), (2, 20, 2), (4, 2, 10), (4, 4, 5), (4, 5, 4), (4, 10, 2), (5, 2, 8), (5, 4, 4), (5, 8, 2), (8, 2, 5), (8, 5, 2), (10, 2, 4), (10, 4, 2), (20, 2, 2)
88	(2, 2, 22), (2, 4, 11), (2, 11, 4), (2, 22, 2), (4, 2, 11), (4, 11, 2), (11, 2, 4), (11, 4, 2), (22, 2, 2)
96	(2, 2, 24), (2, 3, 16), (2, 4, 12), (2, 6, 8), (2, 8, 6), (2, 12, 4), (2, 16, 3), (2, 24, 2), (3, 2, 16), (3, 4, 8), (3, 8, 4), (3, 16, 2), (4, 2, 12), (4, 3, 8), (4, 4, 6), (4, 6, 4), (4, 8, 3), (4, 12, 2), (6, 2, 8), (6, 4, 4), (6, 8, 2), (8, 2, 6), (8, 3, 4), (8, 4, 3), (8, 6, 2), (12, 2, 4), (12, 4, 2), (16, 2, 3), (16, 3, 2), (24, 2, 2)
104	(2, 2, 26), (2, 4, 13), (2, 13, 4), (2, 26, 2), (4, 2, 13), (4, 13, 2), (13, 2, 4), (13, 4, 2), (26, 2, 2)
112	(2, 2, 28), (2, 4, 14), (2, 7, 8), (2, 8, 7), (2, 14, 4), (2, 28, 2), (4, 2, 14), (4, 4, 7), (4, 7, 4), (4, 14, 2), (7, 2, 8), (7, 4, 4), (7, 8, 2), (8, 2, 7), (8, 7, 2), (14, 2, 4), (14, 4, 2), (28, 2, 2)
120	(2, 2, 30), (2, 3, 20), (2, 4, 15), (2, 5, 12), (2, 6, 10), (2, 10, 6), (2, 12, 5), (2, 15, 4), (2, 20, 3), (2, 30, 2), (3, 2, 20), (3, 4, 10), (3, 5, 8), (3, 8, 5), (3, 10, 4), (3, 20, 2), (4, 2, 15), (4, 3, 10), (4, 5, 6), (4, 6, 5), (4, 10, 3), (4, 15, 2), (5, 2, 12), (5, 3, 8), (5, 4, 6), (5, 6, 4), (5, 8, 3), (5, 12, 2), (6, 2, 10), (6, 4, 5), (6, 5, 4), (6, 10, 2), (8, 3, 5), (8, 5, 3), (10, 2, 6), (10, 3, 4), (10, 4, 3), (10, 6, 2), (12, 2, 5), (12, 5, 2), (15, 2, 4), (15, 4, 2), (20, 2, 3), (20, 3, 2), (30, 2, 2)
128	(2, 2, 32), (2, 4, 16), (2, 8, 8), (2, 16, 4), (2, 32, 2), (4, 2, 16), (4, 4, 8), (4, 8, 4), (4, 16, 2), (8, 2, 8), (8, 4, 4), (8, 8, 2), (16, 2, 4), (16, 4, 2), (32, 2, 2)

Table 6. Various of combination groups of Q^c, Q^a, and Q^b at different total gray levels for Eq. (15) where a gray level is equal to Q^c × Q^a × Q^b.

Received: 16 September 2020; Accepted: 20 January 2021

Published online: 10 February 2021

References

- Pickhardt, P. J. *et al.* Assessment of volumetric growth rates of small colorectal polyps with CT colonography: A longitudinal study of natural history. *Lancet Oncol.* **14**(8), 711–720 (2013).
- Kato, S. *et al.* Hyperprogressors after immunotherapy: Analysis of genomic alterations associated with accelerated growth rate. *Clin. Cancer Res.* **23**(15), 4242–4250 (2017).
- Kostis, W. J., Reeves, A. P., Yankelevitz, D. F. & Henschke, C. I. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. *IEEE Trans. Med. Imaging* **22**(10), 1259–1274 (2003).
- Ko, J. P. *et al.* Pulmonary nodules: Growth rate assessment in patients by using serial CT and three-dimensional volumetry. *Radiology* **262**(2), 662–671 (2012).
- Chan, H. P., Hadjiiski, L., Zhou, C. & Sahiner, B. Computer-aided diagnosis of lung cancer and pulmonary embolism in computed tomography—a review. *Acad. Radiol.* **15**(5), 535–555 (2008).
- Li, Q. & Nishikawa, R. *Computer-aided Detection and Diagnosis in Medical Imaging* (CRC Press, Boca Raton, 2015). <https://doi.org/10.1201/b18191>.
- Hu, Y. *et al.* Texture feature extraction and analysis for polyp differentiation via computed tomography colonography. *IEEE Trans. Med. Imaging* **35**(6), 1522–1531 (2016).
- Sareena, A. M. & Kaur, M. Computer-aided-diagnosis in colorectal cancer: A survey of state of the art techniques. In *2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore*, pp. 1–6 (2016). <https://doi.org/10.1109/INVENTIVE.2016.7823260>.
- Aerts, H. *et al.* Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* **5**, 4006. <https://doi.org/10.1038/ncomms5006> (2014).
- Yip, S. & Aerts, H. Applications and limitations of radiomics. *Phys. Med. Biol.* **61**(13), 150–166 (2016).
- Zhang, Y.-H. “Radiomics in cancer prognosis: Applications and limitations of quantitative texture analysis”, PhD Dissertation, Karolinska Institute (Stockholm, Sweden, 2018).
- Cheng, J. *et al.* Computer-aided diagnosis with deep learning architecture: Applications to breast lesions in US images and pulmonary nodules in CT scans. *Sci. Rep.* **6**, 24454 (2016).
- Komeda, Y. *et al.* Computer-aided diagnosis based on convolutional neural network system for colorectal polyp classification: Preliminary experience. *Oncology* **93**, 30–34 (2017).
- Ahmad, O. F. *et al.* Artificial intelligence and computer-aided diagnosis in colonoscopy: Current evidence and future directions. *Lancet Gastroenterol. Hepatol.* **4**(1), 71–80 (2019).
- Zhang, S. *et al.* An investigation of CNN models for differentiating malignant from benign lesions using small pathologically proven datasets. *Comput. Med. Imaging Graph.* **77**, 101645 (2019).
- Wang, H. *et al.* A Hybrid CNN feature model for pulmonary nodule malignancy risk differentiation. *J. X-Ray Sci. Technol.* **26**(2), 171–187 (2018).
- Tan, J. *et al.* 3D-GLCM CNN: A 3-dimensional gray-level co-occurrence matrix based CNN model for polyp classification via CT colonography. *IEEE Trans. Med. Imaging* **39**(6), 2013–2024. <https://doi.org/10.1109/TMI.2019.2963177> (2019).
- Acerbi, I. *et al.* Human breast cancer invasion and aggression correlates with ECM stiffening and immune cell infiltration. *Integr. Biol.* **7**(10), 1120–1134 (2015).

19. Ahmed, H. U., Arya, M., Path, A. F. & Emberton, M. Do low-grade and low-volume prostate cancers bear the hallmarks of malignancy?. *Lancet Oncol.* **13**(11), e509–e517 (2012).
20. Gibbs, P., Liney, G. P., Lowry, M., Kneeshaw, P. J. & Turnbull, L. W. Differentiation of benign and malignant sub-1cm breast lesions using dynamic contrast enhanced MRI. *Breast* **13**(2), 115–121 (2004).
21. Jansen, S. A. *et al.* Differentiation between benign and malignant breast lesions detected by bilateral dynamic contrast-enhanced MRI: A sensitivity and specificity study. *Magn. Reson. Med.* **59**(4), 747–754 (2008).
22. Gity, M. *et al.* Differentiation of benign from malignant axillary masses by dynamic contrast-enhanced MRI (DCE-MRI): Quantitative and semi-quantitative analysis at 3-Tesla MRI. *Asian Pac. J. Cancer Prev. APJCP* **20**(4), 1073–1079 (2019).
23. Ibrahim, A. *et al.* Radiomics for precision medicine: Current challenges, future prospects, and the proposal of a new framework. *Methods* <https://doi.org/10.1016/j.ymeth.2020.05.022> (2020).
24. Callahan, J. J. *Advanced Calculus: A Geometric View* (Springer Science & Business Media, New York, 2010).
25. Denton, P. B., Parke, S. J., Tao, T., & Zhang, X. Eigenvectors from eigenvalues. arXiv:1908.03795v1 (2019).
26. Haralick, R., Shanmugam, K. & Dinstein, I. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **3**(6), 610–621 (1973).
27. Dalal, N. & Triggs, B. Histograms of oriented gradients for human detection. In *IEEE Computer Society International Conference on Computer Vision & Pattern Recognition (CVPR'05)*, vol. 1, 886–893 (2005).
28. Prasanna, P., Tiwari, P. & Madabhushi, A. Co-occurrence of local anisotropic gradient orientations (CoLIAGe): A new radiomics descriptor. *Sci. Rep.* **6**, 37241 (2016).
29. Simonyan, K. & Zisserman, A. Very deep convolution networks for large-scale image recognition. arXiv: 1409–1556 (2014).
30. Wang, F.-Y. *et al.* Where does AlphaGo go: From church-turing thesis to AlphaGo thesis and beyond. *IEEE/CAA J. Autom. Sin.* **3**(2), 113–120 (2016).
31. Monga, O. & Benayoun, S. Using partial derivatives of 3D images to extract typical surface features. *Comput. Vis. Image Underst.* **61**(2), 71–189 (1995).
32. Zhang, Z., Yang, J., Su, X., Ding, L. & Wang, Y. Multi-scale image segmentation of coal piles on a belt based on the Hessian matrix. *Particuology* **11**(5), 549–555 (2013).
33. Li, B., Li, H., & Söderström, U. Scale-invariant corner keypoints. In *2014 IEEE International Conference on Image Processing (ICIP), 29 January 2015, Paris, France*.
34. Song, B. *et al.* Volumetric texture features from higher-order images for diagnosis of colon lesions via CT colonography. *Int. J. Comput. Assist. Radiol. Surg.* **9**(6), 1021–1031 (2014).
35. Cao, W. *et al.* Multilayer feature selection method for polyp classification via computed tomographic colonography. *J. Med. Imaging* **6**(4), 044503. <https://doi.org/10.1117/1.JMI.6.4.044503> (2019).
36. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
37. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **26**(1), 217–222 (2005).
38. Takemura, A., Shimizu, A. & Hamamoto, K. Discrimination of breast tumors in ultrasonic images using an ensemble classifier based on the AdaBoost algorithm with feature selection. *IEEE Trans. Med. Imaging* **29**(3), 598–609 (2010).
39. Pooler, B. *et al.* Volumetric textural analysis of colorectal masses at CT colonography: Differentiating benign versus malignant pathology and comparison with human reader performance. *Acad. Radiol.* **26**(1), 30–37. <https://doi.org/10.1016/j.acra.2018.03.002> (2019).
40. Gerevini, A. E. *et al.* Automatic classification of radiological reports for clinical care. *Artif. Intell. Med.* **91**, 72–81 (2018).
41. Park, S. H. & Han, K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. *Radiology* **286**, 800–809 (2018).
42. Petelli, L., Gerevini, A. E., Lavelli, A., Olivato, M. & Serina, I. Deep learning for classification of radiology reports with a hierarchical schema. *Proc. Comput. Sci.* **176**, 349–359 (2020).

Acknowledgements

This work was partially supported by the NIH/NCI grant #CA206171 and #CA220004.

Author contributions

Z.L. proposed the original notion of the vector texture lesion model and supervised the research, W.C. designed and implemented the vector texture lesion model. W.C., Z.L., Y.G., and M.J.P. designed the experiments. W.C. and M.J.P. performed the data preprocessing and segmentation. F.H. collected and prepared the clinical lung nodule dataset. P.J.P. collected the clinical colon polyp dataset. A.A. reviewed all clinical data and segmentation accuracy. W.C., Z.L., and Y.G. wrote the manuscript. All the authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Z.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021