A Novel Framework Using Graph-Based Clustering for Dental X-Ray Image Search in Medical Diagnosis

Tran Manh Tuan and Le Hoang Son

Abstract—**Dental X-Ray image search is an important process in medical diagnosis for diagnosing exactly dental diseases of a patient. This problem is regarded as the matching of a dental X-Ray image with diseases patterns in the database. In this paper, we propose a novel framework using graph-based clustering for dental X-Ray image search. This framework firstly extracts dental features from an X-Ray image to a dental feature database and then uses a vector quantization algorithm to clarify the principal records from the database. Each record is now regarded as a node in a graph which is classified by a graph-based clustering algorithm according to the disease patterns. The dental X-Ray image is classified having disease or non-disease according to other disease patterns in the same group. The new method is experimental validated on a real dataset of 13 dental X-ray images taken from Hanoi Medical University, Vietnam at the period of 2014-2015. Three variants of the framework namely Prim spanning tree (GCP), Kruskal spanning tree (GCK), and Affinity Propagation Clustering (APC) has been implemented. The experimental results suggest the best variant in term of accuracy.**

Index Terms—**Dental X-ray images, graph-based clustering, medical diagnosis, affinity propagation clustering.**

I. INTRODUCTION

A computerized medical diagnosis system is of great interest to clinicians for accurate decision making of possible diseases and treatments. Because of knowledge re-use and the capability to learn by examples, it has an important role to routine works of clinicians. In recent years, there have been many researches that aim to develop such the system. In 2012, Subgahata Chattopadhyay *et al*. [1] presented an application of Bayesian network to diagnose toothache. Kavitha *et al*. [2] used Support Vector Machine (SVM) to predict the osteoporosis from dental images. In 2014, Kantesh and Xu [3] proposed a fuzzy-based method to predict heart risk. Sutton [4] used Fuzzy Neighbor K-Nearest Neighbour (FKNN) method in different dentistry problems.

In dentistry, dental X-Ray image search is the core process of medical diagnosis for diagnosing exactly dental diseases of a patient. This problem is regarded as the matching of a dental X-Ray image with diseases patterns in the database. If it is similar to a disease pattern then its status is 'disease'; otherwise is 'non-disease'. There have been some achievements in form of image searching based on graph modeling [5]-[7]. Based on these studies, we model the dental

X-Ray image search in a graph and then use a graph-based clustering method to determine the cluster containing the new X-Ray image. From that, diagnosing information is determined via membership degrees of the cluster.

In this paper, we propose a novel framework using graph-based clustering for dental X-Ray image search. The novel parts of this framework in comparison with the relevant researches are: 1) generate a dental feature database of dental X-ray images; 2) use a vector quantization algorithm to clarify the principal records from the database; 3) model the principal records as a graph; 4) apply a graph-based clustering algorithm to classify the graph according to the disease patterns; 5) determine the disease for the new X-Ray image; 6) implement three variants of the framework namely Prim spanning tree (GCP), Kruskal spanning tree (GCK), and Affinity Propagation Clustering (APC); 7) illustrate the methods in the real dataset and evaluate the performance by various validity indices; 8) suggest the best method in term of accuracy.

The next parts of this paper are organized as follows. Section II presents background knowledge of graph-based clustering. Section III describes main mechanism of the proposed method and related concepts. Section IV demonstrates experimental results. Finally, conclusions and further works are given in Section V.

II. GRAPH-BASED CLUSTERING ALGORITHMS

A. Graph Clustering Based on Spanning Tree

A graph-theoretical clustering method [7, 8] is different to many clustering approaches proposed in the literature. Based on a graph, it constructs a minimum spanning tree (MST) and repeatedly removes unsuitable edges until stopping conditions hold (Table I). The criterion to remove edges is one of the following cases:

Criterion 1: Based on distance among nodes, remove the longest node in each of iterations. The stopping condition in this case is threshold δ of edge that is removed. Zahn [9] proposed an average value to delete edges from the graph. This threshold is computed by:

$$\delta = \lambda \times \frac{1}{n-1} \sum_{e \in E'} w(e) \tag{1}$$

where λ is a parameter and n is the number of objects. E' is edge set of MST.

Criterion 2: is used when the distribution of clusters is different. One edge will be removed if its length is f times longer than the average length of its neighbors. Parameter f is defined by users and can be modified in the clustering

progress. But the question is how many neighbors? Point x_i is near to x_j if the path from x_i to x_j is less than or equals to k edges. The challenge is how to select the value of f and k.

Criterion 3: proposed by Zahn [9] for the purpose of separating an input dataset. The author used some validity indices to discover the best algorithm for this problem.

TABLE I: GRAPH-BASED CLUSTERING METHOD USING MST

Input	The graph of the dataset: G(X) = (V,E)
Output	The set of partitioned sub-datasets: S
Algorithm	
1	Construct MST of input graph using Prim or Kruskal algorithm
2	Repeat
3	Remove unsuitable edges to get a set of connected parts.
4	Until the stop condition is satisfied.

We can use Prim or Kruskal to construct a MST. The Prim algorithm is presented as following:

Step 1: (initializations).

$O=\{1\}$ ($V(1)$ is the root of the T tree). $P=\{2,...,n\}$For every j belonging to P: $e(j):=c[e(j,1)]$, $p(j)=1$

By definition of the cost function: $e(j)$ = infinite when $V(j)$ does not connect to V(1)).

Step 2: Choose k: $e(k) \leq e(j)$ for every j belonging to P. In case of tight choose the smaller one.

Step 3: Exchange the O set with the set produced by the union of the O set and $\{k\}$. Exchange the P set with the set produced by the difference of the P set and $\{k\}$: $(P \leftarrow P-\{k\})$ If $P = 0$ then stop.

Step 4: For every j belonging to P compare $e(j)$ with $c[e(kj)]$. If $e(j) > c[e(kj)]$ exchange $e(j) \leftarrow c(e(kj))$. Go back to Step 1.

Similarity, Kruskal algorithm is also given as below:

Step 1: Create a forest F (a set of trees), where each vertex in the graph is a separate tree

Step 2: Create a set S containing all the edges in the graph sorted increasingly by their weights.

Step 3: While S is nonempty and F is not yet spanning:

Step 4: Remove an edge with minimum weight from S

Step 5: If the removed edge connects two different trees then add it to the forest F, combining two trees into a single tree.

At the termination of the algorithm, the forest forms a minimum spanning forest of the graph. If the graph is connected, the forest has a single component and forms a minimum spanning tree.

In the implement results of this paper, we use both algorithms and evaluate how these algorithms influent to clustering results.

B. Affinity Propagation Clustering

Affinity propagation (AP) is a relatively new clustering algorithm introduced by Frey and Dueck [10]. AP can be viewed as exchanging messages between the data points themselves. It has been applied in various fields recently [11], [12] and has been implemented as an R package [13]. AP [12] takes an input matrix of real-value similarities between data points

$$S = \{s_{ij}\}, i = 1,..., N; j = 1,...,N..$$

where s_{ij} indicates the similarity between two objects x_i and x_j.

$$s_{ij} = -\left\| x_i - x_j \right\|^2, i \neq j$$

In the case of $i = j = k$ then we have "preferences" s_{kk}, $k = 1,..., N$ that influence the probability of one point being an exemplar. AP produces two kinds of messages called "responsibility" and "availability" respectively. The former denoted by $r(i,k)$ is sent from data point i to candidate exemplar point k whereas the later denoted by $a(i,k)$ is sent from the candidate representative point k to point i. The formulas to update two kinds of messages are presented as follow:

$$r(i,k) := s(i,k) - \max_{k' \neq k}\{a(i,k') + s(i,k')\} \quad (2)$$

$$a(i,k) := \min \left\{0, r(k,k) + \max_{i' \notin \{i,k\}}\{0, r(i',k)\}\right\} \quad (3)$$

We also have the term "self-availability" that reflects accumulated evidence considered as an exemplar. It is updated differently as follows:

$$a(k,k) := \max_{i' \neq k}\{0, r(i',k)\} \quad (4)$$

After iterative message passing, exemplars can be identified by calculating maximum of $a(i,k) + r(i,k)$ for points. When $k = i$, point i is selected as an exemplar, or point k is the exemplar of point i. To reduce the oscillations when updating messages, damping factor λ is introduced to iteration process (t is the iteration steps):

$$r_{t+1}(i,k) := \lambda.r_t(i,k) + (1-\lambda).r_{t+1}(i,k) \quad (5)$$

$$a_{t+1}(i,k) := \lambda.a_t(i,k) + (1-\lambda).a_{t+1}(i,k) \quad (6)$$

AP have some main advantages such as: AP does not need to pre-assign the number of clusters; the greater value of preference, more number of clusters AP generates; AP also only accepts the collection of similarities as input, which eliminates the need to deal with the raw dataset directly; AP demonstrate its ability of processing large datasets rapidly and effectively from wide-ranging applications [10].

TABLE II. AFFINITY PROPAGATION CLUSTERING (APC)

Input	A set of pairwise similarities $\{s(i,k)\}_{i,k=1,2,...,N}$
Output	Assignments: $\hat{c} = (\hat{c}_1, ..., \hat{c}_N)$
Algorithm	
1	Initialization: set availabilities to zero, $i, k : a(i, k)=0$.
2	Repeat:
3	Update responsibility using equation (5)
4	Update availability using equation (6)
6	Until convergence

However, AP still meets some limitations. Firstly, it is difficult to determine suitable value of preference. Secondly, the damping factor acquires an appropriate value. Frey and Dueck [10] suggested a default damping factor to keep the balance between convergence and oscillation. Denote that $\hat{c}_i = \arg\max_k [a(i,k) + r(i,k)]$ and $\hat{c}_i$ indexes the cluster's exemplar to which point i is assigned. Specifically, if point i is in a cluster with point k serving as the exemplar, then $\hat{c}_i = k$ and $\hat{c}_k = k$. The *Affinity Propagation Clustering* (APC) is described in Table II below.

III. THE PROPOSED FRAMEWORK

In this section, we address the general diagram of the dental image searching problem in Section A. Section B presents some features of dental images. Section C describes the vector quantization method. Lastly, graph representation is shown in Section D.

A. The Cooperative Framework

In Fig. 1, a general diagram of the new framework is presented. From a dental image, feature extraction is used to construct a dental feature database including those of a new dental X-Ray image and of the disease patterns. A vector quantization method is applied to get principal records from the database, which is then formulated as a complete graph where each node is a record. Graph-based clustering is then used to determine which cluster contains the new dental X-Ray image. Diagnosis is made by opting k-nearest neighbor (KNN) method for this cluster with $k=1$. Finally, the diagnostic disease is compared with the real one to compute the validity index value, which shows the performance of the framework.

B. Dental Feature Extraction

This section presents 5 basic dental image features for the feature extraction step.

1) Entropy, edge-value and intensity features

Entropy: is used to define the randomness level of achieved information [14].

$$r(x, y) = -\sum_{i=1}^{L} p(z_i) \log_2 p(z_i) \qquad (7)$$

With a random variable z, a probability of ith pixel $p(z_i)$, $i = 1, 2, ..., L$ (L being the number of pixels). Normalizing this feature, we get the result as in equation (8).

$$R(x, y) = \frac{r(x, y)}{\max\{r(x, y)\}} \qquad (8)$$

+ Edge-value and intensity [17]:

$$e(x, y) = \sum_{p=-\lfloor w/2 \rfloor}^{\lfloor w/2 \rfloor} \sum_{q=-\lfloor w/2 \rfloor}^{\lfloor w/2 \rfloor} b(x, y) \qquad (9)$$

$$b(x, y) = \begin{cases} 1, & \nabla f(x, y) \geq T_1 \\ 0, & \nabla f(x, y) < T_1 \end{cases} \qquad (10)$$

$$\nabla f(x, y) = \sqrt{\left(\frac{\partial g(x, y)}{\partial x}\right)^2 + \left(\frac{\partial g(x, y)}{\partial y}\right)^2} \qquad (11)$$
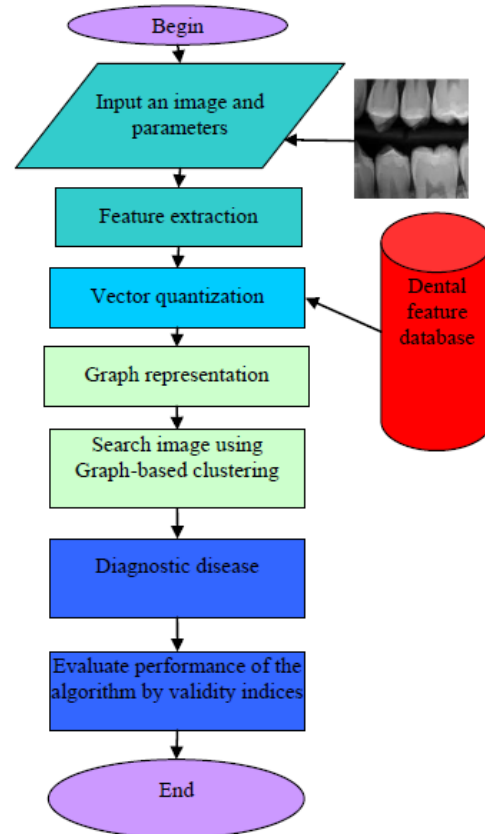

Fig. 1. General diagram.

These features are normalized in the forms:

$$E(x, y) = \frac{e(x, y)}{\max\{e(x, y)\}} \qquad (12)$$

$$G(x, y) = \frac{g(x, y)}{\max\{g(x, y)\}} \qquad (13)$$

2) Local binary patterns — LBP feature

This feature is defined by performing these steps [15]:

1) Select a 3×3 window template from a given central pixel.
2) Compare its value with those of pixels in the window. If this value is greater than mark as 1; otherwise mark as 0.
3) Put all binary values from the top-left pixel to the end pixel by clock-wise direction into a 8-bit string. Convert it to decimal system.

$$LBP(x_c, y_c) = \sum_{n=0}^{7} s(g_n - g_c) 2^n \qquad (14)$$

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & otherwise \end{cases} \qquad (15)$$

where g_c is value of the central pixel (Xc, Yc) and g_n is value of nth pixel in the window.

3) Red-green-blue —RGB feature

This feature is calculated by equation (16):

$$h_{R,G,B}[r,g,b] = N * \Pr ob\{R = r, G = g, B = b\} \quad (16)$$

Or it can be computed in another way using three matrices h_R, h_G and h_B for each element of RGB.

4) Gradient feature

Gradient is used to differentiate various tiny teeth's parts such as enamel, cementum, gum, root canal, etc. [16]. The following steps calculate the Gradient value:

1) Reduce the background noises by applying Gaussian filter
2) Use Difference of Gaussian (DoG) filter to calculate gradient of the image
3) Get normalization form of the gradient vector and receive a 2D vector for each pixel as in equation (17).

$$\theta(z) = [\sin\alpha, \cos\alpha] \quad (17)$$

In wich α is direction of the gradient vector.

5) Patch level feature

Using this feature, we can calculate all gradient vectors of pixels in a patch P, denoted by $\delta(z)$ [3].

$$F_h(P) = \sum_{z \in P} \tilde{m}(z)\delta(z) \quad (18)$$

$$\tilde{m}(z) = \frac{m(z)}{\sqrt{\sum_{z \in P} m(z)^2 + \varepsilon_g}} \quad (19)$$

where $m(z)$ is the gradient amplitude at pixel z, $\tilde{m}(z)$ is the normalization value of m(z) and ε_g is a small constant. δ(z) is often specified by Hard Binning method as below.

$$\delta_i(z) = \begin{cases} 1, & \left\lfloor \dfrac{d\theta(z)}{2\pi} \right\rfloor = i - 1 \\ 0, & otherwise \end{cases} \quad (20)$$

C. Vector Quantization

Vector quantization (VQ) is a lossy data compression method, which met the challenge to the need for multi-dimensional integration. Linde and Gray (LBG) [17] proposed a VQ design algorithm based on training sequences. This bypasses the need for multi-dimensional integration. VQ is a classical quantization technique used in various fields such as image segmentation [1], [18]-[20], detection of cancer [21]. It is used to project a continuous input space on a discrete output space while minimizing the loss of information. It is also used to define zones in space that the set of points contained in each zone being project on a representative vector. Fig. 2 shows the diagram of VQ in which the term "codebook" is a document used for implementing a code. It contains a lookup table for coding and decoding.

In the new framework, K-means clustering algorithm is used for optimization of codebook [22], [23]. It is performed by the following steps.

1) Select k random vectors from the training set (namely code vectors).
2) Compute squared Euclidean distance of all the training vectors with the selected *k* vectors and *k* clusters are formed.
3) Put training vectors X_j in i^{th} cluster using the condition that squared Euclidean distance of X_j with i^{th} code vector is minimal.
4) Compute centroid for each cluster.
5) Compute Mean Squared Error (MSE) for each of k clusters.
6) Compute net MSE.
7) Repeat the above process till the net MSE converges.

Instead of random k-vector selection, we should generate codebook using LBG algorithm which makes the convergence faster.
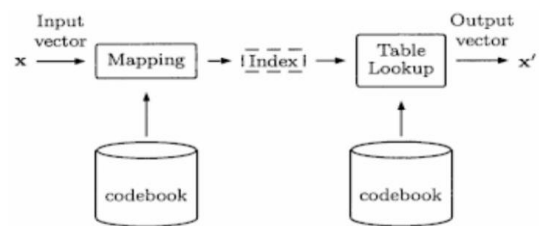

Fig. 2. VQ diagram.

D. Graph Representation

Each representation of vector quantization is set as a node of a complete graph whose weight of edge (*u*, *v*) is defined by $\|u - v\|$ (Fig. 3).
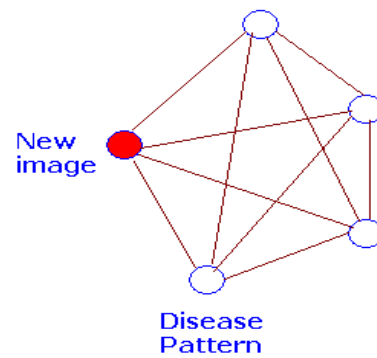

Fig. 3. Graph representation.

IV. EXPERIMENTAL EVALUATION

Three variants of the framework namely Prim spanning tree (GCP), Kruskal spanning tree (GCK), and Affinity Propagation Clustering (APC) were implemented on Matlab 2014 and executed on a PC VAIO laptop with Core i5 processor. The experimental results are the average values after 20 runs. The experimental dataset is from Hanoi Medical University, Vietnam including 13 dental images in the period 2014 – 2015.

In order to evaluate the performance of algorithms, MSE (Mean Squared Error), MAE (Mean Absolute Error) and Accuracy are used. If $\hat{Y}$ is a vector of *n* predictions, and Y is the vector of observed values corresponding to the inputs of the function which generated the predictions, then

MSE of the predictor can be estimated by [6]:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{Y}_i - Y_i\right)^2 = \frac{1}{n}\sum_{i=1}^{n}|e_i| \qquad (21)$$

The best value of MSE index is the smallest one. MAE is a quantity used to measure how close forecasts or predictions are to the eventual outcomes. The mean absolute error is given by [7]:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| = \frac{1}{n}\sum_{i=1}^{n}|e_i| \qquad (22)$$

where f_i is the prediction and y_i is the true value. Note that alternative formulations may include relative frequencies as weight factors. The best value of MAE index is also the smallest one.

The Accuracy index is the simplest measure of diagnostic decision. It is calculated by following formula [7].

$$Accuracy = \frac{a+d}{a+b+c+d} \qquad (23)$$

1) a is the result of diagnostic of the disease that is True Positive.
2) b is the result of diagnostic of the disease that is False Positive.
3) c is the result of diagnostic of the disease that is False Negative.
4) d is the result of diagnostic of the disease that is True Negative.

The larger the Accuracy value is, the better the algorithm is.

TABLE III: Evaluation Performance of All Methods in the Testing Dataset

	GCP	GCK	APC
MSE	2.083	2.084	0.821
MAE	0.917	0.916	0.701
Accuracy (%)	58.3	58.3	89.1

From the experimental dataset, we use the Hold-out method to divide it into the training (9 images) and testing sets (3 images) according to the diseases. The images in the training set are regarded as the disease patterns which are then used as exemplars for new dental images.

Table III shows the performance of all algorithms on the testing dataset with respect to the MSE, MAE and Accuracy values. The results have clearly shown that APC achieves better accuracy than GCP and GCK. Thus, a concluding remark of this paper is that APC should be used for dental X-Ray image search process in medical diagnosis.

V. Conclusions

In this paper, we focused on the dental X-Ray image search in medical diagnosis and proposed a novel framework using graph-based clustering for this problem. Details of this method have been elaborated in Section 3 including the working diagram, dental feature extraction, vector quantization, and a graph representation of the problem. Three variants of the framework namely Prim spanning tree (GCP), Kruskal spanning tree (GCK), and Affinity Propagation Clustering (APC) have been implemented and compared on the real dental dataset of Hanoi Medical University, Vietnam. The results have shown that APC is the best variant in term of accuracy.

This research suggests some further works such as: i) improve parameter selection in criterion 1 and 2 of graph-based clustering method in order to obtain the clusters via the number of diseases; ii) combine these methods using semi-supervised fuzzy clustering to give more exactly results.

References

[1] S. Chattopadhyay, R. M. Davis, D. D. Menezes, G. Singh, R. U. Acharya, and T. Tamura, "Application of bayesian classifier for the diagnosis of dental pain," *Journal of Medical Systems*, vol. 36, no. 3, pp. 1425-1439, 2012.

[2] M. S. Kavitha, A. Asano, A. Taguchi, T. Kurita, and M. Sanada, "Diagnosis of osteoporosis from dental panoramic radiographs using the support vector machine method in a computer-aided system," *BMC Medical Imaging*, vol. 12, no. 1, pp. 2-11, 2012.

[3] K. K. Oad, X. DeZhi, and P. K. Butt, "A fuzzy rule based approach to predict risk level of heart disease," *Global Journal of Computer Science and Technology*, vol. 14, no. 3, pp. 16-22, 2014.

[4] E. Ramírez, O. Castillo, and J. Soria, "Hybrid system for cardiac arrhythmia classification with fuzzy k-nearest neighbors and neural networks combined by a fuzzy inference system," *Soft Computing for Recognition Based on Biometrics*, pp. 37-55, 2010.

[5] D. Cai, X. He, Z. Li, W. Y. Ma, and J. R. Wen, "Hierarchical clustering of WWW image search results using visual, textual and link information," in *Proc. the 12th annual ACM international Conference on Multimedia*, 2004, pp. 952-959.

[6] H. Ding, J. Liu, and H. Lu, "Hierarchical clustering-based navigation of image search results," in *Proc. the 16th ACM International Conference on Multimedia*, 2008, pp. 741-744.

[7] Y. Hu, N. Yu, Z. Li, and M. Li, (2007). "Image search result clustering and re-ranking via partial grouping," in *Proc. IEEE International Conference on Multimedia and Expo*, 2007, pp. 603-606.

[8] H. B. Kekre, T. K. Sarode, and B. Raul, "Color image segmentation using Kekre's algorithm for vector quantization," *International Journal of Computer Science (IJCS)*, vol. 3, no. 4, pp. 287-292, 2008.

[9] Á. Vathy-Fogarassy and J. Abonyi, *Graph-Based Clustering and Data Visualization Algorithms*, 2013.

[10] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, pp. 972-976, 2007.

[11] J. Gao, Z. Ma, Y. Qin, L. Jia, and D. Yao, "Application of affinity propagation clustering algorithm in fault diagnosis of metro vehicle auxiliary inverter," in *Proc. the 2013 International Conference on Electrical and Information Technologies for Rail Transportation*, pp. 3-9, 2014.

[12] K. Zhang and X. Gu, "An affinity propagation clustering algorithm for mixed numeric and categorical datasets," *Mathematical Problems in Engineering*, 2014.

[13] U. Bodenhofer, A. Kothmeier, and S. Hochreiter, "APCluster: An R package for affinity propagation clustering," *Bioinformatics*, vol. 27, no. 17, pp. 2463-2464, 2011.

[14] Y. H. Lai and P. L. Lin, "Effective segmentation for dental X-ray images using texture-based fuzzy inference system," *Advanced Concepts for Intelligent Vision Systems*, pp. 936-947, 2008.

[15] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037-2041, 2006.

[16] K. H. Ghazali, M. M. Mustafa, A. Hussain, M. E. C. Bandar, and G. Kuantan, "Feature extraction technique using SIFT keypoints descriptors," *International Conference on Electrical and Engineering and Informatics Institut technology*, pp. 17-19, 2007.

[17] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Transactions on Communications*, vol. 28, no. 1, pp. 84-95, 1980.

[18] D. Chen, L. Li, and Z. Liang, "A self-adaptive vector quantization algorithm for MR image segmentation," in *Proc. International Society of Magnetic Resonance in Medicine*, 1999, vol. 1, p. 122.

[19] H. B. Kekre, S. Gharge, and T. K. Sarode, "SAR image segmentation using vector quantization technique on entropy images," *Arxiv Preprint arXiv*, 2010.

[20] H. Pujara and K. M. Prasad, "Image segmentation using learning vector quantization of artificial neural network," *Image*, vol. 2, no. 7, 2013.

[21] H. B. Kekre, T. Sarode, M. S. Gharge, and M. K. Raut, "Detection of cancer using vector quantization for segmentation," *International Journal of Computer Applications*, vol. 4, no. 9, 2010.

[22] H. B. Kekre and T. K. Sarode, "Vector quantized codebook optimization using K-means," *International Journal on Computer Science and Engineering,* vol. 1, pp. 283-290, 2009.

[23] B. Kövesi, J. M. Boucher, and S. Saoudi, "Stochastic K-means algorithm for vector quantization," *Pattern Recognition Letters*, vol. 22, no. 6, pp. 603-610, 2001.

Tran Manh Tuan received the bachelor on applied mathematics and informatics at Hanoi University of Science and Technology in 2003 and the master degree on computer science at Thainguyen University in 2007. Now, he is a researcher at Institute of Information Technology, Academic Institute of Science and Technology, Vietnam. He worked as a lecturer at Faculty of Information Technology, School of Information and Communication Technology, Thainguyen University from 2003.

Le Hoang Son obtained the PhD degree on mathematics – informatics at VNU University of Science, Vietnam National University (VNU). He has been working as a researcher and lecturer at the Center for High Performance Computing, VNU University of Science, Vietnam National University since 2007. His major field includes soft computing, fuzzy clustering, recommender systems, geographic information systems (GIS) and particle swarm optimization. He is a member of IACSIT, a member of Center for Applied Research in e-Health (eCARE), and also an associate editor of the International Journal of Engineering and Technology (IJET). Dr. Son served as a reviewer for various international journals and conferences such as PACIS 2010, ICMET 2011, ICCTD 2011, KSE 2013, BAFI 2014, NICS 2014 & 2015, ACIIDS 2015, ICNSC15, GIS-2015, FAIR 2015, International Journal of Computer and Electrical Engineering, Imaging Science Journal, International Journal of Intelligent Systems Technologies and Applications, IEEE Transactions on Fuzzy Systems, Expert Systems with Applications, International Journal of Electrical Power and Energy Systems, Neural Computing and Applications, International Journal of Fuzzy System Applications, Intelligent Data Analysis, Computer Methods and Programs in Biomedicine, World Journal of Modeling and Simulation, Knowledge-Based Systems, Engineering Applications of Artificial Intelligence. He gave a number of invited talks at national conferences namely FAIR 15'. Dr. Son has got 43 publications in prestigious journals and conferences including 15 SCI/SCIE papers and undertaken more than 20 major joint international and national research projects. He has published 2 books on mobile and GIS applications. So far, he has awarded the "2014 VNU Research Award for Young Scientists"