

# HIGH LEVEL SEARCH METHODS IN MEDICAL IMAGE DATABASE USING DEEP VISUAL-SEMANTIC QUANTIZATION

**Nasibeh Esmaeili,**

postgraduate Student of the National Research  
Lobachevsky State University of Nizhni Novgorod,  
Nizhny Novgorod, Russia,  
nasibe.smaeili@gmail.com

## ABSTRACT

**Problem:** Medical images contain useful information for extraction purpose and important in medicine as it can provide important information for diagnosis, monitoring treatment responses and disease management of patients with faster speed. By using the various features can classify images. The problem is feature selection and classification of images in order to retrieve the most relevant image in the fastest time. The problems of image retrieval are becoming widely recognized and the research on more efficient image retrieval systems are still in the development. Challenges of Image Retrieval System is: (a) how to mathematically describe an image which can also be called as feature extraction, and (b) how to assess similarity between a pair of images based on their abstracted descriptions that can also be called as matching. This paper presents a compact coding solution focusing on deep learning in quantization method that improves learning quality through compression coding. Also, the Deep Visual-Semantic Quantization method is used which is the first way to learn deep quantum models of tagged image data and semantic information that is text domains. In this paper, the current methods will be examined in order to retrieve medical image information, in particular deep learning methods. The results of the studies show that the deep learning methods in retrieval medical data yielded favorable results and could solve the challenges of this area very well.

**Keywords:** azimuth; an Autonomous orientation; high-precision orientation; operational orientation; satellite geodesy method higher; method of orientation; space navigation system.

**For citation:** Nasibeh E. High level search methods in medical image database using deep visual-semantic quantization. *I-methods*. 2017. Vol. 09. No. 3. Pp. 24-30.

## Introduction

Techniques and techniques of retrieval image have been investigated in order to satisfy the information needs of users' image content and to manage a large amount of image data [1-6]. There is already a large field research in this area, so that it is known in the medical field as a way of detecting many diseases. Since the presentation of all results goes beyond the scope of this paper, we only provide an overview of the active systems. Generally, CBIR [7] Should work under the two-step. The first one is a feature extraction where the image features are extracted so that they can be distinguished. The second step is to match the features from the first step with images from the database to get similar images. Although some of the existing CBIR systems have succeeded in medical field but there are still some limitations identified and still many challenges should be solved in order to develop an effective CBIRS for medical application. A content-based image retrieval system requires the measurement of visual features. This feature is measured by computer algorithms. The visual features are measured

and then utilized as indices within a database system so that retrieval can be carried out. The visual features are classified into two main areas:

- *Low-level primitives*
- *High-level semantic objects*

Low level or primitive features characterizing image content, such as color, texture, and shape that are automated extracted from images and used in content-based visual query. Higher level semantic features include tumors in medical images and everyday objects like animals, people and houses (in less domain specific image sets). Semantic objects themselves are composed of primitives. Consequently, any form of CBIRS needs the quantification of primitive, low-level visual features.

In this paper, focus on the listed systems and their development, as well as several new systems based on deep learning architecture. Section 2, Provides a related work of the issues involved in the data retrieval. Section 3 discusses the proposed method based on data retrieval, and Section 4 provides the perspective of future work.

### Related Work

Till now there was many researches has been done in the field of medical image retrieval system.

Now we are reviewing the methods that were done by other researchers previously [2, 5, 8-13]. In 2011 Herbert et al are focused on the modeling and development of a CBIR system. The types of images were CT and MRI and the method was metric data structure. The pre-processing initial step converts the image from RGB to gray scale, then adjust the intensity value of the pixels and finally we binarizes the image employing K-means with two channels. In this way, they eliminated the noisy existent in the regions of interest in the image [14]. For feature extraction, they use Border/Interior classification. This approach is proposed for the first time by Stehling [15]. Gabor filter is a two-dimensional Gaussian function modulated with sinusoidal orientations at a particular frequency and direction. They used this technique in order to extract texture information from an image.

In 2012 Panage et al are proposed the Image retrieval system based on dual tree complex wavelet transform (CWT) and support vector machines (SVM) [16]. At the first level, for both texture and color-based features in low level feature extraction, the dual tree complex wavelet transform is used. In order to extract semantic concepts at the second level, they made group medical images with the use of one against all support vector machines. For measuring the similarity between database features and query features are used here Euclidean distance. Also for comparison of SVM distances vectors they used a correlation-based distance metric. Murala uses local ternary co-occurrence patterns for MRI and CT images [17]. After in 2014 Kumaran used Gabor wavelet to extract texture features for MRI images. Also, he used the k-means clustering, progressive retrieval strategy and Euclidean distance are used to retrieve best MRI scan images for the query image in medical diagnosis.

Also, Kingsy in 2014 used apache hadoop framework method for medical images. In this work for efficient image retrieval is used texture based Content Based Image Retrieval algorithm [11]. Hadoop is open source grid architecture, complies with various image formats and can be established among various hospitals to store, share and retrieve images. Various performance metrics such as reliability, accuracy, interoperability, confidentiality and security are improved as with the use of Hadoop. The features that are used in many of the current medical image retrieval systems usually use hand-held features. This limitation may affect the function of image retrieval. To deal with this problem in [13] provide a simple and Indistinguishable feature as Histogram of compact dispersion coefficients of retrieval medical image (HCSCs).

In the proposed work, the dispersion transformation, specific changes to the deep convolution networks, is performed for the first time to produce an abstract experience of a medical image. Then, the predictive operation is performed to compress the dispersion coefficients obtained for processing. Finally, the bag-of-words histogram (BoW) is taken from the

compact dispersion coefficient as a medical image feature. In [18] A new hierarchical approach is described to retrieve the content-based images called Custom Queries (CQA). Unlike a single feature vector approach that tries to classify the query and retrieve similar images in one step, the CQA uses a set of features and a two-step recovery method. The first step is to classify the query according to the class label of the images using the features that best distinguish between classes. The second step, after retrieving the same images in the prediction class, uses the features to recognize the "subclass" in that class. The need to find a subset of custom features for each class force us to explore the feature selection for non-monitoring education. In [19] Provides a framework for image efficient and fast image retrieval for several PET-CT lung scan models. The method is presented in the following steps: Extraction of tissue properties, estimation of lung field, Feature classification, refinement using SVM and similarity measurements.

### Deep learning for Image retrieval

A major challenge in CBMIR systems is the gap between the low level of image information captured by imaging devices and the high level of human perceived semantic information. The effects of such systems are more in the representations of features, which can fully capture top-level information. In [5], we have proposed a deep learning framework for the CBMIR system using deep convolutional neural networks (CNNs) that has been trained to classify medical images. A data set consisting of 24 classes and 5 methods for network training is used. The learned features and classification results are used to retrieve medical images. Finally, in 2016 Chowdhury, Manish, et al proposed the new CBIR system for MRI images [20]. They used a Convolutional Neural Network (CNN) to obtain high level image representations that enable a coarse retrieval of images that are in correspondence to a query image. The retrieved set of images is refined via a non-parametric estimation of putative classes for the query image, which are used to filter out potential outliers in favor of more relevant images belonging to those classes.

### Convolutional Neural Networks

In machine learning, a convolutional neural netrk (CNN, or ConvNet) is a type of feed forward artificial neural network in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex. Individual cortical neurons respond to stimuli in a restricted region of space known as the receptive field. The receptive fields of different neurons partially overlap such that they tile the visual field. The response of an individual neuron to stimuli within its receptive field can be approximated mathematically by a convolution operation [21]. Convolutional networks were inspired by biological processes and are variations of multilayer perceptron's designed to use minimal amounts of preprocessing. LeNet-5 is the latest convolutional network designed for handwritten and machine-printed character recognition. They have wide applications in image and video recognition, recommender systems [22] and natural language

processing [23]. The convolutional neural network is also known as shift invariant or space invariant artificial neural network (SIANN), which is named based on its shared weights architecture and translation invariance characteristics [24].

**Proposed Method**

In retrieval system, is provided a training set of N points and a set of text labels Y which each point with a P-dimensional feature vector. It is displayed with a label set of text tag. The goal of deep learning in quantization is learning hybrid quantizer from the input to binary encoding by deep learning that encodes each of the points x to the compressed B bit binary code  $b = q(x)$ . So, supervising in training data can be maintained in compressed binary codes. This paper provides efficient image retrieval with DVSQ approach. In figure 1 is depicted a structure of deep learning that includes 1) standard conventional neural network (CNN) like AlexNet or GoogLeNet to learn the deep view for each image X and standard Skipgram like Word2Vec to learn embed word V for each tag of text  $y \in Y$  2) A perfectly connected transfer layer for moving some deep images {u} to semantic space labeled with embed labels {v}. 3) An adaptive margin to keep multilayer learning from image viewer {u} and embed labels {v} 4) A new model for semantic image measurement to convert image index to B bit binary codes by minimizing the quantization error from the approximation of the internal multiplication search. Each step is briefly described.

**Deep visual semantic embedding**

Architecture DVSQ includes visual models and a text model. We use alexnet as visual model which is made multilayer convolution filter and MaxPooling. Also is used word2vec as a text model in order to matching. In order to simplification fast learning, is utilized alexnet model that before trained in ImageNet dataset [25] in 2012 and word2vec model that trained before in Google News [26]. The purpose of DVSQ is maintaining the power of influence of semantic knowledge learned in the text. The pre-training example of comparative ranking margin of cost function is described as follows:

$$L_n = \sum_{i \in y_n} \sum_{j \neq y_n} \max \left( 0, \delta_{ij} - \frac{v_i^T z_n}{\|v_i\| \|z_n\|} + \frac{v_j^T z_n}{\|v_j\| \|z_n\|} \right) \quad (1)$$

That is an embedded image for the image and is the embedded word of correct text tag while is the embedded word of wrong text tag of image. Note that specifies a consistent margin that is a key component designed in order to ensure that the internal correlation between embedding and embedding the correct text word must be larger than the paradoxical marginal text with a comparative margin. A consistent margin with the internal factor between the word correct text tag and the wrong text tag is described as follows:

$$\delta_{ij} = 1 - \frac{v_i^T z_n}{\|v_i\| \|z_n\|} \quad (2)$$

**Choosing the paradoxical marginal is important.**

The quantization of visual semantic of inner product

While visual semantic embedding facilitates image retrieval, efficient image retrieval is possible with a qualitative quantization model of visual semantic of inner product. The key insight is to choose the embedding words in the tag of the set Y as a training query set. The reason is that all images are divided into the semantic space by the word embedded in the tag of the Y set, and as a result, this embedding word can be used as modeling the distribution of the basic query. Since Maximum Inner Product Search (MIPS) is widely used in real-world retrieval systems, it's activated by formulating a quantization of inner product of visual semantic. With the error of quantization, the pre-trained example is defined as follows:

$$Q_n = \sum_{i=1}^{|y|} \left( v_i^T z_n - v_i^T \sum_m^M C_m b_{mn} \right)^2 \quad (3)$$

$$Q_n = \sum_{i=1}^{|y|} \left( v_i^T z_n - v_i^T \sum_m^M C_m b_{mn} \right)^2 \quad (4)$$

**Deep visual semantic quantization**

This paper provides the ability to retrieve efficient image in an end-to-end architecture. The DVSQ model is learned by integrating the deep visual-semantic embedding model (1) and the visual semantic inner product quantization model. The problem of common optimization is as follows:

Where  $\lambda > 0$ , between the cost coefficient of the margin L and Q, the quantum cost of the multiplication, and W represent the network parameter set.

**Learning Algorithm**

The DVSQ optimization problem (4) has three sets of variables, network parameters W, M codebooks, and N binary codes. The alternating optimization paradigm is adopted [8] in order to repetitive update a variable and is fixed the rest variables. The network parameter W can be optimized with the help of standard back propagation algorithm via automatic differential techniques in TensorFlow[1]. In order to learn C, is updated codebooks C via correction W and B. For learning B, is used below formulation. Each is independent in

As can be seen this optimization problem is NP-Hard and it is high order Markov random. I can be solved by iterated conditional modes algorithm (ICM).

Figure 1: DVSQ for efficient image retrieval, which consists of four key elements : (1) the standard convolutional neural network (CNN), (2) a completely connected transform layer for converting deep image index {u} into semantic space (3) A comparative margin and (4) a visual-semantic quantization model.

**Implementation and result**

In this paper, the TensorFlow Deep Learning Tool is used in order to develop and train the proposed deep learning framework. The implementation is based on the Asus k550z laptop with Ubuntu 16 with 2.50 GHz AMD A10 processor with 6.00 GB RAM. The proposed method has been evaluated for classification and retrieval. In this paper, we use the data set used for the proposed CBMIR task from public medical databases. The classes are based on the body organ, such as the lungs, the brain, the liver, and so on. Totally in this research 24 classes are used in the data set. Overall 300 images from each class are taken that randomly come to a collection of 7200 images. Data from each class was randomly assigned to training and test sets. Using 70% and 30% of the images for training and a set of tests. In total, 5040 and 2160 images respectively have been used in training and testing. All the images in the DICOM format, except images from Messidor, were included in TIF images (tagged image file). All images from each class were converted to  $256 \times 256$ , and color images of the eye class turned gray. Numeric class tags are assigned to classes for supervised learning. In this paper, only evaluated the classification and used various criteria that exist like average

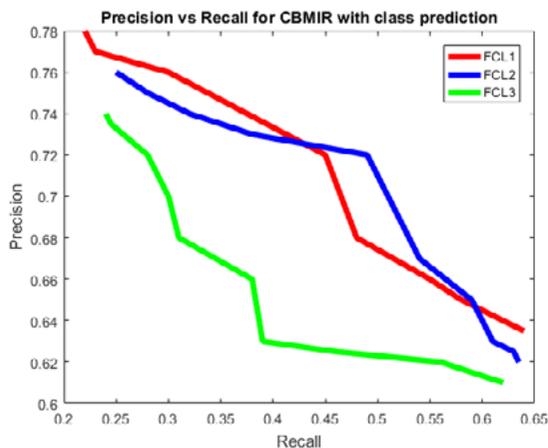
precision (AP), average recall (AR), accuracy and F1 measure. can be significantly reduced by predicting the wrong class, but high accuracy means that this event is very unfavorable.

The classification performance is compared with the classification method presented in reference [27], which is presented in Table 1 for complete comparisons. The proposed system performs better when using polynomial data in classification organs. Although the system trained in [27] is a set of different sets of images, high accuracy in our proposed system, shows the effectiveness of the method in classification work. The better performance in classification of our proposed method encourages us to rely on the idea of class-based retrieval, while class prediction is used to reduce the search area during retrieval. For a retrieval-based class, the retrieval performance

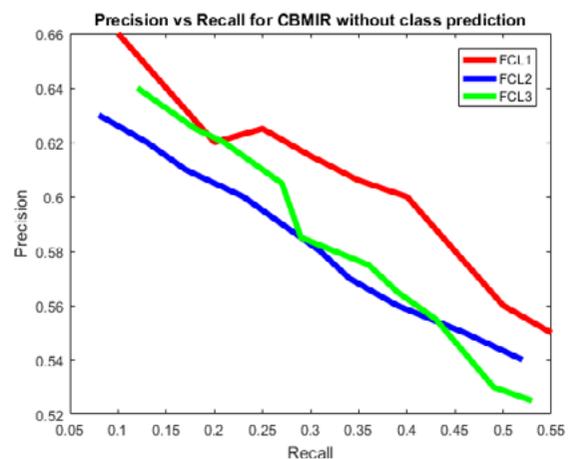
The proposed method for the CBMIR has been tested by Precision and Recall factors which are the performance criterion for the CBIR system. A feature view of the fully-integrated three layers of the trained model is used to retrieve medical images. The analysis of this feature is done in terms of quality of retrieval using both options, using and without using prediction class labels. The recall accuracy diagram is shown

Table for Proposed Method

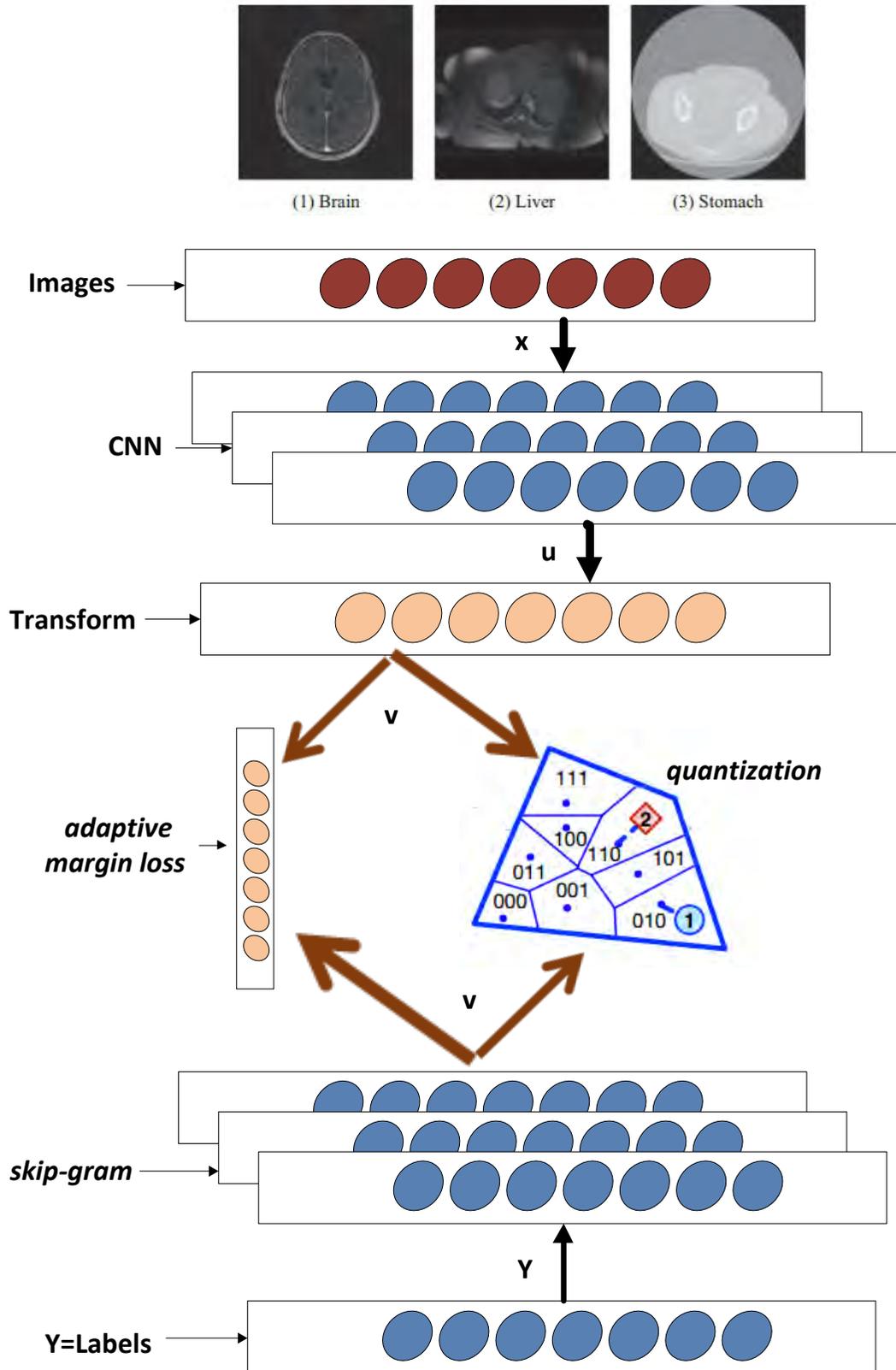
Model	Num. Of images	Training (images)	Testing (Images)	Classes	AP(%)	AR(%)	F1 Measure
Proposed Method	7489	2413	4043	24	99.32	99.3	99.2
Ref [27]	7489	2413	4043	12	98.43	97.28	97.85



**Fig. 2.** Precision and recall for CBMIR with class prediction



**Fig. 3.** Precision and recall for CBMIR without class prediction



**Fig. 1.** :DVSQ for efficient image retrieval, which consists of four key elements : (1) the standard convolutional neural network (CNN), (2) a completely connected transform layer for converting deep image index {u} into semantic space (3) A comparative margin and (4) a visual-semantic quantization model.

in Figures 2 and 3 below to show the feature extracted from FCL1, FCL2, and FCL3 by using class prediction and without using class prediction. The comparison between feature of FCL1 and FCL2 and FCL3 indicate the high accuracy in FCL1. Performance improvement in precision conditions is also evident in using class predictions.

### Conclusion

Nowadays, with the rapid development of computer technology, large number of images which can carry a considerable amount of information stored in computers. Images must be properly stored to be retrieved on demand so Management of this image data systematically is very important for future applications, such as research and patient treatment

In this paper, a deep learning approach is provided for data retrieval. This paper focused on deep learning for the quantization of tagged image data and semantic information extracted from text domains. The DVSQ model that is used in this paper can be learn compressed binary codes by optimizing the relative marginal cost function and the visual-semantic quantization cost function on a deep hybrid network. DVSQ provides better and more efficient image retrieval and better results. It evaluates after a few years of developments the need for image retrieval and presents concrete scenarios for promising future research directions.

### References

- Mukherjea S., Hirata K., Hara Y. Amore: a world-wide web image retrieval engine. In *CHI'99 Extended Abstracts on Human Factors in Computing Systems*. 1999. Pp. 17-18.
- Pilevar A. H. CBMIR: Content-based image retrieval algorithm for medical image databases. *Journal of medical signals and sensors*. 2011. Vol. 1. P. 12.
- Radenović F., Toliaş G., Chum O. CNN image retrieval learns from BoW: Unsupervised fine-tuning with hard examples. *arXiv preprint arXiv:1604.02426*, 2016.
- Lohar N., Chavan D., Arade S., Jadhav A., Chikmurge D. Content Based Image Retrieval System over Hadoop Using MapReduce. 2016.
- Qayyum A., Anwar S.M., Awais M., Majid M. Medical image retrieval using deep convolutional neural network. *Neurocomputing*. 2017.
- Ma L., Liu X., Gao Y., Zhao Y., Zhao X., Zhou C. A new method of content based medical image retrieval and its applications to CT imaging sign retrieval. *Journal of biomedical informatics*. 2017. Vol. 66. Pp. 148-158.
- Jin H. Application of CBIR Technology in Digital Images of Museum Collection. 2015.
- Lehmann T. M., Güld M. O., Deselaers T., Keysers D., Schubert H., Spitzer K. Automatic categorization of medical images for content-based retrieval and data mining *Computerized Medical Imaging and Graphics*. 2005. Vol. 29. Pp. 143-155.
- Malviya N., Choudhary N., Jain K. Content Based Medical Image Retrieval and Clustering Based Segmentation to Diagnose Lung Cancer. *Advances in Computational Sciences and Technology*. 2017. Vol. 10. Pp. 1577-1594.
- Bedo M., Pereira dos Santos D., Ponciano-Silva M., de Azevedo-Marques P. M., Ferreira de Carvalho A., Traina Jr. C. Endowing a Content-Based Medical Image Retrieval System with Perceptual Similarity Using Ensemble Strategy. *Journal of digital imaging*. 2016. Vol. 29. Pp. 22-37.
- Grace R. K., Manimegalai R., Kumar S. S. Medical image retrieval system in grid using Hadoop framework. *In Computational Science and Computational Intelligence (CSCI), 2014 International Conference on*, 2014. Pp. 144-148.
- Nowaková J., Přílepok M., Snášel V. Medical Image Retrieval Using Vector Quantization and Fuzzy S-tree," *Journal of medical systems*. 2017. Vol. 41. Pp. 18-18.
- Lan R., Zhou Y. Medical Image Retrieval via Histogram of Compressed Scattering Coefficients," *IEEE Journal of Biomedical and Health Informatics*. 2017. Vol. PP. Pp. 1-1.
- Chuctaya H., Portugal C., Beltrán C., Gutiérrez J., López C., Túpac Y. M-CBIR: A medical content-based image retrieval system using metric data-structures. *In Computer Science Society (SCCC), 2011 30th International Conference of the Chilean*. 2011. Pp. 135-141.
- Stehling R. O., Nascimento M. A., Falcão A. X. A compact and efficient image retrieval approach based on border/interior pixel classification," in *Proceedings of the eleventh international conference on Information and knowledge management*. 2002. Pp. 102-109.
- Pange S., Lokhande S. Image retrieval system by using CWT and support vector machines," *Signal & Image Processing*. 2012. Vol. 3. 63p.
- Kumaran N., Bhavani R. TEXTURE CONTENT BASED MRI IMAGE RETRIEVAL USING GABOR WAVELET AND PROGRESSIVE RETRIEVAL STRATEGY. *Journal of Theoretical & Applied Information Technology*. 2014. Vol. 63.
- Dy J. G., Brodley C. E., Kak A., Broderick L. S., Aisen A. M. Unsupervised feature selection applied to content-based retrieval of lung images," *IEEE transactions on pattern analysis and machine intelligence*. 2003. Vol. 25. Pp. 373-378.
- Song Y., Cai W., Eberl S., Fulham M. J., Feng D. A content-based image retrieval framework for multi-modality lung images," in *Computer-Based Medical Systems (CBMS), 2010 IEEE 23rd International Symposium on*. 2010. Pp. 285-290.
- Chowdhury M., Bulò S. R., Moreno R., Kundu M. K., Smedby Ö. An efficient radiographic Image Retrieval system using Convolutional Neural Network," in *Pattern Recognition (ICPR), 2016 23rd International Conference on*. 2016. Pp. 3134-3139.
- Matsugu M., Mori K., Mitari Y., Kaneda Y. Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Networks*. 2003. Vol. 16. Pp. 555-559.
- Van den Oord A., Dieleman S., Schrauwen B. Deep content-based music recommendation. In *Advances in neural information processing systems*. 2013. Pp. 2643-2651.
- Collobert R., Weston J. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference*

on Machine learning. 2008. Pp. 160-167.

24. Zhang W., Itoh K., Tanida J., Ichioka Y. Parallel distributed processing model with local space-invariant interconnections and its optical architecture. *Applied optics*. 1990. Vol. 29. Pp. 4790-4797.

25. Donahue J., Jia Y., Vinyals O., Hoffman J., Zhang N., Tzeng E. Decaf: A deep convolutional activation feature for generic visual recognition," in *International conference on*

*machine learning*, 2014. Pp. 647-655.

26. Mikolov T., Chen K., Corrado G., Dean J. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

27. Yan Z., Zhan Y., Peng Z., Liao S., Shinagawa Y., Zhang S. Multi-instance deep learning: Discover discriminative local anatomies for bodypart recognition," *IEEE transactions on medical imaging*. 2016. Vol. 35. Pp. 1332-1343.

## ВЫСОКОУРОВНЕВЫЕ МЕТОДЫ ПОИСКА В БАЗЕ ДАННЫХ МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ГЛУБОКОГО ВИЗУАЛЬНО-СЕМАНТИЧЕСКОГО КВАНТОВАНИЯ

**Насибех Есмаели,**

аспирант Нижегородского государственного университета имени Н. И. Лобачевского, г. Нижний Новгород, Россия, nasibe.smaeili@gmail.com

### АННОТАЦИЯ

Медицинские изображения содержат полезную информацию для целей извлечения и важны в медицине, поскольку она может предоставить важную информацию для диагностики, мониторинга ответов на лечение и лечения заболеваний пациентов с более высокой скоростью. Используя различные функции, вы можете классифицировать изображения. Проблема заключается в выборе функций и классификации изображений для получения наиболее релевантного изображения в самое быстрое время. Проблемы получения изображений становятся все более популярными, и исследования по более эффективным системам поиска изображений все еще находятся в разработке. Вызовы системы поиска изображений: (а) как математически описывать изображение, которое также можно назвать извлечением признаков, и (б) как оценить сходство между двумя изображениями на основе их абстрактных описаний, которые также можно назвать совпадением. В этом документе представлено компактное кодирование, основное внимание уделяется глубокому обучению методу квантования, который улучшает качество обучения с помощью кодирования сжатием. Кроме того, используется метод Deep Visual-Semantic Quantization, который является первым способом изучения глубоких квантовых моделей помеченных данных изображения и семантической информации, которая является текстовыми областями. В этой статье будут рассмотрены текущие методы, чтобы получить информацию медицинского изображения, в частности методы глубокого обучения. Результаты исследований показывают, что методы глубокого обучения в поисках медицинских данных дали благоприятные результаты и могли очень хорошо решить проблемы этой области.

### Сведения об авторе:

Насибех Е., аспирант Нижегородского государственного университета имени Н. И. Лобачевского,

**Для цитирования:** Насибех Е. Высокоуровневые методы поиска в базе данных медицинских изображений с использованием глубокого визуально-семантического квантования // I-methods.2017. Т. 09. №. 3. С. 24-30.