International Journal of Engineering

Journal Homepage: www.ije.ir

A Radon-based Convolutional Neural Network for Medical Image Retrieval

A. Khatami*[a], M. Babaie[b], H. R. Tizhoosh[b], A. Nazari[c], A. Khosravi[a], S. Nahavandi[a]

[a] *Institute for Intelligent System Research and Innovation, Deakin University, Australia*
[b] *KIMIA Lab., University of Waterloo, Ontario, Canada*
[c] *School of Information Technology, Deakin University, Australia*

PAPER INFO

Paper history:
Received 22 August 2017
Received in revised form 24 Januray 2018
Accepted 26 May 2018

Keywords:
Deep Convolutional Neural Network
Image Retrieval in Medical Application
Medical Image Retrieval
Radon Transformation

ABSTRACT

Image classification and retrieval systems have gained more attention because of easier access to high-tech medical imaging. However, the lack of availability of large-scaled balanced labelled data in medicine is still a challenge. Simplicity, practicality, efficiency, and effectiveness are the main targets in medical domain. To achieve these goals, Radon transformation, which is a well-known technology in medical field, is utilized along with a deep network to propose a retrieval system for a highly imbalanced medical benchmark. The main contribution of this study is to propose a deep model which is trained on the Radon-based transformed input data. The experimental results show that applying this transformation as input to feed into a convolutional neural network, significantly increases the performance, compared with other retrieval systems. The proposed scheme clearly increases the retrieval performance, compared with almost all models which use Radon transformation to retrieve medical images.

doi: *10.5829/ije.2018.31.06c.07*

1. INTRODUCTION

Medical imaging is one of the essential components for modern healthcare. The extensive use of radiology images, which is a type of medical imaging, has triggered the creation of several medical image benchmarks. Accessing medical datasets motivates scientists to develop methods to extract proper features and attributes from data, resulting in accurate retrieval and classification systems. These tools play a crucial role in assisting clinicians in diagnosis. Proposing a content-based system to retrieve medical images needs an accurate feature classification. Numerous studies have been conducted on classifying content-based features of X-ray images, resulting in different retrieval systems [1, 2]. Methods such as Local Binary Pattern (LBP) and Histogram of Gradient (HOG) as well as feature extractions techniques such as SIFT, SURF, and ORB have been recently utilized in almost every filed [3-10]; but they may fail to extract robust features in medical imaging [11, 12].

Image Retrieval in Medical Application (IRMA) benchmark is widely used dataset, aiming at classifying and retrieving different parts of body, created by University Hospital of Aachen, Germany [13, 14]. Figure 1 depicts some samples of the benchmark. IRMA dataset includes 12,677 radiology images for training set and 1,733 images for testing set. All images have been divided into 193 categorize/classes. Many studies have been conducted on this dataset.

Radon transformation which captures the parallel projections of images from different directions is well-established in medical domain [15-18]. This transformation is widely used because it is easy to implement and also efficient in matching. Moreover, it enables an efficient retrieval model because it requires less memory and storage. In other words, it provides, in contrast to some of above-mentioned descriptors, a short-length feature vector for further analysis [19]. As examples, Tizhoosh [19] introduced Radon barcode annotations for tagging medical data, resulted in a

*Corresponding Author Email: amin.khatami@deakin.edu.au (A. Khatami)

Please cite this article as: A. Khatami, M. Babaie, H. R. Tizhoosh, A. Nazari, A. Khosravi, S. Nahavandi, A Radon-based Convolutional Neural Network for Medical Image Retrieval, International Journal of Engineering (IJE), IJE TRANSACTIONS C: Aspects Vol. 31, No. 6, (June 2018) 910-915

retrieval system on IRMA and achieved the error of 470.57. In another experiment [20], a small number of equidistant projections of Radon was examined to create a retrieval system.

There are three main findings of this study; (1) applying Radon to original data achieves better performance, compared with utilizing the Radon as a descriptor at the top of a retrieval system; (2) proposing a robust and accurate classification technique; (3) in this specific version of IRMA dataset, which are considered 193 categories, a classification technique can work as a retrieval system, because all images in each category have the same IRMA codes, and a simple search-based scheme can easily retrieve the selected images.

This study is organized as follows: The technical routines are described in section II. The experimental results as well as discussion are reported in Section III, followed by a conclusion in Section IV.

2. METHODOLOGY

Because a deep network is utilized to propose the retrieval system, a large dataset for training is required. Hence, augmentation techniques are utilized to enlarge the training data. Radon transform and a deep CNN are utilized to define the model.

2. 1. Radon Radon transform is an integral conversion, calculating the summation of values of an image from different angles along parallel lines:

$$R(\rho,\theta)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\delta(x\cos\theta+y\sin\theta-\rho)dxdy \quad (1)$$

In this equation, $f(x,y)$ refers to grey-level intensities of image F at position (x,y) and $\delta(.)$ is a Dirac delta operator. Utilizing the Radon transformation is the main benefit of this work. Radon is a well-known method in medical domain, as an image descriptor/feature. This makes the approach more tuned toward medial images as we hope. An important characteristic of Radon is that it represents images with low-dimensional vectors. Figure 2 shows three Radon projections of a matrix, representing an image, from three different angles.

Figure 3 illustrates the reconstructed image by utilizing inverse Radon. It is clear that the reconstructed image has the information and details of the original image; however it loses unnecessary intensities over widely homogenous regions.

Deep *Convolutional Neural Networks (CNNs)* have been inspired from Multilayer Perceptrons (MLPs). A deep CNN includes convolutional blocks which are consisting of convolutional and pooling layers, followed by at least one fully-connected layer.

At the top of this network, a classification layer is inserted to classify the input features.

Figure 1. Some samples of IRMA dataset

The convolutional blocks convert the extracted features at a higher resolution of an image to more complicated information at a coarser level. Each block has different convolutional kernel which is a single matrix showing how to convolute the input data with convolution operations. Feature maps, which are the outputs of convolution blocks, contain the convoluted information of the previous layer.

2. 2. IRMA Dataset The IRMA dataset includes a variety of radiology images which have been randomly selected from radiology routine work at the Department of Diagnostic Radiology, Aachen University of Technology (RWTH), Aachen, Germany [13, 14].

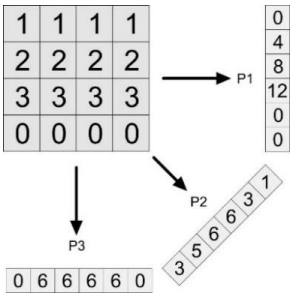

Figure 2. Radon projections in three angles

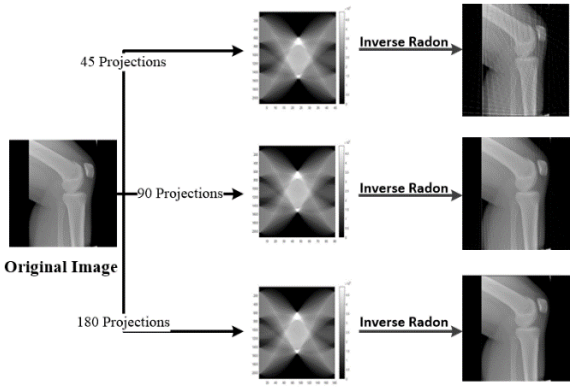

Figure 3. A three-step diagram, applying Radon on an original image

The x-ray images represent various cases with respect to patients' age and gender, viewing positions, and pathologies. All radiology images have been rescaled to a zero-padded 512 × 512 bounding box to make the processing easier. There is a total of 193 image categories. The training set includes 12,677 radiographs with known categories. The test set comprises 1,733 radiographs.

The complete IRMA code shows a string code of 13 characters; TTTT-DDD-AAA-BBB. This string constitutes four mono-hierarchical axes: the technical code T (imaging modality), directional code D (body orientations), anatomical code A (the body region), and biological code B (the biological system examined).

Figure 4 depicts two sample images along with their IRMA code. We utilized the Python code described in literature [13-14], provided by Image CLEFmed09 to evaluate the errors. The total IRMA error can be computed by

$$\sum_{i=1}^{I} \frac{1}{b_i} \frac{1}{i} \delta(l_j, \hat{l}_i) \quad (2)$$

With

$$\delta(l_j, \hat{l}_i) = \begin{cases} 0 & if\ l_j = \hat{l}_j \quad \forall j \le i \\ 0.5 & if\ l_j = * \quad \exists j \le i \\ 1 & if\ l_j \neq \hat{l}_j \quad \exists j \le i \end{cases} \quad (3)$$

where b is the number of possible labels at position i, and δ is a decision measure carrying out 1 for incorrect label and 0 for correct one, when the image Ii is compared with the image $\hat{l}_i$.

3. EMPERICAL RESULTS

IRMA dataset is conducted in this section to evaluate our technique versus the others. Radon transform converts an image to a one-dimensional profile. We use Radon transform to convert the raw radiology images to one-dimensional signals for each direction. Then, the transformed images are fed into a deep-structural feature extraction system to find and classify the most relevant features of each part of body. This strategy differentiates our approach from other techniques using Radon to propose a retrieval system on IRMA.

As depicted in Figures 5 and 6, we follow the illustrated stages to create the retrieval system:

Step 1: a preprocessing phase is performed at the first stage of the model. First, all images are resized to square sizes of 90 × 90 by zero-padding. We also apply the procedure, discussed in literature [21] to remove landmarks and burnt-in annotations of digitalized radiology films. Practically, we need to enlarge the training set because of no availability of big labelled image datasets in medical imaging. Different transformations; flipping, rotating, and scaling, reported in literature [22], are applied to the training set.

Step 2: we use Radon transform to convert the preprocessed x-ray images to one-dimensional signals for each direction. A total of 90 equidistant Radon projections are computed for all images of the t. As depicted in Figures 5 and 6, we followed the illustrated stages to create the retrieval system:

Step 1: A preprocessing phase is performed at the first stage of the model. First, all images are resized to square sizes of 90 × 90 by zero-padding.

We also applied the procedure, discussed in reference [21] to remove landmarks and burnt-in annotations of digitalized radiology films. In addition, because deep networks need many training data for learning, augmentation techniques are utilized to enlarge the training set of the dataset.

Practically, we preprocessed and augmented dataset. These 90 projections are derived from 90 angles, belonging to $\theta_i = \{0, \pi/n, 2\pi/n, \ldots, n\pi/n\}$. Note that the output of this step is the images with a size of 90 × 90 pixels. *Step 3*: accurate setting of parameters of a deep network forces researchers to design and propose different networks, corresponding to a specific problem.

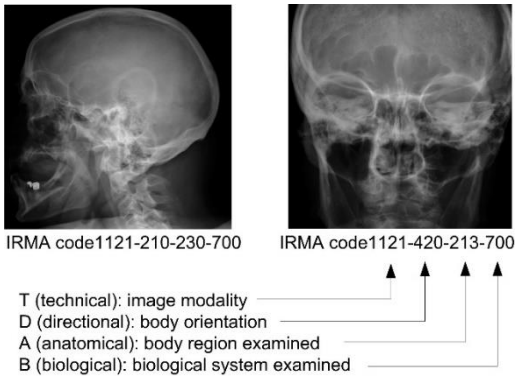

Figure 4. A description of two 13-digit IRMA codes

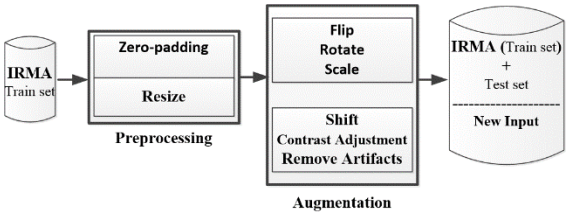

Figure 5. The preprocessing and augmentation steps of the proposed model

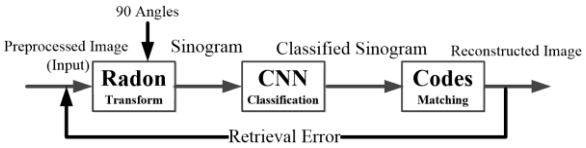

Figure 6. The block diagram of the proposed model

Different sizes of input data are investigated to feed to the deep CNN. As investigated in literature [23], different experiments were conducted based on the network architecture introduced in AlexNet [22], and LeNet [24]. The Theano framework [25] has been utilized for implementation. Different kernel sizes were investigated for convolutional and pooling layers. As reported in literature [23, 26], the size of 3 × 3 was finally selected for this study. It was also found that LeNet performed better on the IRMA data set.

Table 1 presents in more detail the CNN structure used for the transfer learning scenario. It follows the layer pattern of LeNet, Conv-Pool, which has been implemented for MNIST [24].

Note that the AlexNet has a very similar architecture to the LeNet, but it is deeper and larger. Moreover, the features of convolutional layers are stacked on top of each other. However, in the LeNet, each convolutional layer always immediately is followed by a pooling layer. After some investigations, the data, obtained by previous step, are resized to 58×58, feeding to the CNN with three convolutional blocks along with one fully connected and one classifying layer, as explained in Table 1. This procedure categorizes the data into 193 classes. Therefore, classified sinograms are the outputs of the CNN. Code matching step picks an image up from its predicted category, with respect to the test query, and calculates the retrieval IRMA error.

As discussed below, the same investigation is performed to find the best rate for hyper-parameters of the regularization techniques. The following investigation is examined to find the best model, regarding each regularization technique.

As for Drop-Out, the setting from 0 to 60 percentages is considered for the rate of dropping the neurons in the convolution and fully-connected layers.

TABLE 1. The deep CNN details

Name	Filter size	Filter dimension	Stride
Conv1	3	16	1
ReLU1	1		1
Pooling1	2		2
Conv2	3	32	1
ReLU2	1		1
Pooling2	2		2
Conv3	3	64	1
ReLU3	1		1
Pooling3	2		2
Conv4	3	128	1
ReLU4	1		1
Pooling4	2		2
FC	1	625	1

TABLE 2. Comparison the results reported on IRMA dataset. The methods with * were reported in [26]. In gray highlighted results are all achieved with Radon-based methods

Methods	IRMA Error
LBP+Saliency+SVM [23]	146.55
TAUbiomed *	169.50
diap*	178.93
Our Proposed Method	**248.7**
SuperPixels*	249.34
CNN(Radon, no binarization) [27]	270.12
SVM-Radon Barcode [28]	294.83
SP-Radon [21]	311.8
MedGIFT *	317.53
VPA*	320.61
Auto Encoder and Radon [29]	344.08
SP-Radon Barcodes (Binaries) [21]	356.57
IRMA*	359.29
Auto Encoder Radon Barcode [30]	392.09
MedGIFT*	420.91
Radon Barcodes [19]	559.46

In terms of depth analysis, different kernel sizes of 2×2, 3×3, and 4×4 were tested. AlexNet architecture of Conv-Conv-Pool and LeNet structure of Conv-Pool were considered as well.

By setting the kernel size to 3×3, the test prediction rate of 78% (median among five folds, as discussed) is achieved. The investigation shows that if the kernel size is reduced to 2×2, a decrease of 1.78% is noticed in the performance of the test prediction rate, and if it is set to 4×4, a drop of 1% is observed.

Step 4: image retrieval is the main aim of this step. Suppose a test query is selected. This query goes through the proposed chain of processing. Figure 6 depicts, 90 equidistant Radon projections of the computed samples. Then the converted datum is fed to the CNN to obtain the most relevant category. Each IRMA code of the selected category can be chosen as the retrieved code of the original test query. Therefore, the Python code, provided by *ImageCLIFmed09* is used to calculate the error between the original and the retrieved image. We achieved an IRMA error of 248.7 by utilizing the same scenario to all test queries. This outcome significantly surpasses that of the methods that use Radon projections on IRMA dataset.

Table 1 shows the comparison between our proposed method and the other techniques. More specifically, as for the method proposed in reference [19], the error of 559.46 was obtained by proposing Radon-based barcode on IRMA. Liu et al. in [27] conducted another

investigation on IRMA dataset using Radon projections. They applied Radon on the features extracted by a deep CNN and achieved 270.12 IRMA error once no binarization was performed on the features. Sze et al. [29] applied the Radon on the dense features extracted by a deep autoencoders, and obtained an error of 344.08. Accordingly, it is clear that our contribution which applies the Radon transform on raw data, followed by deep extraction for retrieval system, achieves better performance.

4. CONCLUSION

Radon transform has been widely used in medical domain, especially in CT imaging and for image reconstruction. Radon projections can be utilized as features to retrieve images. In contrast to other methods that use Radon projections as a descriptor, we used these features, extracted by Radon transform, as input for a deep network, followed by a procedure to retrieve the selected image. The retrieved image was selected by defining the error function, calculating the weighted distance between the original codes and the retrieved ones. Our experiments showed the superiority of our model compared with other Radon-based techniques. In the future, a further investigation on feeding different types of features to the deep network will be taken into account.

6. REFERENCES

1. Müller, H., Michoux, N., Bandon, D. and Geissbuhler, A., "A review of content-based image retrieval systems in medical applications—clinical benefits and future directions", ***International Journal of Medical Informatics***, Vol. 73, No. 1, (2004), 1-23.
2. Müller, H., Deselaers, T., Deserno, T., Clough, P., Kim, E. and Hersh, W., "Overview of the imageclefmed 2006 medical retrieval and medical annotation tasks", in Workshop of the Cross-Language Evaluation Forum for European Languages, Springer., (2006), 595-608.
3. Ojala, T., Pietikainen, M. and Maenpaa, T., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", ***IEEE Transactions on Pattern Analysis and Machine Intelligence***, Vol. 24, No. 7, (2002), 971-987.
4. Ahonen, T., Hadid, A. and Pietikainen, M., "Face description with local binary patterns: Application to face recognition", ***IEEE Transactions on Pattern Analysis and Machine Intelligence***, Vol. 28, No. 12, (2006), 2037-2041.
5. Dalal, N. and Triggs, B., "Histograms of oriented gradients for human detection", in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, IEEE. Vol. 1, (2005), 886-893.
6. Khatami, A., Mirghasemi, S., Khosravi, A., Lim, C.P. and Nahavandi, S., "A new pso-based approach to fire flame detection using k-medoids clustering", ***Expert Systems with Applications***, Vol. 68, No., (2017), 69-80.
7. Khatami, A., Mirghasemi, S., Khosravi, A. and Nahavandi, S., "A new color space based on k-medoids clustering for fire detection", in Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on, IEEE., (2015), 2755-2760.
8. Khatami, A., Mirghasemi, S., Khosravi, A. and Nahavandi, S., "An efficient hybrid algorithm for fire flame detection", in Neural Networks (IJCNN), 2015 International Joint Conference on, IEEE., (2015), 1-6.
9. Keyvanpour, M., Tavoli, R. and Mozafari, S., "Document image retrieval based on keyword spotting using relevance feedback", ***International Journal of Engineering-Transactions A: Basics***, Vol. 27, No. 1, (2013), 7-14.
10. Ezoji, M. and Iravani, S., "A general framework for 1-d histogram-baesd image contrast enhancement", ***International Journal of Engineering-Transactions A: Basics***, Vol. 29, No. 10, (2016), 1384-1392.
11. Kashif, M., Deserno, T.M., Haak, D. and Jonas, S., "Feature description with sift, surf, brief, brisk, or freak? A general question answered for bone age assessment", ***Computers in Biology and Medicine***, Vol. 68, (2016), 67-75.
12. Sargent, D., Chen, C.-I., Tsai, C.-M., Wang, Y.-F. and Koppel, D., "Feature detector and descriptor for medical images", in SPIE Medical Imaging, International Society for Optics and Photonics., (2009), 72592Z-72592Z-72598.
13. Lehmann, T.M., Gold, M., Thies, C., Fischer, B., Spitzer, K., Keysers, D., Ney, H., Kohnen, M., Schubert, H. and Wein, B.B., "Content-based image retrieval in medical applications", ***Methods of Information in Medicine***, Vol. 43, No. 4, (2004), 354-361.
14. Lehmann, T.M., Güld, M.O., Deselaers, T., Keysers, D., Schubert, H., Spitzer, K., Ney, H. and Wein, B.B., "Automatic categorization of medical images for content-based retrieval and data mining", ***Computerized Medical Imaging and Graphics***, Vol. 29, No. 2, (2005), 143-155.
15. Weisi, L., Tao, D., Kacprzyk, J., Li, Z., Izquierdo, E. and Wang, H. (Eds.), "Multimedia analysis, processing and communications", Springer Science & Business Media, Vol. 346, (2011).
16. Metz, C.E. and Pan, X., "A unified analysis of exact methods of inverting the 2-d exponential radon transform, with implications for noise control in spect", ***IEEE Transactions on Medical Imaging***, Vol. 14, No. 4, (1995), 643-658.
17. Defrise, M. and Clack, R., "A cone-beam reconstruction algorithm using shift-variant filtering and cone-beam backprojection", ***IEEE Transactions on Medical Imaging***, Vol. 13, No. 1, (1994), 186-195.
18. Kuchment, P., "The radon transform and medical imaging, SIAM", (2013), doi: org/10.1137/1.9781611973297.
19. Tizhoosh, H.R., "Barcode annotations for medical image retrieval: A preliminary investigation", in Image Processing (ICIP), 2015 IEEE International Conference on, IEEE., (2015), 818-822.
20. Tizhoosh, H.R. and Rahnamayan, S., "Evolutionary projection selection for radon barcodes", in Evolutionary Computation (CEC), 2016 IEEE Congress on, IEEE., (2016), 1-8.
21. Babaie, M., Tizhoosh, H., Zhu, S. and Shiri, M., "Retrieving similar x-ray images from big image data using radon barcodes with single projections", ***arXiv preprint arXiv:1701.00449***, (2017), https://arxiv.org/abs/1701.00449.
22. Krizhevsky, A., Sutskever, I. and Hinton, G.E., "Imagenet classification with deep convolutional neural networks", in Advances in neural information processing systems., (2012), 1097-1105.
23. Khatami, A., Babaie, M., Khosravi, A., Tizhoosh, H., Salaken, S.M. and Nahavandi, S., "A deep-structural medical image

classification for a radon-based image retrieval", in Electrical and Computer Engineering (CCECE), 2017 IEEE 30th Canadian Conference on, IEEE., (2017), 1-4.

24. LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., "Gradient-based learning applied to document recognition", ***Proceedings of the IEEE***, Vol. 86, No. 11, (1998), 2278-2324.
25. Bastien, F., Lamblin, P., Pascanu, R., Bergstra, J., Goodfellow, I., Bergeron, A., Bouchard, N., Warde-Farley, D. and Bengio, Y., "Theano: New features and speed improvements", ***arXiv Preprint arXiv:1211.5590***, (2012), https://arxiv.org/abs/1211.5590.
26. Khatami, A., Babaie, M., Khosravi, A., Tizhoosh, H. and Nahavandi, S., "Parallel deep solutions for image retrieval from imbalanced medical imaging archives", ***Applied Soft Computing***, Vol. 63, (2018), 197-205.
27. Liu, X., Tizhoosh, H.R. and Kofman, J., "Generating binary tags for fast medical image retrieval based on convolutional nets and radon transform", in Neural Networks (IJCNN), 2016 International Joint Conference on, IEEE., (2016), 2872-2878.
28. Zhu, S. and Tizhoosh, H.R., "Radon features and barcodes for medical image retrieval via svm", in Neural Networks (IJCNN), 2016 International Joint Conference on, IEEE., (2016), 5065-5071.
29. Sze-To, A., Tizhoosh, H.R. and Wong, A.K., "Binary codes for tagging x-ray images via deep de-noising autoencoders", in Neural Networks (IJCNN), 2016 International Joint Conference on, IEEE., (2016), 2864-2871.
30. Tizhoosh, H.R., Mitcheltree, C., Zhu, S. and Dutta, S., "Barcodes for medical image retrieval using autoencoded radon transform", ***arXiv preprint arXiv:1609.05112***, (2016), https://arxiv.org/abs/1211.5590.

A Radon-based Convolutional Neural Network for Medical Image Retrieval

A. Khatami[a], M. Babaie[b], H. R. Tizhoosh[b], A. Nazari[c], A. Khosravi[a], S. Nahavandi[a]

[a] *Institute for Intelligent System Research and Innovation, Deakin University, Australia*
[b] *KIMIA Lab., University of Waterloo, Ontario, Canada*
[c] *School of Information Technology, Deakin University, Australia*

P A P E R I N F O

Paper history:
Received 22 August 2017
Received in revised form 24 Januray 2018
Accepted 26 May 2018

Keywords:
Deep Convolutional Neural Network
Image Retrieval in Medical Application
Medical Image Retrieval
Radon Transformation

چکیده

اخیرا سیستم های دسته بندی و بازیابی تصاویر پزشکی به دلیل دسترسی به موارد مشابه ذخیره شده و پیگیری خط سیر درمان این موارد مشابه، بیشتر از هر زمانی مورد توجه قرار گرفته اند. با این حال، به علت عدم وجود داده های متوازن و برچسب گذاری شده در ابعاد وسیع (به خصوص در تصاویر پزشکی) دسترسی به این سیستم ها را دشوار ساخته است. در این مطالعه به منظور ایجاد یک سیستم بازیابی در یک دیتابیس غیرمتوازن، از تبدیل Radon تصاویر به همراه یک شبکه عصبی عمیق استفاده گردیده است. در بخش نتایج نشان داده شد که اعمال این تبدیل به لایه ورودی یک شبکه عمیق، دقت بازیابی اطلاعات را به طور چشمگیری افزایش می دهد .

***doi**: 10.5829/ije.2018.31.06c.07*