Underground Disease Detection Based on Cloud Computing and Attention Region Neural Network

Pinjie Xu[2], Ce Li[1, 2, *], Liguo Zhang[3, 4], Feng Yang[1, 2], Jing Zheng[1, 5] and Jingwu Feng[2]

Abstract: Detecting the underground disease is very crucial for the roadbed health monitoring and maintenance of transport facilities, since it is very closely related to the structural health and reliability with the rapid development of road traffic. Ground penetrating radar (GPR) is widely used to detect road and underground diseases. However, it is still a challenging task due to data access anywhere, transmission security and data processing on cloud. Cloud computing can provide scalable and powerful technologies for large-scale storage, processing and dissemination of GPR data. Combined with cloud computing and radar detection technology, it is possible to locate the underground disease quickly and accurately. This paper deploys the framework of a ground disease detection system based on cloud computing and proposes an attention region convolution neural network for object detection in the GPR images. Experimental results of the precision and recall metrics show that the proposed approach is more efficient than traditional objection detection method in ground disease detection of cloud based system.

Keywords: Cloud computing, ground penetrating radar, convolution neural network.

1 Introduction

In recent years, with the rapid development of road traffic in China, safety inspection of concrete structures is very crucial to the cost-effective maintenance of Chinese transport facilities, since it is closely related with the structural health and reliability. As an important technology, ground penetrating radar (GPR) transmits and receives high-frequency and broadband electromagnetic waves in microwave band, and determines the location of underground targets by utilizing the characteristics that electromagnetic waves

[1] State Key Laboratory of Coal Resources and Safe Mining, China University of Mining & Technology, Beijing, 100083, China.

[2] School of Mechanical Electronic & Information Engineering, China University of Mining & Technology, Beijing, 100083, China.

[3] Computer Science & Engineering, Harbin Engineering University, Harbin, 150001, China.

[4] Computer Science & Engineering, Hong Kong University of Science and Technology, Hong Kong, China.

[5] Stanford University, Palo Alto, CA 94305-6104, United States.

[*] Corresponding Author: Ce Li. Email: celi@cumtb.edu.cn.

emit at the junction of underground media. It is widely used in railway detection [Hugenschmidt (2000)], road engineering [Saarenketo and Scullion (2000)], mining engineering [Strange and Jecny (2018)], tunnel detection [Du and Deng (2015)] and other fields. However, it is still a challenging task due to data access anywhere, transmission security and data processing on cloud.

Ground Penetrating Radar (GPR) detection needs to store, process and detect data quickly, accurately and effectively. Cloud computing technology with its massive and high-speed processing capability makes it competent for such work effectively. To solve the issues, a novel cloud based GPR data transmission and processing system is in need to be developed with improved processing capability and storage [Liang, Yuan and Yang (2017)]. For large scale GPR data transmission, some researchers have attempted mobile hardware devices for cloud based multimedia data transmission. The recent works [Yang, He, Lin et al. (2017)] designed a security framework of multimedia perception, communication, service with application of Internet of Things. Rosario et al. proposed a high efficiency transmission method about mobile multimedia using wireless link quality and information protocol [Rosário, Zhao, Santos et al. (2014)]. However, there are some research challenges in terms of data access control, cloud storage on multimedia data, and inspection processing service.

For GPR data processing service, the cloud computing and machine learning methods are widely developed [Xia, Ma, Shen et al. (2018); Xia, Xiong, Vasilakos et al. (2017); Xia, Yuan, Lv et al. (2019); Zeiler and Fergus (2014)] and integrated to recognize the disease [Girshick, Donahue, Darrell et al. (2014); Redmon, Divvala, Girshick et al. (2016)]. Various GPR based object detection methods take advantages of signal or image processing techniques as an active research direction to provide accurate results. Recently, convolution neural network attracted more and more attention [Cui, McIntosh and Sun (2018)]. The region convolution neural network is widely applied in image segmentation and objection detection [Zhang, Wang, Li et al. (2018)], however the detection accuracy is affected by the candidate region extraction method. Traditional object detection method get region by slide window [Redmon, Divvala, Girshick et al. (2016)], selective search [Uijlings, Van De Sande, Gevers et al. (2013)], they cannot get the importance of extraction regions. In this paper, we introduce a GPR data transmission protocol to improve the access efficiency, and propose a novel attention region neural network with the squeeze and excitation block [Hu, Shen and Sun (2017)] to get a weighted feature map. Then a region proposal network could get weighted candidate region from the feature map. By the validation, our proposed method brings convenience to municipal security assurance due to the fast transmission and detection. The experimental results validate that the designed system and proposed method can provide more effective detecting performance.

The remainder of this paper is organized as follows. Section 2 introduces the cloud based GPR image detection system. Section 3 explains the proposed method for object detection in GPR data. Experimental result and analysis are given in Section 4, and conclusion in Section 5.

2 Design of the cloud-based GPR image detection system

We have listed three main functional modules, user manager module, data management module, detection service module. The user manager module contains create user, delete user, update user. Radar vehicle and other equipment will upload radar data, GPS data and distance data to the cloud, which will be dumped and processed by the server, and then downloaded and evaluated by users, so the data management module contains data acquisition, data transmission, data storage, data download. The numerical radar signal needs to be transformed into radar image at the cloud end. After object detection, the disease location can be realized by combining distance data and GPS information, so the detection service module contains data visualization, object detection, disease location.

We designed a cloud-based GPR image detection system. Our cloud platform system has four layers of structure, perception and interaction layer, Internet layer, management and service layer, comprehensive application layer, as shown in Fig. 1. In perception and interaction layer, data collection staff upload the original GPR data by GPR vehicle or portable GPR equipment. Professional testing and evaluation personnel monitor the location of disease by downloading the processed data from the cloud. With the help of the Internet, massive radar data are uploaded and analyzed by server providers. The core of our system is the management and service layer. Radar digital signal is transformed into radar image, and the position of the image and the confidence of the disease are calculated by the object detection algorithm on the radar image. Integrating algorithm recognition and expert evaluation, we can build specific applications in the comprehensive application layer. This paper uses road GPR data to verify the feasibility of the system. To ensure the security of information transmission, we adopt the media-aware traffic security architecture which contains four major components key management, batch rekeying, authentication, and watermarking [Xia, Yuan, Lv et al. (2019)]. The architecture considers traffic analysis, security requirements and traffic scheduling of multimedia applications.
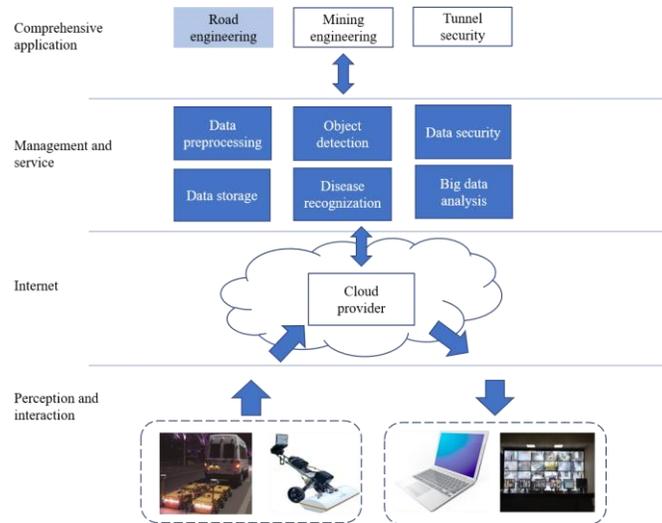
Figure 1: The cloud-based GPR image detection system
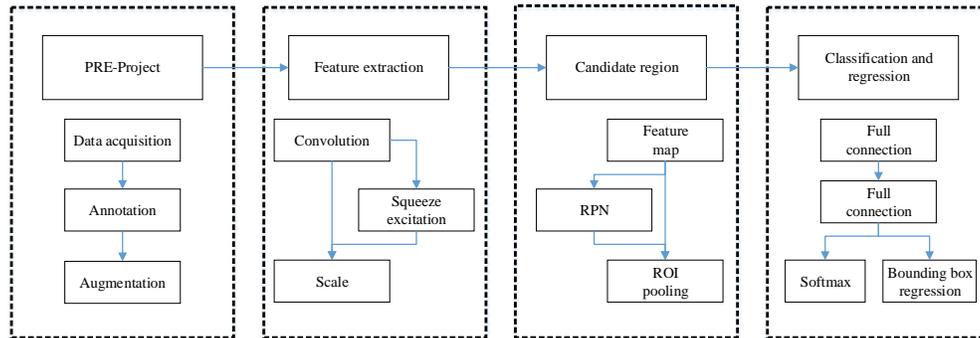
Figure 2: The four steps of disease detection method

3 Proposed method for disease detection

We divided the disease detection method into four steps: pre-project, feature extraction, candidate region extraction, classification and regression, as shown in Fig. 2. In pre-project step, we collected and annotated GPR images. Augmentation has been adopted to prevent the overfit. In features extraction step, the convolution neural network with squeeze and excitation block is used to extract radar image features and obtain feature graphs. The squeeze and excitation block give convolution neural network the attention mechanism which obtains the importance of each convolution feature channel through self-learning, and then increases the useful features and inhibits the unimportant features. In candidate region extraction step, candidate regions are obtained from the feature map through the region proposal network (RPN) [Ren, He, Girshick et al. (2015)]. Because of the attention mechanism, RPN can extract more meaningful candidate regions from feature map. In regression and classification step, the classification and probability are calculated by using the softmax function, and the target position is calculated by the bounding box regression.

3.1 Attention region neural network

3.1.1 Attention region neural network

Each convolution layer of the convolution neural network is computed by multiple convolution kernels to obtain a multi-channel feature map, which can get rich features from the above feature map. If each channel is treated equally, it is detrimental to the learning of important features. We add squeeze and excitation block on traditional convolution layer, given a weight to each channel, make effective channels for higher weight, and poor effective channel for less weight [Hu, Shen and Sun (2017)]. The squeeze and excitation block add an attention mechanism on the convolution neural network. To achieve the attention mechanism, we adopted squeeze mechanism and activation operation. The squeeze operation is realized through global average pooling [Lin, Chen and Yan (2013)].

A channel with specification $H \times W \times C$ is compressed into a channel with specification $1 \times 1 \times C$, which is realized by the following formula:

$$Z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i,j) \tag{1}$$

where Z_c is the output characteristic diagram, U_c is the input characteristic diagram, H is the height of the characteristic diagram, W is the width of the characteristic diagram, and C is the number of channels of the feature map. The activation operation is implemented by a series of full connection layers and activation functions. The full connection layer is used to learn the deep features of each channel weight value, and the activation function gives the activation operation the ability to learn nonlinear features. The global average pooling layer is followed by the full connection layer of dimension C/r and C, which is activated by the ReLu function [Jarrett, Kavukcuoglu, LeCun et al. (2009)] and Sigmoid function respectively. Finally, the weight of size $1 \times 1 \times C$ is multiplied by the channel corresponding to the original characteristic graph as the output of the squeeze and excitation block. The squeeze and excitation block of convolution layer is shown in Fig. 3.
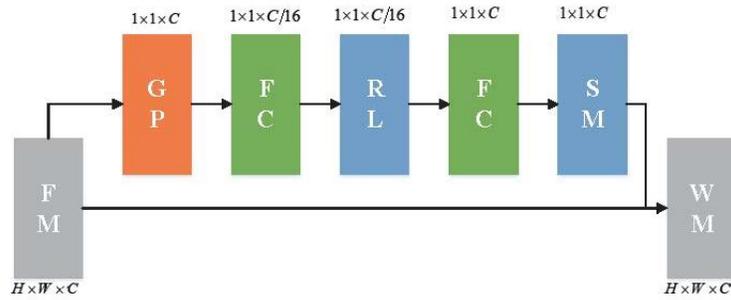
Figure 3: Structure diagram of squeeze and excitation block. FM: the input feature map; GP: global average pooling layer; RL: ReLu activation; FC: full connection layer; SM: Sigmoid activation; WM: the output weighted feature map

3.1.2 Convolution neural network

In this experiment, a five-layer convolution neural network is designed to extract features from radar images, as shown in Tab. 1. The first two convolution kernels are divided into 7×7, 5×5, and the last three convolution kernels are. The five convolution layers will generate the feature maps containing 96, 256, 384, 384 and 256 channels respectively. To reduce network parameters and prevent over-fitting, the pooling layer with size of is used after the first two convolution layers. The last three convolution layers output more channels, to enable the network to learn the importance of each channel independently, a squeeze and activation module is added after each convolution layer. Finally, the convolution neural network output a weighted feature map of 256 channels as the input of the region proposal net.

Table 1: The structure of deep convolution neural network

Layer	Kernel Size	Channel	Stride
Conv1	7×7	96	2
Pool1	3×3	-	2
Conv2	5×5	256	2
Pool2	3×3	-	2
Conv3	3×3	384	1
SE block3	-	-	-
Conv4	3×3	256	1
SE block4	-	-	-
Conv5	3×3	256	1
SE block5	-	-	-

3.2 Region proposal network

The region proposal network is different from the traditional region selection method, which directly extracts candidate regions on the feature map output of the last convolution layer. In this way, object detection and region extraction can share the same features, and the implementation of the algorithm can be completed on GPU. We slide a slide window on the feature map to get candidate regions. Slide windows have nine different size (three size, three scale). Each sliding window is mapped to a 256-dimensional vector. The scores are calculated by softmax, the coordinates are calculated by bounding box regression, and the features are normalized by the pooling layer of interest area [Girshick (2015)].

4 Experimental result and analysis

4.1 Data acquisition and annotation

In this experiment, the original radar data of the roadbed is obtained by the automatic scanning and sampling of the vehicle roadbed detection radar, and then the radar reflection signal sequence at the interface on each scanning road surface is extracted for spectrum analysis, and the spectral grayscale image of strong reflected signals in the ground is obtained. In this experiment, the collected radar image is divided into three types: mud pumping, roadbed subsidence and underground cavity. As shown in Fig. 4.

We collected a total of 483 images of disease radars, including 217 mud pumping, 115 roadbed subsidence and 151 underground cavity. The number of diseases in each image is greater than or equal to 1, a total of 261 mud pumping diseases, 134 roadbed subsidence diseases, 204 underground cavity diseases are marked. It is necessary to perform reasonable data augmentation on the radar image [Van Dyk and Meng (2001)]. We mirrored the disease image to obtain a total of 966 data, which was used as the disease data set. When training the model, use the expansion method such as inversion and random cropping to further expand the sample size. 90% of the disease data set was

selected as the training set to construct the disease detection model; 10% was selected as the test set to evaluate the model performance.

(a) Mud Pumping   (b) Roadbed Subsidence

(c) Underground Cavity   (d) Normal Roadbed
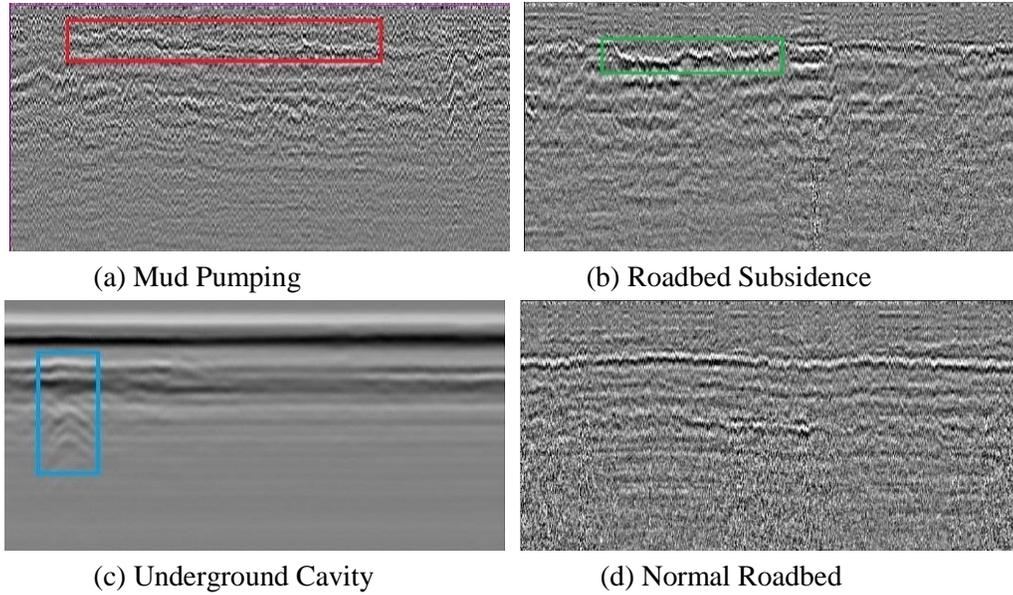
Figure 4: Data description

4.2 Experimental setting

In order to verify the effectiveness of the method, we have done a series of experiments. They are carried out on the client of the developed system prototype. We run the experiments on Intel(R) Core(TM) @3.2GHz CPU unit, 16GB memory, NVIDA GeForce GTX 1050 Ti graphics card, drive with 64-bit windows file system in Microsoft server. We use the weights of the ZF network [Zeiler and Fergus (2014)] to initialize the first five volume layers and the last two full connection layers. The weight is trained by Ren et al. [Ren, He, Girshick et al. (2015)] on the ImageNet dataset [Deng, Dong, Socheret al. (2009)]. The squeeze and excitation block and region proposal net are randomly initialized by a Gaussian distribution with a mean of 0 and a standard deviation of 0.01. The model was optimized by stochastic gradient descent [LeCun, Boser, Denker et al. (1989)]. Each time one sample was used for forward propagation. The initial learning rate was set at 0.01, the learning rate was lowered to 0.001 after 50,000 iterations, and the model converged after 70,000 iterations.

4.3 Model evaluation

In order to further quantify the model performance, Precision and Recall (PR) curve was adopted as the evaluation index. PR curve can show classification and retrieval performance well. Precision reflects the proportion of positive examples determined by the classifier in all predicted positive samples. The recall reflects the ratio of correctly predicted positive samples to true positive samples. Average Precision (AP) was used as

JAI, vol.1, no.1, pp.9-18, 2019

the test score for a certain type of disease, and the AP value was the area under the PR curve. The Mean Average Precision (mAP) was further adopted as the score of the whole model, among which mAP was the Average of all kinds of AP. The evaluation results were shown in Tab. 2 and Fig. 5.

Table 2: Detection performance

Class	AP	mAP
cavity	0.995	
mud	0.924	0.919
subsidence	0.838	
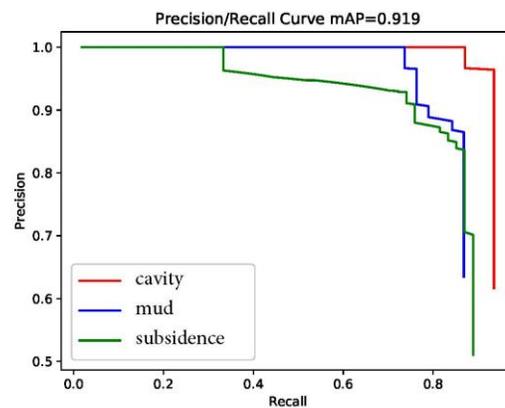
Figure 5: The precision and recall curve

It can be seen from Tab. 2 that the model has best detection effect (0.995) for the underground cavity disease, followed by the mud pumping (0.924) and the roadbed subsidence (0.838), and the overall score of the model is relatively excellent (0.919). It can be seen that there is a large gap between the three categories. From the data distribution, the data amount of roadbed subsidence disease was the least (268). Although the data expansion is processed, there is still over-fitting on training set which resulting in poor performance of the model test set. It can be prevented from over-fitting by collecting more data, using more effective data expansion methods and adding dropout [Krizhevsky, Sutskever and Hinton (2012)] into the network. For the best detection effect of underground cavity, underground cavity detection can be extracted into a separate model and applied to special disease detection. Although the amount of data is basically the same as that of the underground cavity, the detection results are not as good as well as the underground cavity, which is related to the more complicated disease characteristics. The hyperbolic boundary of the mud pumping is more blurred, the size difference is large. The length of some serious mud pumping may be consistent with the image width, which leads to the big difference in distribution between training set, training set and test set. According to the three classes, the model performs well, but there is still room for improvement.

5 Conclusion

To solve the challenge of data access anywhere, transmission security and data processing on cloud, the paper proposed a cloud based GPR data disease detection system. To build a disease detection system, the paper proposed an attention region convolution neural network. The attention mechanism enables the region proposal net to learn the candidate with weights. Experiments show that our method is more effective than traditional object detection cloud based method. However, the feature extractor for the disease of vague outline needs improvement, and the cost remains a decisive factor to embrace the cloud based system. Therefore, future works may be directed to these issues for disease detection system.

References

Cui, Q.; McIntosh, S.; Sun, H. (2018): Identifying materials of photographic images and photorealistic computer generated graphics based on deep CNNs. *Computers, Materials & Continua*, vol. 55, no. 2, pp. 229-241.

Deng, J.; Dong, W.; Socher, R.; Li, L. J.; Li, K. et al. (2009): ImageNet: a largescale hierarchical image database. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248-255.

Du, T.; Deng, X. (2015): Application of GPR in highway tunnel lining quality detection. *Electronic Journal of Geotechnical Engineering*, vol. 20, no. 1, 359-367.

Girshick, R. (2015): Fast R-CNN. *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1440-1448.

Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. (2014): Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580-587.

Hu, J.; Shen, L.; Sun, G. (2017): Squeeze-and-excitation networks. arXiv:1709.01507.

Hugenschmidt, J. (2000): Railway track inspection using GPR. *Journal of Applied Geophysics*, vol. 43, no. 2-4, pp. 147-155.

Jarrett, K.; Kavukcuoglu, K.; LeCun, Y.; LeCun, Y. (2009): What is the best multistage architecture for object recognition? *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2146-2153.

Krizhevsky, A.; Sutskever, I.; Hinton, G. E. (2012): ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, pp. 1097-1105.

LeCun, Y.; Boser, B.; Denker, J. S.; Henderson, D.; Howard, R. E. et al. (1989):

Backpropagation applied to handwritten zip code recognition. *Neural computation*, vol. 1, no. 4, pp. 541-551.

Liang, Y.; Yuan, Y.; Yang, F. (2017): Research of parallel process method of ground penetrating radar (GPR) data based on Hadoop. *Journal of System Simulation*, vol. 29, no. 1, pp. 120-128.

Lin, M.; Chen, Q.; Yan, S. (2013): Network in network. arXiv: 1312.4400.

Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. (2016): You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779-788.

Ren, S.; He, K.; Girshick, R.; Sun, J. (2015): Faster R-CNN: towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, pp. 91-99.

Rosário, D.; Zhao, Z.; Santos, A.; Braun, T.; Cerqueira, E. (2014): A beaconless opportunistic routing based on a cross-layer approach for efficient video dissemination in mobile multimedia IoT applications. *Computer Communications*, vol. 45, pp. 21-31

Saarenketo, T.; Scullion, T. (2000): Road evaluation with ground penetrating radar. *Journal of Applied Geophysics*, vol. 43, no. 2-4, pp. 119-138.

Strange, A. D.; Jecny, Z. (2018): Coal subsurface mapping for selective mining. *Proceedings of the IEEE International Conference on Ground Penetrating Radar*, pp. 1-4.

Uijlings, J. R.; Van De Sande, K. E.; Gevers, T.; Smeulders, A. W. (2013): Selective search for object recognition. *International Journal of Computer Vision*, vol. 104, no. 2, pp. 154-171.

Van Dyk, D. A.; Meng, X. L. (2001): The art of data augmentation. *Journal of Computational and Graphical Statistics*, vol. 10, no. 1, pp. 1-50.

Xia, Z.; Ma, X.; Shen, Z.; Sun, X.; Xiong, N. N. et al. (2018): Secure image LBP feature extraction in cloud-based smart campus. *IEEE ACCESS*, vol. 6, no. 1, pp. 30392-30401.

Xia, Z.; Xiong, N. N.; Vasilakos, A. V.; Sun, X. (2017): Epcbir: an efficient andprivacy-preserving content-based image retrieval scheme in cloud computing. *Information Sciences*, vol. 387, pp. 95-204.

Xia, Z.; Yuan, C.; Lv, R.; Sun, X.; Xiong, N. N. (2019): A novel weber local binary descriptor for fingerprint liveness detection. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*.

Yang, J.; He, S.; Lin, Y.; Lv, Z. (2017): Multimedia cloud transmission and storage system based on internet of things. *Multimedia Tools Application*, vol. 76, pp. 17735-17750.

Zeiler, M. D.; Fergus, R. (2014): Visualizing and understanding convolutional networks. *In European Conference on Computer Vision*, pp. 818-833.

Zhang, Y.; Wang, Q.; Li, Y.; Wu, X. (2018): Sentiment classification based on piecewise pooling convolutional neural network. *Computers, Materials & Continua*, vol. 56, no. 2, pp. 285-297.