Journal of Real-Time Image Processing (2020) 17:1–5
https://doi.org/10.1007/s11554-020-00947-2

GUEST EDITORIAL

Introduction to the special issue on deep learning for real-time information hiding and forensics

Zhili Zhou[1] · Ching-Nung Yang[2] · Cheonshik Kim[3] · Stelvio Cimato[4]

Published online: 28 January 2020

1 Introduction

With rapid development of network technologies and the wide use of digital camera devices, digital multimedia (in particular image and video data) has grown tremendously on the internet. However, by using various powerful multimedia processing tools, the multimedia data is getting easier to be illegally copied, and forged. Consequently, the copyright and privacy protection of multimedia data have been an urgent requirement in the field of multimedia security. To prevent multimedia data from unauthorized use and privacy invasion, two typical approaches have been proposed: information hiding and digital forensics. Both of these approaches have been applied successfully in many digital security applications and they can also complement each other nicely. However, in cloud computing and big data environments, there exist massive amount of multimedia data and new multimedia data keeps growing exponentially. In this context, the information hiding and digital forensics face many new challenges. For example, how to conduct information hiding and digital forensics in real-time, and how to rapidly process a large amount of multimedia data?

Nowadays, with the development of graphic processing unit (GPU) processors and the availability of large-scale training datasets such as ImageNet, the deep learning techniques have gained success in the field of machine learning and computer vision, and have shown to outperform many traditional techniques. Although deep learning techniques require computationally extensive off-line training process, with the help of powerful GPUs, they show high efficiency at online testing stage, which enables achieving real-time computation performance for a variety of computer vision tasks. Also, it is possible to explore the deep learning techniques for various real-time multimedia security applications, such as real-time information hiding and digital forensics in order to achieve desirable performances in the terms of both accuracy and efficiency.

The aim of this special issue is to publish the latest research works in the related areas of real-time information hiding and forensics. It provides the effective and efficient deep learning techniques or other intelligent techniques that have potential to address challenges of real-time information hiding and digital forensics. According to a rigorous review procedure, each manuscript submitted to the special issue was reviewed by two or more anonymous experts. Based on the reviews' comments, a total of 16 papers have been selected to be included in this special issue, which fall into the following three major categories: (1) image forensics (9 papers), (2) steganography and steganalysis (4 papers), and (3) other emerging security topics (3 papers). In the following sections, an overview of these papers is provided.

✉ Zhili Zhou
zhou_zhili@163.com

Ching-Nung Yang
cnyang@gms.ndhu.edu.tw

Cheonshik Kim
mipsan@daum.net

Stelvio Cimato
stelvio.cimato@unimi.it

[1] Jiangsu Engineering Centre of Network Monitoring, School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China

[2] Department of Computer Science and Information Engineering, National DongHwa University, Shoufeng, Hualien 97401, Taiwan

[3] Department of Computer Engineering, Sejong University, Seoul, Republic of Korea

[4] Department of Computer Science, Università degli studi di Milano, Milan, Italy

2 Image forensics

2.1 Image manipulation forensics

Since the non-aligned double JPEG (NA-DJPEG) compression is one of most common image manipulations, the paper entitled "Non-aligned double JPEG compression detection based on refined markov features in QDCT domain", co-authored by Wang et al. [1], focuses on JPEG compression forensics. In this paper, the authors make full use of the color information in images, and propose refined Markov in quaternion discrete cosine transform (QDCT) domain for NA-DJPEG detection. The proposed refinement method not only reduces redundant features, but also makes the acquired features more efficient in detection. Therefore, the refined Markov features can not only capture the intra-block correlation between block QDCT coefficients but also improve computing efficiency in real-time. Finally, support vector machine (SVM) method is employed for NA-DJPEG compression detection. The experiment results demonstrate that the proposed algorithm not only make use of color information of images, but also can achieve better detection performance with small size images (i.e., 64×64) than the state-of-the-arts for NA-DJPEG detection.

In another paper entitled "Real-time estimation for the parameters of Gaussian filtering via deep learning", co-authored by Ding et al. [2], the authors propose a method to estimate the parameters of Gaussian filtering to process images based on convolutional neural networks (CNN). The experiments show that the proposed method can provide excellent real-time performance in estimating the window size and standard deviation of Gaussian filterings. The well-trained model can achieve satisfying estimation accuracy as well as the validation time efficiency.

Instead of focusing on a certain type of manipulations, the paper entitled "A real-time image forensics scheme based on multi-domain learning", co-authored by Yang et al. [3], proposes a generic forensic method using multidomain learning convolutional neural network (MDL-CNN) for the forensics of many kinds of manipulations. In this method, features of modified image extracted from different datasets are then fed into the MDL-CNN in training process. Since the proposed MDL-CNN is trained by many types of image datasets generated by different kinds of manipulations, this method can distinguish many types of modified images. To decrease the computation of proposed scheme, 1×1 kernel convolution layer is used in the second convolutional layer of each network. Furthermore, a multi-domain loss function is developed to enhance the recognition ability of indepth learning features. The evaluation results show that MDL-CNN-based forensic framework not only achieves real-time forensics but also provides significant performance improvement compared with state-of-the-art methods.

Different from the above passive forensic methods, the paper entitled "A real-time reversible image authentication method using uniform embedding strategy", co-authored by Yao et al. [4], proposes an active forensic method for tampered region detection and location. In this method, a uniform embedding strategy is adopted in this paper, in which one AC bit is embedded into each divided image block to ensure they have the same authentication capability. To improve the forgery localization precision, the block size is adaptively sought according to the embedding capacity of the image. In addition, during the image authentication process, the embedding parameters and location map information are verified to increase the process's rigorousness. The experimental results demonstrate the superiority of the detection precision of the proposed method.

2.2 Fingerprint forgery forensics

The main security issue of real-time fingerprint authentication systems is that most fingerprint scanners are vulnerable to presentation attacks by artificial replicas, made from plastic clay, gelatin, silicon, wood glue, etc. Thus, the fingerprints are easily to be forged. To detect forged fingerprints, the paper entitled "Semi-supervised stacked autoencoder-based deep hierarchical semantic feature for real-time fingerprint liveness detection", co-authored by Yuan et al. [5], propose an anti-spoofing attack scheme, called real-time fingerprint liveness detection (RFLD), to discriminate live or fake fingerprints. Different from the most of existing RFLD solutions all relied on handcrafted feature extraction and selection, this scheme applies stacked autoEncoder to RFLD to automatically learn deep hierarchical semantic features representation, and its consists of two parts: parameter pre-training based on unsupervised learning and RFLD based on supervised learning. The performance has been verified on two public fingerprint datasets: LivDet 2011 and 2013, and the experimental results indicate that the proposed approach works well for RFLD as well as the detection performance is satisfactory.

2.3 Forensics on other applications

Besides the above image manipulation forensics and fingerprint forgery forensics, digital forensics are also extended to many other common applications, such as person re-identification, human age estimation, and VPN traffic identification.

Person re-identification aims to detect the specified person across nonoverlapping cameras. It is a difficult forensic task due to the appearance variations caused by occlusion, human pose change, background clutter, illumination variation, and etc. In the paper entitled "Multi-level feature fusion model based real-time person re-identification for forensics",

co-authored by Wang et al. [6], a multi-level feature fusion model (MFFM) is designed to combine both deep features and handcrafted features in real-time. MFFM is firstly utilized to describe person appearance. Then, local binary pattern (LBP) and histogram of oriented gradient (HOG) are extracted to cope with geometric change and illumination variance. Experimental results indicate MFFM can achieve the best performance compared to the state-of-the-art models on the Market1501, CUHK03, and VIPeR datasets.

Human age estimation (AE) is an emerging research topic in computer vision and digital forensics and has attracted increasing amount of research due its wide potential applications. The paper entitled "Real-time human cross-race aging-related face appearance detection with deep convolution architecture" is co-authored by Tian et al. [7]. To specially explore the relationships between aging and facial appearances across races, this paper is devoted to determining the correspondence between facial aging and facial appearances. Specifically, it firstly extracts appearance vector features from facial images with their spatial structure preserved. Then, it selects the aging-related features shared by different races to explore their aging-related common facial regions, while removing redundant features. Thirdly, it improves the proposed model by incorporating potential cross-race relationships in an automated learning manner. Additionally, it extends the model with deep convolution architecture. The proposed methodologies are evaluated on a large face aging database with real-time efficiency.

Since there is a gap between meteorological satellite cloud images and the true information of the pictured clouds, extracting the true atmospheric information from "forged" satellite images in real-time is a challenging task. The paper entitled "A real-time typhoon eye detection method based on deep learning for meteorological information forensics", co-authored by Zhao et al. [8], proposes a real-time typhoon eye detection method from meteorological satellite cloud images based on deep learning. This new approach is the first step in detecting hidden information in satellite cloud images and provides important data support to detect true typhoon information. The simulation experiments and the results show that the proposed method performs well in identifying typhoons. In the testing process, the average time needed to detect each sample is 6 ms, which fulfills the requirement for real-time typhoon eye detection. The proposed method outperforms the k-nearest neighbors (KNN) and support vector machine (SVM) algorithms.

Real-time VPN traffic identification has become an increasingly important task in network management and security maintenance. The paper entitled "Deep learning-based real-time VPN encrypted traffic identification methods", co-authored by Guo et al. [9], proposes two deep learning-based models to classify the traffic into VPN and non VPN traffic. These models utilize convolutional auto-encoding (CAE) and convolutional neural network (CNN) respectively, preprocessing the traffic samples into session pictures, to accomplish the experiment objectives. The CAE-based method, utilizing the unsupervised nature of CAE to extract the hidden layer features, can automatically learn the nonlinear relationship between original input and expected output. The CNN-based method performs well in extracting two-dimensional local features of images. Experimental results show that the models perform better than traditional identification methods.

3 Image steganography and steganalysis

3.1 Image steganography

As a class of new techniques aiming to solve the problem of covert communication under lossy channels, robust steganography has become a new research hotspot in the field of information hiding. The paper entitled "Enhancing reliability and efficiency for real-time robust adaptive steganography using cyclic redundancy check codes", is co-authored by Zhang et al. [10]. To improve the communication reliability and efficiency for current real-time robust steganography methods, a concatenated code, composed of Syndrome-trellis codes (STC) and cyclic redundancy check (CRC) codes, is proposed in this paper. The enhanced robust adaptive steganography framework proposed is this paper is characterized by a strong error detection capability, high coding efficiency and low embedding costs. On this basis, three adaptive steganographic methods resisting JPEG compression and detection are proposed. Then, the fault tolerance of the proposed steganography methods is analyzed using the residual model of JPEG compression, thus obtaining the appropriate coding parameters. Experimental results show that the proposed methods have a significantly stronger robustness against compression, and are more difficult to be detected by statistical based steganalysis methods.

Generally, the traditional image steganography embeds secret information into a cover image by slightly modifying its content, and thus the modification traces will be inevitably left in the cover image, which makes successful steganalysis possible. The concept of "coverless steganography" was proposed in [11–13]. Instead of employing a designated cover image for embedding the secret data, appropriate images that already contain the secret data are chosen as stego-images for secret communication. In the paper entitled "Coverless real-time image information hiding based on image block matching and dense convolutional network", co-authored by Luo et al. [14], a novel coverless information hiding method based on deep learning is proposed. This method chooses a set of real-time stego-images which share one or several visually similar blocks with the given secret image as stego-images. In this method, a group of real-time

images searched online are segmented according to specific requirements. Then, the DenseNet is used to extract the high-level semantic features of each similar block. At the same time, a robust hash sequence with feature sequence, DC and location is generated by DCT. The inverted index structure based on the hash sequence is constructed to attain real-time image matching efficiently. At the sending end, the stego-images are matched and sent through feature matching. At the receiving end, the secret image can be recovered by extracting similar blocks through the received stego-images and stitching the image blocks according to the location information. Experimental results demonstrate that the proposed method provides better robustness and has higher retrieval accuracy and capacity when compared with some existing coverless image information hiding.

3.2 Image steganalysis

Image steganalysis can be regarded as the adversary to steganography, and it aims to detect whether an image contains secret data or not.

The paper entitled "Efficient binary image steganalysis based on ensemble neural network of multi-module", co-authored by Liu et al. [15], is devoted to binary image steganalysis. In this paper, an efficient binary image steganalysis scheme based on CNN which integrates high-pass filters, truncated linear unit and subnetworks is proposed. In the process of binary image steganography, flipped pixels usually scatter on the boundaries of the content in the image. Therefore, the first convolutional layer is constructed with high-pass filters to capture the structure of embedded signals better. Truncated linear unit (TLU) is also adopted after the first convolutional layer for the same purpose. Four truncated linear units with different truncated values are adopted in order to capture embedding signals of different intensities. We also adopt four subnets after the 4 truncated linear units to further boost the performance of the CNN network. The experimental results show that the proposed scheme is efficient and effective on binary steganalysis.

Since deep learning techniques have gained great success in computer vision, many works also employ the deep learning techniques for image steganalysis to achieve higher accuracy and efficiency. The paper entitled "Deep learning for real-time image steganalysis: a survey", co-authored by Ruan et al. [16], gives a survey on real-time image steganalysis based on deep learning. In this paper, an account of preliminary knowledge is described firstly. A brief overview of the deep neural networks (DNN) is also presented. The combination of DNN and real-time image steganalysis is reviewed. For multi-user scenarios, a practical real-time image steganalysis application based on outlier detection methods is analyzed. At last, the future issues of real-time image steganalysis are also prospected.

4 Other emerging security issues

With the development of big data and cloud computing, more and more data-owners store the data in cloud servers. Considering privacy-preserving, images data needs to be encrypted before uploaded to the cloud, which will lead to inefficient image retrieval of ciphertext domain. In the paper entitled "A privacy-preserving image retrieval method based on deep learning and adaptive weighted fusion", coauthored by Jiaohua et al. [17], the authors propose a privacy-preserving image retrieval method based on deep learning and adaptive weighted fusion. Firstly, it extracts low-level feature EHD (edge histogram descriptor), BOW (bag of words) and high-level semantic feature of images. Secondly, the dimension of 1024-dim high-level semantic feature is reduced by PCA (principal component analysis), and the feature is binarized. Then these types of features are adaptively fused. Finally, it constructs a prefilter table for fusion features to improve search efficiency by locality sensitive hashing (LSH) algorithm. K-nearest neighbor (KNN) algorithm and logistic encryption method are used to protect the privacy of fused features and images respectively. The experiments show that the proposed method can not only ensure image security but also improve the retrieval accuracy of encrypted images.

Restricted by local constrained storage space, users can store their images with the assist of the cloud. However, the cloud is a remote semi-trusted party that may extract stored images for adversaries due to monetary reasons. To address the issue, in the paper entitled "Secure real-time image protection scheme with near-duplicate detection in cloud computing", coauthored by Liu et al. [18], a secure real-time image protection scheme is proposed. It can be used to enhance the security of the stored images in cloud computing. Moreover, the convergent encryption is used to construct this scheme, which can provide functionalities of image deduplication checking and near-duplicate detection for the image owner. To improve the efficiency of the near-duplicate detection, deep learning is exploited in this scheme to extract images. Security analysis indicates that the proposed scheme can meet the security requirements of correctness and security. Performance analysis shows that the proposed scheme can be performed with low computational cost.

The high efficiency video coding (HEVC) provides better coding efficiency compared to its predecessors H.264/AVC. However, due to the adoption of a large variety of coding unit (CU) sizes, at RD optimization level, the quadtree partition of the CU still consumes a large proportion of the encoding complexity. Hence, the computational complexity cost remains a critical issue that must be properly considered in the optimization task. The paper entitled "Fast CU partition-based machine learning approach for reducing

HEVC complexity", is co-authored by Bouaafia et al. [19]. In this paper, two machine learning-based fast CU partition methods for inter-mode HEVC are proposed, in order to optimize the complexity allocation at CU level. Firstly, an online support vector machine (SVM) based fast CU algorithm is proposed for reducing HEVC complexity. Secondly, a deep convolutional neural network (CNN) is designed to predict the CU partition, in which large-scale training database including substantial CU partition data is considered. Experimental results demonstrate that, compared to the state-of-the art, the two proposed approaches outperform the related works in terms of both RD performance and complexity reduction at inter-mode.

5 Conclusion

Finally, we would like to appreciate the editors-in-chief of the *Journal of Real-Time Image Processing*, Nasser Kehtarnavaz and Matthias F. Carlsohn, and the entire editorial staff for their support throughout the preparation and publication of this special issue. We would like to thank the authors for their contributions to this issue and reviewers for their valuable comments. We hope that this special issue will have an extensive impact on real-time digital forensics, steganography, steganalysis, secret communication, and deep learning techniques. This work was supported by National Natural Science Foundation of China under Grant 61972205, 61602253, U1836208, U1536206, U1836110, in part by MOST under contracts 108-2634-F- 259-001-through Pervasive Artificial Intelligence Research (PAIR) Labs, Taiwan, in part by the National Key R&D Program of China under Grant2018YFB1003205, in part by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) fund, and in part by the Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET) fund, China.

References

1. Wang, J., Huang, W., Luo, X., Shi, Y.-Q., Jha, S.K.: Non-aligned double JPEG compression detection based on refined Markov features in QDCT domain. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00929-z
2. Ding, F., Shi, Y., Zhu, G., Shi, Y.: Real-time estimation for the parameters of Gaussian filtering via deep learning. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00907-5
3. Yang, B., Li, Z., Zhang, T.: A real-time image forensics scheme based on multi-domain learning. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00893-8
4. Yao, H., Wei, H., Qin, C., Tang, Z.: A real-time reversible image authentication method using uniform embedding strategy. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00904-8
5. Yuan, C., Chen, X., Yu, P., Meng, R., Cheng, W., Jonathan Wu, Q.M., Sun, X.: Semi-supervised stacked autoencoder-based deep hierarchical semantic feature for real-time fingerprint liveness detection. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00928-0
6. Wang, S., Xin, X., Liu, L., Tian, J.: Multi-level feature fusion model based real-time person re-identification for forensics. J. Real Time Image Process. (2019). https://doi.org/10.1007/10.1007/s11554-019-00908-4
7. Tian, Q., Zhang, W., Mao, J., Yin, H.: Real-time human cross-race aging-related face appearance detection with deep convolution architecture. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00903-9
8. Zhao, L., Chen, Y., Sheng, V.S.: A real-time typhoon eye detection method based on deep learning for meteorological information forensics. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00899-2
9. Guo, L., Qianqiong, W., Liu, S., Duan, M., Li, H., Sun, J.: Deep learning-based real-time VPN encrypted traffic identification methods. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00930-6
10. Zhang, Y., Luo, X., Zhu, X., Li, Z., Bors, A.G.: Enhancing reliability and efficiency for real-time robust adaptive steganography using cyclic redundancy check codes. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00905-7
11. Zhou, Z., Sun, H., Harit, R., Chen, X., Sun, X.: Coverless image steganography without embedding. In: Int. Conf. Cloud Comput. Security, pp. 123–132 (2015)
12. Zhou, Z., Mu, Y., Jonathan Wu, Q.M.: Coverless image steganography using partial-duplicate image retrieval. Soft Comput. **23**(13), 4927–4938 (2019)
13. Zhou, Z., Cao, Y., Wang, M., Fan, E., Jonathan Wu, Q.M.: Faster-RCNN based robust coverless information hiding system in cloud environment. IEEE Access. **7**, 179891–179897 (2019)
14. Luo, Y., Qin, J., Xiang, X., Tan, Y., Liu, Q., Xiang, L.: Coverless real-time image information hiding based on image block matching and dense convolutional network. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00917-3
15. Liu, J., Wei, L., Zhan, Y., Chen, J., Zhaopeng, X., Li, R.: Efficient binary image steganalysis based on ensemble neural network of multi-module. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00885-8
16. Ruan, F., Zhang, X., Zhu, D., Zhanyang, X., Wan, S., Qi, L.: Deep learning for real-time image steganalysis: a survey. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00915-5
17. Jiaohua, Q., Jianhua, C., Xuyu, X., Yun, T., Wentao, M., Jing, W.: A privacy-preserving image retrieval method based on deep learning and adaptive weighted fusion. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00909-3
18. Liu, D., Shen, J., Wang, A., Wang, C.: Secure real-time image protection scheme with near-duplicate detection in cloud computing. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00887-6
19. Bouaafia, S., Khemiri, R., Sayadi, F.E., Atri, M.: Fast CU partition-based machine learning approach for reducing HEVC complexity. J. Real Time Image Process. (2019). https://doi.org/10.1007/s11554-019-00936-0