Multimedia Data and Security

Balakrishnan Prabhakaran, *The University of Texas at Dallas, Dallas, TX, 75080, USA*

The detection of manipulation of multimedia data can be categorized as active and passive approaches.[1] The active approaches use source information, such as watermarks and digital signatures. For instance, Bahirat et al.[2] proposed a watermarking-based framework for authentication and localization of tampering in red, green, blue (RGB) and the 3-D point cloud. The watermarking methods, which are computationally expensive, cannot be applied on raw unprocessed data. Passive approaches, or blind forensics,[3] are intended for testing multimedia data where the original or source information is not available. Deep-learning-based tampering detection avoids the need to perform various forensic tests to detect whether multimedia data have been manipulated or not.[4] However, deep-learning-based methodologies need to balance false positives and negatives, as pointed out by Bayar and Stamm.[4] Apart from the need for detecting manipulations in multimedia data, there is also a need for localizing (i.e., identifying the region) where the manipulation occurred.

In this context, several deep learning architectures have been proposed: some requiring extensive feature engineering,[5] some with new layers for better feature learning,[4] a triple network with conditional random fields,[6] and autoencoder-based feature extraction and labeling.[7] Various training methods for these deep learning network architectures have been the topic of research as well: transfer learning; feature extraction using deep networks followed by traditional machine learning for classification; and patch-based learning, especially when the dataset is small. These methods have been demonstrated to be successful in detecting multimedia data manipulations, but they cannot localize the manipulated area.

In terms of localizing the manipulated area of an image, multiple research works hold promise: a mask region-based convolutional neural network (R-CNN) and Sobel edge detection filter were incorporated by Wang et al.[8] to focus on manipulated boundaries, and Salloum et al.[9] used multiple convolution branches and merged them. Some of the localization approaches work only for specific compression methodologies. For instance, the two-stream R-CNN network in Zhou et al.[10] focuses only on JPEG images, with one stream for processing RGB data and the other one working with a steganalysis rich model filter that finds the noise inconsistency between pristine and doctored areas. This helps to find the area of interest by identifying different compression levels within an image. Here, it should be observed that any one particular methodology for compression and decompression is not specified in the JPEG standard. Instead, the standard leaves room for innovation to be incorporated by different implementations. Beneš et al.[11] systematically analyzed JPEG versions since 1998, providing an exhaustive comparison of the outputs of compression and decompression operations for a range of parameter settings.

Addressing the need for content authentication, Fang et al.[12] describe a perceptual image hashing-based technique to extract the features of the image and map them to a hash sequence with a fixed length. Building on this, Xing et al.[13] integrate transformer-based multilayer constraints with a novel perceptual robust image hashing scheme for authenticating the contents of an image. These methods only authenticate the content without identifying the region of multimedia data that has been manipulated. With increasing sophistication in the approaches used for attacking and manipulating multimedia content, the need for advancing the techniques to authenticate, detect, and localize multimedia data manipulations is stronger.

1070-986X © 2024 IEEE
Digital Object Identifier 10.1109/MMUL.2024.3378058
Date of current version 8 April 2024.

REFERENCES

1. A. Piva, “An overview on image forensics,” *Int. Scholarly Res. Notices*, vol. 2013, 2013, Art. no. 496701, doi: 10.1155/2013/496701.
2. K. Bahirat, U. Shah, A. A. Cardenas, and B. Prabhakaran, “ALERT: Adding a secure layer in decision support for advanced driver assistance system (ADAS),” in *Proc. 26th ACM Int. Conf. Multimedia*, 2018, pp. 1984–1992, doi: 10.1145/3240508.3241912.
3. Y. Zhan, Y. Chen, Q. Zhang, and X. Kang, “Image forensics based on transfer learning and

convolutional neural network," in *Proc. 5th ACM Workshop Inf. Hiding Multimedia Secur.*, 2017, pp. 165–170, doi: 10.1145/3082031.3083250.

4. B. Bayar and M. C. Stamm, "A deep learning approach to universal image manipulation detection using a new convolutional layer," in *Proc. 4th ACM Workshop Inf. Hiding Multimedia Secur.*, 2016, pp. 5–10, doi: 10.1145/2909827.2930786.
5. Y. Zhang, J. Goh, L. L. Win, and V. L. Thing, "Image region forgery detection: A deep learning approach," in *Proc. Singap. Cyber-Secur. Conf. (SG-CRC)*, 2016, vol. 2016, pp. 1–11.
6. B. Chen, X. Qi, Y. Wang, Y. Zheng, H. J. Shim, and Y.-Q. Shi, "An improved splicing localization method by fully convolutional networks," *IEEE Access*, vol. 6, pp. 69,472–69,480, 2018, doi: 10.1109/ACCESS.2018.2880433.
7. D. Cozzolino and L. Verdoliva, "Single-image splicing localization through autoencoder-based anomaly detection," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Piscataway, NJ, USA: IEEE Press, 2016, pp. 1–6, doi: 10.1109/WIFS.2016.7823921.
8. X. Wang, H. Wang, S. Niu, and J. Zhang, "Detection and localization of image forgeries using improved mask regional convolutional neural network," *Math. Biosci. Eng.*, vol. 16, no. 5, pp. 4581–4593, 2019, doi: 10.3934/mbe.2019229.
9. R. Salloum, Y. Ren, and C.-C. J. Kuo, "Image splicing localization using a multi-task fully convolutional network (MFCN)," *J. Vis. Commun. Image Representation*, vol. 51, pp. 201–209, Feb. 2018, doi: 10.1016/j.jvcir.2018.01.010.
10. P. Zhou et al., "Learning rich features for image manipulation detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 1053–1061.
11. M. Beneš, N. Hofer, and R. Böhme, "Know your library: How the libjpeg version influences compression and decompression results," in *Proc. ACM Workshop Inf. Hiding Multimedia Secur. (IH&MMSec)*, New York, NY, USA: ACM, 2022, pp. 19–25, doi: 10.1145/3531536.3532962.
12. Y. Fang, Y. Zhou, X. Li, P. Kong, and C. Qin, "TMCIH: Perceptual robust image hashing with transformer-based multi-layer constraints," in *Proc. ACM Workshop Inf. Hiding Multimedia Secur. (IH&MMSec)*, New York, NY, USA: ACM, 2023, pp. 7–12, doi: 10.1145/3577163.3595113.
13. H. Xing, H. Che, Q. Wu, and H. Wang, "Image perceptual hashing for content authentication based on Watson's visual model and LLE," *J. Real-Time Image Process.*, vol. 20, no. 1, 2023, Art. no. 7, doi: 10.1007/s11554-023-01269-9.

BALAKRISHNAN PRABHAKARAN is a professor with the Computer Science Department, The University of Texas at Dallas, Dallas, TX, 75080, USA. Contact him at bprabhakaran@utdallas.edu.