Byte-Level Object Identification for Forensic Investigation of Digital Images

1st Abdul Rehman Javed
National Center for Cyber Security
Air University, Islamabad, Pakistan
https://orcid.org/0000-0002-0570-1813

2nd Zunera Jalil
Department of Cyber Security and NCCS
Air University, Islamabad, Pakistan
https://orcid.org/0000-0003-2531-2564

***Abstract*—Lately, digital data has increased a key role in providing and sharing information. Pictures and video recordings are utilized to pass on convincing messages to be utilized under a few unique situations, from propaganda to coercing. The majority of the effort in the present digital crime investigation network lies in the acquisition, retrieval, and investigation of existing data from digital machines. It is a time consuming and a humanly difficult task to collect, process and analyze each media content manually. In this paper, we provide a novel approach that solves a real-time problem for an investigator while investigating the suspect machine. Our approach acquires all image data at byte level from the suspect machine, perform fast and accurate object detection resorting to the deep learning-based algorithm and present high-level illustration of images containing suspicious object and unique objects that can be presented as evidence. Our approach aims to flag photos where suspicious objects are detected. Performance and time consumption wise, this study confirms the importance of automated object detection in digital forensics.**

***Index Terms*—Digital Forensics, Resent, Byte-code, Object Identification, Crime Investigation .**

I. Introduction

With the advent of broadband internet, smart devices with huge storage capacity and more users on social media platforms, huge amounts of data are being generated [1], [2]. This data is used by intelligent algorithms for beneficial purposes and get people to make smart decisions. A digital human today enjoys much of his physical and mental work by being carried out by machines, to which he is entitled. In addition to that, this data is used by criminals and they exploit users by abusing this data [2]. These days, all around the globe, electronic crime is expanding. Investigations into digital crime are underway, but the speed of the investigation cannot be compared to the speed, complexity, and volume of crimes. This leaves users in chaos and helplessness. Users want to take advantage of technology, but they also want to ensure their privacy and security. An increase in electronic crimes has made it essential to upgrade all strategies that battle this kind of crime. [3].

Recent advances in technology have entered the world into a new digital era where image, video and audio content are the preferred medium to convey messages/information [4], [5]. Smart devices have become the preferred means to capture images and videos, and contents are commonly shared through social media platforms (SMPs, for example, Facebook, YouTube, etc.). The cheap availability and high accessibility of digital multimedia devices have made it difficult to manually analyze multimedia content. The investigators use several tools to perform forensic procedures to obtain inevitable evidence against criminals to hold them responsible in the court of law. According to Quick & Choo [6], three primary contributors to the digital information storm: I) more held onto gadgets; ii) more cases requiring forensic assessment and iii) more information on every person.

A. Problem Statement

The process of digital forensic starts right after the incident is identified or reported as shown in Figure 1.

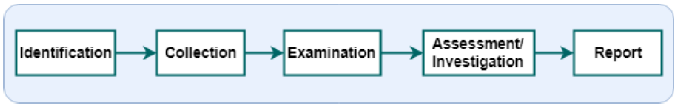

Fig. 1. Process of digital forensic investigation

Acquisition of data from suspect's machines is the second step, examination or analysis is the third step, then the fourth step is the assessment and final reports are arranged to be presented before the court of law [7].

In this process, the examination step is the most time consuming and humanly difficult step to do as suspect devices contain terabytes of data which is hard to analyze manually. The existing tools provide features to present data in a customized fashion. Investigators can have multi-dimensional views but there is a need to automate this process. Several crimes related to cyber harassment and cyberstalking involve images and video data. This data is challenging to investigate due to the time complexity of existing image processing techniques. To address this problem, we use a deep learning method to analyze digital images from the suspect devices.

Below are the key contributions of this work:

- Propose an approach for achieving dramatic improvement in evidence analysis, time complexity, and operational efficiency through the adoption of systematic approaches.
- Utilize ResNet for fast robust and efficient object detection.

978-1-7281-6840-1/20/$31.00 ©2020 IEEE

- Aims to flag photos where suspicious objects are detected.
- Build Connection graphs for object clustering.

The remaining paper is organized as follows: Section II discusses the related research on object detection. Section III discusses the working of deep learning ResNet model and parameter environment. In Section IV, we explain the object detection approach and present the results. Finally, the conclusion is summarized in Section V.

II. Prior and related work

With the development of digital forensics and its applications in industry, various software is rising in the market that started concentrating on various sorts of issues that manage digital investigation [7]. We audit the related work, first concentrating on media authenticity then face recognition and detection, afterward on object detection. Because of the gigantic volume of logical writing on face location and on face acknowledgment, we essentially center around object identification methods applied to computerized crime scene investigation which is the key contribution of this work.

Doctored photos are showing up in newspaper and style magazines, government media, predominant press, internet-based life, online closeout locales, web-based dating destinations, political promotion crusades, and logical diaries. All the more as of late, the coupling of news with counterfeit symbolism has been utilized by people and state-supported substances to upset fair decisions, prompt common and political conflict, and fuel brutality [5], [6], [8], [9]. The author in [1] used the Faster R-CNN network for image forgery and object detection. The author evaluated his approach to image tamper datasets to evaluate the performance. The author in [2] presented a framework that can detect objects using CNN as well as tracking the object. They target CCTV videos to evaluate their approach. They also show a low computational approach for object detection that reduces the cost of GPUs and high computational sources. Authors in [8], [10] presented an overview of the different multimedia forensics techniques. In [8], authors presented a study discussing whether the content is authentic or modified, without the knowledge of any prior information about the image under analysis. There show that there exist some approaches to authenticate the image to be valid.

Recently there are numerous studies focused on deep learning methods for face recognition and detection and object identification. Authors in [11] categorize face recognition methods in two significant classes: I) rigid-templates based methods and ii) Deformable Parts-based Model (DPM) based methods. Same authors further group the unbending layouts into three principle group of calculations: I) Viola-Jones [12] face identification calculation and varieties, comprehensively recognized as Haar-course; ii) techniques dependent on picture recovery, what's more, Histograms of Oriented Gradients (HOG) [13] and, at long last, iii) calculations dependent on Convolutional Neural Networks (CNN) and profound CNN's [14].

Authors discuss how face and object can be identified from still images [15]–[17] as well as real-time [18]–[20]. Some authors discussed forgery detection and tempering techniques [3]–[5], [8]. Authors in [16] used high-capacity convolutional neural networks (CNNs) and domain-based fine-tuning on PASCAL VOC 2012 dataset to boost the performance of object detection. They report a high-performance boost of performance of 30% in comparison with previous works. Authors in [15] used ImageNet classification dataset (2012) for image classification. They report that 152-layer ResNet has less complexity than other deep learning models as well as provides promising results and low error rate.

The closest paper to our work presented by Domingue et al. [9] proposed a semi-automated face detection approach in the field of digital image forensics. They only focused on face detection and face identification. Besides this, we focus on detecting suspicious objects found in a suspect's machine.

III. Deep Residual Learning for Object Identification (ResNet)

ResNet is a deep learning model that improves the classification performance as well as makes it less inclined to overfitting. Figure 2 provides the overview of a Resnet block. It is to be noted, that increasing the layers of a network ought not to be finished by essentially stacking the layers on top of each other since it might prompt the issue of vanishing gradients in the network. The issue is, as the gradient is back-propagated to prior layers, repeated multiplication may make the gradient infinitely small which might result in poor performance [15].

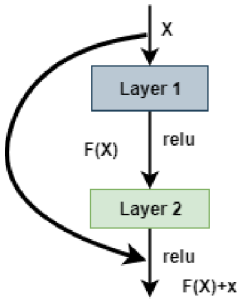

Fig. 2. Illustration of Residual Network block

$$y = f(x, \{W_i\}) + W_s X \tag{1}$$

Here x and y are the input and output vectors of the layers considered. The function $F(x, W_i)$ represents the residual mapping to be learned. It is reported that best performing techniques are complex intricate ensemble frameworks that commonly consolidate numerous low-level image highlights with the high-level setting. There exist a range of ResNet models such as ResNet, ResNetV2, InceptionResNetV2 and different configurations of ResNet model such as ResNet50, ResNet101, ResNet152, ResNet50V2, ResNet50V2, ResNet101V2 and ResNet152V2 [15]. In this work, we use the Renset152 model for object detection.

Table I shows the computing environment in which experiment is conducted and Table II explains the details of the parameters used in this paper.

TABLE I
COMPUTING ENVIRONMENT

Parameter	Value
Operating System	Ubuntu 18.04.2 LTS
CPU	Xeon E5/Corei5
RAM	128GB
GPU	NVIDIA GeFroce 1080
CUDA Verion	9.0
Python Version	3.7

TABLE II
DETAILS OF DEEP RESIDUAL NETWORK

Parameter	Value
Weights	ImageNet
include_top	True
pooling	None
Type of internal layer unit	LSTM
Input dimension	(224, 224, 3)
Learning Optimizer	Adam
Batch size	256
Learning Rate start	0.1
Convolution	True
Batch Normalization	True
Activation Function	ReLU
Decay	0.0001
Momentum	0.9

IV. BYTE LEVEL OBJECT DETECTION

The proposed approach consists of three major and two minor steps as shown in Figure 3: data acquisition, feature byte extraction and detection of objects are the major steps while feature transformation and validation are embedded into feature byte extraction step.

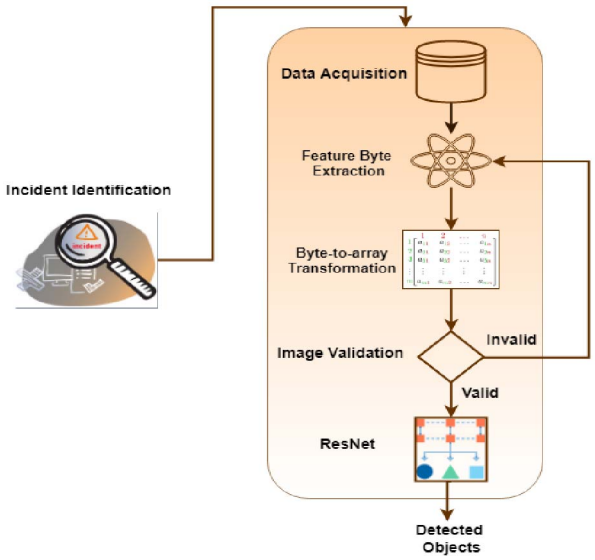

Fig. 3. Proposed Approach of the Object Detection

A. Data Acquisition

The data acquisition step is performed on a suspect hard drive. Figure 4 shows the six levels of the hard drives. Each level provides fine granularity for the previous one. We program the data collection code. The program first checks the volume, then collect the information about the partitions and then check the type of file system. After that, it accesses the memory blocks and collects the data as bytes. We use this byte data to make a Raw disk image of 6 GB using FTK disk imager [23].

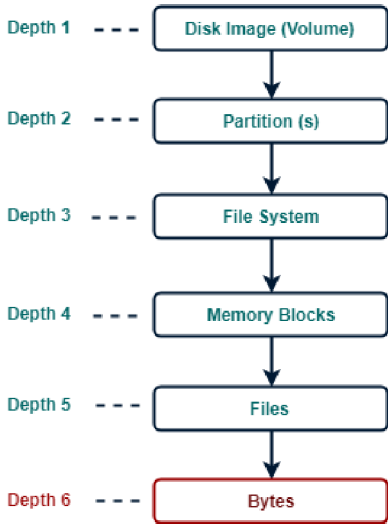

Fig. 4. Taxonomy of the hard disk analysis

B. Feature Byte Extraction

Acquired image contains huge byte data which is preprocessed to obtain only multimedia data. We separate the media files based on the type of the file. Next, we read all the byte images (Image-to-Array step) and then convert it into a Raw image of (224×224) as RGB. Our approach checks whether an input file is valid (validation step) and it has a size of (224×224). Otherwise, other data is discarded. The preprocess_input function is used to the adequate image to the format the model requires and transforms a standard image into an appropriate input.

C. ResNet based Object Detection and Results

We choose the ResNet 152 layers model for object detection. There exist two types of object detection approaches; a proposal based and proposal free approach. Conventional proposal based deep learning methods require high computational costs such as R-CNN [16], Fast-RCNN [17], Faster RCNN [21] and R-FCN [18] while proposal free approaches like YOLO and SSD [19], [22] are fast enough for real-time object detection. After the validation step, the image is fed into the pre-trained ResNet classifier. ResNet is trained on the ImageNet 2012 classification dataset [20] which consists of 1000 classes and has a promising accuracy till time. The model is trained on the 1.28 million training images and evaluated on

the 50k validation images. ResNet pre-trained model shows a promising accuracy of object detection. Figure 5 provides a summarized illustration of the top 18 objects that are found frequently in the 6GB acquired disk image. The suspect machine contains a high range of images related to sports and ammunition. Basketball is the highest detected object, then baseball bat and similarly other. Our approach makes

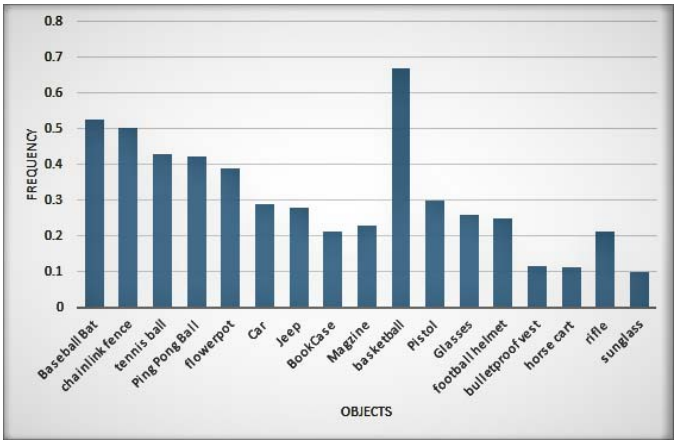

Fig. 5. Frequency graph of the detected objects

a connection graph by clustering similar types of objects as shown in Figure 6. Cluster on the left side (yellow) illustrates the top objects related to sports, cluster on the right side (red) show the objects related to ammunition, cluster on the bottom left side show the objects related to studies and cluster on right bottom side show the object related to fashion. It is seen that the suspect has two priority interest (Sports and Ammunition).

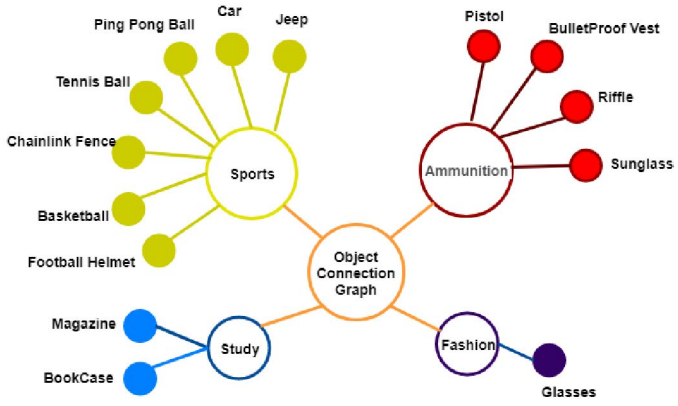

Fig. 6. Object Connection graph

V. Conclusion

The vast majority of the exertion in the present digital crime investigation network lies in the acquisition, retrieval, and investigation of existing data from digital machines. It is a time consuming and a humanly difficult task to collect, process and analyze each media content manually. We proposed a deep learning-based object detection approach that is capable of detecting objects from the acquired disk image of the suspect machine to make the forensic investigation process fast, efficient and robust. We investigate around 6 GB of multimedia content in less then 40 minutes which takes days if done manually by forensic experts. We present the highly detected objects as well as connection graphs of the objects. This work leads towards automated forensic investigation that is the need of time.

References

[1] S. Qu, "An approach based on object detection for image forensics", In 1st International Conference on Industrial Artificial Intelligence (IAI) IEEE, pp. 1-6, 2019.

[2] A. Axenopoulos, V. Eiselein, A. Penta, E. Koblents, E. L. Mattina, and P. Daras, "A Framework for Large-Scale Analysis of Videoïn the Wildto Assist Digital Forensic Examination", IEEE Security & Privacy, vol. 17, no. 1, pp. 23-33, 2019.

[3] R. Khalaf and A. Varol, "Digital Forensics: Focusing on Image Forensics", 2019 7th International Symposium on Digital Forensics and Security (ISDFS), 2019.

[4] A. M. Marshall, "Standards, regulation & quality in digital investigations: The state we are in", Digital Investigation, vol. 8, no. 2, pp. 141–144, 2011.

[5] R. Mercuri, "Courtroom Considerations in Digital Image Forensics", Springer-Verlag, pp. 313–325, 2013.

[6] D. Quick and K. R. Choo, "Impacts of increasing volume of digital forensic data: A survey and future research challenges", Digital Investigation vol. 11, no. 4, pp. 273–294, 2014.

[7] K. Ghazinour, D. Vakharia, K. Kannaji, and R. Satyakumar, "A study on digital forensic tools", 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), 2017.

[8] A. Piva, "An overview on image forensics", ISRN Signal Processing, vol. 2013, 2013.

[9] P. Domingues, and R. A. Frazão, "Deep Learning-based Facial Detection and Recognition in Still Images for Digital Forensics", In Proceedings of the 14th International Conference on Availability, Reliability, and Security, pp. 1-10. 2019.

[10] S. Milani, M. Fontani, P. Bestagini, M. Barni, A. Piva, M. Tagliasacchi, and S. Tubaro, "An overview on video forensics", APSIPA Transactions on Signal and Information Processing, vol. 1, no. 11, 2012.

[11] S. Zafeiriou, C. Zhang, and Z. Zhang, "A survey on face detection in the wild: Past, present and future", Computer Vision and Image Understanding vol. 138, pp. 1–24, 2015.

[12] P. Viola and M. J. Jones, "Robust real-time face detection", International journal of computer vision, vol. 57, no. 2, pp. 137–154, 2004.

[13] C. Shu, X. Ding, and C. Fang, "Histogram of the oriented gradient for face recognition", Tsinghua Science and Technology, pp. 216–224, 2011.

[14] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A Convolutional Neural Network Cascade for Face Detection", In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition", arXiv preprint arXiv:1512.03385, 2015.

[16] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", in CVPR, 2014.

[17] R. Girshick, "Fast r-cnn", in ICCV, 2015.

[18] Y. Li, K. He, J. Sun et al., "R-fcn: Object detection via region-based fully convolutional networks", in NIPS, 2016.

[19] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection", in CVPR, 2016.

[20] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. Imagenet large scale visual recognition challenge. arXiv:1409.0575, 2014.

[21] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks", in NIPS, 2015.

[22] W. Liu, D. Anguelov, D. Erhan et al., "Ssd: Single shot multibox detector", in ECCV, 2016.

[23] "Forensic Toolkit (FTK)", AccessData, 2019. [Online]. Available: https://accessdata.com/products-services/forensictoolkit- ftk. [Accessed: 22- Aug- 2019]