Frame-wise Detection of Relocated I-frames in Double Compressed H.264 Videos Based on Convolutional Neural Network

Peisong He, Xinghao Jiang, Tanfeng Sun, Shilin Wang, Bin Li, Yi Dong

PII: S1047-3203(17)30141-4
DOI: http://dx.doi.org/10.1016/j.jvcir.2017.06.010
Reference: YJVCI 2027

To appear in: *J. Vis. Commun. Image R.*

Revised Date: 6 June 2017
Accepted Date: 20 June 2017

Please cite this article as: P. He, X. Jiang, T. Sun, S. Wang, B. Li, Y. Dong, Frame-wise Detection of Relocated I-frames in Double Compressed H.264 Videos Based on Convolutional Neural Network, *J. Vis. Commun. Image R.* (2017), doi: http://dx.doi.org/10.1016/j.jvcir.2017.06.010

Frame-wise Detection of Relocated I-frames in Double Compressed H.264 Videos Based on Convolutional Neural Network

Peisong He[a,b], Xinghao Jiang[a,b,*], Tanfeng Sun[a,b], Shilin Wang[a,b], Bin Li[c], Yi Dong[a,b]

[a] *School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University*
[b] *National Engineering Lab on Information Content Analysis Techniques, GT036001 Shanghai, China*
[c] *Shenzhen Key Laboratory of Media Security, Shenzhen University, Shenzhen, 518060, China*

Abstract

Relocated I-frames are a key type of abnormal inter-coded frame in double compressed videos with shifted GOP structures. In this work, a frame-wise detection method of relocated I-frame is proposed based on convolutional neural network (CNN). The proposed detection framework contains a novel network architecture, which initializes with a preprocessing layer and is followed by a well-designed CNN. In the preprocessing layer, the high-frequency component extraction operation is applied to eliminate the influence of diverse video contents. To mitigate overfitting, several advanced structures, such as 1×1 convolutional filter and the global average-pooling layer, are carefully introduced in the design of the CNN architecture. Public available YUV sequences are collected to construct a dataset of double compressed videos with different coding parameters. According to the experiments, the proposed framework can achieve a more promising performance of relocated I-frame detection than a well-known CNN structure (AlexNet) and the method based on average prediction residual.

Keywords: Double compression detection, Data-driven methodology, Convolutional neural network, Video forensics

1. Introduction

According to the investigation proposed by Interactive Advertising Bureau, digital videos keep rising in media consumption, content distribution and entertainment over the world [1]. Nowadays, everyone can easily access to digital videos on the Internet and alter video contents using advanced editing software, such as Adobe Premiere. The ease of conducting tampering operations threatens the authentication and the integrity of digital videos, especially when

*Corresponding author
Email addresses: gokeyhps@sjtu.edu.cn (Peisong He), xhjiang@sjtu.edu.cn (Xinghao Jiang), tfsun@sjtu.edu.cn (Tanfeng Sun), wsl@sjtu.edu.cn (Shilin Wang), libin@szu.edu.cn (Bin Li), aa44@sjtu.edu.cn (Yi Dong)