Paired mini-batch training: A new deep network training for image forensics and steganalysis

Jin-Seok Park, Hyeon-Gi Kim, Do-Guk Kim, In-Jae Yu, Heung-Kyu Lee

PII: S0923-5965(18)30434-X
DOI: https://doi.org/10.1016/j.image.2018.04.015
Reference: IMAGE 15384

To appear in: *Signal Processing: Image Communication*

Received date : 30 November 2017
Revised date : 2 March 2018
Accepted date : 29 April 2018

Paired mini-batch training: A new deep network training for image forensics and steganalysis

Jin-Seok Park, Hyeon-Gi Kim, Do-Guk Kim, In-Jae Yu, Heung-Kyu Lee*

School of computing, Korea Advanced Institute of Science and Technology (KAIST), Daehak-ro, Yuseong-gu, Daejeon, 34141, Republic of Korea

Abstract

Deep convolutional neural networks (convnets) have recently become popular in many research areas because convnets can extract features automatically and classify them with high accuracy. Researchers in the image forensics and steganalysis field have proposed methods using convnets to develop technologies that work in practical environments. However, they found that the convnets used for computer vision were not suitable for image forensics and steganalysis because these convnets tend to learn features that represent the contents of images rather than forensic or steganalysis features. To overcome this limitation, researchers have proposed various structures, but there are no studies that take into account other factors related to training neural networks for image forensics and steganalysis. In this paper, we clearly represent the training process for image forensics and steganalysis using a training equation and explain why training convnets with the standard mini-batch is inefficient for image forensics and steganalysis. We then propose a new mini-batch, called the paired mini-batch, which is better suited for image forensics and steganalysis.

Keywords: Deep learning, Deep convolutional neural networks, Mini-batch, Image forensics, Steganalysis

1. Introduction

Deep convolutional neural networks (convnets) have demonstrated excellent performance in applications for various types of computer vision. Since Krizhevsky et al. reported that convnets could be classified in various categories with high accuracy [1], many computer vision methods using convnets have been proposed: for example, face detection [2], pedestrian detection [3], saliency detection [4], super-resolution of images [5], video classification [6], etc. The methods proposed using convnets are simpler and perform better than existing hand-craft-based methods.

Recently, convnets have been employed for computer vision as well as in many research areas because they can generate features automatically and distinguish different categories with high accuracy even in complex environments. In speech recognition [7], natural language processing [8, 9], and other research fields [10, 11], several technologies have already demonstrated good performance using convnets, and researchers in the image forensics and steganalysis fields have also developed technologies using convnets.

Both image forensics and steganalysis involve the classification of normal and manipulated images. Image forensics [12, 13] is aimed at determining whether an image is genuine or forged, which means checking whether the image was directly captured with a camera or manipulated using image editing programs such as Photoshop. Steganalysis [14] is aimed at determining whether an image contains secret messages. The results of steganalysis can be classified into two types: normal or stego, which indicates the presence of secret messages. Fig. 1 shows examples of normal and manipulated images generated through image manipulation and steganography, as well as the differences between the two types of images. These image types contain slight differences, which should be classified through image forensics and steganalysis.

Previous image forensics and steganalysis methods relied on hand-crafted features resulting from the process of image acquisition with a camera or from the image editing process. These features include photo response non uniformity [15], color filter array patterns [16], meaningful noise [17], discrete cosine trans-

*Corresponding author

Email address: heunglee@kaist.ac.kr (Heung-Kyu Lee)