# Recapture Image Forensics Based on Laplacian Convolutional Neural Networks

Pengpeng Yang, Rongrong Ni(✉), and Yao Zhao

Beijing Key Laboratory of Advanced Information Science and Network Technology, Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China
{rrni,yzhao}@bjtu.edu.cn

**Abstract.** Recapture image forensics has drawn much attention in public security forensics. Although some algorithms have been proposed to deal with it, there is still great challenge for small-size images. In this paper, we propose a generalized model for small-size recapture image forensics based on Laplacian Convolutional Neural Networks. Different from other Convolutional Neural Networks models, We put signal enhancement layer into Convolutional Neural Networks structure and Laplacian filter is used in the signal enhancement layer. We test the proposed method on four kinds of small-size image databases. The experimental results have demonstrate that the proposed algorithm is effective. The detection accuracies for different image size database are all above 95%.

**Keywords:** Recapture images forensics · Laplacian Convolution Neural Networks · Laplacian filter

## 1 Introduction

With the rapid development of the multimedia technology, a variety of image processing software are springing up in modern society, such as Adobe Photoshop, Meitu, and CorelDRAW, which make editing images easy. Everyone can tamper an image by means of these software and then deliver it to the Internet, which leads to a serious problem for image content security. In order to deal with the problem, digital image forensics has been studied by a number of scholars. Many algorithms have been proposed focusing on forensic issues of different operations, such as recapture, double JPEG compress, median filter, and copy-move.

As a common and facilitating operation, recapture has drawn much attentions in public security forensics. Recapture operation denotes that the original images are projected on some media, such as computer screen, mobile screen, paper and shot again using a camera. A multimedia content security event related to recapture images should be mentioned. In 2007, a villager lived in Shanxi Province claimed that he found the Wild South China Tiger, a kind of endangered animal, and published several groups of photos containing the Wild


Y.Q. Shi et al. (Eds.): IWDW 2016, LNCS 10082, pp. 119–128, 2017.
DOI: 10.1007/978-3-319-53465-7_9

South China Tiger, which had aroused general concern by Forestry Department of Shanxi Province and the public. However, these photos provided by the villager turned out to be fake and the experts identified that these photos were the recaptured images from a New Year painting.

The key point of recapture image forensics is to discriminate between recaptured and original images. For recapture images obtained from different media, researches have proposed various algorithms based on the differences of image statistical characteristics between the original and the recapture images. For recapture images projected on paper, T.T. Ng et al. [1] discovered that the recapture images showed the mesostructure of the high-frequency spatial variations in the specular component. For recapture images shown on mobile screens, T.T. Ng et al. [2] presented a physical model and extracted a series of features from the model, such as background contextual information, spatial distribution of specularity, surface gradient and so on. For recapture images obtained from the original images shown on computer (LCD) screen, H. Cao et al. [3] proposed three kinds of statistical features to detect good-quality recapture images, namely local binary pattern (LBP), multi-scale wavelet statistics (MSWS), and color features (CF). R.H. Li et al. [4] found the physical traits of recaptured images on LCD screens and proposed two kinds of features: the block effect and blurriness effect due to JPEG compression, screen effect described by wavelet decomposition.

For the case of recaptured Images Shown On Lcd Screen, Despite Of The Good Detect Performance Of The Existing Algorithms, It Is Still A Great Challenge To Detect Small-Size Recapture Images. Because That, In General, There Is Less Information Available For Small-Size Images. In This Work, we propose an effective method based on convolutional neural networks with Laplacian filter layer. Considering that applying the convolutional neural networks model to forensics field directly can not achieve the best detect performance, we put signal enhancement layer into convolutional neural networks structure. Laplacian filter is used in the signal enhancement layer, which strengthens the differences between the original images and the recapture images. We realize the proposed method on four kinds of small-size image databases. The size of the images are 512*512, 256*256, 128*128, 64*64, respectively. The experimental results demonstrate that the proposed method outperforms the Li's [4] method. The rest of the paper is organized as follows: Sect. 2 describes details of our method. Section 3 includes the experimental results and analysis; conclusions are given in Sect. 4.

## 2 Laplacian Convolution Neural Networks

Convolutional Neural Networks (CNNs) have achieved amazingly good performance in computer vision field. From the AlexNet [6] in 2012 to ResNet [7] in 2015, the top-5 error on the ImageNet classification challenge has been decrease from 15% to 3.8%. What's more, CNNs has universality and expansibility. For example, the ResNet in 2015 achieved 3.57% error on the ImageNet test set and won the 1st prize on the ILSVRC 2015 classification task. Besides the classification task, the ResNet also won several 1st places on the tasks of ImageNet detection, ImageNet localization, COCO detection, and COCO segmentation. In addition, it is a remarkable fact that the size of images used in computer vision field is small. For example, the mnist dataset is a dataset of handwritten digits and its images size is $32 \times 32$ only.

Differing from the traditional schemes that extracts feature handcrafted, CNNs automatically learn features and combine feature extraction with feature classification, which make CNNs to get enough information about the image even for small-size image [5]. In general, the unit of Convolutional neural networks structure consists of two parts: feature extraction and classification. For the unit of feature extraction, there are 4 kinds of general layers: convolution layer, pooling layer, ReLU layer, Batch-normalization layer. For the unit of classification, fully connected layer and softmax layer are common layers.

Considering that CNNs model can automatically learn features and use more informations of the images, we utilize the CNNs to create a generalized model for small-size recapture images forensics. However, just putting the CNNs into the problem of recapture image forensics may not be a good way and the experimental results illustrate that it really is. The reason is that, unlike visual recognition task, there is a little difference on two classifications of recapture image forensics. So we add the signal enhancement layer into CNNs structure in order to strengthen the difference.

The architecture of our Laplacian Convolutional Neural Networks model is shown in Fig. 1., which include two parts: signal enhancement layer and general Convolutional Neural Networks structure. Firstly, we put the R, G, B channels of the image into the model. In part of signal enhancement layer, Laplacian filter is executed to amplify the difference between the original images and the recapture images. Then, the five basic units are used in feature extraction layers. Lastly, the softmax layer is used in classification layer.

### 2.1 Signal Enhancement Layer

The recapture operation would introduce some noise into the images. For the case of recaptured images shown on LCD screen, the noise introduced would cause wave effect. We try to stress this effect by detecting edge.

The Laplacian operator of an image finds the regions of rapid intensity change and is often used for edge detection. The Laplacian transform *L(i, j)* of an image $I$ is given by:

$$L(i,j) = \frac{\partial^2 I}{\partial i^2} + \frac{\partial^2 I}{\partial j^2} \tag{1}$$

The Laplacian transform of an image can be achieved using a Laplician filter. The common used Laplician filter is as follows:

$$LF = \begin{bmatrix} 0, & -1, & 0 \\ -1, & 4, & -1 \\ 0, & -1, & 0 \end{bmatrix} \tag{2}$$

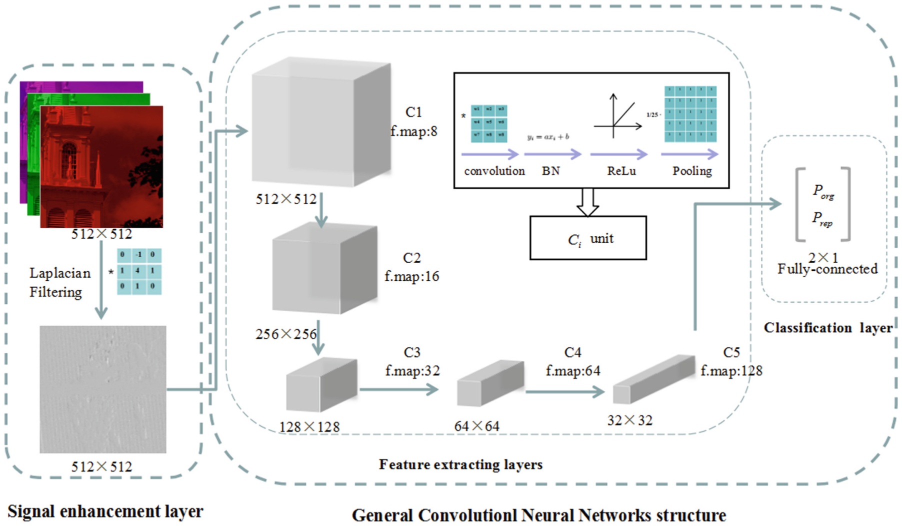


**Fig. 1.** The structure of the proposed method

The Laplacian operator is sensitive to noise, which has an bad influence on edge detection. However, for the detection of recapture images, we consider the attribute of sensitivity quite suitable. Because that after the processing of the Laplacian filter, the wave effect resulted from the recapture operation will be projected and the differences between the original images and the recaptured images will be amplified. To validate the effectiveness of the Laplacian filter, we random select four pairs of images from the image database provided in [4]. The gray versions of the original image and the recapture image after filtering are shown. In Fig. 2, the left column represents the original images and the right column denotes the recapture images. The difference in Fig. 2 between the original images and the recapture images is slight. Subsequently, these images in Fig. 2 are firstly converted to gray images. Then Laplacian filtering is implemented to the gray images, which results in the images in Fig. 3. It is intuitive to find that the difference between the original and recaptured image is amplified through the processing of the Laplacian filtering. The recaptured images after the Laplacian filtering have obvious stripes, which result from the noise introduced by LCD screen during reccapture operation.

## 2.2 General Convolutional Neural Networks Structure

In order to simplify the description, we only show the details of the structure for the case that the size of input images is equal to 512. For the other cases, the same operations will be executed and the difference is the size of input images and feature maps.

The general Convolutional Neural Networks structure consists of two parts: feature extraction layers and classification layer. Feature extraction layers

**Fig. 2.** Images samples. The left column represents the original images and the right column denotes the recapture images

include five basic unit and the softmax layer is applied to classification layer. There are four operations in basic unit: convolution, Batch Normalization, ReLu, and average pooling. The kernels size in convolution layer is 3*3 and step size is 1. In the average pooling layer, the kernels size is 5*5 and step size is 2. In particular, global average pooling is used in average pooling layer of $C_5$. The numbers of feature maps in five basic unit are 8, 16, 32, 64, 128 respectively. The sizes of feature maps are 515*512, 256*256, 128*128, 64*64, 32*32.

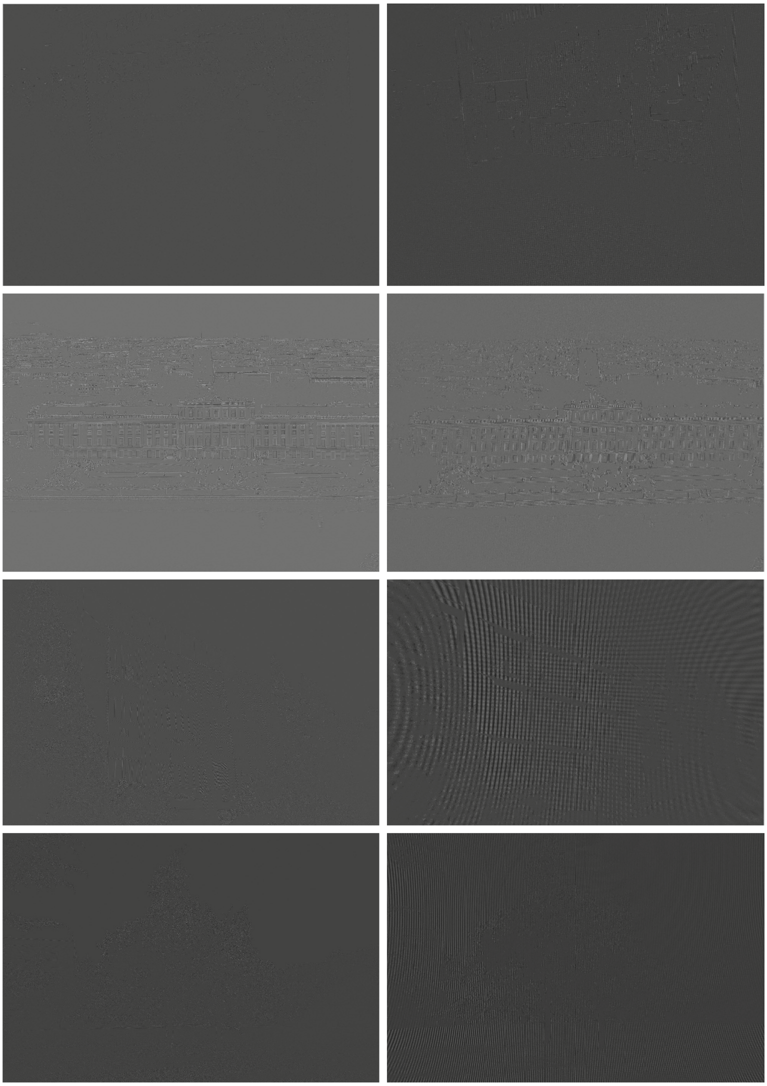

**Fig. 3.** Signal enhancement by Laplacian kernel for the gray version of the images.

## 3 Experimental Results

To evaluate the effectiveness of the proposed algorithms, we construct a image database including 20000 images: 10000 original images and 10000 recapture images. The size of the images is $512 \times 512$. The images derive from the image databases provided in [3,4]. All images from the two image libraries are used. We crop a square with size of 1024*1024 from the center of the images. In particular, if the height or width of the image is less than 1024, we crop the image to 512*1024 or 1024*512. Then the images are cut into non-overlopping blocks of 515*512. We randomly select 10000 original images and 10000 recapture images as the image database used in this work. The training data contains 4000 original images and 4000 recapture images. The validation set consists of 1000 original images and 1000 recapture images. The rest belongs to the test set. We compare the proposed method with the work in [4] and the detection accuracies are averaged over 3 random experiments.

The experiments about the proposed method are conducted on a GPU and performed based on an open source framework of deep leaning: Caffe [8].

### 3.1 Experiment 1

To test the effectiveness of the proposed method for small-size recapture images, the images in the image database that we build are cropped to 256, 128, 64 from the center of the image respectively. The experiments are conducted for four different sizes of images. The results of the experiments are given in Table 1. As shown in Table 1, the detection accuracies of proposed method are higher than LBP method and Li's method in all cases.

**Table 1.** The Detection accuracy (%) for small-size recapture images. The best results are highlighted in bold.

| Image size | Method | Accuracy (%) |
|---|---|---|
| $512 \times 512$ | LBP [3] | 95.02 |
| | Li [4] | 91.9 |
| | Proposed method | **99.74** |
| $256 \times 256$ | LBP [3] | 92.54 |
| | Li [4] | 90.55 |
| | Proposed method | **99.30** |
| $128 \times 128$ | LBP [3] | 89.01 |
| | Li [4] | 88.49 |
| | Proposed method | **98.48** |
| $64 \times 64$ | LBP [3] | 79.5 |
| | Li [4] | 85.41 |
| | Proposed method | **95.23** |

### 3.2 Experiment 2

To verify the effectiveness of the signal enhancement layer, we test the network structure without the signal enhancement layer for th case that image size is equal to $64 \times 64$. The experiment results in Table 2 show that the detection accuary of proposed method using the signal enhancement layer is higher than the method dropped out the signal enhancement layer (in the case of NON in Table 2).

In the other hand, the filter used in the signal enhancement layer has an impact on the detection accuracy. We test five kinds of common filters, namely HF, VF, HVF, GF, LF. The definitions of these filter are as follows.

$$HF = \begin{bmatrix} 1, -1 \end{bmatrix} \tag{3}$$

$$VF = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{4}$$

$$HVF = \begin{bmatrix} 1, & -1 \\ -1, & 1 \end{bmatrix} \tag{5}$$

$$GF = \begin{bmatrix} -1, & 2, & -2, & 2, & -1 \\ 2, & -6, & 8, & -6, & 2 \\ -2, & 8, & -12, & 8, & -2 \\ 2, & -6, & 8, & -6, & 2 \\ -1, & 2, & -2, & 2, & -1 \end{bmatrix} \tag{6}$$

$$LF = \begin{bmatrix} 0, & -1, & 0 \\ -1, & 4, & -1 \\ 0, & -1, & 0 \end{bmatrix} \tag{7}$$

In Fig. 4, the images after different filtering operations are shown. From the second row to the last row, the filter used is HF, VF, HVF, GF, LF respectively. In order to highlight the effect introducted by filtering operation, we randomly chose one image patch. Then it is magnified and displayed in Fig. 4. These five filters both enhence the difference between the original image and the recapture image. The results in Table 2 indicate that the case of LF (using the Laplacian filter) has the best detection accuracy. It is clear that using Laplacian filter in signal enhencement layer is a good choise for recapture image forensics.

**Table 2.** The Detection accuracy (%) for different choise in signal enhencement layer. The best result are highlighted in bold

| Image size | The choise in signal enhencement | Accuracy (%) |
|---|---|---|
| $64 \times 64$ | NON | 91.23 |
| | HF | 93.98 |
| | VF | 93.47 |
| | HVF | 93.78 |
| | GF | 92.61 |
| | LF | **95.23** |

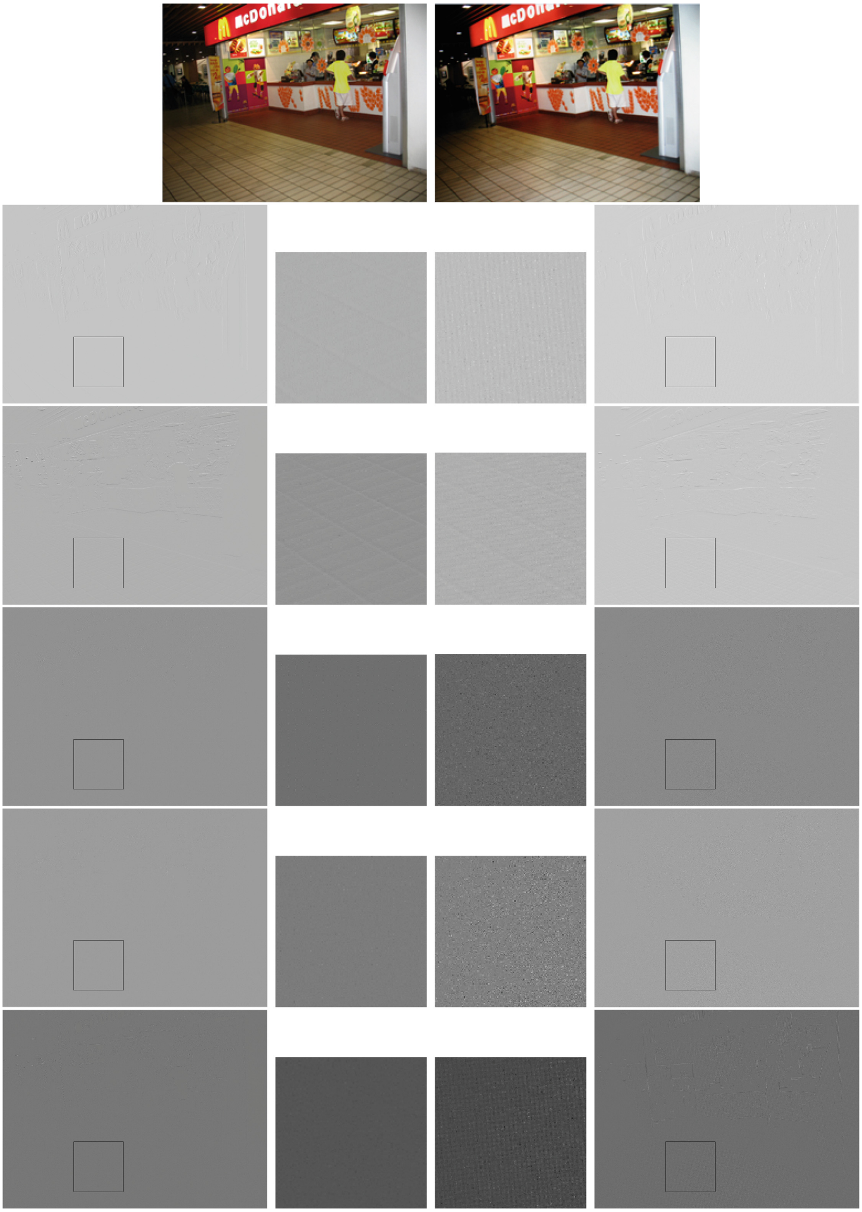

**Fig. 4.** The images are processed by using kinds of filters for the gray version

## 4 Conclusions

In this paper, we propose a generalized model for small-size recapture image forensics based on Laplacian Convolutional Neural Networks. In particular, the signal enhencement layer was added into Convolutional Neural Networks structure and Laplician filter was applied to signal enhencement layer. The Laplician filter used in signal enhencement layer plays a important role for the improvement of the detection performance. The experimental results has proven that the detect performance of our method is great.

**Acknowledgments.** This work was supported in part by National NSF of China (61332012, 61272355, 61672090), Fundamental Research Funds for the Central Universities (2015JBZ002), the PAPD, the CICAEET. We greatly acknowledge the support of NVIDIA Corporation with the donation of the Tesla K40 GPU used for this research.

## References

1. Yu, H., Ng, T.T., Sun, Q.: Recapture photo detection using specularity distribution. In: 15th IEEE International Conference on Image Processing, pp. 3140–3143 (2008)
2. Gao, X., Ng, T.T., Qiu, B., Chang, S.F.: Single-view recaptured image detection based on physics-based features. In: 2010 IEEE International Conference on Multimedia and Expo (ICME), pp. 1469–1474. IEEE (2010)
3. Cao, H., Alex, K.C.: Identification of recaptured photographs on LCD screens. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1790–1793 (2010)
4. Li, R., Ni, R., Zhao, Y.: An effective detection method based on physical traits of recaptured images on LCD screens. In: Shi, Y.-Q., Kim, H.J., Pérez-González, F., Echizen, I. (eds.) IWDW 2015. LNCS, vol. 9569, pp. 107–116. Springer, Heidelberg (2016). doi:10.1007/978-3-319-31960-5_10
5. Chen, J., Kang, X., Liu, Y., Wang, Z.J.: Median filtering forensics based on convolutional neural networks. IEEE Signal Process. Lett. **22**(11), 1849–1853 (2015)
6. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385 (2015)
8. An open source framework of deep learning. http://caffe.berkeleyvision.org/