 CMC, vol.60, no.1, pp.287-303, 2019

# Improved Fully Convolutional Network for Digital Image Region Forgery Detection

**Jiwei Zhang[1], Yueying Li[2], Shaozhang Niu[1,*], Zhiyi Cao[1] and Xinyi Wang[1]**

**Abstract:** With the rapid development of image editing techniques, the image splicing behavior, typically for those that involve copying a portion from one original image into another targeted image, has become one of the most prevalent challenges in our society. The existing algorithms relying on hand-crafted features can be used to detect image splicing but unfortunately lack precise location information of the tampered region. On the basis of changing the classifications of fully convolutional network (FCN), here we proposed an improved FCN that enables locating the spliced region. Specifically, we first insert the original images into the training dataset that contains tampered images forming positive and negative samples and then set the ground truth masks of the original images to be black images. The purpose of forming positive and negative samples is to guide the improved FCN to distinguish the differences between the original images and spliced images. After these steps, we conducted an experiment to verify our proposal, and the results reveal that the improved FCN really can locate the spliced region. In addition, the improved FCN achieves improved performance compared to the already-existing algorithms, thereby providing a feasible approach for digital image region forgery detection.

**Keywords:** Improved FCN, spliced region localization, splicing forgery, image splicing.

## 1 Introduction

With the development of the Internet, a variety of web images that come from cameras, smart phones and tablets are entering into our daily life [Zampoglou, Papadopoulos and Kompatsiaris (2017)]. Re-inspecting into the past decades, the rapid development of image editing techniques has served as a convenience for us to record and improve the quality of moments. However, some photo editing software, such as Photoshop and Beauty Camera, can also bring image forgery, leading to a large number of tampered images around common social media (e.g., Google, Baidu, Twitter, and Facebook). Image splicing [Asghar, Habib and Hussain (2017)], i.e., frequent implementation of images by copying one region of the original image and then pasting it onto another region of the targeted image, is now accepted as one of the most common image

[1] Beijing Key Lab of Intelligent Telecommunication Software and Multimedia, Beijing University of Posts and Telecommunications, Beijing, 100876, China.

[2] College of Arts and Sciences, Boston University, Boston University, Boston, 02215, USA.

[*] Corresponding Author: Shaozhang Niu. Email: szniu@bupt.edu.cn.

tampering behaviors in our daily life. Commonly, the tampered images that come from the Internet involve economic, political, media, military, technology, medical, and judicial fields [Wei, Wang and Ma (2017)] and thus can give a very bad influence on the country and society. In this case, the image splicing detection techniques are of significant scientific importance, but they are traditionally concentrated on using the methods of pattern noise [Siwei, Xunyu and Xing (2014); Yao, Wang, Zhang et al. (2017)], color filter array (CFA) [Ferrara, Bianchi, De Rosa et al. (2012); Varlamova and Kuznetsov (2017)] and blocking [Bianchi and Piva (2012), Bianchi, De Rosa and Piva (2011)] to detect, which have some drawbacks because they need some prior information and can only handle a certain type of forgery. Moreover, the current image splicing detection technologies can only solve whether an image has undergone splicing, without the ability to locate the tampered region. Hence, there are still few algorithms that can simultaneously solve the spliced region localization problem.

In this paper, we propose an improved fully convolutional network (FCN) method, which can be used to locate the spliced region of digital images. Although the related algorithms based on the deep learning technology [Pomari, Ruppert, Rezende et al. (2018)] require a certain amount of time for training, the detection speed is very fast once the model is trained. Since the FCN [Long, Shelhamer and Darrell (2015)] features the capability of precise regional learning, the regional learning ability of FCN is thus used for locating the spliced region and eventually detecting the image tamper. When there is not a person in the non-spliced region of the digital image, the FCN can detect the spliced region [Sundaram and Nandini (2017)] correctly. However, misdetection of the FCN will appear once the non-spliced region involves various people. As can be seen from Fig. 1, there is still a certain difference between the tampered person and the person in the non-spliced region. As a result, this can drive us to distinguish between tampered and non-tampered regions. With these considerations, the improved FCN, which can change the classifications of the original FCN, is also proposed here to locate the spliced region. After that, we train the improved FCN on the training dataset because people of the spliced region are more easily detected than the non-spliced area. We add the original image to the training dataset forming positive and negative samples to increase the difference between the person of the spliced region and the non-spliced area. So the improved FCN can also learn this edge feature of the spliced region rather than just the outline of the spliced region, thereby allowing the improved FCN to have the ability of learning the outline of the spliced region and distinguishing the edges between the tampered and non-spliced regions. This detection capability of locating the spliced region by using the improved FCN is superior to the traditional algorithms.

Our contributions of this work are as follows: (a) The improved FCN method can process a much wider scope of images compared to the existing methods that are based on the handcrafted features; (b) The accuracy of locating the spliced region can be improved significantly; (c) Three different improved FCNs after comparing to the solo FCN have been proved to improve the detection accuracy; and (d) This work can contribute to a person-based database.

## 2 Survey of previous related works

In terms of image tampering identification technologies, several already-reported technologies have become available, but they can only detect the image tampering under certain circumstances, showing some limitations and inadequacies. For example, the traces left by JPEG compression are often used for image tampering identification, and the related methods [Chang, Yu and Chang (2013); Wang, Dong and Tan (2014)] are mainly based on using the quality factor and discontinuity of the JPEG compression grid. As for JPEG quantization, it needs to assume that the original image has undergone a continuous JPEG compression and the spliced region may have lost its original JPEG compression feature due to smoothing and/or resampling. With these inconsistent features, we can detect the location of the spliced area. However, such a technique is only applicable to the JPEG format. On the basis of a fixed imaging device, any subsequent post-processing and compression of an image can produce different unique noises between the images. As a result, the algorithms based on noise pattern [Chierchia, Poggi, Sansone et al. (2014); Pun, Liu and Yuan (2016)] can be used for image tampering identification. In addition, image sensors, which are generally based on some certain modes, can also be used to acquire the image data, and the CFA interpolation process is the most common mode. Moreover, the CFA interpolation process can be used for image tamper identification because the splicing behavior of the image can lead to the destruction of the CFA pattern. Different cameras correspond to different CFA interpolation algorithms, while different image splicing and image scaling may cause discontinuities. In this case, the algorithms based on the CFA interpolation mode [Dirik and Memon (2009)] have also been used for image tamper identification, but they are limited to specific assumptions such as fixed imaging equipment and processing steps.

The algorithms based on the neural network were found in the past few years to have the ability to change the landscape of computer vision significantly. Among the reported literature investigations concerning image tamper identification, some of them are focused mainly on using deep learning and neural network techniques, but they still suffer from some problems. For example, literature investigations [Bayar and Stamm (2016); Flenner, Peterson, Bunk et al. (2018); Cui, McIntosh and Sun (2018)] have applied deep learning techniques for image tamper identification, which were found to have the ability to solve the single tampering problem, but they could not solve the problem of detecting image splicing behavior. Furthermore, some methods were proposed in the literature [Cozzolino and Verdoliva (2016); Bondi, Lameri, Güera et al. (2017)] to solve the image splicing problem, but they were based on some certain assumptions and thus greatly reduced the general applicability of the algorithms. Moreover, the literature [Rao and Ni (2016); Liu, Guan, Zhao et al. (2017)] has proposed an identification method for image splicing, but it did not realize the location of the tampering region. Moreover, the literature Zhang et al. [Zhang, Goh, Win et al. (2016)] proposed a two-stage deep learning method to learn the image corresponding block features to detect tampered images of different image formats. For the first phase, we used the stacked autoencoder model to learn the complex features of each individual patch. For the second phase, we integrated the context information for each patch so that it can be detected more accurately. However, due to the blocking reason, the method has a problem that the complexity of the corresponding algorithm is too high and the

detection speed is slow. The algorithms reported by the literature Salloum et al. [Salloum, Ren and Kuo (2018)] were proposed for solving the problems of image spliced area localizations, but the detection effect of the algorithms for spliced region detection is not perfect, which actually can be optimized by using the improved FCN we propose here. Although the literature Long et al. [Long, Shelhamer and Darrell (2015)] indicates that FCN features precise regional learning capability, it still has some problems for locating image splicing forgery. Specifically, this can be seen from a spliced image shown in Fig. 1. The person on the left side of the image is the spliced person, and the person on the right is the person in the original image (Fig. 1(a)). In Fig. 1(b), we can see an obvious error from the result of the FCN algorithm detection.

(a) (b)

**Figure 1:** (a) A spliced image in which the person on the left side is a spliced person while the person on the right is the person in the original image; (b) The detection results of the FCN

## 3 Proposed method

As for a spliced image without prior knowledge, it is actually difficult to tell which feature belongs to the forgery, and a certain hand-crafted feature can only handle one kind of splicing forgery. Moreover, the image splicing detection algorithms based on the deep learning technique also have some drawbacks, and the algorithms for spliced region detection can be further optimized. One of the key features of the image spliced region is that it exhibits distinguishing features on the boundaries of the non-spliced region and the spliced region. So we use an improved FCN model to capture the distinguishing features between the spliced and non-spliced regions.

### *3.1 Overview of the fully convolutional network (FCN)*

The FCN can perform end-to-end and pixel-by-pixel training, allowing for image segmentation without any preprocessing. It is based mainly on several kinds of structures, such as VGGNet, AlexNet, GoogLeNet, etc. The operation of the convolution layer

corresponding to the above CNN network is as follows.

$$x_j^l = f(\sum_{i \in M_j} k_{ij}^l * x_i^{l-1} + b_j^l) \quad (1)$$

where $l$ represents the number of layers, $k$ is the convolution kernel, $M_j$ represents a choice of the input feature map, and each output map has an offset $b$. The CNN network output layer outputs features, followed by a softmax classifier to distinguish the characteristics of different objects. The classifier is expressed as:

$$L_i = -log(\frac{e^{f_{y_i}}}{\sum e^{f_j}}) \quad (2)$$

where $f$ represents the score and $f_j$ represents the jth element of the score vector $f$, in the case where $L_i$ represents the average losses of all training samples when considering a regular term, and the regular term is used to avoid over-fitting. In addition, Fig. 2 shows the structure of a full convolutional neural network (CNN) being used for semantic segmentation.

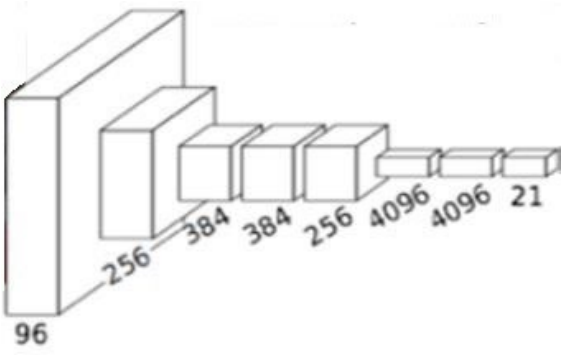


**Figure 2:** The network of the FCN

The five convolutional layers in front of the FCN are the same as the CNN, but the last three layers are replaced by convolutional layers. The specific process of the FCN is an up-sampling process. We can assume that the image size is $W_1*W_1$ in the convolutional network of a certain layer, the image size after up-sampling is $W_2*W_2$, the convolution kernel size is $K$, the step length is $S$, and the edge complement is $P$. The formula for the deconvolution operation is:

$$W_2 = (W_1 - 1) * S + K - 2P \quad (3)$$

where the up-sampling can be further understood, that is, the image with a size of 512*512 is first reduced to $W_1*W_1$, and then an operation is performed to obtain $W_2*W_2$, which corresponds to a deconvolution operation. Differing from one-dimensional output after being fully connected by layers, the output of the convolutional layers is a two-

dimensional matrix, providing the possibility of pixel-level predictions. The FCN uses the ground truth masks of the image that are manually labeled for monitoring the information to train an end-to-end network, thereby allowing the network to make the pixel-level predictions and eventually generate the label images.

As can be seen from Fig. 3, the FCN network has very accurate object contour learning capabilities. As a result, the precise outline learning ability of the FCN makes it possible for the contour learning ability of the FCN to be used for locating the spliced region.

**Figure 3:** (a) Original image; (b) The output of the FCN

### *3.2 Proof of concept for locating the spliced region of digital images by the improved FCN*

Due to the performance of the precise region learning ability of the FCN, we consider that we can improve the existing FCN for image splicing location, which can solve the location of the tampered region. However, as can be seen from Fig. 1, when the non-spliced area of the image contains several people, the FCN is proved to exhibit false detection. Hence, the original FCN cannot solve the key problems of locating the spliced region.

We copy different people into the targeted images from the CASIA v2.0 dataset to form spliced images, and the ground truth masks corresponding to the spliced images were made by Photoshop CC 2015. The spliced region of the ground truth mask was made as red, and the other region was made as black. In order to increase the differences between the people of the tampered and non-tampered regions, we then added the original images that contained the person in the training database, which contains the tampered images forming positive and negative samples. The ground truth mask that corresponds to the original image is shown in Fig. 4. Obviously, it is a completely black image. The positive and negative samples can let the improved FCN learn the difference between the spliced region and the non-spliced area. So the improved FCN network will learn the differences between the spliced region and non-spliced region when the improved FCN is trained.

**Figure 4:** (a) Original image containing the person; (b) Ground truth of the original image

Specifically, to achieve the above target, we first processed the input image to a uniform size of 512*512. According to the VGG-16 configuration, the improved FCN had five max-pooling operations. The height and width of the feature after each max-pooling operation were halved. Thus, the final height and width were 1/32 of the input image. As a result, the improved FCN has the up-sampling operation, which can increase the size of the feature. As a result, applying the 32 x up-sampling operation to the last layer of the VGG network can allow the width and height of the prediction image to be restored to the size of the input image. The FCN 32 x network is also called the 32 x up-sampling operation, and the FCN16 and FCN8 networks are used as different up-sampling filters. As we know, the classification number of the original FCN is 21, which is mainly used to classify the 21 categories of the targeted objects. It is shown in Fig. 6 that we can change the classifications of the FCN. The improved FCN can be used to distinguish between the tampered region and non-tampered region.

There are the spliced region and non-spliced region for a spliced image, defining the spliced region as $R_f$ and the non-spliced region as $R_n$ . All pixel values of the pixel in the spliced region are expressed as:

$$X = [x_1, ..., x_i] \tag{4}$$

where $x_i$ represents the pixel value in an arbitrary position of the spliced region when $i \in R_f$, and $i, j, k$ are the position of the pixel in the image. All pixel values in the non-spliced region are defined as follows:

$$Y = [y_1, ..., y_j] \tag{5}$$

where $y_j$ represents the pixel value of any position in the non-spliced region when $i \in R_f$, and $M_k$ represents the pixel value of the image at any position when $k \in R_f \bigcup R_n$:

$$M_k = \begin{cases} 0 & k \in R_f \\ 1 & k \in R_n \end{cases} \tag{6}$$

Through the distinction between the spliced region and non-spliced region, we can transform the image segmentation into image tampering identification. Moreover, the improved FCN was trained by the positive and negative samples, so that the training improved FCN had the ability to distinguish between the spliced region and non-spliced region.

In summary, the corresponding identification framework is shown in Fig. 5:

- Copied different people from original images into targeted images to form spliced images, and the ground truth masks corresponding to the spliced images were made by Photoshop CC 2015. Moreover, the spliced region of the ground truth mask was made as red, and the other region was made into black.
- Took 1828 spliced images forming the training dataset, and added the original images to the training dataset, which contained the tampered images forming positive and negative samples.
- Improved the original FCN and changed the classification of FCN to 2 for distinguishing between spliced and non-spliced regions, and not just for image segmentation.
- The improved FCN was trained with the training dataset with the original images added, and the training was performed by the SGD algorithm. We initialized the weights and parameters of FCN32 by the weights of the VGG-16 model pretrained on the ImageNet dataset [Russakovsky, Deng, Su et al. (2015)]. The trained FCN32 and FCN16 models were used as an initialization parameter to train FCN16 and FCN8, respectively.
- Detected the image of the test dataset without the corresponding ground truth masks through the trained model, and obtained the detection result.

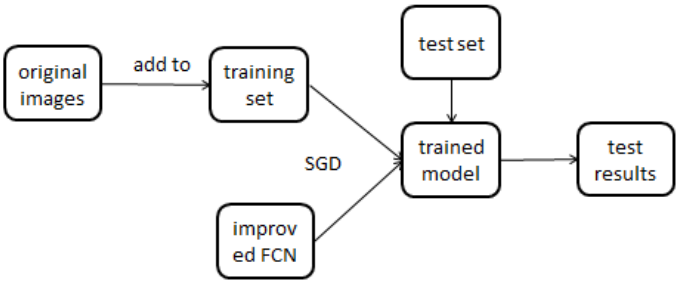


**Figure 5:** Training and testing process of the improved FCN

As can be seen from Fig. 6, the person on the left of the input image is spliced from another image, and the trained FCN network outputs the red region corresponding to the spliced region. From the output of the improved FCN of Fig. 6, the spliced and non-spliced regions have been marked in red and black, respectively. Obviously, the improved FCN has obtained a precise localization of the spliced region compared with that of the solo FCN.

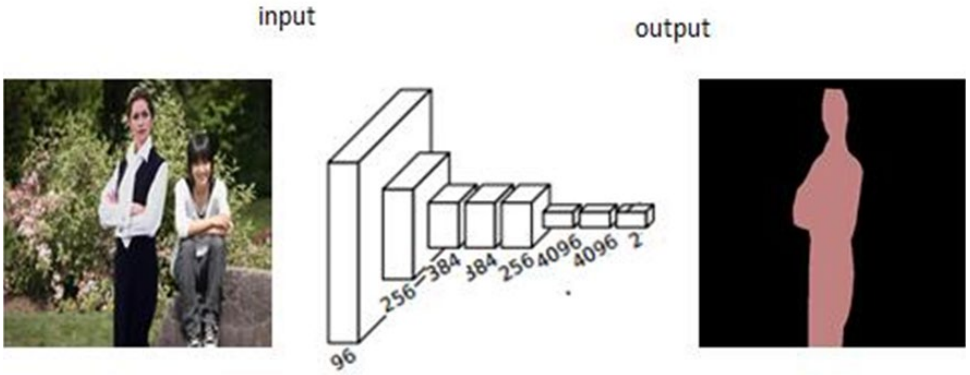


**Figure 6:** Extraction of the tampered region by the improved FCN

### *3.3 Training and testing*

When a normal training is used, the network parameters of the convolutional layers should be initialized to random values that belong to a normal distribution. This needs to take a much longer training time compared to the transfer training. So the transfer training is employed in our experiment. In the transfer learning policy network, the parameters of the pre-trained VGG16 network are used directly, which can speed up the process of network convergence. More specifically, the training of the improved FCN was performed in Caffe [Jia, Shelhamer, Donahue et al. (2014)] using the stochastic gradient descent (SGD) algorithm, with a fixed learning rate of 1e-14, a momentum of 0.99, and a weight decay of 0.0005. We initialized the weights and parameters of FCN32 by the weights of a VGG-16 model pretrained on the ImageNet dataset. The trained FCN32 and FCN16 models were used as an initialization parameter to train the FCN16 and FCN8 networks, respectively.

## 4 Experimental discussion and results

In this section, we trained the improved FCN using the training dataset, and we present our experimental results pertaining to the performance evaluation of the improved FCN, as well as its comparison with the state of the art.

### *4.1 Dataset and evaluation criteria*

We copy different people into the targeted images from the CASIA v2.0 dataset to form spliced images, and the ground truth masks corresponding to the spliced images were made by Photoshop CC 2015. In addition, the spliced region of the ground truth mask was made as red, and the other region was made as black.

The training process is implemented on an Intel Core i7 processor, with 16 GB GPU, and using Caffe. The number of images in the training dataset is 1828; in addition, we also tested the performance of the improved FCN on the test dataset, which included 400 test images. The training dataset contained spliced images that correspond to 20 people in the CASIA v2.0 database. Moreover, the people cover different genders and age groups.

There are 30 to 199 spliced images for each person. Moreover, we spliced the corresponding people to various scenarios, including plant images, architectural images, natural landscape images and images that contain several people. Furthermore, 20 original images containing the person were also added to the training dataset. The dataset used to support this study is available from the corresponding author upon request. Fig. 7 exhibits the scheme of the database.

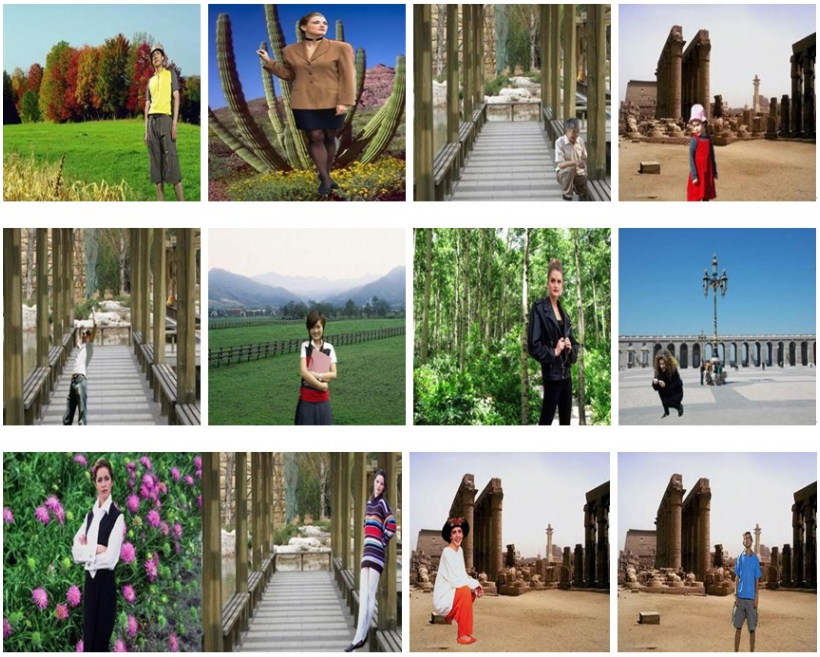

**Figure 7:** Examples of the dataset

We evaluated the performance of the improved FCN model and compared it with the existing algorithms. The evaluation standard is $F_1$ and Matthews Correlation Coefficient (*MCC*) metrics [Salloum, Ren and Kuo (2018)], which are per-pixel localization metrics. Calculating the results of the trained network output and the corresponding ground truth can allow us to obtain $F_1$ and *MCC* metrics. The $F_1$ metric is defined as below:

$$F_1(M_{out}, M_{gt}) = 2TP/(2TP + FN + FP) \tag{7}$$

$M_{out}$ represents the result of the network output, $M_{gt}$ represents the ground truth, *TP* represents the number of pixels classified as true positive where a spliced pixel is correctly classified as spliced, *FN* represents the number of pixels classified as false negative where a spliced pixel is incorrectly classified as authentic, and *FP* represents the number of pixels classified as false positive where an authentic pixel is incorrectly classified as spliced. For a given spliced image, the *MCC* metric is defined as

$$MCC(M_{out,}M_{gt}) = \frac{TP * TN - TP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{8}$$

### *4. 2 Performance comparison*

We trained the improved FCN32, FCN16, and FCN8 models on the training dataset and used the trained models to detect the images of the test dataset. The number of test sets is 500, and the test images that are tampered do not suffer any postprocessing operation. In the process of testing, the test image does not have a corresponding label map. For each model, we computed the average $F_1$ and *MCC* scores across the test dataset. The test results listed in Tab. 1 show that the improved FCN16 and FCN32 have achieved precise detection results for locating the spliced region of the person. In addition, the detection results of the improved FCN8 are superior to that of the improved FCN16 and FCN32 according to both the $F_1$ and *MCC* values, and the following experiments also use the improved FCN8.

**Table 1:** Average $F_1$ and *MCC* scores of the improved FCN8, improved FCN16 and improved FCN32

| Network | Improved FCN8 | Improved FCN16 | Improved FCN32 |
|---|---|---|---|
| $F_1$ score | **0.8473** | 0.8258 | 0.8240 |
| *MCC* score | **0.8527** | 0.8336 | 0.8309 |

Our improved FCN is compared with the existing baseline methods, and these existing baseline methods are derived from ADQ2 [Bianchi, De Rosa and Piva (2011)], NADQ [Bianchi and Piva (2012)], BLK [Li, Yuan and Yu (2009)], CFA1 [Ferrara, Bianchi, De Rosa et al. (2012)], CFA2 [Dirik and Memon (2009)], DCT [Ye, Sun and Chang (2007)], ELA [Zampoglou, Papadopoulos and Kompatsiaris (2017)], NOI1 [Mahdian and Saic (2009)], and MFCN [Salloum, Ren and Kuo (2018)]. The implementation of these existing algorithms is provided in a publicly available Matlab toolbox as described by Zampoglou et al. [Zampoglou, Papadopoulos and Kompatsiaris (2017)]. For each method, we calculated the average $F_1$ and *MCC* scores according to the evaluation criteria, and the results are listed in Tab. 2. Obviously, the proposed method outperforms the existing baseline methods in terms of both $F_1$ and *MCC* scores. When there are many people or individuals in the non-spliced region of the spliced image, the traditional FCN will feature an error detection in the non-spliced region of the image, and it cannot achieve the effect of tampering identification (Fig. 1). Then, we tested the effect on the test dataset. After that, we detected them on the trained improved FCN (Fig. 8). Just a part of the experimental results is shown in Fig. 8. Regardless of whether the non-spliced region contains multiple people or a single person, the improved FCN shows a precise effect on the spliced image.

**Table 2:** Average $F_1$ and *MCC* scores of the baseline methods

| Method | This | MFCN | NOI1 | DCT | CFA2 | BLK | ELA | CFA1 | ADQ2 | NADQ |
|---|---|---|---|---|---|---|---|---|---|---|
| $F_1$ | **0.7468** | 0.5410 | 0.2633 | 0.3005 | 0.2125 | 0.2312 | 0.2136 | 0.2073 | 0.3359 | 0.1763 |
| *MCC* | **0.7608** | 0.5201 | 0.2322 | 0.2516 | 0.1615 | 0.1769 | 0.1337 | 0.1521 | 0.3000 | 0.0987 |

**Figure 8:** Detection results of the spliced image that contains people in the non-spliced region. The top row is the spliced images containing different people in the non-spliced region. The second row is the original images containing the people. The bottom row is the related detection results. It can be seen that the improved FCN can achieve a better localization

### *4.3 Robustness test*

In addition to the image splicing forgery without post-processing operations, the detection of tampered images that are attacked by some post-processing operations is also considered in the proposed scheme. Therefore, a series of experiments have been done to analyze the performance of tampered images that are attacked by some post-processing operations. In order to quantitatively evaluate the robustness of the improved FCN and analyze its ability to resist different image distortions, 200 tampered images are selected from the database, and these tampered images are distorted by different kinds of attacks.

#### *4.3.1 Detection of geometric transforms*

Here, we evaluate the improved FCN for the detection of geometric transforms of rescaling on tampered images. We rescaled the tampered images with the scale factors of [0.5, 0.65, 0.85, 1.05, 1.15]. Tab. 3 presents the detection results for the tampered images attacked by some post-processing operations with different scale factors. When the scale factor of geometric transforms is 0.65, both the $F_1$ and *MCC* scores are the highest, and the $F_1$ and *MCC* scores are the lowest when the scale factor of geometric transforms is 1.15

**Table 3:** Average $F_1$ and *MCC* scores of the improved FCN when we rescaled the tampered images with scale factors. For each result, we highlight in bold the top-performing score

| Scale | 0. 5 | 0.65 | 0.85 | 1.05 | 1.15 |
|---|---|---|---|---|---|
| $F_1$ score | 0.6433 | **0.7165** | 0.7038 | 0.6138 | 0.5357 |
| *MCC* score | 0.6548 | **0.7299** | 0.7245 | 0.6514 | 0.5897 |

*4.3.2 Robustness to additive noise*

In addition to geometric attacks, we evaluate the efficiency of the improved FCN in terms of the detection of other post-processing attacks, viz., the addition of salt and pepper noise and the Gaussian blur of tampered images. In order to consider the spliced image subjected to noise attacks, we added salt and pepper noise to the tampered image to verify the test effect. Then, we used the tampered image for this experiment, filtering by using a signal-to-noise ratio that contains five standard deviation values (i.e., in terms of pixels, a=0.01, 0.007, 0.005, 0.003, and 0.001). As Tab. 4 reveals, when the images are added with the salt and pepper noise, the $F_1$ and *MCC* scores of the proposed methods under different parameters can show a slight degradation, but the performance of the improved FCN is better than the baseline methods.

**Table 4:** Average $F_1$ and *MCC* scores of the improved FCN with different ratios of the salt and pepper noise. For each result, we highlight in bold the top-performing method

| Noise | $a$ = 0.01 | $a$=0.007 | $a$=0.005 | $a$=0.003 | $a$=0.001 |
|---|---|---|---|---|---|
| $F_1$ score | 0.4181 | 0.4972 | 0.5620 | 0.5747 | **0.6862** |
| *MCC* score | 0.5048 | 0.5603 | 0.5994 | 0.6107 | **0.7097** |

*4.3.3 Robustness to additive Gaussian blur*

For this experiment, we performed the blurring of 200 tampered images, with size filters of 3×3, 5×5, 7×7 and 9×9. Tab. 5 shows the performance evaluation results of the improved FCN in terms of robustness to blur attack. From Tab. 5, it is evident that with the increase in filter size, both the $F_1$ and *MCC* scores of the proposed method decrease.

**Table 5:** Average $F_1$ and *MCC* scores of robustness to blurring in tampered images. For each result, we highlight in bold the top-performing method. The Gaussian noise is under different filters

| Filter | 3×3 | 5×5 | 7×7 | 9×9 |
|---|---|---|---|---|
| $F_1$ score | **0.65196** | 0.65121 | 0.65122 | 0.65122 |
| *MCC* score | **0.68110** | 0.68054 | 0.68055 | 0.68055 |

***4.4 Performance on copy-move images***

In this section, we analyze the $F_1$ and *MCC* scores of the proposed technique for copy-move forgery detection. The performance of the proposed method is presented in Fig. 9. Regardless of copy-move forgery or image splicing, the improved FCN shows a good

detection effect.

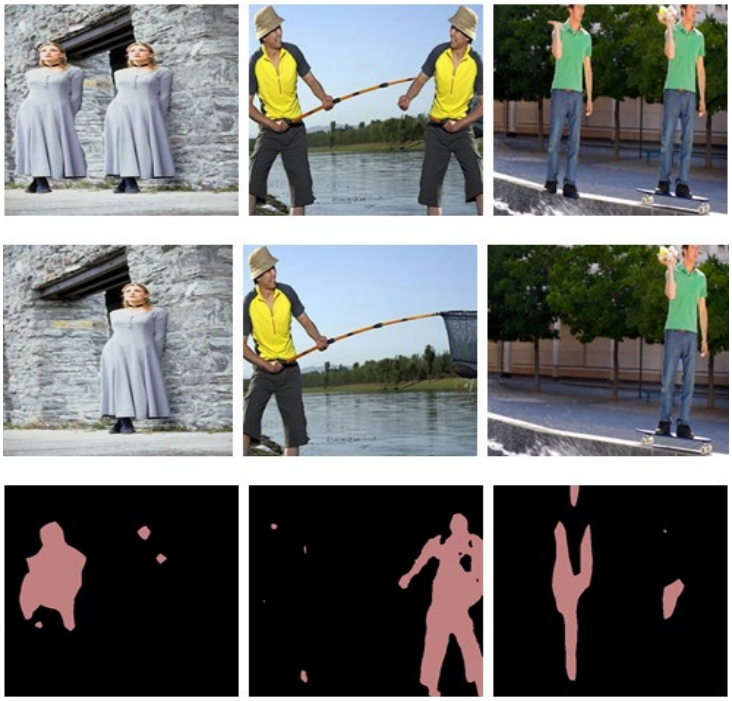

**Figure 9:** Detection results of the copy-move images. The top row is the copy-move images. The second row is the original images. The bottom row is the related detection results. It can be seen that the improved FCN can achieve a better localization

## 5 Conclusions

In this work, we proposed an improved FCN that can locate the image spliced region by changing the classifications of the original FCN. Inspired by the regional learning ability of the FCN, the original images were added to the training dataset, and the ground truth masks that corresponded to the original images were the black images. Our experimental results showed that the proposed improved FCN for locating the spliced region achieved an effect that was better than the existing algorithms on our database. Solving the person-based tamper identification problem more accurately was found to be one of the advantages of our framework. The detection results of people in different postures were also proved to be excellent. When the image contained a person in the non-spliced region, the improved FCN could also achieve a better effect compared to the solo FCN. The improved FCN was proved to have the ability to learn the outline for the spliced region and thus the ability to distinguish between the edges of the tampered and non-tampered regions.

**Acknowledgement:** This work is supported by The National Natural Science Foundation of China (No. 61370195) and the Joint Funds of the National Natural Science Foundation of China (No. U1536121).

**References**

**Asghar, K.; Habib, Z.; Hussain, M.** (2017): Copy-move and splicing image forgery detection and localization techniques: a review. *Australian Journal of Forensic Sciences,* vol. 49, no. 3, pp. 281-307.

**Bianchi, T.; Piva, A.** (2012): Image forgery localization via block-grained analysis of JPEG artifacts. *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 1003-1017.

**Bianchi, T.; De Rosa, A.; Piva, A.** (2011): Improved DCT coefficient analysis for forgery localization in JPEG images. *International Conference on Acoustics, Speech and Signal Processing*, pp. 2444-2447.

**Bayar, B.; Stamm, M. C.** (2016): A deep learning approach to universal image manipulation detection using a new convolutional layer. *Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security*, pp. 5-10.

**Bondi, L.; Lameri, S; Güera, D.; Bestagini, P.; Delp, E. J. et al.** (2017): Tampering detection and localization through clustering of camera-based CNN features. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1855-1864.

**Cui, Q.; McIntosh, S.; Sun, H.** (2018): Identifying materials of photographic images and photorealistic computer generated graphics based on deep CNNs. *Computers, Materials & Continua*, vol. 55, no. 2, pp. 229-241.

**Chang, I. C.; Yu, J. C.; Chang, C. C.** (2013): A forgery detection algorithm for exemplar-based inpainting images using multi-region relation. *Image and Vision Computing*, vol. 31, no. 1, pp. 57-71.

**Chierchia, G.; Poggi, G.; Sansone, C.; Verdoliva, L.** (2014): A Bayesian-MRF approach for PRNU-based image forgery detection. *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 4, pp.554-567.

**Cozzolino, D.; Verdoliva, L.** (2016): Single-image splicing localization through autoencoder-based anomaly detection. *IEEE International Workshop on Information Forensics and Security*, pp. 1-6.

**Dong, J.; Wang, W.; Tan, T.** (2013): Casia image tampering detection evaluation database. *China Summit and International Conference on Signal and Information Processing*, pp. 422-426.

**Dirik, A. E.; Memon, N.** (2009): Image tamper detection based on demosaicing artifacts. *International Conference on Image Processing*, pp. 1497-1500.

**Flenner, A.; Peterson, L.; Bunk, J.; Mohammed, T. M.; Nataraj, L. et al.** (2018): Resampling forgery detection using deep learning and a-contrario analysis. *Electronic Imaging*, pp. 1-7.

**Ferrara, P.; Bianchi, T.; De Rosa, A.; Piva, A.** (2012): Image forgery localization via fine-grained analysis of CFA artifacts. *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 5, pp. 1566-1577.

**Jia, Y.; Shelhamer, E.; Donahue, J.; Karayev, S.; Long, J. et al.** (2014): Caffe: convolutional architecture for fast feature embedding. *Proceedings of the 22nd ACM International Conference on Multimedia*, pp. 675-678.

**Li, W.; Yuan, Y.; Yu, N.** (2009): Passive detection of doctored JPEG image via block artifact grid extraction. *Signal Processing*, vol. 89, no. 9, pp. 1821-1829.

**Long, J.; Shelhamer, E.; Darrell, T.** (2015): Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431-3440.

**Liu, Y.; Guan, Q.; Zhao, X.; Cao, Y.** (2017): Image forgery localization based on multi-Scale convolutional neural networks. *Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Securit*, pp. 85-90.

**Mahdian, B.; Saic, S.** (2009): Using noise inconsistencies for blind image forensics. *Image and Vision Computing*, vol. 27, no. 10, pp. 1497-1503.

**Pomari, T.; Ruppert, G.; Rezende, E.; Rocha, A.; Carvalho, T.** (2018): Image splicing detection through illumination inconsistencies and deep learning. *25th IEEE International Conference on Image Processing*, pp. 3788-3792.

**Pun, C. M.; Liu, B.; Yuan, X. C.** (2016): Multi-scale noise estimation for image splicing forgery detection. *Journal of Visual Communication and Image Representation*, vol. 38, pp. 195-206.

**Rao, Y.; Ni, J.** (2016): A deep learning approach to detection of splicing and copy-move forgeries in images. *IEEE International Workshop on Information Forensics and Security*, pp. 1-6.

**Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S. et al.** (2015): Imagenet large scale visual recognition challenge. *International Journal of Computer Vision,* vol. 115, no. 3, pp. 211-252.

**Siwei, L.; Xunyu, P.; Xing, Z.** (2014): Exposing region splicing forgeries with blind local noise estimation. *International Journal of Computer Vision*, vol. 110, no. 2, pp. 202-221.

**Sundaram, A. M.; Nandini, C.** (2017): ASRD: algorithm for spliced region detection in digital image forensics. *Computer Science On-Line Conference*, pp. 87-95.

**Salloum, R.; Ren, Y.; Kuo, C. C. J.** (2018): Image splicing localization using a multi-task fully convolutional network (MFCN). *Journal of Visual Communication and Image Representation*, vol. 51, pp. 201-209.

**Varlamova, A. A.; Kuznetsov, A. V.** (2017): Image splicing localization based on CFA-artifacts analysis. *Computer Optics*, vol. 41, no. 6, pp. 920-930.

**Wang, W.; Dong, J.; Tan, T.** (2014): Exploring DCT coefficient quantization effects for local tampering detection. *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 10, pp. 1653-1666.

**Wei, J.; Wang, Y.; Ma, X.** (2017): Text image authenticating algorithm based on MD5-hash function and Henon map. *Ninth International Conference on Digital Image Processing*. https://doi.org/10.1117/12.2281572.

**Yao, H.; Wang, S.; Zhang, X.; Qin, C.; Wang, J.** (2017): Detecting image splicing based on noise level inconsistency. *Multimedia Tools and Applications*, vol. 76, no. 10, pp. 12457-12479.

**Ye, S.; Sun, Q.; Chang, E. C.** (2007): Detecting digital image forgeries by measuring inconsistencies of blocking artifact. *IEEE International Conference on Multimedia and Expo*, pp. 12-15.

**Zhang, Y.; Goh, J.; Win, L. L.; Thing, V. L.** (2016): Image region forgery detection: a deep learning approach. *Proceedings of the Singapore Cyber-Security Conference*, pp. 1-11.

**Zampoglou, M.; Papadopoulos, S.; Kompatsiaris, Y.** (2017): Large-scale evaluation of splicing localization algorithms for web images. *Multimedia Tools and Applications*, vol. 76, no. 4, pp. 4801-4834.