Deep fusion network for splicing forgery localization

Bo Liu and Chi-Man Pun

University of Macau, Taipa, Macao, China.
{yb57413,cmpun}@umac.mo

Abstract. Digital splicing is a common type of image forgery: some regions of an image are replaced with contents from other images. To locate altered regions in a tampered picture is a challenging work because the difference is unknown between the altered regions and the original regions and it is thus necessary to search the large hypothesis space for a convincing result. In this paper, we proposed a novel deep fusion network to locate tampered area by tracing its border. A group of deep convolutional neural networks called Base-Net were firstly trained to response the certain type of splicing forgery respectively. Then, some layers of the Base-Net are selected and combined as a deep fusion neural network (Fusion-Net). After fine-tuning by a very small number of pictures, Fusion-Net is able to discern whether an image block is synthesized from different origins. Experiments on the benchmark datasets show that our method is effective in various situations and outperform state-of-the-art methods.

Keywords: Image forensics · Splicing forgery detection · Forgery localization · Deep convolutional network · Fusion network.

1 Introduction

Digital splicing refers to replacing some regions of a digital image with contents from other pictures. It is a common form of image tampering and manipulation. Since the contents of the original image have been altered, the meaning of the image conveyed is changed and sometimes even changed completely. The great development of photo editing software makes high-quality image tampering easily even for non-professional and untrained people. And these intentionally manipulated photos spread rapidly and widely through the Internet, turning to the misleading or fake news. Therefore, there is a strong need for image forensic method which is able to judge whether the contents of an image has been altered, and more specifically, which part of the image has been altered. The latter is indeed an important problem in image forensics: splicing forgery localization.

To discriminate whether a picture has undergone digital splicing, the technique of watermarking can be used [33]. If the watermarking of the picture changes, the picture is regarded as being altered in some ways such as copy-move forgery or splicing forgery. But those methods require the original pictures

to produce the watermark, thus given a test picture without any watermark they cannot make a judgment. Moreover, producing watermarks for every picture taken by cameras is impossible. Therefore, there is a strong need for methods which can detect forgery without a prior. The pictures undergone splicing forgery contains two areas: original background region from the host image and the spliced region from other pictures, therefore there must be some difference between these two areas. Many assumptions on such difference were made to design image forensic algorithms. Noise discrepancies between the spliced area and the host picture can be a cue to locate forgery because an image unavoidably bears a certain type and a certain amount of noise and pictures from different origins may carry different patterns and levels of the noise [32,30]. Another commonly seen assumption is based on the traces left by JEPG compression algorithm [38,26]. In the compression pipeline, an image is divided into fix-sized blocks and quantized by a pre-set table called quantization table. By estimating quantization table used in a test picture, if different tables are found, the picture may be fake [12]. To create a splicing forgery, the manipulation often involves double or more times of compressions. In double compression, the grid of the first compression and the second compression may be not aligned in a spliced area, analysis of such traces can also locate splicing forgery [27,5,2,6]. Under the circumstance multiple compressions, the different number of times of compressions of the different areas is also an indicator of splicing[22]. It is another assumption that the spliced object may be geometrically adjusted, such as rotation and affine transformation. These adjustments produce interpolations, which present periodical patterns in frequency domain thus can be used to expose splicing forgery [31,34]. Apart from these latent discrepancies, perceivable visual patterns can also be utilized. These methods estimate the direction of environmental light beams [19], or model the distortions caused by lenses [16], or detect inconsistent shadows in the image [29].

In recent years, the deep learning shows great power in many research fields and methods based on the deep convolutional neural networks(CNN) outperform traditional methods and achieve huge success in solving problems of computer vision, such as saliency detection [24], semantic segmentation [17] and depth estimation [28]. Some methods utilizing CNN have been proposed to expose image forgeries. One successful attempt is made in [35] that a 10-layer CNN as a classifier to decide a picture is authentic or manipulated by copy-move or splicing operation. A similar method producing a yes-or-no result by CNN is also seen in [20]. Bayar [4] designed a new convolutional network to learn manipulation features, rather than features of image contents as traditional convolutional layers did. To locate splicing forgery, the deep neural network is trained to learn manipulation features proposed in traditional methods, such as JPEG double compression features [37,23,3], resampling features [8,14] and camera-based features [7].

In this paper, we proposed a CNN based framework to deal with the problem of splicing forgery localization. Firstly, several deep convolutional neural networks called base-net are created. Each base-net is trained to be sensitive

to a specific type of discrepancy which exists in synthesized images. Secondly, some layers of each base-net are selected and then combined with selected layers from other base-nets to construct a new network called fusion-net. After the fine-tuning with a few numbers of training images, the fusion-net will be able to detect splicing forgery in digital images. The main contribution of our work is to propose a fusion framework which combines common hypotheses of digital splicing, and obtain the manipulation features by deep learning, which is proven to outperform the hand-crafted features. Our effort in the paper as follows. Firstly, instead of processing a test image as a whole, we process small image blocks from the test image. This is to avoid learning high-level visual features, which are not relevant to the image forensics, by deep neural networks. And in this fashion, our method can handle images in very large size with high resolution. Secondly, we carefully chose two manipulation features and utilized CNN to extract these features in the sufficiently large databases we created. Thirdly, the fusion of different features was via fine-tuning by a small number of training images. Experiments show that our proposed deep fusion network outperformed the state-of-the-art.

2 Related Work

As stated in the former section, in order to expose splicing forgery, a proper assumption will be firstly proposed and then an algorithm is then designed under such assumption. Therefore, a specific method is effective to a certain type of splicing forgery only. In order to detect more types of digital splicing, the fusion method is needed. Before a review of fusion frameworks, we will first describe some algorithms based on a single assumption.

JPEG Compressions. JPEG compression standard has been widely adopted. As a lossy compression scheme, the compression pipeline will unavoidably leave the images some traces which can be used to expose digital splicing. Double compression is a common hypothesis which assumes the forged area have been double compressed while the pristine region has been compressed one-time [26,27,5,6]. Double compression can be detected by finding a derivation of modelled DCT coefficients and generating a likelihood map which presents the probability of each 8×8 image block of being doubly compression [6]. The analysis and model of the work are based on double compression. But in the real situation, forged images are often compressed more than just twice. And the algorithm is not robust to a certain situation when second compression quality is better than the first one. Wang's method [37] based on a seven-layer CNN successfully solves the problem and can deal with such situation. Amerini's work [3] improves detection accuracy by integrating information from spatial domain and frequency domain of pictures into the CNN framework. Their proposed multi-domain neural network includes a seven-layer CNN to extract features from frequency domain and an eight-layer CNN to extract features from spatial domain, followed by two fully connected layers.

Image Noise. Methods based on the hypothesis that the spliced area bears different amount of noise include two steps: local noise estimation and forgery localization. Forgery localization requires accurate local noise estimation in images. Mahdian's work [32] estimate local noise by tilling sub-band HH_1 of the wavelet transformed non-overlapping image blocks. Lyu [30] describes a method based on the phenomenon of kurtosis concentration in natural images. The test image is firstly decomposed into several band-pass filtered channels using AC filters from the DCT decomposition. Then in each band-pass filtered channels, raw moments from the first to the fourth order will be calculated, followed by computing kurtosis and variance for each local window in each band-pass filtered channel. Lastly, noise variance is estimated by the projection from a local window across all band-pass filtered channels. Aforementioned methods firstly evaluate the noise variance, then finding the regions with different noise level from the rest. Therefore, the performance of noise estimation in their methods are crucial. However, blind noise estimation is a difficult task especially when the local window is small. Actually, noise variance estimation is not a must in exposing splicing forgery. Our target is to find the discrepancy of noise, rather than the noise variance.

If we want a single method that can cover more hypotheses so as to deal with more splicing instances, an effective fusion of results from different methods is necessary. Some fusion methods have been proposed so far. Different indicators of splicing forgery can be incorporated by discriminative random field and formulated as a labelling problem [18]. Another fusion method in [15] uses Dempster-Shafer theory of evidence which is regarded as an extension of the Bayesian theory to fuse existing forensic methods. Apart from utilizing the classic probability theories, the fusion can be implemented by pre-defined rules [9,21]. Li's fusion framework [25] firstly uses two existing forensic methods, i.e., statistical feature-based detector and copy-move forgery detector, to produce tampering possibility maps and then project these two scores of each pixel of training images into a two-dimensional plane. A decision curve is then manually determined to distinguish pristine and fake pixels. Although the fusion methods can generate reasonable results, the extension ability of these methods is limited: the fusion scheme must be altered or the computational complexity will increase prominently. Therefore, ideal fusion method should be more flexible and extendible to incorporate new forensic methods.

3 Deep Fusion Network

In this section, we first present the framework our method and then discuss and analyse the network including the Base-net which is used to extract forensic features and the Fusion-Net which fuses forensic features to give predictions.

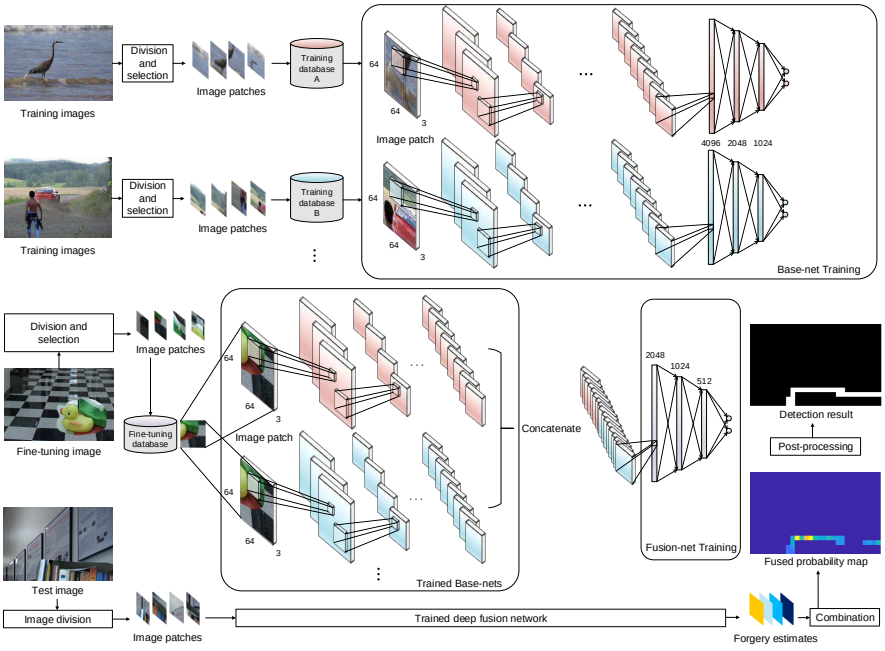


Fig. 1. The framework of proposed method.

3.1 The framework of proposed method

Our proposed digital splicing detection framework is illustrated in Fig.1. The training process for the network has two stages: base-net training and fusion-net fine tuning.

Base-net training. Each base-net is designed to make binary forgery prediction under a certain forgery hypothesis. In our implementation, two forgery hypotheses are used: a fake image patch contains contents from two sources whose noise levels are different, or a fake image patch contains two origins which undergo different JPEG compressions. For each base-net, a particular training database is constructed which consists of image patches of a fixed and same size. These image patches are taken from splicing forged pictures and if an image patch contains both spliced objects and the background image, it will be labelled as forged, if not, it will be labelled as genuine. In order to balance the training images, we selected the image patches to equal the numbers of the forged and the genuine image patches. The structure of VGG-16 [36] is used in our work but other deep convolutional neural networks can also be adopted. When the training for each base-net completes, convolutional kernels except those form fully connected layers will be retained for the next step fine-tuning.

Fusion-net fine-tuning. The construction of fine-tuning database is similar to the base-net training database as the forged images are divided into patches

and proportion of the forged and the genuine controlled to 1 : 1. The trained convolutional kernels from the base-nets are used to construct the fusion-net: the parameters of convolutional kernels from the trained base-nets are fixed and remain unchanged during the fusion-net training while the parameters in the fully connected layers are trained only during fine-tuning. So, an image patch will be sent to these trained base-nets to extract forensic features perspectively and then the features will be concatenated and then goes to fully connected layers. Since the number of parameters in the fully connected layers is smaller than that in convolutional layers, the fine-tuning process will take less time than base-net training and will converge quickly.

Image forensics. The trained deep fusion network now can be used to discern the fake image patches. The test image is firstly divided into non-overlapped image patches and then use deep fusion network to give predictions. The prediction of each image patch will be a probability of being forged. Combining all the predictions of each image patch we get a heat map called fused probability map. From this map, the borders of splicing area show the higher probability of being forged because the border image patches contained contents from different origins and it is perceived by the fusion-net. After simple post-processing such as thresholding, the detection result will be given as a binary map which is a tracing of borders of the spliced area.

3.2 Extractions of forensic features

In the first stage, the revised VGG16 convolutional neural network [36] produces forgery estimates for image patches. In our implementation, the network takes the image patch of fixed size at $64 \times 64 \times RGB$ as input. These input images are non-overlapping patches taken from original sized images in databases. Alternatively, to segment a test image into overlapping patches and then using the network is feasible as well, but we did not see significant improvement of the performance, and it prolongs authentication time. We therefore use non-overlapped image patches for experiments. The network is composed of the following layers in Table 1. *Conv.* and *F.C.* are short for convolutional layers and fully connected layers respectively.

The successive convolutional and max-pooling layers numbered from sequence 2 to sequence 10 are used to extract forensic features, while the function of fully connected layers numbered from sequence 11 to 14 is to classify. Because of this characteristic of deep convolutional networks, we can use several networks to extract different forensic features which relate to different assumptions respectively. One assumption we used in this work is noise discrepancy: comparing to its background host image, the spliced region has a different amount of noise. Therefore, a network is trained to discern this kind of discrepancy caused by the image noise. In order to train this convolutional network and make it noise sensitive, we created a special training dataset in which spliced objects were corrupted by additive noise and the noise variance was adjusted to mimic different situations. Because the noise variance can be controlled, a large dataset can be generated automatically and it covers most of the splicing scenarios where

Table 1. The sequence of layers in base-net.

Sequence	Layer type	Filter size	Output size
1	*Input*	-	[64 64 3]
2	*Conv.+ReLU*	[3 3 3 64]	[64 64 64]
3	*Max-Pooling*	-	[32 32 64]
4	*Conv.+ReLU*	[3 3 64 128]	[32 32 128]
5	*Max-Pooling*	-	[16 16 128]
6	*Conv.+ReLU*	[3 3 128 256]	[16 16 256]
7	*Max-Pooling*	-	[8 8 256]
8	*Conv.+ReLU*	[3 3 256 512]	[8 8 512]
9	*Conv.+ReLU*	[3 3 512 512]	[8 8 512]
10	*Max-Pooling*	-	[4 4 512]
11	*F.C.+ReLU*	[4 4 512 4096]	[1 1 4096]
12	*F.C.+ReLU*	[1 1 4096 2048]	[1 1 2048]
13	*F.C.+ReLU*	[1 1 2048 1024]	[1 1 1024]
14	*F.C.+ReLU*	[1 1 1024 2]	[1 1 2]
15	*Loss*	-	-

noise discrepancy exists. Similarly, a base-net which is sensitive to discrepancies of JPEG compressions is constructed and trained. The dataset consists of untouched image patches and patches undergone splicing forgery. These forged image patches are generated by combining two images with different JPEG compression quality. Apart from the noise and JEPG quality, the visual information such as color, texture and shape gives clues for image forensics. Accordingly, the third base-net is used to extract those forensic features. The only difference is the spliced regions of images in the dataset are not intentionally added with noise or altered JEPG compression quality, and the spliced objects are directly inserted into host images without any processing to form the forged pictures. The details of our three datasets will be introduced in section 4.

Since many deep convolutional networks have a structure of convolutional layers plus fully connected layers like VGG-16, networks in our framework can be replaced with any other deep convolutional networks. This is because the alternating convolutional layers and pooling layers generate features, while fully connected layers classify these features. Fig. 2 visually shows the ability of the network to extract noise features. Gaussian noise was added to the upper half of the image to mimic noise discrepancy in real splicing forgery. And comparing to features maps of untouched version (c)∼(e), there are visible activations in the upper half of feature maps in (f)∼(h). Note that the shown feature maps are from first three convolutional layers of the trained noise base-net and only first 64 feature maps are shown in each layer.

3.3 Features merging by fusion-net

A trained base-net is able to detect splicing forgery especially those are under its assumption, i.e., the noise sensitive base-net is good at discriminating the splicing

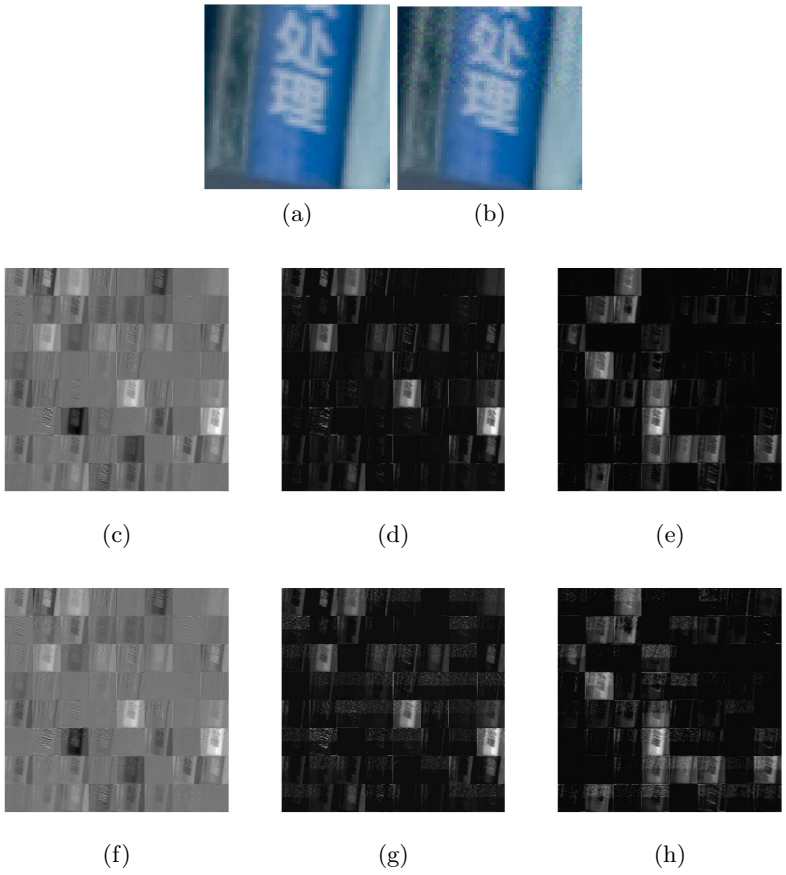


Fig. 2. Comparison between untouched image patch and its Gaussian noise corrupted version: feature maps of first three convolutional layers. (a) untouched image; (b) upper half of image is corrupted by Gaussian noise with variance $\sigma = 0.001$; (c)$\sim$(e) feature maps of untouched image in first three convolutional layers; (f)$\sim$(h) features maps of noise corrupted image in these three layers.

where the noise discrepancy exists, and the JEPG quality sensitive base-net discerns those image patches where JEPG compression quality is inconsistent. But the real situation may contain both of the noise discrepancy and compression inconsistency or either of them or none of them. Therefore, a fusion framework is needed to fuse independent forensic features. As stated in the former section, the alternating convolutional and pooling layers in a base-net extract a certain forensic feature, and when the base-nets have been well trained the parameters of the convolutional layers will be saved. In our work, the structure and parameters of sequence 1$\sim$10 of each base-net will be retained.

To construct the fusion-net, the outputs of sequence 10 from each base-net are concatenated and followed by four fully connected layers. The sequence of layers in fusion-net is shown in Table 2. We created another small database to train the fusion-net in order to make it capable to detect splicing forgery in the real scene. Note that the trainable parameters are from those four fully connected layers while the filters from the base-nets will not be trained because they have been trained already to extract the certain forensic features. The filter size of layer sequenced 2 in the fusion-net should be adjusted according to the number of used base-net. In our work, we trained three base-nets and the output of layer sequenced 10 in the base-net is $8 \times 8 \times 512$, accordingly the third dimension of layers sequenced 2 in the fusion-net will be $512 \times 3 = 1536$.

Table 2. The sequence of layers in fusion-net.

Sequence	Layer type	Filter size	Output size
1	*Concatenate*	-	[4 4 1536]
2	*F.C.+ReLU*	[4 4 1536 2048]	[1 1 2048]
3	*F.C.+ReLU*	[1 1 2048 1024]	[1 1024]
4	*F.C.+ReLU*	[1 1 1024 512]	[1 1 512]
5	*F.C.+ReLU*	[1 1 512 2]	[1 1 2]
6	*Loss*	-	-

The trained deep fusion network produces scores of being pristine and being tampered of an input image patch. This is because we used log-softmax loss function. The test image I will be divided into non-overlapped or overlapped image patches $I(i, j)$. Suppose the *pristine score* of an input image patch $I(i, j)$ is $s_p(i, j)$ and the *tampered score* is $s_t(i, j)$, then we normalized all the pristine scores of the image patches from the picture to be authenticated and obtained normalized pristine score $\hat{s}_p(i, j)$. The normalized tampered score $\hat{s}_t(i, j)$ is calculated in a similar way. Then the fused probability $f(i, j)$ is obtained by

$$f(i, j) = \frac{\hat{s}_t(i, j)}{\hat{s}_p(i, j)}. \tag{1}$$

Combining all fused probability scores $f(i, j)$ produces a fused probability map of the whole image, and giving a threshold τ can easily get the detection result when f is normalized. Morphological opening and closing operation yield better result. The marked area will be the border of the spliced region.

4 Experimental Results and Discussions

4.1 Databases and network training

Three different datasets are constructed to train three base-nets perspectively. The source pictures are from VOC2012 dataset [11] which is used to test algorithms of image segmentation and object detection. Since it provides masks of

objects in the image, we can utilize it to clip objects along with their borders to create fake pictures. To build the dataset to train noise-sensitive base-net, an object in a randomly selected picture in VOC2012 dataset was clipped and then added with the Gaussian noise, then another image was randomly selected and added with Gaussian noise as well, and finally the object was inserted into the later image to create a forged picture. The noise variance of added Gaussian noise was randomly decided from 0 (no noise added) to 0.005 with an interval of 0.001. Then, the fake images were divided into small image blocks sized 64×64. There are 186K images in each set and the ratio of splicing images to pristine images is 1:1 and the number of training images to the number of validation images is 9:1. Creating the dataset to train JPEG-compression-sensitive base-net is in a similar procedure, and the only difference is the compression quality was altered instead of adding noise. The compression quality factor was randomly decided from 10 to 100 with an interval of 5. To construct the third dataset, no additional operation was made and the clipped objects were directly spliced into the host image to create a forgery.

4.2 Detection results and comparisons

We first present the experimental results on a small dataset - Columbia Image Splicing Detection Evaluation Dataset [1]. The pictures in the dataset are uncompressed and before evaluating the fusion net, we converted the files to the JEPG compression format. There are 200 pictures in the dataset and we used 120 pictures to fine-tune the fusion net and left 80 pictures to evaluate the performance.

We presented some pictures and their probability heat maps (m)∼(p) in Columbia dataset in fig. 3. The yellow color indicates a higher probability that an image block is tampered while the blue color indicates a higher probability of being pristine. The borders of the splicing area were successfully detected although there are a few false negative blocks in (m) (o) and (p) and false positive blocks in (n). These unwanted results can be easily corrected by morphological operations in the post-processing, and the final detection results are shown in (q)∼(t). We also presented the pristine probability map (e)∼(h) and fake probability map (i)∼(l) as well. Although the number of images used for training the fusion net is very small. Comparing to traditional deep learning methods which require millions of training images, our deep fusion net only takes a very limited number of pictures to train. This is because most of the convolutional layers are from the trained base-nets and only fully connected layers of the fusion-net are trained in the fine-tuning.

We compared our method with the state-of-the-arts on Columbia dataset. The methods are based on the hand-crafted JPEG and noise forensic features. The ROC curves are presented in fig. 4 and the left curves are from these JEPG based methods while the right curves are noise based methods. Our method outperforms all JEPG based methods and performs well in most cases comparing to noise based methods. A merit of our proposed method is that the fusion-net achieves high true positive ratio while keeping a high value of the true negative.

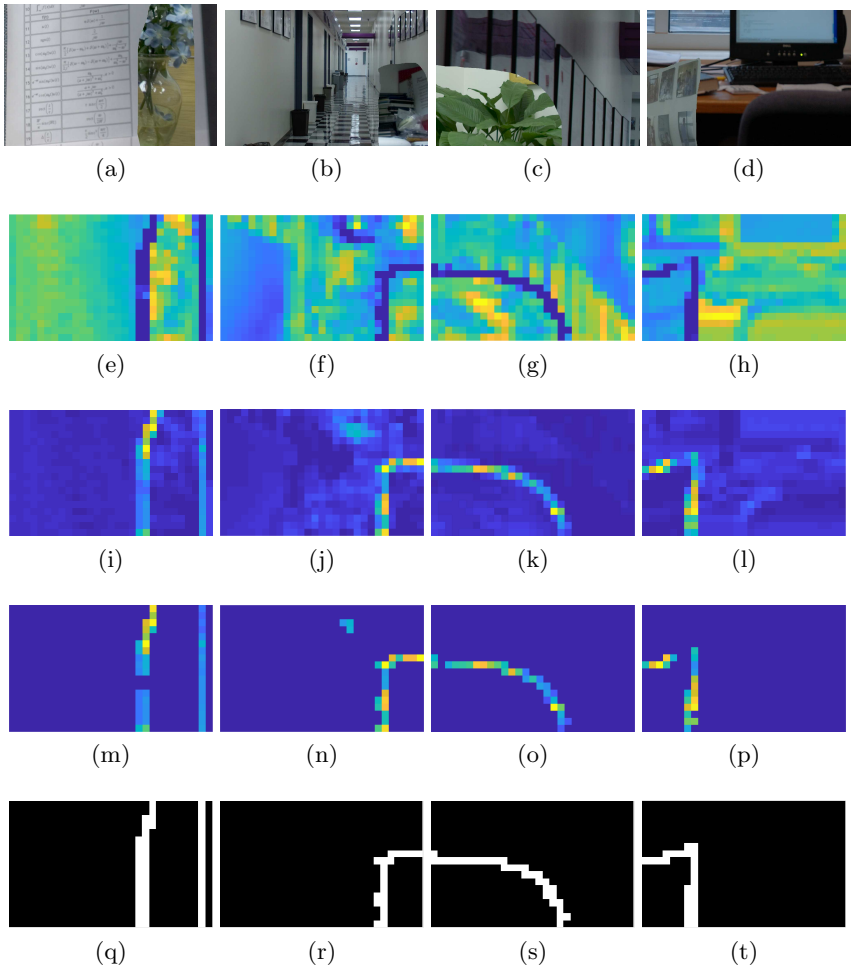
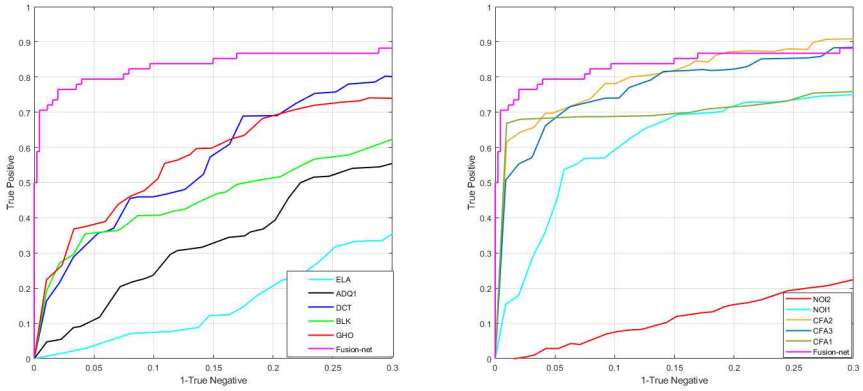


Fig. 3. Some pictures and their probability heat maps f and the final results in the Columbia splicing dataset: (a)∼(d) pictures of splicing forgery in the dataset; (e)∼(h) heat maps of being pristine; (i)∼(l) heat maps of being tampered; (m)∼(p) probability heat maps f produced by the deep fusion network. The yellow color in f indicates a higher probability of being tampered and the deep blue color indicates a higher probability of being pristine. (q)∼(t) the final detection results after proper morphological operations.

This is very important in splicing detection because false alert will significantly affect the observer's judgement. And in this standard, we outperform the other methods.



(a) (b)

Fig. 4. Comparisons of ROC curves with the state-of-the-arts in Columbia dataset: (a) Our fusion-net method and the other JPEG based methods. (b) our method and the other noise based methods. The acronym of algorithms and related works: ADQ1[27], BLK[26],CFA1[13], CFA2 and CFA3[10], DCT[38], ELA[22], GHO[12], NOI1[32], NOI2[30]

5 Conclusions and future works

A new deep fusion network for splicing localization has been presented. The fusion-net consists of convolutional layers from the base-nets and trainable fully connected layers. Since the base-nets have been trained and their convolutional layers have the ability to extract the certain forensic features, only fully connected layers of the fusion-net require training and this dramatically reduces the number of training pictures. Besides, the proposed framework is flexible and can be extended easily. Advanced deep convolutional networks or new forensic assumptions can replace the network used in this work to achieve better performance in the future.

Acknowledgement

This work was supported in part by the Research Committee of the University of Macau under Grant MYRG2018-00035-FST, and the Science and Technology Development Fund of Macau SAR under Grant 041/2017/A1.

References

1. Columbia image splicing detection evaluation dataset. `http://www.ee.columbia.edu/ln/dvmm/downloads/AuthSplicedDataSet/AuthSplicedDataSet.htm`, accessed: 2018-06-28

2. Amerini, I., Becarelli, R., Caldelli, R., Del Mastio, A.: Splicing forgeries localization through the use of first digit features. In: Information Forensics and Security (WIFS), 2014 IEEE International Workshop on. pp. 143–148. IEEE (2014)
3. Amerini, I., Uricchio, T., Ballan, L., Caldelli, R.: Localization of jpeg double compression through multi-domain convolutional neural networks. In: Proc. of IEEE CVPR Workshop on Media Forensics (2017)
4. Bayar, B., Stamm, M.C.: A deep learning approach to universal image manipulation detection using a new convolutional layer. In: Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security. pp. 5–10. ACM (2016)
5. Bianchi, T., De Rosa, A., Piva, A.: Improved dct coefficient analysis for forgery localization in jpeg images. In: Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on. pp. 2444–2447. IEEE (2011)
6. Bianchi, T., Piva, A.: Image forgery localization via block-grained analysis of jpeg artifacts. IEEE Transactions on Information Forensics and Security **7**(3), 1003–1017 (2012)
7. Bondi, L., Lameri, S., Güera, D., Bestagini, P., Delp, E.J., Tubaro, S.: Tampering detection and localization through clustering of camera-based cnn features. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 1855–1864 (2017)
8. Bunk, J., Bappy, J.H., Mohammed, T.M., Nataraj, L., Flenner, A., Manjunath, B., Chandrasekaran, S., Roy-Chowdhury, A.K., Peterson, L.: Detection and localization of image forgeries using resampling features and deep learning. In: Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on. pp. 1881–1889. IEEE (2017)
9. Cozzolino, D., Gragnaniello, D., Verdoliva, L.: Image forgery localization through the fusion of camera-based, feature-based and pixel-based techniques. In: 2014 IEEE International Conference on Image Processing (ICIP). pp. 5302–5306 (Oct 2014). https://doi.org/10.1109/ICIP.2014.7026073
10. Dirik, A.E., Memon, N.: Image tamper detection based on demosaicing artifacts. In: Image Processing (ICIP), 2009 16th IEEE International Conference on. pp. 1497–1500. IEEE (2009)
11. Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A.: The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results. http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html
12. Farid, H.: Exposing digital forgeries from jpeg ghosts. IEEE transactions on information forensics and security **4**(1), 154–160 (2009)
13. Ferrara, P., Bianchi, T., De Rosa, A., Piva, A.: Image forgery localization via fine-grained analysis of cfa artifacts. IEEE Transactions on Information Forensics and Security **7**(5), 1566–1577 (2012)
14. Flenner, A., Peterson, L., Bunk, J., Mohammed, T.M., Nataraj, L., Manjunath, B.: Resampling forgery detection using deep learning and a-contrario analysis. arXiv preprint arXiv:1803.01711 (2018)
15. Fontani, M., Bianchi, T., Rosa, A.D., Piva, A., Barni, M.: A framework for decision fusion in image forensics based on dempstershafer theory of evidence. IEEE Transactions on Information Forensics and Security **8**(4), 593–607 (April 2013). https://doi.org/10.1109/TIFS.2013.2248727
16. Fu, H., Cao, X.: Forgery authentication in extreme wide-angle lens using distortion cue and fake saliency map. IEEE Transactions on Information Forensics and Security **7**(4), 1301–1314 (2012). https://doi.org/10.1109/TIFS.2012.2195492

17. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Region-based convolutional networks for accurate object detection and segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence **38**(1), 142–158 (Jan 2016). https://doi.org/10.1109/TPAMI.2015.2437384
18. Hsu, Y.F., Chang, S.F.: Statistical fusion of multiple cues for image tampering detection. In: 2008 42nd Asilomar Conference on Signals, Systems and Computers. pp. 1386–1390 (Oct 2008). https://doi.org/10.1109/ACSSC.2008.5074646
19. Johnson, M.K., Farid, H.: Exposing digital forgeries in complex lighting environments. IEEE Transactions on Information Forensics and Security **2**(3), 450–461 (2007). https://doi.org/10.1109/TIFS.2007.903848
20. Kim, D.H., Lee, H.Y.: Image manipulation detection using convolutional neural network. International Journal of Applied Engineering Research **12**(21), 11640–11646 (2017)
21. Korus, P., Huang, J.: Multi-scale fusion for improved localization of malicious tampering in digital images. IEEE Transactions on Image Processing **25**(3), 1312–1326 (March 2016). https://doi.org/10.1109/TIP.2016.2518870
22. Krawetz, N., Solutions, H.F.: A pictures worth... Hacker Factor Solutions pp. 1–31 (2007)
23. Li, B., Luo, H., Zhang, H., Tan, S., Ji, Z.: A multi-branch convolutional neural network for detecting double jpeg compression. arXiv preprint arXiv:1710.05477 (2017)
24. Li, G., Yu, Y.: Visual saliency detection based on multiscale deep cnn features. IEEE Transactions on Image Processing **25**(11), 5012–5024 (Nov 2016). https://doi.org/10.1109/TIP.2016.2602079
25. Li, H., Luo, W., Qiu, X., Huang, J.: Image forgery localization via integrating tampering possibility maps. IEEE Transactions on Information Forensics and Security **12**(5), 1240–1252 (2017)
26. Li, W., Yuan, Y., Yu, N.: Passive detection of doctored jpeg image via block artifact grid extraction. Signal Processing **89**(9), 1821–1829 (2009)
27. Lin, Z., He, J., Tang, X., Tang, C.K.: Fast, automatic and fine-grained tampered jpeg image detection via dct coefficient analysis. Pattern Recognition **42**(11), 2492–2501 (2009)
28. Liu, F., Shen, C., Lin, G.: Deep convolutional neural fields for depth estimation from a single image. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 5162–5170 (2015)
29. Liu, Q., Cao, X., Deng, C., Guo, X.: Identifying image composites through shadow matte consistency. IEEE Transactions on Information Forensics and Security **6**(3), 1111–1122 (2011). https://doi.org/10.1109/TIFS.2011.2139209
30. Lyu, S., Pan, X., Zhang, X.: Exposing region splicing forgeries with blind local noise estimation. International Journal of Computer Vision **110**(2), 202–221 (2013). https://doi.org/10.1007/s11263-013-0688-y
31. Mahdian, B., Saic, S.: Blind authentication using periodic properties of interpolation. IEEE Transactions on Information Forensics and Security **3**(3), 529–538 (2008). https://doi.org/10.1109/TIFS.2004.924603
32. Mahdian, B., Saic, S.: Using noise inconsistencies for blind image forensics. Image and Vision Computing **27**(10), 1497–1503 (2009)
33. Podilchuk, C.I., Delp, E.J.: Digital watermarking: algorithms and applications. IEEE Signal Processing Magazine **18**(4), 33–46 (2001). https://doi.org/10.1109/79.939835

34. Popescu, A.C., Farid, H.: Exposing digital forgeries by detecting traces of resampling. IEEE Transactions on Signal Processing **53**(2), 758–767 (2005). https://doi.org/10.1109/TSP.2004.839932
35. Rao, Y., Ni, J.: A deep learning approach to detection of splicing and copy-move forgeries in images. In: Information Forensics and Security (WIFS), 2016 IEEE International Workshop on. pp. 1–6. IEEE (2016)
36. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
37. Wang, Q., Zhang, R.: Double jpeg compression forensics based on a convolutional neural network. EURASIP Journal on Information Security **2016**(1), 23 (2016)
38. Ye, S., Sun, Q., Chang, E.C.: Detecting digital image forgeries by measuring inconsistencies of blocking artifact. In: Multimedia and Expo, 2007 IEEE International Conference on. pp. 12–15. IEEE (2007)