

# Effective images splicing detection based on decision fusion

Anh Thu Phan Ho  
ICD - LM2S Laboratory  
University of Technology of Troyes  
France  
email: anh\_thu.phan\_ho@utt.fr

Florent Retraint  
ICD - LM2S Laboratory  
University of Technology of Troyes  
France  
email: florent.retraint@utt.fr

**Abstract**—Localizing the tampered regions in forgery images is an important and challenging problem in forensics applications. Though there have been an extensive studies on image forgery localization over past decade, each method still has its own limitations. Therefore, it is promising to fuse different forensic approaches in order to obtain better localization performance. In this paper, we propose a framework to aggregate the decision maps of two forensic approaches: Photo Response Non-Uniformity (PRNU) based approach and statistical features based approach using Dempster-Shafer Theory. PRNU noise can be considered as a camera fingerprint thereby being used effectively to localize tampering images. However, the most challenging limitation of this approach is its false identifications on textured, saturated and dark regions. By combining with the statistical feature based approach, we can decrease this false alarm rate on saturated and dark regions. The extensive experimental results demonstrate that the proposed method significantly outperforms the single PRNU based approach.

**Index Terms**—Forgery localization, Dempster-Shafer Theory, Photo Response Non-Uniformity, Decision fusion.

## I. INTRODUCTION

Over past decades, the wide variety of powerful digital image processing tools has made it simpler and easier to tamper with an image while leaving no obvious visual clue. Therefore, verifying the authenticity of an image has become increasingly important.

There has been accordingly an increasing research in forgery image localization [1], [2], [3], [4], [8], [9], [10], [16], [17]. The general idea in forgery image detection is to find the local inconsistencies which is caused by the presence of traces left during tampering processes in the investigated image. For instance, by exploiting the double quantization artifacts hidden among the DCT coefficients in forgery JPEG images, one can localize the tampered regions [5], [6], [7]. By analysing device characteristics such as sensor pattern noise, Photo Response Non-Uniformity noise (PRNU) [8], [9], [10], color filter array (CFA) [11], [12], one is able to detect regions whose camera fingerprint is inconsistent with that of the whole image to accomplish forgery detection. Others detect the forgeries relying on traces such as resampling artifacts [13], lighting condition [14], and so on. However, each approach has its own advantages and disadvantages. For example, approaches based JPEG compression artifacts

require the first compression format is JPEG which may not always be true in practice. Approaches based PRNU can work for forgery RAW images and resaved in JPEG images. However, the PRNU noise can be easily disturbed by the noise in textured, saturated or dark regions leading to false alarm. It is then necessary to fuse several approaches to obtain a better forensics performance.

There have been many strategies of fusion as long as we can incorporate the advantages of each single approach. The authors in [3], [15] fuse multiple candidate tampering maps resulting from the analysis with different sizes of sliding windows to obtain a more reliable tampering map with better localization resolution. They proposed an energy-minimization approach using Markov Random Field to model the prior knowledge about the tampering maps. In [16], [17] the authors proposed a decision fusion framework for the image forensics scenario based on Dempster-Shafer Theory (DST) avoiding the necessity of assigning prior probabilities which would be difficult to estimate. However these methods only considered JPEG compression artifacts. In [2], the authors proposed a more general algorithm which obtained the best score in the first phase of the First IFS-TC Image Forensics Challenge<sup>1</sup>. They used a simple fusion decision OR to combine the detection results of a statistical feature based detector and a copy-move detector to obtain a better performance. Also fusing the tampering possibility maps of statistical feature based detector and copy-move detector, the authors in [1] however designed a more effective fusion strategy. We realize that the statistical feature based detector extracting features such as spatial rich model (SRM) and spatial color rich model (SCRM) adopted in [1], [2] can be considered as a universal approach. It can detect many types of forgery images and stable with saturated and dark regions. We come up with the idea improving the forgery detection performance by incorporating this statistical feature based detector with the PRNU based detector. The PRNU based approach is often stable with the most common forms of image post-processing, such as JPEG compression, filtering, or gamma correction thereby being effective to image forgery detection. Combining with the statistical feature based detector would help to alleviate the problem of false alarm

<sup>1</sup><http://ifc.recod.ic.unicamp.br/>

in PRNU based method. In this paper, we fuse the decision maps of the statistical feature based detector and PRNU based approach using Dempster-Shafer Theory framework. Different from the fusion strategy in [1], fusion based DST does not require a costly training classifier to look for decision curves.

The remainder of this paper is organized as follows. Section II gives a brief overview of the PRNU based approach and the statistical feature based approach. In Section III, we first review the Dempster-Shafer Theory and then describe the fusion method based DST. The experimental results are shown in Section IV. Finally the conclusion and future work are drawn in Section V.

## II. TWO IMAGE FORGERY DETECTION APPROACHES

In this section, we review two image forgery detectors, one based on PRNU and the other based on statistical features.

### A. PRNU-based forgery detection

The PRNU which is a camera imaging sensor imperfection can be considered as a unique sensor pattern of each individual camera thereby being used for forgery detection. The idea is that the tampered regions could destroy or change position of the PRNU in the image. Therefore, by testing on which part of the image the PRNU is changed, one is able to reveal the tampered regions.

We consider a simplified model of the image acquisition pipeline [18].

$$y = (1 + k)x + \eta \quad (1)$$

where  $y$  is a captured image,  $x$  is its idealized noise-free version,  $k$  is the camera PRNU,  $\eta$  is an additive noise term which accounts for all types of disturbances, and products between images, unless otherwise stated, are pixelwise. The PRNU  $k$  can be estimated from  $N$  images obtained by the camera  $y_1, \dots, y_N$  using the maximum likelihood principle.

$$\hat{k} = \frac{\sum_{n=1}^N r_n y_n}{\sum_{n=1}^N y_n^2} \quad (2)$$

where  $r_n = y_n - f(y_n)$  is the noise residual of the image  $y_n$ ,  $f$  is a denoising filter. In the following, for the sake of simplicity, we assume that the estimation of the camera PRNU has no error, i.e.,  $\hat{k} = k$ .

The PRNU of the image under test is compared with the reference PRNU in a sliding-window based manner. The forgery detection at each pixel  $y_{i,j}$  is formulated as a binary hypothesis testing problem applied to a block  $B$  centered around the pixel  $y_{i,j}$ .

$$\begin{cases} H_0 : r_B = \eta_B \\ H_1 : r_B = z_B + \eta_B \end{cases} \quad (3)$$

where  $r_B$ ,  $z_B$  and  $\eta_B$  are the restrictions of  $r$ ,  $z$ , and  $\eta$  respectively, to the block  $B$ ,  $z = yk$  is the signal of interest (also called the reference PRNU). If the PRNU is absent in the block  $B$  (hypothesis  $H_0$ ), its central pixel is labeled as being tampered. If the PRNU is present in  $B$  (hypothesis  $H_1$ ),

its central pixel is labeled as being genuine. The detection test is based on normalized correlation

$$\rho_{ij} = \text{corr}(r_B, z_B) \quad (4)$$

The probability density function (pdf) of  $\rho_{ij}$  under hypothesis  $H_0$  is estimated by correlating the camera PRNU and the noise residuals coming from other cameras. The pdf of  $\rho_{ij}$  under hypothesis  $H_1$  is heavily influenced by the block content. In deed, even in the genuine blocks, the correlation might be very low when these blocks are dark, saturated, or textured. The authors in [18] estimated a predictor based on local images features, such as texture, flatness and intensity, and then computed the expected value  $\hat{\rho}_{ij}$  of the correlation under hypothesis  $H_1$  hoping to reduce the false alarm in these cases.

The decision map  $M^{PRNU}$  is then defined as follows

$$M_{ij}^{PRNU} = \begin{cases} 0 & \rho_{ij} < \gamma_1 \text{ AND } \hat{\rho}_{ij} > \gamma_2 \\ 1 & \text{else} \end{cases} \quad (5)$$

where  $\gamma_1$  is the threshold selected with a Neyman-Pearson approach to obtain the desired false acceptance rate (FAR),  $\gamma_2$  is a threshold chosen heuristically to avoid labeling nontampered ( $M_{ij}^{PRNU} = 1$ ) pixels as tampered ( $M_{ij}^{PRNU} = 0$ ).

### B. Statistical feature-based forgery detection

The statistical feature based forgery detection is an approach in which we first extract some inherent features of image blocks that are likely to be modified when an image undergoes tampering and then use these features to proceed a two-class pristine/forged training procedure.

It can be said that this is a universal approach in which we can detect many types of forgeries though the accuracy is not high. Among various statistical feature sets proposed in steganalysis, in this paper we adopt the statistical features named Spatial Color Rich Model (SCRM) [19] which work quite effectively in forgery detection [1]. SCRM is an extension of SRM. The SRM features from the R, G, and B channel are first added together and then concatenated three dimensional co-occurrences of residuals computed from all three color channels. These features are then used for training procedure. Based on a sliding-window manner, the training samples are extracted from tampered and pristine blocks of size 64x64 pixel with a step of 16-pixel and then fed into an ensemble classifier [20] with linear discriminant analysis base learners for identifying whether an image block is genuine or fake. The image under test  $I$  is divided into 64x64 pixel sliding windows with a step of 16-pixel. For each sliding window, the pre-trained ensemble classifier outputs a vote score  $v \in \{-n_b, -n_b + 1, \dots, n_b - 1, n_b\}$  where  $n_b$  is the number of base learners in the ensemble classifier. The decision map  $M^{SF}$  is computed as follows.

$$M_{i,j}^{SF} = \frac{1}{2n_b} \left( \frac{1}{K} \sum_{k=1}^K v_k + n_b \right) \quad (6)$$

where  $K$  is the number of blocks containing pixel  $I_{i,j}$ , and  $v_k$  is the vote score for the  $k^{th}$  block that contains  $I_{i,j}$ .

### III. FUSION METHOD BASED DEMPSTER-SHAFFER THEORY

#### A. Elements of Dempster-Shafer Theory

DST which is an effective theoretical framework for fusing and reasoning with uncertain and/or imprecise information was introduced by Dempster and Shafer [21], [22]. In this subsection, we briefly review its two main components: the degrees of belief representation and the Dempster's rule for combining such degrees of belief when they are based on independent sources. Let  $X$  be a variable taking values in a finite domain  $\Omega = \{\omega_1, \dots, \omega_n\}$ , called the frame of discernment. Evidence about  $X$  may be represented by a mass function  $m : 2^\Omega \rightarrow [0, 1]$  such that

$$\sum_{A \subseteq \Omega} m(A) = 1 \quad (7)$$

Each number  $m(A)$  denotes a degree of belief attached to the hypothesis that  $X \in A$ .  $m$  is said to be normalized if  $m(\emptyset) = 0$ . This property will be assumed hereafter, unless otherwise specified.

Corresponding to a mass function  $m$ , we can associate belief functions  $Bel : 2^\Omega \rightarrow [0, 1]$  defined as follows

$$Bel(A) = \sum_{B \subseteq A} m(B) \quad (8)$$

Quantity  $Bel(A)$  can be interpreted as the degree to which the evidence supports  $A$ .

In order to combine the evidence coming from multiple independent sources of information, we can use Dempster's combination rule to merge them. Let  $m_1$  and  $m_2$  be two mass functions derived from independent items of evidence. They can be fused to induce a new mass function  $m_{12}$  defined as

$$m_{12}(A) = \frac{1}{1-K} \sum_{B, C \subseteq \Omega, B \cap C = A} m_1(B) m_2(C) \quad (9)$$

where  $K = \sum_{B, C \subseteq \Omega, B \cap C = \emptyset} m_1(B) m_2(C)$ ,  $K < 1$ , measures the degree of conflict between evidence  $m_1$  and  $m_2$ .

#### B. Proposed Method

The framework proposed in this subsection aims at fusing the evidence coming from the PRNU-based forgery detection and the Statistical feature-based forgery detection. We believe that aggregating the evidence from the Statistical feature-based approach will help to decrease the false alarm rate on the saturated and dark regions of images. The fusion procedure can be described as follows.

- Constructing mass functions  $m_1$  and  $m_2$  from the evidence of each approach: the PRNU-based forgery detection and the Statistical feature-based forgery detection.
- Using Dempster's combination rule to induce a fused mass function  $m_{12}$  from  $m_1$  and  $m_2$ .
- Computing the belief function corresponding to the mass function  $m_{12}$ .
- Making final decision bases on the belief function.

What we are interested in this paper is that whether the pixel  $I_{i,j}$  in test image  $I$  is tampered or not tampered. We

can model this scenario by defining a variable  $X$  with frame  $\Omega = \{it, nt\}$  where  $it$  is the proposition "the pixel  $I_{i,j}$  is tampered", and  $nt$  is the proposition "the pixel  $I_{i,j}$  is not tampered". We want to quantify how much we are confident in these propositions.

The mass function  $m_1$  and  $m_2$  are respectively constructed from the decision maps of the PRNU-based forgery detection and the Statistical feature-based forgery detection.

$$m_1(X) = \begin{cases} t_1 & \text{for } X = \{it\} \\ n_1 & \text{for } X = \{nt\} \end{cases} \quad (10)$$

$$m_2(X) = \begin{cases} t_2 & \text{for } X = \{it\} \\ n_2 & \text{for } X = \{nt\} \end{cases} \quad (11)$$

where

$$\begin{aligned} t_1 &= M_{i,j}^{PRNU}, & n_1 &= 1 - t_1 \\ t_2 &= M_{i,j}^{SF}, & n_2 &= 1 - t_2 \end{aligned} \quad (12)$$

The degree of conflict  $K$  and the fused mass function  $m_{12}$  is computed as follows

$$K = t_1 n_2 + t_2 n_1 \quad (13)$$

$$\begin{aligned} m_{12}(\{it\}) &= \frac{t_1 t_2}{1-K} = \frac{t_1 t_2}{1 - t_1 n_2 - t_2 n_1} \\ m_{12}(\{nt\}) &= \frac{n_1 n_2}{1-K} = \frac{n_1 n_2}{1 - t_1 n_2 - t_2 n_1} \end{aligned} \quad (14)$$

The belief function in this case is equal to the fused mass function:  $Bel(\{it\}) = m_{12}(\{it\})$  and  $Bel(\{nt\}) = m_{12}(\{nt\})$ . The quantity  $Bel(\{it\})$  is the degree to which the evidence supports that the pixel  $I_{i,j}$  is tampered. We then make a decision that a pixel is tampered if its degree of belief of tampering is greater than that of non tampering, i.e.  $Bel(\{it\}) > Bel(\{nt\}) + \lambda$ , where  $\lambda$  is a threshold chosen heuristically.

### IV. EXPERIMENTAL RESULTS

In this section, we will report some preliminary experiments meant to support the idea of fusion decision to improve the detection performance. Our experiments were carried out on three cameras, a Canon EOS-100D, a Nikon D5200 and a Panasonic DMC-GM1. We first have estimated the PRNU of each camera over 100 images and then have extracted 25000 correlation samples over 25 images coming from other cameras and 25000 samples coming from the same camera to train the correlation predictor as proposed in [18].

We present in this section results only for one of the cameras, a Canon EOS-100D. For our experiments we used 200 tampered images and 200 pristine ones. The forgeries have been created with a copy-and-paste process and are all rectangular with size of 128x128. We evaluated the percentage of correctly detected forged pixels in the tampered images ( $P_D$ ) and the percentage of falsely identified pixels in the pristine ones ( $P_{FA}$ ), varying the relevant parameters of the

algorithms that is, the  $\gamma_1, \gamma_2$  in the PRNU-based forgery detection algorithm, and the threshold  $\lambda$  in the fusion algorithm.

In Fig. 1, we show the evaluation on several realistic tampered images (second column). The original images (first column) are taken from Canon EOS-100D camera and then are forged by inserting objects using the popular photo editing software GIMP. The third and fourth columns show the output maps of the Statistical feature and the PRNU based approaches. As can be seen, each individual approach has its own limitation. The PRNU based approach correctly detects the tampered regions but the false alarm rate is hard to avoid due to the saturated and dark regions. In contrast, the Statistical feature based approach does not localize tampered regions with high accuracy but it does not have problem with saturated and dark regions. Therefore the integrated map fused from the PRNU and Statistical Feature approaches in the last column shows a significant improvement of the fusion method. It is worth mentioning that we did not apply any morphological operation in the fusion approach.

In Fig. 2 we show the ROCs (receiver operating characteristics) of the PRNU-based approach and of the fusion approach computed on the complete test set (200 forgery images and 200 genuine images). We can see in Fig. 2 that the fusion approach performs better than the single PRNU-based approach.

In Fig. 3 we show the ROC curves of the PRNU-based approach and of the fusion approach computed on the 10 forgery images and 10 genuine images whose saturated and dark regions are considerable. We can see in Fig. 3 that the fusion approach significantly outperforms the single PRNU-based approach.

## V. CONCLUSIONS AND FUTURE WORK

This paper has presented an effective images splicing detection by fusing the decision maps of two forgery detection approaches: PRNU based approach and statistical features based approach using Dempster-Shafer Theory. Preliminary experimental results have shown that this novel method works effectively and significantly improves the forgery detection performance comparing to each single approach. This improvement is mainly due to the fact that the proposed method has significantly decreased the false positive rate on the saturated and dark regions which is one of the most challenging limitation of the PRNU based approach.

Future work shall be devoted to deriving a more advanced fusion framework based on Decision-Making with Belief Functions Theory [23]. We also plan to fuse many forgery detection algorithms. This is an interesting and promising working direction because according to a recent study there are various forgery detection algorithms which have a set of advantages and disadvantages.

## ACKNOWLEDGMENT

This work was supported by ANR project DEFACTO ANR-16-DEFA-0002.

## REFERENCES

- [1] H. Li, W. Luo, X. Qiu, and J. Huang, "Image forgery localization via integrating tampering possibility maps," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 5, pp. 1240–1252, 2017.
- [2] D. Cozzolino, D. Gragnaniello, and L. Verdoliva, "Image forgery detection through residual-based local descriptors and block-matching," in *2014 IEEE international conference on image processing (ICIP)*, pp. 5297–5301, IEEE, 2014.
- [3] P. Korus and J. Huang, "Multi-scale fusion for improved localization of malicious tampering in digital images," *IEEE Transactions on Image Processing*, vol. 25, no. 3, pp. 1312–1326, 2016.
- [4] A. T. P. Ho, K. Wang, and F. Cayre, "An effective histogram-based approach to jpeg-100 forensics," in *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pp. 1–6, IEEE, 2016.
- [5] F. M.-N. Hoai Phuong Nguyen, Florent Retraint and A. Delahaies, "An image forgery detection solution based on dct coefficient analysis," in *Proceedings of Information Systems Security and Privacy Conference*, IEEE, 2019.
- [6] T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of jpeg artifacts," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 1003–1017, 2012.
- [7] J. He, Z. Lin, L. Wang, and X. Tang, "Detecting doctored jpeg images via dct coefficient analysis," in *European conference on computer vision*, pp. 423–435, Springer, 2006.
- [8] G. Chierchia, G. Poggi, C. Sansone, and L. Verdoliva, "A bayesian-mrf approach for prnu-based image forgery detection," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 4, pp. 554–567, 2014.
- [9] G. Chierchia, G. Poggi, C. Sansone, and L. Verdoliva, "Prnu-based forgery detection with regularity constraints and global optimization," in *2013 IEEE 15th International Workshop on Multimedia Signal Processing (MMSp)*, pp. 236–241, IEEE, 2013.
- [10] X. Lin and C.-T. Li, "Refining prnu-based detection of image forgeries," in *2016 Digital Media Industry & Academic Forum (DMIAF)*, pp. 222–226, IEEE, 2016.
- [11] C.-H. Choi, H.-Y. Lee, and H.-K. Lee, "Estimation of color modification in digital images by cfa pattern change," *Forensic science international*, vol. 226, no. 1–3, pp. 94–105, 2013.
- [12] A. Singh, G. Singh, and K. Singh, "A markov based image forgery detection approach by analyzing cfa artifacts," *Multimedia Tools and Applications*, vol. 77, no. 21, pp. 28949–28968, 2018.
- [13] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of resampling," *IEEE Transactions on signal processing*, vol. 53, no. 2, pp. 758–767, 2005.
- [14] M. K. Johnson and H. Farid, "Exposing digital forgeries in complex lighting environments," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3, pp. 450–461, 2007.
- [15] P. Korus and J. Huang, "Multi-scale analysis strategies in prnu-based tampering localization," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 4, pp. 809–824, 2017.
- [16] M. Fontani, T. Bianchi, A. De Rosa, A. Piva, and M. Barni, "A framework for decision fusion in image forensics based on dempster-shafer theory of evidence," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 4, pp. 593–607, 2013.
- [17] P. Ferrara, M. Fontani, T. Bianchi, A. De Rosa, A. Piva, and M. Barni, "Unsupervised fusion for forgery localization exploiting background information," in *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, pp. 1–6, IEEE, 2015.
- [18] M. Chen, J. Fridrich, M. Goljan, and J. Lukás, "Determining image origin and integrity using sensor noise," *IEEE Transactions on information forensics and security*, vol. 3, no. 1, pp. 74–90, 2008.
- [19] M. Goljan, J. Fridrich, and R. Cogranne, "Rich model for steganalysis of color images," in *2014 IEEE International Workshop on Information Forensics and Security (WIFS)*, pp. 185–190, IEEE, 2014.
- [20] J. Kodovsky, J. Fridrich, and V. Holub, "Ensemble classifiers for steganalysis of digital media," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 432–444, 2012.
- [21] G. Shafer, *A mathematical theory of evidence*, vol. 42. Princeton university press, 1976.
- [22] A. P. Dempster, "Upper and lower probability inferences based on a sample from a finite univariate population," *Biometrika*, vol. 54, no. 3–4, pp. 515–528, 1967.

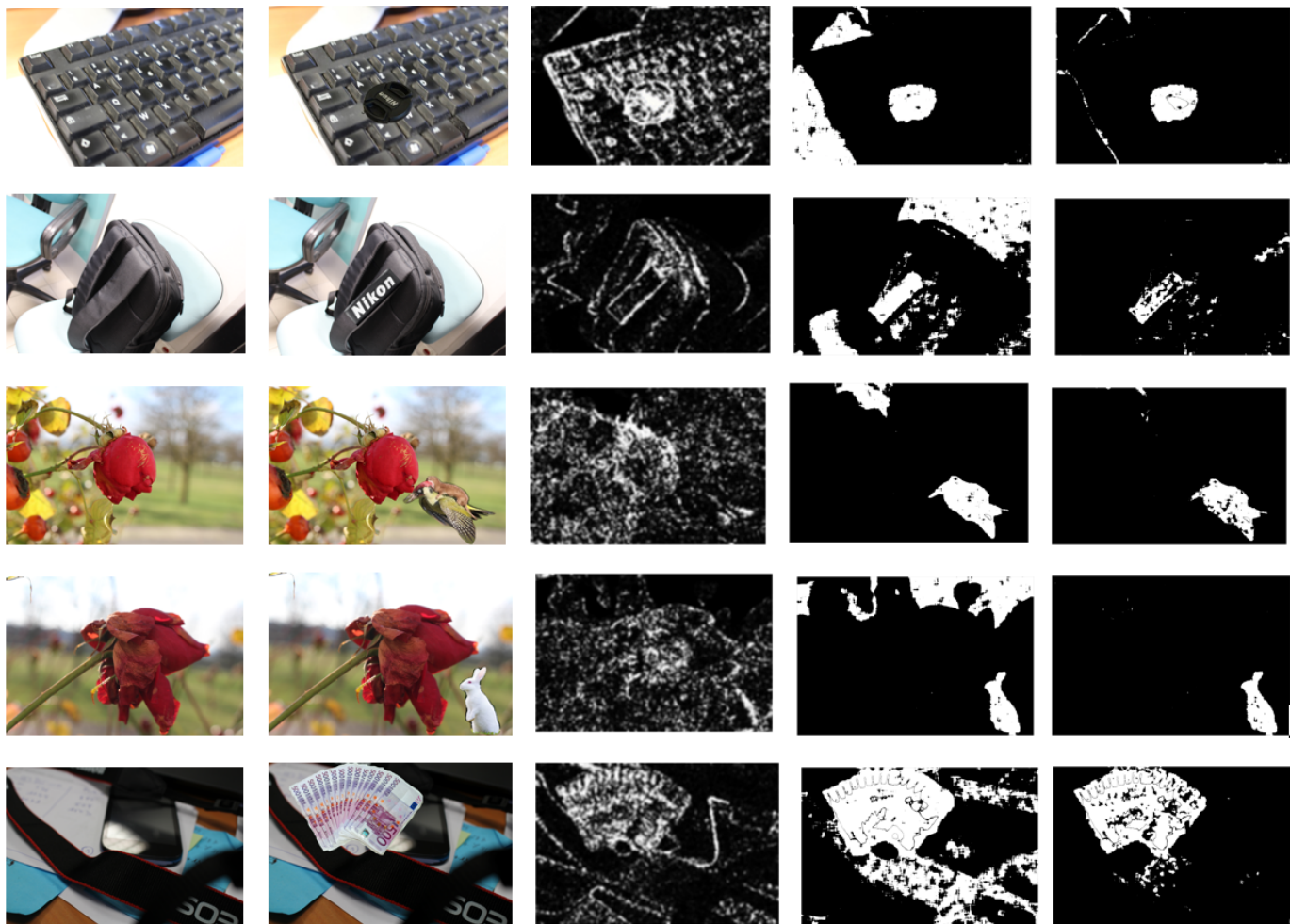


Fig. 1. Realistic examples of localizing tampering images. The first column are original images, the second column are tampered images, the third column are detection maps by Statistical Feature-based approach, the fourth column are maps detected by PRNU-based approach and the last column are maps created by the proposed algorithm.

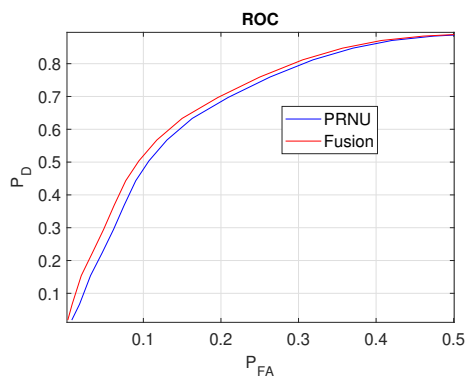


Fig. 2. ROC for PRNU-based forgery detection algorithm and the fusion one on 400 images.

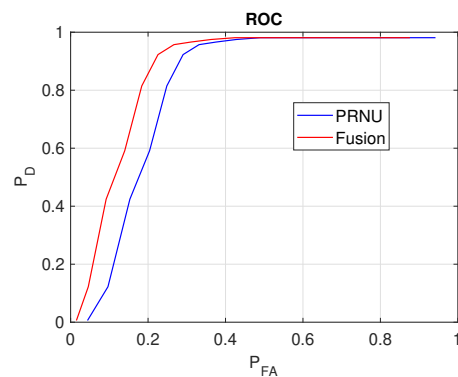


Fig. 3. ROC for PRNU-based forgery detection algorithm and the fusion one on 20 images.

[23] T. Dencux, "Decision-making with belief functions: A review," *International Journal of Approximate Reasoning*, vol. 109, pp. 87–110, 2019.