Knowledge-Based Fusion for Image Tampering Localization

Chryssanthi Iakovidou, Symeon Papadopoulos, Ioannis Kompatsiaris

Knowledge-based fusion for image tampering localization

Chryssanthi Iakovidou[1], Symeon Papadopoulos[1], and Yiannis Kompatsiaris[1]

[1] Information Technologies Institute, Centre for Research and Technology Hellas, 6th km Harilaou - Thermi, 57001, Thessaloniki, Greece
{c.iakovidou,papadop,ikom}@iti.gr, https://mklab.iti.gr/

Abstract. In this paper we introduce a fusion framework for image tampering localization, that moves towards overcoming the limitation of available tools by allowing a synergistic analysis and multiperspective refinement of the final forensic report. The framework is designed to combine multiple state-of-the-art techniques by exploiting their complementarities so as to produce a single refined tampering localization output map. Extensive evaluation experiments of state-of-the-art methods on diverse datasets have resulted in a modular framework design where candidate methods go through a multi-criterion selection process to become part of the framework. Currently, this includes a set of five passive tampering localization methods for splicing localization on JPEG images. Our experimental findings on two different benchmark datasets showcase that the fused output achieves high performance and advanced interpretability by managing to leverage the correctly localized outputs of individual methods, and even detecting cases that were missed by all individual methods.

Keywords: Image tampering localization · late-decision fusion · passive image forensics.

1 Introduction

Image forensics techniques have an important role in determining the authenticity of digital images. This is evident by the plethora of scientific approaches available in the literature that have carefully designed mechanisms to reveal different types of digital manipulations and traces that are expected to be generated during a given tampering process [10]. Producing robust tools for detecting and localizing a specific type of forgery has proven to be challenging, even when testing their effectiveness on benchmark datasets and controlled scenarios [19] and becomes even greater when dealing with real-world scenarios; images are being edited and manipulated in a variety of ways during a single forgery session in order to produce a convincing outcome, and are then forwarded and shared over the Internet, further undergoing transformations (e.g. cropping, resizing, re-compression). These uncontrolled factors inevitably force many of the standalone methods to suffer in terms of detection accuracy and localization robustness, presenting noisy outcomes and higher false positive rates when applied

to new datasets [9, 19, 3]. Thus, researchers have come to realise that there is a true benefit in acquiring different reports from independent tools and evaluating the multiple clues in conjunction in the context of blind/passive image forensics (i.e. forensic analysis where prior knowledge regarding the original capturing circumstances, the manipulations or other post processing transformations is unknown) [9, 12, 7].

Several fusion approaches have been proposed in the literature, aiming at a synergistic analysis that improves the overall robustness and reliability of the forensic report. The different strategies can be roughly categorized based on the level at which fusion is carried out and the traces that are considered. Frameworks proposed for *feature-level fusion* often suffer from drawbacks related with selecting and handling a large number of features and scalability when adding new tools [3, 4], while approaches based on *measurement level fusion* are best suited for the tampering detection problem as they provide a more high-level response in terms of confidence for particular traces being present or not [9, 7]. On the other hand, *pixel-level fusion* is more effective for tampering localization and techniques proposed in this direction usually involve utilizing probability output maps and a fusion model to refine the final output and improve the localization of the tampered region [12, 11]. However, several important issues have not been comprehensively studied, for example, how to select the appropriate "base" forensics approaches, how to fuse their detection results, and how to refine the fused localization map.

In this paper, we introduce an extensible fusion framework for tampering localization and output refinement. The design strategy focuses on analyzing tampering localization approaches from the literature that are selected and categorized based on a multi-criterion ranking process integrating also expert background knowledge regarding their domain of application (types of images and encoding, supported traces, known limitations, etc.). Next we employ a fusion mechanism based on local and cross-tool statistics to produce a single, refined fused heat-map output for tampering localization. We primarily focus on splicing localization which is a very common and effective type of tampering that occurs when parts of the original image are replaced by alien content, while also prioritizing including methods that base their detection mechanisms on JPEG-related traces as it remains the dominant image codec for digital images in devices and on the Internet.

The main objectives of the delivered framework are: i) to exploit tools that are complementary to each other, such that the robustness and reliability of the overall localization system can be improved, and ii) communicate the tampering detection and localization results to end users in a manner that is easier to interpret compared to existing forensics approaches.

2 The proposed tampering localization fusion framework

2.1 Selecting tampering localization base methods

In our effort to integrate different forensic approaches into a single framework, we begin by investigating the properties of state-of-the-art to assess what background information can benefit the fusion scheme. We first start by grouping candidate methods based on: (i) their known domain of application (type of tampering), (ii) their detection mechanisms (types of trace) and (iii) their reported performance (reliability of localization and readability of outputs). In order to limit the possible choices between methods, we primarily focus on splicing localization, a very common and effective type of tampering, and we prioritize passive methods that base their detection on analysis of the JPEG compression given its dominance on the Internet.

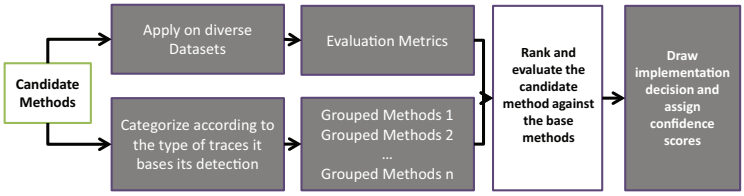

Fig. 1. Diagram of the selection process for localization methods as modules in the image forensics fusion framework.

Figure 1 depicts the general diagram of the selection process for including tampering localization methods as modules in our fusion framework. The various methods are organized in groups depending on the traces they detect, so that, when becoming part of the fusion framework, grouped methods can reinforce each other's results, while results deriving from different groups can be evaluated in a complementary fashion. In parallel, the candidate methods are undergoing a set of evaluation experiments on diverse datasets so as to assess their effectiveness in terms of tampering detection, localization and output readability. For that matter, we utilized the large volume of experiments we conducted in [19] and in [8] as a guide for the selection of the most effective methods. In [19] we evaluated 14 established state-of-the-art methods for image splicing localization that cover the full spectrum of known tampering traces. In [8] we extended the evaluations of seven techniques from [19] (Table 1, rows 1-7) and added a novel algorithm recently developed by us [8] (Table 1, row 8).

The evaluations concern i) the ability of a method to retrieve true positives of tampered images at a low level of false positives (KS@0.05); ii) the ability to achieve good localization of the tampered region within the image (F1); and iii) the readability of the produced heat map, i.e., a high distinction of assigned values for pixels belonging to tampered versus untampered regions, expressed as the range of different binarization thresholds that result in high F1 scores. The

Table 1. Image forensics methods considered as tampering detection modules.

Acronym	Description
DCT [17]	A simple fast detection method for inconsistencies in JPEG DCT coefficient histograms.
ADQ1 [14]	Tampering localization by exploiting the characteristics of double DCT quantization.
BLK [13]	Detection of disturbances of the JPEG 88 block grid in the spatial domain.
NOI1 [16]	Modeling of the local image noise variance using wavelet filtering.
NOI2 [15]	Modeling of the local image noise variance utilizing the properties of the Kurtosis of frequency sub-band coefficients in natural images.
NOI3 [5]	Computes a local co-occurrence map of the quantized high-frequency component of the image and locates inconsistencies in local statistical properties.
CFA1 [6]	Models the Color Filter Array interpolation patterns as a mixture of Gaussian distributions and locates tampering based on detected disturbances.
CAGI [8]	Detects JPEG grid abnormalities in the spatial domain, taking into account the contents of the image. Multiple grid positions are evaluated with respect to a fitting function, and areas of lower (for CAGI) or higher (for inv-CAGI) contribution are identified as tampered.

experiments were performed on three publicly available datasets[1], including both synthetic and real-world tampering cases while their performance robustness when input images are subjected to common post-processing operations was also investigated. Through these evaluations, summarised in Figure 2, we were able to assess, rank and correlate their classification ability and overall performance over a wide spectrum of cases and conditions.

	Fontani et. al			Challenge			Wild Web	
	KS	F1	[thres. range]	KS	F1	[thres. range]	F1	[thres. range]
CAGI	0.70	0.40	**0.3-0.7**	0.17	0.16	**0-0.8**	0.091	0.15-0.6
inv-CAGI	0.31	0.19	0.8-0.95	0.08	0.11	**0-0.95**	0.103	**0.4-0.95**
BLK	0.69	0.21	0.35-0.65	0.21	0.10	0-0.35	0.090	0.3-0.5
NOI3	0.45	0.28	0.05-0.4	0.28	0.18	0.05-0.4	0.092	0.05-0.1
ADQ1	0.48	0.43	**0.05-0.95**	0.13	0.10	0-0.5	0.083	**0.05-0.8**
DCT	0.53	0.33	0.25-0.55	0.25	0.11	0-0.65	0.095	0.3-0.6
NOI1	0.23	0.12	0-0.35	0.21	0.09	0-0.25	0.098	0.1-0.35
NOI2	0.08	0.12	0-0.3	0.04	0.10	0-0.05	0.087	0.05-0.2
CFA1	0.05	0.13	0-0.25	0.01	0.10	0-0.2	0.081	0.1-0.3

(a)

	Fontani et. al	Challenge	Wild Web
CAGI	0.75	0.73	0.51
inv-CAGI	0.21	0.49	0.86
BLK	0.49	0.40	0.31
NOI3	0.47	0.78	0.25
ADQ1	0.89	0.35	0.55
DCT	0.54	0.59	0.50
NOI1	0.18	0.32	0.53
NOI2	0.08	0.07	0.21
CFA1	0.06	0.09	0.11

(b)

	Borda count (scores)			Copeland (scores)			Kenemy-Young (rank)		
	Fon	**Chal**	**WW**	**Fon**	**Chal**	**WW**	**Fon**	**Chal**	**WW**
CAGI	25 (win)	13	12	7	8 (win)	3	2nd	1st	5th
inv-CAGI	9	12	17 (win)	3	6	7 (win)	6th	2nd	1st
BLK	18	12	8	5	4	2	3rd	6th	7th
NOI3	18	13	7	5	3	0	5th	5th	6th
ADQ1	24	10	11	8 (win)	5	1	1st	3rd	2nd
DCT	18	15 (win)	13	5	6	4	4th	4th	3rd
NOI1	10	9	13	2	2	4	7th	7th	4th
NOI2	7	3	5	1	1	0	8th	8th	8th
CFA1	6	3	4	0	0	0	9th	9th	9th

(c)

Fig. 2. Summary of results: a) performance: KS score, max F1 score; threshold binarization range, b) average performance of methods based on normalized KS, F1 and threshold range per dataset; c) rank aggregation results based on Borda count, Copeland and Kenemy-Young voting.

Based on the evaluation results and taking also into account the selection principles described above, a set of five methods were selected as the "base" building blocks of the framework; these include: i) ADQ1 and DCT that both base their detection on analysis of the JPEG compression, in the transform domain; ii) BLK and CAGI that base their detection on analysis of the JPEG

[1] The First IFS-TC Image Forensics Challenge training set [1]; the synthetic Fontani et al. dataset [7]; and the real cases of the Wild Web dataset [18].

compression in the spatial domain; and iii) NOI3 that is a noise-based detector selected as a complementary tool mainly due to its high reported performance and the good interpretability of its produced outputs. Any new candidate method that will be considered for inclusion in the framework, will go through the same evaluation and grouping steps, being additionally ranked against the base methods on these multiple criteria, so as to decide whether it is expected to contribute to the fusion (include or not), how (in which group/what trace), and by how much (confidence weights based on ranks).

2.2 Fusion and refinement of tampering localization masks

The objective of designing a fusion framework is to improve the system's robustness and reliability. If one detector produces noisy or erroneous scores, having other detectors at hand makes it possible to complement, correct and refine the final localization. Figure 3 depicts the block diagram of the proposed fusion framework. For each input image I, we calculate a set of different tampering maps obtained according to the selected subset of detection methods M_k. Based on those, we formulate the fusion task as a labeling problem and we work towards denoting forged pixels with label "0" and authentic pixels with label "1".

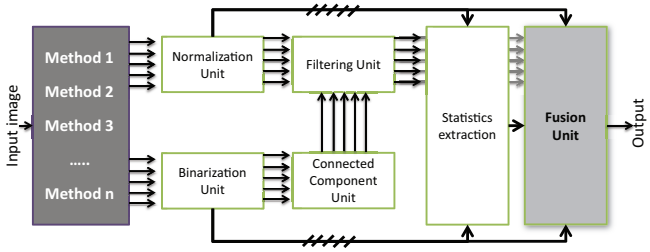

Fig. 3. Block diagram of the proposed fusion framework.

Normalization and Binarization Units: First, output maps are normalized in the [0, 1] range at image level. Next, we are able to cost-effectively automate the binarization of the maps by choosing the appropriate binarization threshold as a value belonging to the respective *safe ranges* per method as these are determined through the analysis that was performed during the selection process (Figure 2). The binarized maps allow easy analysis of their respective spatial and visual properties. We model as valuable and favor outputs that are easy to interpret visually; we are expecting useful maps to have well defined tampered pixel areas that are spatially concentrated and form "blob–like" structures of significant size.

Connected Components Unit: For each binarized map M_k^b, we calculate the center of mass (i.e., centroid) for every 8-connected region that is marked as

tampered:

$$R_c = 1/N \sum_{i=1}^{N} R_i \;\; and \;\; C_c = 1/N \sum_{i=1}^{N} C_i \tag{1}$$

where (R_C, C_C) are the row and column coordinates of the centre of mass of the region under test, R_i, C_i are the i-th pixel coordinates of the region, i.e., matrix elements with zero value, and, N is the total number of pixels in the region. Next we build a feature vector describing the number of the detected connected regions, the location of their centroids (R_C, C_C), the spatial standard deviation of the pixels belonging to a region from their respective centroid, and the image area of each connected region expressed as the smallest possible rectangle (bounding box) containing the pattern of interest. Additionally, for each method, we produce maps of the connected components maps M_k^{cc}, where pixels belonging to each region (hereinafter referred to as blobs) are marked with unique labels.

Filtering Unit: The normalized maps, M_k^n, are forwarded to the Filtering Unit together with the outputs of the Connected Component Unit in order to filter the binary maps, M_k^b. Two types of filtering take place. First, we filter based on findings of each method independently from one another:

- Blobs that present bounding boxes of dimensions bigger than 50% or less that 5% of the image's largest dimension are automatically discarded. This contributes towards fast filtering of spurious, noisy and overall falsely detected results i.e. big blobs that are the result of densely, over-activated maps or isolated small groups of pixels.
- Blobs whose bounding boxes overlap by more than 90% are merged.
- If after the two above steps, the number of blobs is more than five, we calculate the Center of Mass for each M_k^n (as in Eq. 1 but now all pixels are considered and weighted by their actual value in the map) and rank the blobs based on i) their centroids distance from the overall Map centroid (the smaller the distance the better the score), ii) the density of the pixels in the blob (the denser the better the score), and iii) their size (the bigger the size the bigger the score). We then keep the top five based on their mean score in all three criteria.

Second, we perform a content-aware filtering step that depends on the particular methods. Utilizing the content annotation process implemented in CAGI [8] that provides information about areas that are expected to present no noise traces at all (i.e., over and under exposed areas) and the fact that DCT also outputs zero pixel map scores for image blocks of 8-by-8 pixels that share the same intensity value, we are able to filter blobs that may occur as false localizations in BLK and NOI3 outputs; BLK areas that lack any kind of grid pattern are considered tampered; for NOI3 complete lack of noise activates false alarms as they are recognized as inconsistencies in noise distributions. ADQ1 is not triggered by content and thus not affected by this filtering step.

Statistics Extraction: Finally, we extract statistics to automate the evaluation of the outputs' usefulness. These constitute an additional layer of confidence in

selecting from the various intermediate maps the ones that are appropriate for use in the fusion step. We mainly rely on multilevel measurements of the entropy of the data. Image entropy is a quantity used to express the randomness of an image, computed by:

$$E = -\sum_i p_i \log_2 p_i \tag{2}$$

where p_i is the probability that the difference between two adjacent pixels is equal to i. Measuring the entropy of the visual output maps can give us an immediate rough measure of the interpretability of the result. Low entropy corresponds to clear distinction between foreground to background, while noisy outputs with values ranging over many areas will have high entropy. We calculate the following levels of entropy: i) the overall entropy of the normalized map (M^n), per method; ii) the entropy of its binarized counterpart (M^b), and iii) the entropy of each blob region against the entropy of the remaining image.

Additionally, we calculate the Kolmogorov-Smirnov (KS) statistic to compare the value distribution for the different regions of the maps (tampered and untampered) as follows:

$$KS = \max_u |C_1(u) - C_2(u)| \tag{3}$$

where $C_1(u)$ and $C_2(u)$ are the cumulative probability distributions inside and outside the mask, respectively. High KS values indicate that what we have marked as a blob of tampered pixels presents a very different distribution of values compared to the rest of the image.

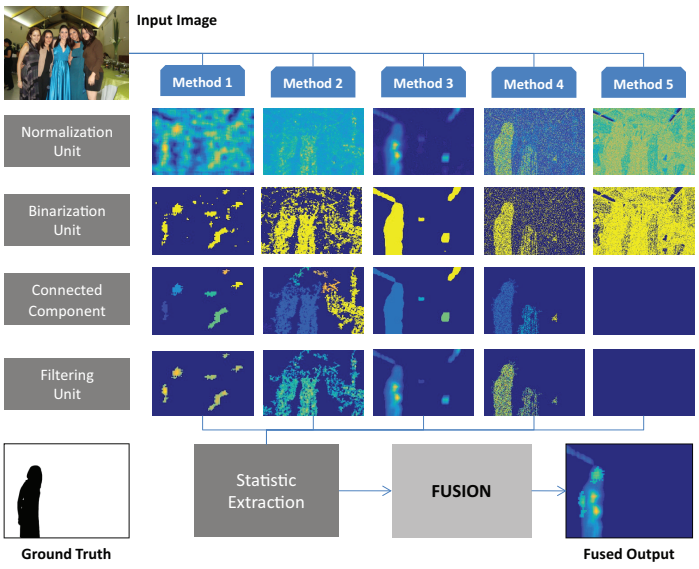

Fig. 4. Intermediate outcomes and final fused output of an example input image (taken from the Challenge dataset [1]).

Fusion Unit: Figure 4 showcases the intermediate steps of the fusion process. This leverages the intermediate calculations to produce a single fused output. To this end, a set of fusion quality properties have been defined:

- *Interpretability of the methods' localization maps:* Maps are ranked and assigned a confidence score, C_i, based on the difference of the entropy before and after the map's binarization.
- *Compatibility between the traces detected by the different methods:* Confidence of a method is reinforced if other methods detecting similar traces also achieve high confidence. Thus, if the C_i is high for more methods from the grouped set of tools (e.g., BLK/CAGI/NOI3 or ADQ1/DCT) the confidence score is boosted.
- *Reliability of the method as measured and assigned after performing extensive evaluations:* The reliability of the tools is also a factor for ranking. All methods are ranked to contribute based on their historical performance (Tables (b) and (c) in Figure as long as their outcome interpretability score surpasses a given threshold.
- *Confidence in the presence or absence of identified tampered regions:* For labeling regions as tampered or not, we also consider the original values of pixels region in the normalized tampering map. The KS statistic is calculated for regions belonging to blobs and background per method. The blobs with highest KS score of the best ranking method serves as our baseline detected tampered region. The refinement of the localization of the blobs is based on comparing it with the blob masks of the other methods in a ranked weighted order.

3 Experimental Evaluation

We tested our proposed fusion framework on two publicly available datasets. The First IFS-TC Image Forensics Challenge training set [1], contains 450 user-submitted forgeries and was designed to serve as a realistic benchmark. Focusing on splicing tampering localization, we excluded cases that were produced by copy-move operations resulting in a set of 306 forgery cases produced through spicing operations only. Tampered images in this dataset are accompanied by Ground Truth (GT) maps. The second dataset is the CASIA V2.0 dataset [2] that contains 5,123 realistically tampered color images of varying sizes. During the tampering process post-processing of spliced boundary regions is also considered. This dataset does not come with GT maps. In order for us to be able to perform localization tests, we manually produced 2,195 reliable GT maps through semi-automated procedures involving image differencing, thresholding and morphological operations. In experiments that follow, when we refer to the CASIA2 dataset we only account for the 2,195 images for which we produced GT binary maps[2].

[2] All produced GT maps are available upon request

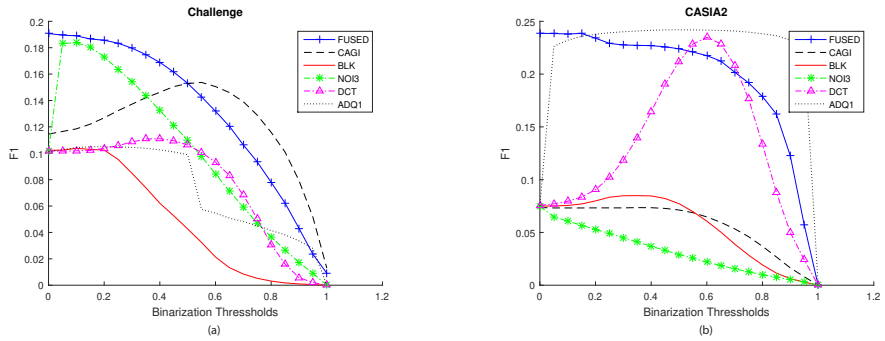

Fig. 5. F1 score curves on the (a) Challenge and (b) CASIA2 datasets for FUSED, CAGI, BLK, NOI3, DCT, ADQ1 of the tampering fusion framework

The overall localization quality and output readability is based on the pixel-wise agreement between the reference mask (GT) and the produced tampering localization heat map and it is measured in terms of the achieved F-score (F1). This evaluation methodology requires the output maps to be thresholded prior to any evaluation. To this end, we first normalize all maps in the [0, 1] range and proceed by successively shifting the binarization threshold by 0.05 increments, calculating the achieved F1 score for every step.

Figure 5 presents the mean F1 scores curves per binarizarion step over the Challenge and CASIA2 collection for the outputs of each individual method along with the fused output. The achieved localization is evaluated by the maximum mean F1 score for each method, at its respective best performing binarization threshold (Table 3).

Table 2. Best mean F1 score and binarization range that allows F1 to remain high ($> 70\%$ of respective maximum F1 score) and reported detections for F1 score $>= 0.7$ at each method's best binarization threshold for Challenge and CASIA2 datasets.

	Challenge				**CASIA2**			
Method	F1 score	Binarization Range	Localizations	Unique Localizations	F1 score	Binarization Range	Localizations	Unique Localizations
FUSED	0.191	0.0-0.4	37	7	0.238	0.0-0.6	490	88
CAGI	0.159	0.3-0.8	16	3	0.073	0.0-0.5	29	4
BLK	0.103	0.0-0.35	8	1	0.085	0.0-0.4	48	3
NOI3	0.183	0.05-0.3	38	15	0.075	0.0-0.05	66	29
ADQ1	0.109	0-0.5	4	1	0.241	0.05-0.9	371	22
DCT	0.105	0-0.65	5	0	0.234	0.55-0.65	488	30

As an indicator of a method's output interpretability we consider the range of the binarization threshold values, where the achieved F1 remains high (> 0.7 of the best reported score). A wide range suggests that the tampered and untampered image regions are characterized by significantly different values in the output maps making the respective heat map easy to interpret. Table 3 also reports the best localized detections achieved per method. The detection threshold was set to 0.7 and the search was performed for the best binarization step

for each method. Unique Localizations corresponds to the number of detections exclusively achieved by that method.

From the experimental results in both datasets we can see that the fused output heat maps achieve high F1 scores over a wide range of thresholds. This verifies that the method produces outputs that exhibit increased localization ability and interpretability. In the Challenge dataset, the next best method (NOI3) achieves similar localization scores but is somewhat worse in terms of interpretability, while all other methods achieve significantly lower F1 scores. In CASIA2, the fusion method is the second best performing method in terms of F1 scores, while it still presents the best interpretability with F1 scores remaining high for a wider range of binarization steps. DCT, which is the leading method in this dataset, is significantly outperforming the rest of the individual methods, which is probably due to the tampering process followed in this specific dataset. The fusion framework manages to produce outputs that generally localize tampering better than most of the individual methods (Figure 5(b)) but, in its current state, does not take full advantage of the very good DCT localizations in building its final output. Instead, while trying to construct hybrid outputs with low risk by collectively examining the various outputs and not heavily relying on only one method, the good DCT localizations were undermined by the many unsuccessful localizations of other methods. Motivated by these findings, assigning better weighting factors and ranking criteria will be at the heart of our next efforts.

Finally, in both datasets the fused method reports a high number of absolute localizations, which is indicating that the fusion criteria set in this framework manage to take advantage of the correctly localized outputs of the individual methods, and more importantly the framework contributes additional unique localizations through fusion and refinement, especially so in the CASIA2 database. Various localization outcomes are depicted in the Figure 6.

Overall, this first set of experimental evaluations verifies the importance of exploiting the available state-of-the-art methods in a manner that improves the robustness and reliability of the system. In our next steps, we will continue to further test and refine the framework, while we also plan to introduce more localization methods in the system.

4 Conclusions

In this paper, we addressed the splicing tampering localization problem focusing on traces and methods that apply to JPEG images. To this end, we proposed an extensible tampering localization fusion and map refinement framework that combines multiple state-of-art techniques by exploiting their complementarities. We performed and took advantage of extensive evaluation experiments with the goal of selecting the most appropriate "base" methods to be fused so as to produce a single refined localization map outcome. Our experimental findings indicate that the fused output achieves high performance and interpretability by managing to exploit the correctly localized outputs of the individual meth-

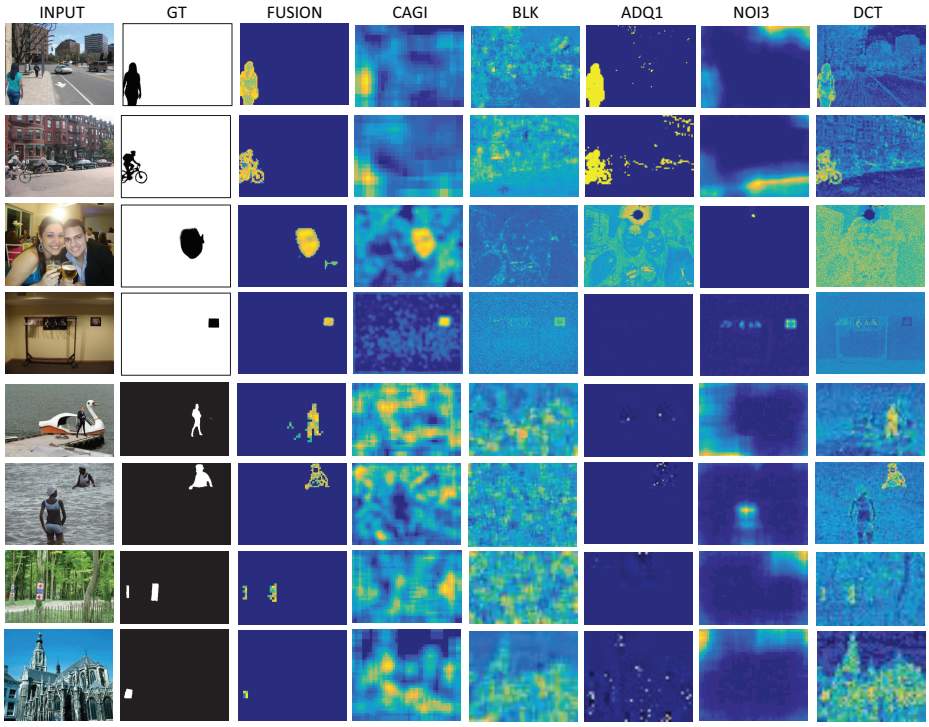

Fig. 6. Various examples of tampering localization outputs from the fusion framework and the individual methods; for the first four rows, images are taken from the Challenge dataset, for the last four rows images are taken from the CASIA2 dataset.

ods while contributing with unique accurate tampering localizations. While we consider the results of our fusion approach promising, we also recognize the fact that the fusion is based on hard-coded expert knowledge that is directly implemented in the fusion criteria and rules. To this end, we plan to also investigate the potential of fusion approaches based supervised learning.

Acknowledgements

This work was partially funded by the European Commission under contract numbers H2020-825297 WeVerify and H2020-700024 TENSOR.

References

1. Report on the IEEE-IFS Challenge. http://ifc.recod.ic.unicamp.br/, accessed: 20-03-2016
2. "CASIA TIDE v2" [online] (2009), http://forensics.idealtest.org/casiav2/, accessed: 20-03-2017
3. Barni, M., Costanzo, A.: A fuzzy approach to deal with uncertainty in image forensics. Signal Processing: Image Communication **27**(9), 998–1010 (2012)

4. Chetty, G., Singh, M.: Nonintrusive image tamper detection based on fuzzy fusion. Inte. Journal of Computer Science and Network Security **10**(9), 86–90 (2010)
5. Cozzolino, D., Poggi, G., Verdoliva, L.: Splicebuster: a new blind image splicing detector. In: Information Forensics and Security (WIFS), 2015 IEEE International Workshop on. pp. 1–6. IEEE (2015)
6. Ferrara, P., Bianchi, T., De Rosa, A., Piva, A.: Image forgery localization via fine-grained analysis of cfa artifacts. IEEE Trans. on Information Forensics and Security **7**(5), 1566–1577 (2012)
7. Fontani, M., Bianchi, T., De Rosa, A., Piva, A., Barni, M.: A framework for decision fusion in image forensics based on dempster–shafer theory of evidence. IEEE Transactions on Information Forensics and Security **8**(4), 593–607 (2013)
8. Iakovidou, C., Zampoglou, M., Papadopoulos, S., Kompatsiaris, Y.: Content-aware detection of jpeg grid inconsistencies for intuitive image forensics. Journal of Visual Communication and Image Representation **54**, 155–170 (2018)
9. Kaur, M., Gupta, S.: A fusion framework based on fuzzy integrals for passive-blind image tamper detection. Cluster Computing **22**(5), 11363–11378 (2019)
10. Korus, P.: Digital image integrity–a survey of protection and verification techniques. Digital Signal Processing **71**, 1–26 (2017)
11. Korus, P., Huang, J.: Multi-scale fusion for improved localization of malicious tampering in digital images. IEEE Trans. on Image Proc. **25**(3), 1312–1326 (2016)
12. 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)
13. Li, W., Yuan, Y., Yu, N.: Passive detection of doctored jpeg image via block artifact grid extraction. Signal Processing **89**(9), 1821–1829 (2009)
14. 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)
15. Lyu, S., Pan, X., Zhang, X.: Exposing region splicing forgeries with blind local noise estimation. International Journal of Computer Vision **110**(2), 202–221 (2014)
16. Mahdian, B., Saic, S.: Using noise inconsistencies for blind image forensics. Image & Vision Com. **27**(10), 1497–1503 (2009)
17. Ye, S., Sun, Q., Chang, E.C.: Detecting digital image forgeries by measuring inconsistencies of blocking artifact. In: 2007 IEEE International Conference on Multimedia and Expo. pp. 12–15. IEEE (2007)
18. Zampoglou, M., Papadopoulos, S., Kompatsiaris, Y.: Detecting image splicing in the wild (web). In: IEEE International Conference on Multimedia & Expo Workshops (ICMEW). pp. 1–6. IEEE (2015)
19. Zampoglou, M., Papadopoulos, S., Kompatsiaris, Y.: A large-scale evaluation of splicing localization algorithms for web images. Multimedia Tools and Applications pp. 4801–4834 (2016)