SPOTTING THE DIFFERENCE: CONTEXT RETRIEVAL AND ANALYSIS FOR IMPROVED FORGERY DETECTION AND LOCALIZATION

Joel Brogan[1], *Paolo Bestagini*[2], *Aparna Bharati*[1], *Allan Pinto*[1,3], *Daniel Moreira*[1]
Kevin Bowyer[1], *Patrick Flynn*[1], *Anderson Rocha*[1,3], *and Walter Scheirer*[1]

[1]Department of Computer Science and Engineering, University of Notre Dame, US
[2]Department of Electronics, Information and Bioengineering, Politecnico di Milano, Italy
[3]Institute of Computing, University of Campinas, Brazil

ABSTRACT

As image tampering becomes ever more sophisticated and commonplace, the need for image forensics algorithms that can accurately and quickly detect forgeries grows. In this paper, we revisit the ideas of image querying and retrieval to provide clues to better localize forgeries. We propose a method to perform large-scale image forensics on the order of one million images using the help of an image search algorithm and database to gather contextual clues as to where tampering may have taken place. In this vein, we introduce five new strongly invariant image comparison methods and test their effectiveness under heavy noise, rotation, and color space changes. Lastly, we show the effectiveness of these methods compared to passive image forensics using Nimble [1], a new, state-of-the-art dataset from the National Institute of Standards and Technology (NIST).

Index Terms— image forensics, forgery detection, splicing detection, context-aware digital forensics, tampering heat maps

1. INTRODUCTION

Now that advanced photo editing software is readily available, image tampering has become ubiquitous, and the traces left behind by such modifications are becoming increasingly hard to detect. Regardless of intention, this trend has undermined the value of images as viable evidence in a number of domains. To examine cases of tampering, a two-fold task can be pursued. First, tampering within an image must be detected without the use of pre-embedded information (*e.g.*, a key). This is known as **Passive Digital Image Forensics (PDIF)** [2]. Second, the tampered area must be accurately localized if it is to be considered for further analysis.

In this paper, we improve upon the image comparisons for contextual-clue-based PDIF offered in Gaborini et al. [3]. In a PDIF scenario, a **contextual clue** can be interpreted as the incongruities between the image in question and images collected from outside sources. While [3] presents a basic method for contextual-clue-based image forensics, our work proposes a fully-automatic, efficient, and scalable search-and-compare framework for image forensics. Additionally, our work offers highly noise-invariant comparison algorithms. This framework treats the image under question as a search probe and does not rely on having well-posed and pre-gathered images, as in [3]. Accordingly, the system described in this paper provides improved performance over traditional PDIF methods, and automatic extraction of alien regions within composite images. These properties make the system potentially useful for subsequent tasks like image provenance analysis, *i.e.*, the study of how and when an image has been modified over time. The proposed system utilizes a fast, light-weight search engine based on KD-Trees [4] optimized for SURF descriptors [5] to retrieve a set of near-duplicate images related to the probe. These images are then

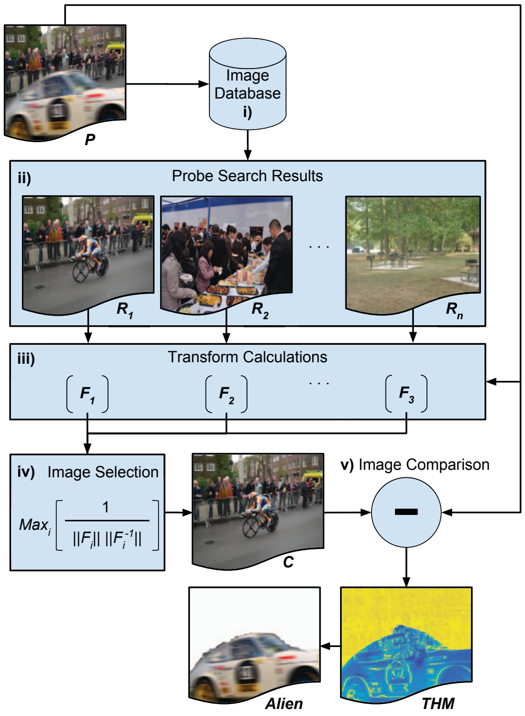

Fig. 1. An overview of the context-based search-and-compare framework. Probe image P is used as input to a database search in step **i)**, which returns a list of results R, shown in step **ii)**. The transforms between P and R are found in step **iii)**. The top-related image C is chosen in step **iv)**. In step **v)**, C is compared to P using one of five proposed algorithms from Section 4, yielding a tampering heat map (THM) used to extract the alien region of P.

This material is based on research sponsored by DARPA and Air Force Research Laboratory (AFRL) under agreement number FA8750-16-2-0173. Hardware support was generously provided by the NVIDIA Corporation. We also thank the financial support of FAPESP (Grant #2015/19222-9), CAPES (DeepEyes Grant) and CNPq (Grant #304472/2015-8).

compared to the probe to produce *contextual* clues as to the location of modifications. An overview of the entire framework is shown in Figure 1.

Three main modifications are typically performed on images: splicing, copy-move, and re-sampling. As a case study, this paper considers only instances of image splicing, also known as image compositing. In general, splicing includes an original **Host Image**, onto which regions from one or more **Donor Images** are added to create a **Composite Image**. Regions from a donor image present on a host image are known as **Alien Regions**. Most classic PDIF approaches exploit the nature of the digital photographic process to find evidence of these alien regions by using the data contained within the image in question. In contrast, our method uses outside information (*i.e.*, context) gathered from a search engine to collect such evidence.

Within our search-and-compare framework, we analyze the performance of five novel signal-processing-based image comparison techniques for extracting **Tampering Heat Maps (THMs)** that indicate regions of the image where tampering might have occurred. We intentionally chose to avoid deep learning-based techniques because we found that sufficient training data that fully captures the variations seen in manipulated images is not readily available, nor is it clear how a comprehensive training dataset could be assembled. Additionally, the computational resources required for deep learning techniques are prohibitive to the massively scalable and efficient system needed to accomplish the proposed task. Using the five signal-processing based comparison techniques, we test the proposed image-search-and-compare algorithms on the Nimble dataset, a dataset newly developed by NIST for the task of image provenance [1], mixed in with a set of 1 million distractor images. Finally, we compare the results of our image search method to a set of 13 state-of-the-art PDIF splicing algorithms to show the effectiveness of the proposed method.

2. RELATED WORK

Forgery detection and localization based on single image analysis can be performed using a variety of traditional PDIF methods [6,7]. For example, one set of techniques exploit JPEG artifacts [8–14]. Another set of approaches utilize Color Filter Array (CFA) footprints [15,16]. Yet another set of methods deal with detecting natural noise inconsistencies within spliced images [17–19]. Additionally, methods based on Error Level Analysis (ELA) [20] can be used.

More recently, image provenance research has shown the possibility of conducting even deeper forensic analysis by jointly considering sets of correlated images [21–23]. Indeed, if multiple images are available, it is possible to achieve robust forgery localization results through image comparison [24]. For instance, using scaled thumbnail meta-data of images to localize forgeries [25] can provide high-accuracy localization maps of image tampering. When thumbnail data is unavailable, near-duplicate images have also been used to build THMs from contextual clues as to where forgeries occur [3]. While these ideas show promise, they provide no automatic method of retrieving contextually relevant images for comparison. Additionally, the comparison methods found in the literature offer relatively poor invariance to color changes, noise, morphing, and compression between images.

By always considering a scenario with multiple images, sophisticated methods for patch comparison using deep learning have been proposed [26,27]. However, the corresponding models were trained for highly specific keypoint matching scenarios. Techniques like these do not capture the variations present in realistic forgeries, and thus cannot be used as-is in a real-world forensic scenario.

3. IMAGE SEARCH ALGORITHM

The first step of the search-and-compare process is image search, as shown in Fig. 1. The search engine must adhere to multiple constraints. First, the system must provide fast and scalable indexing and searching. For the proposed search-and-compare method to be effective, we must have an extremely large database (on the order of one million images in this work) to compare against.

In our proposed search method, we extract 500 SURF keypoints [5] with the relative 64-dimensional descriptors to describe each image. We utilize a KD-tree forest scheme similar to what was used in [4] to scalably index and search the descriptors in the database. This method provides a higher likelihood of returning images that contain objects directly comparable to the probe [28]. According to the described scheme, once a probe image P is queried, the system returns a set of N contextually similar images along with any possible near-duplicate images $R_n, n \in [1, N]$.

4. IMAGE COMPARISON FRAMEWORK

Once the images $R_n, n \in [1, N]$ are retrieved from the image database using probe image P as a query, they must be sorted and filtered to ensure only truly relevant images to our probe are compared. To accomplish this goal, for each image R_n, SURF keypoints and features are re-calculated. Then, the 3×3 affine matrix F_i mapping points of R_n to the coordinate system of P is computed using keypoint matching and the MSAC method, allowing for tighter geometric constraints than RANSAC [29]. To generate a list of images with content that best geometrically matches the probe, we rank each R_n by the Reciprocal Frobenius Condition of its linked affine transform F_n as

$$\text{RFN}_n = \frac{1}{\|F_n\| \|F_n^{-1}\|}, \tag{1}$$

where $\|\cdot\|$ is the Frobenius norm of a matrix.

We assert that the greater the RFN value, the more suitable R_n will be to the comparison task. Therefore, even though multiple images could be used to provide multiple clues, we decided in this work to only select as comparison image C the image R_n with the highest RFN Value warped using the affinity matrix F_n

$$C = \text{warp}(R_{\hat{n}}, F_{\hat{n}}), \quad \hat{n} = \arg\max_n(\text{RFN}_n), \tag{2}$$

where warp applies the affine transform to the image. Once this image has been selected, we must compare the probe image P to the result image C to produce a THM as shown in Fig. 1. To achieve a reliable comparison, an algorithm must overcome differences in image noise, colorspace changes, and slight rotations and translation. For this purpose, we propose the following five algorithms.

1. PSNR of Gaussian Image Residual. (IRPSNR) We define $\mathcal{G}(I, \sigma_\mathcal{G})$ to be the convolution of image I with a Gaussian kernel with standard deviation $\sigma_\mathcal{G}$. We set $\sigma_\mathcal{G} = 4$, as we found it to provide optimal local blurring to allow for invariance to small translations and rotations. To generate a tampering heat map, we compute the pixel-wise Peak Signal to Noise Ratio (PSNR) between the Gaussian blurred versions of P and C as

$$\text{THM}_{\text{PSNR}} = \log_{10} \frac{1}{|\mathcal{G}(P, \sigma_\mathcal{G}) - \mathcal{G}(C, \sigma_\mathcal{G})|^2 + 1}, \tag{3}$$

where all operations are pixel-wise, and the plus 1 in the denominator is used for regularization. Portions of P matching the respective portions of C will contribute to the THM_{PSNR} with high values. Tampered areas should be exposed by low THM_{PSNR} values.

2. Pseudo-PRNU Patch-wise Comparison. Images shot with the same camera are characterized by a multiplicative noise pattern known as Photo-Response Non-Uniformity (PRNU) [30]. This noise residual is characteristic of the capturing device, and can be used for camera attribution [30], for tampering localization [31], or even to assess whether two images come from the same device [32]. As near-duplicate images are acquired by the same camera by definition, we can rely on image noise patch-wise comparison to detect local inconsistencies due to splicing.

Given two corresponding patches of P and C (*e.g.*, the first 64×64 pixel block in the top-left corner of each image), PRNU information extracted from those patches should correlate very well if only global transformations (*e.g.*, color corrections, blurring, compression, etc.) have been applied to P or C. Conversely, PRNU information does not correlate at all if one of the two patches has been spliced from a picture obtained from a different device.

To exploit this property, let us define the noise residuals

$$\tilde{C} = C - \mathcal{W}(C), \qquad \tilde{R} = R - \mathcal{W}(R), \tag{4}$$

where $\mathcal{W}(\cdot)$ is the wavelet-based denoising operation used in [30]. According to [30], the computed $\tilde{C}$ and $\tilde{R}$ contain PRNU traces. Therefore, it is possible to correlate them patch-wise for the THM

$$\text{THM}_{\text{noise}} = \text{average}\left(\frac{\tilde{C} \cdot \tilde{R}}{||\tilde{C}|| \cdot ||\tilde{R}||}\right), \tag{5}$$

where '$\cdot$' represents pixel-wise multiplication, $||\cdot||$ returns the Frobenius norm, and average$(\cdot)$ computes the moving average on 64×64 pixel blocks. The mask $\text{THM}_{\text{noise}}$ should present high values corresponding to areas that are common to P and C, and low values in tampered regions.

3. Structural Similarity Comparison. (SSIM) This method uses the calculated pixel-wise Structural Similarity Index Measure (SSIM) between images P and C. [33] We define the structural similarity-based THM as

$$A = (2\mu_P\mu_C + (0.01D)^2)(2\sigma_{PC} + (0.03D)^2), \tag{6}$$

$$B = (\mu_P^2 + \mu_C^2 + (0.01D)^2)(\sigma_P^2 + \sigma_C^2 + (0.03D)^2), \tag{7}$$

$$\text{THM}_{\text{SSIM}} = \frac{A}{B}, \tag{8}$$

where μ and σ are the local neighborhood means and standard deviations of P and C with a neighborhood radius of 32 pixels, σ_{PC} is the local covariance of the local image patches, and D is the dynamic contrast of the images. Similar to the PRNU-based mask, THM_{SSIM} should assume low values in correspondence of tampered areas.

4. HSV Histogram Patch-wise Comparison. From images P and C, local histogram patches $H_{P_{xy}}$ and $H_{C_{xy}}$ are calculated using a local neighborhood radius of 13 pixels. We use the probability of the two-sample Kolmogorov-Smirnov Test [34] being equal to or more extreme than the observed value of the null hypothesis that $\text{CDF}_{H_P} = \text{CDF}_{H_C}$, where CDF_H is the Cumulative Distribution Function of a given histogram patch. This value is calculated for each corresponding patch to generate a THM:

$$Pr_{xy}(\max_a(|\mathcal{Q}(a, H_{C_{xy}}) - \mathcal{Q}(a, H_{P_{xy}})|) \mid H_{P_{xy}}), \tag{9}$$

$$\text{THM}_{\text{HPC}} = Pr, \tag{10}$$

where $\mathcal{Q}(a, \vec{b})$ is the proportion of $\vec{b}$ less than or equal to a. This allows us to test, in a manner invariant to small rotations, the idea that each patch contains a similar distribution.

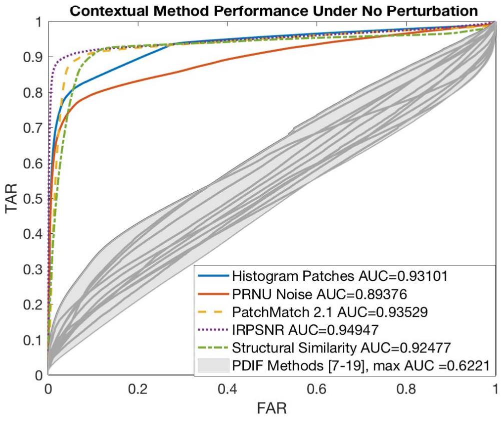

Fig. 2. Five variants of the proposed context-based search-and-compare framework compared against 13 widely-used PDIF techniques. The gray area represents the maximum and minimum performance of each PDIF algorithm. Even the worst performing contextual method, PRNU matching, performs **44%** better than the top PDIF comparison algorithm. For the best performing variants, we see that IRPSNR performs **1.5%** better than PatchMatch 2.1.

5. PatchMatch 2.1. For an additional method, we use the PatchMatch algorithm [35] for image comparison. Specifically, we utilize the rotation and scale-invariant version, using 20 iterations. To generate the relative THM_{PM}, we calculate the L2 match distance of patches within the image. In other words, we associate to each 8×8 patch of P the L2 distance from the patch of C that best approximates it. If a patch of P cannot be well approximated with any patch of C (*i.e.*, in case of tampering), its L2 distance will be high.

5. EXPERIMENTS AND RESULTS

Dataset. For the purpose of our experiments, we utilize a new, state-of-the-art dataset from NIST called Nimble [1]. The dataset contains a subset specifically for splicing operations, dubbed Nimble-Splice, which contains a total of 288 probe images, each having been hand-composited from a gallery of 874 images. A host, donor, and binary tamper mask image are provided with each probe. The masks represent ground-truth data to compare our generated THMs against.

To simulate a real-world scale, we take the 874 gallery images from Nimble-Splice and add one million distractor images provided by RankOne Inc. Medifor program [1]. This allows us to test the effectiveness of indexing and subsequently finding relevant images for comparison. We call this hybrid dataset "Nimble World" (NW).

Framework Setup. Using the method described in Section 3, we extract features and index all 1,000,874 images into a KD-tree forest. For each probe we return the top $I = 100$ results from the KD-forest search. We find that our search algorithm returns relevant results with *Recall at Rank 25* = 99.5%. The top scoring image is registered to the probe using the affine transform described in Section 4. We test all five proposed algorithms to compare the probe and result and generate a THM. Using a sliding threshold to localize forgeries (i.e., image differences), we generate ROCs for pixel-wise classification of our THMs compared to ground-truth masks. [2]

[1] `http://medifor.rankone.io/`

[2] Code for these experiments available at `https://gitlab.com/notredame-provenance/Context_Comparison`

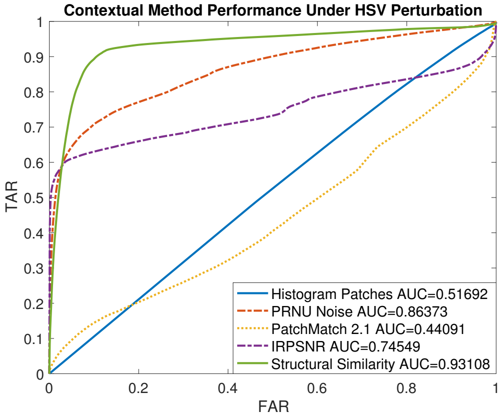

Fig. 3. Performance of all five contextual image forensics methods under random HSV space transformations. The SSIM approach is most invariant to color changes, while most other algorithms suffer. Patchmatch 2.1 performs poorly in such scenarios.

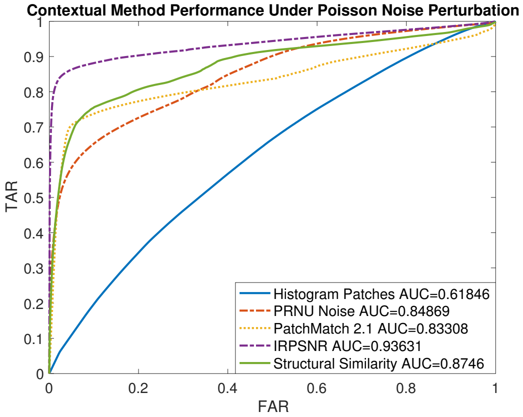

Fig. 4. Performance of all five contextual image forensics methods under the addition of Poisson Noise. The IRPSNR approach is most invariant to noise.

Experiment 1. The first experiment studies the performance of the methods we propose in this paper and 13 current state-of-the-art PDIF splicing-detection algorithms. These algorithms include JPEG artifact analysis [8–14], CFA analysis [15, 16], image noise analysis [17–19], and ELA [20]. The THMs generated by these methods were compared to the ground-truth masks using the same thresholding method to generate ROC curves. In Fig. 2 we see a large performance gap between our five proposed algorithms within the search-and-compare framework, which perform the best, and the set of 13 PDIF methods from the literature.

Experiment 2. The second experiment analyzes the performance of each of the five proposed image-to-image comparison algorithms in the presence of non-negligible noise, color, and rotation perturbations. These perturbations, performed on the result images R that are used in individual comparisons, simulate real-world artifacts likely to be found in images indexed from the web.

To perturb the color space of gallery images in the NW dataset, we randomly fluctuated the HSV channels of each image independently between 0 and 20%. The results for color space perturbing can be seen in Fig. 3. Similarly, to perturb the noise within gallery images, random amounts of Poisson noise were added to each gallery image. Results for noise can be seen in Fig. 4. Lastly, we used random rotations between -15° and 15° *after* the result registration phase, to simulate an incorrect registration caused by erroneous keypoint matches. Results for rotation can be seen in Fig. 5.

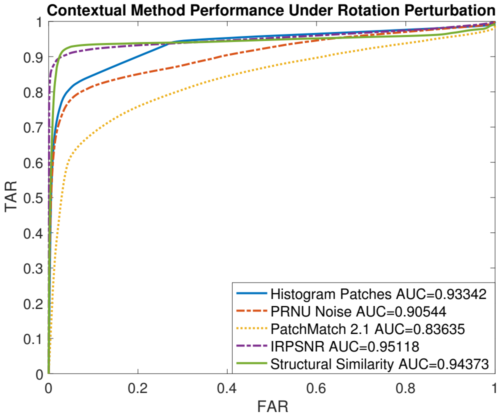

Fig. 5. Performance of all five contextual image forensics methods under random small-angle image rotations. The PatchMatch approach is most negatively impacted; other algorithms perform well.

We concluded that a test on scale-based perturbations was not necessary after observing that algorithm performance was *nearly identical* even for large-scale fluctuations.

6. CONCLUSION

Assuming a search process that provides relevant results in the presence of a large number of distractor images, a context-based search-and-compare framework for image forensics is greatly superior at localizing areas of tampering than traditional PDIF methods. Further, we conclude that of the methods tested, the IRPSNR method provided the most invariance to rotation and noise space perturbations, while the SSIM method had the least performance deterioration under color space perturbation. The PRNU patch-wise comparison algorithm was the most stable over all three perturbation cases, while Histogram patch-wise comparison and PatchMatch had the least invariance to all perturbation methods.

It should be noted that while the search-and-compare paradigm provides improved results over traditional PDIF methods, search-and-compare fails in cases where original or near-duplicate images are not present. However, these instances can be detected simply by testing the maximum Reciprocal Condition of each probe-to-result transform, and thresholding at an empirically determined level.

With respect to further directions for this work, the search-and-compare paradigm introduced in this paper lends itself nicely to the task of image provenance analysis and multimedia phylogeny [22, 23]. To construct accurate provenance graphs that express relationships between tampered images, we must dig down into the localized tampered objects within a composite to further determine each object's origin. The THMs produced by our framework can be easily segmented into tamper regions. These regions can then be directly analyzed to determine the nature of the tampering. Thus, the work we have described in this paper should not be treated as a standalone contribution, but placed in the wider context of digital forensics.

7. REFERENCES

[1] National Inst. of Standards and Technology, "The 2016 Nimble challenge evaluation dataset," `https://www.nist.gov/itl/iad/mig/nimble-challenge`, Jan. 2016.

[2] A. Rocha, W. J. Scheirer, T. E. Boult, and S. Goldenstein, "Vision of the unseen: Current trends and challenges in digital image and video forensics," *ACM Computing Surveys (CSUR)*, vol. 43, October 2011.

[3] L. Gaborini, P. Bestagini, S. Milani, M. Tagliasacchi, and S. Tubaro, "Multi-clue image tampering localization," in *2014 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2014, pp. 125–130.

[4] C. Silpa-Anan and R. Hartley, "Optimised kd-trees for fast image descriptor matching," in *IEEE Conf. on Computer Vision and Pattern Recogn.*, 2008, pp. 1–8.

[5] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European Conference on Computer Vision*, 2006, pp. 404–417.

[6] D. Cozzolino, D. Gragnaniello, and L. Verdoliva, "Image forgery localization through the fusion of camera-based, feature-based and pixel-based techniques," *IEEE International Conference on Image Processing (ICIP)*, 2014.

[7] M. Zampoglou, S. Papadopoulos, and Y. Kompatsiaris, "Large-scale evaluation of splicing localization algorithms for web images," *Mult. Tools and Applications*, pp. 1–34, 2016.

[8] Z. Lin, J. He, X. Tang, and C.-K. Tang, "Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis," *Pattern Recognition*, vol. 42, no. 11, pp. 2492–2501, 2009.

[9] T. Bianchi, A. De Rosa, and A. Piva, "Improved DCT coefficient analysis for forgery localization in JPEG images," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 2444–2447.

[10] I. Amerini, R. Becarelli, R. Caldelli, and A. Del Mastio, "Splicing forgeries localization through the use of first digit features," in *2014 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2014, pp. 143–148.

[11] W. Li, Y. Yuan, and N. Yu, "Passive detection of doctored JPEG image via block artifact grid extraction," *Signal Processing*, vol. 89, no. 9, pp. 1821–1829, 2009.

[12] S. Ye, Q. Sun, and E.-C. Chang, "Detecting digital image forgeries by measuring inconsistencies of blocking artifact," in *IEEE Int. Conf. on Multimedia and Expo*, 2007, pp. 12–15.

[13] T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," *IEEE Trans. on Information Forensics and Security*, vol. 7, no. 3, pp. 1003–1017, 2012.

[14] H. Farid, "Exposing digital forgeries from JPEG ghosts," *IEEE Trans. on Information Forensics and Security*, vol. 4, no. 1, pp. 154–160, 2009.

[15] P. Ferrara, T. Bianchi, A. De Rosa, and A. Piva, "Image forgery localization via fine-grained analysis of CFA artifacts," *IEEE Trans. on Information Forensics and Security*, vol. 7, no. 5, pp. 1566–1577, 2012.

[16] A. E. Dirik and N. D. Memon, "Image tamper detection based on demosaicing artifacts," in *ICIP*, 2009, pp. 1497–1500.

[17] B. Mahdian and S. Saic, "Using noise inconsistencies for blind image forensics," *Image and Vision Computing*, vol. 27, no. 10, pp. 1497–1503, 2009.

[18] S. Lyu, X. Pan, and X. Zhang, "Exposing region splicing forgeries with blind local noise estimation," *International Journal of Computer Vision*, vol. 110, no. 2, pp. 202–221, 2014.

[19] J. Wagner, "Noise analysis for image forensics," `https://29a.ch/2015/08/21/noise-/analysis-for-image-forensics`, Aug. 2015.

[20] N. Krawetz, "A picture's worth... digital image analysis and forensics," *Black Hat Briefings*, pp. 1–31, 2007.

[21] A. De Rosa, F. Uccheddu, A. Costanzo, A. Piva, and M. Barni, "Exploring image dependencies: a new challenge in image forensics," in *IS&T/SPIE Electronic Imaging*, 2010, pp. 75410X–75410X.

[22] Z. Dias, A. Rocha, and S. Goldenstein, "First steps toward image phylogeny," in *2010 IEEE International Workshop on Information Forensics and Security (WIFS),*, 2010, pp. 1–6.

[23] Z. Dias, A. Rocha, and S. Goldenstein, "Image phylogeny by minimal spanning trees," *IEEE Trans. on Information Forensics and Security*, vol. 7, no. 2, pp. 774–788, 2012.

[24] P. Bestagini, M. Tagliasacchi, and S. Tubaro, "Image phylogeny tree reconstruction based on region selection," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016.

[25] M. Kirchner, P. Winkler, and H. Farid, "Impeding forgers at photo inception," in *IS&T/SPIE Electronic Imaging*, 2013, pp. 866504–866504.

[26] S. Zagoruyko and N. Komodakis, "Learning to compare image patches via convolutional neural networks," in *IEEE Conf. on Computer Vision and Pattern Recog.*, 2015, pp. 4353–4361.

[27] X. Han, T. Leung, Y. Jia, R. Sukthankar, and A. C. Berg, "Matchnet: Unifying feature and metric learning for patch-based matching," in *IEEE Conf. on Computer Vision and Pattern Recog.*, 2015, pp. 3279–3286.

[28] A. Krizhevsky and G. E. Hinton, "Using very deep autoencoders for content-based image retrieval," in *ESANN*, 2011.

[29] S. Choi, T. Kim, and W. Yu, "Performance evaluation of ransac family," *Journal of Computer Vision*, vol. 24, no. 3, pp. 271–300, 1997.

[30] J. Lukáš, J. Fridrich, and M. Goljan, "Digital Camera Identification from Sensor Noise," *IEEE Transactions on Information Security and Forensics (TIFS)*, vol. 1, pp. 205–214, 2006.

[31] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš, "Determining image origin and integrity using sensor noise," *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 3, pp. 74–90, 2008.

[32] M. Goljan, M. Chen, and J. Fridrich, "Identifying common source digital camera from image pairs," in *IEEE International Conference on Image Processing (ICIP)*, 2006.

[33] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[34] H. W. Lilliefors, "On the kolmogorov-smirnov test for normality with mean and variance unknown," *Journal of the American statistical Association*, vol. 62, no. 318, pp. 399–402, 1967.

[35] C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman, "Patchmatch: A randomized correspondence algorithm for structural image editing," *ACM Transactions on Graphics-TOG*, vol. 28, no. 3, pp. 24, 2009.