

SHADOW REMOVAL DETECTION AND LOCALIZATION FOR FORENSICS ANALYSIS

*S. K. Yarlagadda, D. Güera,
D. M. Montserrat, F. M. Zhu, E. J. Delp*

Video and Image Processing Laboratory (VIPER),
Purdue University

P. Bestagini, S. Tubaro

DEIB
Politecnico di Milano, Milan, Italy

ABSTRACT

The recent advancements in image processing and computer vision allow realistic photo manipulations. In order to avoid the distribution of fake imagery, the image forensics community is working towards the development of image authenticity verification tools. Methods based on shadow analysis are particularly reliable since they are part of the physical integrity of the scene, thus detecting forgeries is possible whenever inconsistencies are found (e.g., shadows not coherent with the light direction). An attacker can easily delete inconsistent shadows and replace them with correctly cast shadows in order to fool forensics detectors based on physical analysis. In this paper, we propose a method to detect shadow removal done with state-of-the-art tools. The proposed method is based on a conditional generative adversarial network (cGAN) specifically trained for shadow removal detection.

Index Terms— Image forensics, shadow removal, CNN, cGAN

1. INTRODUCTION

Image editing tools are widely available. It is possible to download professional image manipulation tools (e.g., Photoshop), to use image editing operations directly from web interfaces (e.g., Pixlr), or even more easily to automatically forge a picture using completely unsupervised tools (e.g., FaceSwap). If maliciously edited images are shared online or distributed through broadcast channels, their impact in terms of opinion formation and fake news distribution can cause serious social consequences.

Many blind image forensic tools have been developed in the literature through years [1, 2, 3]. Among these techniques, many focus on verifying image digital integrity. These methods typically exploit statistical traces left by alterations of digital signals and can be used for a wide variety of applications (e.g., detecting the originating device or camera model [4, 5], general forgeries [6, 7], re-sampling [8, 9, 10] and multiple compressions [11, 12]). The main issue behind many of these methods is that they rely on a strict set of assumptions that cannot always be verified and they suffer from multiple editing operations being applied altogether. Many methods can be fooled if “laundering” operations that scramble image statistics

This material is based on research sponsored by DARPA and Air Force Research Laboratory (AFRL) under agreement number FA8750-16-2-0173. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA and Air Force Research Laboratory (AFRL) or the U.S. Government.

are used to edited images (e.g., small resizing and cropping, subtle global operations, and re-compressions).

Other forensics techniques rely on verifying image physical integrity. This means detecting whether an image is authentic by checking physical consistencies in reflections [13], lightning [14, 15], shadows [16, 17, 18], and other constraints that must be verified in a real-world photoshoot. As an example, by knowing the direction of light illuminating a scene, it is possible to estimate shadow directions. In the same way, it is possible to estimate whether all objects in the scene present a shadow coherent with the other ones [16]. The drawback of this technique is that they are often semi-supervised (e.g., the analyst should manually check for shadow cast, lightning directions, etc.). However, they are innately robust against laundering. Indeed, as long as the image semantic content remains unchanged, it is still possible to verify physical inconsistencies despite resizing, rotations, or image re-compression operations. To fool these techniques, an expert manipulator must take into account the laws of physics, and retouch the picture accordingly.

In the past fooling physical integrity detection was considered a challenging task, nowadays this might be only partly true. Indeed, thanks to the advancement in machine learning and signal processing, many image editing operations can be used in an almost automatic fashion, not always requiring the hand of a professional. Among these, many methods for shadow removal have been proposed in the literature [19, 20, 21, 22, 23]. These can be readily used to remove incorrectly cast shadows from edited images to fool physical integrity detectors leveraging shadows to assess image authenticity [17, 24].

In this paper, we propose a method to detect whether an automatic shadow removal technique has been used to edit an image. If shadow modification is detected, we also propose a way to partly recover the location of the missing shadow. In doing so, we can help shadow-based image forensics detectors. The proposed solution is based on the use of a specific class of convolutional neural network (CNN) known as conditional generative adversarial network (cGAN). The architecture is trained on purpose for the problem under analysis on a dataset of images whose shadows have been removed with a very accurate yet easy-to-use state-of-the-art technique [23].

The rest of the paper is structured as follows. Section 2 provides the reader with some background on shadow removal algorithms, and provides the formal problem definition. Section 3 is devoted to the explanation of the proposed methodology for shadow removal detection and localization. Section 4 contains all the details of the performed experimental campaign. Finally, Section 5 ends the paper providing some conclusive remarks.

2. BACKGROUND AND PROBLEM STATEMENT

In this section we introduce the reader to state-of-the-art techniques for automatic shadow removal. Then, we provide the formal definition of the shadow removal detection and localization problem.

2.1. Shadow Removal

The problem of shadow removal involves inconspicuously relighting the shadow pixels while leaving the non shadow pixels unchanged. Over the years, many methods have been proposed to address this problem [19, 20, 21, 22, 23]. These methods can be classified as automatic [22, 21, 20] or user-aided [19, 23], and the criterion for this classification relies solely on how shadows are detected before removal. User aided methods rely on input from humans to detect shadows and then they proceed to remove shadows automatically. Automatic shadow removal methods aim to directly go from an input image to its shadow free counterpart. Both automatic and user aided methods have their potential downsides. For automatic methods, errors in shadow detection could severely hamper the effectiveness of shadow removal, while user aided methods could prove to be tedious.

In this paper we choose to use the shadow-removal technique proposed in [23]. This choice is driven by the following considerations. Despite being a user aided method, it requires only two rough strokes from a user as input. This makes it very easy to use for non-expert image manipulators and the visual results are very pleasant. We use the authors own implementation [23] of the shadow removal method.

2.2. Problem Formulation

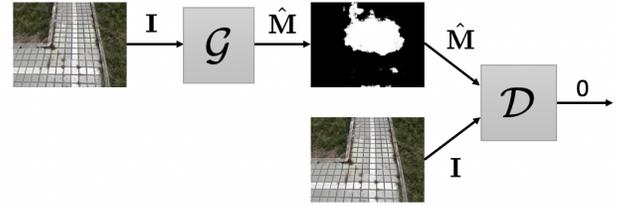
Let us define a natural image under analysis as \mathbf{I} . A pixel with coordinate (x, y) is denoted as $\mathbf{I}(x, y)$. Let us also define a shadow forgery mask \mathbf{M} , being a matrix the same size of the image, whose entries indicate which pixels are affected by the shadow removal algorithm. In other words the sample with coordinate (x, y) in \mathbf{M} is defined as

$$\mathbf{M}(x, y) = \begin{cases} 1, & \text{if a shadow has been deleted in } \mathbf{I}(x, y). \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

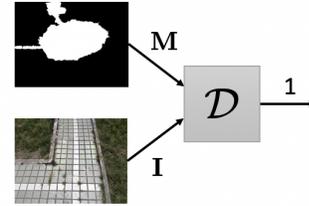
The goal of our method is twofold. First, to detect whether any shadow has been removed in image \mathbf{I} . Second, if a shadow has been removed, to estimate the locations of pixels that originally contained shadow traces. In order to solve both problems, we compute $\hat{\mathbf{M}}$ being an estimate of \mathbf{M} . If $\hat{\mathbf{M}} \approx \mathbf{0}$, we conclude that no shadows have been removed and the image is authentic. Conversely, if $\hat{\mathbf{M}} \neq \mathbf{0}$, we conclude that the image has been edited, and the original shadow was located in pixel at locations $\{(x, y) : \hat{\mathbf{M}}(x, y) = 1\}$.

3. PROPOSED METHOD

Our proposed method for shadow removal detection and localization is based on the following pipeline: (i) the image \mathbf{I} under analysis is adapted to fit a given resolution; (ii) a convolutional neural network (CNN) trained to generate a heatmap that indicates the likelihood of shadow removal traces for each patch pixel is used; (iii) the heatmap is thresholded to estimate $\hat{\mathbf{M}}$ and to make a decision concerning shadow detection and localization. In the following, we provide a detailed description about each step of the proposed procedure.



(a) Generator Training



(b) Discriminator training

Fig. 1. cGAN overall architecture. The generator (a) is trained to fool the discriminator. The discriminator (b) is trained to detect ground truth and estimated masks.

3.1. Image Size Adaptation

One of the problems that arises when processing high resolution images with CNNs, is that image resolution rarely matches the CNN input size. Therefore, a common strategy consists in to split the full resolution image into smaller patches, analyze each patch separately, and finally aggregate the results. This rationale has been already successfully used by other recently proposed forensic detectors [25, 26].

One of the issues in doing this, is the trade-off between accuracy and computational complexity. To obtain a fine-grained solution, patches must be extracted with a large overlap. This increases the number of patches to cover the whole image area, thus a higher computational time.

In order to compromise and reduce the required computational power, we propose to use a slightly different solution. As a matter of fact we train our CNN in order to work on images that have been downsized by a factor of almost 2 with respect to their original resolution. This has two major positive effects. First, the CNN becomes naturally resistant to resize laundering. Second, when a high resolution image is under analysis, the analyst can extract larger (thus less) patches, resize them for CNN analysis, and finally upsample the results back to the original image size.

3.2. CNN Architecture

To estimate $\hat{\mathbf{M}}$ from an image \mathbf{I} resized to the correct CNN resolution, we learn a mapping function that goes from the resized \mathbf{I} to $\hat{\mathbf{M}}$ using a cGAN. This cGAN is based on pix2pix [27]. The architecture of the cGAN is composed by two different CNNs namely, the Generator G and the Discriminator D , coupled together as shown in Figure 1.

Generator G is a U-net [28] containing more than 10 convolutional layers with skipped connections. This network turns the input image into the estimated mask as defined $\hat{\mathbf{M}} = G(\mathbf{I})$. Discriminator D is a simpler and shallower network composed by a

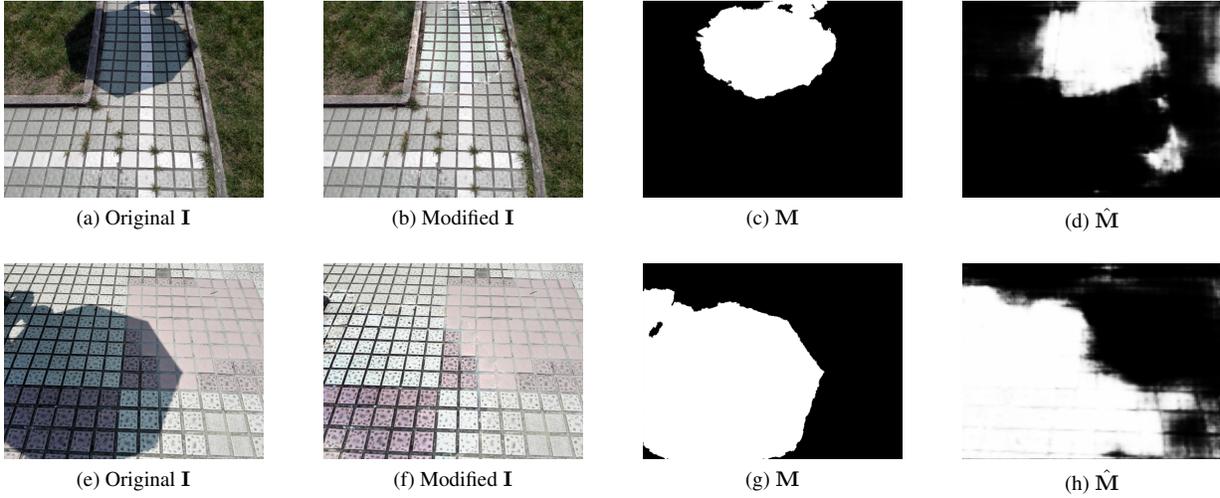


Fig. 2. Example of shadow removal from image \mathbf{I} , ground truth shadow mask \mathbf{M} and estimated shadow mask $\hat{\mathbf{M}}$.

series of convolutional, pooling and fully connected layers. This network acts as a binary classifier on shadow masks, trying to distinguish whether they are a ground truth mask, or a mask estimated by the generator G . The discriminator is trained to output either 1 or 0 depending on the nature of the inputs, i.e. $\mathcal{D}(\mathbf{I}, \mathbf{M}) = 1$ and $\mathcal{D}(\mathbf{I}, \hat{\mathbf{M}}) = 0$. Both the generator and discriminator are coupled together using a loss function $\mathcal{L}_{CGAN}(G, D)$ (please refer to [27] for details on $\mathcal{L}_{CGAN}(G, D)$). In addition to $\mathcal{L}_{CGAN}(G, D)$ the generator is also trained to reduce a reconstruction loss between the predicted mask $\hat{\mathbf{M}}$ and the true mask \mathbf{M} , denoted as $\mathcal{L}_R(\mathbf{M}, \hat{\mathbf{M}})$. The loss function of cGAN denoted by \mathcal{L} , is defined as

$$\mathcal{L} = \mathcal{L}_{cGAN} + \lambda \cdot \mathcal{L}_R \quad (2)$$

By coupling the two loss functions as shown in Eq.2, we force the generator to not only generate $\hat{\mathbf{M}}$ that is close to \mathbf{M} but also fool the discriminator in the process. This additional constraint results in a G that better maps \mathbf{I} to \mathbf{M} as opposed to just training G to reduce \mathcal{L}_R without the discriminator D .

We chose \mathcal{L}_R to be the binary cross-entropy (BCE) between $\hat{\mathbf{M}} = G(\mathbf{I})$ and \mathbf{M} . This is different with respect to the classic pix2pix network, which makes use of $l1$ -norm. However, as our goal is to estimate a binary mask, cross-entropy seems like a more natural choice (as we verify in the results presentation).

Once the network has been trained, the discriminator is not considered anymore, and the generator is used to turn new images under analysis \mathbf{I} into estimated shadow masks as $\hat{\mathbf{M}} = \mathcal{G}(\mathbf{I})$.

3.3. Shadow Removal Detection and Localization

Depending on the image size adaptation strategy, one might need to splice together (with possible overlaps) all estimated masks $\hat{\mathbf{M}}$ coherently with the image patch extraction policy. If the image under analysis already fit the network input size, there is no need to perform additional steps and the estimated mask $\hat{\mathbf{M}}$ can be directly used. However, as we did not constrain the network output to be boolean, the mask $\hat{\mathbf{M}}$ is estimated in a real domain. To construct a binary mask, we need to threshold $\hat{\mathbf{M}}$ using a value Γ , which can be learned upon a validation set of images. An example of original

image, manipulated image, ground truth mask, and estimated mask $\hat{\mathbf{M}}$ is reported in Figure 2.

4. EXPERIMENTS AND RESULTS

In this section we describe our experimental evaluation. We first describe the image dataset. We then report details about the use CNN training policy. Finally, we present the achieved numerical results.

4.1. Image Datasets

To correctly evaluate the proposed method, we constructed a dataset containing both natural non-manipulated images and image whose shadows have been removed. We started with the publicly available Image Shadow Triplets Dataset (ISTD) proposed in [22]. ISTD consists of 1870 color image pairs from 135 different natural scenes. Each image pair is defined as $\{\mathbf{I}^S, \mathbf{I}^{SF}\}$, where \mathbf{I}^S denotes an image with a shadow, and \mathbf{I}^{SF} denotes a shadow-free image depicting the same scene of \mathbf{I}^S . Each image has a resolution of 640×480 pixels. A couple of examples are shown in Figure 2.

For each pair, we used the selected shadow removal [23] for each image \mathbf{I}^S to obtain the manipulated shadow-free image denoted by $\hat{\mathbf{I}}^{SF}$. The binary forgery mask \mathbf{M}^S is obtained by checking which pixels have been actually modified

$$\mathbf{M}^S(x, y) = \begin{cases} 0, & \text{if } \mathbf{I}^S(x, y) = \hat{\mathbf{I}}^{SF}(x, y), \\ 1, & \text{if } \mathbf{I}^S(x, y) \neq \hat{\mathbf{I}}^{SF}(x, y). \end{cases} \quad (3)$$

In our experimental scenario $\hat{\mathbf{I}}^{SF}$ is a forged image whose shadow has been removed, and its binary forgery mask is \mathbf{M}^S . The effectiveness of any forensic method is not only determined by how well it works on forged images but also on how effective it is on authentic images. In our scenario each non-manipulated image \mathbf{I}^{SF} serves as an authentic image with a forgery mask $\mathbf{M}^{SF} = \mathbf{0}$.

In summary, our image dataset \mathcal{D} consists of 1870 pairs of forged images with masks denoted by $\{\hat{\mathbf{I}}^{SF}, \mathbf{M}^S\}$, along with 1870 pairs of authentic images with masks denoted by $\{\mathbf{I}^{SF}, \mathbf{M}^{SF}\}$.

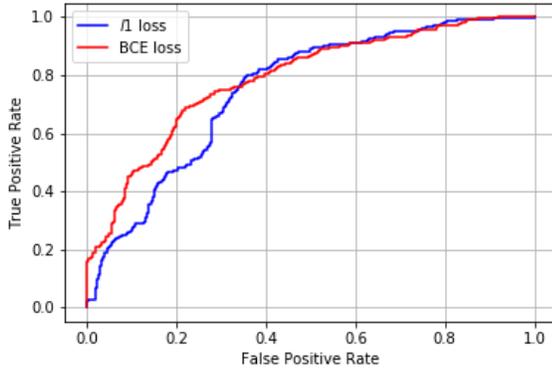


Fig. 3. ROC showing shadow removal detection performance. BCE loss is the proposed one, whereas $l1$ loss is the standard pix2pix one.

4.2. Training Strategy

Prior to training, the forged and authentic images along with their masks are resized to a resolution of 256×256 pixel to match the CNN input. The dataset \mathcal{D} is then split into training \mathcal{D}_{train} , validation \mathcal{D}_{val} and test \mathcal{D}_{test} . Dataset \mathcal{D}_{train} consists of 1130 forged images and the corresponding 1130 authentic images. Similarly \mathcal{D}_{val} consists of 200 forged and 200 authentic images. Finally, \mathcal{D}_{test} consists of the remaining 540 forged and 540 authentic images. While the entire dataset \mathcal{D} contains images from 135 different scenes, the images for training and validation come from 90 of 135 different scenes, whereas images used for testing come from the remaining 45 scenes. In doing so, we ensure that the method is not merely learning how to distinguish between scenes.

In order to train the used CNN minimizing the proposed loss function, we used Adam optimizer [29] for both the discriminator and the generator. We set $\lambda = 10$, and trained the model for 200 epochs, selecting for test the model minimizing loss on \mathcal{D}_{val} .

4.3. Numerical Analysis

To show some examples of masks obtained by running our algorithm, Figure 2 reports two sets of images composed by the original picture with shadow, the edited picture whose shadow has been removed, the ground truth mask \mathbf{M} , and the estimated mask $\hat{\mathbf{M}}$. From this visual example it is possible to notice that our proposed method is able to correctly pinpoint shadow removal even on images that have been manipulated in a visually plausible manner. Moreover, the method provides information about the original shadow location, which can be helpful to perform some shadow-based forensic analysis.

To numerically evaluate our proposed method in terms of shadow-removal detection, we tested the constructed dataset. Specifically, after estimating each mask $\hat{\mathbf{M}}$, we computed its average value, and compared it against a threshold Γ to detect removed shadows. The idea is that the average value of $\hat{\mathbf{M}}$ should be zero (or approximately so) for non-manipulated pictures. Figure 3 reports the receiver operating characteristic (ROC) curve obtained by changing the value of the used threshold Γ . Figure 3 shows two curves: one obtained using the proposed binary cross-entropy (BCE) loss; one obtained using the conventional pix2pix $l1$ -norm loss. It is possible to see that the proposed loss modification makes the algorithm more

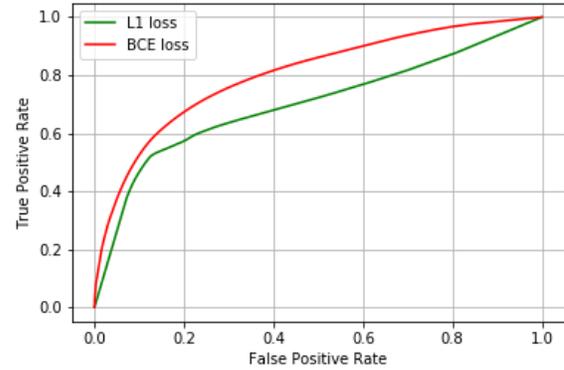


Fig. 4. ROC showing shadow removal localization performance. BCE loss is the proposed one, whereas $l1$ loss is the standard pix2pix one.

precise, as the achieved area under the curve (AUC) increase to 0.788 (using BCE) from 0.751 (using conventional $l1$ -norm loss).

If an image has been detected as manipulated, we also want to estimate the manipulated pixels. To this purpose, we compared each thresholded estimated mask $\hat{\mathbf{M}}$ to the ground truth mask \mathbf{M} in a pixel-wise fashion. By changing the threshold used to binarize $\hat{\mathbf{M}}$ we computed two ROC curves. In terms of localization, it is possible to note that AUC reaches 0.803 using the proposed BCE loss, whereas it only reached 0.701 using the conventional one.

As final experiment, we compared our method against a set of general-purpose image forensic techniques. Specifically, we considered the toolbox presented in [30], which contains a set of more than 10 algorithms. Additionally, we also tested the technique proposed in [25], which is considered among the best splicing detection and localization tools to be used when no a priori information about the kind of manipulation are available. None one of these techniques was able to provide localization AUC > 0.6 on the proposed dataset. However, this behavior is somehow expected as none of these methods are specifically tailored to this type of manipulation and probably needs some more tuning.

5. CONCLUSIONS

In this paper we proposed a shadow removal detector for forensic image analysis. Given an input image, the proposed CNN outputs a mask showing: (i) whether the image was manipulated by means of a shadow removal technique; (ii) the pixel locations where a shadow was possibly present before removal. This detector can be used as additional tool in an analyst's asset in order to counter anti-forensic attacks tailored to shadow-based forensics detectors.

The proposed solution has been tested against a shadow removal method that has good performance despite been very easy to use by non-experts. As a byproduct of our investigation, we noticed that many digital integrity detectors that appear to be extremely accurate in many situations, did not achieve the same performance in our analysis. It is possible that manipulation traces left by more unconventional image processing methods (as shadow removal) are different in nature by more classical and well-studied image editing operations.

6. REFERENCES

- [1] A. Rocha, W. Scheirer, T. Boulton, and S. Goldenstein, "Vision of the unseen: Current trends and challenges in digital image and video forensics," *ACM Computing Surveys*, vol. 43, no. 4, pp. 26:1–26:42, Oct. 2011.
- [2] A. Piva, "An overview on image forensics," *ISRN Signal Processing*, vol. 2013, p. 22, Jan. 2013.
- [3] M. C. Stamm, Min Wu, and K. J. R. Liu, "Information Forensics: An Overview of the First Decade," *IEEE Access*, vol. 1, pp. 167–200, May 2013.
- [4] M. Kirchner and T. Gloe, "Forensic camera model identification," in *Handbook of Digital Forensics of Multimedia Data and Devices*, A. T. S. Ho and S. Li, Eds. Wiley-Blackwell, 2015, pp. 329–374, Hoboken, NJ.
- [5] A. Tuama, F. Comby, and M. Chaumont, "Camera model identification with the use of deep convolutional neural networks," *Proceedings of the IEEE International Workshop on Information Forensics and Security*, pp. 1–6, Dec. 2016, Abu Dhabi, United Arab Emirates.
- [6] L. Verdoliva, D. Cozzolino, and G. Poggi, "A feature-based approach for image tampering detection and localization," *Proceedings of the IEEE International Workshop on Information Forensics and Security*, pp. 149–154, Dec. 2014, Atlanta, GA.
- [7] L. Gaborini *et al.*, "Multi-clue image tampering localization," *IEEE International Workshop on Information Forensics and Security*, pp. 125–130, Dec. 2015, Atlanta, GA.
- [8] A. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of resampling," *IEEE Transactions on Signal Processing*, vol. 53, pp. 758–767, Jan. 2005.
- [9] M. Kirchner, "Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue," *Proceedings of the ACM Workshop on Multimedia and Security*, Sep. 2008, Oxford, UK.
- [10] D. Vázquez-Padín and F. Perez-Gonzalez, "Prefilter design for forensic resampling estimation," *Proceedings of the IEEE International Workshop on Information Forensics and Security*, pp. 1–6, Nov. 2011, Iguacu Falls, Brazil.
- [11] 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, Jun. 2012.
- [12] M. Barni *et al.*, "Aligned and non-aligned double JPEG detection using convolutional neural networks," *Journal of Visual Communication and Image Representation*, vol. 49, pp. 153–163, Nov. 2017.
- [13] E. Wengrowski, Z. H. Sun, and A. Hoogs, "Reflection correspondence for exposing photograph manipulation," *Proceedings of the IEEE International Conference on Image Processing*, pp. 4317–4321, Sep. 2017, Beijing, China.
- [14] M. K. Johnson and H. Farid, "Exposing digital forgeries by detecting inconsistencies in lighting," *Proceedings of the ACM Workshop on Multimedia and Security*, pp. 1–10, 2005, New York, NY.
- [15] T. J. d. Carvalho *et al.*, "Exposing digital image forgeries by illumination color classification," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 7, pp. 1182–1194, Jul. 2013.
- [16] Q. Liu, X. Cao, C. Deng, and X. Guo, "Identifying image composites through shadow matte consistency," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1111–1122, Sep. 2011.
- [17] E. Kee, J. F. O'Brien, and H. Farid, "Exposing photo manipulation with inconsistent shadows," *ACM Transactions on Graphics*, vol. 32, no. 3, pp. 28:1–28:12, July 2013.
- [18] V. Tuba, R. Jovanovic, and M. Tuba, "Digital image forgery detection based on shadow hsv inconsistency," *Proceedings of the IEEE International Symposium on Digital Forensic and Security*, pp. 1–6, Apr. 2017, Tirgu Mures, Romania.
- [19] H. Gong, D. Cosker, C. Li, and M. Brown, "User-aided single image shadow removal," *Proceedings of the IEEE International Conference on Multimedia and Expo*, pp. 1–6, Jul. 2013, San Jose, CA.
- [20] R. Guo *et al.*, "Paired regions for shadow detection and removal," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 12, pp. 2956–2967, Dec. 2013.
- [21] Q. Yang, K. Tan, and N. Ahuja, "Shadow removal using bilateral filtering," *IEEE Transactions on Image Processing*, vol. 21, no. 10, pp. 4361–4368, Oct. 2012.
- [22] J. Wang, X. Li, and J. Yang, "Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1788–1797, Jul. 2018, Salt Lake City, UT.
- [23] H. Gong and D. Cosker, "Interactive shadow removal and ground truth for variable scene categories," *Proceedings of the British Machine Vision Conference*, Sep. 2014, Nottingham, UK.
- [24] Y. Ke, F. Qin, W. Min, and G. Zhang, "Exposing image forgery by detecting consistency of shadow," *The Scientific World Journal*, vol. 2014, p. 9, Mar. 2014.
- [25] D. Cozzolino, G. Poggi, and L. Verdoliva, "Splicebuster: A new blind image splicing detector," *Proceedings of the IEEE International Workshop on Information Forensics and Security*, pp. 1–6, Nov. 2015.
- [26] L. Bondi *et al.*, "Tampering detection and localization through clustering of camera-based cnn features," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1855–1864, Jul. 2017, Honolulu, HI.
- [27] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5967–5976, Jul. 2017, Honolulu, HI.
- [28] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Oct. 2015, Munich, Germany.
- [29] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *Proceedings of the International Conference on Learning Representations*, May 2015, San Diego, CA.
- [30] M. Zampoglou, S. Papadopoulos, and Y. Kompatsiaris, "Large-scale evaluation of splicing localization algorithms for web images," *Multimedia Tools and Applications*, vol. 76, no. 4, pp. 4801–4834, Feb. 2017.