A Novel Image Forgery Classification and Detection Using Residual Deep Learning

D.Bhargavi[1], K.V.Satyanarayana[2], B.S.Panda[3]
[1]*Student, Computer Science Engineering, Raghu Engineering College, A.P, India*
[2]*Associate Professor, Computer Science Engineering, Raghu Engineering College, A.P, India*
[3]*Professor, Computer Science Engineering, Raghu Engineering College, A.P, India*

ABSTRACT

Acquiring images has been increasingly popular in recent years, owing to the widespread availability of cameras. Images are essential in our daily lives because they contain a wealth of information, and it is often required to enhance images to obtain additional information. A variety of tools are available to improve image quality; nevertheless, they are also frequently used to falsify images, resulting in the spread of misinformation. This increases the severity and frequency of image forgeries, which is now a major source of concern. Numerous traditional techniques have been developed over time to detect image forgeries. In recent years, convolutional neural networks have received much attention, and RESNET + CNN has also influenced the field of image forgery detection. However, most image forgery techniques based on RESNET+CNN that exist in the literature are limited to detecting a specific type of forgery. As a result, a technique capable of efficiently and accurately detecting the presence of unseen forgeries in an image is required. In this paper, we introduce a robust deep learning-based system for identifying image forgeries in the context of double image compression. The difference between an image's original and recompressed versions is used to train our model. The proposed model is lightweight, and its performance demonstrates that it is faster than state-of-the-art approaches. The experiment results are encouraging, with an overall validation accuracy of 98.23%.

Keywords:- *Convolutional neural network, Neural networks, Forgery detection, Image compression, Image processing, TensorFlow, Keras, Python, Resnet.*

1. Introduction:

Due to technological advancements and globalization, electronic equipment is now widely and inexpensively available. As a result, digital cameras have grown in popularity. There are many camera sensors all around us, and we use them to collect a lot of images. Images are required in the form of a soft copy for various documents that must be filed online, and a large number of images are shared on social media every day. The amazing thing about images is that even illiterate people can look at them and extract information from them. As a result, images are an integral component of the digital world, and they play an essential role in storing and distributing data. There are numerous tools accessible for quickly editing the images. These tools were created with the intention of enhancing and improving the images. However, rather than enhancing the image, some people exploit their capabilities to falsify images and propagate falsehoods. This is a significant threat, as the damage caused by faked images is not only severe, but also frequently irreversible. There are two basic types of image forgery: image splicing and copy-move, which are discussed below:

Image Splicing: A portion of a donor image is copied into a source image. A sequence of donor images can likewise be used to build the final forged image.

Copy-Move: This scenario contains a single image. Within the image, a portion of the image is copied and pasted. This is frequently used to conceal other objects. The final forged image contains no components from other images.

The primary purpose in both cases of image forgery is to spread misinformation by changing the original content in an image with something else Earlier images were an extremely credible source for the information exchange, however, due to image forgery, they are used to spread misinformation. This is affecting the trust of the public in images, as the forging of images may or may not be visible or recognizable to the naked eye. As a result, it is essential to detect image forgeries to prevent the spread of misinformation as well as to restore public trust in images. This can be done by exploring the various artifacts left behind when an image forgery is performed, and they can be identified using various image processing techniques. Researchers have proposed a variety of methods for detecting the presence of image forgeries Conventional image forgery detection techniques detect forgeries by concentrating on the multiple artifacts present in a forged image, such as changes in illumination, contrast, compression, sensor noise, and shadow. CNN's have gained popularity in recent years for various computer vision tasks, including image object recognition, semantic

segmentation, and image classification. Two major features contribute to CNN's success in computer vision. Firstly, CNN takes advantage of the significant correlation between adjacent pixels. As a result, CNN prefers locally grouped connections over one-to-one connections between all pixel. Second, each output feature map is produced through a convolution operation by sharing weights. Moreover, compared to the traditional method that depends on engineered features to detect specific forgery, CNN uses learned features from training images, and it can generalize itself to detect unseen forgery. These advantages of CNN make it a promising tool for detecting the presence of forgery in an image. It is possible to train a CNN-based model to learn the many artifacts found in a forged image. Thus, we propose a very light CNN-based network, with the primary goal of learning the artifacts that occur in a tampered image as a result of differences in the features of the original image and the tampered region. The major contribution of the proposed technique are as follows: A lightweight CNN-based architecture is designed to detect image forgery efficiently. The proposed technique explores numerous artifacts left behind in the image tampering process, and it takes advantage of differences in image sources through image recompression. While most existing algorithms are designed to detect only one type of forgery, our technique can detect both image splicing and copy-move forgeries and has achieved high accuracy in image forgery detection. Compared to existing techniques, the proposed technique is fast and can detect the presence of image forgery in significantly less time. Its accuracy and speed make it suitable for real-world application, as it can function well even on slower devices. The rest of the paper is organized as follows. Section2 provides a literature review of image forgery detection methodologies. Section3 introduces the proposed framework for detecting the presence of forgeries in an image. Section4 contains a discussion of the experimentation and the results achieved. Finally, in Section5, we summarize the conclusions.

2. Literature Review:

Various approaches have been proposed in the literature to deal with image forgery. The majority of traditional techniques are based on particular artifacts left by image forgery, whereas recently techniques based on CNNs and deep learning were introduced, which are mentioned below. First, we will mention the various traditional techniques and then move on to deep learning-based techniques.

In, the authors' proposed error level analysis (ELA) for the detection of forgery in an image. In based on the lighting conditions of objects, forgery in an image is detected. It tries to find the forgery based on the difference in the lighting direction of the forged part and the genuine part of an image. In, various traditional image forgery detection techniques have been evaluated. In Habibi et al., use the contourlet trans- form to retrieve the edge pixels for forgery detection. In Dua et al., presented a JPEG compression-based method. The discrete DCT coefficients are assessed independently for each block of an image partitioned into non-overlapping blocks of size 8 8 pixels. The statistical features of AC components of block DCT coefficients alter when a JPEG compressed image tampers. The SVM is used to classify authentic and forged images using the retrieved feature vector. Ehret et al. in introduced a technique that relies on SIFT, which provides sparse keypoints with scale, rotation, illumination invariant descriptors for forgery detection. A method for fingerprint faking detection utilizing deep Boltzmann machines (DBM) for image analysis of high-level characteristics is proposed in Balsa et al. in compared the DCT, Walsh–Hadamard transform (WHT) and discrete Fourier transform (DFT) for analog image transmission, changing compression and comparing quality. These can be used for image forgery detection by exploring the image from different domains. Thanh et al. proposed a hybrid approach for image splicing in [high they try to retrieve the original images that were utilized to construct the spliced image if a given image is proven to be the spliced image Myung-Joon introduced CAT-Net, to acquire forensic aspects of compression artifact on DCT and RGB domains simultaneously. Their primary network is HR-Net (high resolution). They used the technique proposed which tells us that how we can use the DCT coefficient to train a CNN, as directly giving DCT coefficients to CNN will not train it efficiently. Ashraful et al. in proposed DOA-GAN, to detect and localize copy-move forgeries in an image, authors used a GAN with dual attention. The first-order attention in the generator is designed to collect copy-move location information, while the second-order attention for patch co-occurrence exploits more discriminative properties. The affinity matrix is utilized to extract both attention maps, which are then used to combine location-aware and co-occurrence features for the network's ultimate detection and localization branches.

Yue et al. in proposed Buster Net for copy-move image forgery detection. It has a two-branch architecture with a fusion module in the middle. Both branches use visual artifacts to locate potential manipulation locations and visual similarities to locate copy- move regions. Yue et al. in employed a CNN to extract block-like characteristics from an image, compute self-correlations between various blocks, locate matching points using a point-wise feature extractor, and reconstruct a forgery mask using a deconvolutional network. Yue et al. in designed ManTra-Net that is s a fully convolutional network that can handle any size image and a variety of forgery types, including copy-move, enhancement, splicing, removal, and even unknown forgery forms. Liu et al. in proposed PSCC-Net, which analyses the image in a two-path methodology: a top-down route that retrieves global and local features and a bottom-up route that senses if the image is tampered and predicts its masks at four levels, each mask being constrained on the preceding one. In Yang et al., proposed a technique based on two concatenated CNNs: the coarse CNN and the refined CNN, which extracts the differences between the image itself and splicing regions from patch descriptors of different scales. They enhanced their work in and proposed a patch-based coarse-to-refined network (C2RNet). The coarse network is based on VVG16, and the refined network is based on VVG19. In Xiuli et al., proposed a ringed residual U-Net to detect the

splicing type image forgery in the images. Younis et al. in [utilized the reliability fusion map for the detection of the forgery. By utilizing the CNNs, Younis et al. in classify an image as the original one, or it contains copy-move image forgery. In train four models at the same time: a generative annotation model GA, a generative retouching model GR, and two discriminators DA and DR that checks the output of GA and GR. Mayer et al. in system that maps sets of image regions to a value that indicates if they include the same or different forensic traces. In Minyoung et al., designed an algorithm that leverages the automatically recorded image EXIF metadata for training a model to identify whether an image has self-consistency or if its content may have been generated from a single image. In Rongyu et al., proposed a UNet that consists of a dense convolutional and deconvolutional networks. The first is a down-sampling method for retrieving features, while the second is an up-sampling approach for recovering feature map size. In Lui et al., introduced the CNN segmentation-based approach to find manipulated regions in digital photos. First, a uniform CNN architecture is built to deal with various scales' color input sliding windows. Then, using sampling training regions, they meticulously build CNN training processes. In an unfixed encoder and a fixed encoder are used to build a Dual-encoder U- Net (D-Unet). The unfixed encoder learns the image fingerprints that distinguish between genuine and tampered regions on its own. In contrast, the fixed encoder offers direction data to facilitate the network's learning and detection. In [Francesco et al., tested the efficiency of several image forgery detectors over image-to-image translation, including both ideal settings and even in the existence of compression, which is commonly performed when uploading to social media sites. Kadam et al. in Proposed a method based on multiple image splicing using Mobile Net V1.Jaiswal et al. in proposed a framework in which images are fed into a CNN and then processed through several layers to extract features, which are then utilized as a training vector for the detection model. For feature extraction, they employed a pre-trained deep learning resnet-50. Hao et al. in proposed using an attention method to analyse and refine feature maps for the detection task. The learned attention maps emphasize informative areas to enhance binary classification and illustrate the altered regions. In Nguyen et al., developed a CNN that employs a multi-task learning strategy to detect altered images and videos while also locating the forged areas. The information received from one work is shared with the second task, improving both activities' performance. To boost the network's generability, a semi-supervised learning strategy is adopted. An encoder and a Y-shaped decoder are included in the network. Li et al. introduced a deepfake detection method in The DeepFake techniques can only create fixed-size images of the face, which must be affinely warped to match the source's face arrangement. Due to the resolution disparity between the warped face area and the surrounding context, this warping produces different artifacts. As a result, DeepFake Videos can be identified using these artifacts. Komodakis et al. in suggested a method for learning image features by training CNNs to recognize the two-dimensional rotation that is applied to the picture that it receives as input. The method proposed in is composed of three parts: single image super-resolution, semantic segmentation super- resolution, and featps to fully examine the visual-semantic relationships and enhance the level of produced sentences. For more details about image forgery and media, forensics readers may refer to the state-of-the-art techniques available for detecting the presence of tampering in the images generally take a very long time to process the images. Most of them can detect either image splicing forgery or copy-move type of forgery, not both. Another major issue with them is that they detect the forgery with low accuracy. Hence, there is a need for a better framework that is fast and more accurate. To address this, we presented a novel image recompression-based system. Apart from achieving better image forgery detection accuracy, our proposed framework has also achieved faster response time. This makes it suitable for real-life applications, as it is more accurate and can be utilized even by slower machines. The proposed framework is detailed in the next section.

3. Proposed Methodology:

We can propose resnet+cnn combination architecture for the problem statement.

3.1 Dataset:

Before going into the dataset overview, the terminology used will be made clear

- Fake image: An image that has been manipulated/doctored using the two most common manipulation operations namely: copy/pasting and image splicing.
- Pristine image: An image that has not been manipulated except for the resizing needed to bring all images to a standard size as per competition rules.
- Image splicing: The splicing operations can combine images of people, adding doors to buildings, adding trees and cars to parking lots etc. The spliced images can also contain resulting parts from copy/pasting operations. The image receiving a spliced part is called a "host" image. The parts being spliced together with the host image are referred to as "aliens".

The entire dataset for both the first and second phase can be found here. For this project, we will be using only the train set. It contains 2 directories — one containing fake images and their corresponding masks and the other containing pristine images. Mask of a fake image is a black and white (not grayscale) image describing the spliced area of the fake image. The black pixels in the mask represent the area where manipulation was performed in the source image to get the forged image, specifically it represents the spliced region. The dataset consists of 1050 pristine and 450 fake images. Color images are usually 3 channel images one channel for each red, green and blue colors, however sometimes the

fourth channel for yellow may be present. Images in our dataset are a mix of 1, 3 and 4 channel images. After looking at a couple of 1 channel images i.e. grayscale images, it was evident that these images

1. were very few in number
2. were streams of black or blue color

The challenge setters added these images on purpose as they wanted solutions robust to such noise. Although some of the blue images can be images of a clear sky. Hence some of them were included while others discarded as noise. Coming to four channel images — they too didn't have any useful information. They were simply grids of pixels filled with 0 values. Thus, our pristine dataset after cleaning contained about 1025 RGB images. Fake images are a mix of 3 and 4 channel images, however, none of them are noisy. Corresponding masks are a mix of 1, 3 and 4 channel images. The feature extraction we will be using requires information from only one channel of the masks. Thus, our fake image corpus has 450 fakes. Next up we did a train-test split to keep 20% of 1475 images for final testing.

3.2 Feature extraction on the train set:

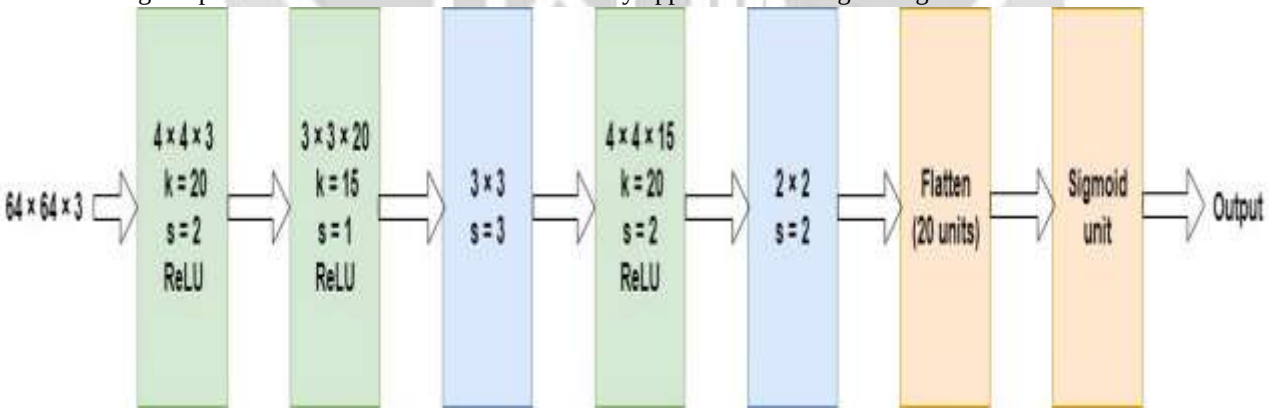
The dataset in its present state is not apt for training a model. It must be transformed into a state which is well suited for the task at hand i.e. detection of anomalies at the pixel level introduced due to forging operations. Taking ideas from here we designed the following methodology to create relevant images from the given data.

For every fake image, we have a corresponding mask. We use that mask to sample the fake image along the boundary of the spliced region in such a way so as to ensure at least a 25% contribution from both forged part and unforged part of the image. These samples will have the distinguishing boundaries that would be present only in fake images. These boundaries are to be learned by the CNN we design. Since all 3 channels of the mask contain the same information, we need only 1 channel to extract samples.

To make boundaries even more distinct, the grayscale images were converted to binary using thresholding. Implemented in opencv, after denoising using a Gaussian filter. After this operation, sampling was merely a matter of moving a 64×64 window (with a stride of 8) through the fake image and counting 0 valued pixels in the corresponding mask and sampling in case the value lies in a certain interval.

3.3 Custom CNN Architecture:

The first architecture we tried was inspired by the architecture given in the original paper They had input images of size 128×128×3 and hence a large network. Since we have half the spatial size, our network was also smaller. This is the first tried architecture. Residual Network a.k.aResNet50 is a variant of the ResNet model which consists of 48 Convolution layers along with 1 MaxPool and 1 Average Pool layer. It is capable of 3.8 billion Floating-point operations. Out of all other variants of residual network with different capabilities, this one widely used ResNet model and we have shown ResNet50 architecture in detail in Figure 4. Because of this framework, it is possible to train ultra DNN (deep neural networks) i.e. Now, the network can contain thousands of layers and still achieve great performance. The ResNets were initially applied to the image recognition task but as is mentioned in



the paper that the framework can be used for non-computer vision tasks also to achieve better accuracy. Many people argued that simply stacking more layers also gives us better accuracy why was there a need for Residual learning for training ultra- deep neural networks but stacking more layer arises a serious problem of vanishing/exploding gradients, that is why ResNet is used in this paper so that we can assess its effectiveness in deepfake detection problem.

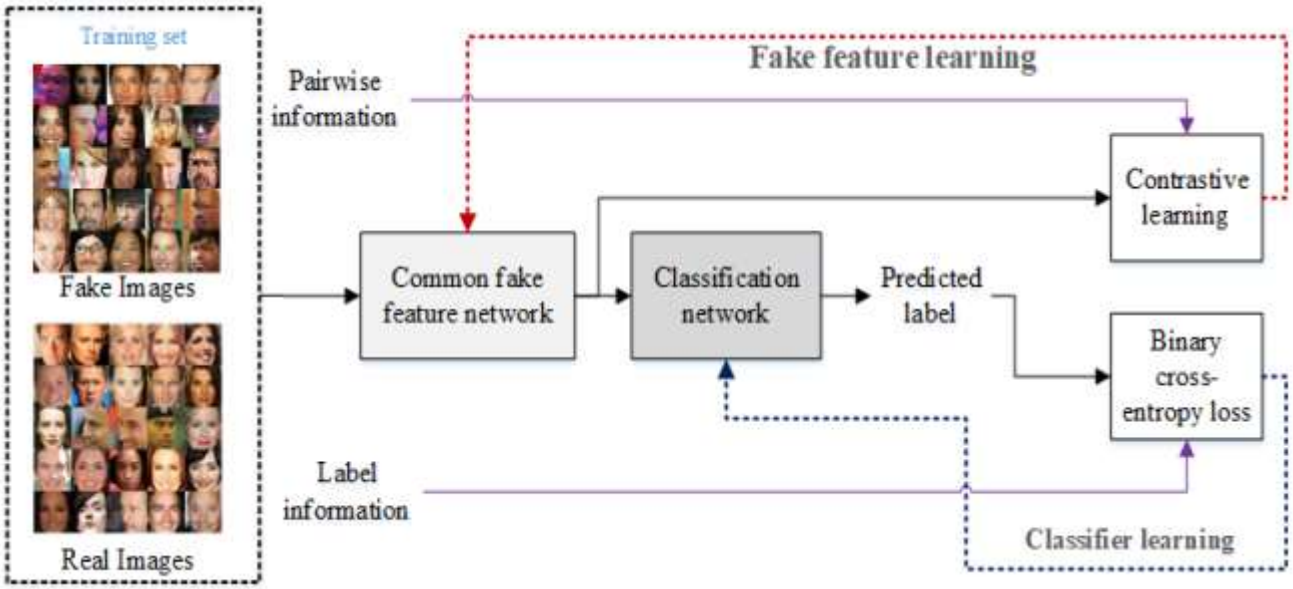
Here layers are convolutional and blue ones are Max pool. This network was trained on 150,000 train samples (for testing purpose) and 25,000 validation samples. The network had 8,536 parameters which were relatively less compared to the train samples, hence avoiding the need for a more aggressive dropout. A dropout rate of 0.2 was applied to the flattened output of 20 units. We used Adam optimizer with a default value of learning rate (0.001) and beta_1, beta_2. After about ___ epochs the results were as follows

Train accuracy: 96.13%, Train loss: 0.4678

Validation accuracy: 98.68%, Validation loss: 0.0521

These numbers are not very impressive given the fact that in 2012 a CNN beat yearlong researched features developed by experts, by a huge margin. However, these numbers are not very bad either given the fact that we used absolutely no knowledge of image forensics to get a best accuracy on unseen data.

Training pipeline:



4. Results:

The residual deep learning classification accuracy on both datasets after the 10-fold cross-validation is presented below:

Dataset	Accuracy
CASIA2	96.82% ± 1.19%
NC2016	84.89% ± 6.06%

Input image:

Output Image:

5. Conclusion & Future scope:

The increased availability of cameras has made photography popular in recent years. Images play a crucial role in our lives and have evolved into an essential means of conveying information since the general public quickly understands them. There are various tools accessible to edit images; these tools are primarily intended to enhance images; however, these technologies are frequently exploited to forge the images to spread misinformation. As a result, image forgery has become a significant problem and a matter of concern. In this paper, we provide a unique image forgery detection system based on neural networks and deep learning, emphasizing the CNN architecture approach. To achieve satisfactory results, the suggested method uses a CNN architecture that incorporates variations in image compression. We use the difference between the original and recompressed images to train the model. The proposed technique can efficiently detect image splicing and copy-move types of image forgeries. The experiments results are highly encouraging, and they show that the overall validation accuracy is 92.23%, with a defined iteration limit.

We plan to extend our technique for image forgery localization in the future. We will also combine the suggested technique with other known image localization techniques to improve their performance in terms of accuracy and reduce their time complexity. We will enhance the proposed technique to handle spoofing as well. The present technique requires image resolution to be a minimum of 128 128, so we will enhance the proposed technique to work well for tiny images. We will also be developing a challenging extensive image forgery database to train deep learning networks for image forgery detection.

6. References:

1. Kwon, M.J.; Yu, I.J.; Nam, S.H.; Lee, H.K. CAT-Net: Compression Artifact Tracing Network for Detection and Localization of Image Splicing. In Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 5–9 January 2021.
2. Mirsky, Y.; Lee, W. The Creation and Detection of Deepfakes: A Survey. ACM Comput. Surv. 2021.
3. Castillo Camacho, I.; Wang, K. A Comprehensive Review of Deep-Learning-Based Methods for Image Forensics. J. Imaging 2021.
4. Habibi, M.; Hassanpour, H. Splicing Image Forgery Detection and Localization Based on Color Edge Inconsistency using Statistical Dispersion Measures. Int. J. Eng. 2021.
5. Xiao, B.; Wei, Y.; Bi, X.; Li, W.; Ma, J. Image splicing forgery detection combining course to refined convolutional neural network and adaptive clustering. Inf. Sci. 2020.
6. Ali, S.S.; Baghel, V.S.; Ganapathi, I.I.; Prakash, S. Robust biometric authentication system with a secure user template. Image Vis. Comput. 2020.
7. Jing, L.; Tian, Y. Self-supervised Visual Feature Learning with Deep Neural Networks: A Survey. IEEE Trans. Pattern Anal. Mach. Intell. 2020.
8. Verdoliva, L. Media Forensics and DeepFakes: An Overview IEEE J. Sel. Top. Signal Process. 2020.

9. Matern, F.; Riess, C.; Stamminger, M. Gradient-Based Illumination Description for Image Forgery Detection. IEEE Trans. Inf. Forensics Secur. 2020.

10. Dua, S.; Singh, J.; Parthasarathy, H. Image forgery detection based on statistical features of block DCT coefficients. Procedia Comput. Sci. 2020.

11. Balsa, J. Comparison of Image Compressions: Analog Transformations, Proceedings 2020.

12. Wu, Y.; Abd Almageed, W.; Natarajan, P. ManTra-Net: Manipulation Tracing Network for Detection and Localization of Image Forgeries with Anomalous Features. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019.

13. Zheng, L.; Zhang, Y.; Thing, V.L. A survey on image tampering and its detection in real-world photos. J. Vis. Commun. Image Represent. 2019.

14. Zheng, L.; Zhang, Y.; Thing, V.L. A survey on image tampering and its detection in real-world photos. J. Vis. Commun. Image Represent. 2019.

15. Meena, K.B.; Tyagi, V. Image Forgery Detection: Survey and Future Directions. In Data, Engineering and Applications: Volume 2; Shukla, R.K., Agrawal, J., Sharma, S., Singh Tomer, G., Eds.; Springer: Singapore, 2019.

16. 10. Rony, J.; Belharbi, S.; Dolz, J.; Ayed, I.B.; McCaffrey, L.; Granger, E. Deep weakly-supervised learning methods for classification and localization in histology images: A survey. arXiv 2019.

17. Zhang, M.; Zhou, Y.; Zhao, J.; Man, Y.; Liu, B.; Yao, R. A survey of semi- and weakly supervised semantic segmentation of images. Artif. Intell. Rev. 2019.

18. Ehret, T. Robust copy-move forgery detection by false alarms control. arXiv 2019.

19. de Souza, G.B.; da Silva Santos, D.F.; Pires, R.G.; Marana, A.N.; Papa, J.P. Deep Features Extraction for Robust Fingerprint Spoofing Attack Detection. J. Artif. Intel. Soft Comput. Res. 2019.

20. Pham, N.T.; Lee, J.W.; Kwon, G.R.; Park, C.S. Hybrid Image-Retrieval Method for Image-Splicing Validation. Symmetry 2019.

21. Lu, Z.; Chen, D.; Xue, D. Survey of weakly supervised semantic segmentation methods. In Proceedings of the 2018 Chinese Control Furthermore, Decision Conference (CCDC), Shenyang, China, 9–11 June 2018.

22. Bunk, J.; Bappy, J.H.; Mohammed, T.M.; Nataraj, L.; Flenner, A.; Manjunath, B.; Chandrasekaran, S.; Roy-Chowdhury, A.K.; Peterson, L. Detection and Localization of Image Forgeries Using Resampling Features and Deep Learning. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017.

23. Christlein, V.; Riess, C.; Jordan, J.; Riess, C.; Angelopoulou, E. An Evaluation of Popular Copy-Move Forgery Detection Approaches. IEEE Trans. Inf. Forensics Secur. 2012.

24. Luo, W.; Huang, J.; Qiu, G. JPEG Error Analysis and Its Applications to Digital Image Forensics. IEEE Trans. Inf. Forensics Secur.2010.