Effective Deep Features for Image Splicing Detection

Conference Paper · November 2021

DOI: 10.1109/ICSET53708.2021.9612569

3 authors:

Ismail T. Ahmed
University of Anbar
29 PUBLICATIONS **42** CITATIONS
SEE PROFILE

Baraa Hammad
University of Anbar
24 PUBLICATIONS **62** CITATIONS
SEE PROFILE

Norziana Jamil
Universiti Tenaga Nasional (UNITEN)
103 PUBLICATIONS **458** CITATIONS
SEE PROFILE

Effective Deep Features for Image Splicing Detection

Ismail Taha Ahmed
College of Computer Sciences and Information Technology
University of Anbar
Anbar, Iraq
ismail.taha@uoanbar.edu.iq

Baraa Tareq Hammad
College of Computer Sciences and Information Technology
University of Anbar
Anbar, Iraq
baraa.tareq@uoanbar.edu.iq

Norziana Jamil
College of Computing and Informatics
Universiti Tenaga Nasional
Malaysia
Norziana@uniten.edu.my

2021 IEEE 11th International Conference on System Engineering and Technology (ICSET) | 978-1-6654-3766-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICSET53708.2021.9612569

Abstract—In the last few years, Digital image forgery (DIF) detection has become a prominent subject. Image splicing is a frequent approach for making digital image forgeries. Image splicing creates forged images that are hard to detect immediately. The detection accuracy of most existing image splicing detection algorithms is low, thus there is room for improvement. Therefore, this research provides an image splicing detection (ISD) method based on deep learning. The proposed image splicing detection has three stages: (1) RGB image conversion and image size fitting are examples of image pre-processing. (2) Using the pre-trained CNN AlexNet model, we extract the final discriminative feature for a preprocessed image. (3) Finally, the generated feature representation is used to train a Canonical Correlation Analysis (CCA) classifier for binary classification (authentic/forged). The accuracy of the proposed approach using a pre-trained AlexNet model based deep features with CCA classifier is equal to 98.79 % when evaluated on the CASIA v1.0 splicing image forgery database. In comparison, the proposed surpassed existing methods. In the future, the proposed could be applied to other types of image forgery, such as image retouching.

Keywords— Digital image forgery (DIF), Image Splicing, Deep Features, AlexNet model, Canonical Correlation Analysis (CCA) classifier.

I. INTRODUCTION

Images are now used in a variety of applications, including medical, military, and law enforcement. The majority of the images have been manipulated to increase their quality. However, on occasion, the manipulative individual will strive to create a fake image from the original image and upload it to websites without leaving any trace. As a result, DIF detection has been a hot topic in recent years.

Among the most prevalent methods for generating digital image forgeries is image splicing. Image splicing is the process of joining two or more fragments of various photos to make a new one with no post-processing [1]. Image splicing is a challenging task due to the absence of a replicated region in the case of image splicing. As a result, ISD approaches that are based on the evidence left after the manipulation operation are required. The original and spliced image forgery are shown in Fig. 1. As shown in the figure, Image splicing is a combination of two or more images.

(a) Original image 1 (b) Original image 2 (c) Spliced image

Fig. 1. (a) First Original Image, (b) Second Original Image, (c) Splicing Image [1].

Traditional features approaches and deep learning features approaches are used in the majority of current Splicing image detection (ISD) methods. Traditional tamper detection approaches (both spatial and transform based methods) rely on hand-crafted features.

Alahmadi et al. [2] suggested a Local Binary Pattern (LBP) and Discrete Cosine Transform-based picture splicing detection approach (DCT). A support vector machine (SVM) is utilized for classification. Li et al. [3] presented an ISD method depending on Markov features and DCT domain. Bunk et al. [4] presented an ISD method depending on Radon transform and deep learning. The entry image is broken into tiny parts in this method, and Radon resampling features are determined from each patch. The Random Walker segmentation approach was utilized to find the forgery. Rachna and Navneet [5] presented an ISD method depending on merging all three domains' Markov characteristics, which are retrieved separately in the spatial, DCT, and DWT domains. For classification, an efficient Ensemble classifier is used. Yıldırım and Ulutas [6] presented an ISD method depending on the hybrid image's statistical and textural properties in the SWT domain. The image is classified using an SVM classifier. Because of the statistical and textural characteristics combined, the suggested method has a high-dimensionality feature vector size, which is one of its flaws. Shilpa et al. [7] presented an ISD method depending on DCT coefficients. The SVM is used to classify authentic and counterfeit images using the retrieved feature vector. Experiments are carried out on a standard dataset of CASIA v1.0 and v2.0 pre- and post-processed faked images. Jaiswal et al. [8] presented an ISD method depending on combining four features HoG, LTE, DWT, and LBP. The judgement about the image's authenticity was predicted using a logistic regression classifier.

All of the approaches for detecting image splicing forgeries listed above are traditional methods that rely on handcrafted features. These methods have a number of flaws, including (1) a high computational complexity. (2) Recognize a certain form of tampering by looking for specific traits in an image. (3) Low detection precision.

The high role of DL approaches in the computer vision fields, as well as the development of GPU technology, prompted researchers to use DL models for image forgery detection. The features extraction and classification phase are merged in DL. Because of its ability to automatically learn abstract and complicated features, these methods were found to be more accurate than traditional methods in identifying modified areas. Furthermore, it saves the time and effort required to locate hand-crafted characteristics in manipulated images. Convolutional Neural Networks (CNN), Deep Neural Networks (DNN), and Recurrent Neural Networks (RNN) are examples of DL models. Among these DL models, CNN are popular. The convolution layer of CNN serves as both a discriminator and a feature extractor.

Rao and Ni [9] presented an ISD method depending on CNN. CNN's first convolution layer is utilized for preprocessing in order to determine the impacts of tampering operations. The pre-trained CNN was then applied to each test image, and an SVM classifier was utilized to detect tampering. The final assessment was carried out using the public CASIA v1.0 and CASIA v2.0 datasets. Salloum et al. [10] presented an ISD method depending on Multi-task Fully Convolutional Network (MFCN). As the basis network, FCN VGG-16 with skip connection was used. Because SFCN only delivers coarse localization output in a few circumstances, it has been proven that Multi-task FCN (MFCN) outperforms single-task FCN (SFCN). Experiments are carried out using a standard dataset like CASIA v1.0 and v2.0. Jaiswal and Srivastava [11] presented an ISD method depending on the pretrained residual network (ResNet-50). Three classifiers are used in order to classify the pretrained feature: SVM, naive Bayes, and KNN. For the final assessment, the public CASIA v2.0 datasets were used. The results of their investigation revealed that SVM has the highest accuracy. Bi et al. [12] presented an ISD method depending on Ringed Residual U-Net (RRU-Net) CNN.RRU-Net intends to increase CNN's learning abilities by utilizing the human brain's memory and consolidation mechanisms. CASIA and COLUMBIA datasets were used to test the RRU-Net. The authenticated and tampered photos are trained and tested using three distinct classifiers: Nave Bayes, K-nearest neighbor, and Multi-Class Model utilizing SVM Learner. Xiao et al. [13] presented an ISD method depending on a coarse-to-refined convolutional neural network (C2RNet). The features at various scales are derived from small patches of the given image via course-CNN and refined-CNN. Ahmed et al. [14] presented an ISD method depending on deep neural network ResNet-Conv. To generate a preliminary feature map of RGB images, two ResNet CNN variants, ResNet-50 and ResNet-101, were employed. Almawas et al. [15] presented an ISD method depending on VGG16, GoogLeNet, and DenseNet201, which are all pretrained CNN models. To find the best one, three classification algorithms are used: SVM, naive Bayes, and KNN.Each model employs a varied number of layers, allowing them to learn various feature representations for a variety of images. The layers in the VGG16 network model are 16, and the image matrix entry size is 224*224. For a GoogLeNet network model, its tailored design enables for greater network depth and width while keeping the computation budget constant. The layers in the DenseNet201 network model are 201, and the image matrix entry size is 224*224. Experiments are carried out using a standard dataset like CASIA v1.0 and v2.0. For all classification techniques, when compared to other approaches, the CASIA v1.0 dataset yielded the lowest results.

Despite the fact that the preceding methods have solved the majority of ISD problems, the existing methods have low detection accuracy, thus there is need for improvement. To address these issues, this paper presented ISD based on deep learning. The proposed image splicing detection is composed of three stages: (1) Image pre-processing includes RGB image conversion and image size fitting. (2) Using the pre-trained CNN AlexNet model, we extract the final discriminative feature for a preprocessed image. (3) Finally, a Canonical Correlation Analysis (CCA) classifier is trained for binary classification (authentic/forged) using the obtained feature representation.

The rest of this paper is organized in the following manner. Section 2 defines the proposed technique. Section 3 discusses the experimental results and analysis. Finally, conclusions can be formed in Section 4.

II. The Proposed Methods

This research proposes an ISD method based on a CNN-based pre-trained AlexNet model to extract deep features with little training time. The classification using canonical correlation analysis was utilized.

Deep Learning reduces the amount of time and effort required to extract hand-crafted characteristics from manipulated images. Deep learning model training, on the other hand, is difficult and takes a lot of computational power and a lot of data. As a result, pre-trained CNN models such as AlexNet [16] were employed as feature extractors to reduce the training effort. Krizhevsky et al. [16] proposed the AlexNet model concept. There are 25 layers in the AlexNet model. The AlexNet model has five convolutional, pooling, three fully connected, and softmax layers, as well as a ReLU activation function. AlexNet is a relatively simple deep CNN architecture that can be trained and optimized easily, compared to other complex deep CNN architectures such as GoogleNet, VGG16 network (consisting of 13 convolutional layers and three fc layers), VGG19 (contains 16 convolutional layers and three fc layers), and DenseNet201.

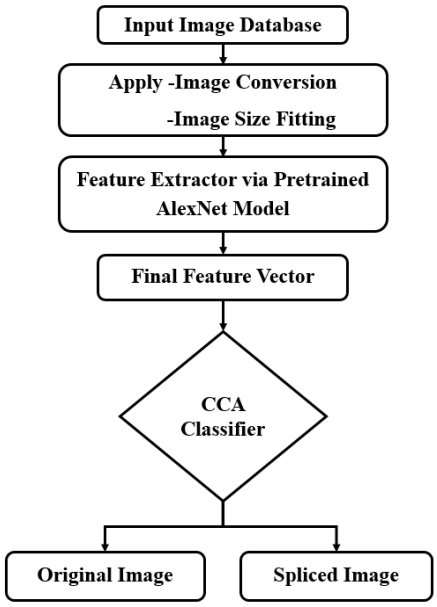

Fig. 2. Flow diagram of a proposed ISD Detection.

There are three components to the proposed techniques phases. The proposed work is depicted in Fig. 2 as a block diagram. The details for each step are as follows:

Phase 1: Pre-processing Stage

To reduce the total computational complexity, the images in the datasets were preprocessed by transformed into gray-scale images. The images of the datasets were then downsized to 227 227 pixels in accordance with the AlexNet model's first input layer.

Phase 2: Feature Extraction Stage

In any ISD method, this phase is essential. In order to obtain the features, the extraction were applied by using the AlexNet model that had been pretrained. After that, the fully connected fc7 layer's 4096-dimensional output was extracted.

Phase 3: Classification Stage

DIF detection is a two-class problem, with original and forged images. Canonical Correlation Analysis (CCA) is currently able to produce satisfied and effective classifier results in a variety of fields [17]. H. Hotelling [18] invented canonical correlation analysis in 1935 as a statistical tool for identifying and quantifying correlations between two sets of variables. For discriminative tasks, Canonical Correlation Analysis (CCA) is widely and recently employed. When the correlation between two linear combinations is greatest, the CCA classifier employs the canonical vector of a linear combination of variables as the discriminant function.

III. Experimental Results and Analysis

This section is dedicated to evaluating the proposed method. First, the CASIA v1.0 dataset is described. Performance measures are briefly discussed after that. After then, the experimental findings are examined, and the proposed method is compared to other methods. Table 1 lists the properties of the equipment that was used to conduct experiments.

TABLE I. The Properties of Experimentation

Hardware	Properties
PC	HP laptop
Operating system	Microsoft Windows 10 64-bit (OS)
RAM	8 GB
Processor	Intel(R) Core(TM) i7-6500U CPU @ 2.50GHz 2.60 GHz
Software	MATLAB version R2020a
Graphic Card	Intel® HD Graphics 520 (NVIDA GTX 950M)

A. Data Set

Public CASIA v1.0 image dataset [19] is used to test the performance of the proposed method. The CASIA v1.0 contains 1721 images, including 800 original images, and 921 forgeries. These color images range in resolution from 384×256 pixels. JPEG is the format of these images without any post-processing. Some of images from the CASIA v1.0 image dataset are shown in Fig. 3.

Fig. 3. CASIA v1.0 image dataset samples. The first row has original photos, while the second row contains spliced photos.

B. Performance Evaluation Measures

To test their suggested image splicing detection methods, the majority of the researchers used some of detection accuracy measures. As a result, we evaluated the proposed method using the accuracy metric (also known as recognition rate). The percentage of images accurately derived from a mixture dataset of original and forged images is known as accuracy. The detection accuracy is defined as (1) [20]:

$$Accurcy = \frac{(TP + TN)}{(TP + TN + FP + FN)} x\ 100\ \% \qquad (1)$$

TP stands for True Positive, TN stands for true Negative, FP stands for False Positive, and FN stands for False Negative.

C. *Proposed Performance Evaluation*

The k-fold cross-validation (CV) assessment is one of the most well-known in the area of machine learning. When the training dataset is modest, this evaluation scheme produces reliable findings. As a result, using a 10-fold CV method, the dataset is separated into training and testing. The CCA classifier is trained by randomly selecting features from 70% of the photos in the data sample. The remaining 30% of image features are used for testing purposes. The technique has been repeated ten times, with each test set being distinct. Finally, the presented method's detection accuracy is calculated as the average accuracy of all 10 evaluations. Under the CASIA-1 Database, the accuracy of our suggested ISD, which is relied on a pre-trained AlexNet with a CCA classifier, is 98.79 percent. It is possible to conclude that the features collected from the Pretrained AlexNet model are effective and simple to classify using CCA.

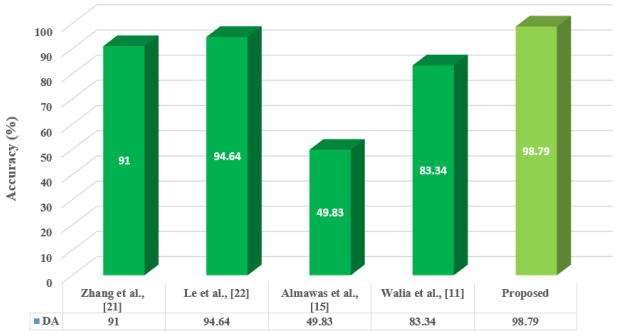

Fig. 4. Comparative Results Description Graph.

D. *Result Comparison with Existing Methods*

In order to further illustrate the efficiency of the proposed method, its results are compared to those of other existing ISD that are currently available. For comparison, the methods [21], [22], [15], and [11] were utilized. It should be mentioned that the proposed method, as well as others, was tested on the CASIA v1 image dataset. The comparison findings in Table 2 show that the proposed method is superior when it comes to detection precision.

Table 2 shows that the proposed method has a detection accuracy of 98.79 percent, while Le et al., [22] have the second-best detection accuracy of 94.64 percent. As a result, the feature of the VGG16 model that has been pre-trained exhibits strong robustness. However, by utilizing the features of a pre-trained AlexNet model, we are able to achieve excellent accuracy. It's worth mentioning that Almawas et al., [15] approach is the least effective of all the ways. The reason for this is that the combined feature of Pretrained VGG16, GoogLeNet, and DenseNet201 Models produces poor results. The proposed method surpasses the others by attaining great accuracy in discriminating original and spliced images, as shown in the Table 2. In comparison to other current strategies based on various pretrained CNN models, the proposed method has a highest precision of roughly 98.79 percent, as shown in Fig. 4.

TABLE II. THE COMPARATIVE RESULTS WITH THE CURRENT SPLICING DETECTION ALGORITHMS.

Techniques	**Feature Extraction**	**Classifier**	**DB**	**Accuracy (%)**
Zhang et al., [21]	Deep Feature (CNN)	Threshold Classifier	CASIA v1.0	91
Le et al., [22]	Pre-trained VGG16 Model	CNN	CASIA v1.0	***94.64***
Almawas et al., [15]	Pretrained VGG16, GoogLeNet, and DenseNet201 Models	SVM	CASIA v1.0	49.83
Walia et al., [11]	Pre-trained ResNet-18 Model	Shallow Neural Network	CASIA v1.0	83.34
Proposed	Pretrained AlexNet Model	CCA	CASIA v1.0	***98.79***

IV. CONCLUSION

This paper proposes a method for detecting image splicing that uses a CNN-based AlexNet model that has been pre-trained to extract deep features without requiring extensive training. In the presented method, the CCA is also used as a classifier. There are three stages to the proposed image splicing detection: (1) RGB image conversion and image size fitting are examples of image pre-processing. (2) We extract the final discriminative feature for a preprocessed image using the pre-trained CNN AlexNet model. (3) Finally, the resulting feature representation is used to train a CCA classifier for binary classification (authentic/forged). According to the results of the testing, the proposed method has a highest accuracy of around 98.79 percent. On the CASIA v1 dataset, the proposed method surpasses other current methods, with an accuracy of 98.79 percent. The proposed method performs well because it employs a deep CNN AlexNet-based feature extractor, and the system's transfer learning component can learn the differences between original and spliced images. The features gathered from the Pretrained AlexNet model can be concluded to be effective and straightforward to classify using CCA. The technique could be used against additional types of image forgery in the future, such as image retouching. It can be achieved by fine-tuning the model with datasets that include forgery examples.

ACKNOWLEDGMENT

This research is supported by Uniten BOLD Publication Fund 2021.

REFERENCES

[1] K. B. Meena and V. Tyagi, "A Deep Learning based Method for Image Splicing Detection," in Journal of Physics: Conference Series, 2021, vol. 1714, no. 1, p. 12038.

[2] A. A. Alahmadi, M. Hussain, H. Aboalsamh, G. Muhammad, and G. Bebis, "Splicing image forgery detection based on DCT and Local Binary Pattern," in 2013 IEEE Global Conference on Signal and Information Processing, 2013, pp. 253–256.

[3] C. Li, Q. Ma, L. Xiao, M. Li, and A. Zhang, "Image splicing detection based on Markov features in QDCT domain," Neurocomputing, vol. 228, pp. 29–36, 2017.

[4] J. Bunk et al., "Detection and localization of image forgeries using resampling features and deep learning," in 2017 IEEE conference on computer vision and pattern recognition workshops (CVPRW), 2017, pp. 1881–1889.

[5] R. Mehta and N. Agarwal, "Splicing detection for combined dct, dwt and spatial markov-features using ensemble classifier," Procedia Comput. Sci., vol. 132, pp. 1695–1705, 2018.

[6] E. O. Y\ild\ir\im and G. Uluta\cs, "Augmented features to detect image splicing on SWT domain," Expert Syst. Appl., vol. 131, pp. 81–93, 2019.

[7] S. Dua, J. Singh, and H. Parthasarathy, "Image forgery detection based on statistical features of block DCT coefficients," Procedia Comput. Sci., vol. 171, pp. 369–378, 2020.

[8] A. K. Jaiswal and R. Srivastava, "A technique for image splicing detection using hybrid feature set," Multimed. Tools Appl., pp. 1–24, 2020.

[9] Y. Rao and J. Ni, "A deep learning approach to detection of splicing and copy-move forgeries in images," in 2016 IEEE International Workshop on Information Forensics and Security (WIFS), 2016, pp. 1–6.

[10] R. Salloum, Y. Ren, and C.-C. J. Kuo, "Image splicing localization using a multi-task fully convolutional network (MFCN)," J. Vis. Commun. Image Represent., vol. 51, pp. 201–209, 2018.

[11] A. K. Jaiswal and R. Srivastava, "Image splicing detection using deep residual network," 2019.

[12] X. Bi, Y. Wei, B. Xiao, and W. Li, "RRU-Net: The ringed residual U-Net for image splicing forgery detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2019, p. 0.

[13] B. Xiao, Y. Wei, X. Bi, W. Li, and J. Ma, "Image splicing forgery detection combining coarse to refined convolutional neural network and adaptive clustering," Inf. Sci. (Ny)., vol. 511, pp. 172–191, 2020.

[14] B. Ahmed, T. A. Gulliver, and S. alZahir, "Image splicing detection using mask-RCNN," Signal, Image Video Process., vol. 14, no. 5, pp. 1035–1042, 2020.

[15] L. Almawas, A. Alotaibi, and H. Kurdi, "Comparative performance study of classification models for image-splicing detection," Procedia Comput. Sci., vol. 175, pp. 278–285, 2020.

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105.

[17] G. Yang and H. Zhang, "The relationship between canonical correlation analysis and minimum squared error classifier," in 2008 9th International Conference on Signal Processing, 2008, pp. 1647–1651.

[18] C. E. Heckler, "Applied multivariate statistical analysis." Taylor \& Francis, 2005.

[19] J. Dong, W. Wang, and T. Tan, "Casia image tampering detection evaluation database," in 2013 IEEE China Summit and International Conference on Signal and Information Processing, 2013, pp. 422–426.

[20] X. Zhao, S. Li, S. Wang, J. Li, and K. Yang, "Optimal chroma-like channel design for passive color image splicing detection," EURASIP J. Adv. Signal Process., vol. 2012, no. 1, pp. 1–11, 2012.

[21] Y. Zhang, J. Goh, L. L. Win, and V. L. L. Thing, "Image Region Forgery Detection: A Deep Learning Approach.," SG-CRC, vol. 2016, pp. 1–11, 2016.

[22] T. Le-Tien, P. X. Hanh, N. Pham-Ng-Quynh, and D. Ho-Van, "A combination of Super-resolution and Deep Learning Approaches applied to Image Forgery Detection," in 2020 International Signal Processing, Communications and Engineering Management Conference (ISPCEM), 2020, pp. 244–249.