Image Splicing Detection using Deep Residual Network

Article *in* SSRN Electronic Journal · January 2019

DOI: 10.2139/ssrn.3351072

2 authors:

Ankit Jaiswal
Indian Institute of Technology (Banaras Hindu University) Varanasi
9 PUBLICATIONS **21** CITATIONS

SEE PROFILE

Rajeev Srivastava
Indian Institute of Technology (Banaras Hindu University) Varanasi
156 PUBLICATIONS **864** CITATIONS

SEE PROFILE

Image Splicing Detection using Deep Residual Network

Ankit Kumar Jaiswal[1] and Rajeev Srivastava[2]

Abstract—Forgery using images are common nowadays. This may result in misleading the court, changing the mindset of people and defaming an individual. It is the need of the hour to design a tool that can detect forged and authenticated images. Image forgery detection schemes may be active or passive. Tampering detection schemes come into the category of passive or blind image forgery detection schemes. Deep Learning is a technique used to recognize or classify images into multiple class. Images are used as input for the convolutional neural network and processed through various layers for feature extraction and these extracted features are used as a training vector for the classifier model. This paper uses a pre-trained Deep Learning model resnet-50 for feature extraction from CASIA 2.0 dataset and three different classifiers for classification purpose.

Keywords: Image Splicing Detection, Deep Learning, Deep Residual Network Classification Model

I. INTRODUCTION

With the availability of the technology on a fingertip, an increase in number crime is observed in the cyber world. The most prevalent crime is cyber defamation, forgery and obscene or offensive content [1]. As the images, being of the most convenient tool for expressing one's views, thoughts and idea, also act as a weapon for defaming an individual, misleading the court system by altering and manipulating evidence in form of an image. Manipulation of the images previously was done to enhance the quality of images, but nowadays this is done either for character assassination or for defamation of famous personality. For manoeuvring of digital images, several techniques are there like copy-move, cloning, cropping, image enhancement, image morphing, retouching, etc. In order to detect this, there are two main classes under which the detection technique is classified that is active detection technique and passive detection technique. The classified Fig. 1 explains the forgery detection techniques.

In Active detection technique, the images are tested by checking the watermarks or digital signature embedded on images whereas passive detection is also known as blind detection technique, no such pre-embedded symbols are required [2]. Passive detection involves cloning, copy-move forgery and compositing [3]. Image splicing is the smallest operation performed in image compositing. Image splicing is the fundamental step of image compositing. In Image splicing fragments of different images are put together for image manipulation. In this, no further processing of images is made like removing the blur effect, smoothening the boundaries, which makes it difficult for recognizing the manipulation or tempering. Image splicing is the basic operation involved in the digital photomontage in which photos are produced by pasting images using different editing tool like photoshop. In image splicing, parts of different images are being joint together without any further processing of the tampered image. An example of image splicing is shown in Fig. 2. In this image first, two photographs tamper where India's Prime Minister Narendra Modi is touching feet of Akbaruddin Owaisi and Saudi King and last one is original where he is touching feet of L.K. Advani. In the spliced photographs L.K. Advani is replaced by Akbaruddin Owaisi and Saudi King [4].

Figure 1: Types of image forgery detection

[1]*Computing and Vision Lab,*
Department of Computer Science and Engineering,
Indian Institute of Technology (BHU), Varanasi, India
[2]*Computing and Vision Lab,*
Department of Computer Science and Engineering,
Indian Institute of Technology (BHU), Varanasi, India
E-mail: [1]akjiitbhu@gmail.com, [2]rajeev.cse@iitbhu.ac.in

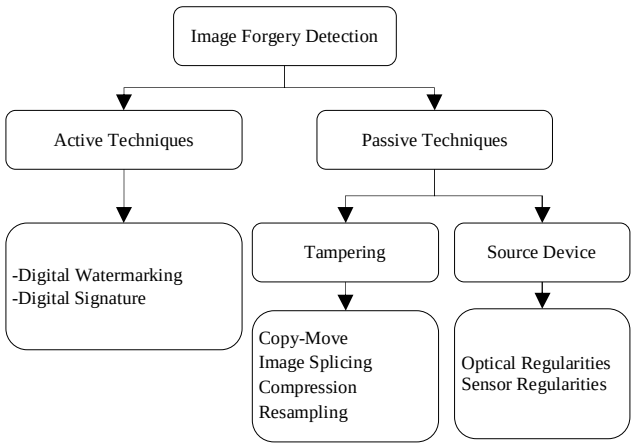

Figure 2: Example of image splicing

Thus, there should be an automatic detection scheme for image splicing detection so that on can assure about the authenticity of an image. In this case, images can be classified into two categories- one is spliced or tampered, and another is authenticated or original. This is a type of binary class classification problem. Some state-of-the-art techniques are already there to classify images. These techniques are based on features extracted from images and machine learning classification techniques.

A deep learning convolutional neural network (CNN) model is used in this paper to predict forged images. A large number of input images are given to pre-trained residual neural network (RESNET-50) to train the classifier model and predict for other images [5]. In this paper, three different classifiers are used to train and test the authenticated and tampered images.

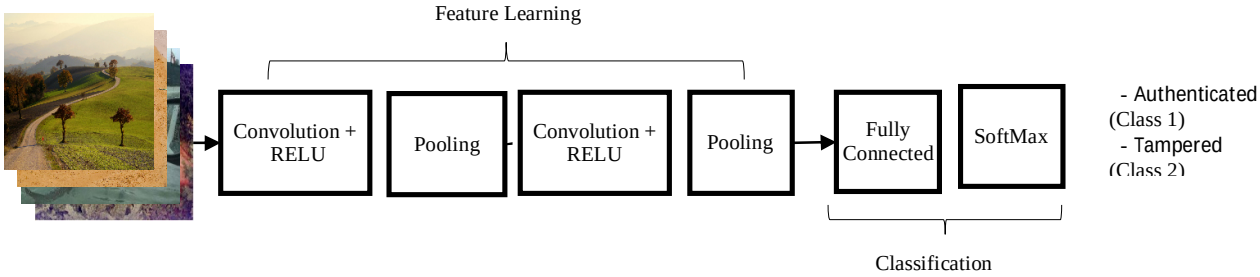

Figure 3: Deep learning model

This paper is divided into five different sections, the first section is all about the background and introduction part. The second section is a brief description of related works in this field. The proposed approach is given in the third section. Section four discusses the experiment and result. Section five concludes the pa-per.

II. Related Work

Various state of the art techniques are there to classify images into two classes as authenticated or forged. Machine learning classification techniques are there for binary class classification. Here, some of those techniques are discussed–A novel approach for image splicing detection is given by [6]. This approach is basically based on chrominance of an image where features from chroma channel are extracted. Local binary patterns (LBP) of both the chroma are calculated and from then Discrete Cosine Transform (DCT) of these LBP are transformed in the form of 16x16 blocks. From these DCT coefficients, standard deviation features were calculated. These features are trained using Support Vector Machine (SVM) classifier.

A machine learning approach based on Gray-Level Cooccurrence Matrix (GLCM) features is given in [7]. In this approach, image is pre-processed into YCbCr color space. As chrominance is more sensitive than luminance approach is based on chrominance channel of YCbCr color space. GLCM features are extracted from chrominance channels and to reduce these features Boost Feature Selection (BFS) techniques is used then these extracted features are trained using LIBSVM classifier.

Another image splicing detection technique is proposed in [8]. This approach is based on features of DCT and Discrete Wavelet Transform (DWT). Markov features from the coefficients of DCT and DWT are extracted and to reduce the computational cost SVM-RFE is used to minimize the feature set. For the classification, SVM is used to train and test the dataset. Three datasets are used in this approach COLUMBIA, CASIA1 and CASIA2.

A method using run-length encoding feature has been proposed in [9]. This technique suggests that if detection of image splicing is tough in one color space then another color space can also be used for feature extraction. In this proposed method 4 grey level run length numbers are extracted from the chrominance channel of the image for four different directions. These features are trained using SVM classifier. The method also explains that extracted features from the chrominance channel have better performance than extracted features from individual red, green and blue channel.

III. Proposed Model

In the field of image recognition or image classification, deep learning is widely used. For the classification of images, CNN takes images as input, process it on

different layers and classify it into different categories. Technically, images are passed through a series of convolution layers as input and process with different types of filter, pooling and fully connected layers. Using these layers features are extracted, these features are used for training and test purpose. Training features are used to train the classifier model and test features are used to predict the class using the trained classifier model.

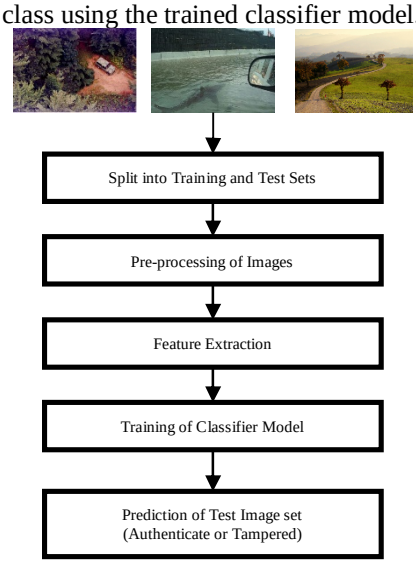

Figure 4: Flow diagram of proposed model

The proposed method is based on pre-trained residual network (resnet-50). Classification of images into different classes using deep learning is similar to a machine learning algorithm, three steps are pre-processing, feature extraction and classification. Input data is first split into two parts, one is the training image set and another is test image set. Training and test both image set are pre-processed first to resize the images according to the pre-trained network size ( ). Then images are passed through various layers of the network till fully connected layer (FC-1000) to extract features from image set. Now trained features are passed to the classifier to train the classifier model. Using this trained model pre-diction of test image sets are done. This method uses three different classifiers Naïve Bayes, K-nearest neighbour and Multi-Class Model using SVM Learner. The results from all these classifier models are given in experiment and result section.

IV. Experiment and Result

The experiment is performed on the server with Xeon processor, 16 GB ram and on Ubuntu Linux server operating system. Tool for the experiment is used MATLAB R2017b. As deep learning CNNs use GPU for a large number of images (dataset) to perform the experiment, the proposed experiment is done on 12613 number of images known as CASIA 2.0 dataset [10]. This dataset contains 7491 authenticated images and 5122 tampered images with datatype .jpg and .tif. Size of images are Images are of different categories like animals, architecture, article, character, plant, nature, scene and textures, except these indoors images are also there. For the purpose of evaluation, classification algorithms use confusion matrix, where classifier distinguish dataset into different classes. In this case, there are two classes one is Authenticated, and another is tampered. Confusion chart is divided into two columns and two rows. Based on this confusion matrix accuracy, specificity and sensitivity is analysed. Receiver Operating Characteristics (ROC) curve is also shown in the result and discussion section.

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total} \quad (1)$$

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive} \quad (2)$$

$$Sensitivity = \frac{True\ Positive}{False\ Negative + True\ Positive} \quad (3)$$

$$Total = True\ Positive + False\ Positive + True\ Negative + False\ Negative$$

As mentioned above, the experiment is done on CASIA 2.0 dataset, it is divided into two parts one is authenticated and another is tampered. In the proposed approach whole dataset is taken for the experiment purpose. This proposed method is based on pre-trained network RESNET-50. Training of images is done till 1000 fully connected layers. Images are split into two parts one is for training purpose and another is for test purpose. Mixed 8829 images are used for training purpose and 3784 images are used for test purpose. Classifiers used in this approach are Naïve Bayes, Multiclass model using SVM Learner and K-Nearest Neighbour. Using multi-class model classifier correct predicted images are 1684 while using Naïve Bayes and K-NN are 1215 and 1555 respectively. Comparison of all three classifiers with their accuracy, specificity and sensitivity are given in below table.

Table 1: Comparison of three different classification model

Classifier Model	*Accuracy*	*Sensitivity*	*Specificity*
Multiclass Model	0.7026	0.6339	0.7497
Naïve Bayes	0.5991	0.5047	0.7147
K-Nearest Neighbor	0.5991	0.5071	0.6533

ROC curve of all three classifier are shown in below graph:

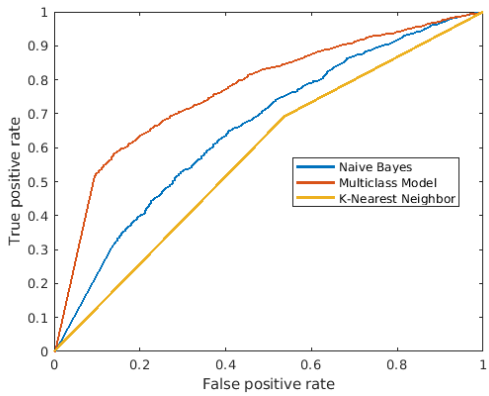

Figure 5: ROC curve for three different Classification Model

V. CONCLUSION

Image splicing is a technique where parts of two or more images are composited together. The purpose may be different, but this is a very important issue. To detect spliced images various state-of-the-art machine learning techniques are there which are already discussed in related works. In this paper, a deep learning CNN approach has been discussed which is based on pre-trained residual network. In this approach three classifiers Multiclass Model using SVM Learner, K-NN and Naïve Bayes are used to train the classifier model. With the help of an experiment on MATLAB accuracy for all three classifiers are 70.26%, 59.91% and 59.91%.

REFERENCES

[1] Birajdar, G.K., Mankar, V.H.: Digital image forgery detection using passive techniques: A survey. Digit. Investig. 10, 226–245 (2013).

[2] Korus, P.: Digital image integrity—a survey of protection and verification techniques. Digit. Signal Process. 71, 1–26 (2017).

[3] Al-Qershi, O.M., Khoo, B.E.: Passive detection of copy-move forgery in digital images: State-of-the-art. Forensic Sci. Int. 231, 284–295 (2013).

[4] Swati: Viral photo of modi touching feet of saudi king and akbaruddin owaisi is fake, https://www.thelallantop.com/jhamajham/viral-photo-of-modi-touching-feet-of-saudi-king-and-akbaruddin-owaisi-is-fake/, (2018).

[5] Zhong, Z., Li, J., Ma, L., Jiang, H., Zhao, H.: Deep Residual Networks for Hyperspectral Image Classification. 3–6 (2017).

[6] Alahmadi, A.A., Hussain, M., Aboalsamh, H., Muhammad, G., Bebis, G.: Splicing Image Forgery Detection Based on DCT and Local Binary Pattern. Glob. Conf. Signal Inf. Process. IEEE. 253–256 (2013).

[7] Wang, W., Dong, J., Tan, T.: Effective Image Splicing Detection Based on Image Chroma. In: IEEE International Conference on Image Processing. pp. 1257–1260 (2009).

[8] He, Z., Lu, W., Sun, W., Huang, J.: Digital image splicing detection based on Markov features in DCT and DWT domain. Pattern Recognit. 45, 4292–4299 (2012).

[9] Zhao, X., Li, J., Li, S., Wang, S.: Detecting digital image splicing in chroma spaces. Int. Work. Digit. Watermarking. 12–22 (2010).

[10] Dong, J., Wang, W.: CASIA v1.0 and CASIA v2.0 Image Splicing Dataset, http://forensics.idealtest.org.