Advances in Social Science, Education and Humanities Research, volume 529
Proceedings of the International Conference on Engineering, Technology and Social Science (ICONETOS 2020)

Effectiveness of Deep Learning Architecture for Pixel-Based Image Forgery Detection

Hisyam Fahmi[1,*], Wina Permana Sari[2]

[1] *Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim, Indonesia*
[2] *Computer Science Department, School of Computer Science, Bina Nusantara University, Indonesia*
[*]*Corresponding author. Email: hisyam.fahmi@uin-malang.ac.id*

ABSTRACT
Digital image forgery or forgery is easy to do nowadays. Verification of the authenticity of images is important to protect the integrity of the images from being misused. The use of a deep learning approach is state-of-the-art in solving cases of pattern recognition, the one is image data classification. In this study, image forgery detection was carried out using a deep learning-based method, the Convolutional Neural Network (CNN). The analysis of the different architecture of CNN has been done to show the effectiveness of each architecture. Two architectures were tested to know which one is more effective, architecture 1 has three convolution and pooling layers with 256 × 256 × 3 image input. While the other architecture has two convolution layers and pooling with 128 × 128 × 3 image input. The results show that the accuracy rate of the image forgery detection model in each architecture is around 80%. However, the validation accuracy is not more than 70%.

Keywords: *convolutional neural network, copy-move forgery, deep learning, digital image forensics*

1. INTRODUCTION

In the digital era today, image/photo data is very vulnerable to forgery [1]. It can be done using sophisticated image editing software that is easily available today, not only on personal computers and laptops but also on mobile devices. The results of this image forgery are widely used by people on social media, in the commercial field, and some even use them for criminal purposes. The use of image forgery for matters that violate the law needs to be of great concern and can pose a threat to society, government and business. Therefore, the images around us need to be verified for authenticity. Protecting the integrity of digital images is important. So that in this condition, the authenticity of digital images can be ascertained by utilizing an image forgery detection algorithm [2]. Digital image authenticity detection, both in terms of the integrity of the image content and its source, is the field of Digital Image Forensics (DIF).

Algorithms for detecting forgery image in DIF are classified as active and passive forgery detection approaches [3]. The passive forgery detection approach does not require prior knowledge of the image content. In contrast, the active approach involves the process of authenticating the image by extracting the watermark and digital signature embedded. So, any forgery operation performed on the image can break the embedded watermark and digital signature, helping to detect the authenticity of the image. The passive image forgery detection that most influences the original image is the copy-move (cloning) forgery method [4], [5].

Research on deep learning is the current trend for solving problems in computer vision, such as image classification. This is because the architecture in deep learning, one of which is the Convolutional Neural Network (CNN), can extract complex statistical features from high-dimensional data. Deep learning has also been applied to passive image forgery detection applications [6], [7], [8]. However, conventional deep learning frameworks should not be used directly because fake images are difficult to distinguish from the original images with many current image forgery tools, so it is necessary to modify the input and architecture used [9]. Therefore, in this study, image forgery detection, especially copy-move, was carried out using a deep learning approach.

2. PROPOSED METHOD

Liliana and Basaruddin [10] apply numerical computation techniques to detect fake images. The

method used is the singular value decomposition (SVD). The experimental results on images with various conditions were successful in detecting the forgery image with a threshold value of 0.2.

Dehnie, et al. [11] discussed digital image forensic techniques to distinguish images captured by digital cameras from computer-generated images. This difference is captured in terms of the residual image properties extracted by a wavelet-based denoising filter. The results of this study indicate that the two types of residues obtained from different digital camera images and computer-generated images have some general characteristics that are not present in other types of images.

Warbhe and Dharaskar [12] present an active approach to identify and authenticate original digital images from forged or tampered with. The experimental results show how the Independent Component Analysis (ICA) method is successful in extracting and detecting image forgery if it is in the image. While this method is good at detecting adulteration in images, the main limitation of this method is that it requires both a faked image as well as an original forged image. This limitation can be overcome by using and applying single-channel ICA to a single spurious image to extract the fakes.

Rao and Ni [6] proposed a new image forgery detection method based on a deep learning technique, which leverages the Convolutional Neural Network (CNN) to automatically learn a hierarchical representation of an input RGB colour image. The weights in the first layer of the network are initialized with the basic high-pass filter set used in the calculation of the residual map in the spatial rich model (SRM), which functions as a regularizer to efficiently suppress the effects of image content and capture invisible forgeries introduced by tampering operations.

Bayar and Stamm [13] have developed a new layer type on CNN called the Constrained Convolutional Layer which is adaptively capable of learning features to detect image manipulation. The experiment shows that the CNN architecture can detect several different forgery operations with an accuracy of 99.97%.

We proposed the deep learning approach for image forgery detection. The analysis of the different architecture of CNN has been done to show the effectiveness of each architecture. Two architectures were tested to know which one is more effective.

2.1. Digital Image Forgery Detection

Passive image forgery detection techniques can be divided into five categories: pixel-based, format-based, camera-based, physical environment-based, and geometry-based techniques [14]. Pixel-based techniques detect statistical anomalies that exist at the pixel level; format-based techniques make use of statistical correlation introduced by lossy compression schemes; camera-based techniques exploit images processed by camera lenses, sensors, or on-chip post-processing; physical environment-based techniques explicitly model and detect anomalies in three-dimensional interactions between physical objects, light, and cameras; and geometry-based techniques for measuring objects and their position relative to the camera.

Pixel-based techniques emphasize the processing of digital image pixels. These techniques can be categorized into four types: cloning (copy-move), splicing, resampling (resize, stretch), and statistical. The types of copy-move and splicing techniques are the most common image forgery detection techniques. In the copy-move technique, parts of the image are copied and pasted elsewhere in the image. In the splicing technique, 2 or more images are combined into one composite image [15].

2.2. Convolutional Neural Network (CNN) Architecture

In this research, deep learning methods, which is CNN, is used to detect image forgeries. Technically, this network is designed to extract the relevant features for classification, namely those that minimize the loss function. Network parameter-kernel weight trained by the Gradient Descent method to produce the most discriminating features of the rendered image to the network [16]. These features are then assigned to the fully connected layer to perform classification [17].

The architecture used is inspired by the architecture given in Rao and Ni's research [6]. The image size used in their study is 128 × 128 × 3 with ten convolutional and pooling layers. Whereas in this study, the image resized to a size of 256 × 256 × 3. The CNN architecture consists of three convolution layers with a 5 × 5 kernel and three pooling layers. Figure 1 shows the architecture in the experiment. Also, the experiment compared against CNN with two convolutional layers with a 3 × 3 kernel and two pooling layers (Figure 2).

2.3. Dataset

The dataset used in this study is photo image data taken by mobile phone cameras. From the original photo, a copy-move forgery and/or splicing is made. The number of original photos and their forgeries are 20 images each. Added a dataset of manipulated images from the Pattern Recognition Laboratory, Computer Science, Friedrich-Alexander-Universität Erlangen-Nürnberg (https://lme.tf.fau.de/) totalling 48 images for each category of original and modified images. Also added is the dataset from the CASIA ITDE database [18] which was also used in the study of Warif et al. [19], [20] with a total of 510 images, 255 images for each category,

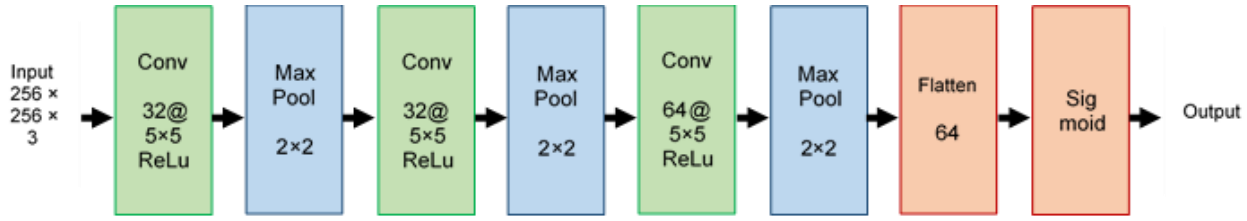


Figure 1 The first architecture of CNN for image tampering detection

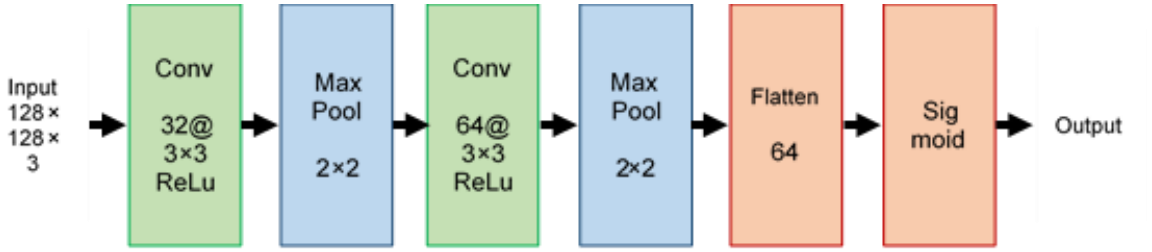


Figure 2 The second architecture of CNN for image tampering detection

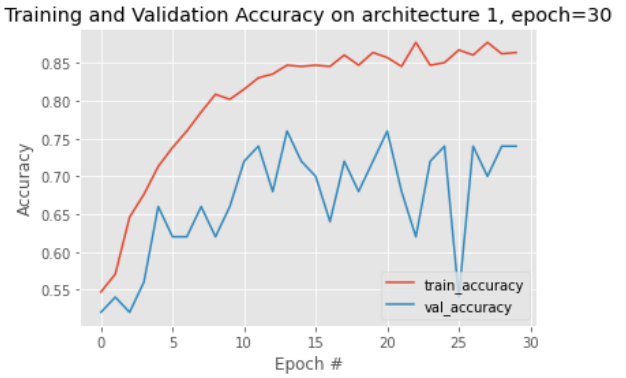
original and forgeries. So that the total number of datasets used is 323 original images and 323 forgeries images, a total of 646 data. From the dataset, 596 images were used for training data, and the remaining 50 images were used for validation. Examples of original and forgeries images can be seen in Figure 3.

3. RESULTS AND DISCUSSION

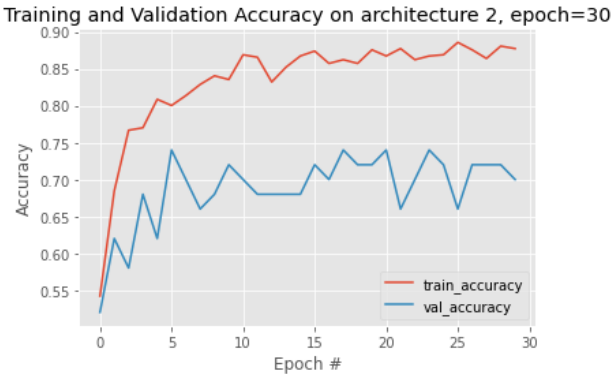
Experiments were conducted to demonstrate the effectiveness of a deep-learning approach using CNN for image forgery detection. The device used for testing has an Intel Core-I7 4702MQ specification with 12GB RAM. The implementation of the CNN architecture uses the *Keras* library in Python.

The image is processed on the CNN architecture by conducting a training and validation process. Cross-validation is used to evaluate the performance of the proposed image forgery detection scheme. From the dataset, 596 images were used for training data, and the remaining 50 images were used for validation.

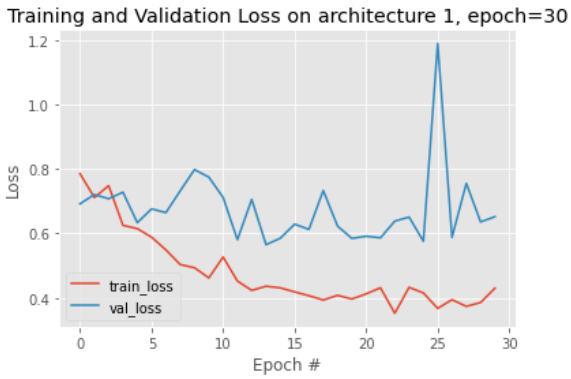
Figure 3 Examples of original (first row) and modified image (second row) from three datasets



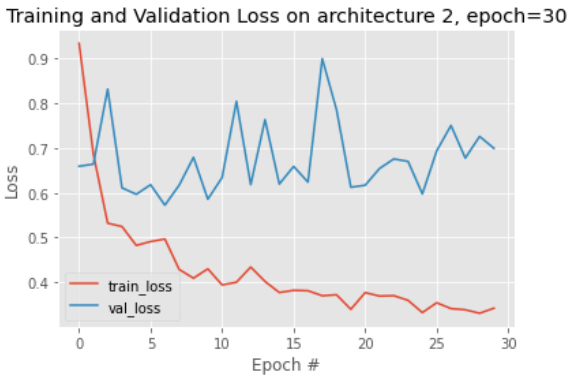
(a) Accuracy for CNN architecture 1



(b) Accuracy for CNN architecture 2



(c) Loss for CNN architecture 1



(d) Loss for CNN architecture 2

Figure 4 Graph of accuracy and loss of the training and validation dataset for both CNN architecture

Extensive experiments on several image datasets were carried out with epoch 30 times and batch size 25. In the training stage, a classification model was obtained, and the accuracy and loss values were calculated using the same data. Then, the model validates with different data to see the accuracy and loss of validation from the new data.

Table 1. Result comparison between CNN architecture 1 and 2

	CNN arch. 1		CNN arch. 2	
	Average	Std. Deviation	Average	Std. Deviation
Accuracy	0.8032	0.0872	**0.8344**	0.0691
Validation Accuracy	0.6680	0.0724	**0.6873**	0.0488
Loss	0.4792	0.1135	**0.4260**	0.1212
Validation Loss	0.6775	0.1148	**0.6741**	0.0779
Time (s)	55.4667	1.5217	**21.9**	1.2477
Time/Step (s)	2	0	**0.9111**	0.0514

Graphs of accuracy, loss, validation accuracy and validation loss for each epoch are shown in Figure 4. Figure 4(a) shows a graph of the accuracy in architecture 1, with a maximum accuracy of 0.8775, while architecture 2 (Figure 4(b)) produces a maximum accuracy of 0.8859. In the architecture 1, the model converges with accuracy above 0.85 after 25th epoch, and architecture 2 converge after 15th epoch.

Table 1 shows the comparison of the average and standard deviation with 30 epochs between CNN architecture 1 and 2. The accuracy of the model using architecture 2 with two convolution layers is better than architecture 1 which has three convolutional layers. Architecture 2 excels in all aspects, from model accuracy, validation accuracy, to training execution time.

4. CONCLUSIONS

Detection of image forgeries can be done using a deep learning approach with the convolutional neural network (CNN) method. In this research, 2 architectures of CNN were tested, architecture 1 has three convolution and pooling layers with 256 × 256 × 3 image input. While the other architecture has two convolution layers and pooling with 128 × 128 × 3 image input.

From the results, it was found that CNN could produce an image forgery detection model with an average accuracy above 80%. The comparison of the architecture used shows that with architecture 2 can recognize modified images quite effectively. In terms of training accuracy, validation, and time, it shows the advantages of architecture with only 2 convolutional layers.

Both architectures have a low level of recognition with new data, as seen from the value of validation accuracy, which averages less than 70%. This is an indication of overfitting. It is recommended in future studies to use a greater number of image data to improve validation accuracy. It is also necessary to detect the pixel location in the modified image.

ACKNOWLEDGMENTS

Thanks to the Republic of Indonesia Ministry of Religious Affairs for funding this work.

REFERENCES

[1] R. Ahmed and R. V Dharaskar, "Study of Mobile Botnets: An Analysis from the Perspective of Efficient Generalized Forensics Framework for Mobile Devices General Terms," in *National Conference on Innovative Paradigms in Engineering & Technology (NCIPET)*, 2012, pp. 5–8, [Online]. Available: http://us.norton.com/theme.jsp?themeid=botnet,.

[2] A. D. Warbhe, R. V. Dharaskar, and V. M. Thakare, "Computationally Efficient Digital Image Forensic Method for Image Authentication," in *Procedia Computer Science*, 2016, pp. 464–470, doi: 10.1016/j.procs.2016.02.089.

[3] V. Conotter, G. Boato, and H. Farid, "Active and Passive Multimedia Forensics," 2011.

[4] M. Mishra and M. C. Adhikary, "Digital Image Tamper Detection Techniques - A Comprehensive Study," *Int. J. Comput. Sci. Bus. Informatics*, vol. 2, no. 1, 2013.

[5] S. Sadeghi, S. Dadkhah, H. A. Jalab, G. Mazzola, and D. Uliyan, "State of the art in passive digital image forgery detection: copy-move image forgery," *Pattern Anal. Appl.*, vol. 21, no. 2, pp. 291–306, May 2018, doi: 10.1007/s10044-017-0678-8.

[6] Y. Rao and J. Ni, "A deep learning approach to detection of splicing and copy-move forgeries in images," Jan. 2017, doi: 10.1109/WIFS.2016.7823911.

[7] Y. Abdalla, M. Iqbal, and M. Shehata, "Convolutional Neural Network for Copy-Move Forgery Detection," *Symmetry (Basel).*, vol. 11, no. 10, p. 1280, Oct. 2019, doi: 10.3390/sym11101280.

[8] A. Kuznetsov, "Digital image forgery detection using deep learning approach," in *Journal of Physics: Conference Series*, Nov. 2019, vol. 1368, no. 3, p. 32028, doi: 10.1088/1742-6596/1368/3/032028.

[9] Y. Zhang, J. Goh, L. L. Win, and V. Thing, "Image region forgery detection: A deep learning approach," in *Cryptology and Information Security Series*, 2016, vol. 14, pp. 1–11, doi: 10.3233/978-1-61499-617-0-1.

[10] D. Y. Liliana and T. Basaruddin, "Deteksi Pemalsuan Citra Berbasis Dekomposisi Nilai Singulir," *MAKARA Sci. Ser.*, vol. 13, no. 2, pp. 180–184, 2010, doi: 10.7454/mss.v13i2.422.

[11] S. Dehnie, T. Sencar, and N. Memon, "Digital image forensics for identifying computer generated and digital camera images," in *Proceedings - International Conference on Image Processing, ICIP*, 2006, pp. 2313–2316, doi: 10.1109/ICIP.2006.312849.

[12] A. D. Warbhe and R. V Dharaskar, "An Active Approach based on Independent Component Analysis for Digital Image Forensics." [Online]. Available: www.ijcsit.com.

[13] B. Bayar and M. C. Stamm, "Constrained Convolutional Neural Networks: A New Approach Towards General Purpose Image Manipulation Detection," *IEEE Trans. Inf. Forensics Secur.*, vol. 13, no. 11, pp. 2691–2706, Nov. 2018, doi: 10.1109/TIFS.2018.2825953.

[14] M. Dilshad Ansari, S. P. Ghrera, and V. Tyagi, "Pixel-Based Image Forgery Detection: A Review," *IETE J. Educ.*, vol. 55, no. 1, pp. 40–46, 2014, doi: 10.1080/09747338.2014.921415.

[15] Z. Qu, G. Qiu, and J. Huang, "Detect digital image splicing with visual cues," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2009, vol. 5806 LNCS, pp. 247–261, doi: 10.1007/978-3-642-04431-1_18.

[16] B. Soni and D. Biswas, "Image Forensic using Block-based Copy-move Forgery Detection," 2018, doi: 10.1109/SPIN.2018.8474287.

[17] V. Singh, "Image forgery detection - Using the power of CNN's to detect image manipulation." https://towardsdatascience.com/image-forgery-detection-2ee6f1a65442 (accessed Aug. 14, 2019).

[18] 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, ChinaSIP 2013 - Proceedings*, 2013, pp. 422–426, doi: 10.1109/ChinaSIP.2013.6625374.

[19] N. B. A. Warif, A. W. A. Wahab, M. Y. I. Idris, R. Salleh, and F. Othman, "SIFT-Symmetry: A robust detection method for copy-move forgery with reflection attack," *J. Vis. Commun. Image Represent.*, vol. 46, pp. 219–232, Jul. 2017, doi: 10.1016/j.jvcir.2017.04.004.

[20] N. B. A. Warif *et al.*, "Copy-move forgery detection: Survey, challenges and future directions," *Journal of Network and Computer Applications*, vol. 75. 2016, doi: 10.1016/j.jnca.2016.09.008.