Image Tampering Detection with ELA and Deep Learning

SAADHVI HOSMANE[1], PUNYASHREE M [2], ADITI LADIA [3], ANIRUDHA MALPANI[4]
[1, 2, 3, 4] *Student, Department of Information Science & Engineering, Global Academy of Technology.*

Abstract— Photographs are the foremost powerful and trustworthy media of expression. At present, digital images not only give forged information but also work as agents of secret communication. Users and editing professionals manipulate digital images with various objectives. In fact, images are often considered as evidence of a fact or reality, therefore, fake news or any form of publication that uses images that have been manipulated in such a way as to have the capability and greater potential for misleading. To detect falsification of the image, image data is required in large quantities multiple, and a model that can process every pixel in picture. In addition, efficiency and flexibility in data training is also needed to support its use in everyday life. Deep learning concepts like Convolutional Neural Network (CNN) with Error Level Analysis is the perfect solution for this problem.

Index Terms— Image forgery Detection, Convolutional Neural Network, Error Level Analysis, Deep Learning.

I. INTRODUCTION

In recent times, digital image tampering is easier due to easy access of commercial image editing software, free or paid. For example, these software’s have made it easier to duplicate and manipulate the image's content without (significantly) demeaning its quality or leaving any visible suggestions to an untrained eye (depending on the skills of the user, the software used, etc.).

Image manipulation, often known as image editing, is any type of action performed on digital images using any software. Image forgery is a technique for altering the content of an image to make it contradict a historical truth. Image used by humans to reproduce reality, and is often used as evidence of a news, publication, or fact. There are many types of digital image tampering. These cases can be categorized into three major types, based on the process involved in creating the tampered image. The types are Image Retouching, Image Splicing as well as Copy-Move Attack.

Fake news that has supporting images, tend to be accepted and trusted by the public. To determine an image is genuine or fake, it is very difficult to see with the naked eye, special techniques and certain accuracy are needed in order to know for sure an image is an original image or has been modified.

For ordinary people, this may be difficult to do. For this reason, this image forgery detection technology needs to be developed, so that it can be used as a means to assist people in determining the authenticity of an image. This technology requires a lot of image data, and each image has a lot of constituent pixels. With ordinary machine learning, this technology would be difficult to develop. Thus, Deep learning along with Error level analysis is the right solution for image tampering detection.

II. LITERATURE SURVEY

In digital forensics, the detection of the presence of tampered images are important. The main take through of this literature is that majority of them identify certain features in images tampered by a specific tampering method (such as copy-move, splicing, etc). This implies that the tactic doesn't work reliably across various tampering methods. Additionally, in terms of tampered region localization, most of the work targets only JPEG images because of the exploitation of double compression artifacts left during the re-compression of the manipulated image. However, in reality digital forensics tools mustn't be specific to any image format and can even be ready to localize the region of the image that was modified.

In [1], the authors have proposed a two stage Deep learning approach to seek out features in order to detect tampered images in numerous image formats. For the first stage, they utilized a Stacked Autoencoder model to be told the complex feature for each individual patch. In the second stage, they integrated the contextual information of each patch thus the detection was conducted more accurately. In their experiments, they were able to obtain an overall tampered region localization accuracy of about 91.09% over both TIFF and JPEG images from CASIA dataset, with a fall-out of 4.31% and a precision of 57.67% respectively. The accuracy over the JPEG tampered images was around 87.51%, which outperforms the 40.84% and 79.72% that were obtained from two state of the art tampering detection approaches. The authors in [2] proposed a Deep learning-based approach to detect object-based forgery within the advanced video. The presented deep learning approach uses a convolutional neural network (CNN) to automatically extract high-dimension features from the input image patches. Different from the quality CNN models utilized in computer vision domain, they let video frames undergo three pre-processing layers before being fed into the CNN model. They include a frame absolute difference layer to cut down temporal redundancy between video frames, a max pooling layer to reduce computational complexity of image convolution, and a high-pass filter layer to enhance the residual signal left by video forgery. Additionally, an asymmetric data augmentation strategy has been established to urge a similar number of positive and negative image patches before the training. The experiments have demonstrated that the proposed CNN-based model with the pre-processing layers has achieved excellent results.

A customized convolutional neural network, named CGFace was proposed by the authors in [3]. It was specifically designed for the computer-generated face detection task by customizing the number of convolutional layers, so it performs well in detecting computer-generated face images. Later on, an imbalanced framework (IF-CGFace) is formed by altering CGFace’s layer structure to manage to the imbalanced data issue by extracting features from CGFace layers and use them to teach AdaBoost and eXtreme Gradient Boosting (XGB). Further on, they explained about the tactic of generating an outsized computer-generated dataset supported the state-of-the-art PCGAN and commenced model. Followed by these various experiments were carried out to the means that the proposed model with augmented input yields the absolute best accuracy at 98%. Finally, they provided comparative results by applying the proposed CNN architecture on images generated by another GAN research.

In [4], the authors have proposed image forgery check system supported SURF features, it is most often a pixel-based technique where after pre-processing the photographs, relevant features are extracted and compared with an outlined estimated threshold value. According to the demonstrated results it's decided whether the image has been forged or not and if it’s, then the part where tampering has been done is displayed as a forged part. The proposed algorithm was tested using an open source CASIA image dataset. The presented result shows that SURF feature-based authentication provide forgery detection accuracy of 97%. The result was then compared with other techniques in similar domain to prove the novelty of the work. The author A Kuznetsov in [5] has presented an algorithm for detecting one of the foremost commonly used types of digital image forgeries - splicing. The algorithm is based on the use of the VGG-16 convolutional neural network. Here, image patches are taken as input and obtains results for each patch i.e., original or forgery. During the training stage the author selected patches from original image regions and on the borders of embedded splicing. The obtained results approximately have high classification accuracy such as 97.8% accuracy for fine-tuned model and around 96.4% accuracy for the zero-stage trained for a bunch of images containing artificial distortions in comparison with existing solutions and also the experimental research was conducted using the CASIA dataset.

The authors in [6] proposed an effective and efficient technique for detecting the copy-move forged image supported deep learning. They proposed an algorithm that initializes the tampered image because the input to the system to determine the tampered region. The system includes processes like segmentation, feature extraction, dense depth reconstruction, and eventually identifying the tampered areas. The proposed Deep learning-based system can save on computational time and detect the duplicated regions with more accuracy. The understanding and extensive literature review of state-of-the-art techniques of deep learning within the detection of copy-move image forgery was presented by the authors of [7]. Because of this development of sophistication of tools and software like Adobe

Photoshop, Pixir, and Affinity, digital images content is typically simply manipulated, and thus forged images are produced. Thus, the process authenticating a digital image becomes difficult such as to differentiate between manipulated images and actual images through the naked eyes. And also, the importance of digital image forensics has attracted many researchers who are deeply involved during this area and has established many techniques for forgery detection in image forensics. Lately, Deep learning approach features a high interest among researchers across the sector and has shown good end in its application. Thus, forensic researchers plan to apply deep learning approach as a way for detecting forgery image.

[9] In this paper, the author proposed an innovative image forgery system that has been supported by Discrete Cosine Transformation (DCT) and native Binary Pattern (LBP) and a replacement feature extraction method using the mean operator. First, images are divided into non-overlapping fixed size blocks and 2D block DCT is applied to capture changes because of image forgery. Also, LBP is applied to the magnitude of the DCT array to reinforce forgery artifacts. Finally, the mean of a particular cell across all LBP blocks is computed, which yields a tough and fast number of features and presents a more computationally efficient method. Using Support Vector Machine (SVM), the proposed method has been extensively tested on four documented publicly available Gray scale and color image forgery datasets, and additionally on an IoT based image forgery dataset that was built. Experimental results reveal the prevalence of the proposed method over recent state-of-the-art methods in terms of widely used performance metrics and computational time and demonstrate robustness against low availability of forged training samples. [10] Due to availability of many software's like Photoshop, GIMP, and Coral Draw, it is very hard to differentiate between original image and tampered image. Traditional methods for image forgery detection often use handcrafted features. The matter with the traditional approaches of detection of image tampering is that most of the methods can identify a selected sort of tampering by identifying a particular feature in image. Currently Deep learning methods are used for image tampering detection. These methods reported better accuracy than traditional methods due to their capability of extracting complex features from image. In this paper, the author presents an in-depth survey of deep learning-based techniques for image forgery detection, outcomes of survey in form of analysis and findings, and details of publicly available image forgery datasets.

GoogleNet deep learning model to extract the image features and use Random Forest machine learning algorithm to detect whether the image is forged or not was implemented in [11]. The proposed approach was implemented on the publicly available dataset MICC-F220 with k-fold cross validation approach to separate the dataset into training and testing dataset and compared with the state-of-the-art approaches. In [12] a mask regional convolutional neural network (Mask R-CNN) approach for patch-based inpainting detection was proposed. [13] In recent years, many tampering operations were performed on the image and post-processing is done to erase the traces left behind by the tampering operation, making it more difficult for the detector to detect the tampering. It was found that to detect image manipulation are often supported by Deep learning methods. In this paper, the authors had more focus on the study of various recent image manipulation detection techniques. Authors also examined various image forgeries that can be performed on the image and various image manipulation detection and localization methods. In [14] a Deep learning-based method was proposed to detect image splicing within the images. At the start, the input image is pre-processed employing a technique called 'Noiseprint' to urge the noise residual by suppressing the image content. Then he favoured ResNet-50 network is employed as a feature extractor. Finally, the obtained features are classified as spliced or authentic using the SVM classifier. The experiments performed on the CUISDE dataset show that the proposed method outperforms other existing methods. The proposed method achieves a mean classification accuracy of 97.24%.

[15] In contrast with another recent survey, this paper covers significant developments in passive image forensic analysis methods adopting deep learning techniques. Existing methodologies are studied concerning benefit, limitation, the dataset used, and type of attack considered. The paper further highlights future challenges and open issues, and also provides the

possible future solution in building efficient tampering detection mechanism using deep learning technique. Experiment outcomes show good performance in reference to TPR, FPR, and F1-Score.

III. METHODOLOGY

There are two main methods in this project, namely Error Level Analysis (ELA) and machine learning with deep learning techniques in the form of Convolutional Neural Network (CNN).

A. Error Level Analysis (ELA)

Error Level Analysis is one of the techniques that used to detect image manipulation with how to re-save an image at a certain quality level and calculate the ratio between the compression levels. In general, this technique is performed on images that have a lossy format (lossy compression). Image type used in this data mining is JPEG. On JPEG images, compression is performed independently for every 8x8 pixels in the image. If an image is not manipulated, every 8x8 pixel in the image must have the same error rate.

B. Convolutional Neural Network (CNN)

CNN is a type of network based on feedforward, which the flow of information is only one way, namely from input to data output. Although there are several types of CNN architectures, in general, CNN has some convolutional layers and pooling layers. Then, followed by one or more fully connected layer. In image classification, input on CNN is in the form of images, so each pixel can be processed.

In short, a convolutional layer is used as a feature extractor that studies the representation of these features from images that are input on CNN. Meanwhile, the pooling layer is tasked with reducing spatial resolution of feature maps. Generally, before fully connected layer, there are several convolutional and pooling layer that is used for extract representation for more abstract features. After that, fully connected layer will interpret these features and perform the functions that require high-level reasoning. The classification at the end of CNN will use the function SoftMax.

IV. DESIGN AND IMPLEMENTATION

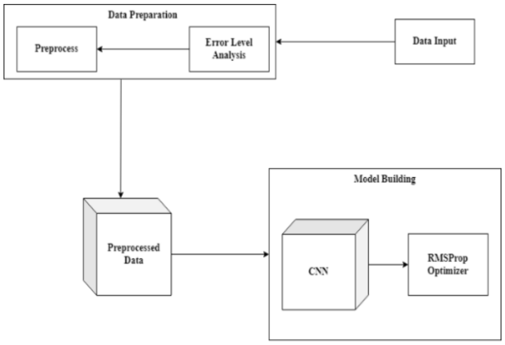
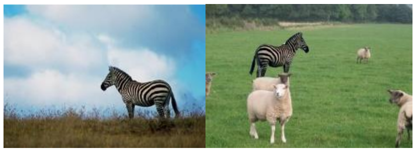


Figure 1: Architecture design.

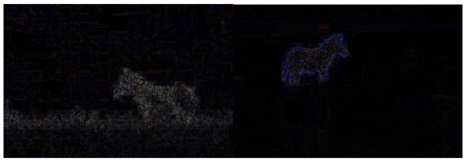
In general, architectural design is divided into two parts large, namely data preparation and model building. At the initial stage, the input data consisting of images with the format ".jpg", with the following details: 1771 images with label tampered and 2940 images labelled real, is passed to the data preparation stage. Data preparation stage is the stage where each image that is inputted is converted into an Error Level Analysis (ELA) result image. Then, the ELA image will be resized into an image with a size of 128 x 128.

The conversion of raw data to the ELA result image is a method used to increase the training efficiency of the CNN model. This efficiency can be achieved because the results of the ELA image contain information that is not as redundant as the original image. The features generated by the ELA image are focused on the part of the image that has a level error above the limit. In addition, the pixels of an ELA image tend to have colours that are similar to or in sharp contrast to the pixels nearby, so training the CNN model is becoming more efficient.



a) b)

Figure 2: a) An example of an original image and b) An example of tampered image.



a) b)

Figure 3: a) ELA converted image of original image and b) ELA coverted image of tampered image.

After that conversion, the image size changes. In the next step, each RGB value is divided by the number 255.0 to normalize so that CNN converges faster (reaching the global minimum of loss values belonging to validation data) because the value of each RGB value only ranges between 0 and 1. The next step is by changing the label of the data, where 1 represents tampered and 0 represents real in to a categorical value. After it was done by dividing training data and validation data using the division of 80% for training data and 20% for validation data. The next step is to use training data and validation.

Next, with the use of training data and validation data the model is trained by deep learning concept CNN. RMSProp optimizer is applied during training for optimization. The below figure shows the complete architecture of CNN model building.

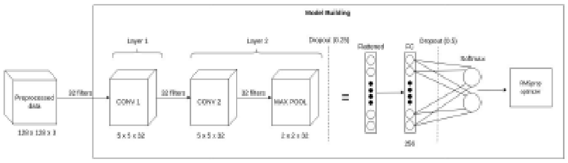


Figure 4: CNN model development architecture.

In the deep learning model used, the first layer CNN consists of convolutional layers with kernel size of 5x5 and the number of filters is 32. The second layer of CNN consists of a convolutional layer with the size of the kernel of 5x5 and the number of filters as many as 32 and a MaxPooling layer with a size of 2x2. The second Convolutional layer is used using kernel initializer and the ReLU activation function to make neurons that are convolutional, the layer selects so that it can receive useful signals from input data.

After that, the MaxPooling layer added a dropout of 0.25 to prevent overfitting. The next layer is a Fully connected layer with the number of neurons as many as 256 and the ReLU activation function. After a fully connected layer, a dropout of 0.5 is added to prevent overfitting. The output layer uses an activation function called softmax.

In the architecture used, only two convolutional layers are needed, because the results are generated from the conversion process to an ELA image can highlight the important features of knowing whether an image is original or has been properly modified.

V. RESULT AND ANALYSIS

The results obtained from the proposed method have maximum accuracy of 99.94 %. The figure below shows the accuracy curve and loss curve.

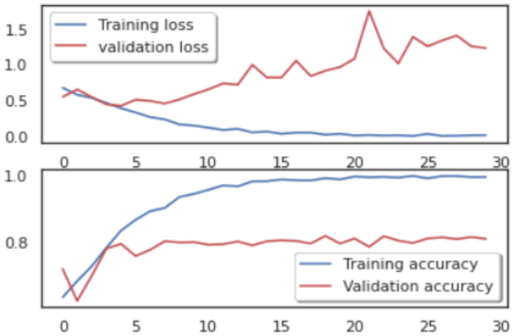


Figure 5: Accuracy curve and Loss curve for training data and validation data.

It can be seen in the picture above that the best accuracy found in the 20th epoch. Value of validation loss after the 5th epoch started to flatten and finally increased, which is a sign of overfitting. With the early stopping, training will be stopped when validation accuracy value starts to decrease, or the validation loss value starts to increase. The number of training epochs required is small for achieve convergence, due to the use of ELA converted image makes model training so much more efficient, and the normalization performed on the RGB values for each pixel also accelerates the convergence of CNN model.

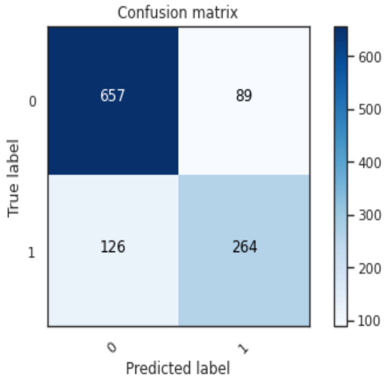


Figure 6: Confusion matrix from validation data(1 represents tampered and 0 represents the original image).

The accuracy results obtained by the model in classifying can be said to be high. This is an indication that the ELA converted image can be used to classify whether the image is the original image or has experienced modification.

CONCLUSION

In most of the research papers, researchers have clarified that image tampering detection may be a very complicated procedure due to the vacuity of different software packages. All features are very sensitive to operations within the interference process. So, features in the image tampering process plays a pivotal part in the process of tamper discovery.

In the last decade, the utilization of convolutional neural networks (CNN) has spread within the image forensic community. These algorithms have focused on training the CNN to see the most effective features to classify camera models.

Here, we have proposed a methodology that uses CNN and ELA for detecting the tampered image. One advantage of using CNN is that the features are extracted directly from the image dataset. The principal advantage of these CNN based approaches is that they are capable of learning classification features directly from image data. It is also found that CNN-based tampering detection methodologies are highly efficient in detecting multiple tampering with high accuracies. The use of ELA can increase efficiency and reduce the computational cost of the training process.

REFERENCES

[1] Ying Zhang, Jonathan Goh, Lei Lei Win & Vrizlynn Thing, “Image Region Forgery Detection: A Deep Learning Approach”, Proceedings of the Singapore Cyber-Security Conference (SG-CRC) 2016, doi:10.3233/978-1-61499-617-0-1.

[2] Ye Yao, Yunqing Shi, Shaowei Weng and Bo Guan, “Deep Learning for Detection of Object-Based Forgery in Advanced Video”, MDPI, Symmetry,26 December 2017, doi:10.3390/sym10010003.

[3] L. Minh Dang, Syed Ibrahim Hassan, Suhyeon Im, Jaecheol Lee, Sujin Lee and Hyeonjoon Moon, “Deep Learning Based Computer-Generated Face Identification using Convolutional Neural Network”, MDPI, Applied Sciences,13 December 2018, doi:10.3390/app8122610.

[4] Payal Srivastava, Manoj Kumar, Vikas Deep and Purushottam Sharma, “A Technique to Detect Copy-Move Forgery using Enhanced SURF”, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volum-8, Issue-6S August 2019, doi: 10.35940/ijeat. F1133.0886S19.

[5] A Kuznetsov, “Digital image forgery detection using deep learning approach”, Journal of Physics: Conference Series, ITNT 2019,doi:10.1088/1742 6596/1368/3/032028.

[6] Ritu Agarwal and Om Prakash Verma, “An efficient copy move forgery detection using deep learning feature extraction and matching algorithm”, Springer Science+Business Media, LLC, part of Springer Nature 2019, 23 December 2019.

[7] Arfa Binti Zainal Abidin, Azurah Binti A Samah, Hairudin Bin Abdul Majid and Haslina Binti Hashim, “Copy-Move Image Forgery Detection Using Deep Learning Methods: A Review”, 978-1-7281-6726-8/19/$31.00 2019 IEEE.

[8] Gul Muzaffer and Guzin Ulutas, “A new deep learning-based method to detection of copy-move forgery in digital images”, 978-1-7281-1013-4/19/$31.00 2019 IEEE.

[9] Mohammad Manzurul Islam, Gour Karmakar, Joarder Kamruzzaman and Manzur Murshed, “A

Robust Forgery Detection Method for Copy–Move and Splicing Attacks in Images", MDPI, electronics,12 September 2020, doi:10.3390/electronics9091500.

[10] Zankhana J. Barad and Mukesh M. Goswami, "Image Forgery Detection using Deep Learning: A Survey", 2020 6th International Conference on Advanced Computing & Communication Systems (ICACCS), 978-1-7281-5197-7/20/$31.00 2020 IEEE.

[11] Amit Doegar, Maitreyee Dutta and Gaurav Kumar, "Image Forgery Detection Using Google Net and Random Forest Machine Learning Algorithm", Journal of University of Shanghai for Science and Technology, Volume 22, Issue 12, December – 2020, doi - 10.51201/12508.

[12] Xinyi Wang, He Wang and Shaozhang Niu, "An Intelligent Forensics Approach for Detecting Patch-Based Image Inpainting", Hindawi, Mathematical Problems in Engineering, Volume 2020, Article ID 8892989, 10 pages, 28 October 2020.

[13] Rahul Thakur and Rajesh Rohilla, "Recent Advances in Digital Image Manipulation Detection Techniques: A brief Review", Forensic Science International, 24 April 2020, Published by Elsevier.

[14] Kunj Bihari Meena and Vipin Tyagi, "A Deep Learning based Method for Image Splicing Detection", Journal of Physics: Conference Series, CONSILIO 2020.

[15] Manjunatha S and Malini M Patil, "Deep learning-based Technique for Image Tamper Detection", 2021 Third International Conference on Intelligent Communication Technologies & Virtual Mobile Networks (ICICV), 2021 IEEE.

[16] Marra, Francesco & Gragnaniello, Diego & Verdoliva, Luisa & Poggi, Giovanni. A Full-Image Full-Resolution End-to-End-Trainable CNN Framework for Image Forgery Detect ion, 2019.

[17] Y. Rao and J. Ni, "A deep learning approach to detection of splicing and copy-move forgeries in images," 2016 Workshop on Information Forensics and Security (WIFS), Abu Dhabi.

[18] J. Fridrich, B. D. Soukal, and A. J. Luks, Detect ion of copy-move forgery in digital images, in Proceedings of Digital Forensic Research Workshop, Citeseer 2003.

[19] R. Dixit, R. Naskar, and A. Sahoo. Copy-move forgery detection exploiting statistical image features, 2017 International Conference on Wireless Communications, Signal Processing, and Networking (WiSPNET), Chennai, 2017,

[20] Belhassen Bayar, Matthew C. Stamm. A Deep Learning Approach To Universal Image Manipulation Detection Using A New Convolutional Layer. ACM. ISBN 978-1-4503-4290-2/16/06.