Forged Face Detection using ELA and Deep Learning Techniques

Qurat-ul-ain,
Dept. of computer Science
University of Engineering and Technology
Taxila, Pakistan
quratul.ain2@students.uettaxila.edu.pk

Nudrat Nida
Department Of Computer Science,
Air University Islamabad, Aerospace & Avaition Campus, Kamra, Pakistan
Nudrat.Nida@aack.au.edu.pk

Aun Irtaza,
Dept. of computer Science
University of Engineering and Technology
Taxila, Pakistan
aun.irtaza@uettaxila.edu.pk

Nouman Ilyas
Dept. of computer Science
University of Engineering and Technology
Taxila, Pakistan
17-CS-47@students.uettaxila.edu.pk

2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST) | 978-1-6654-0516-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/IBCAST51254.2021.9393234

***Abstract*— The paper presents real and fake facial image recognition using Deep learning and CNN. Image forgery recognition becomes a difficult task to find the authenticity of an image with the naked eye. This research aims to evaluate the working of different deep learning techniques in the novel "Real and Fake Face detection" dataset by Computational Intelligence Photography Lab, Yonsei University. For the detection of forged faces, the first step of the proposed method is image normalization for real and fake image recognition. Normalized images are then preprocessed using Error Level Analysis (ELA) and train to different pre-trained deep learning models. We finetune these models for categorization of 2 classes that are forged and real to evaluate these models' performance. From all tested models, VGG models give the best training accuracy of 91.97% and 92.09% on VGG-16 and VGG-19, whereas VGG-16 shows the good test set Accuracy using a smaller number of epochs, which is competitively better than all other techniques. Results of these models were evaluated using confusion matrix evaluation measures and compared with state of the art techniques.**

Keywords*— *Transfer learning, augmentation, CNN, ELA, Forgery, VGG, ResNet-50, InceptionV3.

I. Introduction

Image forensic is a way to verify the authenticity of an image that either the image is real or altered. It became difficult to recognize the truth behind each image when images are in thousands of numbers. In image processing and computer vision, many feature extraction techniques work well by extracting features from related images and training them with deep learning models.

Deep learning is a machine learning technique that uses the gained knowledge for solving related problems. Some most typical approaches of deep learning are pre-trained models. These models are trained on large-scale datasets, which makes them intelligent to improve the model's efficiency. However, in this research, we use these pre-trained models for the "Real and Fake Face detection" dataset, which has comparatively fewer images. Deep learning models are pre-trained on millions of images to classify thousands of different categories; if we talk about real and forged image classification, then we have 2 categories to classify. Deep learning also provides a feature to reuse and finetune the pre-trained model by refining it. By finetuning the pre-trained model, we change its output layers with CNN layers according to classification required by available dataset categories.

Several methods are being used to recognize the image authenticity using machine learning techniques. Convolutional neural network is very productive deep learning approach which classify the data available for the task of interest. CNN's are versions of dense networks in which fully connected layers are used to forward information to next layer.

In past many feature extraction methods such as LBP(local binary pattern) and HOG(histogram of gradient) were used to find the originality of image, but we normalize images by resizing and then process them with ELA(Error level analysis). ELA compress the images of both real and forged images at certain level of quality and then take their difference with original images, then pass these images to deep learning models. ELA extract features accurately and brighten the digitally modified portion of images. All these techniques are discussed in coming sections.

II. Literature Review

Sudiatmika [1] proposed an image forgery classification technique using Error Level Analysis to determine the compression ratio between the original image and the fake image. When images were compressed, the original image and the fake images were different. Results of performed experiment showed training accuracy but no test accuracy. Meera [2] used CASIA TIDE v.1 Dataset for forgery detection using Gabor Wavelets and Local Phase Quantization. Hakimi [3] proposed a method in which they used chromatic components to enhance image forgery detection. The performance of Cb is almost like that of Cr on CASIA v2.0 dataset. The results are comparable to each other, which shows the consistency of their proposed research. Jing [4] used self-evaluation and the purpose of creating CASIA Dataset for image tampering recognition. Li [5] used Forgery detection using the block artifact grid technique on images downloaded directly from NASA.

Muhammad [6] proposed method Passive copy move image forgery detection in which they analyzed image forgery using dyadic wavelet transform. Birajfar [7] used passive technique for fake image detection. Akhtar [8] suggested an effective method for handling the Digital Image Forgery Detection problem. LBP and HOG was used for feature extraction. Muhammad [9] proposed a new method for detecting forgery based on ELA; the method was evaluated on CASIAv2.0

978-0-7381-0535-2/21/$31.00 ©2021 IEEE

publicly available datasets. These experiments showed different accuracies using different learning rates. Gypsy Nandi [10] discussed some of the most effective forgery detection techniques that helped forgeries identify images, either single or composite. Mankar [11] compared different classification approaches to find image forgery, consisting mainly of active and passive approaches. Mahale, V. H [12] presented a unique approach based on Local Binary Pattern (LBP) for evaluating image inconsistency and detection, the system was tested on COMOFOD dataset. Mohamad [13] used an efficient method of combined un-decimated wavelet transform from which scale-invariant feature transform were judged from evaluation measures. Wu-Chi [14] summarized some forgery methods based on watermarking and alpha mattes to analyze images.

Many state-of-art techniques worked on image detection and recognition using different feature extraction techniques and CNN's. This research presents forgery recognition on a novel "Real and Fake Face detection" dataset using deep learning models and convolutional neural networks. We also compared our results of each deep learning model with other and with state-of-art techniques.

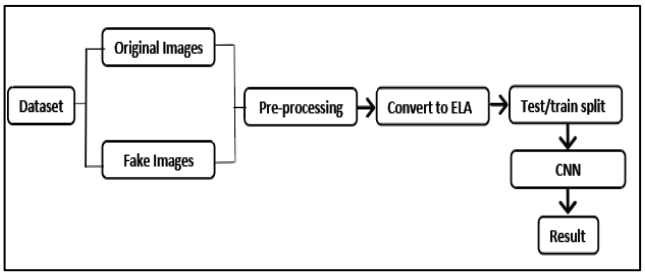
III. Proposed Method

Our dataset consists of RGB images with an extension of (.jpg) which were classified into real and fake classes. After preprocessing the images of the whole dataset, it was converted into ELA images. These images were split up into training and test sets, which were then forwarded to the Deep CNN model to recognize real and fake images. The proposed method is shown in Fig.1.

A. Pre-processing

The first step of the proposed approach was preprocessing in which the whole dataset was resized into (128*128) pixels, which make the whole dataset normalized.

Fig. 1. Proposed method



B. Error Level Analysis

For feature extraction, we used the Error Level Analysis (ELA). This technique identified that the images were either real or manipulated by restoring images at a certain quality level. This technique works well on images with lossy compression. Suppose we have the original image (Io), and the same image was resaved at 90% compression level, so the resaved image is (Ir). Based on preprocessed and resaved image ELA of image measured the difference between them, as in (1).

$$ELA = Io - Ir \quad (1)$$

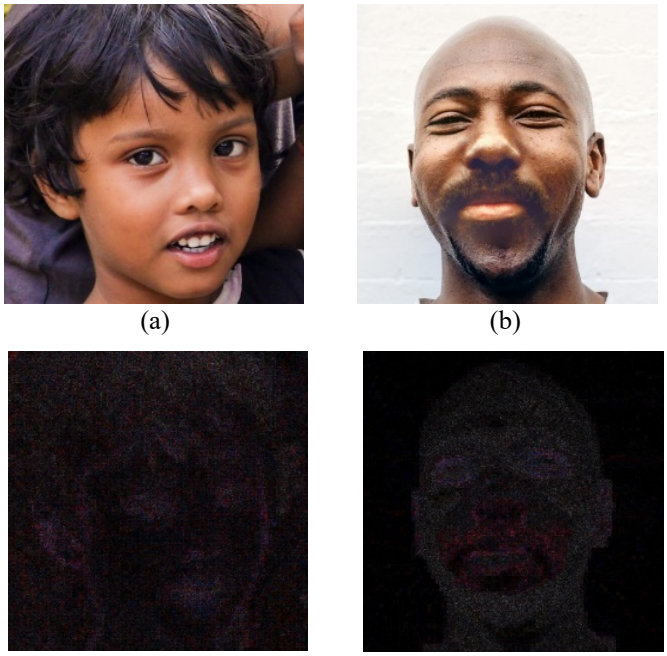
1) Resaved images: Dataset contained (JPG) lossy images. To convert image into ELA first the preprocessed images were resaved at specific level of quality. As we resaved both real and forged preprocessed images into 90% compression level with .jpg extension. This process make the modified part of image white or brighter. For images resaving following steps are included:

- Take both real and forged preprocessed images.
- Resave images at 90% level of compression.

2) Taking difference: In this step images were compared preprocessed and resaved to investigate the absolute difference which include the following steps:

- Choose preprocessed images.
- Choose resaved images.
- Take difference of preprocessed and resaved images.

Fig. 2.Example of some images from dataset(a) Real image,(b) Fake image, (c) ELA of Real image, (d) ELA of Fake image



The absolute difference of images was taken by comparison using ELA processing. It takes the difference of each block, 8x8 pixels of both images. In Fig.2. the image (a) and (b) shows real and forged the images from the dataset, whereas the image (c) and (d) are ELA generated images. The forged ELA-generated image shows modified parts of the image brighter the corresponding original ones.

C. Image Recognition using Deep Learning

Once preprocessed images were converted into ELA, all images were saved in NumPy array and split into train set and test set and then the train set was passed to the finetuned pre-trained deep learning models. 4 pre-trained models were used to evaluate our dataset results, which are VGG-16, ResNet-50, InceptionV3 and VGG-19. These models were already pre-trained. We changed these models' input layers as per the size of pre-trained image size (128*128) and the last layers were changed according to the number of dataset classes. We modified these 4 models by adding sequential model layers as output layers.

1)Training parameters for deep learning models: We used the object recognition models, which were already pre-trained on the many image datasets. Using deep learning, we finetune these models for "Fake and Real face detection" dataset for image recognition. RMSprop was used as the optimization function with 0.0001 learning rate and batch size 5 per epoch.

D.Convolutional neural network

Using Keras libraries, we used the flatten layer and dense layer as a sequential CNN model. After images were trained from a finetuned preprocessed model, they were passed to the flatten layer, which transformed the feature map by flattening them into the feature vector and forwarding them to a fully connected layer.

The fully connected layer was used for pattern recognition in the last dense layer of models, SoftMax activation function was used to convert the feature vector in a probabilistic manner. Based on SoftMax activation, the training set was compared with the test set and return a probability distribution on real and forged images. These layers were used to finetune all the deep learning models used in this research.

IV. Experiments and Results

A. Dataset

"Real and Fake Face detection" by Computational Intelligence Photography Lab, Yonsei University. This novel dataset consists of 2041 facial images that were further divided into a fake (960) and real (1081). Sample images from both types are shown in Fig.3.

B. Experiments

This section presents the experiments performed using different deep learning pre-trained techniques on "Real and Fake Face detection" dataset.

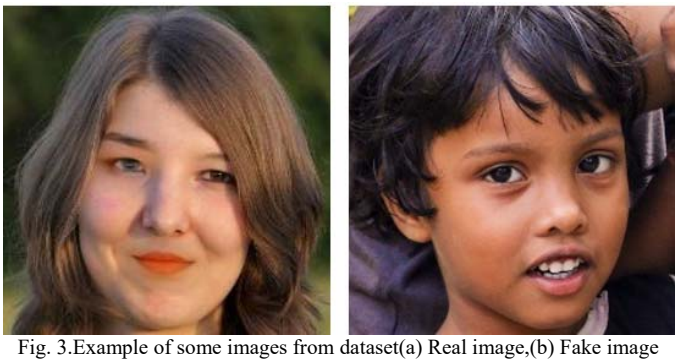


Fig. 3.Example of some images from dataset(a) Real image,(b) Fake image

(a) (b)

1)VGG-16: In the first experiment, ELA converted training set images were passed to finetuned VGG-16, the simplest model consisted of 22 layers, we changed its input layer according to our dataset images size of 128 and replaced its last two layers with sequential CNN layers, then used early stopping to prevent the model from over-fitting. The confusion matrix of VGG-16 model is shown in Fig.5.

Training parameters were used for the evaluation of results produced by the model. This technique gave Accuracy of 91.97% training and 64.49% test using 9 epochs. Model truly predict 267 images from 414 test set images, from which 116 were fake and 151 real images.

2)Inception V3: In the second experiment, the same parameters were used for finetuning InceptionV3. This finetuned model worked by factorized convolution and batch normalization shows Accuracy of 57.13% training and 57.25% test using 15 epochs after early stopping, which was not as good as produced by VGG-16. From 414 test images, this model predicted 237 images, 69 fake and 169 real, and other images were predicted false.

3)ResNet-50*:* This pre-trained model worked on the skip connection technique was finetuned using same parameters as previous ones. This finetuned model shows Accuracy of 53.74% on the training set and 53.62% on test set using 15 epochs after early stopping. This model gives less accurate results than the previous two models although parameters passed to this model were used in other models. This model predicts only 7 real and 168 fake images from a set of 414 images.

4)ResNet: The last experiment VGG-19 model with 24 layers was used to check the Accuracy of the model on a dataset. We finetuned this model with the same parameters. Model truly predicted 125 real and 126 fake images out of 414 images. This experiment model shows Accuracy of 92.09% training and 60.63% test using 8 epoch. The training results were comparatively the same as the VGG-16 model used in the first experiment.

Fig . 4. Train and Validation Accuracy

C.Results and Discussions

The results of experiments and discussion were stated and compared in this section.

1)Forgery recognition: Dataset used in this experiment contain 2041 images for deep analysis. In preprocessing, images were resized and then featured were extracted with ELA. ELA was good at finding handcrafted features and make recognition more accurate. Dataset was divided into train and test sets, from which train set was further divided into the train set and validation set. On the basis of parameters learned by our models during training process, test images were examined. Our experiments proved that VGG-16 showed the best test accuracy of 64.49% on the novel "Real and Fake Detection" dataset. The traning and validation accuracy of the best model is shown in Fig.4.

The confusion matrix of the test set of VGG-16 is shown in Fig.5, in which the upper left square shows true negative samples, which are predicted true by the model and belong to the fake class. In contrast, the lower left is false negative samples that belonged to a fake class, but the model predicted them real. The lower right square shows accurate positive samples that are predicted true by the model and belong to the real class. In contrast, the upper right is false positives that negative samples belong to real class, but the model predicted them fake.

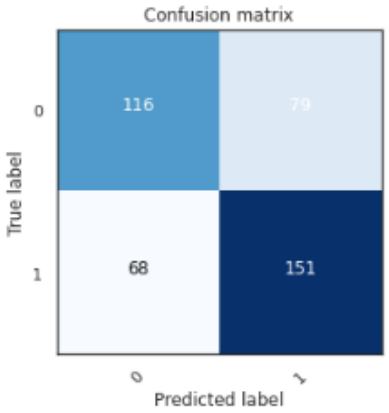


Fig . 5. Confusion matrix of VGG-16 Test

2)Evaluation measures: The evaluation parameters were used to measure the working of proposed techniques were Precision, recall, F-1 Score, Accuracy, and error rate metrics are used to measure the performance.

$$\text{Precision} = \frac{TP}{FP+TP} \quad (1)$$

$$\text{Recall} = \frac{TP}{FN+TP} \quad (2)$$

$$\text{F1-score} = \frac{2*Re*Pr}{Re+Pr} \quad (3)$$

$$\text{Accuracy} = \frac{TN+TP}{\text{Total samples}} \quad (4)$$

$$\text{Error} = \frac{FP+FN}{\text{Total samples}} \quad (5)$$

TP denotes true positive samples, TN are true negative samples, FP are false positive samples, FN as false negative samples, and the sum of all these represent total samples. The test results obtained from the proposed method are presented in Table 1. The results can be observed that VGG-16 provides better classification results on the used dataset for real and forged image recognition.

Measures / Models	Precision	Recall	Accuracy	F1-score	Error
VGG-16	0.65	0.68	0.64	0.66	0.35
ResNet-50	0.57	0.76	0.57	0.65	0.42
InceptionV3	0.53	0.98	0.53	0.68	0.46
VGG-19	0.64	0.57	0.60	0.60	0.39

TABLE I. Evaluation measures of Models

Various evaluation criteria were used to determine the Accuracy of the experimental results. In this paper, the finetuned deep learning models' performance metrics were evaluated using these measures on the basis of the confusion matrix as shown in Figure 6. The test accuracies and error rates of all models are shown in Figure 7.

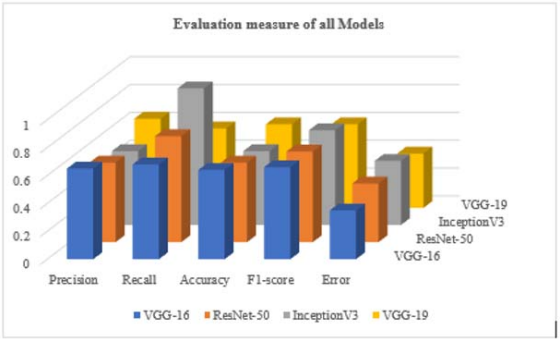


Fig.6. Evaluation measures of different models

D. Comparative analysis

Methods	Dataset	Technique	No. of Epochs	Traning Acc	Test Acc
[1]	CASIA V.2	ELA and CNN	100	60-90%	-
[9]	CASIA V.2	ELA and CNN(Lr=0.01)	500	84%	60%
		ELA and CNN(Lr=0.05)	500	84%	62%
[5]	NASA images	BAG	-	-	47.21%
[4]	CAISA Tide	CNN	-	90%	58.7%
Proposed method	Real and Fake Face detection	ResNet- 50	15	53.74%	53.64%
		InceptionV3	15	57.13%	57.25%
		VGG-19	15	**92.09%**	60.63%
		VGG-16	9	**91.97%**	**64.49%**

TABLE II. Comparative Analysis

*1)Comparisons with state-of-the-art techniques:*We compared the performance of our experiment results with the state-of-the-art method. [1] uesd ELA and CNN on CASIA v.2 dataset using 100 epaochs and there model graph shows traning accuracy between 60-90 % but test set accuracy was not mentioned. [9] used ELA for feature extraction for image recognition and perform recognition using different learning rates. [4] used CASIA TIDE Dataset for forgery recognition using CNN, [5]used image recognition using images from NASA. Comparison of our best technique result on "Real and Fake Detection" dataset are shown in Table 2, on minimal dataset our finetuned model perform well using less number of epochs.

V. CONCLUSION

This paper compares different state-of-the-art deep learning pre-trained models to distinguish between real and forged images. For feature extraction, we used ELA to recognize and verify images' authenticity by utilizing original and resaved images from the dataset "Real and Fake Face detection" by Computational Intelligence Photography Lab, Yonsei University. We found that both VGG models 16 and 19 give good training accuracy of 91.97% and 92.09%, while VGG-16 gives 64.49% test set accuracy, which is comparatively better than other pretrained models on the same datasets. In our future research, we plan to generalize our work. We will devise a CNN architecture to get the best Accuracy on multiple datasets by using various feature extraction approaches to recognize the original image and forged image.

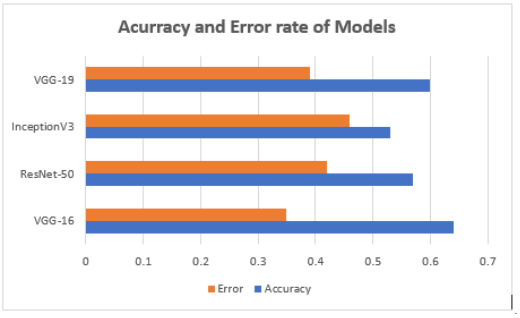


Fig. 7. Accuracies and errors of different models

V. REFERENCES

[1] Sudiatmika, I. B. K., & Rahman, F. (2019). Image forgery detection using error level analysis and deep learning. *Telkomnika*, *17*(2), 653-659.

[2] Isaac, M. M., & Wilscy, M. (2015). Image forgery detection based on Gabor wavelets and local phase quantization. *Procedia Computer Science*, *58*, 76-83.

[3] Hakimi, F., Zanjan, I., & Hariri, I. (2015). Image-splicing forgery detection based on improved lbp and k-nearest neighbors algorithm. *Electron Inf Plan*, *3*.

[4] Dong, J., Wang, W., & Tan, T. (2013, July). Casia image tampering detection evaluation database. In *2013 IEEE China Summit and International Conference on Signal and Information Processing* (pp. 422-426). IEEE..

[5] Li, W., Yuan, Y., & Yu, N. (2009). Passive detection of doctored JPEG image via block artifact grid extraction. *Signal Processing*, *89*(9), 1821-1829..

[6] Muhammad, G., Hussain, M., & Bebis, G. (2012). Passive copy move image forgery detection using undecimated dyadic wavelet transform. *Digital investigation*, *9*(1), 49-57.

[7] Isaac, M. M., & Wilscy, M. (2015). Image forgery detection based on Gabor wavelets and local phase quantization. *Procedia Computer Science*, *58*, 76-83.

[8] Akhtar, F., & Qayyum, H. (2018). Two Fold Image Forgery Detection System Using Combined Key point based method and Block based method. *Journal of Information Communication Technologies and Robotic Applications*, 62-70.

[9] Villan, M. A., Kuruvilla, A., Paul, J., & Elias, E. P. (2017). Fake Image Detection Using Machine Learning. *IRACST—International Journal of Computer Science and Information Technology & Security (IJCSITS)*.

[10] Sarma, B., & Nandi, G. (2014). A Study on Digital Image Forgery Detection. *International Journal of Advanced Research in Computer Science and Software Engineering*, *4*(11).

[11] Mankar, S. K., & Gurjar, A. A. (2015). Image forgery types and their detection: A review. *International Journal of Advanced Research in Computer Science and Software Engineering*, *5*(4), 174-178.

[12] Mahale, V. H., Ali, M. M., Yannawar, P. L., & Gaikwad, A. T. (2017). Image inconsistency detection using local binary pattern (LBP). *Procedia computer science*, *115*, 501-508.

[13] Hashmi, M. F., Anand, V., & Keskar, A. G. (2014). Copy-move image forgery detection using an efficient and robust method combining un-decimated wavelet transform and scale invariant feature transform. *Aasri Procedia*, *9*, 84-91.

[14] Hu, W. C., Chen, W. H., Huang, D. Y., & Yang, C. Y. (2016). Effective image forgery detection of tampered foreground or background image based on image watermarking and alpha mattes. *Multimedia Tools and Applications*, *75*(6), 3495-3516.