JPEG Tamper Detection Using Error Level Analysis and Hybrid Transfer Learning

[1]Monish K, [2]JayaShankar G S, [3]Rithik G, [4]Amith Kumar R, [5]Dr. N Sreenivasa

[1, 2, 3,4Student, [5]Associate Professor
Department of Computer Science, Nitte Meenakshi Institute of Technology, Bangalore, India

Abstract
Image manipulation is becoming a serious concern in many fields nowadays as the available software to manipulate images is increasing, at the same time it is becoming increasingly hard to authenticate between original and duplicate images. To address this problem, we discuss a few of the famous transfer learning architectures in image classification and how it is used to authenticate between original and manipulated images in the JPEG format, with the use of lossy double compression for preprocessing. The classification is further improved by the new Deep learning architecture, a similar combination of AlexNet and InceptionNet. The paper mainly focuses on detecting passive image tampering with the help of the CASIA V2 dataset. And have provided good results both on test and validation data.

Keywords: Image Classification, JPEG, Error Level Analysis (ELA), Lossy compression, Transfer learning, LeNet-5, AlexNet, VGG, Inception block, CNN.

I. INTRODUCTION

Digital media nowadays which includes images has become an important medium of communication because of its ability to express, acquire, distribute, and store easily. The significance of digital images in describing data has made them preferable to text information as a means of transmission. Image forgery implies the manipulation of the digital image to camouflage some significant or valuable data about the image. [1] JPEG can play a significant role in the field of digital image forensics by understanding inherited features of the JPEG format image. Advanced computers and image editing software have helped to manipulate images very easily. So, with the increase in tampering, several techniques and models are being developed for image tampering detection.A convolution neural network (CNN) is an artificial neural network that is used to recognize images and process the pixel data of the image. The program can be used to detect tampered images by making use of the convolution neural network. It can be used by the users to verify whether a forgery is made or not on a particular image. Image tampering detection has had the latest trend in recent times with an increase in the forgery of images. Many different algorithms were proposed and much Architecture were proposed for image classifications some of which were referred to as LeNet, AlexNet, VGG, and Inception network.

II. BACKGROUND

[2] Deep learning-based convolution neural network (CNN) with error level analysis of JPEG is one such that is proposed in this project. There are various types of file formats; some formats can be lossless while few are lossy. If lossy formats are used then there is a chance of data loss. Adjusting the quality level happens in JPEG formats, by removing some colors it can be compressed. That is the reason why the color of the image changes whenever saved in an image format, like JPEG. It is hard to identify the damaged part of the image by human vision. The Error Level Analysis is performed on the JPEG images to find the modified sections of the images. The JPEG image is inserted onto a different JPEG image that is of high quality, it checks whether a part of the image is of lower quality compared to other parts of the same image. Error level analysis analyzes compression artifacts in lossy compressed data, like JPEG image format. Normally, compression artifacts stay at a stable level in one image. Therefore, if certain parts of the image undergo deformation or another kind of lossy compression, the data in those regions may occur differently from other data in the image. If all of the data are at the same rate then the image is not modified. If the part of the image data value is different from other parts then the image is altered. The error level analysis (ELA) technique is for passive authentication in image forensics which involves copy-move image forgery, JPEG compression, and image retouching. [3]The accuracy of the Lossy Image compression is analyzed. The investigation is done via several techniques which are Target Registration Error (TRE), Mean Square Error (MSE), and Mutual Information (MS). The outcome of this investigation exhibited MS performs well in lossless compression, but does not perform well in lossy compression.

III. LITERATURE SURVEY

As we know, there were many image classification architectures, and LeNet-5 a CNN architecture was the first of the architectures to become famous, developed by LeCun et al. (1998), for the recognition of handwritten digits. The LeNet model was developed especially to determine handwritten symbols.[1]The classic LeNet-5 model is enhanced by removing the 1*1 convolutional layer and adding the moving average model. [4]The base LeNet-5 CNN architecture was modified by changing the neurons in each layer of a CNN and also modifying the way that connects across various layers. The CNN outputs are set to error-correcting codes, therefore CNN can deny recognition results. To train the CNN, a reinforcement learning strategy with

error-samples-based is developed, and they choose the model which gives good robustness and better performance. For traditional Competitive Learning, outputs for LeNet are placed with place code. By using these place codes it's not able to reject the recognition results. LeNet-5 could reject illegal samples in the printed character recognition. Compared with other methods, CNN has provided an encouraging solution for offline HECR.[5]Mentioned problems that arise while classifying images in a cloud computing environment to achieve good stability of image classification and good effectiveness and robustness and improved AlexNet architecture was proposed and designed. To obtain stability of algorithm structure, the convolutional nerve is introduced and a brand new image training model is designed in combination with the AlexNet network model. The AlexNet network model firstly simplifies the image processing and describes the image features in a simple geometric form, and then by introducing the convolution nerve, the image obtained from the front layer of the convolution nucleus is trained by convolution training. The improved AlexNet network model is implemented which reduces the training time for the image classification process. [6]Transfer learning with AlexNet Convolutional Neural Network (AlexNet CNN) for the recognition established on human ear images. Transfer learning is an effective way to solve classification problems that will contain a very small amount of details.The Rectified linear unit(ReLU) can be added to enhance the non-linearity in the problem-solving capacity of the network. AlexNet is tuned to determine 10 classes only using 250 training images and 50 testing images, and the model performed 100% accuracy.

[7]A modified VGG-16 Network was designed which was able to automatically classify the three types of corneal ulcers. Data preprocessing steps normalization, masking, and data augmentation to find ulcer images were done before changing into the VGG-16 model. For the loss function of the training model, the weighted categorical cross-entropy can be added, and experimentations show that a modified VGG-16 network will have fewer parameters and promising performance than the classical CNN network.Compared with the classic VGG-16 network, the altered VGG-16 network with CCM layers addition of GAP, and added feature fusion layers. The experimentations show that the altered VGG-16 network achieved better than the AlexNet and VGG-16 networks. [8]Evaluates the VGG-19 and VGG-16 architecture and identify four different classes of dementia by adding a densely connected layer to the end of the network. The classic architecture is changed by the addition of a densely connected layer at the rear of the earlier un-trained network. With the change in the loss function, categorical cross-entropy can be used because of its capability to categorize into multi classes. Using VGG-16 and VGG-19 architecture 4 classes of Deep Convolutional Neural Network methods used to predict and classify dementia have been improved by adding a fully connected layer at the conclusion. [9]Implementation of a neural style transfer model consisting of VGG-19 and AlexNet architectures. Neural style transfer is used to generate an output hybrid image which is a mixture of both styled and the original image. By using VGG-19 and AlexNet the neural style transfer model is executed, and the major aim is to generate an output-styled image with high accuracy. In both models, 1000 iterations were performed. In VGG-19, Relu is used for the activation function, which helped to produce more accurate output-styled images than Alexnet's Relu-activated architecture. After comparing both architectures for 1000 iterations, the quality of the output image for VGG-19 is high than Alexnet architecture.

[10]A novel image is an image forgery detection scheme that will detect the copy-move by making use of the DenseInceptionNet. DenseInceptionNet is a multi-dimensional densely featured connection, DeepNeuralNetwork. Regardless, Inception Net and VGG-16 are the classical feed-forward network in which each layer will receive a previous layer's state and writes over the current layer. By keeping high detection performance the Dense-InceptionNet model can aim to get better efficiency. [11]Here architecture is designed by combining the Inception-like blocks into DenseNet, which is called Inception-DenseNet architecture. A new activation called hybrid activation was introduced, which is different from previous inception blocks. Visualization experimentations showed hybrid activation modes are giving more flexible responses to object semantic regions. The experimentation results suggested that InceptionDenseNet can get the same or more acceptable classification results while using a smaller number of trainable parameters.

[12] To detect the tampered regions in a JPEG format image a convolutional neural network-based solution was developed. Here, DCTcoefficients will be input to CNN. The output will be in the form of a binary-segmented image that contains white and black pixels that depict tampered and original regions. While Corresponding with the previous MDBD approach, in which features value is established through know-how, a new approach is given which optimizes by using the CNN and reaches a more increased detection accurateness. [13]Convolutional Neural Networks establish optimistic results when detecting faked images that emerge from the exact type of manipulation they are trained on. It is determined that not all techniques fetch good accuracy for all kinds of image tampering such as splicing, compression, etc. It is required to develop an effective deep learning-based architecture for noticing manipulations efficiently. It is also important to design architecture based on a characteristic extraction mechanism that comprehends correlation better among pixels. The conventional approach can classify a particular type of manipulation by identifying a definite feature in that image. In image formatting, various techniques like copy-move, splicing, etc., are the most known manipulation techniques that are found. Photocopying a part of the image, giving any distortion into this fragment, and infiltrating the altered piece into another area of the same image. [14]The image in the format of JPEG is a famous one. Because its higher compression rate is higher than that compared with the other formats. A novel method with DCT-Coefficient is used to find the tampered images. This method will decompose an input image by saving the image again with other JPEG formats. The tampered regions can be found by making use of the contrast image acquired by an input image and the image saved again in the DCT domain. The significance of the presented technique was proved by testing results.

[15]The copy-paste effect on JPEG images can be found using these methods. The technique to detect will be executed by removing and examining blocking-artifact grids, presented through block-processing during compression of JPEG. The investigation is based upon the actuality that BAGs generally do not match after achieving copy-paste processes. The method is

displayed on two faked images. This process allows adequate finding of impersonated regions on faked images. [16]With the help of singular value decomposition, the hash generation method is used to detect image tampering. To detect and localize image tampering an efficient hash vector need to be designed. This technique is powerful against content preserving but too liable to even minute structural tampering. [17] Few details regarding image, image forensics, image features, and the way they're stored in devices, can be deleted and recovered can be observed. also emphasized a few components in forensic digital images, digital forensics, and some benefits for digital forensic pictures. [18]To detect the presence of non-aligned double JPEG compression (NA-JPEG) in compressed images a simple dedicated algorithm is developed. This model is based on the integer periodicity of the blockwise discrete cosine transform (DCT) coefficients which considers a single feature when the DCT is calculated according to the grid of the last JPEG compression. [19] The model with regular effects in dual quantize is diagnosed, and a possibility of quantization of DCT coefficients in every block is estimated on the whole image. The considerable rear chance of individual blocks is calculated pertaining to Bayesian theory and of the effects cited in the initial position. Afterward, the variance and mean rear possibility are utilized for evaluating if the target block has meddled. Exploratory results indicate that the process can precisely indicate the altered part, and via experimentation, it is also found that for detecting the meddled parts, the better the two contraction grades are the more different the detected efficiency.

IV. APPROACH OVERVIEW

The whole idea for JPEG image tamper detection rests on the Error level obtained from double compression of the same image. The image is compressed to 80% of its original value. As JPEG is a lossy compression, recompression of the image will degrade the intensity of the RGB grid(pixel) value. Using this, the original image pixel values are subtracted, then converted back to an image of the difference in the value, this resultant image is referred to as an Error level image. See. Figures 1 & 2, for working.

```
function ELA_Image(Image):
        image = open(Image).convert('RGB')
        image.save('resave_path.JPEG', quality=80)
        resaved = open('resave_path.JPEG')
        diff = Numpy.array(image) - Numpy.array(resaved)
        diff.save('ELA.JPEG')
```

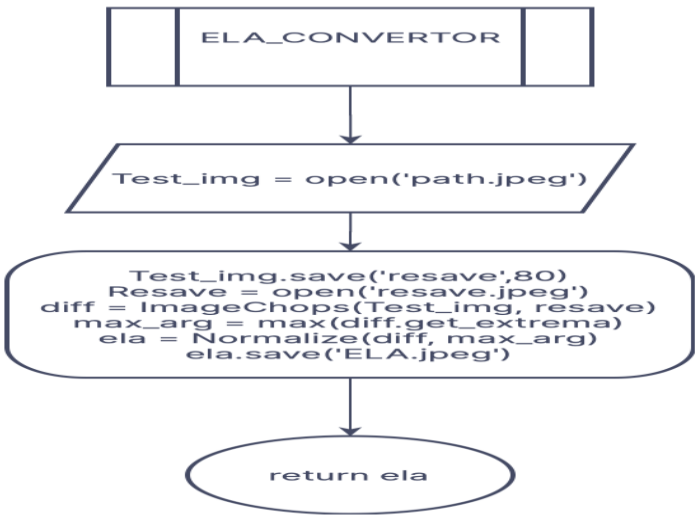

Figure 1. Flowchart of Error level analysis.

This ELA image is converted into a labeled dataset and fed to the deep neural network for training

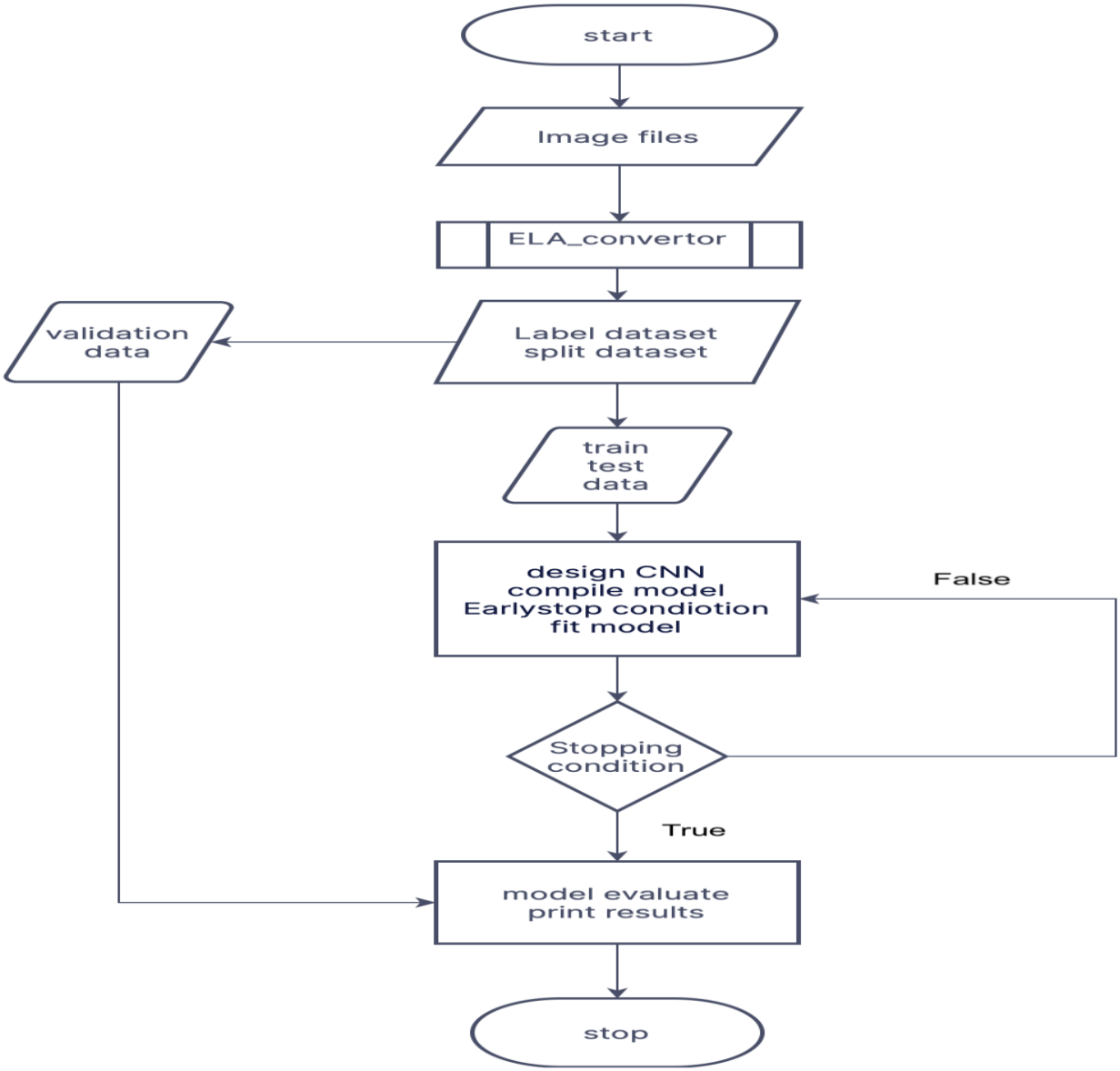

Figure 2. Flowchart of model training

V. COMPARATIVE STUDY OF TRANSFER LEARNING MODELS

In the field of CNN, especially in Image classification, Deep Learning models such as LeNet-5, AlexNet, VGG-16, Inception Block(single block in GoogleNet), ResNet, etc, have given excellent results. But the main focus of these algorithms came when they could be transferred to multiple different cases with or without pre-trained weights, this led to the area of transfer learning.

During the Implementation of these architectures, many of the parameters had to be fine-tuned, for example, the input dimensions, number of filters, size of the filter, etc, so that it satisfies the needs of this experiment.

Training set size - 7483
Test set size - 1497
Validation set size - 2494
Early stopping condition - ‘Validation loss’
Restore best weights - ‘Least loss’
Activation - Relu and Softmax
Optimizer - ADAM

Table 1 shows how the architectures have performed on multiple runs, tuning and the experimental results obtained from the validation data.

SL. NO.	Models	Tuning	Test Results
1	LeNet-5	Parameters - 650978 Layers - 5 Convolutions - 3	Accuracy - 0.8096 Precision - 0.8127 Recall - 0.8068 F1 score - 0.808
2	AlexNet	Parameters - 1349754 Layers - 8 Convolutions - 5	Accuracy - 0.8049 Precision - 0.8108 Recall - 0.8108 F1 score - 0.8108
3	VGG-16	Parameters - 395870 Layers - 16 Convolutions - 13	Accuracy - 0.8116 Precision - 0.8138 Recall - 0.8080 F1 score - 0.8093
4	Inception Block	Parameters - 295162 Layers - 27 Convolutions - 22	Accuracy - 0.8424 Precision - 0.8480 Recall - 0.8408 F1 score - 0.8421

Table 1. Comparative results

VI. EXPERIMENTAL MODEL AND RESULT

With the learning from the comparative study, a hybrid architecture similar to Inception block and AlexNet was developed to improve the results. See Figure 3 for the architecture overview. The major change apart from hyper-parameter tuning is seen in the use of 'Leaky_relu' as the activation function in initial layers. The advantage of using this is to avoid dead neurons. By convection, the final layer activation function is still 'Softmax'.

The model result is as follows:
Accuracy on the test set - 0.8570
Precision on the validation set - 0.8609
Recall on the validation set - 0.8545
F1-score on the validation set - 0.8556

The above results show the improvement of the model in both TPR and FPR categories for both tampered and non-tampered classes. The results are obtained on validation set over multiple iterations during evaluation. During the training process we have made use of early stopping conditions to avoid overfitting. During the training, it was noticed that the below experimental hybrid model displayed faster learning compared to the other transfer learning model which such as VGG-16 and Inception block. Also, unlike the compared transfer model where they drastically decreased in training loss while the accuracy was not increasing, the hybrid model displayed the back and forth fall in gradient towards global minima.

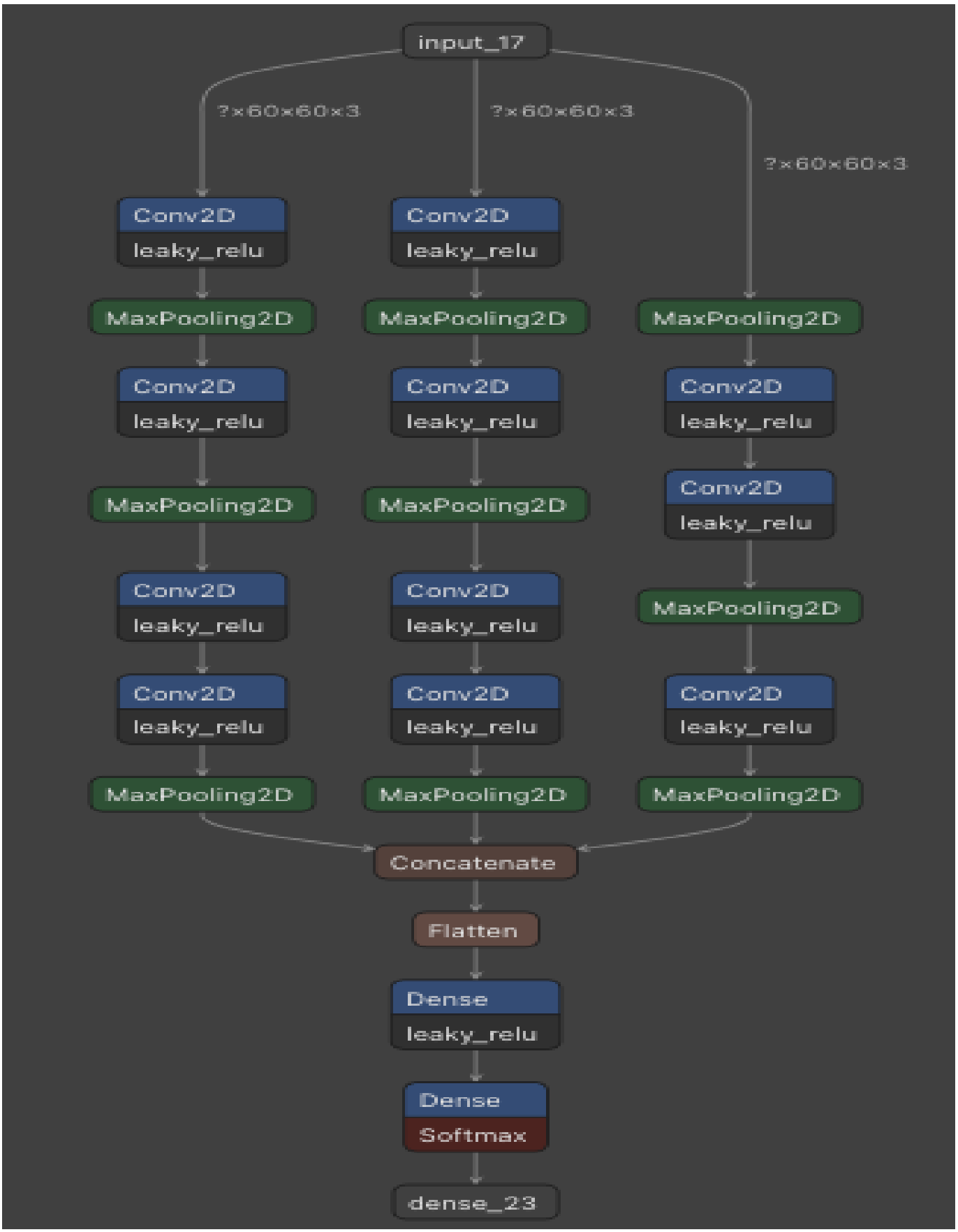

Figure 3. Experimental Model(Hybrid Model)

VII. FUTURE SCOPE

The present working of the model was built and tested upon passive image tampering, like copy-move-paste using the CASIA V2 dataset. Limiting the converted image to a size of 60x60. It can be further improved on different scale size images ranging from 120x120 to 1080x720. Also, on the note for detecting active image manipulation such as gaussian blur, etc.

Apart from increasing the classification boundary, we could also try and compare other advanced transfer learning architectures like Xception, Noisy-Student, etc. As well as to use the latest state of art image classifier Coca model, and try to combine that with Faster RCNN and other RPN-based networks.

VIII. CONCLUSION

As everyday new image editing products and applications have become easier and simpler to use, the negative effect of manipulating images for personal and financial gain have been increasing. This is a critical issue in the field of crime investigation as well as internet pornography. But, it is noticeable that most of these image manipulations is done on mobile devices upon JPEG images.To address this problem and to authenticate images in criminal justice. We have developed a new Deep learning architecture to classify between tampered and non-tampered images. The results of this model were better than the transfer learning model as shown in the comparative study.Finally, we like to conclude by saying that it was in our best interest to try and help all the concerned people in the field of image forensics.

REFERENCES

[1] SeungJu Cha, UiJung Kang, EunJung Choi, "The Image Forensics Analysis of Jpeg Image Manipulation", International Conference on Software Security and Assurance (ICSSA), IEEE 2018.
[2] Nor Bakiah Abd Warif, Mohd. Yamani Idna Idris, Ainuddin Wahid Abdul Wahab, Rosli Salleh, "An Evaluation of Error Level Analysis in Image Forensics", 5th International Conference on System Engineering and Technology, IEEE 2015.
[3] Monagi H. Alkinani, "Effects of Lossy Image Compression on Medical Image Registration Accuracy", International Conference on Consumer Electronics (ICCE), IEEE 2021
[4] Shuai Tan, Zhi Tan, " Improved LeNet-5 Model-Based On Handwritten Numeral Recognition", 2019 Chinese Control And Decision Conference (CCDC), IEEE 2019.
[5] Aiquan Yuan, Gang Bai, Lijing Jiao, Yajie Liu, "Offline Handwritten English Character Recognition based on Convolutional Neural Network", 10th IAPR International Workshop on Document Analysis Systems, IEEE 2012.
[6] Yung Lu, "Image Classification Algorithm Based on Improved AlexNet in Cloud Computing Environment", International Conference on Industrial Application of Artificial Intelligence (IAAI), IEEE 2020.
[7] Ali Abd Almisreb, Nursuriati Jamil, N. Md Din, "Utilizing AlexNet Deep Transfer Learning for Ear Recognition", Fourth International Conference on Information Retrieval and Knowledge Management, IEEE 2018
[8] Ningbiao Tang, Keqiang Yue, Xueying Yue, Hao Liu, Wenjun Li, "Automatic classification for corneal ulcer using a modified VGG network", International Conference on Artificial Intelligence and Computer Engineering (ICAICE) , IEEE 2019.
[9] Abitya Bagaskara, Muhammad Suryanegara, "Evaluation of VGG-16 and VGG-19 Deep Learning Architecture for Classifying Dementia People", 4th International Conference of Computer and Informatics Engineering (IC2IE), IEEE 2021.
[10] S. Kavitha, B. Dhanapriya, G. Naveen Vignesh, K.R.Baskaran, "Neural Style Transfer Using VGG19 and Alexnet", International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), IEEE 2021.
[11] Jun-Liu Zhong and Chi-Man Pun, "An End-to-End Dense-InceptionNet for Image Copy-Move Forgery Detection", TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, IEEE 2020.
[12] Jinming Zhang, Zuren Feng, "Inception DenseNet With Hybrid Activations For Image Classification", 6th International Conference on Systems and Informatics (ICSAI 2019), IEEE 2019.
[13] Kunihiko Taya, Nobutaka Kuroki, Nobutaka Kuroki, Nobutaka Kuroki, Masahiro Numa, "Detecting tampered regions in JPEG images via CNN", 18th International New Circuits and Systems Conference (NEWCAS), IEEE 2020.
[14] S Manjunatha, Malini M Patil "Deep learning-based Technique for Image Tamper Detection", Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), IEEE 2021.
[15] Mamoru Nakahara, Nariaki Imamura, Shota Furukawa, "Tampered Regions Detection for JPEG Images by Using DCT Coefficients of Re-Saved Image", 20th International Symposium on Communications and Information Technologies (ISCIT), IEEE 2021.
[16] Vinod Mall, Anil K. Roy, Suman K. Mitra, "Digital Image Tampering Detection and Localization Using Singular Value Decomposition Technique", Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), IEEE 2013.
[17] Rayan Sulaiman Khalaf, Asaf Varol, "Digital Forensics: Focusing on Image Forensics", 7th International Symposium on Digital Forensics and Security (ISDFS), IEEE 2019.
[18] Tiziano Bianchi, Alessandro Piva, "Detection of Nonaligned Double JPEG Compression Based on Integer Periodicity Maps", IEEE Transactions on Information Forensics and Security, IEEE 2012.
[19] - Wang Junwen, Liu Guangjie, Dai Yuewe, Wang Zhiquan, "Detecting JPEG image forgery based on double compression", Journal of Systems Engineering and Electronics, IEEE 2009.