iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50770

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

Brief Cognizance into Tampered Image Detection

Raghav Gupta[1], Nikhil[2], Aditya Nagyal[3], Nidhi Mishra[4]
[1, 2, 3]*Btech CSE, ASET, Amity University, Noida, UP, India*
[4]*Dept. of CSE, Amity University, Noida, UP, India*

Abstract: ***This paper is a testament to the research carried out to detect tampered images using deep learning and machine learning techniques, including the implementation of various standard and hybrid convolution neural networks, thus creating a comparative study for the neural networks. Tampered images can be found on various social media networks and sites thus creating a false impression in the minds of people consuming the information from them. Hence, to filter out these fake images the following research work and implementation are carried out.***
Keywords: CNN, coco, Kerass, splicing, region removal, retouching, lightning tampering.

I. INTRODUCTION

With this research, the aim is to develop and train a convolutional neural network that will be able to detect tampering in images. Further, it will also be able to classify the type of tampering in the given image using standard and hybrid deep learning and machine learning techniques. Multiple open-source image databases, such as CASIA V2.0, MICC-F2000, CoMoFod were found on searching for a large enough tampered image database. However, these databases were not large enough to train an accurate network. Thus, to increase the size of the dataset, the use of multiple tampering scripts and algorithms like splicing, Retouching etc. were made. Research work included study and deep understanding of these above-mentioned image transformation techniques using various approaches like gaussian filters, seam carving etc. Further research will use the convolution algorithm using convolution neural networks which aim to classify meddled images. To develop, train and test neural networks use an open-source keras implementation of convolution and pooling operations. The aim of this paper is to create a comparative study between different architectures like Resnet, Inception Net, NASnet etc.

II. LITERATURE REVIEW

Splicing is one of the most popular methods for manipulating digital images; it involves copying a specific section from one image and pasting it into another. The identification of image forgeries is thought to be a trustworthy method of confirming the veracity of digital photographs. In this paper, we suggested a method based on the cutting-edge ResNet50v2 deep learning architecture. The suggested model uses the ResNet50v2 architecture and the YOLO convolutional neural network (CNN) weights to process image batches as input.**[1]**

The DeepFake face image, which was built on image noise and image augmentation, may be detected using an upgraded VGG network called NA-VGG, which is shown in this research. First, the SRM filter layer is utilized to highlight the image noise features in order to learn the tampering artefacts that might not be seen in RGB channels; second, the image noise map is augmented to weaken the face features. In order to train the network and determine whether the image is fake, the augmented noise images are finally fed into it. The experimental outcomes utilizing the Celeb-DF dataset demonstrated that NA-VGG significantly outperformed other state-of-the-art methods. **[2]**

Convolutional Neural Network (CNN) and Support Vector Machine (SVM), two powerful classifiers that have demonstrated success in identifying various patterns, are combined into a hybrid model in this paper. In this model, SVM serves as a recognizer and CNN serves as a trainable feature extractor. This hybrid model creates predictions by automatically extracting information from unprocessed photos. The well-known MNIST digit database has been the subject of experiments. In comparison to prior research on the same database, this fusion produced better results, with a recognition rate of 94.40% with 5.60% rejection and 99.81% without rejection. These performances have been examined in light of those given by human test subjects. **[5]**

Deep learning-based algorithms do remarkably well in detecting picture modification. However, the majority of them have poorly standardized handcrafted or predefined features. Meanwhile, they exclusively focus on manipulation localization and disregard manipulation classification. We suggest a coarse-to-fine architecture called Constrained R-CNN for thorough and precise picture forensics to overcome these problems.

First, a unified feature representation is directly learned from data via the learnable manipulation feature extractor. Second, with the following manipulation classification and coarse localization, the attention region proposal network successfully discriminates altered regions. After that, the skip structure combines high- and low-level data to enhance the global manipulation features. The model is then guided by the coarse localization information to learn the finer local features and segment out the tampered region. **[6]**
Convolutional neural networks (CNNs) are used to automatically create hierarchical representations from the input RGB-colored photographs in a new deep learning-based method for detecting false images. For applications like copy-move detection and picture splicing, the recommended CNN is created specifically. The basic high-pass filter set used to calculate residual maps in the spatial rich model (SRM) is used as the initialization of the weights at the first layer of our network as a regularizer to effectively suppress the effect of image contents and capture the subtle artefacts introduced by the tampering operations. In order to extract dense features from the test images, the pre-trained CNN is next employed as a patch descriptor in a feature fusion technique. **[8]**
This research suggests a novel median filtering detection approach based on CNN to address this issue. In particular, a brand-new network structure called MFNet is built. The first step in preprocessing is upsampling the small-size images using the closest neighbour interpolation approach. The up-sampling process can effectively preserve the median filtering property, allowing for increased contrast between the original image and its median-filtered counterpart. The first and second levels of the MFNet then use the well-known mlpconv structure. The nonlinear classification capability of the suggested method can be improved with mlpconv layers. **[9]**

Table I. Reference paper table

S.No.	Author	Methodology	Result/Implication	Limitations
1.	Barad, Z. J. & Goswami, M. M. [1]	The suggested model uses the ResNet50v2 architecture and the YOLO convolutional neural network (CNN) weights to process image batches as input.	For tamper detection, traditional approaches include hand-crafted features. the conclusion from the survey was that the standard procedures do not consistently counteract different tampering techniques.	These datasets' inadequate size restricts the use of DL-based tampering detection methods. Deep network training is challenging and demands powerful computers and a sizable dataset.
2.	Chang, X., Wu, J., Yang, T. & Feng, G. [2]	The experimental outcomes utilizing the Celeb-DF dataset demonstrated that NA-VGG significantly outperformed other state-of-the-art methods.	The findings demonstrate that NA-VGG much improved at identifying DeepFake face photos.	NA-VGG gave an average AUC performance accuracy of only 85.7%.
3.	Niu, X. X., & Suen, C. Y [5]	The model is then guided by the coarse localization information to learn the finer local features and segment out the tampered region.	The handwritten digit recognition issue has been addressed by the development of a new hybrid CNN-SVM model. This model used SVM as the output predictor and CNN as an automatic feature extractor.	Based on factors like the input layer's size, the amount of feature layer maps in layers 2 to 4, the model's kernel functions, etc., improvements might be achieved.

4.	Yang, C., Li, H., Lin, F., Jiang, B., & Zhao, H. [6]	The model is guided by the coarse localization information to learn the finer local features and segment out the tampered region.	The technique is more generic and successful for complicated picture forensics than earlier approaches because it can extract alteration clues directly from data without relying on any constructed components.	The features they use are always specifically established for one type of modification technique, these methods cannot be applied to forensics and are tested for only four benchmarks.
5.	Rao, Y., & Ni, J. [8]	A feature fusion strategy is then investigated to extract dense features from the test pictures using the pre-trained CNN as a patch descriptor.	Comprehensive experiments on a number of public datasets have been conducted, showing that the suggested CNN-based methodology performs better than existing cutting-edge image forgery detection techniques.	While forgeries photos frequently resemble real ones statistically and visually when using elaborately constructed techniques, the typical deep learning architecture may not be directly applicable to image tampering detection.
6.	Tang, H., Ni, R., Zhao, Y., & Li, X. [9]	The first and second levels of the MFNet then use the well-known mlpconv structure. The nonlinear classification capability of the suggested method can be improved with mlpconv layers.	Presented an efficient CNN model called MFNet as well as nearest neighbour interpolation to enlarge the small-size testing images.	Results don't verify whether the magnifying idea is useful for image data forensics.

III. METHODOLOGIES

A. *Gaussian Blur*

It is a common effect in graphics software, usually used to lessen detail and visual noise. This blurring technique produces a smooth blur that looks like you're seeing through a translucent screen, which is noticeably different from the bokeh effect that is created by an out-of-focus lens or the shadow of an object under normal lighting.

The Gaussian blur is a type of picture-blurring filter that determines the change to apply to each pixel in the image using the Gaussian function, which also describes the normal distribution in statistics. the one-dimensional Gaussian function formula.

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

Equation 1: Gaussian Function in one-dimension

B. *Python Pillow*

A library for image processing called Pillow diverged from PIL (Python Image Library). The pillow is currently widely utilised as PIL is no longer being developed. Simple image processing operations like resizing (scaling), rotation, and trimming (partial cutoff) can be carried out, but advanced image processing (facial recognition, optical flow, etc.) like OpenCV cannot be done. We have used the PIL library for various purposes like loading images, applying random crop operations, changing the brightness of the region of an image, and drawing over an image to create copy-paste and cut-paste forgery effects.

Basic operations of python image libraries used in the project –

1) Open – To load the image file into the memory so that we can apply various transformations.
2) Save – To export or save the transformed image in a new directory.
3) New – To create a black channel image that will be used as a mask to superimpose the image over another image.
4) FromArray – To convert the numpy array format to PIL image format so that it can be displayed and saved.
5) Close – To remove unused images from memory thus making memory free to load more images simultaneously.
6) Draw – Draw the numpy array over the existing image using the mask image as an area reference since the mask contains all 0 pixels in both dimensions.
7) Pie Slice – To create a circular-shaped crop region randomly for copy-paste and cut-paste forgery effect.
8) Gaussian blur – To apply a gaussian blur algorithm over an area of the image, so that we will get more augmented data and our model will be generalized for more types of images.
9) Copy – To create a deep copy of an existing pixel region so that any operation applied to the image doesn't alter the original image.

C. *Random*

Using a random module's randrange() and randint() functions, we are able to produce a random integer in a certain range. The following functions to generate random numbers in Python are covered in this function. It is used to find the random area and location of the image on which we are going to apply transformation operations.

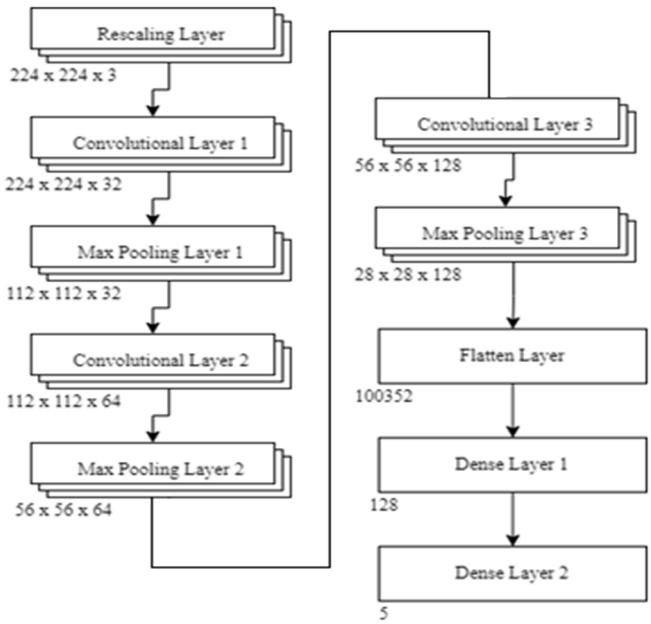
D. *Numpy*

The Python programming language now has support for massive, multi-dimensional arrays and matrices, as well as a wide range of high-level mathematical operations to work on these arrays, thanks to the NumPy module. Since the image is represented in matrix form we have used the numpy library for applying various matrix transformations to modify the image as numpy is fast and memory efficient as it's based on python which is implemented in c language which is close to the hardware and provides direct memory access thus making it fast and size of integers are smartly allocated in numpy that makes computation and storage more optimised that usual python lists.

Basic operations used –

- Array – to get an array object from a PIL object so that we can overlap other regions of the image to create a copy-paste forgery effect.

E. *Architecture of Simple CNN*



ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

IV. DIFFERENT NEURAL NETWORK MODELS

A. *Resnet*

Residual Network is abbreviated as Reset. It addresses the problem of disappearing gradients in bigger neural networks and was first introduced in 2016. Utilizing residual connections is the core idea underlying Resnet. By bypassing one or more layers, information can move directly from earlier layers to later layers. This makes it possible to train intense neural networks accurately and effectively.

B. *InceptionNet*

InceptionNet, commonly called GoogleNet, was introduced in 2014. The utilization of various convolutional and pooling methods is InceptionNet's primary area of expertise. There are several "inception modules" comprising various shapes and sizes of convolution layers. InceptionNet performs excellently on computer vision problems. Speech recognition and natural language processing both use it.

C. *XceptionNet*

Extreme Inception is abbreviated as XceptionNet. It was unveiled that year. Deep neural networks' effectiveness and precision are enhanced by it.
It performs depth-wise separable convolutions, dividing the channel-wise and spatial convolution into two independent operations. Hierarchical blocks that include depth-wise separable convolutions make up XceptionNet.

D. *NASNET*

"Neural Architecture Search Network" is the abbreviation for NASNet. It was created to use the method of neural architecture search to automate the process of constructing the neural network architecture. The same holds true for computer vision tasks.

E. *MobileNet*

In 2017, MobileNet was unveiled. It is concentrated on offering a computationally effective answer for embedded and mobile devices. Additionally, depthwise separable convolutions were applied. It is compact and has lower performance demands, and it offers good computational accuracy.

V. REQUIREMENTS

A. *Hardware Requirements*

1) Dedicated GPU (Nvidia/AMD)
2) Intel i5 or better
3) 8GB RAM or better
4) 50 GB Free Space

B. *Software Requirements*

1) VS Code
2) Python 3.8+
3) PyCharm
4) Windows / Linux
5) Tensorflow APIs
6) Keras APIs

VI. RESULT AND DISCUSSION

We had initially searched existing databases, although we --were not satisfied with the ones we found. We compared the databases and finally created a custom database for our training purposes. The following table contains all the demerits of the popular and open image databases that contain fake images.

Table II. Comparison Chart

Model Used	Dataset	Demerits Observed
ResNet50v2	CASIA_v1 CASIA_v2	CASIA dataset consists of only 5000 tampered images, which is much less for a deep learning network to learn from data.
NA-VGG	Celeb-DF	Trained only on fake faces.
MANFA HF-MANFA	Manual Face Dataset	Trained only on fake faces.
R-CNN	NIST16 COVERAGE Columbia dataset	Works well with digit datasets, experiencing difficulties in identifying complex handwriting.
SVM + CNN	Public datasets	-
CNN + SVM	MSNIT	Works well with digit datasets, experiencing difficulties in identifying complex handwriting.

From the above-discussed methodologies sample dataset of around 5000 images had been generated by scripts and various algorithms to induce morphing in images, this dataset is created for testing purposes of the scripts.

We successfully created our own database containing 25000 images. We did this by automating the process of tampering with Python scripts. Our scripts used various methodologies, such as Gaussian Blur, Python Pillow etc, to achieve the desired tampering.

We took untampered images from the COCO dataset, which is a large, open-source database for the purpose of object detection. We ran our scripts on these images to form different types of tampering.

We introduced multiple types of tampering-

1) *Region Removal:* In region removal, a certain part of the image is removed entirely, replaced with either nothing or a particular colour.
2) *Lightning Tampering:* In Lightning Tampering, we change the lighting of the image or a particular area of the image. This can be used to alter the objects visible in an image.
3) *Retouching:* In retouching, the entire image or a part of an image is put through a filter, which changes the contrast, colour or sharpness of the image.
4) *Splicing:* It consists of pasting a small part of an image into itself or a different image. This is used to add to remove objects in an image.

The above-discussed tampering methodologies are illustrated in the following table: -

Table III. Sample Tampered Images on COCO dataset using custom scripts

Original 	Original 
Splicing (Before Tampering) 	Splicing (After Tampering) 

Lightning Tampering (Before Tampering) 	Lightning Tampering (After Tampering) 
Retouching (Before Tampering) 	Retouching (After Tampering) 
Region Removal (Before Tampering) 	Region Removal (After Tampering) 

Table IV. Comparative table for all neural network Models

Neural Network Model	Result and Outcome	Validation Accuracy
ResNet	Resnet had one of the lowest training and validation accuracy rates during our training. It achieved a training accuracy of 80.01%.	66.43%
InceptionNet	Under our training, InceptionNet did well, achieving a training accuracy of 94.91%.	89.82%
XceptionNet	During our training, XceptionNet delivered the most accurate results. XceptionNet achieved a training accuracy of 96.39%.	94.67%
NasNet	Nasnet achieved a training accuracy of 94.06%.	63.51%
MobileNet	It achieved good training accuracy but lower validation accuracy. It had 94.42% training accuracy	84.60%

VII. CONCLUSION AND FUTURE SCOPE

Aim of this research are to perform a study and create a model that will be robust enough to identify fake news on social media websites and identify tampered images to curb the spread of false news across the internet. Combinations of different methodologies are used to achieve this task; Convolution neural networks are used to extract the features from images and map these features to a feature map. Further in the architecture fully connected layers or dense layers were used to classify images. A simple neural network architecture with 3 convolution layers, 3 max-pooling layers, 1 rescaling layer, 1 flatten layer, and 2 dense layers with softmax activation in the last layer. The network consists of around 13 million parameters with the last layer having an output size of 5. The network is built using the TensorFlow framework and its definition of layer provides an easy implementation of sequential models. Also, after comparing all the neural network models i.e. ResNet, InceptionNet, XceptionNet, NasNet, and MobileNet we saw that XceptionNet gave the best accuracy with a training accuracy of 96.39% and validation accuracy of 94.67%.

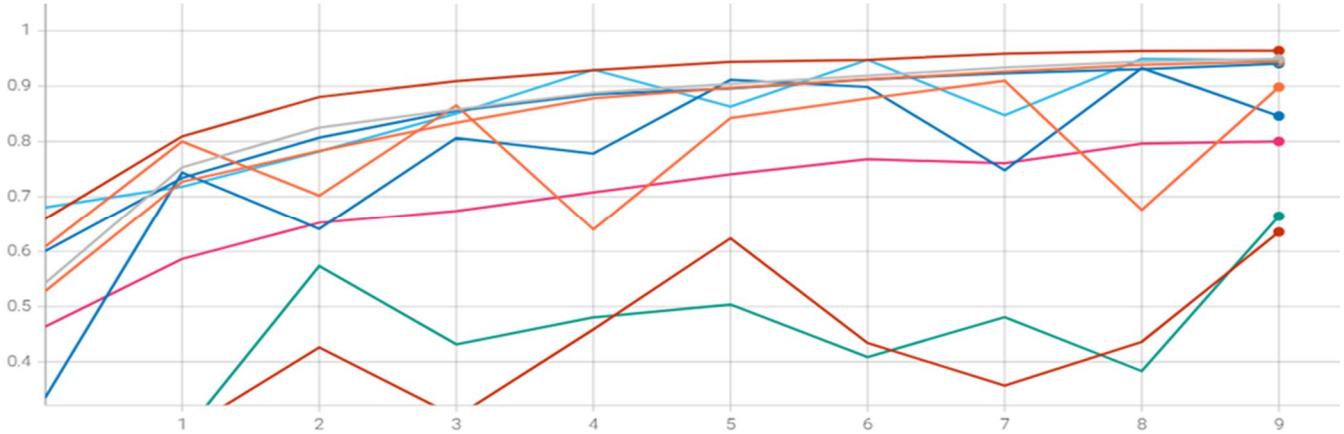


Image 1: Graph shows the comparison of various models used in the study

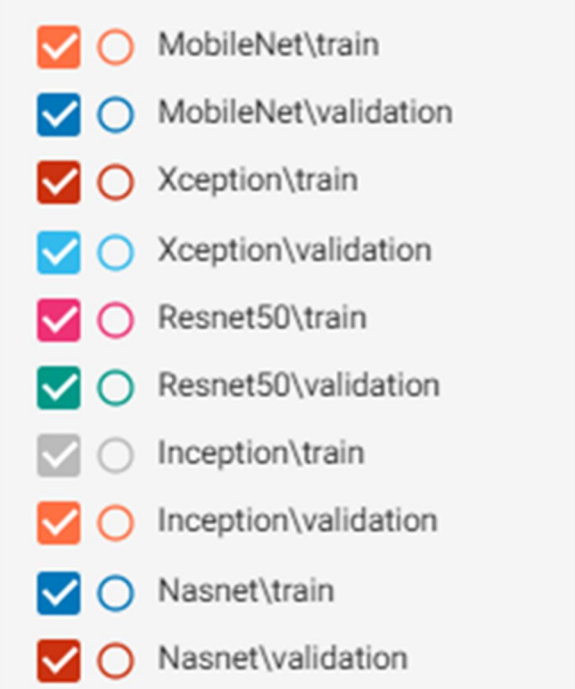
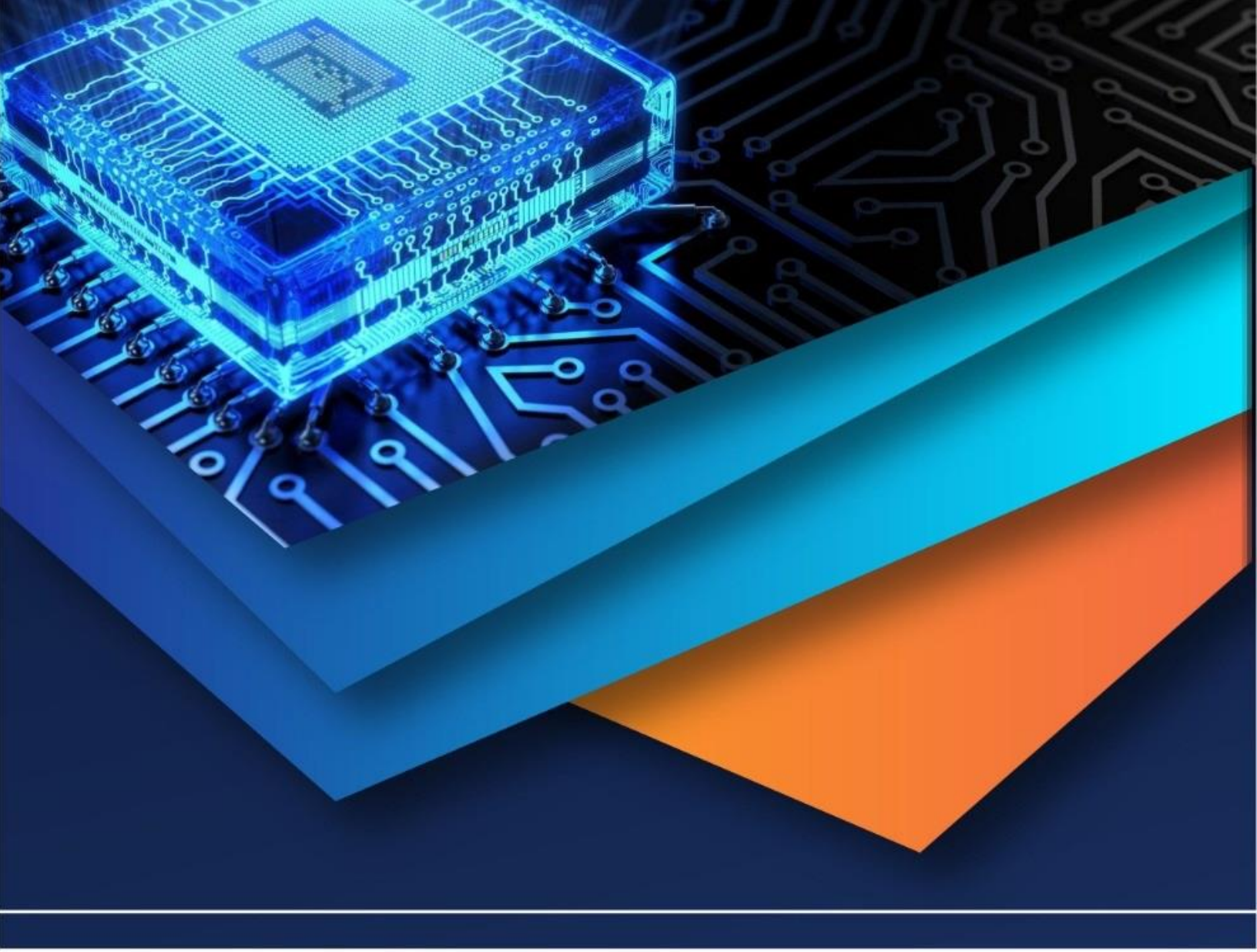


Image 2: Color coding of each plot line on graph

REFERENCES

[1] Barad, Z. J., & Goswami, M. M. (2020, March). Image forgery detection using deep learning: a survey. In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS) (pp. 571-576). IEEE.
[2] Chang, X., Wu, J., Yang, T., & Feng, G. (2020, July). Deepfake face image detection based on improved VGG convolutional neural network. In 2020 39th Chinese Control Conference (CCC) (pp. 7252-7256). IEEE.
[3] Dang, L. M., Hassan, S. I., Im, S., & Moon, H. (2019). Face image manipulation detection based on a convolutional neural network. Expert Systems with Applications, 129, 156-168.
[4] A Survey on Image Tampering and Its Detection in Real-world Photos - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Popular-image-tatampering-datasets_tbl3_329518852 [accessed 19 Oct 2022].
[5] Niu, X. X., & Suen, C. Y. (2012). A novel hybrid CNN–SVM classifier for recognizing handwritten digits. Pattern Recognition, 45(4), 1318-1325.
[6] Yang, C., Li, H., Lin, F., Jiang, B., & Zhao, H. (2020, July). Constrained R-CNN: A general image manipulation detection model. In 2020 IEEE International conference on multimedia and expo (ICME) (pp. 1-6). IEEE.
[7] Manjunatha, S., & Patil, M. M. (2021, February). Deep learning-based Technique for Image Tamper Detection. In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 1278-1285). IEEE.
[8] Rao, Y., & Ni, J. (2016, December). A deep learning approach to detection of splicing and copy-move forgeries in images. In 2016 IEEE international workshop on information forensics and security (WIFS) (pp. 1-6). IEEE.
[9] Tang, H., Ni, R., Zhao, Y., & Li, X. (2018). Median filtering detection of small-size images based on CNN. Journal of Visual Communication and Image Representation, 51, 162-168



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY