Exposing Digital Image Manipulation

K. Nikitha[1], G. Rohitha[2], Fazal Ur Rahman[3], K. Kushal[4], Dr. Padmaja Pulicherla[5]
[1,2,3,4] *Student of Computer Science and Engineering, HITAM Hyderabad, India*
[5]*Professor and HOD, Department of CSM, HITAM, Hyderabad, India*

***Abstract*—This paper talks about the increasing sophistication of image manipulation software have made it difficult to distinguish between authentic and altered photographs, requiring new detection methods. Traditional techniques often fall short against advanced fakes, leading researchers to explore Machine Learning (ML) in Artificial Intelligence (AI) for innovative solutions. A promising approach involves segmenting images into smaller regions and analysing them with ML algorithms trained on large datasets of real and altered images. These algorithms detect inconsistencies like pixel intensity and texture patterns, assigning manipulation likelihood scores. Convolutional Neural Networks (CNNs), a deep learning algorithm, are particularly effective in extracting complex visual patterns, making them a significant tool in identifying subtle signs of forgery and counterfeiting techniques. This method identifies subtle signs of manipulation, often undetectable by traditional means. By training on large, diverse datasets, these ML models recognize various counterfeiting techniques, making CNNs a significant advancement in the fight against image forgery.**

***Index Terms*—Image Forgery Detection, Machine Learning (ML), Deep Learning, Convolutional Neural Networks (CNNs), Image manipulation, Pixel intensity, Visual data analysis.**

I. INTRODUCTION

Image Alteration is an intentional alteration of the digital pixels for forensic purposes. They might resize and crop photos, or they might make adjustments to a picture to hide something. It's like adding a layer on top of a painting to alter its appearance. Because editing tools have advanced, it has become more difficult to spot forgeries. Using picture forgeries has some significant disadvantages. It can spread false information and erode confidence in real photographs. Legally, it could lead to problems with copyright and privacy. The main goal of image forgery detection is to identify and assess altered images using CNN and Machine Learning to confirm their authenticity and integrity. It is easy for fake photos to expand on social media, undermining public confidence in information. Fake images can incite violence, change people's perceptions and damage people's reputations. This work uses machine learning to identify photo forgeries uploaded to address these problems.

The main objective of this study is to design and implement a CNN based Image Forgery Detection System. CNNs, a subset of deep learning algorithms can automatically extract hierarchical features from unprocessed image data, they are a good fit for image analysis tasks.

II. RELATED WORK

Now in this related work part, we will discuss some work that has been done in this field.

Based on Dense Local Descriptors and ML approaches, this work presents a new approach to image forgery detection, using descriptors that have been effective in the past for texture classification, steganalysis and forgery detection. The suggested detector performs well by combining and fine-tuning these characteristics and using an SVM classifier on the training set. The research effectively lowers the rate of missed detections by developing a straightforward but highly specific copy-move detector based on region matching and decision fusion to overcome this constraint. Overall, the findings are encouraging for the development of forgery detection techniques.[1]

The rise of mobile devices with cameras and image editing software has made it easier and easier to produce phony photos as social media integration gets more and more integrated into daily life. Politicians frequently use these fake images to further their own agendas and damage the reputations of famous figures. Image forgery detection refers to the difficulty of identifying and localizing such manipulated pictures. However, using deep learning methods without the

need for human feature engineering is but using deep learning methods without human feature engineering is difficult because of how little the altered regions are in comparison to the size of the entire image. The binary classification of real and fake images, this research paper primarily focuses on photographs while emphasizing the application of deep learning models to get the best results possible in tasks involving the classification of forged images [2].

To identify intentionally blurred edges inside manipulated photos, this research presents a novel image forensics technique. The Non subsampled contourlet Transform (NSCT) is first used to analyse the edges of the image. Next, by studying phase congruency and prediction-error images, differences between normal and blurred edges are extracted. Blurred edges can be distinguished by using these attributes to train a Support Vector Machine (SVM). To differentiate between manual blur and defocus, local definition is also included. According to experimental results, the approach can effectively identify possible image blurring and locate tampered borders with a reasonable degree of precision [3].

Due to the increasing use of cameras, image modification technologies are useful for improving images but also raise concerns because they can be used to propagate false information by creating photos that are altered. While conventional methods have been developed to detect image forgeries, CNN has garnered attention recently. Many CNN-based techniques, however, they are only applicable to particular kinds of forgeries. With a focus on double image compression, this work presents a lightweight deep learning system that is intended to effectively identify invisible forgeries. The suggested model offers potential performance gains over existing methods by taking use of the variations between original and recompressed images [4].

This work investigates digitally modified documents and offers a method for distinguishing original from morphed documents in the context of the rise in cybercrimes made possible by readily accessible photo editing software. To counteract fraudulent operations in the digital sphere, a Graphical User Interface (GUI) for detecting digitally manipulated photos is developed [5].

This article describes an approach that uses the VGG-16 Convolutional Neural Network (CNN) to identify splicing, a common form of digital image forgeries. Upon receiving image patches as input, the suggested network architecture determines if they are authentic or counterfeit. Patches are chosen from the borders of embedded splicing as well as from the original picture regions throughout training. When the algorithm is compared to other solutions, the experimental results show that it is effective in obtaining high classification accuracy. The study validates its experimental findings using the CASIA dataset [6].

It is intense to authenticate digital images due to sophisticated forgery techniques such as, but not limited to, the copy-move forgery whereby digital assets are manipulated by taking a portion of an image and altering it then merging the section back to the original picture. Other detection methods are further impeded by factors such as additive noise, image data jpeg compression, and rotation. The implementation of the proposed strategy involves Censure key point detection neurologically incorporated into the architecture to assist in identifying and localizing the copy-move forged region. The approach overcomes various attacks including but not limited to edge transformation, scale variation, rotation, undergoes JPEG compression, adds noise and brightening/contrasting images even when a lot of textures are in the image due to CNNs data driven learning and key points' supplementation with CNN features. The disadvantages, however, include reliance on large number of and diverse kinds of training datasets, more expensive computational power, and difficulty or loss of accuracy in glaringly overprocessed images or extremely covered up forgeries [7].

The investigation of copies-move forgery detection is difficult because ground truth (GT) is ambiguous resulting from convolutional neural networks with different models producing different yet similar patch-based patterns. The solution proposed in this work addresses this problem by coming up with an ingenious ground truth image in three steps; creating ground truth images from the models which are referred to in the scheme as GTnet, combining the images created into a single image mist GTconv, and again reconstructing the combined image GTdecomp using the threshold filter which aims to provide a compromise between classification and segmentation information. Accordingly, this provided an improvement in the accuracy and overall F1 score of 0.4% and 0.2% respectively, where graded AUC value

was 0.9 rated as "Excellent". The adverse aspects involve, however, attention to being issues that include reliance on expensive operations, excessive dependence on threshold values, and difficulty in the preservation of generalization over varying datasets [8].

In this age of image manipulation software, it has become very important to detect unsuspecting manipulations. In this paper, we put forward a copy-move forgery detection (CMFD) technique based on using SURF and PCET. The image is segmented into non-overlapping blocks by the use of super pixel segmentation technique, which is then classified into profiles consisting of smooth and textured regions. Key points are located using SURF and features incorporated into the feature matching with PCET. False matches are removed with the help of the RANSAC algorithm together with filtering techniques, while the areas of the image that have been altered are further improved through the use of mathematical morphology and iterative methods. The approach is well suited for use in the area of concern counterfeiting works using smooth or visually homogenous areas and is resistant to a number of strains such as rotation, scale, blur, lossy jpeg compression and noise among others. Nevertheless, weaknesses are identified in the lack of effectiveness in very busy scenes leading to processing times of a high order owing to heavy usage of iteration and finally poor performance in image correction after excessive image enhancement [9].

Even though digital images can be easily posted on social media, the ability to forge such images is dangerous because of the scenario of misrepresentation. In this paper, a deep learning-based approach is introduced using transfer learning to ensure that image splicing and copy-move forgery detection is done simultaneously. The method analyses the difference in compression artifacts that is present between the forged and authentic regions and creates a featured image to be fed into a pre-trained network with a classifier model that has been fine-tuned. Results from comparisons of eight pre-trained networks indicate that MobileNetV2 is the best in detecting forgery with accuracy at 95%, fewer parameters, and overall speed of training. One of the limitations is the need to be dependent on compression-based differences, lower detection accuracy for very compressed or low-quality images, and its transfer learning that in some cases, extensive training datasets are prerequisite [10].

One of the most difficult tasks in digital image forgery detection, especially copy-move forgery, is that several methods are used to hide the evidence of forgery, including geometric transformations, scaling, rotation, JPEG compression, and Additive White Gaussian Noise (AWGN). This work proposes a novel solution that effectively utilizes the self-supervised Super Point detector for key point detection and descriptor extraction, which helps achieve precise localization of forgeries. The method is confirmed to be robust across different textures and types of attacks providing reasonable results and cost efficiency for real-time operation. The drawbacks, on the other hand, include those related to overuse in the presence of heavy transformations or noise, ineffective key point extraction, and diminished performance in the presence of too much visual clutter or uniform areas [11].

Modern deep learning models tend to insist on the resizing of input images when they are trained due to limited computational resources, which may lead to the loss of fine details that are important in image forensics. In this work, we present a forgery detection system based on CNN that does not require input image resizing during training as it works on unedited full-resolution images and utilizes gradient checkpointing, thus allowing end-to-end training under limited memory and weak supervision. The framework achieves better performance than the baseline methods on several forensic datasets. Nevertheless, drawbacks are the fact that training is more computationally expensive, it may be affected by noise when dealing with high-resolution images, and it is based on weak supervision, which could affect the ability to recognize as many types of forgeries as possible [12].

This study extends the existing definition of fractional Zernike moments (FrZMs) for the purpose of introducing fractional quaternion Zernike moments (FrQZMs) as developments directly related to the field of quaternion signal processing. This facilitates the efficient computation of each quaternion component independently. The colour image copy-move forgery detection method proposed here is based on the use of

FrQZMs as features and a modified Patch Match for feature matching. Experiments performed on the FAU and GRIP datasets demonstrate superior performance in all visual detection tasks compared to the current state of the art techniques, especially with extra operations included. However, there are some limitations such as complicated calculations when processing high order quaternions as well as susceptibility to excessive changes or noise that can compromise the effectiveness in real life applications [13].

Copy-move forgery is another widely known example of image manipulation, and it is possible to combat its effects using key point-based detection methods. In this paper we propose an CMFD methodology which groups SIFT key points according to their scale and colour so that the number of matching operations is considerably less by working with smaller key point clusters. Interestingly, there is a new localization method that draws the regions suspected of being tampered by measuring and comparing similar neighbourhoods iteratively every time by using two different similarity metrics. We show the results of experiments against three datasets, and these experiments reveal that the time efficiency, detection performance, and forgery localization ability of the proposed approach are better as compared with the existing methods. On the other hand, easiness of implementation has pros and cons, such as possible dependence on the threshold during the process of key point clustering, lower precision in the areas with little or no texture, and problems with very sharp or very noisy pictures [14].

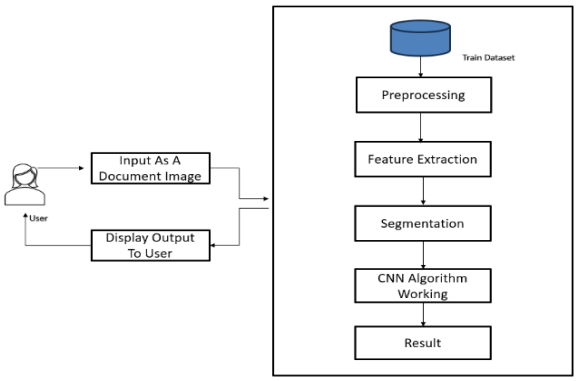
III. PROBLEM STATEMENT

The growing advancement of image forgery software has made it increasingly difficult to reliably distinguish between authentic and altered photographs, as traditional detection techniques are often ineffective against sophisticated forgeries. Modern digital manipulations can be highly subtle, evading detection by the human eye and conventional forensic methods. One promising approach leverages Machine Learning (ML), specifically Convolutional Neural Networks (CNNs), which are highly effective at analysing complex visual data. By segmenting images into smaller regions and examining features such as pixel intensity, texture patterns, and colour inconsistencies, CNNs can detect minute signs of manipulation that traditional methods miss. These models are trained on large, diverse datasets of both authentic and manipulated images, enabling them to learn the nuances of various forgery techniques and assign manipulation likelihood scores with high accuracy. However, the primary challenge lies in developing CNN-based models that can generalize across a wide range of image manipulation methods and remain effective in real-world applications. Success in this area is critical for domains such as digital forensics, journalism, and legal evidence verification, where the ability to detect image tampering is essential for ensuring trust, credibility, and the integrity of visual content.

IV. PROPOSED METHODOLOGY

The proposed framework is designed to develop a robust and efficient fake image recognition application that uses deep learning especially CNN, to analyse and classify images true. The project includes front-end and back-end development, focusing on easy-to-use, user-friendly web applications for end users. Because of their ability to learn complex hierarchical features directly from raw pixel data, CNNs are particularly effective for image mesh network recognition. This capability allows CNNs to detect subtle changes in images, distinguishing between real and cropped areas through a variety of scholarly features. These features range from low edges to high resolution structures, which are important in lie detection applications

CNNs use spatial algorithms that incorporate convolutional and pooling layers, which help preserve the spatial structure and local trust of the input image These spatial algorithms are useful for detecting splicing, cloning, copy-moving manipulation, etc Despite the emphasis on inconsistencies, CNNs' architecture is flexible towards modifications such as their scaling, rotation, and translation. This robustness is important in real-world scenarios where images may have been zoomed in, slightly rotated, or otherwise manipulated before reaching the required range. The flexibility of CNNs further enhances their usefulness in fraud detection. Their settings can be changed and optimized to suit a particular networker.



Preprocessing of Dataset:

- Preprocess Active Error Level Analysis ELA based on decompression difference between the original image and its different compressed versions to create tampering sensitive representations.
- Resize to a uniform size (for instance: 128x128 pixels) and normalize pixel values to the range [0, 1].

Feature Extraction: Collect the real dataset and the fake image dataset present at their specific directories: mini/Au for real and mini/Tp for fake images. Labelling: real image '1', fake image '0', along with the use of balancing for collection dataset, so that it only Favors one class.

Segmentation: Processed data shall be split into train, validate, and test data for enhanced effective evaluation using,"train_test_split".

Convolutional Neural Network (CNN)- Architectures: Build a sequential model of CNN with these layers:

Convolutional Layers: Their role is mainly extracting the spatial features from the filters used to detect edges, textures, patterns, etc.

Max-pooling Layers: The size of the spatial dimensions is reduced, but the important features on which computation relies are preserved.

Dropout Layers: This prevents overfitting because neurons are dropped out randomly during the training process.

Fully connected layers: They are obtained by merging the extracted features, which are used in prediction.

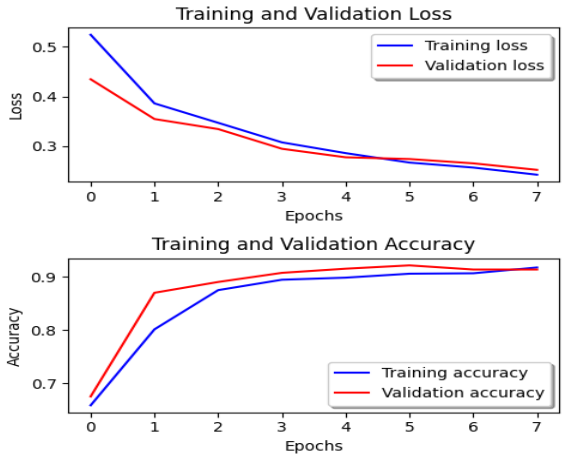
Training the CNN Model:

Compile the model with:

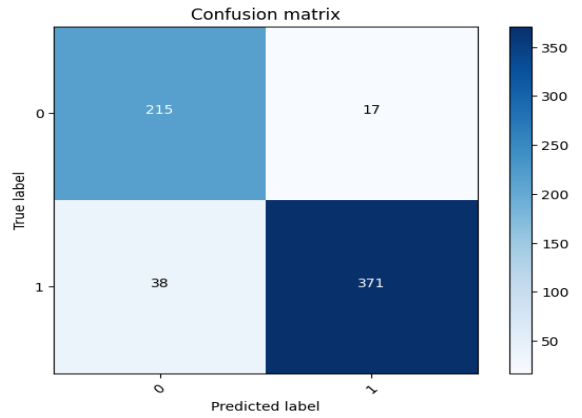
- Loss Function: binary_cross entropy for binary classifications.
- Optimizer: Adam for adaptive learning rate optimization.
- Metric: accuracy for cherishing improvement.

V. RESULT AND DISCUSSION

In this project, we successfully implemented a CNN-based approach to detect image forgeries by examining pixel-level anomalies, texture patterns, and colour inconsistencies. The model provides a quantitative assessment by assigning a manipulation likelihood score, ultimately determining whether an image is real or fake. The final output presents the authenticity of an image in terms of percentage values—indicating both the probability of the image being real and the probability of it being fake.



Training and Validation Accuracy of Epochs




```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 74ms/step
Real Percentage: 16.71%
Fake Percentage: 83.29%
```

Final Output - Real and Fake Percentage of the test image

The output windows confirm the model's efficiency in real-time image forgery detection. The system processes incoming images, analyses key features, and classifies them as real or fake instantaneously. At the end of each analysis, the system presents the percentage likelihood of an image being real or fake.

REFERENCES

[1] Davide Cozzolino, Diego Gragnaniello, Luisa Verdoliva 2013 - Image Forgery Detection Based on The Fusion of Machine Learning and Block-Matching Methods.

[2] Md. Taksir Hasan Majumder, A. B. M. Alim Al Islam 2019 - A Tale of a Deep Learning Approach to Image Forgery Detection.

[3] Junwen Wang, Guangjie Liu, Bo Xu, Hongyuan Li, Yuewei Dai, Zhiquan Wang 2010 - Image Forgery Forensics Based on Manual Blurred Edge Detection.

[4] Mesut kartal, Osman Duman 2022 - Ship Detection from Optical Satellite Images with Deep Learning.

[5] shruti Ranjan, Prayati Garhwal, Anupama Bhan, Monika Arora, Anu Mehra 2018 - Framework for Image Forgery Detection and Classification Using Machine Learning.

[6] A Kuznetsov 2019 - Digital Image Forgery Detection Using a Deep Learning Approach.

[7] Anjali Diwan, and Anil K. Roy. - CNN-Key point Based Two-Stage HybridApproach for Copy-Move Forgery Detection

[8] Kang Hyeon R Generation of Novelty Ground Truth Image Using Image Classification and Semantic Segmentation for Copy-Move Forgery Detection

[9] Chengyou Wang, Zhi Zhang, Qianwen Li, and Xiao Zhou. - An Image Copy-Move Forgery Detection Method Based on SURF and PCET

[10] Ashgan H. Khalil, Atef Z. Ghalwashi, Hala Abdel-Galil El Sayed, Gouda I. Salama, and Haitham A. Ghalwash - Enhancing Digital Image Forgery Detection Using Transfer Learning

[11] Anjali Diwan, Dinesh Kumar, Rajesh Mahadeva, H. C. S. Perera, and Janaka Alawatugoda.- Unveiling Copy-Move Forgeries: Enhancing Detection with Super Point Keypoint Architecture

[12] Fransesco Marra, Diego Gragnaniello, Luisa Verdoliva, and Giovanni Poggi - A Full-Image Full-Resolution End-to-End-Trainable CNN Framework for Image Forgery Detection

[13] Beijing Chen, Ming Yui, QingTangSu, Hiuk Jae Shim, and Yun-Qing Shi. - Fractional Quaternion Zernike Moments for Robust Color Image Copy-Move Forgery Detection

[14] Haipeng Chen, Xiwen Yang, and Yinga Lyu - Copy-Move Forgery Detection Based on Key point Clustering and Similar Neighbourhood Search Algorithm