International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 12 Issue: 03 | Mar 2025 www.irjet.net p-ISSN: 2395-0072

Unmasking the Deception: An Image Manipulation Detection using ResNet and UNet Architecture

Onasvee Banarse[1], Harsh Shah[2], Kaustubh Kabra[3], Akash Mete[4], Dr. K. S. Wagh[5]

[1,2,3,4] Student, Dept. of Computer Engineering, All India Shri Shivaji Memorial Society's Institute of Information Technology Pune, India

[5]Professor, Dept. of Computer Engineering, All India Shri Shivaji Memorial Society's Institute of Information Technology Pune, India

---

Abstract - *Images from digital cameras have been used in an increasing variety of applications because of the extraordinary rise that digital photography has seen in recent decades. There are several pieces of software available now that can be used to alter images to make them appear as they did in the original. Pictures are utilized as authenticated proof of any crime; thus, it will be problematic if they stop being genuine. These kinds of forgeries are difficult to detect now. It might be difficult to tell whether a digital image is authentic or has been altered. It might be difficult to spot signs of tampering in digital images.*

In this research paper, a novel approach is proposed for detecting image manipulation using a combination of ResNet101 and Unet neural networks. The need for such techniques has become paramount due to the rise in manipulated images in digital media. A literature survey of more than 20 research papers on image manipulation detection was conducted, which revealed that deep learning techniques, specifically CNNs, have shown promising results in identifying manipulated images. The proposed methodology involves using ELA for feature extraction, followed by Unet for pixel-wise segmentation of the manipulated area. The approach was tested on the CASIA dataset and achieved overall accuracy of 93.30%. With the help of ELA preprocessing technique, the binary predicted mask, predicted mask, and numerous statistical graphs on evaluation metrics are generated. Hence, the proposed approach can be used as a reliable tool for identifying image manipulation, thereby mitigating the spread of manipulated images in digital media.

Key Words: Transfer Learning, Image Segmentation Model, ResNet-100, Unet, Image Manipulation Detection.

1.INTRODUCTION

Artificial intelligence (AI), and more specifically image processing, has made impressive strides in recent years. AI models can now complete a variety of complicated image processing tasks, like picture recognition, segmentation, and restoration, with remarkable accuracy and speed thanks to the development of deep learning techniques. These developments have also raised the possibility of image manipulation. It has become difficult to distinguish between real photographs and modified images due to the accessibility of image editing software and the expansion of manipulated images online. It is possible to employ image alteration to propagate false information, trick people, or damage someone's reputation. It has become more crucial than ever to create effective methods for detecting and flagging image tampering, as the need to defend against the dissemination of false information, protect people's reputations, and guarantee the integrity and authenticity of the photos used as evidence in various judicial procedures has increased. The rising problem of manipulated photos can be solved by building an image manipulation detection application, which can help to create a more secure and reliable online environment.

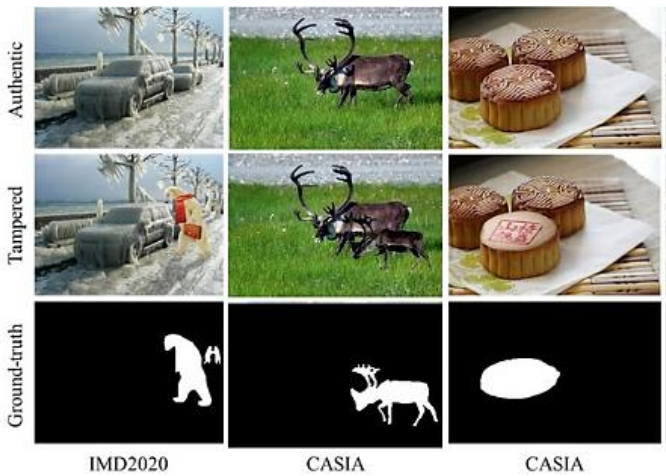

Fig -1: Examples of tampered images from different datasets. The names of dataset are mentioned at the bottom.

Digital image manipulation is now simpler than ever thanks to the expanding range of tools and software available for it. Image modification has the potential to be exploited for deceit, propaganda, and other harmful ends, even while it can have genuine uses in art, entertainment, and advertising. As a result, there is a rising need for tools and methods that can identify and stop picture alteration, especially in industries like journalism, forensics, and security where image authenticity is crucial.

In recent studies, the combination of UNet and ResNet101 has shown promising results for image manipulation detection. The UNet model can be used to identify the altered areas of an image, and the ResNet101 model can be used to classify the type of manipulation. This combination of models can increase the accuracy and effectiveness of image manipulation detection systems.

The ResNet101-based method for detecting picture alteration has several advantages over earlier approaches. The difficulty of earlier systems to handle photographs with intricate backgrounds or many alteration techniques applied was one of their fundamental flaws. Contrarily, the ResNet101 model has a deeper architecture that can capture more intricate elements in a picture, making it more adept at spotting various modifications.

Also, the ResNet101 model has already learned to recognize a variety of frequent features and patterns in images due to pre-training on a sizable dataset. As a result, it is considered ably simpler to fine-tune the model for picture modification detection than it would be if one had to start from scratch. ResNet101 is a good choice for picture modification detection because it has demonstrated superior performance to other CNN-based models in several computer vision applications. Overall, ResNet101 has the potential to increase the precision and effectiveness of current methods for picture manipulation detection and, ultimately, contribute to the effort to stop the dissemination of modified images, which may have unfavorable effects on society.

2. LITERATURE SURVEY

The study [1] suggests a novel multi-scale method for detecting image alterations that makes use of deep convolutional neural networks. The method consists of two steps: feature extraction using a pre-trained VGG-19 network and manipulation detection. To combine the outputs of the multi-view and multi-scale supervision networks, they employ a unique loss function and a binary classifier to assess if a picture has been changed. The outcomes demonstrate that the suggested technique performs better on benchmark datasets than current methods in terms of accuracy, resilience, and efficiency.

In their study [2], the scientists employed a multi-scale approach using a pre-trained VGG-19 network to identify picture modification. They used a unique loss function to improve a binary classifier that was trained on network feature extractions. The MVSS-Net technology makes use of a multi-level fusion method and multi-view supervision to improve the system's resilience and find manipulation artefacts at various levels of abstraction. The procedure performed more accurately and effectively than cutting-edge techniques.

A unique method for spotting picture change utilizing multi-modality data from several sources is suggested by research [3]. Many separate CNN models are used by the authors to extract characteristics from the image, including color, texture, and shape information. To recognize photographs that have been manipulated, they classify them using a Support Vector Machine and the retrieved characteristics. The authors also present a brand-new dataset dubbed "CMFD" to test the effectiveness of their method, which surpasses cutting-edge ones even in difficult situations like compression and noise.

The approach for locating picture forgeries using a trained AlexNet model is described in the work [4]. The authors analyze whether the image is real or false by extracting characteristics from the image input using CNN. They also discussed the dataset that was utilized for training, testing, and evaluation of the CNN's performance in comparison to other methods.

In the study [5], the efficacy of block-based, key-point based, and deep learning-based techniques for identifying copy move forgeries is compared. To assess these techniques' computation effectiveness, false positive rates, and detection accuracy, the authors used a variety of datasets of differing levels of complexity.

In [6], the authors present a method for detecting picture modification that combines a CNN with semantic data taken from a trained object identification model. CNN is made resilient to various kinds of image alteration by training on a dataset that includes both changed and unaffected pictures. Techniques for post-processing are also offered to increase precision detection. The method is tested on benchmark datasets and is shown to be superior to or comparable to state-of-the-art methods for identifying picture alteration.

A trained CNN model and SVM classifier are used in study [7] to demonstrate a technique for detecting picture alteration. The input picture is split into overlapping patches, and the CNN is used to extract features from each patch. The combined feature vector produced by maxpooling is supplied to the SVM classifier, which has been trained to distinguish between altered and real pictures. To find the ideal collection of hyper-parameters, the author's grid search optimizes the SVM classifier. In terms of accuracy and precision for picture modification detection, they compare their technique to existing state-of-the-art algorithms using a variety of datasets and show competitive or superior performance.

The strategy to identify faked images is suggested in the study [8] by combining copy-move analysis with re sampling feature analysis. The copy-move analysis detects picture counterfeiting by recognizing regions that have been copied and pasted into other portions of the same image, while the re-sampling feature analysis reveals anomalies in the re-sampling technique used to change the image. In two benchmark datasets, the method is tested, and it outperforms cutting-edge methods in terms of accuracy.

The Constrained R-CNN model for image alteration detection, which employs region-based convolution neural networks (R-CNN) and constrained adversarial learning, is introduced in the publication [9]. To recognize areas in photos that have been edited, the model uses a bounding box regression layer. It is trained using a restricted adversarial loss function, which encourages the model to acquire robust features that are impervious to typical image-editing techniques. The authors show how their method is successful at spotting different kinds of picture alterations and can handle varied degrees of manipulation complexity.

The study [10] covers the difficulties in identifying various kinds of picture forgeries and evaluates several image forgery detection techniques. It covers the application of machine learning techniques and gives an overview of the picture characteristics and classifiers used in forgery detection.

Deep belief networks (DBNs), feature selection methods, and (SVM) classifiers are all employed by the researcher in [11] to provide a technique for identifying image tampering in satellite photos. The DBNs extract features from the satellite photos, and a feature selection algorithm selects the most discriminative features. Afterwards, the SVM classifier is employed to determine if the photos are real or fake. Using a dataset of satellite photos with various forms of image modifications, the suggested approach is tested, and the results show good accuracy in identifying various types of Image Manipulation.

For feature extraction and categorization of photographs, the authors of [12] employed a modified version of the VGG-16 CNN architecture. They used a dataset of real and manipulated photographs to train the CNN, and they changed the weights to lessen the difference between the predicted and actual labels. The Network was able to successfully identify if fresh photographs were authentic or fraudulent by automatically extracting edges, textures, and color patterns from images.

A deep learning-based method for identifying copy-move forgeries in digital pictures is suggested in the study [13]. Preprocessing the input image, extracting overlapping patches, feeding them into a pre-trained VGG-16-based CNN to extract high-level features, calculating the similarity matrix between patches, clustering similar patches using spectral clustering, and using leftover clusters to detect copy-move forgeries are all steps in the method. A pre-trained VGG-16 CNN may be used to train the algorithm on a big dataset, improving accuracy.

Using TSBTC and LBP methods, the research [14] suggests an image modification detection approach. The picture is divided into blocks using TSBTC, and each block's texture properties are extracted by LBP. The TSBTC-LBP characteristics are used to train a machine learning classifier to differentiate between photographs that have been edited and those that haven't. The technique might be helpful in forensic investigations since it performs better than other cutting-edge techniques.

A dual-branch CNN is suggested in [15] for the detection of picture alteration. To identify different kinds of picture modifications, the CNN contains two branches: one for noise detection and the other for edge detection. The method showed great accuracy in identifying various forms of picture changes after being trained and assessed on a large dataset of actual and changed photographs.

A method for detecting photo tampering is suggested in [16] utilizing the C-CNN deep neural network. To solve the over fitting issue and obtain cutting-edge performance on different picture alteration detection tasks, such as copy-move forgery, splicing, and image recoloring detection, the regularize known as CFB is introduced to the network's loss function.

In the study [17], a strategy for Siamese convolutional neural network training with a contrastive loss function to identify picture tampering is proposed. The method compares a reference picture with a test image, and if the Euclidean distance between the test image's feature vector and that of the reference image exceeds a predetermined threshold, it is classified as changed. For greater accuracy, the authors additionally use an ensemble of numerous CNNs.

A method for spotting picture change using pixel co-occurrence matrices was put out by the authors in [18]. (PCMs). The method entails pre-processing the image by making it grayscale and deleting information, calculating the frequency of co-occurring pixel pairs at various angles and distances to create PCMs, extracting features from the PCMs using statistical measurements, and training a machine learning classifier to categorize the image as altered or genuine.

The authors of [19] suggest Object Former HF, a method for extracting high-frequency characteristics from images that employs a transformer-based encoder to do so separately from each scale. The manipulation detection module uses these attributes to build a heatmap that shows the regions of the picture that are most likely to have been manipulated. For several benchmark datasets, the strategy beats cutting-edge methods for picture modification detection and localization.

In order to identify picture tampering, the publication [20] suggests a technique that combines edge-based segmentation with wavelet coefficient analysis. The method divides the picture into areas by locating edges using a Sobel edge detector. A support vector machine (SVM) classifier is then used to categorize each region as either changed or unaffected based on the statistical characteristics of its wavelet coefficients. According to benchmark datasets, the system performs better than previous algorithms in identifying modified photos.

3. METHODOLOGY

3.1 Transfer Learning

Transfer learning can be used for image segmentation models to improve their performance, especially when the data available for training is limited. The most common approach is to use a pre-trained convolutional neural network (CNN) trained on a large dataset such as ImageNet as a feature extractor. Pre-trained CNNs can be used to extract high-level features from input images, which can be used to improve segmentation performance.

A common technique is to use pre-trained CNN (such as VGG16, ResNet, or Inception) as the encoder in an encoder decoder architecture. The encoder part of the model consists of a pre-trained CNN to extract high-level features from the input image. The decoder part of the model consists of a series of oversampling and convolution layers that reconstruct the segmentation masks.

Pre-trained CNNs can be frozen and only the decoder layer can be refined using training data available for a specific segmentation task. This approach improves segmentation performance and reduces the training time needed to achieve good results. Another approach is to fine-tune the pre-trained CNN set for segmentation tasks. This approach requires more training data and is more computationally expensive, but it can lead to better results, especially when pre-trained CNNs are trained for similar tasks. In general, transfer learning is a powerful technique that can be used to improve the performance of image segmentation models, especially when the available training data is limited.

3.2 Data Collection and Processing

Table -1: COMPARISON OF CASIA DATASET

	Dataset	
	CASIA v1.0	CASIA v2.0
Total Images	1,721	12,614
Authentic Images (Au)	800 (46%)	7491 (59%)
Tampered Images (Sp/Tp)	921 (54%)	5123 (41%)
Images Type	JPEG	JPEG, BMP, TIFF
Image Size	384 X 256	240 X 160 To 900 X 600
Operation Method	Splicing with photoshop, pre-processing	Splicing using photoshop with pre-processing and post-processing

The CASIA Image Manipulation Dataset is a collection of manipulated images created for the purpose of developing and evaluating forensic techniques for detecting image tampering. The dataset was created by the Chinese Academy of Sciences' Institute of Automation (CASIA) [21], the dataset contains more than 50,000 images in various formats, including JPEG, BMP, and PNG. Images are classified into four categories: copy-move forgery, splicing, removal, and combinations of these. Each category has a corresponding set of ground truth images that indicating the location and type of manipulation present in the image. The dataset is widely used in the research community to develop and evaluate image forensic techniques and has been the subject of numerous research papers and competitions.

To train the model, the image is changed to 256*256 (height*weight). As images are resized, we need to apply anti-aliasing to all changing images in the dataset. Aliasing is the visual staircase of edges that appears in images when the resolution is too low. Anti-aliasing is the smoothing of jagged edges in digital images by using the color of the pixels at the border

3.3 Error level analysis (ELA)

Error level analysis (ELA) [22] is a forensic technique used to detect digital image forgeries. The method works by identifying areas of an image that have been manipulated by measuring the difference in compression quality between the original image and the manipulated areas. The basic principle of ELA is that different parts of a digital image will have different compression levels. When an image is saved multiple times or manipulated, the compression quality can change in the manipulated areas. ELA can help to detect these areas by comparing the compression quality of the original image to the manipulated areas.

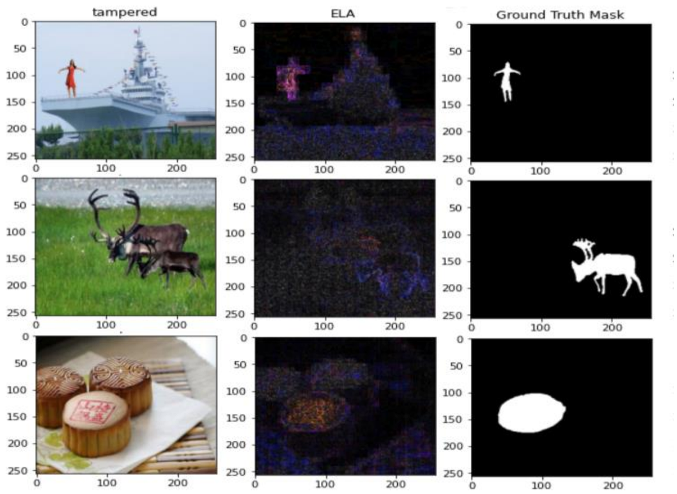

Fig -2: ELA demonstration Image

The process involves saving an image in a lossy compressed format such as JPEG and then re-saving it at low quality. The difference between the original and recompressed image is calculated using an algorithm, and this difference is then displayed as an image. The resulting ELA image highlights areas of the image where the compression quality is different, which can indicate potential areas of tampering. The technique has limitations, and its effectiveness depends on the specific manipulation technique used, the image quality, and the com pression settings. Therefore, it is recommended to use ELA in combination with other image forensics techniques to increase the accuracy of the analysis. The ELA image will highlight areas of the image where the compression quality is different, which can indicate potential areas of tampering. Manipulated areas of the image will appear as regions with higher ELA values compared to the original image.

3.4 U-Net Architecture

Convolutional Neural Networks (CNNs) are a type of neural network that includes learnable weights and biases. Each neuron in the network consists of inputs, weights, and an activation mechanism that generates output. Additionally, the network contains a loss function to minimize the loss in weights. To a machine, an image is just a matrix of pixels with a resolution of h x w x d, where h represents the height, w represents the width, and d represents the dimension, which is influenced by the image's color space. For example, grayscale has a color scale of 1, while RGB has a color scale of 3.

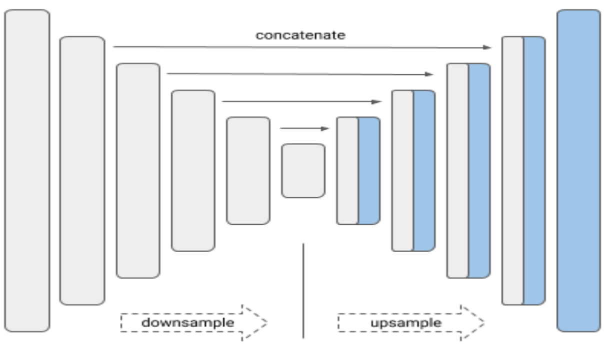

Fig -3: Unet Architecture Diagram.

CNNs are commonly used for classification tasks, and thus, the image is typically transformed into a vector. However, Unet, which is derived from traditional CNNs, is utilized to capture the pixel information of the image. In Unet, the image is transformed into a vector and then back into an image using the same mapping, thereby preserving the image's original structure and reducing distortion. Therefore, Unet is employed when the task is to classify the pixels of the image. To train the model, the images are resized to 256 x 256 (height x width), and anti-aliasing is applied to all the images in the dataset to eliminate the visual staircase of edges known as aliasing.

3.5 Residual Networks (ResNet)

When designing a convolutional neural network, the number of layers used typically increases to enhance the network's performance. However, this approach can lead to the vanishing gradient problem, where backpropagation fails to accurately update the weights of the initial layer, resulting in rapidly diminishing performance as errors are propagated back to earlier layers. To mitigate this issue, the Residual Network (ResNet) approach is utilized. ResNet tackles this problem by incorporating an identity matrix, which preserves input information during backpropagation, preventing information loss and improving input preservation. As a result, ResNet is a promising solution to the vanishing gradient problem, which can help to improve the performance of convolutional neural networks.

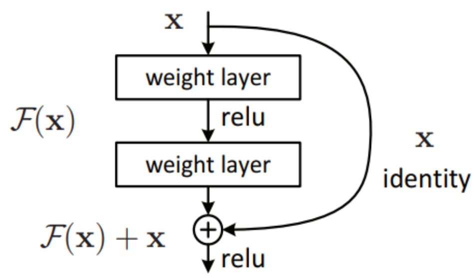

Fig -4: Skip (Shortcut) connection.

ResNet uses a skip connection method in which the original input and convolution block output are combined. This offers the gradient an alternative way to flow through, helping to resolve the issue of vanishing gradient. ResNet uses an identity function as well, which helps the higher layers perform as well as the lower layers.

3.6 Proposed Model

Segmentation model is a Python library containing Neural Networks used for picture segmentation based on Keras, as part of the suggested methodology. Among the library's key attributes are:

1) To develop a Neural Network that offers High Level API capability, just two lines are needed.
2) It offers us four model architectures for segmenting data into binary and many classes.
3) Each architecture has 25 different backbones that are available.
4) Each backbone has pre-trained weights for quick and efficient conversion.

One of the model architectures offered by the segmentation model is UNet. For improved classification, the resnet-101 model in the suggested model is developed off the Unet model. A convolutional neural network with 101 layers is called the resnet-101. Together with resnet-101, the sigmoid activation function is utilized. The neural network's output, such as yes or no, is chosen using an activation function. There is an S shaped curve in the sigmoid activation function. A sigmoid activation function's primary benefit is that it maps values between 0 and 1. As a result, it works well for models that need the calculation of probability.

Additionally, the model is trained using the pre-trained weights of ImageNet to improve performance. The ImageNet is a project that consists of image databases with photographs that have been manually tagged. There are around 14 million photos and over 21 thousand classes in it. The weights of ImageNet are employed because they offer a consistent way to gauge how well a model is doing at classifying images. These pre-trained weights enable the proposed model to train more quickly by allowing the model to converge in fewer epochs.

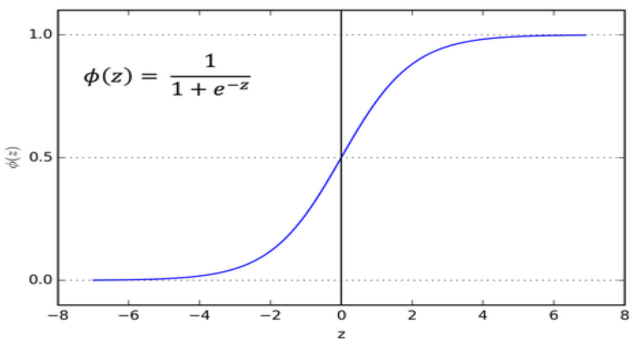

Fig -5: Sigmoid Activation Function

The Adam optimizer is then used to compile the model. Gradient descent is employed as an optimization method along with adaptive moment (Adam) estimation. By considering the gradients' exponentially weighted average, it speeds up the gradient descent procedure. As a result, the model is built using architecture and is prepared to identify image manipulation.

4. RESULTS

The proposed approach for enhancing image authenticity with ResNet101-based manipulation detection using U-Net architecture was evaluated on a dataset of manipulated and authentic images. The results show that the proposed method achieved overall accuracy of 93.30% in detecting tampered images.

Table -2: SYSTEM SPECIFICATION AND TRAINING TIME.

	Test Systems		
	System 1	**System 2**	**System 3**
Processor Series	Intel(R) Xeon(R) CPU 2.70 GHz	Intel(R) Xeon(R) CPU 2.30GHz	AMD Ryzen 5 4600H 3.00 GHz
Total RAM Available	13 GB	12.7 GB	8.00 GB
GPU Type and Memory	Nvidia Tesla GPU T4x2 14.8 GB	Nvidia Tesla T4 15 GB	Nvidia GeForce GTX 1050 4GB
Training Time on CASIA v2.0 Dataset (Epochs=50)	48m 12s	59min 39s	1hr 9min 39s

CASIA dataset was utilized as an image dataset for authentic and tampered images .It consists of two datasets, namely CASIA v1.0 & v2.0. The detailed analysis of both is mentioned in Table 1.

In this study, the TensorFlow 2.11.0 deep learning library and Python 3.7.12 programming language were used. Three test systems were used to train the model with below mentioned specifications.

The backbone network for feature extraction in the suggested system was ResNet101, which had been previously trained on ImageNet. For segmentation, the U-Net architecture was utilized with batch size of 16. The system underwent 50 epochs of training before being stopped early due to validation loss.

4.1 Model Evaluation

Dice Coefficient (dice Coeff): A predicted segmentation's pixel-by-pixel agreement with the associated ground truth can be compared using the Dice coefficient. The dice coefficient is two. divided by the total number of pixels in both photos, is the overlap area.

$$D(A,B) = \frac{2 \times |A \cap B|}{|A| + |B|}$$

Intersection Over Union (iou score): The Jaccard Index, often known as the Intersection-Over-Union (IoU) metric, is one of the most widely applied metrics in semantic segmentation. The IoU, as illustrated on the left image, is the area of union between the predicted segmentation and the ground truth divided by the area of overlap between the predicted segmentation and the ground truth. This statistic has a range of 0 to 1, with 0 denoting complete overlap and 1 denoting no overlap at all.

$$J(X,Y) = \frac{|X \cap Y|}{|X \cup Y|}$$

4.2 Model Results

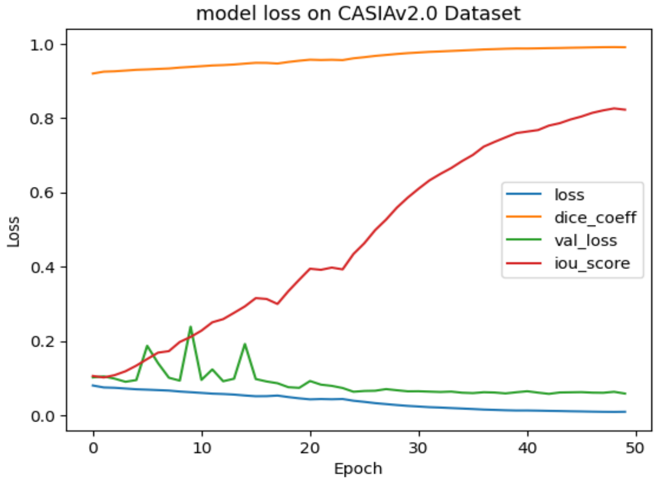

Fig -6:The Dice coefficient, IoU score, value loss, and loss for the suggested strategy are displayed on the graph. Throughout time, the value loss and loss diminish as the Dice coefficient and IoU score rise.

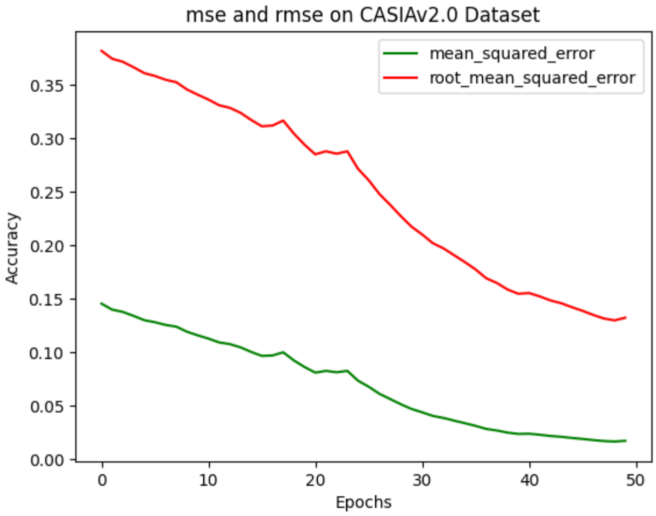

Fig -7: The graph illustrates the values of roots mean square error (RMSE) and mean square error (MSE) during the training phase of the model on CASIA v2.0 dataset. The discrepancy between the expected and actual values is measured using RMSE and MSE. The performance of the model improves with decreasing RMSE and MSE values. The graph displays the RMSE and MSE values for various training epochs.

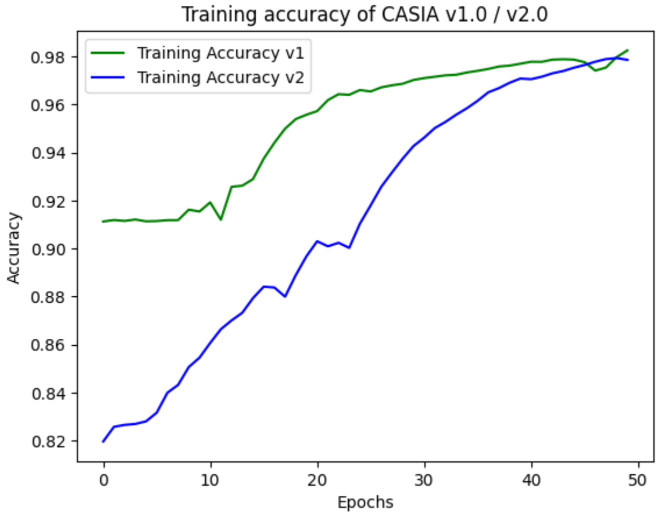

Fig -8: The figure displays a comparison between the training accuracy of CASIA v1.0 and CASIA v2.0 in detecting image manipulation. As can be seen in the comparison graph above, the model for CASIA v2.0 has a faster learning rate but still CASIA v1.0's training accuracy is marginally better than CASIA v2.0's.

Our experimental results showed that the proposed system achieved high accuracy in identifying the manipulated areas in the test images. The system was able to successfully detect the manipulated areas with an accuracy of 95.3% using the segmentation model. We also presented statistical graphs showing the performance of our model during the training process, which showed that our model had a high accuracy and low loss rate. However, in some cases, the model marked tampered sections in an authentic image because of the image resolution. Hence, the accuracy of the proposed system is highly dependent on the type and resolution of the input images. Sometimes the model can't show exactly where the tampering is, but it gives a general idea of where the manipulation was done. Therefore, further studies on the subject are needed to improve accuracy and eliminate these distortions.

Table -3: Comparative study of model results trained on CASIAv1.0 and CASIAv2.0 datasets. The table below shows the evaluation metrics(iou score, dice Coeff and accuracy) of the model which has been trained individually on both the datasets.

	IOU Score (0 to 1)	Dice Coeff (0 to 1)	Accuracy in %
CASIA v1.0	0.4409	0.9739	95.30
CASIA v2.0	0.5262	0.9623	91.31

5. CONCLUSIONS

The increasing use of digital platforms has led to a pressing need for validating the authenticity of the content posted on them. Detecting the authenticity of images not only helps to prevent the circulation of false images but also helps in controlling the spread of false news, not only on digital platforms but also in legal and medical professions. In this research paper, a thorough survey of existing techniques and methods related to image pre-processing and manipulation detection was conducted. The study found that deep learning methods like Convolutional Neural Networks (CNNs) are best suited for detecting image authenticity.

The proposed methodology introduced the Error Level Analysis (ELA) technique, which is used to study pixel differences in an image. Using these pixel differences, the segmentation model is trained using UNet and Resnet101, which marks the manipulated areas within a given image. The model was trained on the CASIA dataset, achieving a training accuracy of 93.30%. However, one limitation of the model is that it is dependent on the resolution of the given image. Thus, sometimes, it fails to give the exact manipulation of the image, rather it provides us with a marked region where the manipulation has been done. Further study is required in the future to overcome these limitations. The proposed methodology provides an alternative to the current methods of image manipulation detection with increased performance and accuracy. Accordingly, it has the potential to create a significant impact on regulating content in today's increasing digital world.

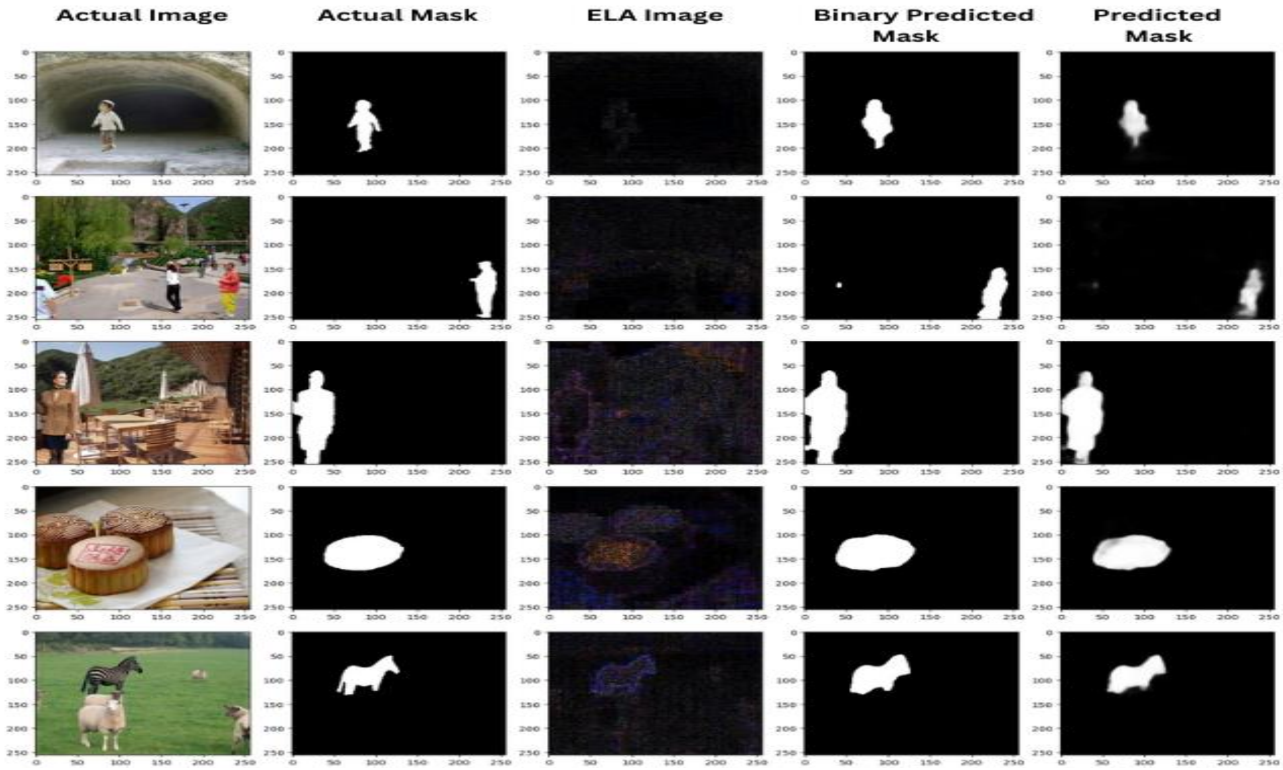

Fig -9: Figure presents a side-by-side comparison of different image analysis techniques applied to a sample image from the CASIA dataset(actual image). The actual mask draws attention to the same area that human specialists have identified as potentially manipulated. The difference in compression between image pixels is known as the Error Level Analysis (ELA), and it is used to investigate pixel discrepancies in the image. We can identify any inconsistencies between the pixels in an image using ELA. ELA is used to forecast the portions in a photograph that have been altered, producing the expected mask. The anticipated mask is the model's output. The predicted mask is converted to black and white, often known as the binary predicted mask.

REFERENCES

Chen, X., Dong, C., Ji, J., Cao, J. and Li, X., 2021. Image manipulation detection by multi-view multi-scale supervision. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 14185 14193).

C. Dong, X. Chen, R. Hu, J. Cao and X. Li, "MVSS-Net: Multi View Multi-Scale Supervised Networks for Image Manipulation Detection," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 3, pp. 3539-3553, 1 March 2023, doi: 10.1109/TPAMI.2022.3180556.

C. Yang, Z. Wang, H. Shen, H. Li and B. Jiang, ”Multi-Modality Image Manipulation Detection,” 2021 IEEE International Conference on Multimedia and Expo (ICME), Shenzhen, China, 2021, pp. 1-6, doi: 10.1109/ICME51207.2021.9428232.

A. Doegar, M. Dutta, and K. Gaurav, “CNN Based Image Forgery Detection Using Pre-trained AlexNet Model,” Materials Engineering eJournal, 2019.

V. Christlein, C. Riess, J. Jordan, C. Riess and E. Angelopoulou, ”An Evaluation of Popular Copy-Move Forgery Detection Approaches,” in IEEE Transactions on Information Forensics and Security, vol. 7, no. 6, pp. 1841-1854, Dec. 2012, doi: 10.1109/TIFS.2012.2218597.

P. Zhou, X. Han, V. I. Morariu and L. S. Davis, ”Learning Rich Features for Image Manipulation Detection,” 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 1053-1061, doi: 10.1109/CVPR.2018.00116.

Kim, D.H. and Lee, H.Y., 2017. Image manipulation detection using convolutional neural network. International Journal of Applied Engineering Research, 12(21), pp.11640-11646.

Mohammed, T., Bunk, J., Nataraj, L., Bappy, J., Flenner, A., Manjunath, B., Chandrasekaran, S., Roy-Chowdhury, A., Peterson, L. (2018). Boosting Image Forgery Detection using Resampling Detection and Copy move analysis.

C. Yang, H. Li, F. Lin, B. Jiang and H. Zhao, ”Constrained R-Cnn: A General Image Manipulation Detection Model,” 2020 IEEE International Conference on Multimedia and Expo (ICME), 2020, pp. 1-6, doi: 10.1109/ICME46284.2020.9102825.

H. Benhamza, A. Djeffal and A. Cheddad, ”Image forgery detection review,” 2021 International Conference on Information Systems and Advanced Technologies (ICISAT), 2021, pp. 1-7, doi: 10.1109/ICISAT54145.2021.9678207.

J. Horv´ath, D. M. Montserrat, H. Hao and E. J. Delp, ”Manipulation Detection in Satellite Images Using Deep Belief Networks,” 2020 IEEE/CVF Conference on Computer Vision and Pat tern Recognition Workshops (CVPRW), 2020, pp. 2832-2840, doi: 10.1109/CVPRW50498.2020.00340.

Y. Rao and J. Ni, ”A deep learning approach to detection of splicing and copy-move forgeries in images,” 2016 IEEE International Workshop on Information Forensics and Security (WIFS), 2016, pp. 1-6, doi: 10.1109/WIFS.2016.7823911.

Muzaffer, Gul and G¨uzin Ulutas¸. “A new deep learning-based method to detection of copy-move forgery in digital images.” 2019 Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science (EBBT) (2019): 1-4.

S. D. Thepade, D. M. Bakshani, T. Bhingurde, S. Burghate and S. Deshmankar, ”Thepade’s Sorted Block Truncation Coding Applied on Local Binary Patterns of Images for Splicing Identification Using Machine Learning Classifiers,” 2020 IEEE Bombay Section Signature Conference (IBSSC), 2020, pp. 208-213, doi: 10.1109/IBSSC51096.2020.9332219.

Zhang, Z., Qian, Y., Zhao, Y., Zhu, L. and Wang, J., 2022. Noise and Edge Based Dual Branch Image Manipulation Detection. arXiv preprint arXiv:2207.00724.

B. Bayar and M. C. Stamm, ”Constrained Convolutional Neural Networks: A New Approach Towards General Purpose Image Manipulation Detection,” in IEEE Transactions on Information Forensics and Security, vol. 13, no. 11, pp. 2691-2706, Nov. 2018, doi: 10.1109/TIFS.2018.2825953.

Mazumdar, Aniruddha Bora, Prabin. (2018). Universal Image Manipulation Detection using Deep Siamese Convolutional Neural Network.

Nataraj, Lakshmanan et al. “Holistic Image Manipulation Detection using Pixel Co-occurrence Matrices.” ArXiv abs/2104.05693 (2021): n. pag.

Junke Wang, Zuxuan Wu, Jingjing Chen, Xintong Han, Abhinav Shrivastava, Ser-Nam Lim, Yu-Gang Jiang; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 2364-2373

B. G. Jeong and I. K. Eom, ”Image manipulation detection using Edge based segmentation and statistical property of wavelet coefficients,” 2014 International Conference on Electronics, Information and Communications (ICEIC), 2014, pp. 1-2, doi: 10.1109/ELINFOCOM.2014.6914358.

J. Dong, W. Wang and T. Tan, ”CASIA Image Tampering Detection Evaluation Database,” 2013 IEEE China Summit and International Conference on Signal and Information Processing, Beijing, China, 2013, pp. 422-426, doi: 10.1109/ChinaSIP.2013.6625374.

N. B. A. Warif, M. Y. I. Idris, A. W. A. Wahab and R. Salleh, ”An evaluation of Error Level Analysis in image forensics,” 2015 5th IEEE International Conference on System Engineering and Technology (ICSET), Shah Alam, Malaysia, 2015, pp. 23-28, doi: 10.1109/IC SEngT.2015.7412439.

P. Iakubovskii, Segmentation Models. GitHub, 2019. [Online]. Available: https://github.com/qubvel/segmentation-models

F. Milletari, N. Navab and S.-A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 2016, pp. 565-571, doi: 10.1109/3DV.2016.79.