IMAGE AUTHENTICITY VERIFICATION

Mrs.B.Mahalakshmi[1], Panchvae Akshad Singh[2], Laasya Deshagani[3], G.satya Prakash[4], Muppu Surya[5]

[1]Associate Professor, Dept of CSE, Sreyas Institute of Engineering and Technology

[2]Ug scholar, Sreyas Institute of Engineering and Technology

[3]Ug scholar, Sreyas Institute of Engineering and Technology

[4]Ug scholar, Sreyas Institute of Engineering and Technology

[5]Ug scholar, Sreyas Institute of Engineering and Technology

Corresponding Author: - *B.Mahalakshmi*

Associate professor, Dept of CSE, Sreyas Institute of Engineering and Technology

Abstract: ***The goal of this project is to create a robust system that uses advanced learning methods and image analysis techniques to detect fa ke images. As digital media continues to evolve and image management becomes easier, the need for reliable methods to identify or iginal images and prevent misinformation is increasing. The project addresses this need by collecting data from real and fake images, improving their quality through preprocessing, and using con volutional neural networks (CNN) to detect conflicts in art. The performance of the system will be evaluated using metrics such as accuracy, precision, recall, and F1score, and will be compared with existing methods to demonstrate its beauty.***

I. INTRODUCTION

In today's digital age, the rapid growth of images on social media platforms, news sites, and other online sites has led to the emergence of an important form of communication today. These images have tremendous power to inform, persuade, and influence public discourse. However, the emergence of efficient and easy to use photo editing software has reduced the possibility of creating b ad or fake photos. The increase in surveillance on social media has led to serious problems such as the spread of false information, the undermining of public trust, and the use of social narratives.

In journalism, fake images can deceive the public by creating a false message. In the context of law, tampering with evidence can lead to a miscarriage of justice. Additionally, the use of altered images for commercial or personal criticism can result in financial or reputational damage. Social media platforms make the problem worse by allowing this content to spread rapidly, making it difficult to find and mitigate. Advanced automated methods are needed to verify the ac curacy of images. Manual measurement, while sometimes useful, does not measure or reliably measure the complexity or productivity of the work. The project aims to use the power of machine learning g and image analysis technology to solve these problems, focusing on neural networks (CNNs). CNNs are well suited to this task because they are good at analysing image data and detecting patterns and defects that are not visible to the human eye. Many sources have contributed to digital forensics. The project aims to increase trust in digital media and create a foundation for future growth in the sector by developing powerful tools and strategies s for detecting the authenticity of images.

II.OBJECTIVES AND METHODOLOGY

The main goal of the project is to create a competition that can be trusted to distinguish between real and fake images. This includes creating a framework that can detect anomalies that indicate image manipulation. To achieve this goal, the project focuses on sever all key objectives: generating and presenting high-quality data on real and fake images, analysing models using neural l networks (CNN), and measuring performance metrics such as accuracy, precision, recall, and F1 Goal. The system's capabilities will also be compared with existing systems to demonstrate its advantages and adaptability. This difference is important for training mode ls optimized for real situations. This step makes the input CNN model similar. CNNs are particularly good for this task because they can learn locations in image data. Metrics such as accuracy, precision, recall, and F1 score are calculated to evaluate whether it can correctly recognize false images. This comparison demonstrates the development and quality of the new method and evaluates its contribution to popularization. This is useful for verifying the correctness of images.

III. LITERATURE SURVEY

The rapid development of digital media has led to advances in the creation and presentation of art. These fake images pose a serious threat to the accuracy and reliability of online content. Therefore, re searchers are actively seeking new technologies to solve this problem. The main contributions in this field can be divided into statistic al models, machine learning models, deep learning models, and hybrid models. Each method has its advantages and limitations, as dis cussed below. For example, they can detect lighting inconsistencies s, noise patterns, or colour shifts. These techniques are particularly useful for analysing simple forms of interference, such as clustering or propagation. However, their effectiveness decreases when dealing with multiple variations, such as those generated by deepfakes or advanced algorithms such as generative adversarial networks (GA Ns). The limitations of traditional methods suggest that a more diverse approach is needed. These techniques often rely on inhouse developed methods (such as edge detection, colour analysis, and texture models) to identify patterns. While these methods are promising in terms of discovering specific controls, they require extensive engineering that can be time consuming and prone to human error. They also have limited ability y to adapt to a variety of tasks, making them less effective at changing and evolving. Neural networks (CNNs) have transformed image analysis using subtraction techniques. CNNs can learn spatial hierarchies directly from raw images, making them useful for identifying complex functional patterns that traditional methods may miss. Recent studies have demonstrated the superiority of CNNs in comp lex search tasks, including those generated by GANs. CNNs have become the backbone of modern image recognition due to their ability to recognize small details and patterns. Model. For example, traditional statistics can be used as input to improve the ability of deep learning models to detect defects. This integration increases the accuracy of traditional methods and supports the development of deep learning to increase accuracy and robustness. Hybrid methods are e specially promising for situations where computational efficiency and accuracy must be balanced. Analysing advanced tasks such as t hose generated by state-of-the-art GANs remains a challenging task. Sustainability of invisible products is another important challenge because the actual performance often differs from the information shown. Also, measuring the time spent sharing on social media platforms or other crowded websites requires more research and innovation. While working on these issues, take a few photos every day to demonstrate the problems. The program focuses on continuous efforts to ensure the accuracy of digital images by focusing on advanced technologies and solving current limitations.

IV. Proposed System

The proposed system is built around a convolutional neural network (CNN)-based architecture specifically designed to detect artifacts in digital photography. CNNs are well suited for this task because they can learn and extract important features from object images. The system's architecture consists of the following components:

Input layer:

Using pre-processed images as a network concept. Images are standardized and transformed to ensure size and type consistency and provide design ideas for models. For example, they can detect jagged edges , irregular patterns, or irregularities that indicate manipulation. Con volutional layers apply filters to the input data to extract those features that are important for distinguishing real images from fakes. . This helps preserve the most important features while reducing the complexity of the computer. These layers allow the structure to focus on the main structure and not make it redundant. Make predictions. This layer allows the system to isolate features to better understand the image. The output is usually a reliable representation of the image. It consists of the following steps:

Input information: The system starts with a profile containing the correct and complete image. The images are carefully collected and compiled to create a library with many activities. Advanced technology. These include:

Normalization: Normalization of pixel values to a standard range (typically [0, 1]) to help speed up and stabilize learning. Dimensions according to the CNN policy. The processed images are fed in to the CNN model for training. During this phase, the model learns patterns and features that distinguish real images from manipulate d images. The training process involves optimizing test examples to minimize classification errors. During inference, the model processes the input image and produces a confidence score that estimates the accuracy of the image and measures the probability that the image is correct or incorrect. Know the image you want to create. The combination of powerful CNN tools and efficient algorithms lays the foundation for reliable and scalable deployment while also solving the challenges of today's image analysis

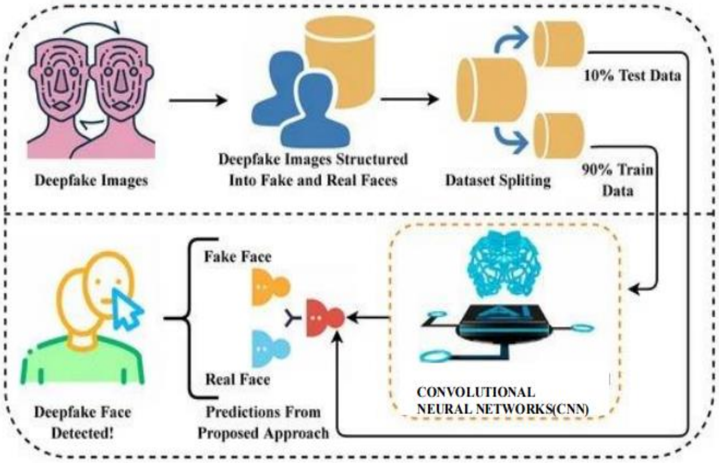

Deepfake Images
Deepfake Images Structured Into Fake and Real Faces
Dataset Spliting
10% Test Data
90% Train Data
Fake Face
Real Face
CONVOLUTIONAL NEURAL NETWORKS(CNN)
Deepfake Face Detected!
Predictions From Proposed Approach

Fig 1: Architecture Diagram

This diagram represents a *Deepfake detection system* using a *Convolutional Neural Network (CNN). The work is divided into several stages:

1. *Introductory level:*
- Post the text from the Deepfake images as comments. The images are divided into two groups: *fake faces* and *real faces*.

2. *Dataset structuring and segmentation:*
- The dataset is created and organized for training and testing purposes. To ensure good evaluation of the model, the data is split into *90% training data* and *10% testing data*.

3. *Using CNN for Deepfake Detection:*
- Use CNN model to process the data. The trained network can determine whether the displayed image is a *fake face* or a *real face* based on the learned features.

4. *Reporting Process:*
- CNN's prediction is used to classify faces as "fake" or "real". If a deep hole is detected, it will be flagged accordingly.

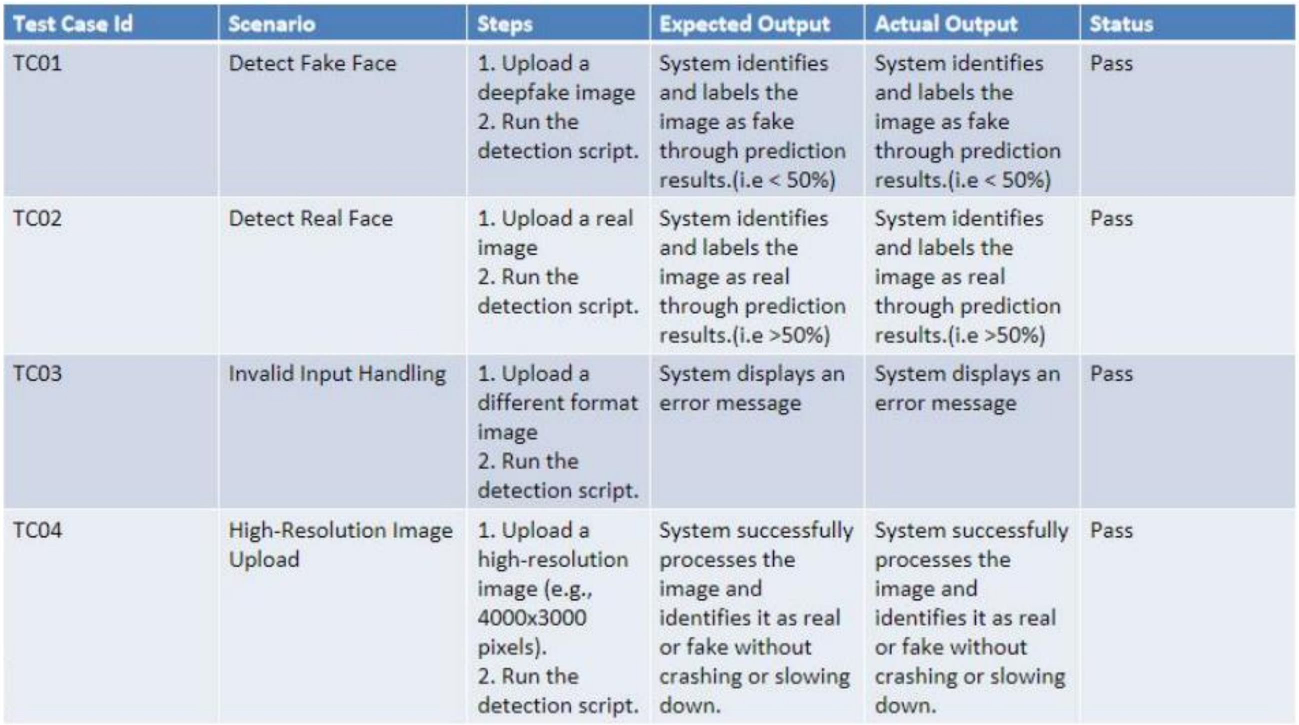
The architecture uses the power of CNNs to extract useful patterns and detect subtle differences in faces, making it more efficient in identifying details in images.

.

V. IMPLEMENTATION

Test Cases:

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product. It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of tests. Each test type addresses a specific testing requirement. The testing process involves several stages, including unit testing, integration testing ,application testing and security testing.



Test Case Id	Scenario	Steps	Expected Output	Actual Output	Status
TC01	Detect Fake Face	1. Upload a deepfake image 2. Run the detection script.	System identifies and labels the image as fake through prediction results.(i.e < 50%)	System identifies and labels the image as fake through prediction results.(i.e < 50%)	Pass
TC02	Detect Real Face	1. Upload a real image 2. Run the detection script.	System identifies and labels the image as real through prediction results.(i.e >50%)	System identifies and labels the image as real through prediction results.(i.e >50%)	Pass
TC03	Invalid Input Handling	1. Upload a different format image 2. Run the detection script.	System displays an error message	System displays an error message	Pass
TC04	High-Resolution Image Upload	1. Upload a high-resolution image (e.g., 4000x3000 pixels). 2. Run the detection script.	System successfully processes the image and identifies it as real or fake without crashing or slowing down.	System successfully processes the image and identifies it as real or fake without crashing or slowing down.	Pass

Tools and methods

The use of the scheme requires a combination of programming languages, libraries and hardware to facilitate development and completion. The main tools and technologies used are:

Programming languages:

Python: chosen for its extensive library and ease of use in machine learning and photography. and frameworks:

TensorFlow and PyTorch: These deep learning techniques are used to design, train and fine-tune CNN models. They provide prebuilt modules for convolutional layers, activation functions and opt imitation algorithms. :

GPU performance: The speed of the CNN model training process is crucial, especially when dealing with large data and complex processes. Preparation: The first step is to create a correct and
complete image archive. Images are collected from public repositories and libraries to provide a variety of content and ideas. Each image is carefully labelled as “real” or “fake” to create a good tracking rec ord. The first step is to use:

Resize: resize all images to the size required by the CNN mode l. Increase the power and speed of the training process. This makes the model more diverse and helps prevent overfitting by exposing the model to different data during training. Layer :

Convolutional layer: captures spatial features and patterns representing the process. . Iteratively optimize biases to reduce classification error. This allows for a fair assessment of its performance. Th e main metrics used to evaluate the model include:

Accuracy: percentage of all images classified. Ratio: The ratio of the true estimated quality to each quality. This approach makes the model more robust, better for invisible objects, and can complete the picture.

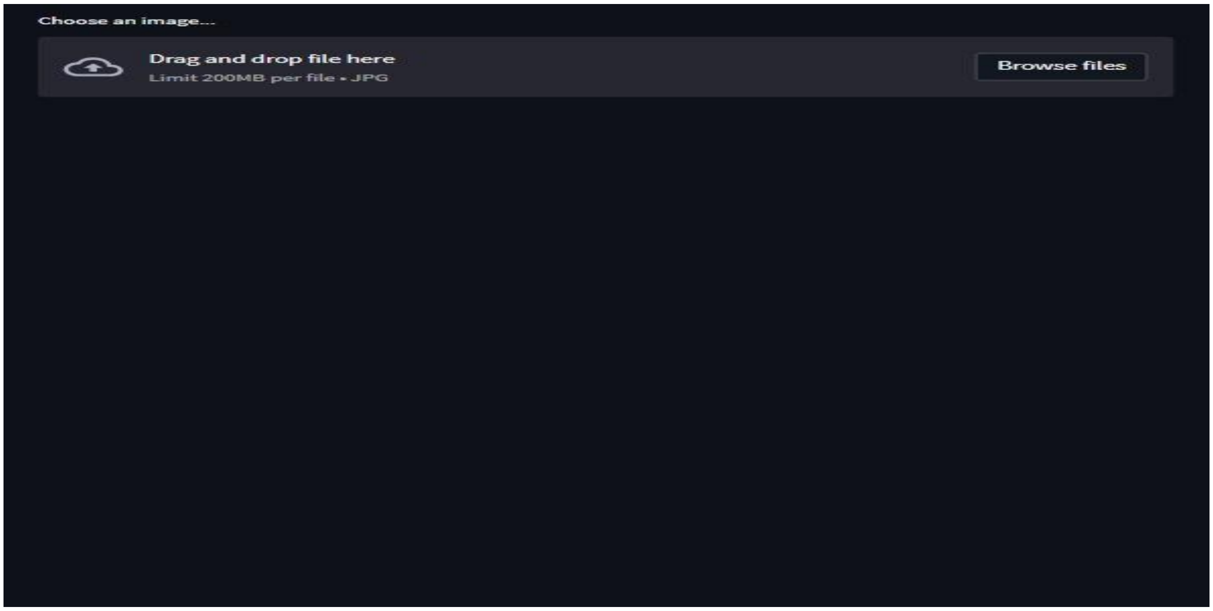


Fig 1: Input for the image Authenticity Verification

The Image Authentication Project's input includes a large number of images in formats such as JPEG, taken from various platforms such as social media, news sites or personal devices. These images will contain metadata (EXIF, IPTC, XMP) that provides information about the camera, date and location they were taken. The images may also be accompanied by content information such as a title or URL, which can help assess their authenticity. The project aims to analyze these images for changes, including manipulation or distortion, by analyzing pixel data, compression characteristics, inconsistent lighting and image quality, and valid metadata. Additionally, the system can use inverse image or image forensics to compare the image to existing data and assess its reliability, giving a confidence score about its authenticity.

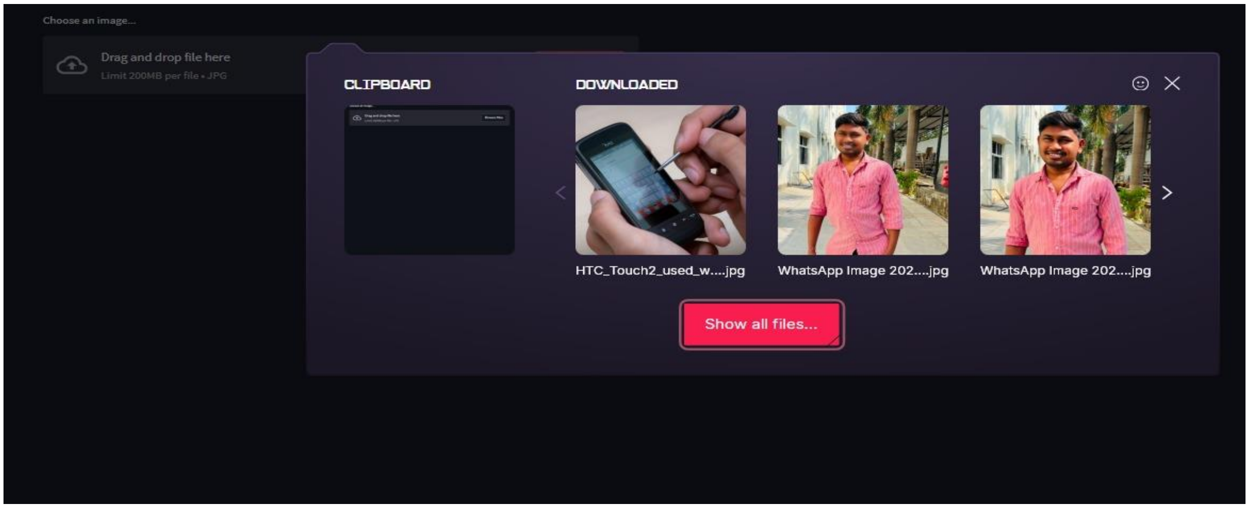
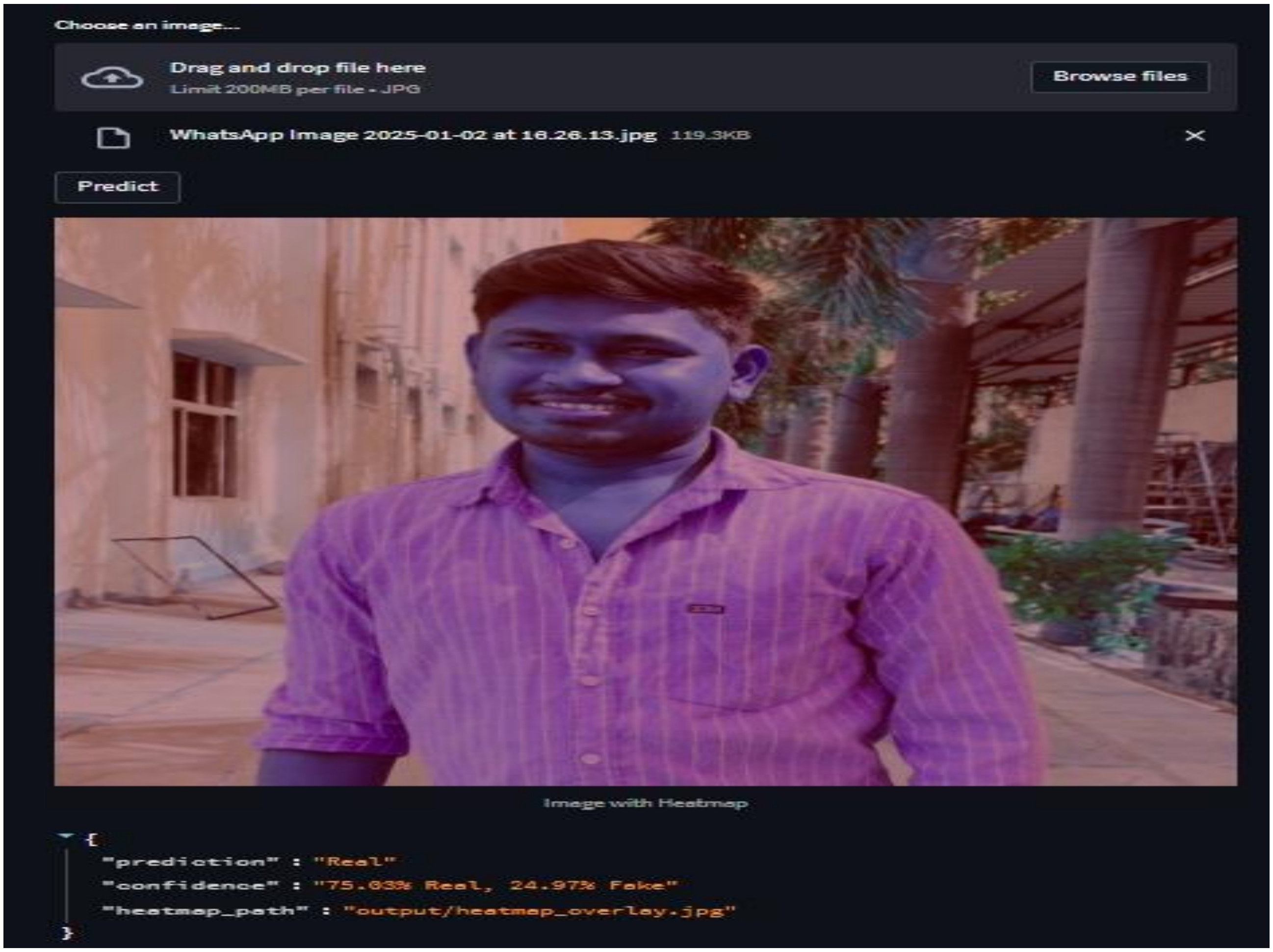


Fig 2 : selection of file

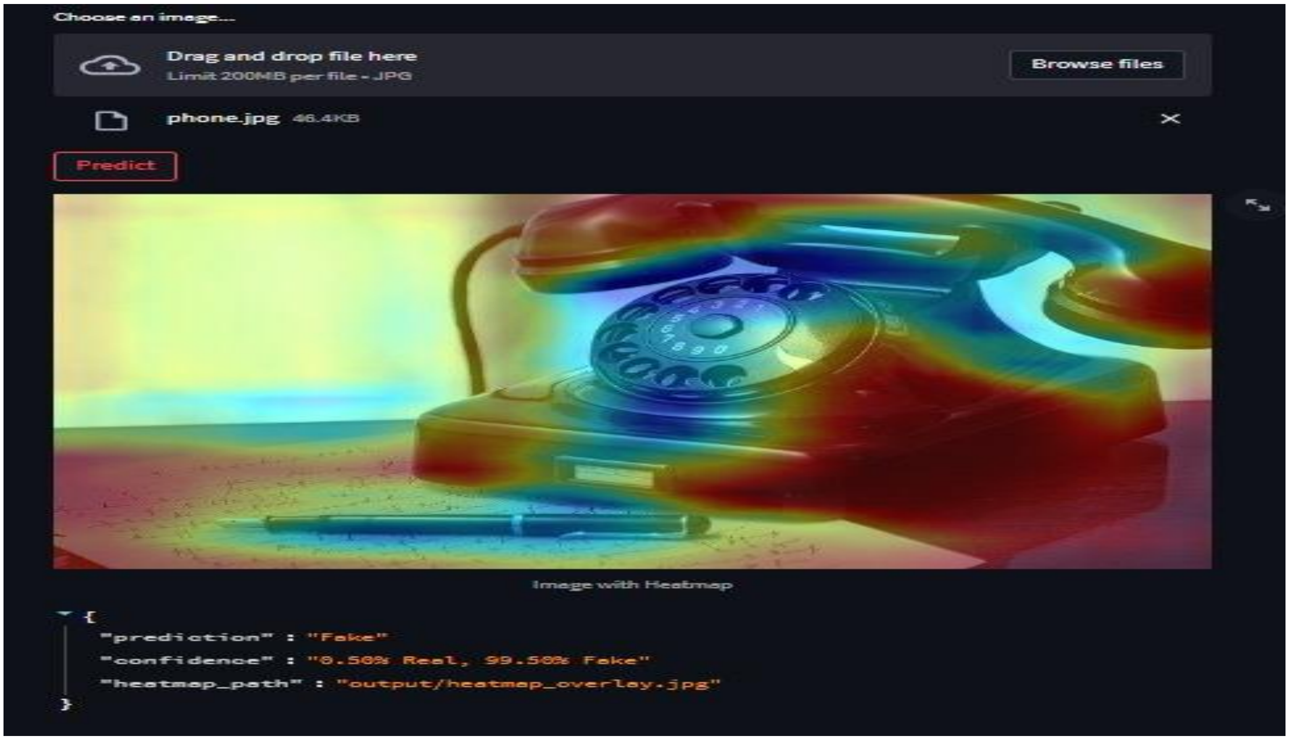
The output from the Image Authentication Project provides an integrity assessment of the image to determine if the image is authentic or has been tampered with. This includes a detailed description of any changes in the image that may indicate an error in the digital editing, such as pixellevel manipulation, uneven lighting, shadows, or negative patterns. The system can also flag specific areas that have changed, indicating areas that may need to be changed (such as adding, replacing, or copying elements). In addition to visual analysis, the output includes analysis of image metadata (such as EXIF or IPTC files) for signs of interference or inconsistencies, such as change capture date, location, or device information. The system can also crossreference images with known data from image acquisition or forensic tools and compare them to other sources to determine source or usage history. The result is usually expressed as a confidence score, which indicates the probability that the image is real, and can help the user assess whether the image is normal or complete. This may also include a detailed report that indicates any discrepancies, providing a transparent view of the verification process and results.



WhatsApp Image 2025-01-02 at 16.26.13.jpg 119.3KB

```
{
  "prediction" : "Real"
  "confidence" : "75.03% Real, 24.97% Fake"
  "heatmap_path" : "output/heatmap_overlay.jpg"
}
```

Fig 3 : Detection of Authentic images



phone.jpg 46.4KB

```
{
  "prediction" : "Fake"
  "confidence" : "0.50% Real, 99.50% Fake"
  "heatmap_path" : "output/heatmap_overlay.jpg"
}
```

Fig 4: Detection of Fake Image

VI. Discussion

Advantages:

High sensitivity: The system demonstrates the ability to accurately identify and manipulate images. By leveraging the feature extraction capabilities of convolutional neural networks (CNNs), it is possible to detect subtle inconsistencies and signs of tampering that are invisible to the human eye. Manual analysis is often time-consuming, laborintensive, and prone to human error. This leads to faster and more consistent results. Its scalability allows it to be sent to multiple sites with multiple image files. Ensuring their performance against complex systems such as those built by artificial neural networks (GANs) remains a challenge. GAN-based operations are often realistic and push the boundaries of machine perception. Limited or biased information can cause performance to degrade when faced with actual performance that differs from reported data. The effect exploits weaknesses in the model and bypasses detection. Addressing this vulnerability requires effective countermeasure training and continuous improvement of structures that can withstand these countermeasures. This provides benefits while acknowledging the challenges that need to be addressed for real-world deployment. This information forms the basis for future development and improvements to ensure the effectiveness and reliability of the system in a variety of situations.

VII. Conclusion and future scope

This project successfully implemented the use of convolutional neural networks (CNN) to identify real images. By utilizing the advanced feature extraction capabilities of CNNs, the system achieves high accuracy in analyzing manipulated images. It provides a broad and effective solution to the growing problem of misinformation in digital media. This application demonstrates the effectiveness of automated processes over manual processes, making it useful for real-world applications in journalism, forensics, and social security.

Expand the power of the system by participating in the education system. This approach allows models to identify and counter attacks designed to bypass detection. A wider range of data improves the generalizability of the model, allowing it to perform well in a variety of real-world situations. Social media platforms and mobile devices. This will allow for instant focus on images and give an immediate indication of their authenticity. The hybrid model can leverage the performance of both methods, which can increase detection accuracy and robustness against complex manipulations. This makes it an important tool in the fight against misinformation and digital manipulation.

VIII. References

1. Goodfellow, I., et al. “Explaining and Harnessing Adversarial Examples.” 2015.
2. Simonyan, K., & Zisserman, A. “Very Deep Convolutional Networks for Large-Scale Image Recognition.” 2014.
3. Rossler, A., et al. “Face Forensics++: Learning to Detect
4. Manipulated Facial Images.” 2019.
5. OpenCV Documentation: https://docs.opencv.org
6. TensorFlow Documentation: https://www.tensorflow.org
7. Zhou, P., Han, X., Morariu, V.I., & Davis, L.S. "Learning Rich Features for Image Manipulation Detection." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
8. Bayar, B., & Stamm, M.C. "A Deep Learning Approach to Universal Image Manipulation Detection Using a New Convolutional Layer." ACM Workshop on Information Hiding and Multimedia Security, 2016.
9. Dong, J., Wang, W., & Tan, T. "CASIA Image Tampering Detection Dataset." IEEE Transactions on Information Forensics and Security, 2013.
10. He, K., Zhang, X., Ren, S., & Sun, J. “Deep Residual Learning for Image Recognition.” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
11. Chollet, F. "Exception: Deep Learning with Depth wise Separable Convolutions." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
12. Dang, H., Liu, F., Stehouwer, J., Liu, X., & Jain, A.K. "On the Detection of Digital Face Manipulation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.
13. Li, Y., Chang, M.C., & Lyu, S. "In Ictu Oculi: Exposing AI Created Fake Videos by Detecting Eye Blinking." IEEE International Workshop on Information Forensics and Security (WIFS), 2018.
14. Mandelli, S., Gauravaram, P., & Simionato, M. "GAN-Based Synthetic Image Detection." Forensic Science International: Digital Investigation, 2021.
15. PyTorch Documentation: https://pytorch.org/docs/
16. Keras Documentation: https://keras.io