Library Progress International *Print version ISSN 0970 1052*
Vol.44 No.3, Jul-Dec 2024: P. 27470-27479 *Online version ISSN 2320 317X*

Original Article *Available online at www.bpasjournals.com*

Deep Fake Detection

Mr. VIJAYABHASKARAREDDY V[1], P. SAI NIVAS[2], S. VENU[3], T. DEEPAK[4], K. INDUMATHI[5]

1 Assistant professor, Department of Computer Science and Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India.
[2,3,4,5] Students, Department of Computer Science and Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India.
[1] vijayabhaskara559@gmail.com,[2] sainivaspeddineni@gmail.com, [3]sriramavenu09@gmail.com
[4] deepakthalla123@gmail.com,[5]katukuriindumathi2@gmail.com

How to cite this article: VIJAYABHASKARAREDDY V, P. SAI NIVAS, S. VENU, T. DEEPAK, K. INDUMATHI (2024) Deep Fake Detection. *Library Progress International,* 44(3), 27470-27479

Abstract:
In this paper, we demonstrate our automatic method for detecting deep fake images. With the growing threat of digital misinformation, the ability to accurately identify deepfake content is critical. It is getting harder to discern real photos from deepfake ones produced with sophisticated deep-learning algorithms. This study uses convolutional neural networks (CNNs) to evaluate and categorize real and fake photographs. The system's goal is to construct a deepfake detecting system. Users can upload photographs through an intuitive interface, and the CNN model processes them for detection.In order to attain an excellent classification performance, the model was given training using a wide range of original and deepfake images.This technique can be applied in numerous areas, for instance social platforms, media validation, and any other form of content restriction among other measures to reduce the levels of fake digital content.
With the assistance of streamlit, an interface for the deepfake detection system created with TensorFlow and Python was designed.

Keywords: Deepfake detection, ML, GAN, Feature Extraction, Deep Learning, Image processing, CNN, TensorFlow, Keras, Digital forensics, Binary Classification.

I. Introduction

The purpose of this paper is to construct an automatic instrumentation for the detection of deepfake images. The tools for machine and deep learning have rapidly evolved within the past few years, enabling the development of AI based hyper-realistic modifications\fakes of pictures that are hard to tell if they are real. The contentious issue of deep fakes is that they can be harmful in numerous sectors for instance media and politics, and the Internet in general where the audience can be easily misinformed because of the presence of digital content.
Distorted images known as deepfake images are produced by altering one or both faces or elements of an image and making it appear realistic, as this is often done with the help of fully developed technologies, namely, generative adversarial networks (GANs). Adjusting these changes, of course, is not an easy thing to do by ourselves. There are many reasons why people may want the use of a deepfake and one of them is the advancements in the deepfake technologies available today. Therefore, there is a great urgency in the need and importance of resorting to a deepfake detection technique that will be done with the help of machines.
Automating deepfake detection would not only relieve individuals from the manual and often error-prone process of detecting fake media but also enhance the reliability of media verification. Although there are currently a number of tools that can detect simple photo alterations, more complicated processes such training an ML model

from scratch to detect deep fakes requires even more sophisticated techniques. The objective of the project is to use CNNs to determine if an image is real or a deepfake.

The deepfake image identification application has a user-friendly design and enables the uploading of pictures instantly to determine if a picture is authentic or not within seconds. The system is to have been designed to be applicable in one or more of the following areas: media verification and regulating content in social media platforms.With the growth of the application, the upcoming releases may incorporate deep fake video detection capabilities and fortify the system against more sophisticated attempts of compromise. As the application goes live, there is also an easy-to-use interface wherein users can upload the images and see the outputs while the system is doing the image processing with an already trained CNN model in the background.

Some notable tools and platforms using Generative Adversarial Networks (GANs) to create deepfake images and videos:

DeepFaceLab
FaceSwap
Prisma
ArtBreeder
FaceRig

II. Literature Survey

The research area of detecting deepfakes and its capacity to go beyond the art has become very important as manipulated content continues to increase and its dangers become even more real. Many approaches have been suggested for verification of the genuineness of images or videos and reducing the impact and effects of such vices. The present paper examine many kinds of techniques and approaches that have been used in relation to this problem.

Gironi et al. [1] introduced a forensic tool for video manipulation, which is intended for the researchers in film forensics and aids in spotting media file tampering. Specifically, Thies et al. [2] presented the Face2Face approach that allows for face capturing and re-enacting of colored videos in real time. Besides, these techniques were essential for media manipulation detection because they provided the means to trace the continuity of the series of the video images.

Nguyen et al. [3] advanced the existing model of detecting video and image forgeries by putting forward the idea of multi-task learning where detection and segmentation of faked regions was done concurrently. This multifactorial layer made it straightforward to detect modified regions in pictures and movies, thus optimizing the performance of deepfake detection.

Li and Lyu [4] were concerned more about how to tell the difference between a fake video or an original by analyzing face warping artifacts, which are frequently found in altered videos. Artifacts such as these often form the basis of how realistic the content is and aids much in deepfake deterrence.

Media forensics has also taken the opportunity to explore the more advanced machine learning models such as transformers. In the article, ‘Attention is All You Need’, Vaswani et al. [5] proposed the design of the Transformer model, which later attracted attentions in many areas of deep learning particularly artificial intelligence in language translation and images deceptive identification known as deep fake detection.

Machine learning techniques to deepfakes have advanced to the extent of biological imitation. Media Forensics [6] has also extended research towards multi-angled research such as transformers and their history. The introduction of machine learning and AI to visual imagery shapes that employ deep fakes has therefore yanked away from the studies in linguistics and evolved to a host of other fields including vision research.

As noted by Rafique et al. [7], Error Level Analysis paired with deep learning techniques was advanced for deepfake detection. Sharp discrepancies at the pixel level caused due to the degradation of the image, are examined in this approach that helps pinpoint areas affected in deepfakes. To take a step further in the direction of detecting image forgeries, Shad et al. [8] also presented a review of deepfake detection CNNs and in addition, showcasing their development of image forgery detection based on convolutional neural networks.

Despite this, the progression of ways of detecting deepfakes has also attracted Convolutional Vision Transformers (CVT), which Wodajo and Atnafu [9] worked on. These transformed the traditional CNN’s image feature extraction and incorporated it into the attention attention of the transformer thus enhancing the detection of deepfake videos.

According to Guarnera et al. [10], deepfakes can also be detected by analyzing the convolutional traces deposited

in images during their creation. This method relies on the presence of CNN-specific image generation artifacts, llowing the distinction between real and fake images. As an example of the detection of tampered images, Durall et al. [11] addressed the issue of generative deep neural networks (GNNs) when it comes to spectral distributions preservation during up-convolutions.

Mittal et al. [12] proposed analyzing emotional incongruities for the purposes of deepfake detection employing audio-visual cues while discrepancies between faces and their background components were the focus of Nirkin et al. [13]in relation to deepfake detection.

All these works are aimed at the rise of the deepfake detection field that uses CNNs, error level analysis, and advanced architectures like transformers to aid in enhancing the performance of detecting deepfakes.

When combined, these studies enhance the development of various techniques that can easily detect deepfakes, especially with regard to the application of CNNs and high-level techniques that are currently embraced such as error-level analysis and transformers. In this regard, content manipulation detection systems are increasingly becoming effective owing to the incorporation of these approaches in almost every field. This particular review examines the emerging patterns in the timeline of detection techniques development, whereas the authors assert the increasing need to create additional techniques in these developments on account of the changing deception techniques.

A. Deepfake Detection Techniques

The first techniques available for deepfake detection dealt with facial analysis and the geometry of images. For instance, the study done by Afchar et al. in 2018 came up with a deepfake detection approach that is based on how the cunningly designed convolutional neural networks are used to distort structures of human faces. This particular technique can also nowadays be taught using the more sophisticated algorithm that studies the content, motion patterns and more time and space factors.

Korshunov and Marcel (2019) presented a broad review of the potential applications of various biometric markers, including information on eye blinks or the motion of facial features, in relation to the early detection of deepfake technology. With the improvement of certain types of deepfake without such clues, these techniques have been shown to perform well on selective deepfake types, but faster than the development of such telltale signs have come developments in GAN technology.

The last decades have been devoted to the combination of deep learning techniques where CNNs and recurrent neural networks (RNN) are employed in the processing of individual frames and video files in turn. Rössler et al. (2019) introduced the FaceForensics++ dataset as a benchmark for building and evaluating deepfake detection systems, where improvement of the detection rates has been observed. An experiment with several architectures was conducted and they concluded that large datasets fed into deep learning models could offer a solution in identifying image-based and video-based deep fake contents.

B. Evaluation and Metrics

A number of benchmarks and criteria have been developed to assess how effective deepfake detection algorithms are. Li et al. (2020) for example suggested the computing precision, recall and F1 scores with respect to the given media manipulations. Their study emphasized the need for well-established detection approaches over a range of deepfake types, which requires both normal and edited media in the datasets. Also, the advent of deepfake-detector adversarial attacks has brought about an interest in the survivability of such detection systems.

Zhang et al. (2020), for instance, sought to enhance the performance of deep learning models but also make them adversarially robust, that is, able to withstand the attacks of manipulators whose intention is to fool the detectors. These trends in deepfake detection literature point out that the topic remains a work in progress.

The earlier designs desired compromises in simple facial features but with the reality of deepfakes deep learning methods have come to pull the uncrossable wall. Research is still ongoing to find better, stronger, and more resourceful model that would be effective in the face of extreme deepfake tendencies.

III. Existing System

Currently available deep fake detection systems many require a human assessment or simple algorithms to detect the image manipulation. Early detection techniques were based on visual imperfections or anomalies in certain facial features, such as blinking (too fast or too slow), lighting (unnatural) or distortions in the pixels.

These techniques, albeit for the purpose of targeting early deepfake content, are not effective anymore with the improved technology on creating deepfakes. There are also tools available that apply conventional machine learning approaches where images are classified as real or fake. There performance, however, is not sufficient,

particularly where high quality deep fake contents created using complex GANs are concerned. They mostly do not support real-time detection and do not perform well against deepfakes that are very high resolution/quality, or that contain minor edits.

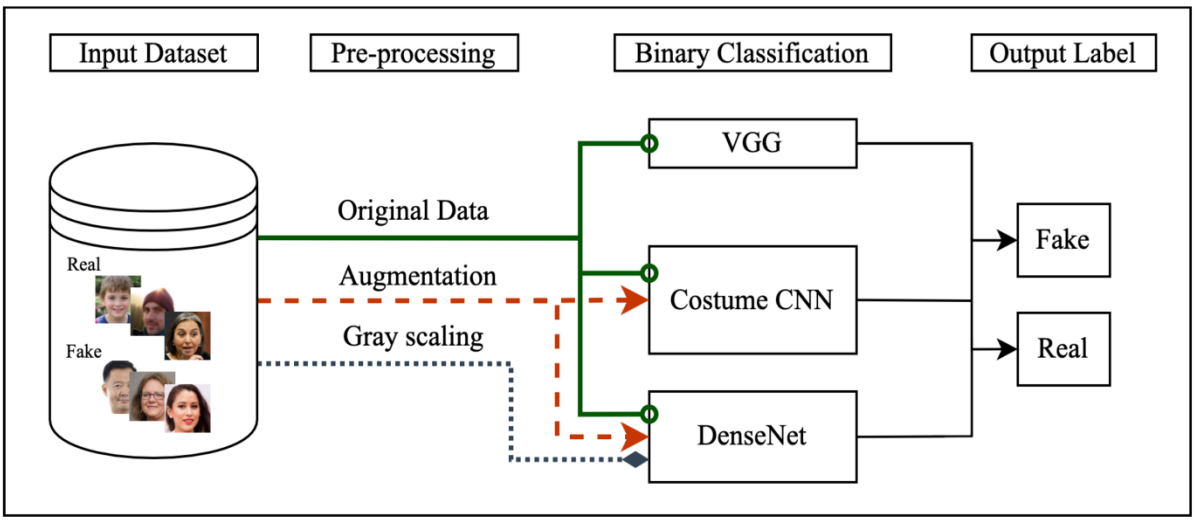


Fig. 1. Architecture of Existing System

Existing systems usually have an overall structure that involves some primitive processing of the image followed by feature extraction and classification. These types of systems, though practical for detection operations of lower complexity, do not have the required resilience and agility to perform detection on high-quality GAN deepfakes. This is where more advanced techniques that are leaning more on artificial intelligence come into play that are more effective and have better scaling properties.

From the point of view of such users content moderators, fact-checkers, the existing deepfake detection tools are very sluggish and often produce unreliable results. These drawbacks cause the problem of further computerized manipulation resolution systems to arise.

IV. Methodology:

The proposed system for detecting deepfake images involves several key steps: image preprocessing, deep learning-based classification, and result output.

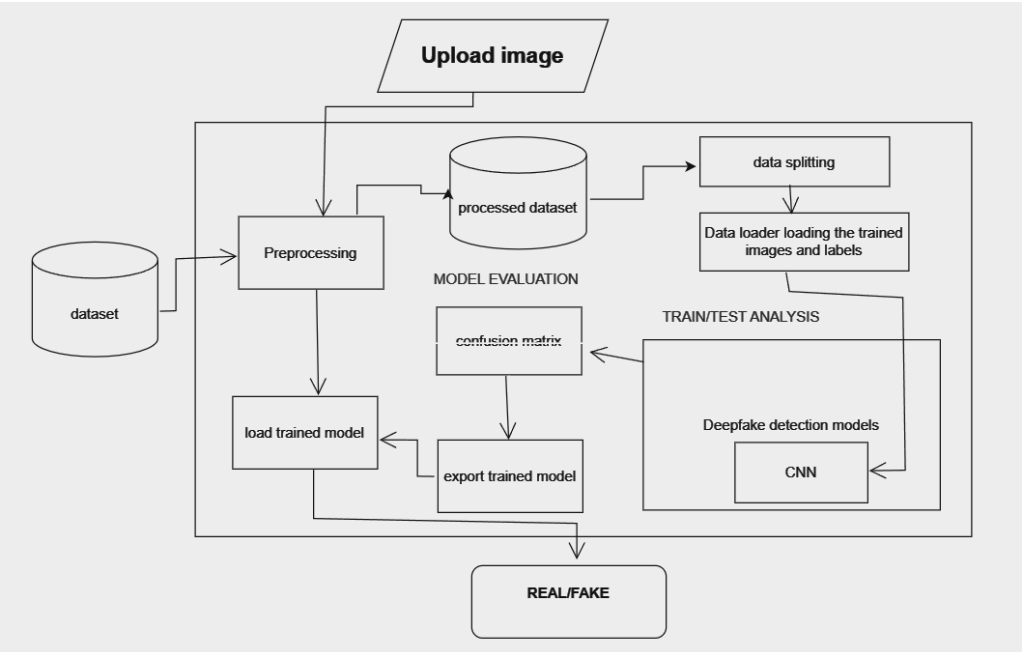


Fig. 2. Proposed System Architecture

The architecture includes image input through a user-friendly interface, where users upload images suspected of

being deepfakes. These images are processed using a Convolutional Neural Network (CNN) model, trained on a large dataset of real and deepfake images. The system outputs whether the image is classified as 'Real' or 'Deepfake', along with an accuracy score.

A. Key Components

1) Image Preprocessing:

The stage of image acquisition or collection invariably includes the downscaling and/or adjusting the sample images to a specific scale for the sake of uniformity of inputs to the model. This enables the system to accommodate different image formats and resolutions.

2) CNN Based Detection:

The methodology employs a systematic control mechanism in the use of convolution neural networks (CNNs) and various image processing methods in detecting deepfake images. The dataset used for this case is composed of two diagonally opposite classifications: real photographs, which have not been altered, and fake images. This dataset is sub-divided into three divisions: a training dataset with 140,002 records, a validation dataset of 39,428 records, and a testing dataset of 10,905 records, all balanced across the two classes. Each image goes through preprocessing steps before the training commences, one of them involves changing the images dimensions to create uniformity in the dataset, in this case changing them to (224 x 224) pixels. Normalization is also done, where pixel intensities are brought down to between 0 and 1 range, in order to quicken the convergence rate in training.

A convolutional neural network (CNN) is used in the task of deepfake image detection with the positive aspect being that it knows how to take advantage of the spatial hierarchies and features within the images. The architecture includes convolutional layers to gather main elements from the images such as lines, angles and textures, interspersed with pooling layers that help reduce image size while most absolute paramount details are preserved. The model has 12 convolutional layers:

These layers are in effect followed by fully connected layers which perform the role of classifying the obtained features and making decisions on which classes to assign. The last output layer outputs two probabilities, one for real images and the other for fake images, with the model predicting the class whose probability is higher.

3) Result Prediction:

Upon submitting an image, the system analyses the image uploaded and predicts if the image is original or a fake which is shown on the user interface. A confidence score indicates to the users the degree of reliability of such results.

B. Training the Model

To create the system, the first phase entailed gathering actual and doctored images. This data set was then used to train a CNN model that could identify the characteristic features used to differentiate real images from fake ones. The model was developed in TensorFlow environment and robust data augmentation strategies were applied in order to increase the performance of the model.

As this was a binary classification problem, softmax binary cross-entropy is used to train the model. The Adam optimizer is employed due to its capabilities, namely variable learning rates, which ensure that training will converge efficiently and effectively. For training the model, the dataset is divided into training and validation subsets. Model performance is assessed on the validation set and hyperparameters are tuned as appropriate. Strategies such as early stopping and model checkpointing are incorporated in order to avoid model overfitting and to enable the model to generalize well on unseen data.

The CNN model is designed in such a way that it extracts features at different levels of the model making it easier to discriminate between the real and the fake images. First, in the fundus layers, low-level features take shape in form of fundamental elements such as edges, corners, lines, and textured surfaces. Mid-level features are elaborated in the intermediate layers where CNN starts spotting features such as components of different objects, outlines, and regular – irregular symmetry of faces. High-level features have to do with advanced features precision and understanding of high-level distortions, moods and changes, and other aspects of in-depth analysis of the given face. Lastly, the model constructs the so-called global features, which serve as an abstract way to describe the given image, depicting its overall shape along with the important parts of it.

In the end, the judgement of a convolutional neural network is reliant on the assessment of the congruity of the facial image along with its features, the recognition of any artifacts related to the alteration of the image, and finally a discrepancy in the illumination and the overall appearance of the image that may suggest some alterations

in the image have taken place. With this in place, the project seeks to effectively curb the menace of deepfakes by incorporating modern day imaging and convolutional neural network structures to solve the problem.

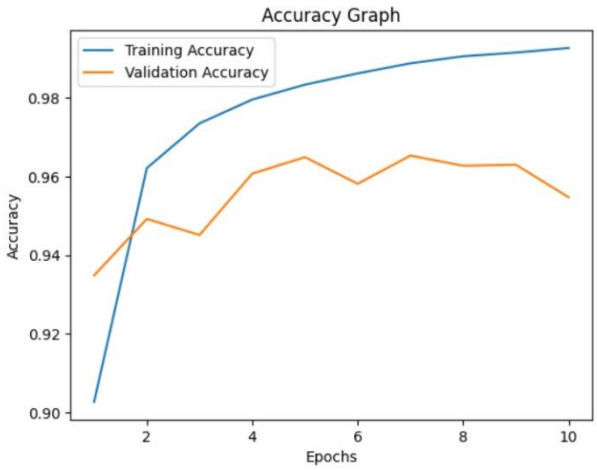


Fig: Training acccuracy and validation accuracy v/s Epochs

C. User Interface

The application has been designed with a thoughtful user experience in mind which is enabled using Streamlit. Users have the ability to upload images, see them, and load them for processing. The system is capable of operating also on low-quality images with additional capability for very high-quality images broaden their usage.previous training.

D. Automatic Detection

Upon completing the upload of the image, the next stage is that the image gets fed into the CNN model for classification. The image - before it is fed into the model - goes through various changes, thanks to the backend, which is quite responsible for carrying out all the necessary processes such as resizing, normalizing, and others. Afterward, the model determines if the image is a real one or a deepfake based on the things it has learned through previous training.

LAYER(type)	**Output Shape**	**Param#**
Conv2d (Conv2d)	(None, 224, 224, 16)	448
Conv2d_1 (Conv2d)	(None, 224, 224, 16)	2320
Max_pooling2d (MaxPooling2d)	(None, 112, 112, 16)	0
Conv2d_2 (Conv2d)	(None, 112, 112, 32)	4640
Conv2d_3 (Conv2d)	(None, 112, 112, 32)	9248
Max_pooling2d_1 (MaxPooling2d)	(None, 56, 56, 32)	0
Conv2d_4 (Conv2d)	(None, 56, 56, 64)	18496
Conv2d_5 (Conv2d)	(None, 56, 56, 64)	35928
Max_pooling2d_2 (MaxPooling2d)	(None, 28, 28, 64)	0
Conv2d_6 (Conv2d)	(None, 28, 28, 128)	73856
Conv2d_7 (Conv2d)	(None, 28, 28, 128)	147584
Max_pooling2d_3 (MaxPooling2d)	(None, 14, 14, 128)	0
Conv2d_8 (Conv2d)	(None, 14, 14, 256)	295168
Conv2d_9 (Conv2d)	(None, 14, 14, 256)	590080

Max_pooling2d_4 (MaxPooling2d)	(None, 7, 7, 256)	0
Conv2d_10 (Conv2d)	(None, 7, 7, 512)	1180160
Conv2d_11 (Conv2d)	(None, 7, 7, 512)	2359808
Max_pooling2d_5 (MaxPooling2d)	(None, 3, 3, 512)	0
Flatten (Flatten)	(None, 408)	0
Dense (Dense)	(None, 1024)	4719616
Dense_1 (Dense)	(None, 2)	2050

Table 1: CNN Classification Layers

The architecture of the CNN based mechanism comprises of several layers for the purpose of feature extraction and classifying images as well. This model has a structure of convolutional layers to perform the function of pattern recognition and pooling layers to perform the function of compressing and finally thin layers to enable the model differentiate between real images and deepfake ones.

E. Result Display

The last phase is to present the outcome of the detection to the end user. The system employs a color-coded representation (green for the label 'Real' and red for the label 'Deepfake') for better comprehension. The system also shows a confidence score whereby the users can tell to which extent the model is confident in making such a prediction.

F. Data Storage and Future Improvements

The outcomes are kept in a database system for a better understanding and enhancement of the quality in subsequent stages. With the advancements of deepfake technologies, the system will also require modifications of its model and datasets in order to sustain effective detection.

V. Results and discussions:

The system has managed to attain quite an impressive prediction performance when it comes to that total, standing at 91%. Even though it comes with a cost, the macro averaged precision, recall and F1 scores were calculated and found to be 0.91 indicating that the model does not favor any one of the existing classes. This was also clear in the weighted averages where precision, recall and F1 scores were noted to be 0.91.

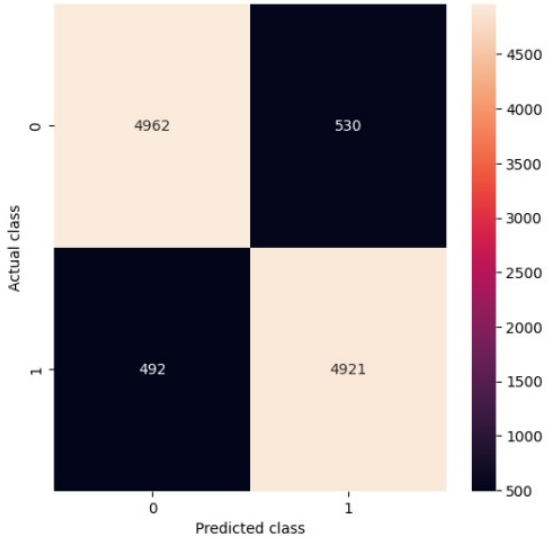
This paper uses a publicly available dataset that is compiled by Kaggle. Kaggle real and fake face database is a dataset that contains images of both real and fake human faces. In this study,we achieved a deepfake detection accuracy of 91% with the use of Deep Learning and convolution neural networks(CNNs), which are drastically improved from existing systems. For instance, the system in "Deep fake detection and classification using deep learning," which was based on GoogLeNet, produced an accuracy of 81%. Current implemented system is efficient but has drawbacks such as relatively lower rate of detection and poor performance with high-dimensional data due to reliance on GoogLeNet, which is not a state-of-the-art model applied for the purpose.

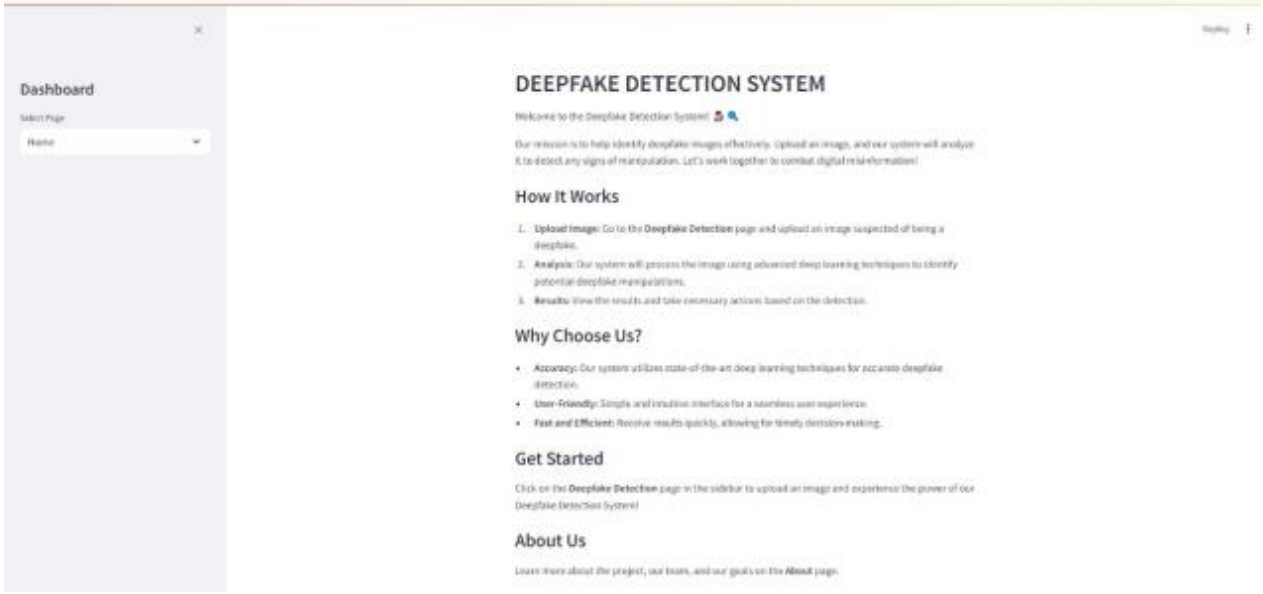
As opposed to this, our model makes use of the strengths of Deep Learning and CNN - aggregated residual transformations that enable enhanced visualization of tricky patterns in visual data. In this way, our system is relatively better at distinguishing between the slightest differences between manipulated images and hence has overall better performance.

These results, although underlining the advances made by our approach, actually open the way to further optimizations in future iterations.

	PRECISION	RECALL	F1-SCORE	SUPPORRT
0	0.91	0.90	0.91	5492
1	0.90	0.91	0.91	5413
ACCURACY			0.91	10905
MACRO AVG	0.91	0.91	0.91	10905
WEIGHTED AVG	0.91	0.91	0.91	10905

In the confusion matrix it was also seen that the values of true positives, true negatives, false positives and false negatives were given and it was about identifying deepfake images that: these other challenges in the form of false negatives tend to occur.


4962
530
492
4921
0
1
Actual class
0
1
Predicted class
4500
4000
3500
3000
2500
2000
1500
1000
500


Dashboard
DEEPFAKE DETECTION SYSTEM
How It Works
Why Choose Us?
Get Started
About Us

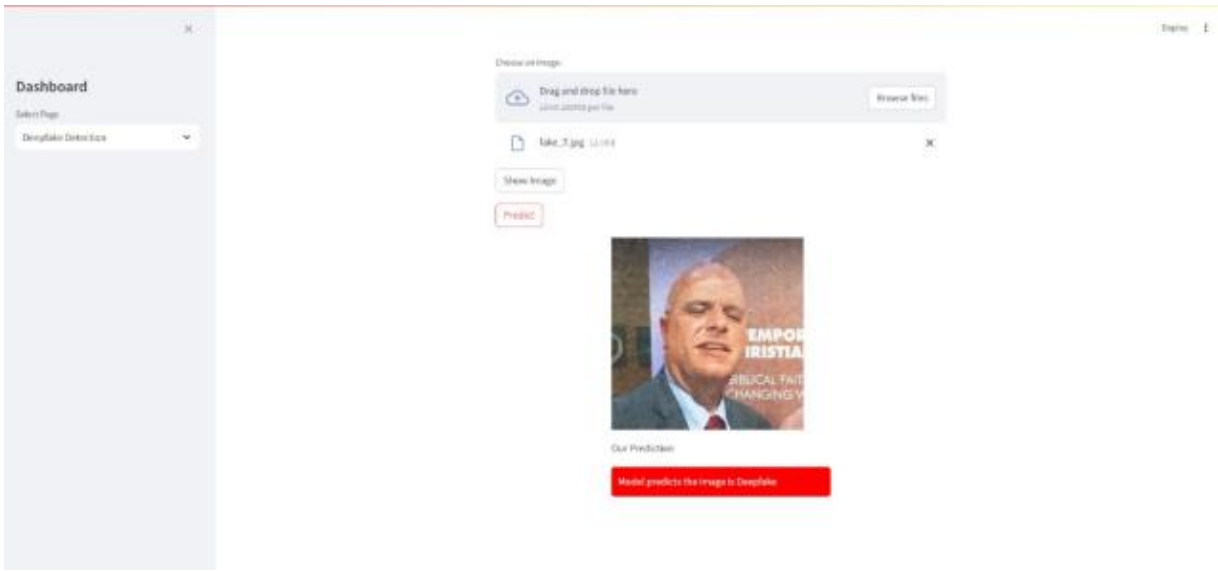

Dashboard
Drag and drop file here
Browse files
Show Image
Predict
Our Prediction

Fig : Platform home page and detection of images whether it is fake or not

VI. Conclusion:

This research confirms the appropriateness of using CNNs for deep fake image analysis with an impressive rate of 91% accuracy alongside proportional precision, recall, and F1-score measures. The model features showed it was able to differentiate an original image from a fake: from the simplest edges to complex facial features that were impossible in real images. These results are very important for combating deepfake technology and propose solutions for the difficult issue of the authentication of digital material for the purpose of information security.

However, the study, while relevant, is not without its drawbacks including limitations in the datasets used and some unwanted classifications. Future dimension of this work should aim at investigating cutting – edge structures and other datasets so as to enhance generalization and turn up resilience. This can be incorporated in social and media platforms to enhance the intergity of the content posted.

In summary, this study presents a timely response to an emerging threat to the integrity of media and provides evidence for why its initial objectives have been achieved and deepfake detection focus has expanded. But since deepfakes are an evolving threat, more work will be necessary to keep up with all this technology to maintain the trust in digital content.

VII. References:

[1] A. Gironi, M. Fontani, T. Bianchi, A. Piva and M. Barni, “A video forensic technique for detecting frame deletion and insertion,” 2014.

[2] J. T. M. Zollhofer, M. Stamminger and C. Theobalt, “Face2face: Real-time face capture and reenactment of RGB videos,” 2016.

[3] H. H. Nguyen, F. Fang, J. Yamagishi and I. Echizen, “Multi-task learning for detecting and segmenting manipulated facial images and videos,” vol. abs/1906.06876, 2019.

[4] Y. Li and S. Lyu, “Exposing deepfake videos by detecting face warping artifacts,” in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019.

[5] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin, “Attention is all you need,” in *Advances in Neural Information Processing Systems (NIPS)*, 2017.

[6] M. J. Turek, “Media Forensics (MediFor),” Defense Advanced Research Projects Agency (DARPA), 2019.

[7] R. Rafique, M. Nawaz, H. Kibriya and M. Masood, “DeepFake detection using error level analysis and deep learning,” in *4th International Conference on Computing & Information Sciences (ICCIS)*, Lahore, 2021.

[8] H. S. S. e. al., “Comparative Analysis of Deepfake Image Detection Method Using Convolutional Neural Network,” 2021.

[9] D. Wodajo and S. J. Atnafu, “Deepfake Video Detection Using Convolutional Vision Transformer,” arXiv, 2021.

[10] L. Guarnera, O. Giudice and S. Battiato, “DeepFake detection by analyzing convolutional traces,” in *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, 2020.

[11] R. Durall, M. Keuper and J. Keuper, “Watch your up-convolution: CNN based generative deep neural networks are failing to reproduce spectral distributions,” in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 2020.

[12] T. Mittal, U. Bhattacharya, R. Chandra, A. Bera and D. Manocha, “Emotions don’t lie: An audio-visual deepfake detection method using affective cues,” in *Proceedings of the 28th ACM International Conference on Multimedia*, Seattle, 2020.

[13] Y. Nirkin, Y. Keller and T. Hassner, “FSGAN: Subject agnostic face swapping and reenactment,” *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV),* pp. 7184-7193, 2019.

[14] L. A. e. al., “Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures,” *Sensors,* vol. 21(5), pp. 1688-1689, 2021.

[15] S. Lee, S. Tariq, Y. Shin and S. S. Woo, “Detecting handcrafted facial image manipulations and GAN-generated facial images using Shallow-FakeFaceNet,” *Applied Soft Computing,* vol. 105, p. 107256, 2021.

[16] L. Jiang, R. Li, W. Wu, C. Qian and C. Change Loy, “DeeperForensics-1.0: A large-scale dataset for real-world face forgery detection,” arXiv, 2020.

[17] C. Öhman, "Introducing the pervert's dilemma: A contribution to the critique of deepfake pornography," *Ethics and Information Technology,* vol. 22, pp. 133-140, 2020.

[18] R. Thakur and R. Rohilla, "Recent advances in digital image manipulation detection techniques: A brief review," *Forensic Science International,* vol. 312, p. 110311, 2020.

[19] N. Carlini and H. Farid, "Evading deepfake-image detectors with white-and black-box attacks," arXiv, 2020.

[20] S. Agarwal, H. Farid, O. Fried and M. Agrawala, "Detecting deep-fake videos from phoneme-viseme mismatches," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, 2020.

[21] D. Güera and E. J. Delp, "Deepfake video detection using recurrent neural networks," in *IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)*, 2018.

[22] X. Yang, Y. Li and S. Lyu, "Exposing deep fakes using inconsistent head poses," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* p. 8261–8265, 2019.

[23] K. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies and M. Nießner, "FaceForensics++: Learning to detect manipulated facial images," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.

[24] T. Isola, P. Zhou and S. A. Efros, "Image-to-image translation with conditional adversarial networks," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 1125-1134, 2017.

[25] D. Cozzolino, G. Poggi and L. Verdoliva, "Recasting residual-based local descriptors as convolutional neural networks: An application to image forgery detection," *Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security,* pp. 14-19, 2017.

[26] A. Tolosana, R. Vera-Rodriguez, J. Fierrez and J. Ortega-Garcia, "Deepfakes and beyond: A survey of face manipulation and fake detection," in *arXiv preprint arXiv:2001.07668*, 2020.

[27] H. Farid, "Digital image ballistics from JPEG quantization," *IEEE Transactions on Information Forensics and Security,* vol. 5(2), pp. 480-487, 2009.