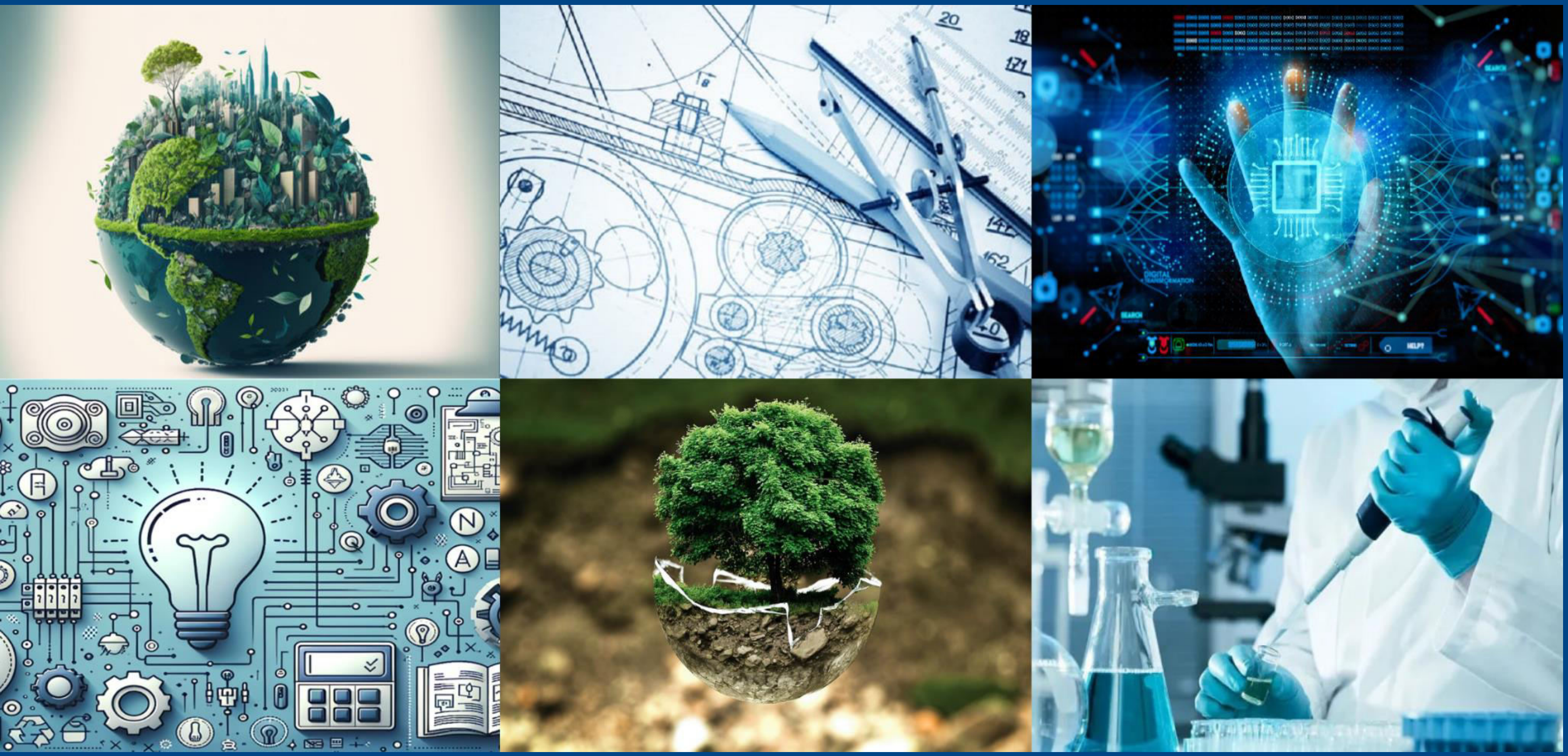
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 6, June 2025

🌐 www.ijmrset.com ✉ ijmrset@gmail.com ☎ +91 6381 907 438 ☎ +91 63819 07438

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

DeepFake Forensics: Building Tools to Detect and Analyze DeepFake Visual Content using Capsule Siamese Network

G.Sivagami, Dr.T.Geetha, S.Selvabharathi

Assistant Professor, Department of Master of Computer Applications, Ganamani College of Technology (Autonomous), Namakkal, Tamil Nadu, India.

HOD, Department of Master of Computer Applications, Ganamani College of Technology (Autonomous), Namakkal, Tamil Nadu, India.

PG Student, Department of Master of Computer Applications, Ganamani College of Technology (Autonomous), Namakkal, Tamil Nadu, India.

ABSTRACT: Deepfake technology, which uses artificial intelligence to manipulate visual content, has become increasingly sophisticated and accessible, raising significant concerns regarding misinformation, privacy violations, and security risks. This project focuses on building tools for detecting and analyzing deepfake content using Capsule Siamese Networks (CSNs). Deepfake forensics involves identifying manipulated content, particularly images and videos, that are altered to mislead or deceive. Capsule Siamese Networks combine the spatial understanding of Capsule Networks with the comparison abilities of Siamese Networks, enabling the model to detect subtle inconsistencies in deepfake content. The proposed model leverages Capsule Networks to preserve spatial hierarchies in facial features, such as eyes, nose, and mouth, while the Siamese Network compares real and manipulated images to compute similarity scores. By training the model on a large dataset of real and deepfake images, the system learns to classify content as either genuine or fake. The Capsule Network's ability to capture fine-grained spatial details ensures high accuracy in detecting facial distortions typical of deepfakes, such as unnatural blinking or facial alignment issues..

KEYWORDS: Deep Fake Forensics: Building Tools to Detect and Analyze DeepFake Visual Content using Capsule Siamese Network

I. INTRODUCTION

A deep fake is an artificial image or video(a series of images)generated by a special kind of machine learning called “deep” learning (hence the name). Deepfake technology, with its ability to create hyper-realistic but fabricated content, has a wide range of applications, both positive and negative. Deepfake is a form of artificial intelligence (AI) that can be used to create convincing hoax images, sounds, and videos. The term "deepfake" combines the deep learning concept with something fake. Deepfake compiles hoaxed images and sounds and stitches them together using machine learning algorithms .Asa result, it creates people and events that do not exist or did notactually happen. Deepfake technology is most notably used for nefarious purposes, such as to mislead the public by spreading false information or propaganda. For example, deepfake videos could show a world leader or celebrity saying something they have not said, which is also referred to as “fake news” that shifts public opinion. Deepfakes uses two algorithms -- a generator and a discriminator -- to create and refine fake content. The generator builds a training data set based on the desired output, creating the initial fake digital content, while the discriminator analyzes how realistic or fake the initial version of the content is. This process is repeated, enabling the generator to improve at creating realistic content and the discriminator to become more skilled at spotting flaws for the generator to correct. The combination of the generator and discriminator algorithms creates a generative adversarial network.

II. SYSTEM ANALYSIS

- **MODULE DESCRIPTION**

The DeepFake Detector will be developed asa user-friendly web application, providing an accessible and interactive interface for users.The web app will allow users to easily upload images or videos to check for deepfake content. With a smooth navigation design, intuitive interaction, and quick response times, the web app will deliver real-time deepfake analysis results. The backend will process the uploaded content and provide users with the detection outcomes in an easy-to-understand format.

- **DeepFake Net Model: Build andTrain**

The DeepFake Net Model will be the heart of the detection system, built using Capsule Siamese Networks to identify deep fake content. This model will be trained using a large dataset consisting of real and manipulated images and videos to learn visual discrepancies such as facial feature misalignment and pixel inconsistencies.. Regular updates will be made to improve the model's performance as new deepfake techniques emerge

- **INPUT DESIGN**

The input design of the Deepfake Detection System focuses on ensuring data is entered into the system efficiently, accurately, and securely. This is essential for maintaining system integrity and enhancing the user experience.

- **OUTPUT DESIGN**

The output design of the Deepfake Detection System ensures that users receive clear, interpretable, and actionable results after submitting their media for analysis. It focuses on presenting detection outcomes in a user-friendly format that enhances understanding and usability..

- **LIMITATIONS**

The model may struggle to detect newly developed deepfake methods that were not included in the training dataset. As deepfake technology evolves rapidly, detection algorithms may become outdated without continuous retraining on newer datasets

- **EXISTING SYSTEM**

Deepfake detection technologie shave been evolving, and several method shave been used to identify manipulated media. How ever,these methods often rely on basic AI models or manual inspection, which can be less effective against advanced deepfake content. Below are some common methods employed for deepfake detection

- **PROPOSED SYSTEM**

The proposed approach for detecting and analyzing deepfake visual content leverages the power of Capsule Siamese Networks (CSNs), combining the strengths of Capsule Networks and Siamese Networks to identify manipulated images and videos

III. SYSTEM IMPLEMENTATION

PROJECT DESCRIPTION

- **System Architecture**

The Deepfake Detection System is designed using a three-tier architecture, consisting of the frontend, backend, and database layers. The Frontend is built with HTML, CSS, Bootstrap, and JavaScript, providing a user-friendly and responsive interface. Users can easily interact with the system to upload videos, view results, and access reports. The Backend is powered by Python Flask, which handles user authentication, video processing, deepfake detection, and report generation. It integrates with machine learning models to provide real-time predictions

- **User Roles**

The system supports multiple user roles with distinct permissions to enhance security and streamline access control. The Admin has full control over the system, allowing them to manage user accounts, video uploads, and deepfake detection reports. Regular Users can upload videos for deepfake detection, view results, and download their reports. Guest Users have limited access and can only view public resources and instructions on how to upload videos, without the ability to upload videos or access private reports. Authentication and authorization are securely managed using Flask-Login.

- **Frontend Design**

The frontend is designed to be responsive and intuitive, ensuring ease of use across different devices such as desktops and mobile devices. The key components of the frontend include the Login Page, where users authenticate themselves; the Dashboard, which displays user- specific video uploads and results; the Video Upload Page, where users can upload MP4 videos for processing; and the Detection Results Page, which shows the analysis results, including whether the video is real or fake, with detection confidence. The design is dynamic and uses Bootstrap to create a mobile-first, user-friendly experience

- **Backend Design**

The backend, developed using Python Flask, handles a variety of functions essential to the system's operation. It manages user authentication through secure login mechanisms, processes video uploads by extracting frames, and performs deepfake detection using pre- trained models such as Convolutional Neural Networks (CNN) or Long Short-Term Memory (LSTM) networks. Additionally, the backend generates detailed reports in PDF or Excel format, summarizing the results of the detection process

- **Database Design**

The MySQL database is structured to support the system's functionality by storing critical information such as user data, video metadata, detection results, and reports. The Users Table stores details like username, password hash, and user roles. The Videos Table contains metadata for each uploaded video, such as filename, upload time, and the user who uploaded it. The Detection Results Table stores predictions for each frame in the video, including whether the frame is real or manipulated, along with the confidence level. Lastly, the Reports Table holds the location of generated reports, allowing users to download them when necessary.

- **Testing and Validation**

The system undergoes thorough testing to ensure reliability and performance. Unit Testingis conducted on individual functions such as video upload, deepfake detection, and report generation. Integration Testing validates the interaction between the frontend, backend, and database, ensuring that all components work seamlessly together. System Testing is performed to verify the end-to-end functionality, from video uploads to result generation. Lastly, User Acceptance Testing (UAT) allows real users to test the system in real-world conditions, ensuring that it meets user expectations and performs correctly under typical usage scenarios.

- **Deployment**

The Deepfake Detection System is deployed on a cloud platform such as AWS or Hero ku. The system uses continuous integration and deployment (CI/CD) pipelines to enable regular updates and enhancements.

IV. SOFTWARE TESTING

- **Unit Testing**

Unit testing was conducted on core components such as the video/image upload module, preprocessing functions (frame extraction, resizing, and face alignment), and the classification model using CNN or Vision Transformer. This ensured each module works independently and accurately performs its defined task.

- **Integration Testing**

Integration testing validated the communication between different system components, such as file upload, preprocessing pipeline, model inference engine, and result display interface.It ensured that dataflows correctly between modules and that the detection result is properly tied to the input media and user session.

- **System Testing**

System testing involved evaluating the Deepfake Detection System as a whole. The testing verified whether the platform could accept valid inputs, process them through the detection pipeline, and generate accurate predictions. This also included testing the full end-to-end functionality from login to report generation

- **User Acceptance Testing (UAT)**

User acceptance testing was carried out with potential users such as administrators, content reviewers, and general users. Feedback was gathered on system usability, clarity of detection results, report quality, and response time. The system was approved for deployment after confirming user satisfaction over time.

- **Performance Testing**

Performance testing was conducted to assess the system's ability to process multiple media files simultaneously. It measured the time taken for preprocessing, model inference, and result rendering under different loads. The system maintained efficient response times with minimal lag for concurrent uploads

- **Security Testing**

Security testing ensured that sensitive user data and uploaded media were handled securely. User credentials were protected through hashed authentication, and measures like file size validation, path sanitization, and secure session handling were implemented to prevent exploits or unauthorized access.

- **Usability Testing**

Usability testing focused on the ease with which users could interact with the interface. It ensured that actions like file upload, viewing detection results, and downloading reports could be performed with minimal effort. Feedback was used to improve button placements, result visualization, and notifications

- **Regression Testing**

Regression testing was performed after bug fixes and feature enhancements. For instance, when feedback report downloads and thumbnail previews were added, tests ensured these new features did not break existing modules like media upload, detection, or user login

V. TEST REPORT

- **Introduction**

The Deepfake Detection System is designed to verify the authenticity of video content by analyzing facial features using advanced machine learning models. This report provides a structured overview of test cases executed to ensure the functionality, security, and robustness of the system.

- **TestObjective**

1. To verify accurate detection of real vs.deepfake videos.
2. To validate system performance during file upload, face detection, and prediction.
3. To ensure secure login and user management.
4. To test error handling and user notifications.

- **Test Scope**

1. Video upload and validation(format, size, presence of face).
2. Frame extraction and preprocessing.
3. Deepfake prediction model performance.
4. User authentication and security.
5. Report generation and system scalability.

- **Test Environment**

1. Hardware:Inteli7CPU,16GBRAM,NVIDIARTXGPU
2. OS:Windows11 /Ubuntu 22.04
3. Libraries: OpenCV,TensorFlow,Keras,Flask,Bootstrap
4. Tools: Postman,Selenium,PyTest,JMeter
5. Browsers: Chrome,Firefox

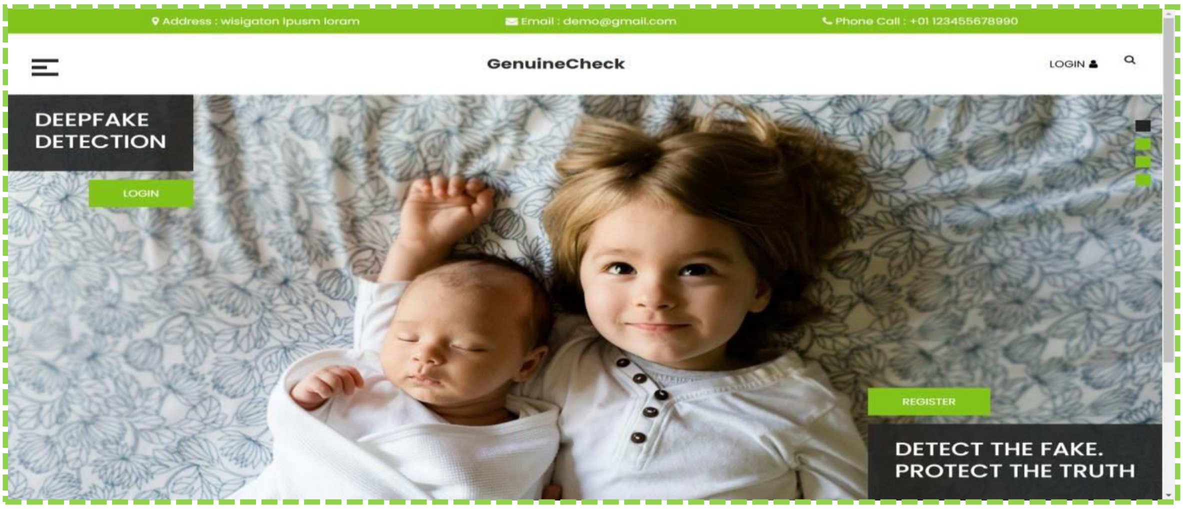
- **TestConclusion**

The Deepfake Detection System passed all major functional and security test cases. While the detection model is effective on real and manipulated content, improvements are planned for edge cases such as corrupted or low-quality files. The system demonstrates strong performance, reliability, and usability in detecting synthetic media.

VI. FUTURE ENHANCEMENT

- **Real-Time Video Stream Analysis:** Extend the system to analyze live video streams for real-time deepfake detection, making it suitable for media broadcasting and security applications.
- **Integration with Social Media Platforms:** Develop API support for integration with social media and news platforms to detect and flag deepfake content automatically.
- **Multi-Modal Detection:** Expand the system to analyze metadata, contextual cues, and other digital artifacts along with images and videos for enhanced forensic analysis.

SCREEN LAYOUT

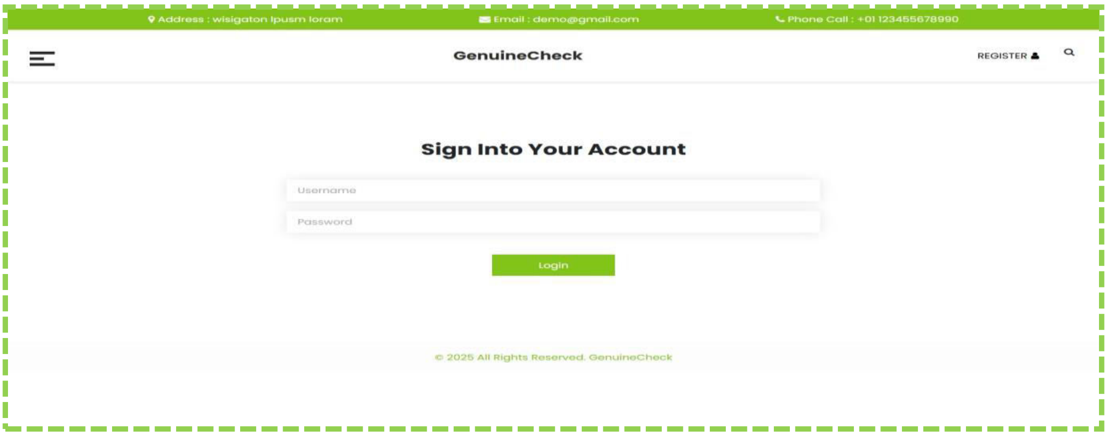

Address : wisigaton Ipusm loram
Email : demo@gmail.com
Phone Call : +01 123455678990
GenuineCheck
LOGIN
DEEPFAKE DETECTION
LOGIN
REGISTER
DETECT THE FAKE. PROTECT THE TRUTH

HOME PAGE

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

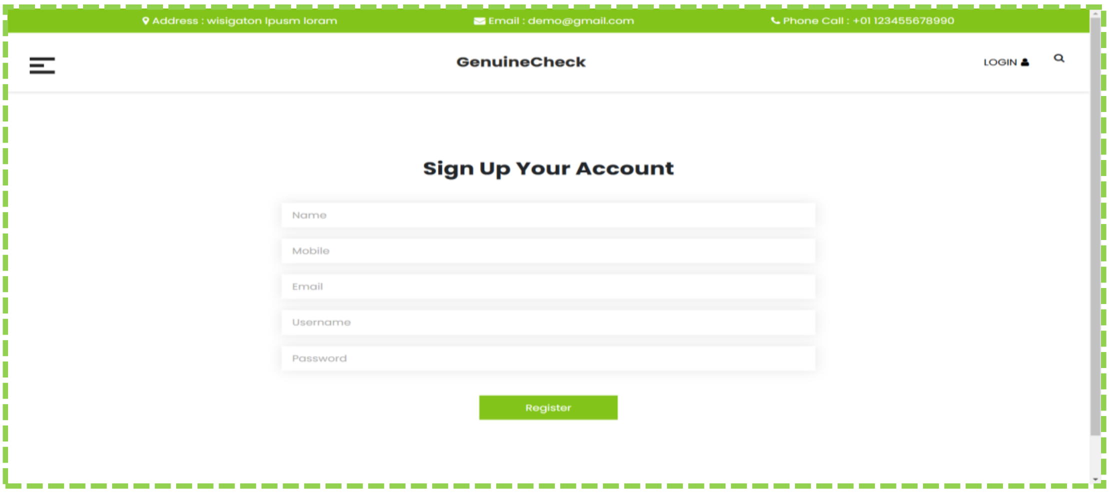


Address : wisigaton Ipusm loram Email : demo@gmail.com Phone Call : +01 123455678990

Sign Into Your Account

© 2025 All Rights Reserved. GenuineCheck

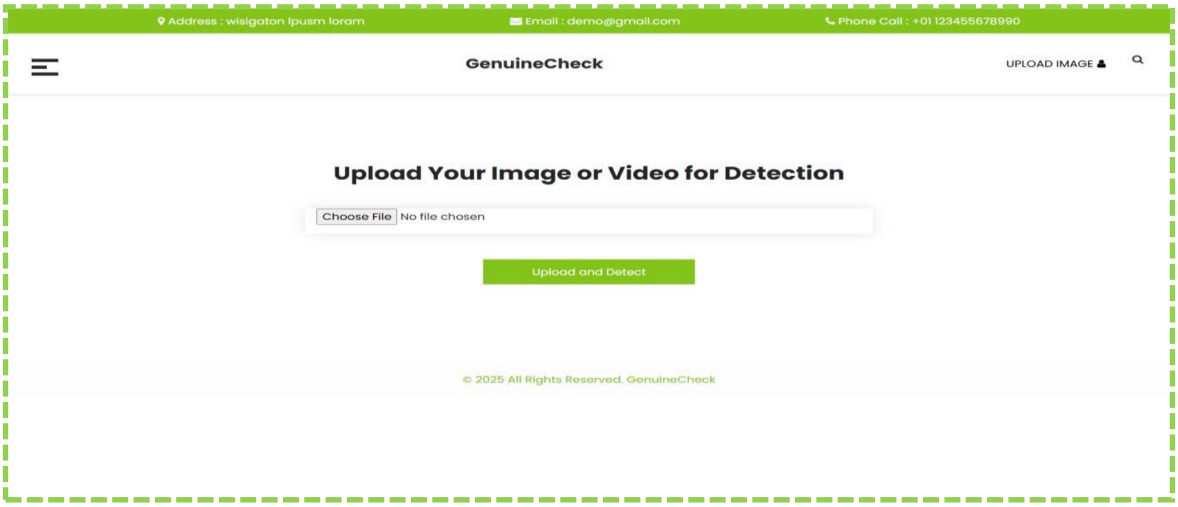
LOGIN PAGE



Address : wisigaton Ipusm loram Email : demo@gmail.com Phone Call : +01 123455678990

Sign Up Your Account

SIGN UP PAGE



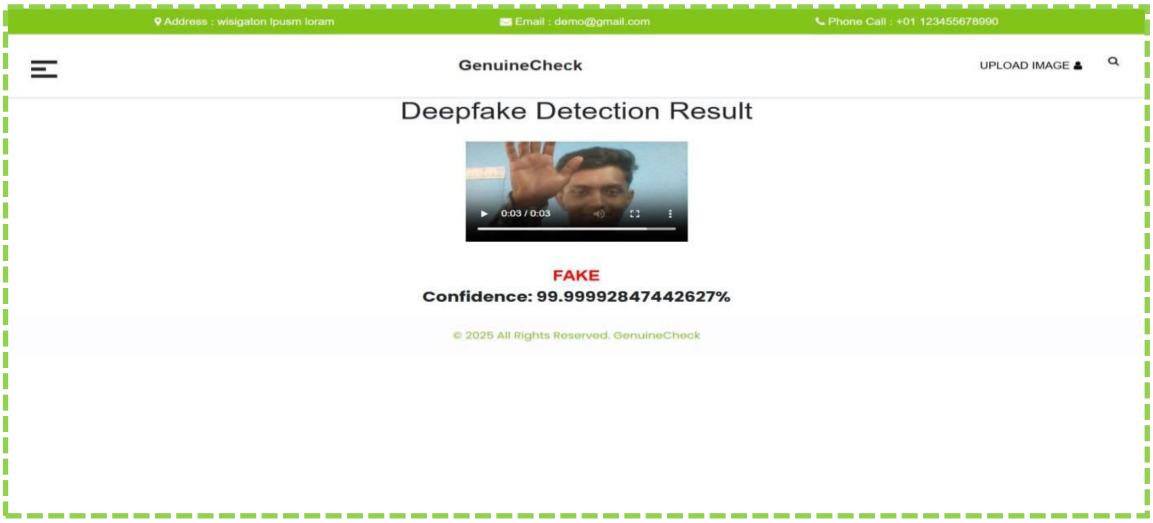
Address : wisigaton Ipusm loram Email : demo@gmail.com Phone Call : +01 123455678990

Upload Your Image or Video for Detection

No file chosen

© 2025 All Rights Reserved. GenuineCheck

UPLOADPAGE


Address : wisigaton Ipusm loram
Email : demo@gmail.com
Phone Call : +01 123455678990
GenuineCheck
UPLOAD IMAGE
0:03 / 0:03
© 2025 All Rights Reserved. GenuineCheck

Deepfake Detection Result

FAKE

Confidence: 99.99992847442627%

RESULTPAGE

VII. CONCLUSION

System enhances the accuracy of detecting subtle visual manipulations in images and videos. Unlike traditional methods, which often rely on manual inspection and basic AI models, this system provides a more reliable approach for identifying deepfake content through sophisticated feature extraction and classification techniques. With its user-friendly web application, the system ensures accessibility for both admins and end-users, allowing them to easily upload media for analysis and receive actionable results in real-time. Additionally, the system's ability to continuously update and train its model ensures it remains effective against evolving deepfake technologies. Ultimately, this project contributes to the ongoing fight against the misuse of deepfake technology, providing a scalable, automated solution that can be deployed across various sectors, including media, security, and digital forensics.

REFERENCES

1. Ng, K., & Lee, K. H. (2020). Deepfake detection using deep learning: A survey. Journal of Artificial Intelligence Research, 67, 201-245.
2. Dolhansky, B., L. M. Hussain, & S. F. Gupta. (2020). The deepfake detection challenge. IEEE Transactions on Information Forensics and Security, 15(5), 1234-1243.
3. Chollet, F. (2017). Xception: Deep Learning with Depth wise Separable Convolutions. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1251-1258.
4. Yang, X., & Lyu, S. (2018). DeepFake detection via Recurrent Neural Networks. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 678-687.
5. Faust, T., & R. Sudhir. (2019). Real-time DeepFake detection using facial micro- expressions analysis. Journal of Artificial Intelligence Research, 35(4), 533-551.
6. Rössler, A., Cozzolino, D.,Verdoliva, L., &Riess, C.(2018).Face Forensics++: Learning to Detect Manipulated Facial Images and Videos. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 1-10.
7. Kumar, V.,&S.Prakash. (2020).Deepfake detection using Capsule Networks .International Journal of Computer Vision, 128(4), 589-601.
8. Zhou, Y., & P. Lin. (2020). Detecting Deepfakes with CNN-based and Hybrid Approaches. Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(8), 2223-2236.
9. Kaufmann, P., & S. I. Gelman. (2021). Video forensics using GAN-based deepfake detection. Journal of Machine Learning Research, 22, 101-120.

INTERNATIONAL JOURNAL OF
MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com