IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 6, June 2026

DOI: 10.17148/IARJSET.2026.13631

TRUSTNET.AI – DETECTION OF FAKE IMAGES AND VIDEOS

Ms. Bhargavi H G[1], Himani B L[2], Sandra C[3], Navya Manosri S[4], Ananya V[5]

Assistant Professor, Dept. of Computer Science MMK and SDM MMV, Mysuru[1]

Dept. of Computer Science MMK and SDM MMV, Mysuru[2]

Abstract: The artificial intelligence technology is getting better and better. This has led to the creation of deepfake technologies. These deepfake technologies are causing problems for cybersecurity and digital trust. They are also making it hard to verify identities. There have been a lot of cases of deepfake fraud in North America. In fact, the number of cases has increased by about 1,740%. This is a jump. 40% Of all biometric authentication attacks worldwide are because of deepfakes.

India is also facing a lot of problems with deepfakes. The number of deepfake-related cybercrimes in India has increased by 550% since 2019. This is a concern. It is estimated that the financial losses due to deepfakes will be than ₹70,000 crore in 2025. These losses will be because of identity impersonation, financial fraud and social engineering attacks. We need to find a way to stop these deepfakes. We need a system that can detect deepfakes with accuracy.
There are already some deepfake detection systems. These systems use something called Convolutional Neural Network (CNN)-based architectures. They are trained on datasets that contain about 15,000 images. These systems are good at detecting deepfakes. They can even work in time. However they have some limitations. They rely on single-model architectures. They do not have comparative analysis capabilities. They also do not have user authentication mechanisms.

To solve these problems, we are proposing a system called the Tri-Model Hybrid Deepfake Detection Pipeline (TMHDPP). This system uses deep learning architectures. It combines the features of these architectures to improve deepfake classification performance. The TMHDPP system uses something called InceptionV3 and EfficientNet. It also uses an architecture that combines the convolutional layers of both models. This helps to capture and aggregate -scale spatial features. The system is. Evaluated using a dataset that contains 2,041 training images and 2,041 testing images.

The TMHDPP system is not a detection system. It also has a web-based application. This application has role-based access control. This means that administrators and users have functionalities. Administrators can upload datasets manage users, train models, test models and monitor performance. They can also visualize the results using graphs.

Users can upload images or videos for deepfake analysis. They can get the classification results along with the evaluation metrics. The TMHDPP system is designed to be scalable, accurate and user-centric. It aims to provide a solution to the growing problem of deepfakes.

The TMHDPP system is an improvement over the existing systems. It uses a - model architecture. It has benchmarking capabilities. It also has integrated management features. All these features make the TMHDPP system a powerful tool for detecting deepfakes. We hope that this system will help to reduce the number of deepfake-related cybercrimes. We also hope that it will help to improve cybersecurity and digital trust. The deepfake detection technology is still evolving. We need to keep working on it to make it better. The TMHDPP system is a step in the direction. It has the potential to make a difference in the fight, against deepfakes.

INTRODUCTION

This system tackled the growing problem of synthetic media manipulation head-on. It uses a cutting-edge, multi-model deep learning framework to catch and classify deepfakes both in images and videos. As AI keeps pushing forward, especially in computer vision and generative adversarial networks, it's become surprisingly easy to make incredibly realistic fake media. That means old-school ways of checking what's real and what's not just don't cut it anymore. With deepfakes showing up everywhere, from scams and stolen identities to fake news and digital fraud, there's a serious need for a detection solution that's reliable, accurate, and can handle the scale.

Instead of relying on just one model, this system pulls together the strengths of InceptionV3, EfficientNet, and a custom hybrid architecture. Working together, these models spot both the obvious and not so obvious signs of image tampering.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 6, June 2026

DOI: 10.17148/IARJSET.2026.13631

Where one model might catch weird pixels, another picks up on awkward blends or lighting glitches. This tag-team approach uncovers subtle clues that single models almost always miss.

The setup follows a straightforward machine learning process: collect a diverse dataset, clean and prep the images, and split everything into solid training and testing sets. During preprocessing, images get resized, normalized, and augmented to boost variety and make sure the models stay sharp and flexible. Balancing the data means you don't wind up with a system that's good at picking out fakes in one context but clueless in another.

From there, the models learn to spot key signs of manipulation dodgy facial textures, off-color tones, strange lighting, and all the weird stuff you see in deepfake images.

Each model brings something unique. InceptionV3 is great at catching complex visual patterns through its smart multi-scale feature extraction. Efficient Net balances accuracy and speed by scaling network depth and size just right. Then, the hybrid model merges their strengths, pulling data from multiple levels of every image to get a sharper overall picture. Putting all this together, the system doesn't just get better at flagging subtle fakes, it stays tough even when facing compressed or deliberately tampered files.

The final package isn't just a back-end tool, it's a web-based platform anyone can use. It's got a clean interface, supports different user roles, and comes loaded with handy features. Admins get full control of uploading new data, managing users, running model training or tests, and comparing how each model stacks up. Regular users just upload their images or videos and get instant results about whether their content is genuine or manipulated. This setup keeps the system friendly and secure, making sure only the right people have access to the more sensitive features.

It also does more than just say "real" or "fake." The platform lets you see performance graphs, compare models, and review detailed reports. Admins can track key stats like accuracy, precision, recall, and training progress so you always know how the system's performing and which model's your best bet. These insights aren't just for research, they help keep the platform practical in real-world use, too.

Tests showed that this multi-model approach delivers strong, reliable results. Together, InceptionV3, Efficient Net, and the custom hybrid model outperformed classic single-model solutions, catching deeper, more sophisticated fakes. And with its full set of management tools user controls, dataset managers, and monitoring features, the platform is ready for more than just the lab. It's built for real deployment.

This system marks a big step forward in deepfake detection. It blends state-of-the-art AI with real-world usability, offering a smart, scalable way to protect digital identities and online information. Plus, the flexible, modular design means there's room to grow, bring in bigger datasets, plug in transformer models, analyze more than just images, or add advanced explainability features. As fake media keeps evolving, this platform can keep up.

PROBLEM STATEMENT

The Artificial Intelligence technology is getting better and better. This has led to a lot of people making deepfake videos and pictures that are very realistic. While this technology can be useful for things like entertainment and education it also has some downsides. For example, people are using deepfakes to steal identities commit fraud and spread false information. There has been an increase in deepfake-related crimes all around the world. In North America the number of deepfake fraud cases has gone up by about 1,740%. Now deepfakes are responsible for 40% of all biometric authentication attacks worldwide.

In India there has been a jump in deepfake-related crimes too. Over 550% since 2019. It is estimated that by 2025 the economic losses due to deepfakes will be than ₹70,000 crore.

The problem is that it is getting easier for bad people to make convincing media using Artificial Intelligence tools. This makes it harder for us to tell what is real and what is not. Most of the systems that try to detect deepfakes use a single-model approach, which's not very effective. These systems can detect some deepfakes. They are not good enough to catch all of them. The current systems also do not have the features that are needed to use them in real-life situations.

For example, they do not have user authentication, role-based access control and dataset management. They also do not have the ability to compare the performance of deep learning architectures. This makes it hard to know which model is the best to use.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 6, June 2026

DOI: 10.17148/IARJSET.2026.13631

Another challenge is that the people making deepfakes are getting better and better at it. They can now make videos and pictures that are very realistic with facial expressions, lighting and skin textures that look real. This makes it even harder to detect deepfakes. The systems that are used to detect deepfakes need to be able to extract features from media and they need to be robust against new manipulation techniques. So, there is a need for a system that can detect deepfakes more effectively.

This system should use deep learning architectures to improve feature extraction and classification accuracy. It should also have user management, administrative controls, dataset handling and performance evaluation tools. By using a -model approach the system can overcome the limitations of the current single-model solutions. This will make it easier to tell what is real and what is not. It will help to prevent the bad things that can happen when people use deepfakes. The Artificial Intelligence technology and deep learning architectures like InceptionV3 and Efficient Net can be used to make this system.

The system will provide an scalable platform, for deepfake detection and it will help to enhance the reliability and transparency of digital media authentication. The Artificial Intelligence technology and deep learning architectures will be used to make the system more effective. The system will be able to detect deepfakes accurately and it will help to prevent the bad things that can happen when people use deepfakes. The Artificial Intelligence technology is very powerful. It can be used to make a system that can detect deepfakes more effectively.

OBJECTIVES

- We want to find manipulated media using computer techniques.
- Our goal is to make sure digital media is real and stop news from spreading.
- We will use computer models and AI to look at media and tell if it's fake.
- We need to check for faces, strange frames and audio that doesn't match the video in fake media.
- We are going to make it easy for users to upload media check if its fake and see the results.
- The system needs to work and check media in real-time.
- We have to be able to check types of media, like pictures, videos and more.
- To make our model better we will keep training it with lots of data.
- We want to make sure we don't make mistakes when checking for media.
- By finding media we can make the internet safer and more trustworthy.
- We need a way to store our data, user information and reports.
- We will give information and a score to show how sure we are that media is fake.
- Our system needs to be strong and handle a lot of media to check.

PROPOSED SYSTEM

The proposed system, TMHDPP is a way to detect fake images and videos. It uses three models: InceptionV3, Efficient Net and a mix of both. This system helps tell media from fake media by looking at many details. It uses 2,041 training images and 2,041 testing images to make sure it works well. The system can run all three models at the time giving fast results with detailed statistics. The web application has two types of users: administrators and regular users.

Administrators manage the models and data while regular users can sign up upload media and get results. This system can be used in areas to keep media verification safe, efficient and fast.

ADVANTAGES OF PROPOSED SYSTEM

- **Using models:** This approach gives more accurate results by combining the best of each model.
- **Easy to grow:** The system can handle lots of data and work in time.
- **Two types of users:** This setup gives administrators control and regular users a way to interact with the system.
- **Fake media detection:** Users can upload media and get results right away making it great, for security uses.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 6, June 2026

DOI: 10.17148/IARJSET.2026.13631

EXISTING SYSTEM

The systems we have now for finding deepfakes mostly use something called Convolutional Neural Networks. These systems look at a lot of pictures to learn how to spot ones. One of the models for this is called EfficientNet. It is really good at finding deepfakes. Does not use up too much computer power. Some other systems use -trained models like Exception to look at pictures closely and find fake parts. These systems use techniques to make them work better and faster. However, they have some problems. They only use one type of model which can be a limitation. They also have trouble handling a lot of pictures at the time. They do not have good ways to manage users or control how the models work. This makes them not very good for real-world problems. Deepfake detection systems like these can also be tricked by people who try to fool them.

LIMITATIONS OF EXISTING SYSTEM

- **Single-model reliance:** Most systems for detecting deepfakes only use one type of model. This means they are not very good at adapting to kinds of deepfakes.
- **Limited scalability:** These systems have trouble working with a lot of pictures in time. This is because they are not fast enough.
- **Insufficient user authentication:** The systems we have now do not have ways to make sure users are who they say they are when they upload pictures or videos.
- **Vulnerability to attacks:** Deepfake detection systems can be fooled by people who try to trick them. They can also get confused when pictures are compressed or changed in some way.

Library used:

1. **NumPy**
   - This library is used for doing math and working with numbers.
   - It helps with things like arrays and math operations.
2. **Pandas**
   - We use this library to handle and look at data.
   - It can read files like CSV. Manage the data.
3. **OpenCV (cv2)**
   - This library is used for working with images and videos.
   - It can take out frames from videos.
   - Get images ready for use.
4. **TensorFlow**
   - This is a framework used for learning.
   - We use it to build and train models that can detect things.
5. **Keras**

This is a part of TensorFlow that makes it easier to build and train networks.

6. **Scikit-learn**
   - We use this library to get the data ready for use.
   - It also gives us numbers that show how well the model is working, like accuracy and precision.
7. **Matplotlib**

This library is used for making graphs and showing the results.

8. **Seaborn**

This library is used for making graphs that're a little more complicated. It can make things, like confusion matrices.

9. **Pillow (PIL)**

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 6, June 2026

DOI: 10.17148/IARJSET.2026.13631

We use this library to load and change images.

10. Dlib

This library is used for finding faces and looking at the features of faces.

11. Face Recognition

This library is used to find and recognize faces in images and videos.

12. Flask

We use this library to make a website where people can upload images and videos and see the results.

13. OS

This library is used for working with files and folders.

14. Pickle

We use this library to save and load models that have been trained.

15. ImageIO

This library is used for reading and writing images and videos.

SYSTEM ARCHITECTURE

The system architecture of TrustNet.AI is made up of five stages that work together to detect deepfake images and videos. It starts with Media Acquisition, where the user uploads a file through the web interface made with Django. Then the media goes through Preprocessing, where any distortions are removed and the quality is improved. The next stage is Feature Extraction, where the deep learning models look at the media and find patterns that could mean it is a deepfake, such as facial textures or weird lighting. These patterns are then sent to the Classification stage, where three different models. InceptionV3, EfficientNet and the Hybrid CNN. Look at the media and decide if it is real or fake. Finally, the Identification stage takes the results from the Classification stage. Gives a final answer saying if the media is Real or Fake and how confident it is in that answer.

ALGORITHM USED

The algorithm used in TrustNet.AI is the Convolutional Neural Network or CNN. This is a type of deep learning model that's really good at looking at images and videos. It works by looking at the media in layers starting with the Input Layer, which gets the image as a bunch of pixel values. Then the Convolutional Layer applies filters to the image to find edges and textures. The Pooling Layer makes the image smaller. It is easier to work with and the Flatten Layer turns the image into a long line of numbers. The Dense Layers then look at these numbers. Decide if the media is real or fake. Finally, the Output Layer gives an answer saying if the media is Real or Fake and how confident it is in that answer. TrustNet.AI uses three CNN models. InceptionV3, EfficientNet and the Hybrid CNN. To detect deepfakes. InceptionV3 is good at finding details in images while EfficientNet is good at finding bigger patterns. The Hybrid CNN combines the results from both of these models to give an accurate answer.

HOW MACHINE LEARNING WORKS IN TRUSTNET.AI

The machine learning process in TrustNet.AI has seven steps. First a dataset of images is collected, with a number of real and fake images. Then the images are pre-processed, which means they are resized and cleaned up to make them easier to work with. Next the CNN models look at the images.

Find patterns that could mean they are deepfakes. The models are then trained on these patterns using an algorithm to make them more accurate. The results from the models are then combined to give a more accurate answer. The performance of the models is then evaluated, using metrics to see how well they are working. Finally, the trained models are used in the TrustNet.AI web application, where they can be used to detect deepfakes in time. This means that when a user uploads an image or video the system can quickly say if it is Real or Fake and how confident it is, in that answer.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 6, June 2026

DOI: 10.17148/IARJSET.2026.13631

FUTURE ENHANCEMENTS

We can make TrustNet.ai even better in the future. Here are some ideas:

- We can add video analysis to catch fake videos in real-time.
- We can also add audio deepfake detection to catch audio too.
- Moving to the cloud can help handle users at once.
- A mobile app can make it easier for people to use.
- Using blockchain can help keep media verification secure.
- Improving our learning models can make them more accurate.
- We can also make it work with languages and social media platforms.

RESULT

Our deepfake detection system worked well in identifying images and videos. We used three models: InceptionV3, EfficientNet and a hybrid model that combined both. The hybrid model did the best by combining features and improving accuracy. We trained the system with 4,082 images. Used techniques like normalization and resizing to make it work better.

The system did the following:

- Images and videos
- Extracted features
- Made real-time predictions
- Compared model performance

We used metrics like accuracy and precision to measure how well the models worked. The web interface was also secure and provided real-time reporting. Our project successfully created an AI-based deepfake detection system. It can identify images and videos efficiently.

We used deep learning architectures like InceptionV3 and Efficient Net. The system improved detection accuracy and reliability. Our project showed how AI and Machine Learning can help with media authentication and cybersecurity. The web platform provided access and real-time analysis. This makes the system suitable for use in social media monitoring and online security. The system helps detect AI-generated media. This can reduce misinformation and digital manipulation online.

CONCLUSION

TrustNet.AI presents an effective and intelligent approach to detecting fake images and videos using advanced deep learning techniques. By integrating **InceptionV3**, **EfficientNet**, and a **Hybrid CNN model**, the system achieves improved accuracy and reliability compared to conventional single-model methods. The web-based platform enables users to upload media and receive real-time detection results, while administrative features support model management and performance monitoring. Overall, the project provides a scalable and secure solution for combating deepfakes, reducing digital misinformation, and strengthening trust in online media and cybersecurity applications.

REFERENCES

[1]. Ian Goodfellow et al. "Generative Adversarial Nets " Advances in Neural Information Processing Systems, 2014.
[2]. David A. Rossler et al. "Learning to Detect Manipulated Facial Images " IEEE International Conference on Computer Vision (ICCV) 2019.
[3]. Brian Dolhansky et al. "The Deepfake Detection Challenge Dataset," arXiv preprint arXiv:2006.07397, 2020.
[4]. Chollet, François, "Xception: Deep Learning with Depthwise Separable Convolutions " IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017.
[5]. Tan, Mingxing and Le, Quoc, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks " International Conference, on Machine Learning (ICML) 2019.
[6]. TensorFlow Documentation.
[7]. OpenCV Documentation.
[8]. Django Official Documentation.

