iJRASET
International Journal For Research in Applied Science and Engineering Technology



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 12 **Issue:** V **Month of publication:** May 2024

**DOI:** https://doi.org/10.22214/ijraset.2024.62191

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 12 Issue V May 2024- Available at www.ijraset.com*

# Video Manipulation Detection Using DenseNet and GoogLeNet

Anamika K P[1], Angel Saju[2], Merine George[3], Sumi Joy[4], Roteny Roy Meckamalil[5]
*Department of Computer Science and Engineering, Mar Athanasius College of Engineering, Kothamangalam, Kerala*

***Abstract:*** ***Recent advancements in video technology have democratized facial manipulation techniques, such as FaceSwap and deepfakes, enabling realistic edits with minimal effort. While beneficial in various applications, these tools pose societal risks if misused for spreading disinformation or cyberbullying. The widespread availability of this technology through web and mobile apps has heightened the threat, making the detection of manipulated content a critical challenge. It aims to address the ethical and societal risks associated with the widespread avail- ability of user-friendly facial manipulation tools, such as deepfakes, by focusing on detecting deepfake manipulation in videos. This emphasizes the significance of dataset selection and training methodology for optimal results, employing ensemble learning with DenseNet and GoogLeNet models. The proposed methodology involves training models from scratch and with pre-trained weights, using Celeb- DF (v1) and Face Forensics++ datasets to enhance model generalization, and analyzing video frames to extract features for detecting manipulated faces in videos. The overall goal is to fortify societal defenses against the misuse of facial manipulation tools and promote responsible technological innovation.***

***Index Terms: Deep learning, DeepFake, CNNs, GANs***

## I. INTRODUCTION

Deepfakes, a aggregate of the phrases “deep studying” and “fake,” are a form of marketing in which the picture or likeness of someone is replaced with a photo or video of every other man or woman that is already there. while the exercise of fake records is not new, deepfakes use artificial intelligence and machine learning to edit or create visible and audio content that is simple to faux. the primary deep learning-primarily based system gaining knowledge of used to make deep connections calls for training generative hostile networks (GANs) or autoencoders, two kinds of generative neural network architectures. although the era become evolved through experts, it’s miles now available to every person thru internet and telephone programs, allowing each person to create personalized pix and videos. each person can edit faces in video sequences and this can be done effortlessly thanks to this generation. although those gear are useful in many regions, they are able to negatively have an effect on people when used incorrectly. developing deep gaining knowledge of algorithms has advantages and drawbacks. Deepfakes created the use of deep mastering algorithms each have benefits and downsides. The purpose of the undertaking is to perceive video modifications referred to as ”deepfakes”. The power and accuracy of detection tools can be progressed the usage of a multi-layered technique that combines listening, vision and nature. reading AI descriptions can also help deepen selection-making algorithms and boom confidence in their results. Collaboration among service companies and policymakers is essential to prevent the spread of the virus. The design and implementation of a comprehensive investigative system need to be guided via moral issues, together with balancing the proper to freedom of expression with the want to block negative content. The complex troubles because of deepfakes require a very good strategy, which in the long run makes use of advances in public awareness, regulation and technology.

## II. RELATED WORKS

### *A. Correlation Sequences for Deepfake Detection*

Using DenseNet to analyze video clips to ensure physical inconsistency and identify movement patterns indicative of manipulation. The dense connections in Densenet help preserve different parts of the frame, thus improving the model’s ability to detect subconscious flanking inconsistencies. This approach includes a monitoring system that detects areas with abnormal physical behavior, thereby improving overall detection accuracy. Experimental results demonstrate the effectiveness of DenseNet in distinguishing real and practical videos even in difficult situations. Adaptive learning is used to modify the pre-trained DenseNet model for specific tasks in deep learning. The study highlights the importance of time elements in in-depth analysis and provides insight into the reduction in physical activity reported by management.

©IJRASET: All Rights are Reserved 

*B. Adversarial Defense in DenseNet Based Forgery Detection*

Examine the vulnerability of DenseNet adversarial at-tack and propose to improve the strength of the anti-counterfeiting model. By analyzing the impact of compet-ing against Densenet's intense belief system and orderliness. Evaluation of countermeasures shows that DenseNet-based fraud detection systems are better protected against sophis-ticated attacks. This work investigates the trade-off between detection accuracy and computational complexity in anti-aliasing techniques applied to deep learning models. Ideas for using the powerful DenseNet model in real-world situations are discussed, highlighting the need for effective prevention strategies in video forensics.

*C. Multi-modal Spoof Detection using Deep Fusion Networks*

Using DenseNet in a fusion architecture to combine visual and auditory cues to improve detection performance of complex video tasks. The deep formula combines dense dense mesh network with multi-modal learning, which can obtain different models of different models. This approach leverages interactive interactions to improve discrimination and robustness to different processes. Validation experiments show improved performance compared to a single method and highlight the importance of using multiple cues for fraud detection. This study investigates a new listening technique in deep interaction to prioritize important information with each change, thus making it easier to manage more accurately. The benefits of using different types of benchmarking methods in terms of efficiency and immediate performance are discussed.

*D. Fine-grained Video Analysis with DenseNet*

Complete DenseNet blocks for detailed video manipulation detection by focusing on local feature extraction and highlighting subtle changes. Using dense connections, dense, dense, dense, dense, dense, dense, dense dense structure can retain fine-grained spatial information in the video image, thus facilitating the identification of the true nature of artificial objects. DenseNet's hierarchical feature representation allows the discovery of subtle differences expressed by fakes, thus increasing scientific analysis data. This study includes a monitoring system to identify key areas for monitoring, interpretation and development of regional capacity. Experimental results demonstrate the effectiveness of using DenseNet for video quality analysis of various functions and data. Concepts for using DenseNet-based models in real-world video forensic applications are discussed, highlighting the need for computational efficiency and scalability.

*E. Temporal Learning using DenseNet*

Explore DenseNet's ability to learn physical features in spatiotemporal fraud detection, improving the model's ability to detect dynamic activity.

This study investigates a 3D convolutional extension of DenseNet for modeling the time between video images, enabling a qualitative analysis of spatiotemporal activity patterns. DenseNet's dense connections facilitate time dilation, making the model more sensitive to the physical inconsistencies that characterize work. This method uses modified new unemployment to analyze physical parameters that enable the model to distinguish between natural and operational conditions. Experimental analysis demonstrates the effectiveness of DenseNet-based physical learning in detecting complex video forgery across different data sources.

This study discusses the benefits of integrating anatomical studies into forensic video analysis techniques, highlighting the importance of anatomical details in management activities.

*F. Effective Video Forensics using Lightweight GoogLeNet*

A lightweight variant of GoogLeNet is designed for effective video forensics and provides fast transmission while maintaining accuracy. This work optimized the architecture and parameters of GoogLeNet to reduce computational complexity without affecting the search.

By leveraging lightweight Inception modules, the model achieves scalability and efficiency, making it suitable for use in resource-constrained environments. The results attempt to demonstrate the effectiveness of different GoogLeNet methods in exploring various video games with minimal budget. This work provides insight into optimizing deep learning for video forensics applications by discussing the trade-off between sample size, fast inference, and predictive accuracy. Practical requirements for integrating GoogLeNet's weighting model into video analysis and content verification are discussed, highlighting the importance of efficiency in real-world deployments.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

### *G. Deep Learning Based Forgery Detection Using GoogLeNet*

A deep learning based forgery detection is recommended using GoogLeNet in terms of precise extraction of spatial and physical features. Thanks to the integration of the Inception module, the model captures a hierarchical representation of the video content, enabling the detection of fraud in different types of work. This work investigates adaptive learning strategies to adapt the GoogLeNet model before a specific error detection task, thereby increasing the comprehensiveness and performance of the model. Analysis of test data demonstrates the effectiveness and efficiency of GoogLeNet-based error detection. This study helps understand deep learning in video forensics by discussing the translation and interpretation of GoogLeNet features in forensic science. The benefits of using GoogLeNet-based fraud detection systems in real-world applications are discussed, challenges and future directions in this field are presented.

## III. PROPOSED METHODOLOGY

The program was created and used as a way for the community to solve the problem caused by deepfake videos. First, two datasets are selected: Face Forensics++ and Celeb-DF (v1), which provide a variety of real and manipulated face data. Learning options include DenseNet and GoogLeNet, which are known for their ability to detect complex features in images and videos. To train the model, in addition to training from scratch and using pre-training weights to take advantage of the learning curve, additional data strategies

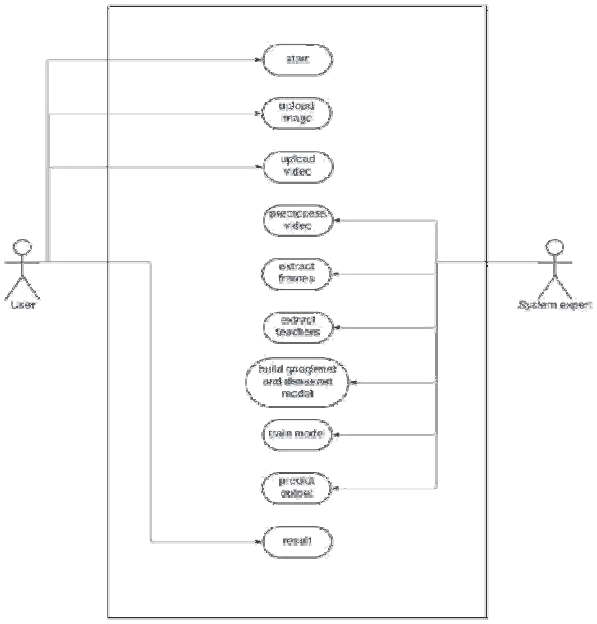



Fig. 1. Use Case Diagram of Our Proposed Model

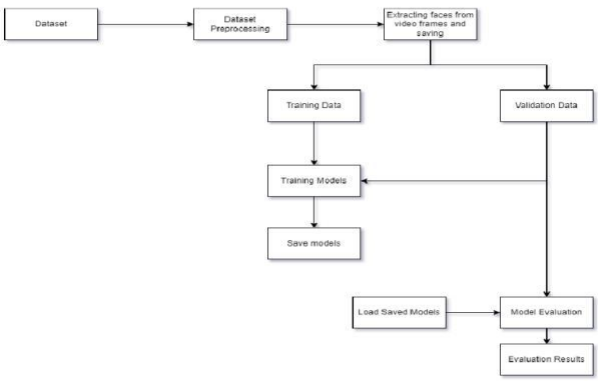



Fig. 2. Data Flow Diagram of Our Proposed Model

should be used to improve proficiency standards. Deepfake content can be easily identified by extracting relevant features from video analysis used to identify facial alterations. Measure performance. The learning process is used to combine the predictions of individual models to improve the accuracy of the overall detection. The project aims to create a good model for verifying the truth through comprehensive evaluation of results and performance evaluation, thus strengthening social influence against improper use of face adjustment tools.

### *A. Data Collection*

The main task of data collection is to collect large files containing real and in-depth videos. A realistic movie requires multiple locations, lighting and camera angles to provide realistic content. When it comes to deepfake videos, many aspects of management, software tools, and data need to be involved to make them look different from the real thing. Educational monitoring should record videos to understand whether they are real or fake. To prevent bias in the training model, real and deep videos must be classified. First of all, ethical issues such as obtaining informed consent and protecting privacy rights during data collection should be taken into account in the first place.

### *B. Data Preprocessing*

In pre-processing, frames are taken from the video, converted to its original size, and normalized so that the pixel values fall into a predetermined order. Additionally, a tag will be used to indicate the accuracy of the post (true or false). Ensuring that the preliminary process checks the consistency and integrity of the material is essential for the CNN model to learn during training and theory and be able to distinguish real from fake videos. By accurately processing prior information, the CNN model can accurately and adaptively predict whether previously unseen videos are legitimate.

### *C. Recommendations and Results*

For deep networking, dense networking outperforms dense networking due to its dense structure that supports reuse and usage with less. It is very good to capture the smallest details of the orientation work. In contrast, GoogleNet's Inception module can extract features from multiple parameters that help identify different parts of the image. The decision of the two designs depends on many factors such as computing resources and data size, although both have proven effective in computing work. By experimenting with deep search data, the most appropriate visualization can be found.

### *D. Data Augmentation*

This is crucial for training deep learning models for deep learning. Adding noise, translation, rotation, scaling, and other techniques to real and fake videos can improve the overall capabilities of the model and make it better for editing. In addition to reducing over fitting, augmentation also increases model accuracy in analyzing edited videos.

### *E. Data Partitioning and Architecture*

Data classification separates data into training, validation, and testing methods for deep video detection. Models are trained according to the training process; During training of the validation process, hyperparameters are tuned and the per-formance of the model is used to evaluate the performance of the training model on untested data. Commonly used models include ResNet, VGG, and custom network design. Activations and release mechanisms for distribution, such as ReLU, are often placed behind multiple connections, links, and layers in these networks. Supervised techniques or neural networks (RNN) can be used to continue capturing physical connections in the video. The complexity of the task, computing power, and the specific needs of the monitoring system influence the chosen architecture.

## IV. RESULTS

The result of our study underscore the urgent need for deeper research techniques to counter the threat of synthetic media. Problems with compressing video frames and extending the device deeper into the audio are speed-related operations for detecting different types of media. Our research demonstrates the effectiveness of deep learning in identifying subtle differences between real and practical context and highlights the importance of generating unique data to improve algorithm training and accuracy. Addressing complex systems such as artificial neural networks (GAN) requires specific strategies for the development of deep learning technologies. In the future, our researchers have prioritized the development of search algorithms, the use of deep learning techniques, the management of complex data, ethics and law to effectively deal with the risks from deepfake videos and protect ourselves from abuse and fraud.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

## V. FUTURE SCOPE

*1)* Detection Expansion to Audio:Deep perception involves analyzing sound in addition to visual content. Explore ways to identify audio variables that can be used with existing video detection techniques.

*2)* Real-Time identification:Instantly detect and reduce deep-fake content across multiple online sites and applications and improve the efficiency and speed of deepfake detec-tion algorithms.

*3)* Multimodal Fusion Techniques: Explore ways to combine multiple variables (e.g. text, audio, and video) into a single framework to get deep insights. Discover how inte-grating data from multiple sources can increase discovery accuracy and reliability.

*4)* Privacy-Preserving Deepfake Detection: Create a system for in-depth discovery that prioritizes customer privacy while providing accurate information. Explore methods like privacy differences or state training to provide group training without compromising personal information.

*5)* Hardware-Embedded Solutions: To achieve this, to smol-der, to smolder, to check out how deep learning models can be integrated into edge devices to see into the deep.

*6)* Cross-Modal Deepfake Detection:Check the process to identify a deep sync state where many processes (such as

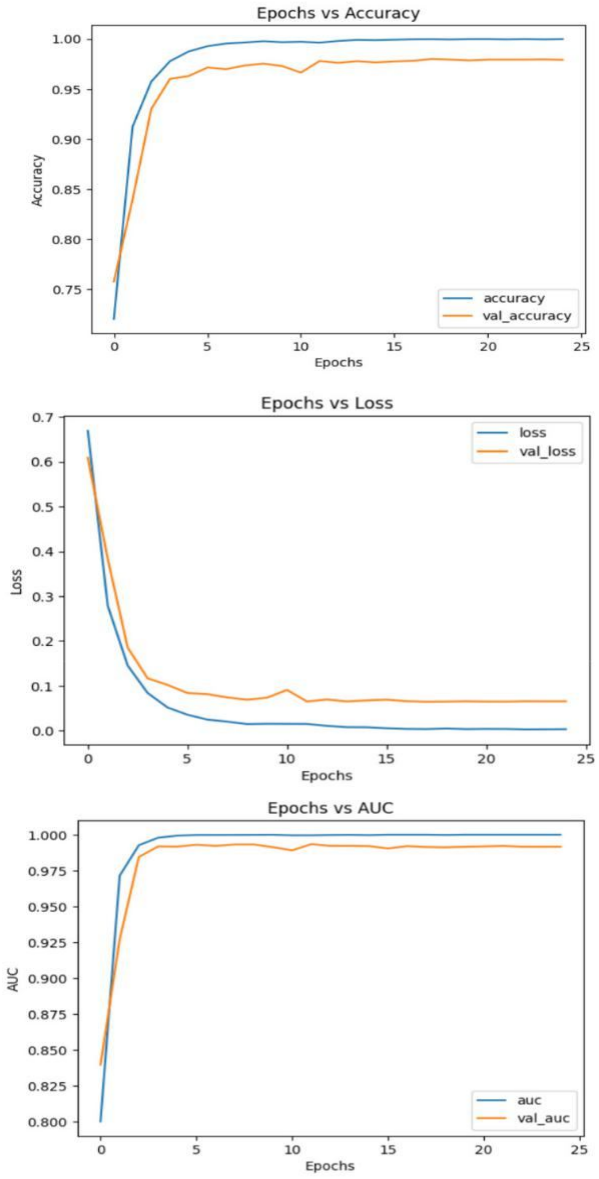



Fig. 3. Loss, Accuracy and AUC graphs of DenseNet(pre-trained)

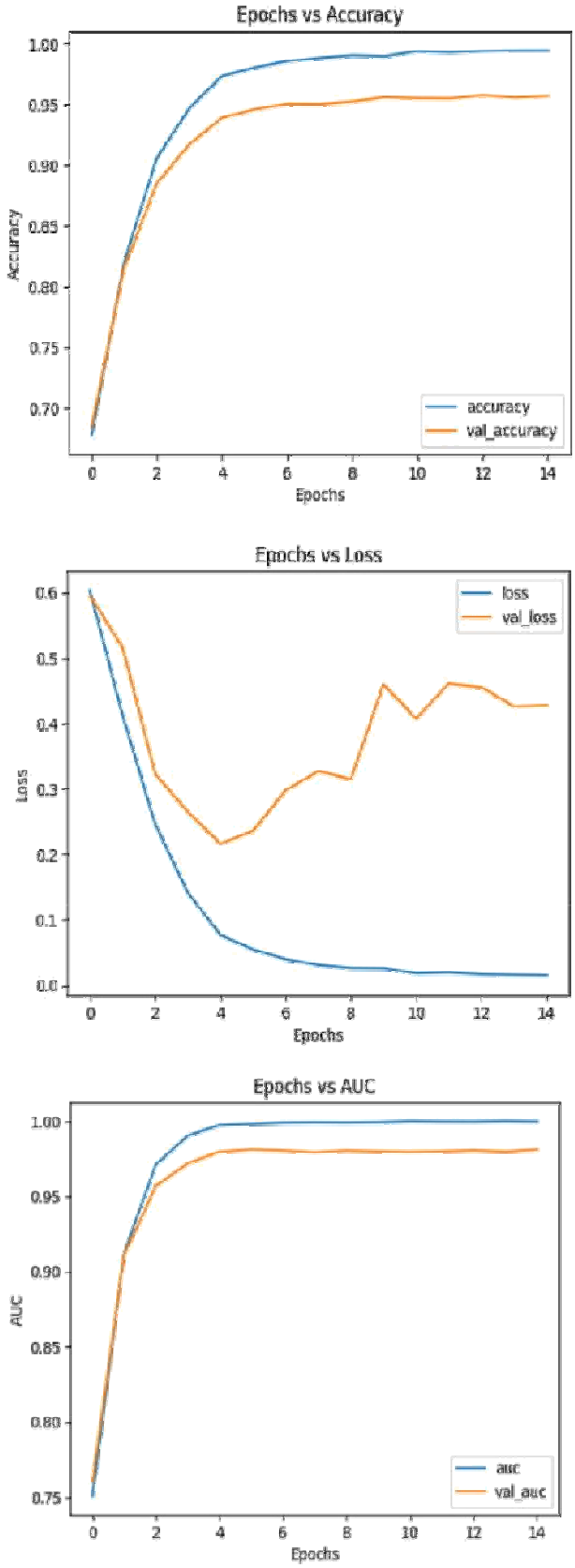



Fig. 4. Loss, Accuracy and AUC graphs of GoogLeNet(pre-trained)

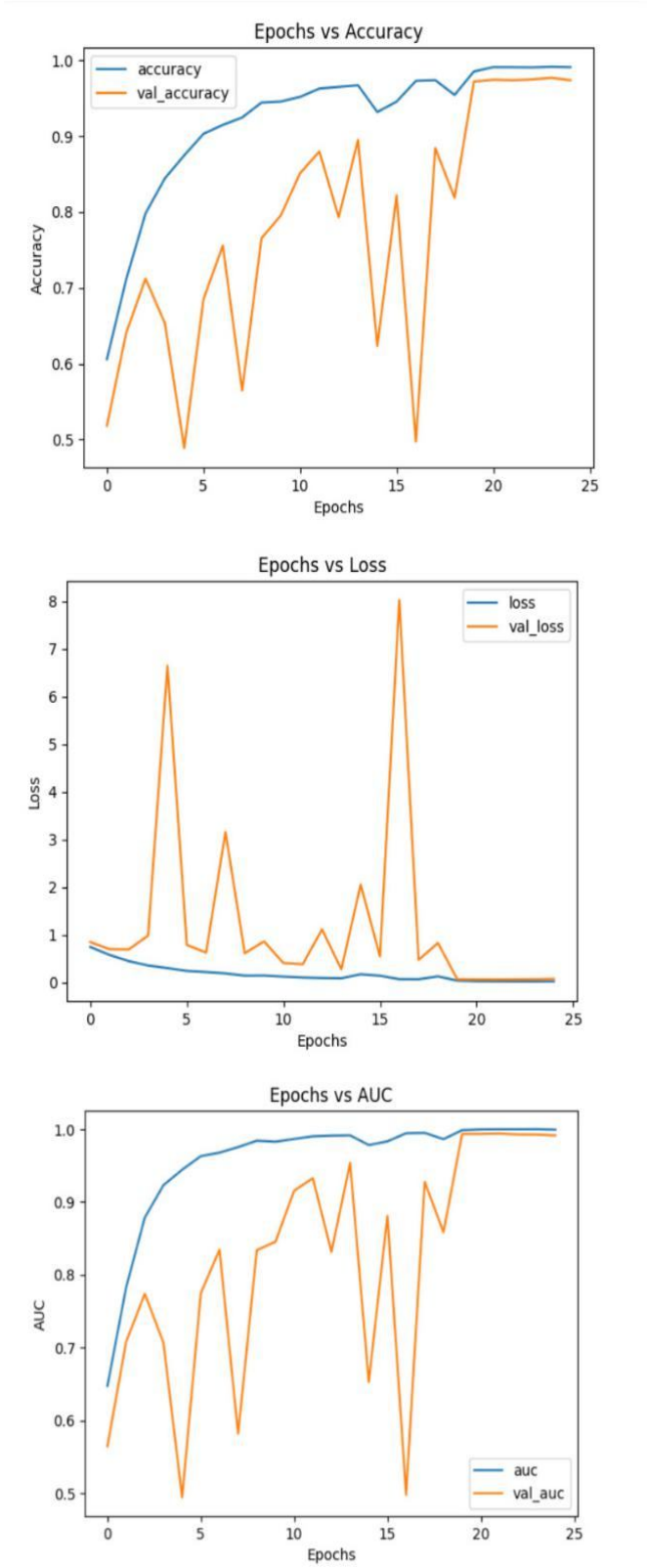



Fig. 5. Loss, Accuracy and AUC graphs of DenseNet(Scratch)

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

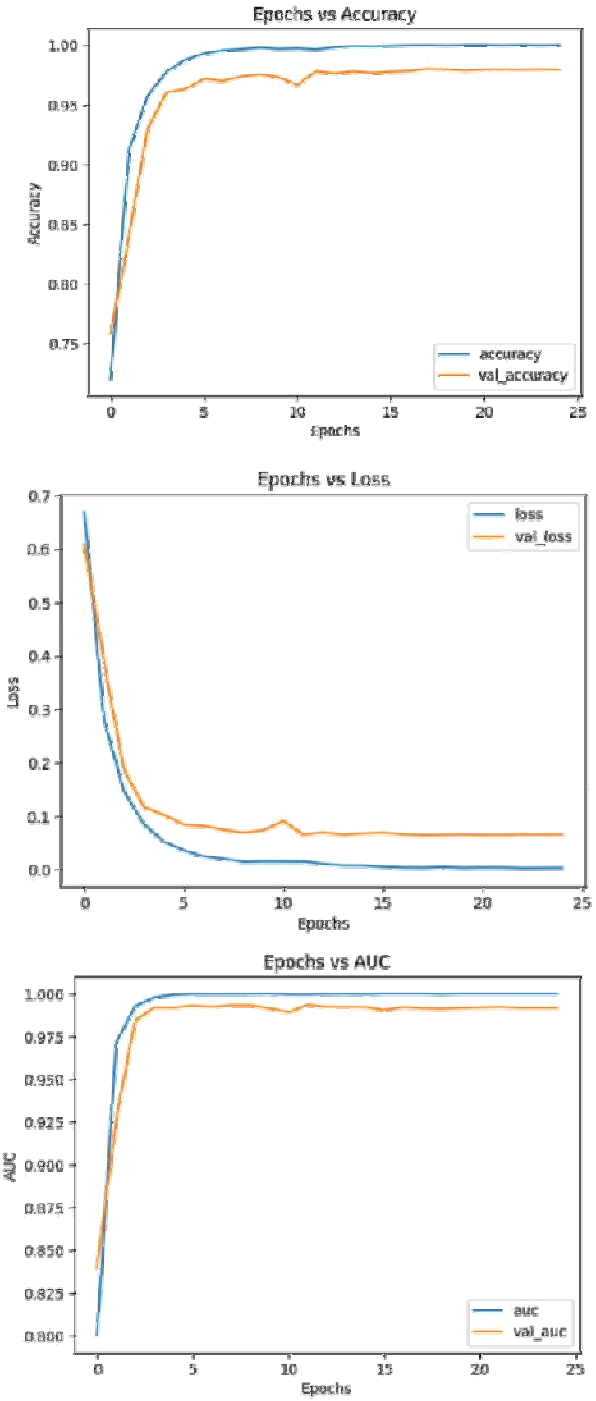



Fig. 6. Loss, Accuracy and AUC graphs of GoogLeNet(Scratch)

| Model | Loss | Accuracy | Val Loss | Val Accuracy |
|---|---|---|---|---|
| DenseNet(pre-trained) | 0.0034 | 0.9997 | 0.0643 | 0.9800 |
| GoogLeNet(pre-trained) | 0.0765 | 0.9730 | 0.2161 | 0.9387 |
| DenseNet(scratch) | 0.0184 | 0.9905 | 0.0600 | 0.9733 |
| GoogLeNet(scratch) | 0.0276 | 0.9882 | 0.0696 | 0.9700 |

Fig. 7. Best validation loss epoch

text, audio, and video) are running simultaneously. De-velop methods to detect discrepancies and discrepancies between various models to improve the accuracy of the analysis.

In summary, future directions of deep search point to important research to combat the growing threat of synthetic media. Challenges such as processing compressed video frames and extending deep processing to audio imply the need for advanced search that can handle different types of media. Deep learning algorithms are promising at detecting subtle differences between real and practical content, but progress depends on iterative refinement of deep learning data for effective learning. Addressing technologies such as news generators (Gans) Future research should focus on improving search algorithms, using deep learning, generating data details, and consider ethical and legal measures to reduce the risks associated with fake videos and prevent abuse and fraud.

## VI. CONCLUSION

In this paper, we examine the effectiveness of combining DenseNet and GoogLeNet models in face detection in videos, focusing on Deepfake detection. We aim to achieve comprehensive and robust results by training these models from scratch and using pre-training weights on Celeb-DF (v1) and Face Forensics++ datasets. Our results show that the pre-trained model performs well on the data, while the training model performs better when tested on different datasets (e.g., DFDC data). The actual accuracy of GoogLeNet, which has been working on it for a while, is 0.9700. Therefore, we conclude that the training model for face detection examined in this paper is more effective, and we point out that the pre-training model selection and training model should be based on data features and decisions. This research provides insight into how deep learning can be used to combat fake videos such as deepfakes.

## REFERENCES

[1] Z. Akhtar, M. R. Mouree and D. Dasgupta, ”Utility of Deep Learning Features for Facial Attributes Manipulation Detection,” 2020 IEEE International Conference on Humanized Computing and Communication with Artificial Intelligence (HCCAI), Irvine, CA, USA, 2020, pp. 55-60, doi: 10.1109/HCCAI49649.2020.00015.

[2] M. Patel, A. Gupta, S. Tanwar and M. S. Obaidat, ”Trans-DF: A Transfer Learning based end-to-end Deepfake Detector,” 2020 IEEE 5th International Conference on Computing Communication and Au tomation (ICCCA), Greater Noida, India, 2020, pp. 796-801, doi: 10.1109/ICCCA49541.2020.9250803.

[3] S. Suratkar, F. Kazi, M. Sakhalkar, N. Abhyankar and M. Kshir sagar, ”Exposing DeepFakes Using Convolutional Neural Networks and Transfer Learning Approaches,” 2020 IEEE 17th India Council International Conference (INDICON), New Delhi, India, 2020, pp. 1-8, doi: 10.1109/INDICON49873.2020.9342252.

[4] B. Zi, M. Chang, J. Chen, X. Ma, and Y.-G. Jiang, WildDeep fake: A challenging real-world dataset for deepfake detection, 2021, arXiv:2101.01456.

[5] M. Dordevic, M. Milivojevic, and A. Gavrovska, “DeepFake video anal ysis using SIFT features,” in Proc. 27th Telecommun. Forum (TELFOR), Nov. 2019, pp. 1–4.

[6] Deepfake Images Detection and Reconstruction Challenge—21st Inter national Conference on Image Analysis and Processing. Accessed: Jan. 5, 2023. [Online]. Available: https://iplab.dmi.unict.it/Deepfake chal-lenge

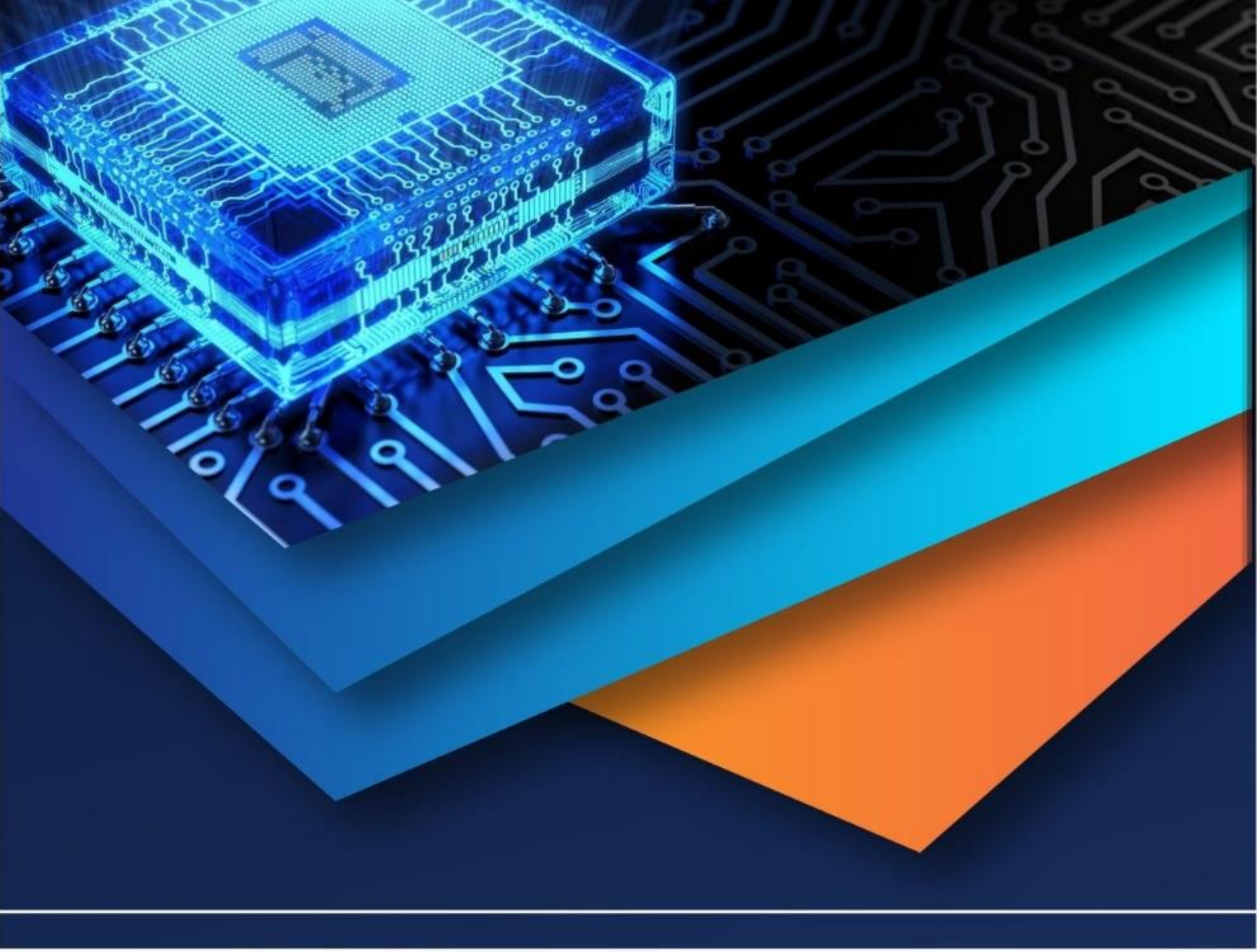
[7] J. Hui. How Deep Learning Fakes Videos (Deepfake) and How to Detect it. Accessed: Jan. 4, 2021. [Online]. Available: https://medium. com/how-deep-learning-fakes-videos-deepfakes-and-how-to-detect-it c0b50fbf7cb

[8] FaceApp. Accessed: Jan. 4, 2021. [Online]. Available: https://www. faceapp.com/

[9] H. Kim, P. Garrido, A. Tewari, W. Xu, J. Thies, M. Niessner, P. P´erez, C. Richardt, M. Zollh¨ofer, and C. Theobalt, Deep video por traits, ACM Trans. Graph., vol. 37, no. 4, pp. 114, Aug. 2018, doi: 10.1145/3197517.3201283.

[10] J. Hernandez-Ortega, R. Tolosana, J. Fierrez, and A. Morales, DeepFakesON-phys: DeepFakes detection based on heart rate estima tion, 2020, arXiv:2010.00400.

[11] ”DeepFakesON-phys: DeepFakes detection using heart rate estimation,” 2020, arXiv:2010.00400.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)