VERITY VISION: Holistic Anti- Spoofing and Fabricated Content Recognition System Implementing Mobile Net and Resnext

Archana J R[1], Chaithra S K[2], Kavana B M[3], Shreya N R[4], Prof. Nikita Kunachi[5]
[1,2,3,4] UG Students, Department of Computer Science and Engineering,
Bahubali College of Engineering, Shravanabelagola
[5]Assistant Professor, Department of Computer Science and Engineering
Bahubali College of Engineering, Shravanabelagola

***Abstract*—Verity Vision is a deep-learning anti-spoofing and content authentication system that detects fake or manipulated media. It couples the speed and real-time feature extraction capability of Mobile Net with the strong classification capability of ResNeXt to set apart genuine from spoofed inputs. It can defend against printed photos, replayed videos, 3D masks, and AI-generated deepfakes, making it suitable for secure biometrics, surveillance, and digital content verification. The system runs on mobile and cloud platforms for flexibility and scalability. It emphasizes security, reliability, and user-centric design with secure media handling, strict content validation, and finally, structured logging in order to prevent misuse and aid audits in banking, e-governance, and remote identity verification. Predictions and authenticity reports are presented through an intuitive web interface and API for uncomplicated integration. Its lightweight architecture, platform-agnostic and extensible, allows for continuous updates with new datasets and attack patterns to make it resilient against evolving spoofing and deepfake techniques.**

***Index Terms*—Anti-spoofing, Fake Media Detection, Hybrid CNN Architecture, Deepfake detection, Mobile Net, ResNeXt.**

I. INTRODUCTION

The authentication systems and multimedia platforms play a vital role in today's digital world for secure identification and information exchange. However, with growing dependence on the digital operating systems, the spoofing attacks and spread of fabricated contents have also started to grow. An attacker may manipulate vulnerabilities in face recognition systems using printed photographs, replayed videos, 3D masks, and hyper-realistic deepfakes generated with AI.

These attacks compromise the security of authentication systems, besides damaging the credibility of digital media by allowing misinformation, identity theft, and fraud. Most traditional methods of detecting spoofing or digital manipulation rely on rule-based techniques, which are narrow in scope and non-viable against rapidly changing attacker tactics. There is, therefore, an emerging need for intelligent, adaptive, and real-time solutions that accurately detect and block such threats before actual damage can occur.

It leverages two advanced deep convolutional neural network architectures, namely Mobile Net and ResNeXt, toward efficient and effective detection of fake or manipulated content. Being a lightweight model, Mobile Net allows for feature extraction in a much faster and resource-friendly way, thus making it perfectly suitable for a mobile and real-time application. Correspondingly, ResNeXt enhances the accuracy and generalization of the system with its powerful classification capabilities due to multi-path aggregated transformations.

Verity Vision combines the efficiency of Mobile Net with the representational strength of ResNeXt to set up an all-inclusive framework that can handle both biometric spoofing and recognition of fabricated content in one single pipeline. The system enhances security in authentication while reinforcing trust in digital communication through the detection and mitigation of misinformation. This project ultimately focuses on creating a solution that is scalable, intelligent, and future-ready, safeguarding identity verification and content authenticity in an era increasingly challenged by synthetic media. With a dual emphasis on both speed and accuracy, it can find widespread adoption in highly sensitive sectors such as banking, e-governance, and social media moderation in order to ensure resilience against ever-evolving threats. Therefore, the system sets a new benchmark for secure and trustworthy digital ecosystems with respect to integrating state-of-the-art deep learning with practical deployment strategies.

II. METHODOLOGY

A. SYSTEM OVERVIEW

Verity Vision is a deep learning–based anti-spoofing and fabricated content recognition system tailored for digital identity verification and multimedia platforms. It detects spoofing attacks such as printed photos, replayed videos, 3D masks, and AI-generated deepfakes. The system adopts the hybrid CNN architecture that combines Mobile Net for fast and light feature extraction and ResNeXt for accurate classification. This combination assures real-time performance with high detection accuracy and can operate both on mobile and in the cloud environment. Verity Vision processes images and video frames through preprocessing, feature extraction, and classification, yielding predictions regarding the authenticity of the content. Secure media handling, structured logging, and API-based integration enhance the reliability and auditability of the solution. Verity Vision has a scalable and extensible design to cater to a wide array of industry needs, including banking, e-governance, surveillance, and remote identity verification, keeping it resilient against emerging spoofing and deepfakes threats.

B. SYSTEM ARCHITECTURE

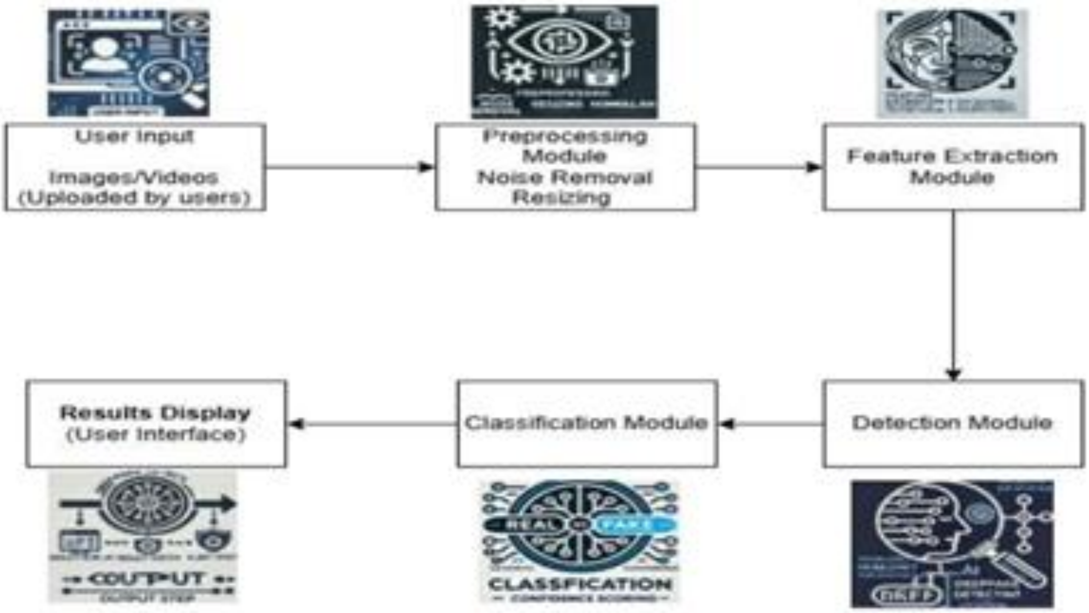


FIG. 1: THE ARCHITECTURE OF THE SYSTEM AND FLOW OF DATA IN THE PROPOSED VERITY VISION FRAMEWORK. THE MODULAR PIPELINE STARTS FROM DATASET COLLECTION AND PREPROCESSING AND PASSES THROUGH ENSEMBLE MODEL TRAINING AND EVALUATION, CLASSIFICATION MODULE AND DISPLAY RESULTS [11]

C. DATASET AND FEATURE EXTRACTION

The training and evaluation of the Verity Vision system are performed with a diverse dataset that includes both real and spoofed media, such as printed photos, replayed videos, 3D masks, and AI-generated deepfakes. The dataset like FaceForensic++, Deepfake Detection Challenge datasets (DFDC) [3] and ASVspoof [9]. For consistency and robustness, all input images and videos will undergo preprocessing, including face detection, cropping, resizing, normalization, and video frame extraction. The dataset is split into training, validation, and testing subsets to avoid biased model evaluation. Feature extraction is by Mobile Net, a lightweight CNN designed for real-time applications. Mobile Net extracts facial texture patterns, reflectance variations, motion cues, and spatial inconsistencies that distinguish real faces from spoofed or manipulated content. These high-level feature vectors are then fed into the ResNeXt classifier, which, through its multi-path transformation architecture, serves as an accurate classifier. This approach provides efficient processing, high accuracy in detection, and suitability for both mobile and cloud deployment, hence it effectively resolves several spoofing and deepfake attack scenarios.

D. MODELS OF DEEP LEARNING

Mobile Net is a light-weight convolutional neural network that is optimized for fast feature extraction. It is optimized for fast computation. Mobile Net has the ability to extract important spatial and textural information for the input image or video. In the context of Verity Vision, Mobile Net is used to extract the face input to provide higher-level features of face textures, even reflectance, motion cues, as well as the slightest inconsistencies that might detect face spoofing. Mobile Net is well suited in the context of face anti-spoofing since the system is able to perform feature extractions in real time. It is efficient even in the context of light-weight hardware. Mobile Net is able to achieve the aforementioned through its innovative depthwise separation of convolutions. In this concept, regular convolutions are comprised of two smaller computations—depthwise convolution as well as pointwise convolution.

ResNeXt is a CNN-powered classifier, which is an effective complement to Mobile Net's feature extraction capabilities. ResNeXt's architecture is designed based on a multi-path, cardinality-focused philosophy, according to which various parallel paths related to convolution operations are grouped together for more advanced feature extractions. The ability of ResNeXt to identify complex features, with subtle differences between original and spoofed samples, is innovative. In the Verity Vision system, ResNeXt takes feature representations produced by Mobile Net, classifying them as genuine or spoofed samples. Its applicability and robustness enable the system to generalize well towards different attack modes, whether it is printed picture attacks, video attacks, 3D face mask attacks, or AI-generated deepfakes.

The hybrid integration of Mobile Net and ResNeXt constitutes a pipeline that maintains a good trade-off between efficiency and accuracy. Mobile Net is capable of extracting crucial features from the input very quickly, whereas ResNeXt is tasked with precise classification of these features. This hybrid model makes it possible for Verity Vision to work in real-time, address various spoofing attacks concurrently, and be scalable to multiple platforms such as mobile and cloud. The model will also be easily updatable to address various forms of attacks in the future.

E. TRAINING AND PERFORMANCE EVALUATION

These include several key performance metrics, such as accuracy, precision, recall, and frame processing time. Accuracy shows how often the system makes correct predictions overall, while precision indicates how many of the instances labelled as fake are actually fake. Recall represents how well the system can identify all the fake cases correctly; it shows the effectiveness in capturing spoofing or deepfake attempts. Frame processing time is further analysed to find out how effectively the system processes individual images or video frames, which is considered crucial for real-time applications. To validate robustness even more, unseen datasets should be tested. Careful dataset selection ensures that the training and evaluation performed for the Secure Vision system used only widely recognized datasets, aiming for robustness as well as practical applicability. The ASVspoof dataset was used for training Mobile Net because of the extensive range of spoofing attacks included, such as printed photo and video replay attacks, which can closely model real-world scenarios where biometric authentication is concerned. ResNeXt was trained using the FaceForensics++ and DFDC datasets since they contain diverse and high-quality examples of deepfake manipulations, a great variation of the techniques used, and levels of sophistication therein. These are very representative datasets under real-world conditions and are thus very commonly used within the community, allowing the system to generalize well to unseen threats and varying environmental factors.

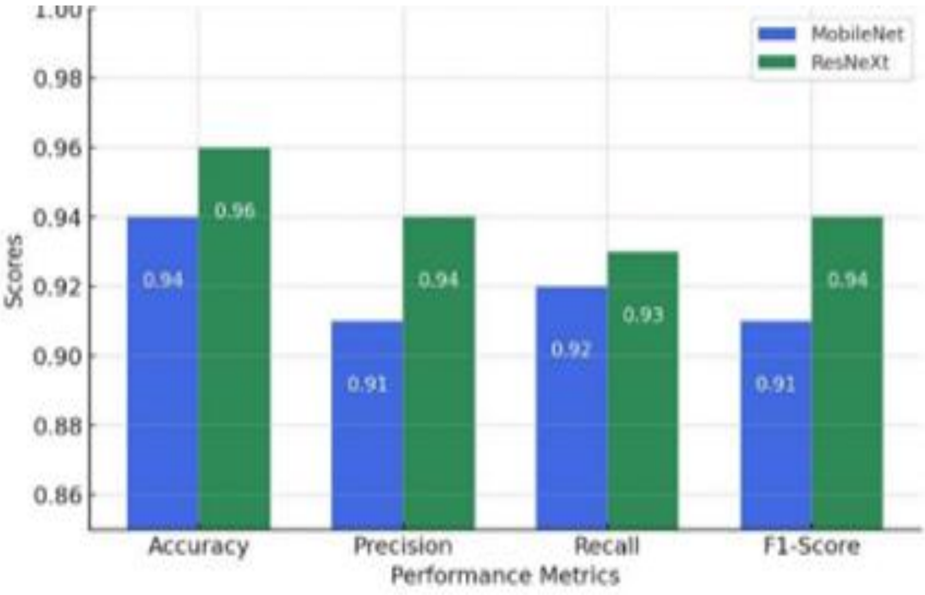
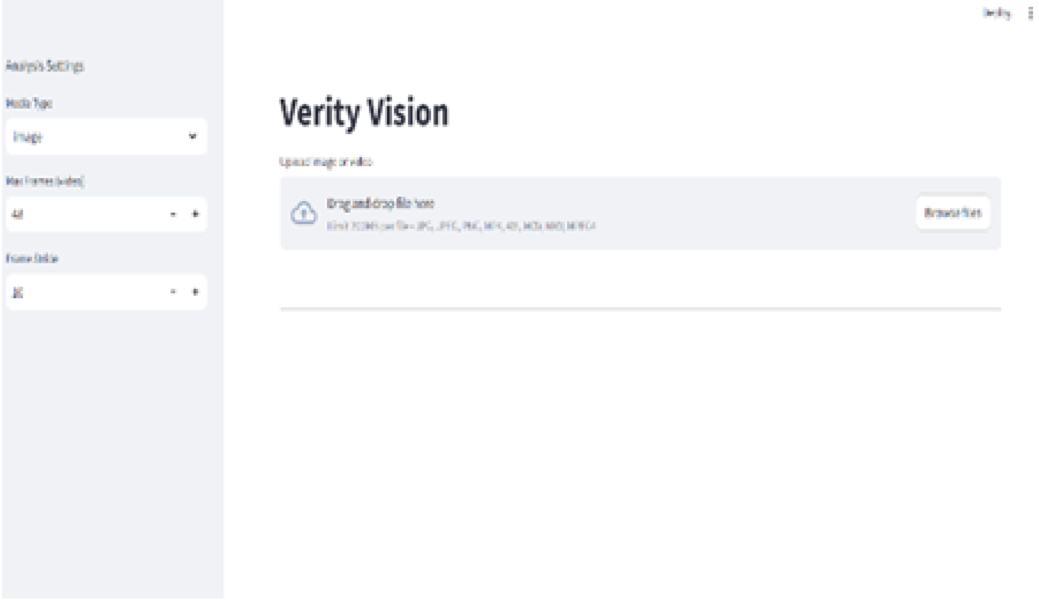


FIG. 2. PERFORMANCE METRICS AMONG MOBILE NET AND RESNEXT [11]

III. RESULTS


Analysis Settings

Media Type

Image

Max Frames (video)

Frame Delay

Verity Vision

Upload image or video

Drag and drop file here

FIG. 3. USER INTERFACE FOR UPLOADING IMAGES AND DISPLAYING REAL/FAKE IMAGES.

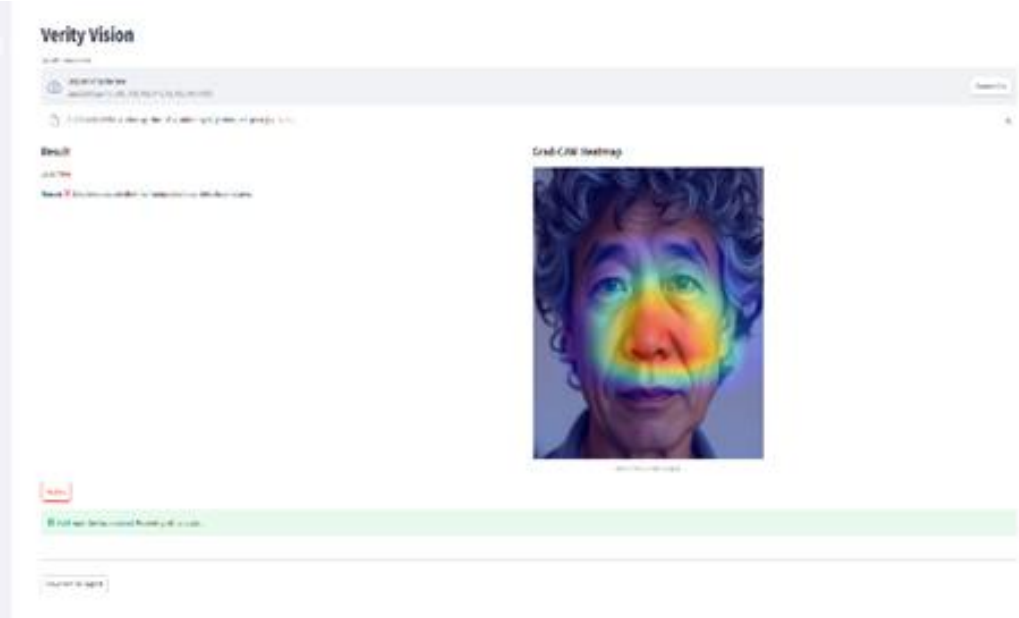
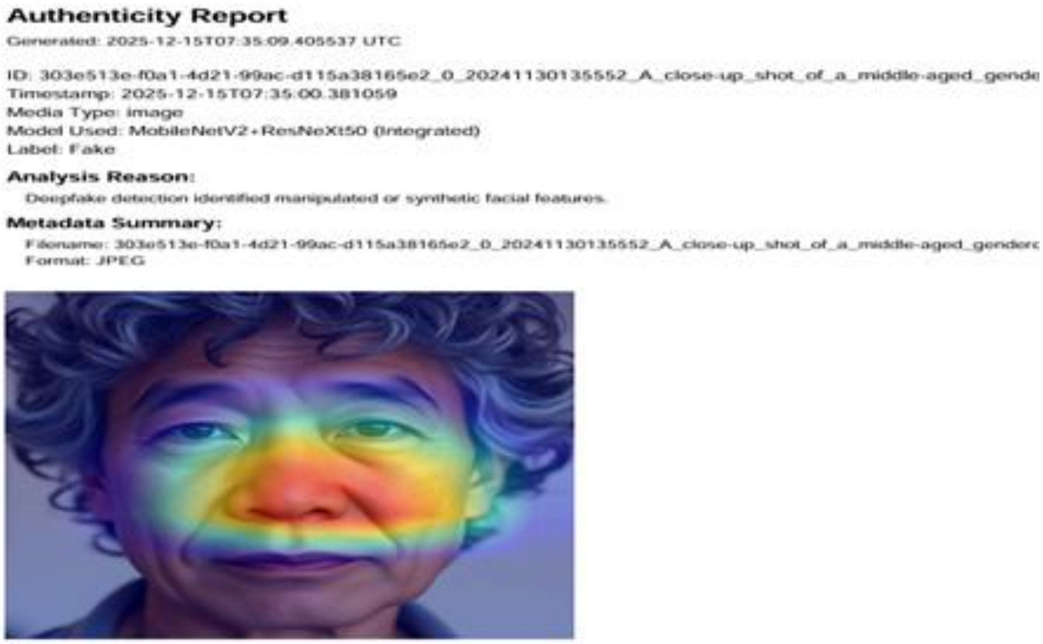


FIG. 4. OUTPUT OF THE VERITY VISION SYSTEM SHOWING THE UPLOADED IMAGE IS FAKE.



Authenticity Report

Generated: 2025-12-15T07:35:09.405537 UTC

ID: 303e513e-f0a1-4d21-99ac-d115a38165e2_0_20241130135552_A_close-up_shot_of_a_middle-aged_gende

Timestamp: 2025-12-15T07:35:00.381059

Media Type: image

Model Used: MobileNetV2+ResNeXt50 (Integrated)

Label: Fake

Analysis Reason:

Deepfake detection identified manipulated or synthetic facial features.

Metadata Summary:

Filename: 303e513e-f0a1-4d21-99ac-d115a38165e2_0_20241130135552_A_close-up_shot_of_a_middle-aged_genderc

Format: JPEG

FIG. 5. GENERATE THE AUTHENTICITY REPORT AFTER ANALYSIS WHICH CONSISTING OF IMAGE ID, TIMESTAMP, MEDIA TYPE, LABEL, ANALYSIS REASON AND METADATA SUMMARY WITH THE GRAD-CAM HEATMAP.

IV. CONCLUSION

The anti-spoofing and deepfake detection system uses Mobile Net and ResNeXt to identify the fake image by analyzing facial features and texture patterns. Harnessing such a deep learning architecture result in high accuracy and real-time performance against unauthorized access via printed photos, replayed videos, or altered faces. This further assists in increasing face recognition reliability and authenticating all applications to protect digital identity.

Mobile Net guarantees effective deployment on resource-constrained devices, while the strong feature extraction by ResNeXt increases precision in spotting real and fake images. They come together to increase speed and provide higher accuracy in detection, thereby making the system appropriate for biometric security, online identity verification, and financial systems; hence, showcasing the potential of deep learning toward secure, trustworthy digital ecosystems.

REFERENCES

[1] Reshma Sunil, Parita Mer, Anjali Diwan, Rajesh Mahadeva, Anuj Sharma, "Exploring autonomous methods – a detailed survey paper known about the different algorithms". In "Heliyon", pp. no. 23, Vol. 11, 2025.

[2] Jianyu Xiao, Wei Wang, Lei Zhang, Huanhua Liu, "A MobileFaceNet-Based face anti-spoofing algorithm for low-quality images". In "Electronics", pp. no.14, Vol. 13, 2024.

[3] Trupti Kularkar, Tanvi Jikar, Vansh Rewaskar, "Deepfake detection using LSTM and ResNeXt". In "International Journal of Creative Research Thoughts", pp. no. 9, Vol. 11, 2025.

[4] Sayyam Zahra, Mohibullah Khan, Kamran Abid, Naeem Aslam, Ejaz Ahmad Khera, "Face anti-spoofing and liveliness detection using Mobile Net and Haar cascade algorithm". In "International Journal of Scientific Research in Science and Technology", pp. no. 14, Vol. 10, 2023.

[5] Sayyam Zahra, Mohibullah Khan, Kamran Abid, Naeem Aslam, Ejaz Ahmad Khera, "A novel face spoofing detection using hand crafted Mobile Net". In "VFAST", pp. no. 10, Vol. 11, 2023.

[6] Gourav Gupta, Kiran Raja, Manish Gupta, Tony Jan, Scott Thompson Whiteside, Mukesh Prasad, "A Comprehensive review of Deepfake detection using advanced machine and fusion methods". In "Electronics", pp. no. 27, Vol. 13, 2024.

[7] Leandro A. Passos, Danilo Jodasa, Kelton A. P. Costaa, Luis A. Souza Junior, Douglas Rodrigues, Javier Del Ser, David Camacho, Joao Paulo Papa, "A review of Deep learning-based approaches for Deepfake content detection". In "arXiv", pp. no. 84, Vol. 3, 2024.

[8] Azeddine Benlamoudi, Salah Eddine Bekhouch, Maarouf Korichi, Khaled Bensid, Abdeldjalil Ouahabi, Abdenour Hadid, Abdelmalik Taleb-Ahmed, "Face Presentation Attack Detection Using Deep Background Subtraction". In "sensors", pp. no. 17, Vol. 22, 2022.

[9] Farhan, Md Huzaif Shah, Aditya S, Archana S, "Secure Vision: Integrated Anti Spoofing and Deepfake Detection System Using Mobile Net and ResNeXt" In "International Journal of Emerging Technologies and Innovative Research", pp. no. 5, Vol. 11, 2024.

[10] Hala S. Mahmood, Salah Al-Darraji "Face Anti-Spoofing Detection with Multi-Modal CNN Enhanced by ResNet: Face Anti-Spoofing Detection" In "Journal of Basrah Researches (Sciences)", pp. no. 11, Vol. 50, 2024.

[11] Mathan Sri Prasath E, Mahesh Kumar D, Naveen G V "Integrated Anti-Spoofing and Deepfake Detection system using Mobile Net and ResNeXt" In "International Journal of Engineering Development and Research", pp. no. 12, Vol 13, 2025.