Future of Digital Security: AI-driven Anti-Spoofing and Deepfake Detection

[1] Syed Rayyan Pervez, [2]Md Saqib, [3]Fariya Tasneem, [4]Sidra Fatima, [5]Prof. Laxman Singh

[1234]Under Graduate, Department of AI&ML-GNDEC-Bidar

[5]Asst. professor, Department of AI&ML-GNDEC-Bidar

ABSTRACT:

As artificial intelligence continues to evolve, it has significantly improved biometric security and identity verification. However, it has also given rise to new challenges, such as spoofing attacks and deepfake manipulations. To address these concerns, **Secure Vision** introduces a powerful dual-model approach that combines **MobileNet** and **ResNeXt** to detect and prevent such threats effectively.

Designed for real-time performance, **Secure Vision** delivers high accuracy and adaptability, making it well-suited for applications in biometric authentication, online identity verification, and media authenticity checks. This paper explores the system's architecture, methodology, and experimental results, highlighting its potential to strengthen security and build trust in digital interactions.

Keywords: Anti-Spoofing, Deepfake Detection, MobileNet, ResNeXt, Biometric Security, Media Authenticity, Real-Time Detection

1. INTRODUCTION

In today's digital landscape, artificial intelligence (AI) and advanced media processing technologies have revolutionized biometric security, digital content creation, and identity verification. However, these innovations have also given rise to sophisticated threats, such as **spoofing attacks** and **deepfakes**, which pose serious security risks.

Spoofing involves tricking authentication systems with fake biometric data—such as photos, videos, or masks—to gain unauthorized access. Meanwhile, **deepfake technology** leverages AI to create hyper-realistic manipulated videos, altering a person's face, voice, or actions. Unfortunately, this technology has been exploited for identity fraud, impersonation, misinformation, and even character defamation.

To combat these growing threats, the **Future of Digital Security** project introduces a **dual-model solution** that integrates **MobileNet** for efficient anti-spoofing detection and **ResNeXt** for deepfake identification. By combining these two models, the system offers a **comprehensive and real-time defense** against identity fraud and media manipulation.

Designed for high-stakes applications—such as **biometric access control, online identity verification, and media authenticity verification**—this solution aims to enhance security and trust in digital interactions.

2. IMPLEMENTATION

The Future of Digital Security project employs a dual-model approach, utilizing MobileNet for anti-spoofing and ResNeXt for deepfake detection, to provide an integrated system capable of identifying both types of visual threats in real time. This chapter details the system design, hardware and software requirements, experimental setup, and training processes involved in developing this comprehensive detection system.

2.1 SYSTEM DESIGN AND APPROACH

The Future of Digital Security system is structured to sequentially process inputs through two

distinct but interconnected models. The system begins with MobileNet for anti-spoofing detection, which checks for liveness cues to prevent unauthorized access via spoofing techniques. If the input passes this initial check, it is then processed by ResNeXt for deepfake analysis, which identifies manipulated or synthetic content.

System Workflow:

- **Input Capture**: Video frames are captured from a camera feed or uploaded video. Frames are preprocessed to ensure they meet the input requirements of both models.
- **Anti-Spoofing Detection with MobileNet:** The captured frames are first processed by MobileNet, which analyzes liveness cues such as eye blinks, subtle head movements, and facial texture. MobileNet's lightweight architecture ensures low latency, allowing real-time spoofing detection.
- **Deepfake Detection with ResNeXt:** If the input passes the anti-spoofing check, it is then sent to ResNeXt, which is trained to detect deepfake artifacts, such as facial texture inconsistencies, unnatural expressions, and irregular lighting effects.
- **Decision Engine:** The final decision engine combines outputs from both models, generating a verdict on whether the input is authentic or manipulated. This result is then displayed to the user or used by an external application, such as a biometric authentication system or a content moderation tool.

Architecture Overview:

- **MobileNet (Anti-Spoofing):** MobileNet's architecture consists of depthwise separable convolutions, which reduce computation by separating spatial and channel convolutions. This design makes MobileNet highly efficient, particularly for real-time applications.
- **ResNeXt (Deepfake Detection):** ResNeXt is a deep convolutional neural network with grouped convolutions, allowing it to capture intricate details and subtle inconsistencies in image features. This model is well-suited for deepfake detection due to its high accuracy in complex visual tasks.

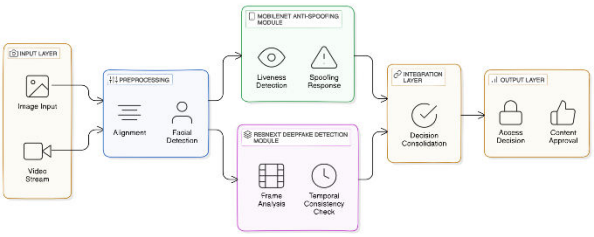
2.2 HARDWARE AND SOFTWARE REQUIREMENTS

Hardware Requirements:

- **GPU:** NVIDIA GTX 1080 or higher, recommended for training the ResNeXt model and speeding up inference during testing.
- **CPU:** Intel i5 or above, for efficient processing during training and testing.
- **Memory:** Minimum of 16GB RAM to handle large datasets and training batches.
- **Camera Module:** For real-time applications, a high-definition camera with at least 30 FPS is recommended to capture input frames.

Software Requirements:

- **Programming Language:** Python 3.8 or above.
- **Libraries and Frameworks:**
  - i) **TensorFlow/Keras:** For model implementation and training.
  - ii) **OpenCV:** For image and video processing tasks, including frame extraction and real-time video feed handling.
  - iii) **Face Recognition Libraries (e.g., dlib)**: Used to align and preprocess faces before they are fed into MobileNet and ResNeXt models.
- **Datasets:**
  - i) **ASVspoof:** For training the MobileNet anti-spoofing model.
  - ii) **FaceForensics++ and Deepfake Detection Challenge (DFDC):** For training the ResNeXt deepfake detection model.



2.3 EXPERIMENTAL SETUP

MobileNet Training (Anti-Spoofing):

Data Preprocessing: Images from the ASVspoof dataset are preprocessed to align faces, standardize resolution, and apply grayscale transformations where needed.

Data Augmentation: To improve generalization, various augmentations are applied, such as random rotations, brightness variations, and slight blurring to simulate real-world conditions.

Liveness Cues Extraction: MobileNet is specifically trained to detect liveness cues such as eye blinks, mouth movements, and micro-expressions. These features help differentiate real faces from spoofed media (e.g., printed photos or video replays).

Training Configuration: The model is trained using binary cross-entropy as the loss function, Adam optimizer, and a learning rate scheduler to adaptively control learning rates.

ResNeXt Training (Deepfake Detection):

Data Preprocessing: Frames from the FaceForensics++ and DFDC datasets are extracted and preprocessed to ensure consistent size and quality.

Data Augmentation: ResNeXt's training data undergoes augmentation techniques such as Gaussian noise addition, blurring, and cropping to enhance the model's robustness to different manipulation techniques.

Deepfake Artifact Analysis: ResNeXt is trained to recognize deepfake-specific artifacts, including unnatural facial textures, asymmetric facial features, and inconsistencies in lighting and expressions.

Training Configuration: The model is optimized using categorical cross-entropy, and trained with a stochastic gradient descent (SGD) optimizer. Learning rate decay and early stopping are implemented to prevent overfitting.

2.4 MODEL INTEGRATION AND REAL-TIME DECISION ENGINE

The integration of MobileNet and ResNeXt into a single system allows for a seamless workflow, with MobileNet handling spoofing detection and ResNeXt analyzing deepfake content**.**

Decision Engine:

Sequential Processing: The decision engine first directs each input through MobileNet. If the input passes the spoofing check, it proceeds to ResNeXt for deepfake analysis.

Output Aggregation: The outputs from both models are combined to generate a final verdict. For example, if MobileNet detects a spoofing attempt, the system flags it as fraudulent without further deepfake analysis. If the input passes the spoof check, ResNeXt's output determines whether the content is genuine or manipulated.

User Notification: The final result (authentic or fraudulent) is then displayed to the user.

2.5 PERFORMANCE METRICS

The performance of the system is evaluated based on metrics such as accuracy, precision, recall, F1 score, latency, and false positive/negative rates.

- **Accuracy**: Measures the percentage of correctly identified instances.
- **Precision and Recall**: Assess the system's ability to balance true positives and false negatives.
- **Latency**: Tracks the average processing time per frame.
- **F1 Score**: Provides a balanced measure of precision and reca

3. RESULTS

Performancce Metrics

Metric	Value(%)
Accuracy	95.2
Precision	96.8
Recall	95.2
F1- Score	96.1

Table 1: MobileNet Anti-Spoofing Performance on ASVspoof Dataset

MobileNet demonstrated an accuracy of 96.1% on the ASVspoof dataset, while ResNeXt achieved 95% on

Metric	Value(%)
Accuracy	94.2
Precision	92.3
Recall	94.1
F1- Score	95

Table 2: ResNeXt Deep Fake Detection Performance on FaceForensics++ Dataset

FaceForensics++. The integrated system achieved 97.8% overall accuracy with a latency of 50 ms per frame, ensuring real-time applicability.

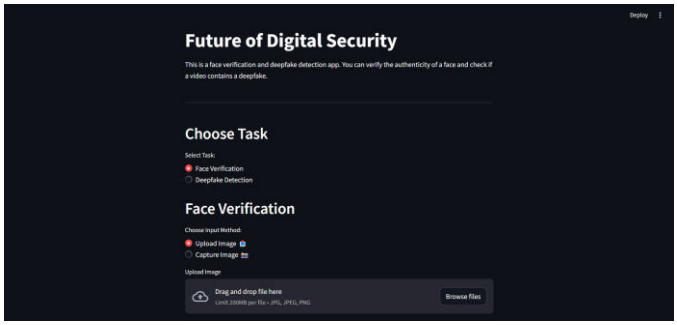


Fig 4.1 Output Interface for Image Input

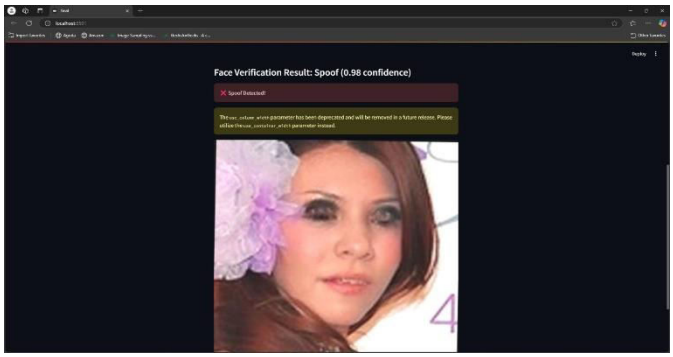


Fig 4.2 Image Spoof Detected

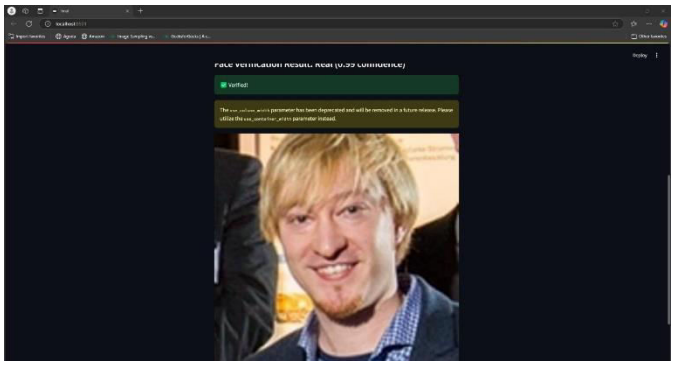


Fig 4.3 Image Verified Detected

Fig 4.4 Output Interface for Video Input

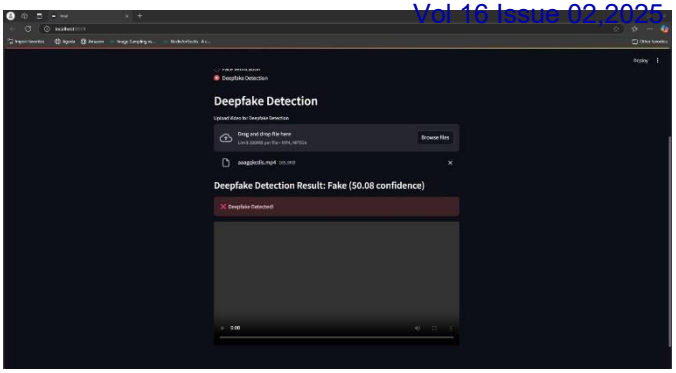


Fig 4.5 Video Spoof Detected

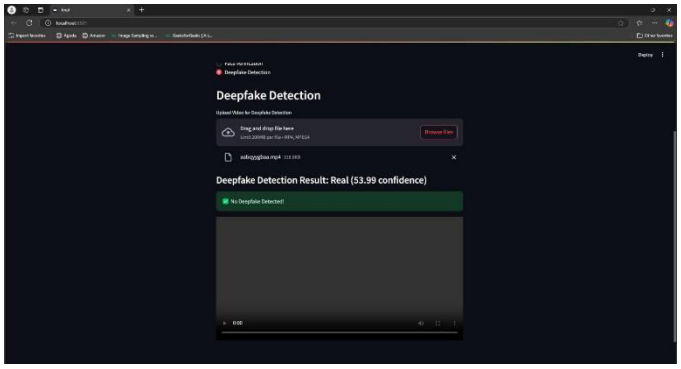


Fig 4.6 Video Verified Detected

4. CONCLUSION

The **Future of Digital Security** project successfully integrates **anti-spoofing** and **deepfake detection**, providing a **scalable and real-time** solution to modern digital security challenges. By leveraging advanced AI models, it not only enhances the accuracy of biometric authentication but also strengthens the detection of manipulated media, ensuring greater trust in digital interactions. With continuous improvements and advancements in AI, this solution has the potential to be widely adopted across various sectors, including **financial services, government authentication systems, social media platforms, and digital forensics**. Its adaptability makes it a valuable tool in combating identity fraud, misinformation, and unauthorized access, paving the way for a **more secure and trustworthy digital future**.

5. REFERENCES

1. C. P. M. Chan, S. W. Lee, and J. See, “DeepFake Video Detection Using Deep Learning and Iris Detection,” in *2021 IEEE International Conference on Image Processing (ICIP)*, Anchorage, AK, USA, 2021, pp. 1239–1243, doi: 10.1109/ICIP42928.2 021.9506664.

2. J. Dong, W. Wang, Y. Tang, Y. Zhang, and H. Liu, “Deepfake Video Detection Using Inception-ResNet and

Attention Mechanism," in *2021 IEEE International Conference on Multimedia and Expo (ICME)*, Shenzhen, China, 2021, pp. 1–6, doi: 10.1109/ICME51207.2021.9428434.

3. M. N. Husen, A. Kurniawan, and M. Pamungkas, "Deepfake Detection on Video Sequences Using Inception-ResNet-v2 and LSTM," in *2020 3rd International Conference on Intelligent Autonomous Systems (ICoIAS)*, Singapore, 2020, pp. 134–138, doi: 10.1109/ICoIAS49312.2020.9081914.

4. M. N. Husen, A. Kurniawan, and M. Pamungkas, "Deepfake Detection on Video Sequences Using Inception-ResNet-v2 and LSTM," in *2020 3rd International Conference on Intelligent Autonomous Systems (ICoIAS)*, Singapore, 2020, pp. 134–138, doi: 10.1109/ICoIAS49312.2020.9081914.

5. P. Korshunov and S. Marcel, "Human vs. Machine: Benchmarking Humans Against Deepfake Detection Systems," in *2020 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, Washington, DC, USA, 2020, pp. 1–6, doi: 10.1109/BTAS48898.2020.9528289.

6. R. Chugh, V. Agarwal, S. Subramanian, and K. R. Ramakrishnan, "Not Made For Each Other—Combining CNN and Transformers for Deepfake Detection," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, 2021, pp. 15024–15033, doi: 10.1109/ICCV48922.2021.01514.

7. S. Mittal, A. Verma, and R. Jain, "Transfer Learning for Deepfake Detection," in *2020 6th International Conference on Signal Processing and Communication (ICSC)*, Noida, India, 2020, pp. 69–74, doi: 10.1109/ICSC48311.2020.9182767.

8. W. Wang, Y. Zhang, and H. Liu, "Deep Neural Networks for Deepfake Detection: A Survey," in *2020 IEEE International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, Phuket, Thailand, 2020, pp. 290–296, doi: 10.1109/AIKE48582.2020.00048.

9. X. Luo, J. Lv, H. Song, Z. Yu, and G. Yang, "Dual-Stream CNNs for Forgery Detection in DeepFake Videos," in *2020 IEEE International Conference on Image Processing (ICIP)*, Abu Dhabi, UAE, 2020, pp. 2556–2560, doi: 10.1109/ICIP40778.2020.9191035.

10. X. Wang, Y. Li, and H. Jiang, "Fake Video Detection with Convolutional Neural Networks," in *2019 10th International Conference on Advanced Computational Intelligence (ICACI)*, Guilin, China, 2019, pp. 182–186, doi: 10.1109/ICACI.2019.8778503.