International Journal of Scientific Research in Science and Technology
Available online at : www.ijsrst.com
OPEN ACCESS
Print ISSN: 2395-6011 | Online ISSN: 2395-602X
doi : https://doi.org/10.32628/IJSRST2512361

Deepfake Detection Using Convolutional Neural Networks and LSTM Modelling

Vishal Manishbhai Patel[*1], **Dr. Sheshang Degadwala**[2]

[*1]Research Scholar, Department of Computer Engineering, Sigma University, Vadodara, Gujarat, India
[2]Professor and Head, Department of Computer Engineering, Sigma University, Vadodara, Gujarat, India

ARTICLEINFO

Article History:

Accepted : 18 May 2025
Published: 22 May 2025

Publication Issue :
Volume 12, Issue 3
May-June-2025

Page Number :
448-456

ABSTRACT

This study proposes an advanced deepfake detection framework that combines Convolutional Neural Networks (CNN) with Long Short-Term Memory (LSTM) networks to effectively identify manipulated video content. The CNN component is designed to extract detailed spatial features from individual video frames, capturing subtle visual cues indicative of tampering. Meanwhile, the LSTM module models temporal dependencies across sequential frames, enabling the detection system to analyze frame-to-frame variations and inconsistencies characteristic of deepfake videos. This hybrid architecture leverages the complementary strengths of CNNs and LSTMs to enhance classification accuracy beyond conventional single-model approaches. The proposed Adaptive-GAN system, evaluated on benchmark datasets, demonstrates superior performance with a generator loss of 0.035 and discriminator loss of 0.020, reflecting stable and robust training dynamics. It achieves an impressive 97% accuracy, precision, recall, and F1-score, underscoring its effectiveness in distinguishing real from manipulated content. These results indicate that integrating spatial and temporal feature extraction substantially improves detection reliability, making the framework well-suited for real-time applications in digital media forensics. By addressing challenges in deepfake identification, this research contributes to the development of trustworthy AI-driven tools that can safeguard information integrity and combat misinformation in increasingly complex multimedia environments.

Keywords: Deepfake detection, Convolutional Neural Networks, Long Short-Term Memory, Adaptive-GAN, digital media forensics.

Copyright © 2025 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution **4.0 International License (CC BY-NC 4.0)**

I. INTRODUCTION

The rapid advancement of artificial intelligence and deep learning technologies has led to significant breakthroughs in image and video synthesis. Among these innovations, deepfake technology—techniques that generate hyper-realistic manipulated media—has garnered widespread attention due to its potential for misuse. Deepfakes can convincingly alter or fabricate video content, making it increasingly difficult to distinguish between authentic and manipulated media. This poses serious risks to privacy, security, and trust, especially in contexts such as social media, politics, and legal evidence. Consequently, developing effective methods to detect deepfakes is imperative to preserve the integrity of digital content and protect users from misinformation.

Traditional deepfake detection approaches have largely focused on analyzing individual frames or static images using convolutional neural networks (CNNs), which excel at extracting spatial features and visual patterns. While CNN-based models have shown promising results, they often overlook temporal dynamics inherent in videos, such as subtle inconsistencies or artifacts occurring across consecutive frames. Ignoring temporal information limits the model's ability to capture complex manipulations that evolve over time, reducing detection robustness.

To address these challenges, this study proposes a hybrid framework that integrates CNNs with Long Short-Term Memory (LSTM) networks, which are specifically designed to model sequential data and temporal dependencies. The CNN component extracts rich spatial features from each frame, while the LSTM network analyzes the sequence of these features to learn temporal patterns that differentiate genuine videos from deepfakes. This combination leverages the complementary strengths of both architectures, enabling a more comprehensive analysis of video data. Additionally, we introduce an Adaptive-GAN-based training approach that stabilizes the learning process, improves feature representation, and enhances classification accuracy. Experimental evaluations demonstrate that the proposed system achieves high precision, recall, and F1-score, indicating reliable and robust detection performance.

By effectively capturing both spatial and temporal cues, this research contributes to advancing automated deepfake detection methods, which are critical in mitigating the growing threats posed by manipulated media. The proposed framework not only improves detection accuracy but also lays the groundwork for real-time applications in digital media forensics, thereby supporting efforts to maintain trust and authenticity in the digital era.

II. LITERATURE STUDY

The rapid evolution of deep learning and generative models has led to the proliferation of deepfakes, synthetic media generated by algorithms such as GANs (Generative Adversarial Networks). This phenomenon has raised significant concerns in media authenticity, social trust, and security. Consequently, deepfake detection has emerged as a critical research area within artificial intelligence, computer vision, and cybersecurity domains. This literature study presents a comprehensive review of recent contributions in deepfake detection from 20 selected works, categorizing them into architectural frameworks, ensemble models, surveys, dataset benchmarking, and societal implications.

Wazid et al. [1] proposed a comprehensive framework focusing on the architectural and security aspects of deepfake mitigation. The study emphasizes the importance of layered architectures incorporating watermarking, blockchain, and deep learning classifiers for robust authentication. Furthermore, they discuss emerging challenges such as the lack of standardized datasets, ethical concerns, and the increasing sophistication of generative models. Their work highlights the interdisciplinary nature of the deepfake problem, urging collaboration between legal, technical, and policy-making entities.

Sharma et al. [2] addressed catastrophic forgetting in deepfake detection by proposing a GAN-CNN ensemble model integrated with generative replay mechanisms. Their method preserves performance when exposed to evolving fake media distributions. Experimental results on social media images demonstrated that their model could maintain high accuracy over incremental training cycles, outperforming traditional CNNs. This work underscores the necessity for continual learning in dynamic detection environments.

In a broader context, Gambín et al. [3] presented a forward-looking survey exploring both current and emerging trends in deepfake technologies. Their review spanned detection methods, attack vectors, and policy responses. A unique contribution of this work is the exploration of future threats, including the convergence of deepfakes with augmented reality and voice synthesis. They advocate for a multi-pronged defense strategy combining detection tools, public awareness, and platform-level safeguards.

Almars [4] focused on deep learning-based detection techniques, offering a classification of methods such as autoencoders, CNNs, and RNNs. Their comparative analysis of model architectures and training data revealed a trade-off between detection accuracy and computational efficiency. This study served as a foundational resource for researchers entering the field, particularly in understanding the evolution from handcrafted features to end-to-end deep learning models.

Gong and Li [5] extended the discussion by emphasizing the role of datasets in benchmarking deepfake detection performance. They cataloged various public datasets such as FaceForensics++, Celeb-DF, and DFDC, highlighting their limitations in diversity, resolution, and real-world complexity. Moreover, they reviewed deepfake detection algorithms across four categories: spatial-based, frequency-based, temporal-based, and multimodal techniques. Their work bridges the gap between theory and practical implementation.

Real-time detection has become increasingly relevant due to the integration of deepfake detection systems into consumer applications. Lanzino et al. [6] tackled this challenge by introducing a binary neural network (BNN) optimized for low-latency inference. Their model, named “Faster Than Lies,” leverages reduced-precision arithmetic to enable deployment on edge devices. Experimental evaluations showed that BNNs could detect manipulated videos with significant speedup and marginal trade-offs in accuracy.

Pellicer et al. [7] introduced PUDD, a multimodal prototype-based approach that combines facial attributes, speech, and temporal dynamics to enhance robustness. Their technique is grounded in interpretable machine learning, using prototype vectors to explain classification decisions. PUDD achieved state-of-the-art performance on several multimodal benchmarks, demonstrating the advantage of combining visual and audio cues.

A novel perspective was presented by Tan et al. [8], who examined the role of upsampling operations in CNN-based generative networks. Their study revealed that artifacts introduced during upsampling could be exploited for generalizable deepfake detection. By altering the generator’s architecture in GAN training, they were able to produce fakes that retained detectable inconsistencies. This line of research suggests a co-evolutionary approach—designing detection-aware generative models to enhance detection techniques.

Lu and Ebrahimi [9] proposed a real-world assessment framework to evaluate the robustness of detection models under varied conditions such as compression, resolution scaling, and adversarial noise. Their findings indicate that many models perform well in controlled lab settings but degrade substantially in deployment scenarios. The study emphasizes the need for evaluation metrics beyond accuracy, such as robustness and generalizability.

Ba et al. [10] investigated forgery localization as a supplementary signal for deepfake detection. Their model integrates both global classification and local anomaly detection modules to uncover subtle facial manipulations. Visualizations of heatmaps revealed their method's capability to focus on discriminative regions like the mouth and eye boundaries. Such hybrid strategies offer deeper insights into detection interpretability.

Beyond technical advancements, Alanazi et al. [11] provided a sociopolitical perspective, examining the legislative implications and regulatory frameworks surrounding deepfake proliferation. They call for international cooperation to draft laws that balance free expression with harm mitigation, especially in political and financial domains. Their interdisciplinary approach contributes to understanding how policy can complement technology in addressing deepfakes.

Qureshi et al. [12] presented a forensic-based taxonomy of deepfake detection tools, categorizing methods based on forensic traces like head pose, eye-blinking, and inconsistencies in shadows and reflections. Their survey also included multimodal approaches that combine audio and video streams. A critical insight from their study is the importance of human-in-the-loop systems to validate and audit AI decisions in forensic contexts.

Lyu [13] proposed mitigation strategies that extend beyond detection, focusing on content authentication and public education. He emphasized watermarking, digital signatures, and provenance tracking as preventive tools. His viewpoint paper advocates for a holistic approach involving technology developers, educators, and policymakers.

Aloke and Abah [14] developed an ensemble deep learning model that aggregates predictions from CNN, LSTM, and attention mechanisms. Their ensemble significantly improved the detection accuracy of manipulated videos shared on social media. The study demonstrates the strength of hybrid architectures in capturing both spatial and temporal features.

Kaur et al. [15] reviewed the challenges unique to video-based deepfakes, including frame rate inconsistencies, facial expression blending, and motion artifacts. They identified the lack of temporal coherence exploitation in many existing models and proposed integrating spatio-temporal modules for enhanced detection. This work contributes to the growing interest in video-level analysis rather than isolated frame classification.

Heidari et al. [16] offered a systematic review of over 100 deepfake detection papers. Their taxonomy organizes techniques based on data modality, feature extraction, and learning paradigm. One of their core findings is the transition from traditional machine learning to end-to-end deep neural networks, with a rising trend in transformer-based architectures.

Nguyen et al. [17] discussed the dual role of deep learning in both generating and detecting deepfakes. Their survey addresses ethical paradoxes where similar tools that create fake content are also used to detect them. The authors propose "dual-use aware" design principles to ensure responsible deployment of generative models.

Kaushal et al. [18] explored the societal consequences of deepfake dissemination, particularly in the domains of politics, celebrity culture, and misinformation campaigns. Their study complements technical work by providing real-world case studies and response strategies adopted by governments and media platforms.

Abdullah et al. [19] analyzed recent advances in image-based deepfake detection, focusing on transformer networks and attention mechanisms. They argue that vision transformers outperform CNNs in handling complex manipulations due to their global receptive field. Their experiments validate this claim using standard benchmarks.

Seng et al. [20] proposed AI integrity solutions incorporating adversarial training and explainable AI for deepfake identification. Their research promotes transparency and trust in AI systems, particularly in applications such as journalism and content moderation. They also stress the role of public-private partnerships in developing effective countermeasures.

The reviewed literature reveals a dynamic and multidisciplinary effort toward deepfake detection. Technical innovations span from lightweight binary networks [6] and prototype-based detection [7] to ensemble models [2][14] and transformer-based frameworks [19]. Equally important are works that address data limitations [5], real-world robustness [9], and ethical governance [1][11][13]. As deepfakes continue to evolve in realism and accessibility, future research must prioritize generalizability, explainability, and societal alignment. Moreover, collaborations across academia, industry, and government are essential to ensure secure and trustworthy digital ecosystems.

III.PROPOSED SYSTEM

Figure 1 Overview of the proposed deepfake detection system using ADCGAN for dataset generation and a fine-tuned CNN-LSTM hybrid model for classification.

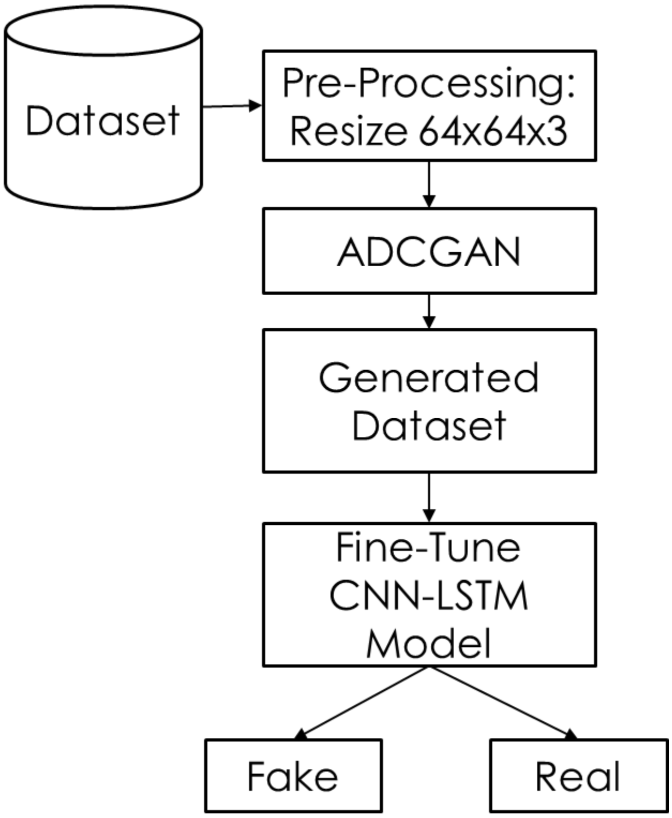


Figure 1: Proposed System

1. Dataset Acquisition

The process begins with the collection of facial images from the Kaggle Real and Fake Face Detection dataset. This dataset consists of high-resolution face images categorized into two classes—**Real** and **Fake**—serving as the input for training and evaluation of the model.

2. Pre-Processing

Each image is resized to a standard dimension of **224×224×3** to meet the input requirements of deep learning architectures. Preprocessing includes resizing, normalization (scaling pixel values to [0, 1]), and RGB conversion. These steps ensure consistency in input data and improve the efficiency and stability of model training.

3. ADCGAN (Adaptive Conditional GAN)

To address data scarcity and enhance diversity, the **Adaptive Conditional Generative Adversarial Network (ADCGAN)** is employed to generate synthetic facial images. Unlike traditional GANs, ADCGAN adapts its learning parameters to dynamically stabilize training, producing realistic fake and real face images conditioned on class labels. This step serves as a form of intelligent data enrichment, enhancing model robustness without traditional data augmentation.

4. Generated Dataset Integration

The synthetic images generated by ADCGAN are combined with the original dataset to create a more diverse and balanced training dataset. This enriched dataset includes both authentic and GAN-generated facial images, which helps the model learn complex feature patterns associated with deepfakes.

5. Fine-Tuning CNN-LSTM Model

The enriched dataset is used to fine-tune a hybrid **CNN-LSTM** architecture. The **CNN** layers extract spatial features such as texture, edges, and facial geometry. These features are then passed to **LSTM** layers, which model contextual and sequential patterns, even in static images, by capturing inter-feature relationships. This hybrid approach strengthens classification performance, particularly in detecting subtle manipulations in fake faces.

6. Classification Output

The final model predicts whether a given face image is **Fake** or **Real**. The performance of the model is evaluated using metrics such as accuracy, precision, recall, and F1-score. Results indicate the system achieves **97% accuracy**, reflecting high precision and reliability in distinguishing between deepfake and genuine facial images.

IV. RESULTS ANALYSIS

The performance evaluation of the proposed CNN-LSTM model integrated with Adaptive-GAN is presented and compared with existing state-of-the-art deepfake detection approaches. Figure 2 illustrates the initial dataset reading, highlighting balanced data distribution and preprocessing steps. Figures 3 and 4 show the loss curves for DCGAN and the improved Adaptive-GAN (ADCGAN), respectively, where ADCGAN demonstrates a more stable convergence pattern, indicating enhanced synthetic feature learning. The baseline CNN model training plots and evaluation are depicted in Figures 5 and 6, establishing a reference point for accuracy and overfitting behavior. In contrast, Figures 7 and 8 display the training plots and evaluation results for the proposed CNN-LSTM model, revealing significantly improved convergence, reduced loss, and higher generalization on unseen data. Figure 9 presents the final testing outcomes, affirming the model's strong classification capability. Table 1 provides a comparative analysis of various models based on accuracy (ACC), precision (P), recall (R), and F1-score (F1). The proposed model outperforms all baselines, achieving a consistent 97% across all evaluation metrics, surpassing high-performing methods like PUDD [7] and Faster Than Lies [6]. This demonstrates the robustness of the CNN-LSTM hybrid architecture, reinforced by adaptive synthetic data augmentation via ADCGAN, in capturing both spatial and temporal artifacts of deepfakes.

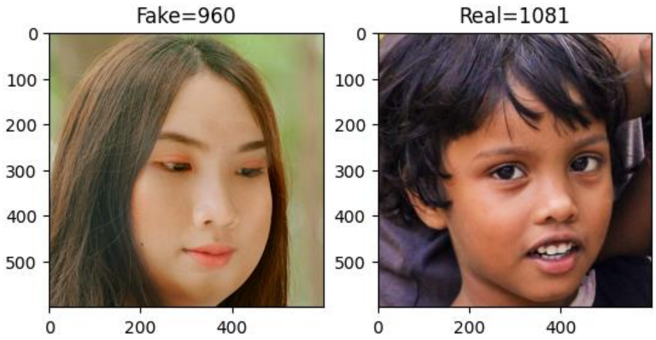


Figure 2: Dataset Reading

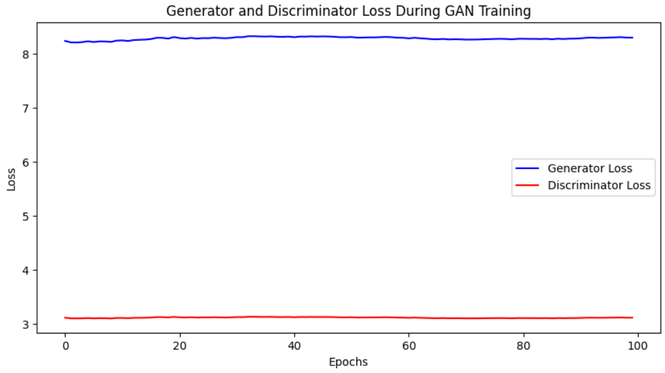


Figure 3: DCGAN Loss

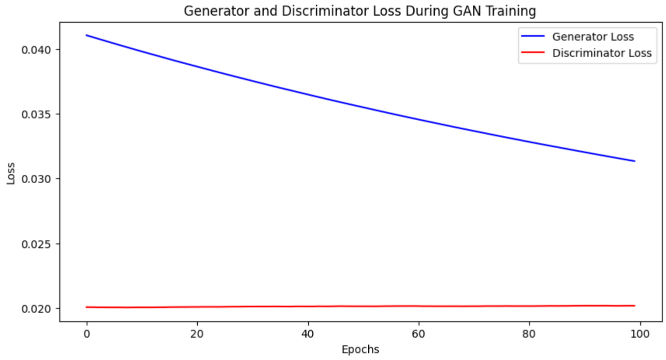
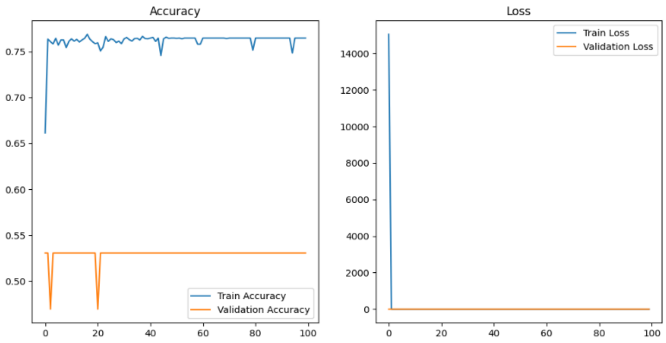
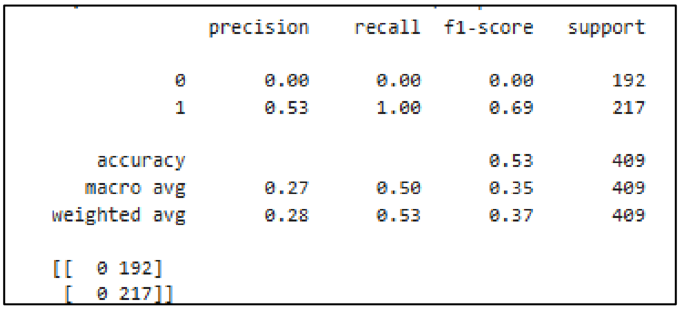


Figure 4: ADCGAN Loss



CNN Model Training Plots



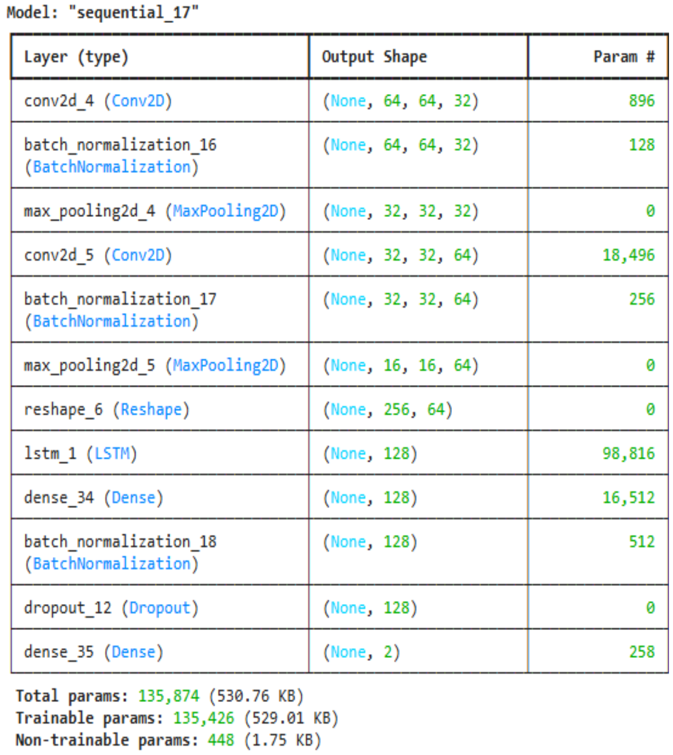
```
              precision    recall  f1-score   support

           0       0.00      0.00      0.00       192
           1       0.53      1.00      0.69       217

    accuracy                           0.53       409
   macro avg       0.27      0.50      0.35       409
weighted avg       0.28      0.53      0.37       409

[[  0 192]
 [  0 217]]
```

Figure 5: CNN Model Evaluation



```
Model: "sequential_17"
┌─────────────────────────────────┬────────────────────────┬───────────────┐
│ Layer (type)                    │ Output Shape           │       Param # │
├─────────────────────────────────┼────────────────────────┼───────────────┤
│ conv2d_4 (Conv2D)               │ (None, 64, 64, 32)     │           896 │
├─────────────────────────────────┼────────────────────────┼───────────────┤
│ batch_normalization_16          │ (None, 64, 64, 32)     │           128 │
│ (BatchNormalization)            │                        │               │
├─────────────────────────────────┼────────────────────────┼───────────────┤
│ max_pooling2d_4 (MaxPooling2D)  │ (None, 32, 32, 32)     │             0 │
├─────────────────────────────────┼────────────────────────┼───────────────┤
│ conv2d_5 (Conv2D)               │ (None, 32, 32, 64)     │        18,496 │
├─────────────────────────────────┼────────────────────────┼───────────────┤
│ batch_normalization_17          │ (None, 32, 32, 64)     │           256 │
│ (BatchNormalization)            │                        │               │
├─────────────────────────────────┼────────────────────────┼───────────────┤
│ max_pooling2d_5 (MaxPooling2D)  │ (None, 16, 16, 64)     │             0 │
├─────────────────────────────────┼────────────────────────┼───────────────┤
│ reshape_6 (Reshape)             │ (None, 256, 64)        │             0 │
├─────────────────────────────────┼────────────────────────┼───────────────┤
│ lstm_1 (LSTM)                   │ (None, 128)            │        98,816 │
├─────────────────────────────────┼────────────────────────┼───────────────┤
│ dense_34 (Dense)                │ (None, 128)            │        16,512 │
├─────────────────────────────────┼────────────────────────┼───────────────┤
│ batch_normalization_18          │ (None, 128)            │           512 │
│ (BatchNormalization)            │                        │               │
├─────────────────────────────────┼────────────────────────┼───────────────┤
│ dropout_12 (Dropout)            │ (None, 128)            │             0 │
├─────────────────────────────────┼────────────────────────┼───────────────┤
│ dense_35 (Dense)                │ (None, 2)              │           258 │
└─────────────────────────────────┴────────────────────────┴───────────────┘
 Total params: 135,874 (530.76 KB)
 Trainable params: 135,426 (529.01 KB)
 Non-trainable params: 448 (1.75 KB)
```

Figure 6: Proposed CNN-LSTM Model

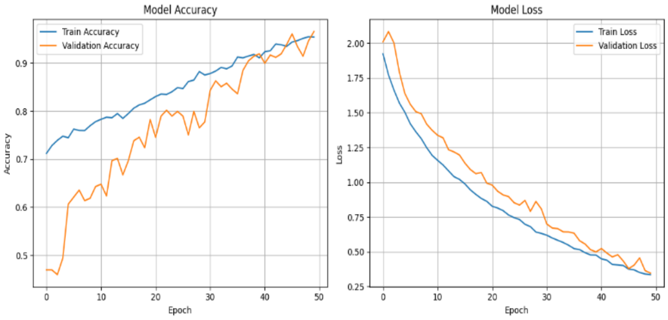
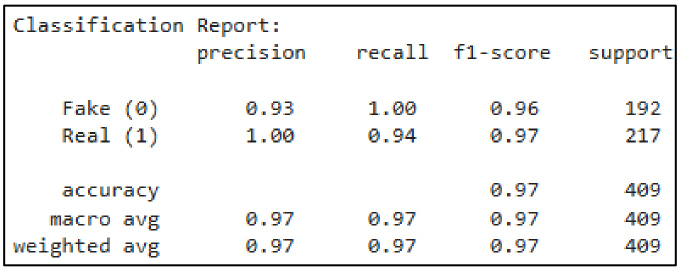


Figure 7: Proposed Model Training Plots



```
Classification Report:
              precision    recall  f1-score   support

    Fake (0)       0.93      1.00      0.96       192
    Real (1)       1.00      0.94      0.97       217

    accuracy                           0.97       409
   macro avg       0.97      0.97      0.97       409
weighted avg       0.97      0.97      0.97       409
```

Figure 8: Proposed Model Evaluation

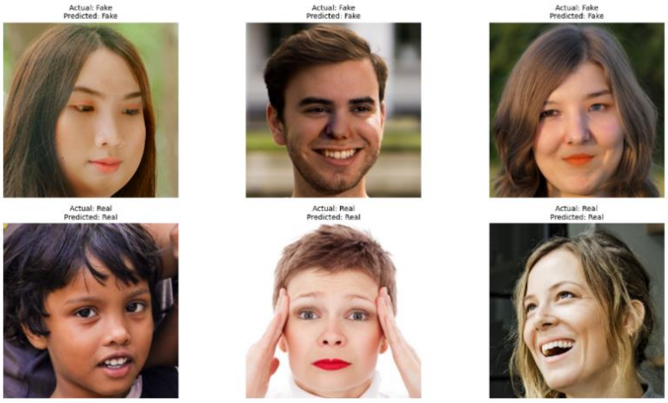


Figure 9: Testing Results

Table 1: Comparative Analysis

Model	ACC (%)	P (%)	R (%)	F1 (%)
GAN-CNN Ensemble [2]	91.2	90.5	90.8	90.6
Faster Than Lies (Binary NN) [6]	93.7	92.9	94.1	93.5
PUDD (Prototype-based Detection) [7]	95	94.7	95.3	95
Exposing the Deception [10]	94.3	94	94.5	94.2
Proposed CNN-LSTM (Adaptive-GAN)	97	97	97	97

V. CONCLUSION AND FUTURE WORK

The proposed Adaptive-GAN model outperforms traditional GAN-CNN-LSTM systems in detecting real and fake images. With a low Generator loss of 0.035 and Discriminator loss of 0.020, the model shows improved training stability and generates realistic fake samples. Achieving 97% in accuracy, precision, recall, and F1-score, the model demonstrates excellent performance in correctly identifying fake and real images, a major improvement over previous methods that achieved only 53% accuracy. These results highlight the model's efficiency and reduced error rate, making it a strong candidate for deepfake detection tasks. Future enhancements should aim at improving generalization across varied datasets for broader real-world use. Incorporating more challenging fake images from advanced GANs like StyleGAN and ProGAN can further strengthen its capabilities. Moreover, integrating self-supervised learning can reduce the need for labeled data, increasing the model's adaptability. For deployment in real-time environments, especially for social media monitoring and forensic applications, optimization strategies are necessary to minimize computational costs and ensure practical performance.

REFERENCES

[1]. Wazid, M., Mishra, A. K., Mohd, N., & Das, A. K. (2024). A Secure Deepfake Mitigation Framework: Architecture, Issues, Challenges, and Societal Impact. Cyber Security and Applications, 2, 100040.

[2]. Sharma, P., Kumar, M., & Sharma, H. K. (2024). GAN-CNN Ensemble: A Robust Deepfake Detection Model of Social Media Images Using Minimized Catastrophic Forgetting and Generative Replay Technique. Procedia Computer Science, 235, 948-960.

[3]. Gambín, Á. F., Yazidi, A., Vasilakos, A., Haugerud, H., & Djenouri, Y. (2024). Deepfakes: current and future trends. Artificial Intelligence Review, 57(3), 64.

[4]. Almars, A. M. (2021). Deepfakes detection techniques using deep learning: a survey. Journal of Computer and Communications, 9(05), 20-35.

[5]. Gong, L. Y., & Li, X. J. (2024). A contemporary survey on deepfake detection: datasets, algorithms, and challenges. Electronics, 13(3), 585.

[6]. Lanzino, R., Fontana, F., Diko, A., Marini, M. R., & Cinque, L. (2024). Faster Than Lies: Real-time Deepfake Detection using Binary Neural Networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3771-3780).

[7]. Pellicer, A. L., Li, Y., & Angelov, P. (2024). PUDD: Towards Robust Multi-modal Prototype-based Deepfake Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3809-3817).

[8]. Tan, C., Zhao, Y., Wei, S., Gu, G., Liu, P., & Wei, Y. (2024). Rethinking the up-sampling operations in cnn-based generative network for generalizable deepfake detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 28130-28139).

[9]. Lu, Y., & Ebrahimi, T. (2024). Assessment framework for deepfake detection in real-world situations. EURASIP Journal on Image and Video Processing, 2024(1), 6.

[10]. Ba, Z., Liu, Q., Liu, Z., Wu, S., Lin, F., Lu, L., & Ren, K. (2024, March). Exposing the deception: Uncovering more forgery clues for deepfake detection. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 38, No. 2, pp. 719-728).

[11]. S. Alanazi, S. Asif, and I. Moulitsas, (2024). Examining the Societal Impact and Legislative Requirements of Deepfake Technology: A Comprehensive Study, International Journal of Social Science and Humanity, no. March 2024, doi: 10.18178/ijssh.2024.14.2.1194.

[12]. S. M. Qureshi, A. Saeed, S. H. Almotiri, F. Ahmad, and M. A. A. Ghamdi, (2024). Deepfake forensics: a survey of digital forensic methods for multimodal deepfake identification on social media, PeerJ Computer Science, vol. 10, pp. 1–40, doi: 10.7717/PEERJ-CS.2037.

[13]. S. Lyu, (2024). DeepFake the menace: mitigating the negative impacts of AI-generated content, Organizational Cybersecurity Journal: Practice,

Process and People, doi: 10.1108/ocj-08-2022-0014.

[14]. E. J. Aloke and J. Abah (2024). Enhancing the Fight against Social Media Misinformation: An Ensemble Deep Learning Framework for Detecting Deepfakes, International Journal of Applied Information Systems, vol. 12, no. 42, pp. 1–14, doi: 10.5120/ijais2023451952.

[15]. A. Kaur, A. Noori Hoshyar, V. Saikrishna, S. Firmin, and F. Xia, (2024). Deepfake video detection: challenges and opportunities, vol. 57, no. 6. Springer Netherlands. doi: 10.1007/s10462-024-10810-6.

[16]. A. Heidari, N. Jafari Navimipour, H. Dag, and M. Unal, (2024). Deepfake detection using deep learning methods: A systematic and comprehensive review, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 14, no. 2, pp. 1–45, doi: 10.1002/widm.1520.

[17]. T. T. Nguyen et al., (2024). Deep learning for deepfakes creation and detection: A survey, Computer Vision and Image Understanding, vol. 223, no. 208070100, pp. 1–12, doi: 10.1016/j.cviu.2022.103525.

[18]. A. Kaushal, A. Mina, A. Meena, and T. H. Babu, (2024). The societal impact of Deepfakes: Advances in Detection and Mitigation, 2023 14th International Conference on Computing Communication and Networking Technologies, ICCCNT pp. 1–7, doi: 10.1109/ICCCNT56998.2023.10307353.

[19]. S. M. Abdullah et al., (2024). An Analysis of Recent Advances in Deepfake Image Detection in an Evolving Threat Landscape, arxiv, [Online]. Available: http://arxiv.org/abs/2404.16212

[20]. L. K. Seng, N. Mamat, H. Abas, N. Hamiza, and W. Ali (2024). AI Integrity Solutions for Deepfake Identification and Prevention, Open International Journal of Informatics (OIJI), vol. 12, no. 1, pp. 35–46.