Deepfake Video Detection: Comprehensive Report

1 Introduction

Deepfake technology has advanced rapidly, enabling the creation of highly realistic manipulated videos. While these tools showcase the potential of AI, their misuse has led to challenges like misinformation, identity theft, and cyber fraud. This project aims to address these issues by detecting deepfake videos using advanced machine learning techniques. By leveraging the **Tall & Tall++** architecture and training on a curated dataset, the model achieves an accuracy of over 75%.

2 Objectives

- Develop a reliable deepfake video detection system.
- Train and evaluate the model on a sample dataset of 5,000 images.
- Ensure the system achieves scalability and real-time efficiency.

3 Dataset

The project utilized the DFDC Facial Cropped Videos Dataset from Kaggle: Dataset Link.

- **Subset Used:** 5,000 images.
- **Preprocessing:** Frames resized, normalized, and augmented.
- **Augmentation Techniques:** Rotation, flipping, brightness adjustment.

4 Challenges in Deepfake Detection

- Dynamic improvement of deepfake tools makes detection complex.
- Scarcity of diverse and high-quality datasets.
- Real-time detection demands high computational power.

5 Model Architecture

The **Tall & Tall++** models are designed for spatial and temporal analysis of video frames:

1. **Feature Extraction:** Convolutional layers for spatial feature analysis.
2. **Temporal Analysis:** Sequential inconsistencies detection.
3. **Classification:** Binary output (Real or Fake) using a softmax function.

6 Implementation

6.1 Steps

1. Preprocessing: Frames resized, normalized, and split (70:30).
2. Model Training:
   - Optimizer: Adam, learning rate = 0.0001.
   - Batch Size: 32, Epochs: 15.
3. Evaluation: Accuracy, Precision, Recall, F1-Score.

6.2 Performance Analysis

- **Accuracy:** >75%.
- **Precision:** High true positive rate.
- **Recall:** Low false negatives.
- **F1-Score:** Indicates balanced metrics.

7 Applications

- Digital Media Forensics.
- Social Media Content Authentication.
- News and Media Verification.
- Educational Tools for AI awareness.

8 Future Enhancements

- Use larger datasets for improved generalization.
- Implement real-time detection capabilities.
- Integrate Vision Transformers for enhanced performance.
- Conduct robustness testing against adversarial inputs.

9 Conclusion

This project addresses the growing concern of deepfake videos by employing advanced AI models. With an accuracy of over 75%, it provides a practical solution for real-world applications and sets the stage for future advancements in the field.