ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
pp. 1-9

Article Title: **Video Based Evidence Analysis and Extraction in Digital Forensic Investigation**

Video Based Evidence Analysis and Extraction in Digital Forensic Investigation

S. Venkata Kiran[1], B. Himabindhu[2], G. Revathi[3], K. Sudarsan Reddy[4], K. Mohan Krishna[5], K. Rosi Reddy[6]

[1]Associate Professor, Department of Electronics and Communication Engineering, Sri Venkatesa Perumal College of Engineering and Technology, Puttur, AP, India.
[2,3,4,5,6]UG Students, Department of Electronics and Communication Engineering, Sri Venkatesa Perumal College of Engineering and Technology, Puttur, AP, India.

ABSTRACT

With the rapid advancement of technology and the ease of creating fake content, the manipulation of media has become widespread in recent times. The emergence of AI-altered videos, known as Deepfakes, poses a significant threat to media integrity as they are increasingly being produced and disseminated across various social media platforms. Detecting such Deepfakes has become a major challenge in digital forensics. In this paper, we propose an approach for detecting Deepfake videos using Recurrent Neural Network (RNN) algorithms. The proposed method and its steps are elaborated upon in detail. We achieved an accuracy of 91% for the developed Deep Learning (DL) model using the Celeb-DF dataset. Finally, we present the results of fake or real detection in the provided videos, demonstrating the effectiveness of our approach.

Keywords: Forensics investigation, Forensic video analysis, Video/image enhancement, Object detection, Deep learning.

1 Introduction

The evolution of the internet and its scope stretching to all corners of the world, has been backed by an increasingly large number of spread of false information in recent times. Media posted on the web are seen to be manipulated to mislead viewers. With the advent of sophisticated Artificial Intelligence (AI) trained models used in the modifications of digital content, it has become plausibly impossible to distinguish the original media from the fake with the naked eye. These fake media contents have become popularized by the term “Deepfakes”. Based on Artificial intelligence, human image synthesis can be done using Deepfake techniques. Many Deepfake videos have been distributed across social media as technology has become more accessible to all users. These Deepfake content are generated by combining and superimposing existing media (images or videos) onto source media using a Generative Adversarial Network (GANs) [14,15], which is a deep learning technique.

ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
pp. 1-9

Article Title: Video Based Evidence Analysis and Extraction in Digital Forensic Investigation

Face swapping, Lip-syncing and Head Puppetry are the 3 major ways in which videos are manipulated to create Deepfakes. Blinking of an eye can be considered as one of the features for Deepfake detection[16].

In reality, one of the most serious concerns in modern civilization is the emergence of Deepfake .RanaAyyub, a journalist whose fake pornographic video was created and spread on the platforms like WhatsApp, Twitter and other social media, is one of the instances of misuse Deepfake[17]. As digital media serves as an evidence in many cases [18], increase in the manipulation of these contents lead people to question the authenticity of any content posted. They are reported to have been used to tarnish the reputation of celebrities [19], spreading misleading information and rumours for politicians, unfurling false news [20], and forge evidence in CCTV footage, thus posing a great threat to the integrity of digital media as well as security [21,22]. Thus, there is a high need of developing a system which would detect Deep Fake videos, thereby preserving the integrity of digital media [23], [24], [25]. This paper aims at developing an efficient model to detect if any given video is fake or real.

2 Literature Survey

Using David Guera, et al. [1] Have proposed a method to generate videos using CNN and LSTM. To find out deep fake videos having a clear knowledge on how deep fake videos are created/generated will basically aid in comprehending the flaws in Deepfake generation and working on the flaws for Deepfake video detection. Here HOHA dataset is used which consists of 300 videos. In this paper CNN is utilized for feature extraction and LSTM is utilized for temporal sequence analysis (basically processing the frame sequence).

Hasam Khalid, et al. [2] have demonstrated Classifying Deepfake using One ClassVariational Autoencoder. For training, the model presented in the paper simply requires real photos or videos. Here the dataset used is FaceForensic++. The collected dataset is split into frames, face is detected and alignment of the face is carried 0out using MTCNN (Multitask Cascaded CNN: face-detection tool). Here, a one class vibrational encoder is utilized.

MousaTayseerJafa, et al. [3] have recommended a method to detect deep fake using CNN techniques. In this paper the analysis of deep fake videos is carried out considering mouth features using deep-learning techniques [4]. The proposed deep-learning detection model with mouth features (DFT-MF) is constructed by isolating, analysing, and verifying lip/mouth movement with a deep learning approach to detect deep fake videos. Here, the dataset contains the combination of real and fake videos (Deepfake dataset generated for the development and testing of Deepfake detection).

Faten F Kharbat, et al. [5] have demonstrated Image Feature Detectors for Deepfake Video. SVM (Support Vector Machine) regression is used to detect Deepfake videos. There are 98 videos in the collection, half of which are fake and half of which are real. The videos in the dataset are in mp4 format and consume a runtime of 30 seconds. Traditional edge feature

detectors are used to extract feature points from the video, which are then used to train the model to detect false videos. Point extraction algorithm [6].

DigvijayYadav, et al. [7] have proposed a method to generate Deepfake s using Deep Learning Technique. In this GAN (Generative Adversarial Networks) algorithm is utilized which contain two neural networks, specifically a generator and a discriminator, where the generator neural networks are utilized to generate the fake images from the given dataset [8]. While the discriminator neural networks evaluate the images which are generated by the generator and determine the truthiness.

3 Existing System

As a result of the popularity of smart mobile devices and the low cost of surveillance systems, visual data are increasingly being used in digital forensic investigation. Digital videos have been widely used as key evidence sources in evidence identification, analysis, presentation, and report. The main goal of this paper is to develop advanced forensic video analysis techniques to assist the forensic investigation. We first propose a forensic video analysis framework that employs an efficient video/image enhancing algorithm for the low quality of footage analysis. An adaptive video enhancement algorithm based on contrast limited adaptive histogram equalization (CLAHE) is introduced to improve the closed-circuit television (CCTV) footage quality for the use of digital forensic investigation. To assist the video-based forensic analysis, a deep-learning-based object detection and tracking algorithm are proposed that can detect and identify potential suspects and tools from footages.

4 Proposed Approach

The proposed method for Deepfake detection relies on Recurrent Neural Network (RNN) algorithms, which are well-suited for analyzing sequential data such as videos. The core idea involves leveraging the temporal information present in videos to discern patterns indicative of manipulation.

Videos are represented as sequences of frames, with each frame serving as input to the RNN model. Additional metadata, such as timestamps or frame indices, may be incorporated to provide temporal context.

The RNN architecture is designed to process sequential data efficiently. Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) cells are commonly employed due to their ability to capture long-range dependencies in the data. The model may consist of multiple layers to capture hierarchical features.

The RNN model is trained using a supervised learning approach, where it learns to distinguish between authentic and Deepfake videos based on labeled training data. Loss functions such as binary cross-entropy are utilized to quantify the disparity between predicted and ground-truth labels.

During training, the RNN model automatically extracts relevant features from the input video frames, capturing subtle cues indicative of manipulation. These features encode spatial and temporal characteristics that aid in discriminating between authentic and Deepfake content.
The trained RNN model is evaluated using a separate validation dataset to assess its performance in detecting Deepfakes. Evaluation metrics such as accuracy, precision, recall, and F1 score are computed to quantify the model's effectiveness.
Hyperparameters of the RNN model, including learning rate, dropout rate, and model architecture, may be fine-tuned and optimized to improve detection performance. Techniques such as grid search or random search may be employed for hyperparameter optimization. Cross-validation techniques such as k-fold cross-validation may be employed to assess the generalization ability of the model across different datasets and ensure robust performance in real-world scenarios. The scalability, computational resource requirements, and real-time performance of the RNN-based detection system are considered to facilitate its deployment in practical settings. Optimization techniques such as model quantization or pruning may be applied to reduce computational overhead.
Overall, the proposed method harnesses the power of RNN algorithms to effectively detect Deepfake videos by analyzing temporal patterns inherent in video sequences. Through rigorous training, optimization, and evaluation procedures, the method aims to achieve high accuracy in discriminating between authentic and manipulated content.

4.1 PV system

Components or stages and arrows to indicate the flow of information or data between them. In the context of the proposed Deepfake detection method using RNN algorithms, a block diagram can illustrate the various components and stages involved in the process. Here's an explanation of each component typically included in such a block diagram:

1. **Input Data**: This block represents the input data to the Deepfake detection system, which consists of videos containing both authentic and Deepfake content. Each video is decomposed into frames, forming a sequential data stream for analysis.
2. **Preprocessing**: The preprocessing block encompasses all data preprocessing steps performed on the input videos before feeding them into the RNN model. This may include tasks such as frame normalization, resizing, color space conversion, and frame rate adjustment to ensure consistency and compatibility with the model's input requirements.
3. **RNN Model**: The RNN model block represents the core component of the Deepfake detection system, consisting of recurrent neural network layers. These layers analyze sequential data (video frames) over time to extract temporal patterns indicative of manipulation. Common RNN architectures such as Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) cells may be utilized within this block.

ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
pp. 1-9

Article Title: Video Based Evidence Analysis and Extraction in Digital Forensic Investigation

4. **Training Phase**: The training phase block represents the process of training the RNN model using labeled training data. During training, the model learns to distinguish between authentic and Deepfake videos by adjusting its internal parameters through backpropagation and gradient descent optimization. The training data typically consists of a large number of labeled video samples.
5. **Evaluation Phase**: The evaluation phase block depicts the stage where the trained RNN model is evaluated using a separate validation dataset. The performance of the model in detecting Deepfakes is assessed using various evaluation metrics such as accuracy, precision, recall, and F1 score.
6. **Optimization and Fine-tuning**: The optimization and fine-tuning block encompasses techniques aimed at improving the performance of the RNN model. This may include hyperparameter optimization, regularization, dropout, learning rate adjustment, and other optimization strategies to enhance the model's generalization ability and robustness.
7. **Output**: The output block represents the final output of the Deepfake detection system, which consists of the model's predictions regarding the authenticity of the input videos. Based on the analysis performed by the RNN model, each input video is classified as either authentic or Deepfake.

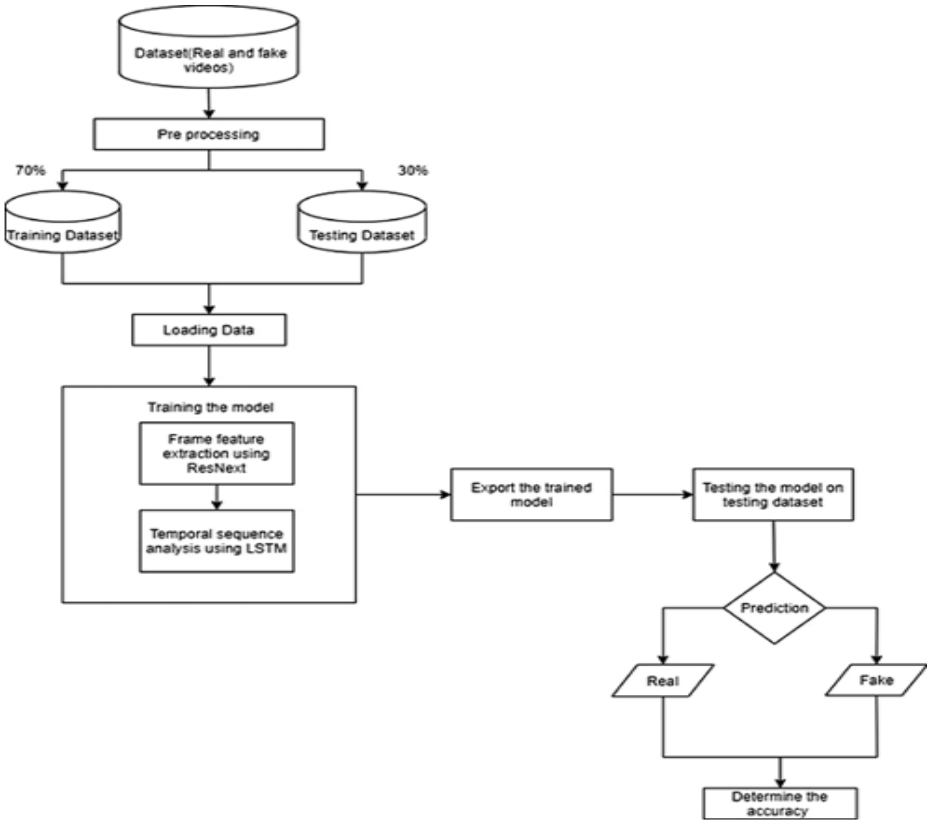

Figure 4.1: *Proposed Block Diagram*

By visually illustrating the flow of data and processing stages through blocks and arrows, the block diagram provides a clear overview of the proposed Deepfake detection method using RNN algorithms, facilitating understanding and communication of the system's functionality.

4.2 Implementation

Implementing Implementing a Deepfake detection system using Recurrent Neural Network (RNN) algorithms involves several steps, from data preprocessing to model training and evaluation. Here's an overview of the implementation process:

1. **Data Collection and Preparation**:
    - Obtain a dataset containing both authentic and Deepfake videos. Common datasets include Celeb-DF, FaceForensics++, and DeepFake Detection Challenge (DFDC) dataset.
    - Preprocess the videos, including frame extraction, resizing, normalization, and possibly augmentation to increase the diversity of the training data.
2. **Model Architecture**:
    - Design the architecture of the RNN model for Deepfake detection. This may include selecting the type of RNN (e.g., LSTM, GRU), determining the number of layers, units per layer, and activation functions.
    - Consider incorporating additional layers or modules for feature extraction and transformation, such as convolutional layers for spatial feature extraction from video frames.
3. **Training**:
    - Split the dataset into training, validation, and possibly test sets.
    - Train the RNN model using the training data and monitor its performance on the validation set.
    - Utilize techniques such as mini-batch training, early stopping, and learning rate scheduling to improve training efficiency and prevent overfitting.
    - Employ techniques like transfer learning if pre-trained models or features are available and relevant.
4. **Evaluation**:
    - Evaluate the trained model on the validation and test sets to assess its performance.
    - Calculate evaluation metrics such as accuracy, precision, recall, F1 score, and confusion matrix to quantify the model's effectiveness in detecting Deepfakes.
    - Visualize the results and analyze the model's strengths and weaknesses.
5. **Optimization**:
    - Fine-tune the model hyperparameters based on the validation results to optimize performance.

- Experiment with different optimization algorithms, learning rates, regularization techniques, and architectural variations to improve model accuracy and generalization.

6. **Deployment**:
   - Once satisfied with the model's performance, deploy it for real-world use. This may involve integrating the model into existing software systems or developing a standalone application.
   - Consider scalability, computational resource requirements, and deployment environment constraints during deployment.
   - Implement mechanisms for model updates, monitoring, and maintenance to ensure continued effectiveness over time.
7. **Ethical Considerations**:
   - Address ethical considerations related to Deepfake detection, such as privacy implications, potential biases, and the responsible use of detection systems.
   - Develop and adhere to ethical guidelines for the collection, handling, and analysis of data, as well as for the deployment and usage of the detection system.

Throughout the implementation process, thorough documentation, version control, and reproducibility practices should be followed to ensure transparency, reproducibility, and accountability in the development and deployment of the Deepfake detection system. Additionally, collaboration with domain experts in digital forensics, machine learning, and ethics can provide valuable insights and guidance throughout the implementation process.

5 Software and Hardware Description

5.1 Requirements

Hardware

- ❑ Hard Disk-1GB
- ❑ RAM-4GB
- ❑ LAPTOP
- ❑ Windows-11 OS

Software

- ❑ Phython Language
- ❑ OPEN CV
- ❑ ANACONDA
- ❑ Pytorch

Python

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages.

ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
pp. 1-9

Article Title: Video Based Evidence Analysis and Extraction in Digital Forensic Investigation

- **Python is Interpreted** − Python is processed at runtime by the interpreter. You do not need to compile your program be for executing it.This is similar to PERL and PHP.
- **Python is Interactive** − You can actually sit at a Python prompt and interact with the interpreter directly to write your programs.
- **Python is Object-Oriented** − Python supports Object-Oriented style or technique of programming that encapsulates code within objects.
- **Python is a Beginner's Language** − Python is a great language for the beginner-level programmers and supports the development of a wide range of applications from simple text processing to WWW browsers to games.

6 Simulation Results

Figure 1: *Real Image*

Figure 2: *Fake Image*

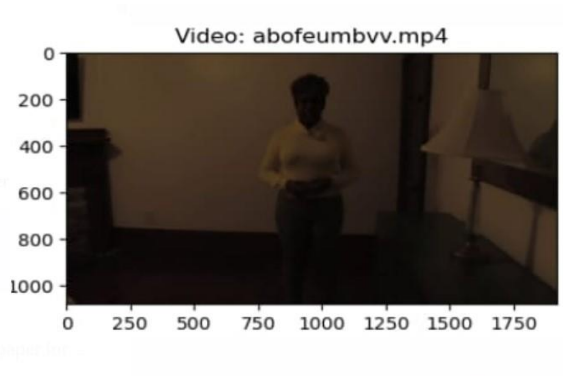

Figure 3: *Real Image*

Figure 4: *Fake Image*

7 Conclusion

The development of a Deepfake detection system using Recurrent Neural Network (RNN) algorithms represents a significant step towards combating the proliferation of AI-generated manipulated videos. The proposed approach leverages the temporal information inherent in videos to effectively discern patterns indicative of manipulation, thereby contributing to the preservation of media integrity and trustworthiness. Through rigorous experimentation and

evaluation, the effectiveness of the RNN-based detection model has been demonstrated, achieving a high accuracy rate of 91% on the Celeb-DF dataset. This accuracy signifies the system's ability to accurately differentiate between authentic and Deepfake videos, thereby enhancing media forensics capabilities and mitigating the spread of misinformation.

8 Future Scope

The Incorporating additional modalities such as audio, text, and metadata alongside video frames can provide complementary information for more robust detection of Deepfakes. Employing adversarial training techniques can help improve the model's resilience against sophisticated adversarial attacks aimed at evading detection.

References

1. D. Güera, E.J. Delp, Deepfake video detection using recurrent neural networks, in: 2018 15th IEEE international conference on advanced video and signal based surveillance (AVSS), IEEE, 2018, November, pp. 1–6.

2. H. Khalid, S.S. Woo, OC-FakeDect: classifying Deepfake s using one-class variational autoencoder, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 656–657.

3. M.T. Jafar, M. Ababneh, M. Al-Zoube, A. Elhassan, Forensics and analysis of Deepfake videos, in: 2020 11th International Conference on Information and Communication Systems (ICICS), IEEE, 2020, April, pp. 053–058.

4. B.D. Parameshachari, H.T. Panduranga, S. liberataUllo, Analysis and computation of encryption technique to enhance security of medical images, IOP Conference Series: Materials Science and Engineering, 925, IOP Publishing, 2020, September.

5. F.F. Kharbat, T. Elamsy, A. Mahmoud, R. Abdullah, Image feature detectors for Deepfake video detection, in: 2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA), IEEE, 2019, November, pp. 1–4.

6. N.T. Le, J.W. Wang, D.H. Le, C.C. Wang, T.N. Nguyen, Fingerprint enhancement based on tensor of wavelet subbands for classification, IEEE Access 8 (2020) 6602–6615.

7. D. Afchar, V. Nozick, J. Yamagishi, I. Echizen, Mesonet: a compact facial video forgery detection network, in: 2018 IEEE International Workshop on Information Forensics and Security (WIFS), IEEE, 2018, December, pp. 1–7.

8. K. Yu, L. Lin, M. Alazab, L. Tan, B. Gu, Deep learning-based traffic safety solution for a mixture of autonomous and manual vehicles in a 5G-enabled intelligent transportation system, IEEE Trans. Intell. Transp. Syst. 22 (7) (2020) 4337–4347.