INTEL AI HACKATHON

Team Nooglers

Deep fake Detection shield

SAI NIVEDH V	cb.sc.u4aie23062@cb.students.amrita.edu	Amrita Vishwa Vidyapeetham
Hari Heman	cb.sc.u4aie23062@cb.students.amrita.edu	Amrita Vishwa Vidyapeetham
Roshan T	cb.sc.u4aie23062@cb.students.amrita.edu	Amrita Vishwa Vidyapeetham
Baranidharan Selvaraj	cb.sc.u4aie23062@cb.students.amrita.edu	Amrita Vishwa Vidyapeetham

Problem statement

With the advancements in deep learning and computer vision techniques, the proliferation of deepfake technology poses a significant threat to the authenticity and integrity of information in today's digital age.

Deepfakes, which are artificially created videos or images manipulated using machine learning algorithms, can deceive viewers by replacing or altering parts of an original video or image, such as a person's face.

Detecting and distinguishing these manipulated media from real content is crucial for maintaining trust and accuracy in our digital landscape.

Solutions

We're building a deepfake detection system that utilizes machine learning and multimodal analysis (visual and audio) to identify manipulated content.

Our interface provides clear results and explanations, while the system is scalable and adaptable to future deepfake advancements.

We prioritize responsible development, addressing ethical concerns like fairness, transparency, and data security. Together, we aim to empower users and combat online manipulation with this innovative solution.

Methods :

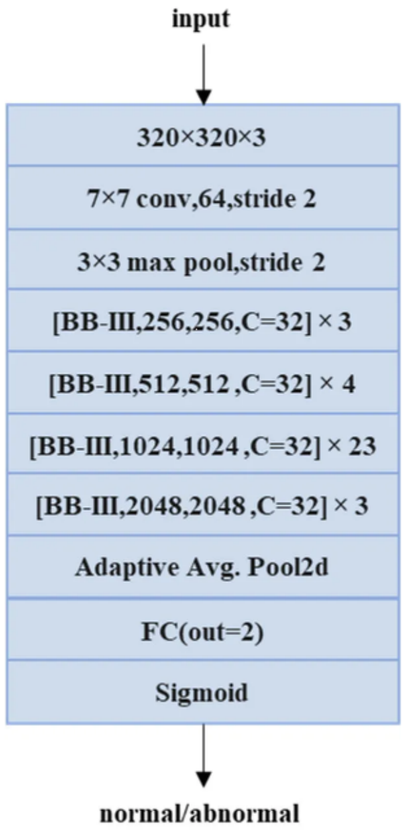
ResNext CNN and LSTM

Data Collection and preprocessing :

- Gather a large dataset of both real and fake videos, ensuring diversity in lighting conditions, facial expressions, backgrounds, etc.
- Preprocess the videos by extracting frames and converting them to a suitable format for input into the neural networks.

Methods :

Basic Arichitecture of ResNext 150



Methods :

Feature Extraction

- Utilize a pre-trained ResNext CNN model (such as ResNext-50 or ResNext-101) to extract high-level features from each frame.
- Fine-tune the ResNext CNN on the dataset to enhance its ability to discriminate between real and fake frames.

Methods :

Sequence Modeling with LSTM

- Construct an LSTM-based architecture to model the temporal dependencies in the sequence of extracted features from the frames.
- Design the LSTM to take sequences of features extracted by the ResNext CNN as input and learn to capture the dynamic patterns characteristic of deepfake videos.

Methods :

Model fusion and Decision Making

- Combine the output of the ResNext CNN and the LSTM in a fusion layer, which may involve concatenation or another merging technique.
- Train the fusion layer along with the rest of the model to learn optimal decision boundaries for distinguishing real and fake videos.

Methods :

Training and Evaluation

- Split the dataset into training, validation, and test sets.
- Train the entire model end-to-end on the training set, using techniques like regularization and data augmentation to prevent overfitting.
- Evaluate the model's performance on the validation set, adjusting hyperparameters as necessary.
- Finally, assess the model's generalization ability on the test set to determine its effectiveness in detecting deepfake videos.

Methods :

Post-processing and Deployment

- Implement any necessary post-processing steps, such as thresholding or confidence filtering, to refine the model's predictions.
- Deploy the trained model in a real-world setting, potentially integrating it into a larger platform for automated deepfake detection or forensic analysis.

Methods :

MesoNet (alterntive)

Data Collection and preprocessing :

- Gather a diverse dataset of both real and deepfake videos under clear conditions and we’ve standardize the format ,resolution and frame rate

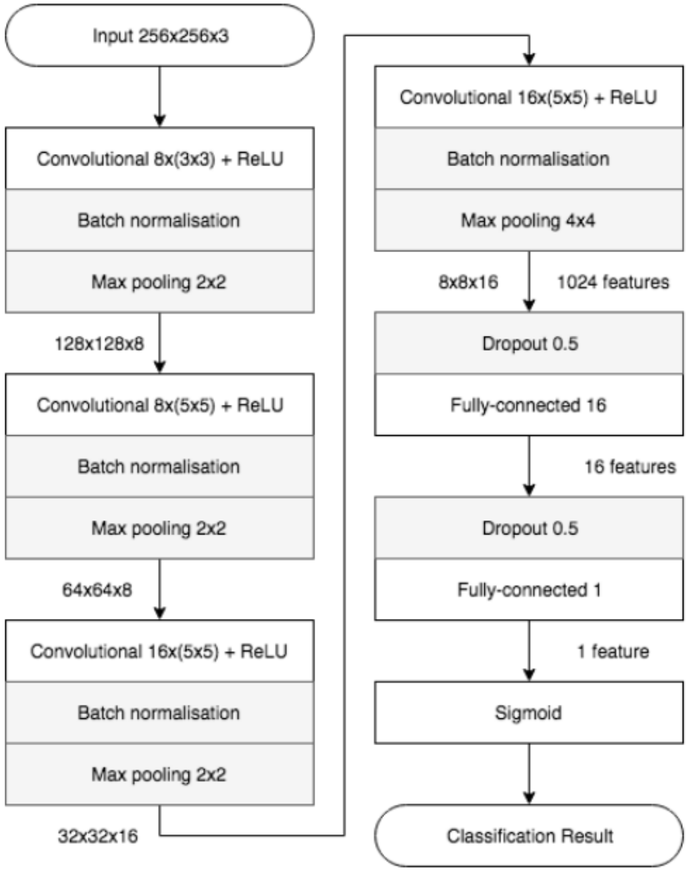
Methods :

MesoNet Architecture:

- Implement the MesoNet architecture, which typically consists of a deep convolutional neural network (CNN). MesoNet is specifically designed for detecting deepfake images and videos by analyzing subtle artifacts and inconsistencies introduced during the generation process.

Methods :

MesoNet Architecture:



Methods :

Training

- Train the MesoNet model using the preprocessed dataset. During training, the model learns to distinguish between real and deepfake videos by identifying features indicative of manipulation, such as unnatural facial expressions, inconsistent lighting, or artifacts introduced by the generation process.

Methods :

Validation :

- Validate the trained model using a separate validation dataset to assess its performance accurately. This step helps in identifying potential overfitting and ensures that the model generalizes well to unseen data.

Methods :

Fine-tuning and Optimization:

- Fine-tune the model and optimize hyperparameters to enhance its detection accuracy further. Techniques such as data augmentation, regularization, and learning rate adjustments may be employed to improve the model's performance.

Methods :

Evaluation

- Evaluate the performance of the trained MesoNet model on a test dataset comprising unseen real and deepfake videos. Metrics such as accuracy, precision, recall, and F1-score are commonly used to assess the model's effectiveness in detecting deepfakes.

Methods :

Deployment

- Deploy the trained MesoNet model for real-world applications, such as social media platforms or video hosting sites, to automatically identify and flag potential deepfake content. Continuous monitoring and updates may be necessary to adapt to evolving deepfake generation techniques.

Why we use ?

ResNext CNN and LSTM

- The wide residual networks in ResNeXt imply a strategic increase in the width (number of neurons per layer) rather than just depth (number of layers). This tactic enhances the network's learning capability without a proportional increase in computational complexity. Each layer in ResNeXt consists of a set of transformations, which are aggregated at the end.

- ResNeXt introduces a new dimension, called "cardinality," in addition to depth and width, commonly used in network design. This innovation enables the model to achieve higher accuracy

Why we use ?

MesoNet

- The wide residual networks in ResNeXt imply a strategic increase in the width (number of neurons per layer) rather than just depth (number of layers). This tactic enhances the network’s learning capability without a proportional increase in computational complexity. Each layer in ResNeXt consists of a set of transformations, which are aggregated at the end.

Intel ToolKit

- HPC Toolkit
- Intel open image denoise
- Intel OSpray
- Intel OpenVino
- Intel DevCloud
- DeepLearning WorkBench

References

MesoNet: a Compact Facial Video Forgery Detection Network

https://arxiv.org/pdf/1809.00888.pdfpact Facial Video Forgery Detection Network

FakeCatcher: Detection of Synthetic Portrait Videos using Biological Signals

https://arxiv.org/pdf/1809.00888.pdfpact Facial Video Forgery Detection Network