Documentation: Track 1- b

-Team 2

1. Setup Instructions

Prerequisites:

- Python 3.8 or later.
- A GPU-enabled system with CUDA installed (optional but recommended).

Dependencies: Install the required Python libraries by running the command:

pip install gradio torch torchvision matplotlib pandas opencv-python-headless

This will install:

- **Gradio**: For creating the interactive web application.
- **Torch and Torchvision**: For handling video data and implementing deep learning models.
- **Matplotlib**: For visualizing results.
- **Pandas**: For creating tabular reports.
- **OpenCV**: For video processing.

---

2. How to Use

1. **Launch the Interface:** Run the notebook cells sequentially. The script will initialize a Gradio interface. You can also run it as a standalone Python script.
   If running in Google Colab, the interface will generate a public link when launched.

2. **Input Requirements:**
    - **Real Videos**: Upload real videos (e.g., Sims4Action dataset).
    - **Synthetic Videos**: Upload synthetic video files.
3. **Process:**
    - Upload the video pairs in the respective input fields.
    - Click **Submit** to evaluate.
4. **Outputs:**
    - **Results Summary**: A table detailing the FID score for each video pair.
    - **Visualization**: A bar plot showing the FID scores.

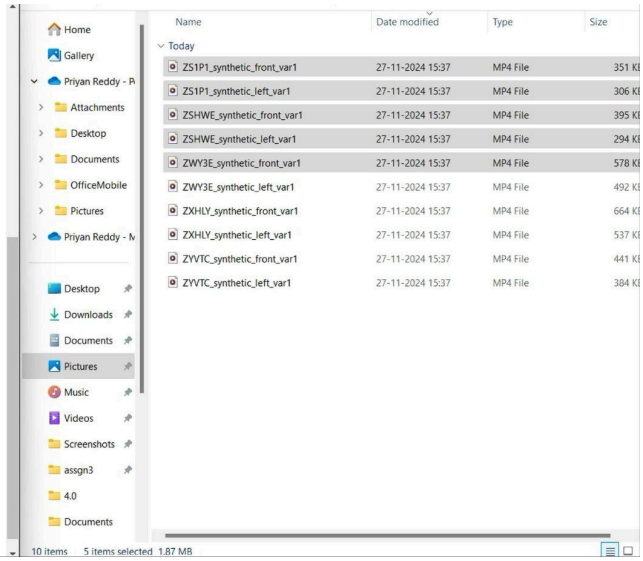
---

3. Code Logic and Functionality

1. **Core Functions and Classes:**
    - **VideoDataset**: A PyTorch dataset class that:
        - Loads video frames using OpenCV.
        - Ensures videos are sampled to a fixed sequence length.
        - Converts frames to tensors normalized to [0, 1].
    - **evaluate_videos(real_videos, synthetic_videos)**: The main evaluation function that:
        - Loads real and synthetic videos.
        - Uses a pre-trained Inception model to extract features.
        - Computes FID metrics using these features.
        - Creates a DataFrame summarizing FID scores and visualizations.
2. **Interactive Interface:**
    - Powered by **Gradio**, the interface accepts real and synthetic video uploads and outputs textual results and a plot.
3. **Dependencies on External Models:**
    - Relies on **FID Metrics** functions (imported but assumed pre-defined):
        - build_inception(): Loads a 2D Inception model for feature extraction.

- calculate_fid(): Computes the FID score.

Expected Outputs



---

5. Error Handling

- **No Videos Provided**: Outputs a message prompting the user to upload videos.
- **Video Loading Issues**: Logs detailed errors if frames cannot be extracted from a video.
