Evaluation Framework for Synthetic Human Action Recognition

Submitted By:

Chethan Maram (Se21uari082)
Mani Krishna Divi (Se21uari081)
Madhusudan Reddy (Se21uari091)
Rupen Akula (Se21uari121)

1. Introduction

This project is designed to evaluate the quality of synthetic human action videos using FID-VID and FVD metrics. It leverages a Gradio interface to allow users to:

- Compute FID-VID and FVD values for given video pairs.
- Visualize the metric results through a scatter plot.
- The evaluation framework makes use of the fid-metrics repository for metric computations.

2. System Requirements

Hardware:

- A system with sufficient computational resources (preferably with a GPU for faster computation).
- Installed Python (≥3.8).

The following dependencies are required for the project:

- gradio==3.42.0
- matplotlib==3.8.0
- Pillow==10.0.0
- PyYAML==6.0
- subprocess.run

Install the dependencies using:

- pip install -r requirements.txt

3. Installation Process

This project requires the fid-metrics repository for FID and FVD computations.

1. **Clone the repository:**

   git clone https://github.com/npurson/fid-metrics.git
2. **Navigate into the fid-metrics directory:**

   cd fid-metrics
3. **Set the environment variable for PYTHONPATH:**

   export PYTHONPATH=$(pwd):$PYTHONPATH
4. **Launch the Gradio Interface**:

   Run the main script: python main.py

   This will open the Gradio interface in your browser.

4. How to Use the Framework

1) **Select a Metric**
2) **Choose one of the following metrics:**
    a) FID-VID
    b) FVD
3) **Upload Videos**
    a) Upload two videos (synthetic and real) for evaluation.
4) **View Outputs**
    a) **Metric Results**: The interface will display FID-VID and FVD values.
    b) **Visualization**: A scatter plot comparing the computed metrics.

5. Outputs

- **Metric Results**:
  FID-VID Score
  FVD Score
- **Visualization**:

  A scatter plot comparing the metric values.

6. Results and Observations

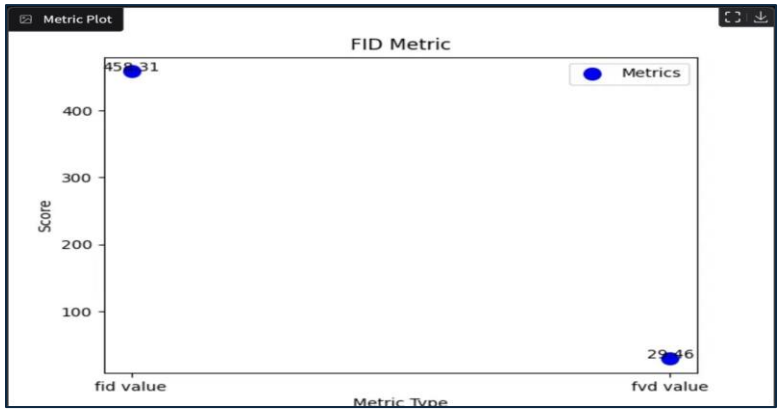
6.1 Sample Output Table:

- Measures perceptual similarity between real and synthetic videos. **FVD (Frechet Video Distance)**:
- Evaluates temporal coherence in videos

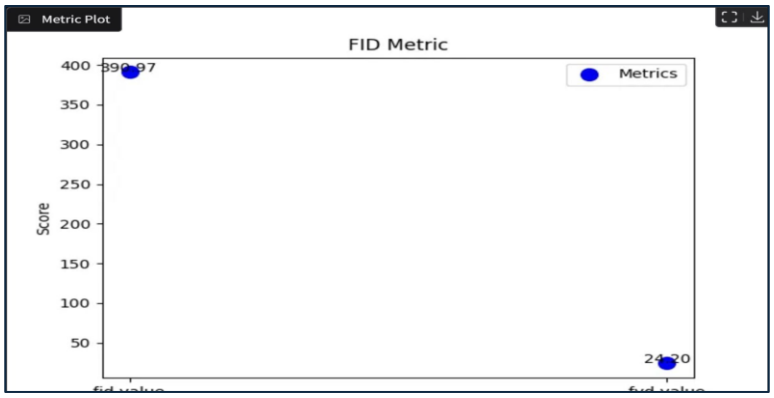
METHOD	FID-VID	FVD
VIDEO1	458.308	29.46
VIDEO2	390.97	24.20
VIDEO3	372.345	28.688
VIDEO4	431.778	25.68
VIDEO5	459.266	19.945

6.2 Graphical Representation:

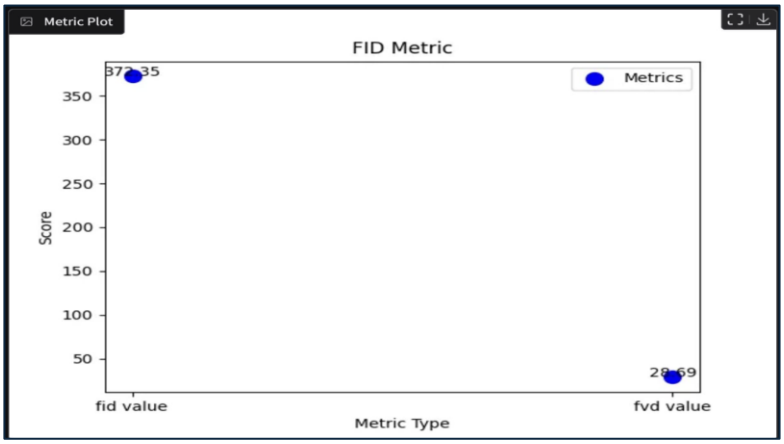
- **Video-1:**



- **Video-2:**

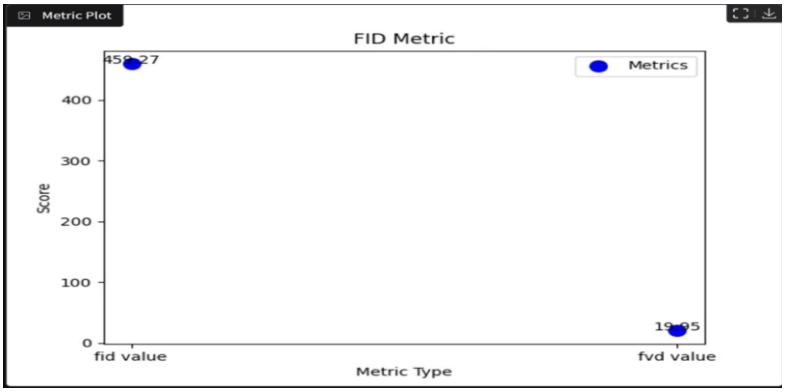


- **Video-3:**



- **Video-4:**

- **Video-5:**



7. Demo Video

- A demonstration video has been created to show the workflow and output of the framework.
- YouTube Link (Unlisted):

https://youtu.be/z1J4cyMaayU?si=pidEtYN5ifa1QxN8

8. Repository link:

https://github.com/dmknaidu/NvidiaAssignment