Deepfake Datasets: User Guide

ANAND M K - 242CS008

1 Datasets Overview

The following deepfake datasets are available for download and usage:

- **UADFV Dataset**
  Description: The UADFV (Unaltered and DeepFake Video) dataset contains deepfake videos and images created using various face-swapping techniques. This dataset is useful for performing statistical analysis of the quality and characteristics of deepfake images and videos.
  Download Link: `https://www.kaggle.com/datasets/adityakeshri9234/uadfv-dataset/data`

- **FaceForensics++**
  Description: FaceForensics++ is a large-scale dataset with manipulated videos containing real and deepfake faces. It includes several manipulation techniques, allowing for statistical comparisons and analysis of deepfake video attributes.
  Download Link: `https://www.kaggle.com/datasets/hungle3401/faceforensics`

- **OpenForensics**
  Description: OpenForensics is a dataset containing manipulated face images, primarily aimed at studying the statistical properties of deepfake images. It provides a diverse set of manipulations to analyze various deepfake techniques.
  Download Link: `https://github.com/ltnghia/openforensics`
  Google Drive Link: The dataset is hosted on Google Drive for easier access. You can download the data directly using the following link: `https://drive.google.com/drive/folders/1RjHbark1MKihDtAN3HJovzBCi5PGPW6E`

2 How to Use the Datasets

Follow the instructions below to download and use each dataset for statistical analysis:

1. **UADFV Dataset:**
   - Visit the dataset page on Kaggle using the provided link.

- Click on the "Download" button to get the data files.
- The dataset includes both real and deepfake videos/images, which can be used for statistical analysis of deepfake characteristics.

2. **FaceForensics++:**
   - Go to the FaceForensics++ dataset page on Kaggle.
   - After signing in to your Kaggle account, click on "Download."
   - Make sure you have enough storage space as this dataset is quite large.
   - The dataset contains deepfake videos that can be used for various statistical comparisons, such as temporal and quality analysis.

3. **OpenForensics:**
   - The OpenForensics dataset is hosted on GitHub, and the repository contains a link to download the dataset via Google Drive.
   - You can access and download the data directly from the Google Drive link provided in the GitHub repository. Ensure you have sufficient storage for the data.