Vision Transformer-Based Classification of Authentic vs. AI-Generated Human Faces

This repository contains the code and resources for a project focused on detecting AI-generated human faces using a fine-tuned Vision Transformer (ViT). The project aims to combat the proliferation of deepfakes and misinformation by providing a robust classification model.

Brief Summary

The project utilizes a ViT-Base/16 model, trained on a large dataset of 200,001 real and AI-generated human faces sourced from Kaggle. The model achieves a high accuracy of 93% on an unseen test set, demonstrating its effectiveness in distinguishing synthetic content from authentic imagery.

Data Source

The dataset used for this project is the **"200k Real vs AI Visuals"** dataset, which can be found on Kaggle.

- **Dataset URL:**
  https://www.kaggle.com/datasets/muhammadbilal6305/200k-real-vs-ai-visuals-by-mbilal

Repository Contents

- README.md: This file, providing an overview of the project.
- requirements.txt: Lists the Python libraries required to run the code.
- classification_model.py: The main Python script containing the code for the model, training, and evaluation.
- LICENSE: The license file for the project.
- index.html: A standalone web application to demonstrate the final product.

Requirements

To run this project, you will need to install the necessary Python libraries. It is highly recommended to use a virtual environment.

pip install -r requirements.txt

How to Run

1. **Download the Dataset:** Download the dataset from the Kaggle URL provided above and place the real_vs_ai_visuals folder in your project directory.
2. **Run the Script:** Execute the main Python script.

python classification_model.py

Final Output

The final output of the model is a binary classification label (Real or Fake) for an input image. The project report details the model's performance on the test set.

- **Accuracy:** 93%
- **AUC:** 0.99
- **F1-Score:** 0.93

Final Product

A simple web application is included to demonstrate the model's capabilities in a user-friendly interface. You can run this file directly in a web browser.

License

This project is licensed under the MIT License. See the LICENSE file for details.