Celebrity Image Classifier Documentation

INTRODUCTION

This Python script implements a Convolutional Neural Network (CNN) using TensorFlow and Keras to classify images of celebrities. The dataset comprises images of five celebrities: Lionel Messi, Maria Sharapova, Roger Federer, Serena Williams, and Virat Kohli. The script involves data loading, preprocessing, model building, training, evaluation, and prediction on individual images.

CNN

CNNs are widely used in computer vision tasks like image classification, object detection, and segmentation due to their ability to automatically learn and extract hierarchical representations of features from images. Convolution Neural networks trains on a set of images and classifier a newly given image data.

Requirements

- Python 3.x
- TensorFlow
- OpenCV
- NumPy
- PIL (Python Imaging Library)
- Scikit-learn

1. **Data Loading and Preprocessing**:

   - Images from respective directories are loaded, resized to 128x128 pixels, and stored along with assigned labels.

2. **Data Splitting**:

   - The dataset is split into training and testing sets using **train_test_split** from Scikit-learn.

3. **Data Normalization**:

   - Image pixel values are normalized using **tf.keras.utils.normalize**.

4. **Model Architecture**:

   - A CNN model is defined using Keras' Sequential API with convolutional layers, pooling layers, dropout layers, and dense layers.

5. **Model Training**:

   - The model is compiled using 'adam' optimizer and 'sparse_categorical_crossentropy' loss. Training is conducted over 35 epochs.

6. **Model Evaluation**:

- The model's performance is evaluated on the test set, and accuracy metrics are calculated. Additionally, a classification report with precision, recall, and F1-score is generated.

7. **Model Prediction**:

   - A function **make_prediction** is provided to make predictions on individual images using the trained model.

Key Findings

Accuracy

The accuracy of the CNN model is 76.47. The accuracy is highly dependent on the number of epocs in this case.

The model was able to predict the celebrity accurately when a new image data was given.