Sports Celebrity Image Classification Project Documentation

Overview

The "Sports Celebrity Image Classification" project aims to classify images of sports celebrities into different categories. The project utilizes image processing techniques and machine learning to achieve this task. The classification is based on detecting and analyzing facial features, particularly the eyes, using a Support Vector Machine (SVM) classifier.

Project Contributors

Special thanks to **Debjyoti Paul**, a data scientist at Amazon, for contributing to the development of this project.

Table of Contents

- The dataset is prepared by iterating through all images, detecting faces, and creating cropped images with at least two visible eyes. The resulting dataset is organized into folders for each celebrity.

6. **Model Training**

- A Support Vector Machine (SVM) classifier with a Radial Basis Function (RBF) kernel is chosen for model training. The training set consists of both raw images and wavelet-transformed images.

7. **GridSearch for Model Fine-Tuning**

- GridSearch is employed to explore different models and fine-tune their parameters. The goal is to identify the best model with optimal parameters.

8. Model Evaluation

- The trained model is evaluated on the test set, and a classification report is generated to assess its performance.

9. Confusion Matrix

- A confusion matrix is created to visualize the performance of the model in terms of true positive, true negative, false positive, and false negative predictions.

10. Save Trained Model

- The trained model is saved as a pickle file, and the class dictionary mapping celebrity names to numerical labels is stored as a JSON file.

11. Conclusion

- The "Sports Celebrity Image Classification" project demonstrates the use of image processing and machine learning techniques to classify images based on facial features. The documentation provides a step-by-step guide on preprocessing, data preparation, model training, and evaluation. Users can utilize the saved model and class dictionary for predicting the sports celebrity category of new images.