

# Celebrity Image Classification with CNN

## Dataset

The dataset comprises cropped images of five celebrities: Lionel Messi, Maria Sharapova, Roger Federer, Serena Williams, and Virat Kohli. Each celebrity has a dedicated directory containing PNG images.

## Data Preprocessing

Images are loaded using OpenCV and resized to a uniform size of (128, 128) using the Pillow library.

The dataset is constructed by appending images and corresponding labels to lists (dataset and label).

The dataset is converted to NumPy arrays for further processing.

## Train-Test Split

The dataset is split into training and testing sets using the `train_test_split` function from scikit-learn.

## Data Normalization

Pixel values of images are normalized to the range [0, 1].

## Model Architecture

A Convolutional Neural Network (CNN) is implemented using TensorFlow's Keras API. The architecture consists of:

- Convolutional layer with 32 filters, kernel size (3,3), and ReLU activation.
- MaxPooling layer with a pool size of (2,2).
- Flatten layer to convert the 2D feature maps to a vector.
- Dense layers with ReLU activation and Dropout for regularization.
- Output layer with 5 units (number of classes) and softmax activation.

## Model Compilation

The model is compiled with the Adam optimizer, categorical crossentropy loss, and accuracy as the metric.

## Model Training

The model is trained for 50 epochs with a batch size of 32, using a 20% validation split.

## Model Evaluation

The trained model is evaluated on the test set, and accuracy along with a classification report is printed.

## Model Prediction

The model is used to predict the celebrity class for a set of sample images, and the predicted labels are printed.