Assignment 4

Problem Statement:

We have been given datasets containing images of Pokemon and various statistics about them. We have been tasked with finding the following values:

1. Type of Pokemon (ID Number)
2. CP Detection
3. HP Detection
4. Stardust Detection
5. Level of the Pokemon

Solution:

Initially, a model is created using training data. This model is to be used to identify imaged based on image features. This model is formed using Principal Component Analysis (PCA) .

Training:

a. Preprocess (using **imgaussfilt()**, **rgb2gray()** and **imresize()**) and crop (using **imcrop()**) the images to only the Pokemons' body
b. Represent each image as feature vectors
c. Calculate the average of the feature vectors to create the face vector. For this we use the **mean()** function.
d. Find the difference between the mean face vector and individual face vectors
e. Create a covariance matrix based on these values using **cov()**
f. Find the eigenvectors of this covariance matrix and the eigen vector feature to be used as the model in the approach. Eigenvectors can be found using the **eig()** function.

Now we can move on to finding the values required by the problem statement

1. Type of Pokemon
   In this step, we use the model created in the training phase to match and find the identities of the Pokemon.
   a. First, we preprocess and crop the images to the same dimensions used in the training phase. The same functions are used as in the training phase.
   b. Subtract the mean face vector from the image
   c. Next, we multiply the model's eigenvector and the difference calculated in the previous step to find the most similar image present in the training set.
   d. Check each image in the model and find the most similar image and its index.
   e. This index is our predicted ID value.

2. CP Detection

In this step, we find the CP of the Pokemon. The CP value is shown on the top center of the image.

   a. Crop the image using **imcrop()** so that only the letters CP and its corresponding values are in the image.
   b. Perform preprocessing on the image. Use **imgaussfilt()**, **imbinarize()** and **imcomplement()**
   c. Use the **ocr()** function to find out the characters in our image. These characters are stored as words, stored in cell arrays.
   d. Find which cell array contains the string 'CP'
   e. Extract the numbers from the string using **regexp()** function.
   f. Assign value to CP.

3. HP Detection
   This step is similar to CP Detection.
   a. Crop the image using **imcrop()** so that only the HP value is in the image.
   b. Perform preprocessing on the image. Use **imbinarize()**
   c. Use the **ocr()** function to find out the characters in our image. These characters are stored as words, stored in cell arrays.
   d. Find which cell array contains the string numbers separated by a '/' symbol.
   e. Use **split()** function to split the string at the '/' symbol.
   f. The first half of the string is the HP value

4. Stardust Detection
   a. This step is similar to the previous steps.
   b. Sharpen the image using **imsharpen()**
   c. Crop the image using **imcrop()** so that only relevant parts are visible
   d. Perform preprocessing on the image. Use **imbinarize()**
   e. Use the **ocr()** function to find out the characters in our image. These characters are stored as words, stored in cell arrays.
   f. Use **regexp()** to find the value for stardust.
   g. Assign this value to stardust.

5. Finding the Level of the Pokemon.

   The level of the Pokemon is represented by a dot on a semicircle in the top half of the screen. This involves two steps: finding the the dot and finding the semicircle.

   a. Preprocessing is done on the image.
   b. Initially, we use **imfindcircles()** to find the circles using Hough Transforms.
   c. If this fails, we use **regionprops()** to find the blobs corresponding to a circle.
   d. Next, we use three random points on the semi-circle arc to find its center.
   e. Finally, we use **imfindcircles()** to find the dot on top of the semicircle.

Results

The accuracy for the values found are

1. ID: 14.71%
2. CP: 33.82%
3. HP: 47.79%
4. Stardust: 30.88%