Universidad Politécnica de Madrid Information retrieval, extraction and integration

Implementation of a 'toy' CBIR

Béla Bönte, Daniel Marouf

Academic year 2022/2023

Contents

1 Proposal of CBIR application in a real problem

In the last 2 years, the hype around Pokemon has increased rapidly. Driven by Youtubers like Logan Paul, a first edition Charizard can now reach 400,000 or more.(GeeksforGeeks, 2022) So it's only logical that this hype has also introduced many new people to Pokemon. Each card is characterized by an illustration, but also properties (e.g. Type, Species, Abilities). Therefore, the goal is to provide Pokémon interested people with the appropriate information about the Pokemon, simply by using an image search.

2 Technical proposal of Pokemon Checker

In order to effectively search for pokemons based on the image, data is needed on the one hand and methodologies to automatically recognize images on the other. Therefore, the foundation is created by collecting the data. First, the images of the Pokemons must be present in a data set. In the following dataset there are 809 Pokemon images.

Main dataset link

In addition to the image, the data set also contains the name and type of the pokemon. But to show the user the properties of a Pokemon, more data is needed. Therefore, another data set is used:

Additional dataset link

This dataset contains additional properties, then both datasets can be merged, based on the one-to-one assignment of Pokemon names.

The second step is to match the searched Pokemon with the stored images in the database. For this purpose, the image must first be described in a computer understandable form. The description of the image can be done by different methods. These methods are called discriptors. In the following, three discriptors are presented and explained.

- Histogram Distance for color discriptor
- Histogram Distance for greyscale discriptor
- Distance with Zernicke Distance

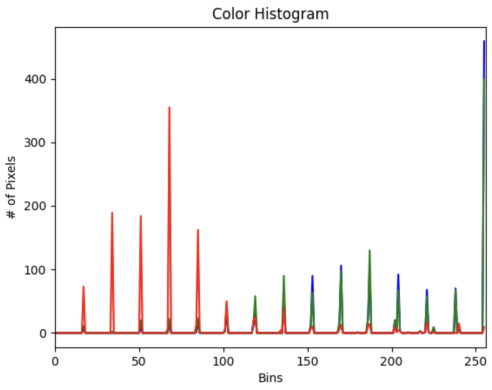

Figure 1: Color histogram.

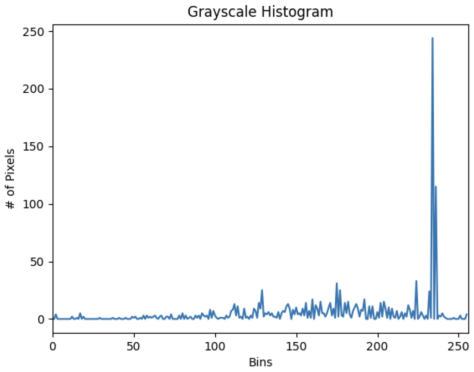

Figure 2: Gray histogram.

3 Implementation of the color and grey Histogram descriptor

The basic idea was that pokémon come in many colours and different families have different colour compositions. So given a pokémon, it's quite likely that the most similar species will come from the pokémon family based on colour.(Guiguilegui, 2016) In order to show that colours are good attributes for the task, we implemented the grey histogram, which we expected to produce worse results, since we store less information about an image and therefore it is easier to find matches with other elements.

First of all, we scanned the images into the system, then for the colour image, we split the stack into four parts (the three main colours and transparency). Then in both cases we removed the zeros from the render because without this the system would have taken the background black and so in both cases we cleaned the pictures so that the similarity in the later comparison. Two histograms of the selected pokemon were also produced. Two histograms of the selected pokemon were also produced.

Finally, to compare with the other images, we went through all the images in both cases and indexed the resulting matrices to the image name.

4 Implementation of the Histogram distance for the color and grey descriptors

Two types of distance methods were used to measure similarity.

4.1 Hellinger distance

The Hellinger distance is a limited metric on the set of probability distributions over a probability space that is defined by a given probability field. When P assigns probability zero to any set to which Q assigns a positive probability, and vice versa, the maximum distance 1 is attained, the maximum distance 1 is known as the maximum distance 1. (Wikipedia, 2022)

4.2 Manhattan distance

Manhattan distance is a distance metric between two points in an N-dimensional vector space that is defined as the distance between them. It is the total of the lengths of the projections of the line segment between the points onto the coordinate axes that makes up the line segment between the points. In layman's words, it is the total of the absolute differences in the measurements of two points over all of their dimensions. (Opengenus, 2022)

5 Implementation of a shape descriptor

Another way to distinguish images from each other and examine them for similarities is to compare their shape. This can be achieved with various methodologies, one of which is Zernicke moments. For this purpose, Zernike polynomials are calculated as an "orthogonal basis set (a set of functions for which the integral of the product of any pair of functions is equal to zero)."(GeeksforGeeks, 2022) "However, since the Zernike polynomials are orthogonal to each other, there is no redundancy of information between moments." (Rosebrock, 2021)

Since Zernike's moment is used here to characterize the outline of a Pokemon, the

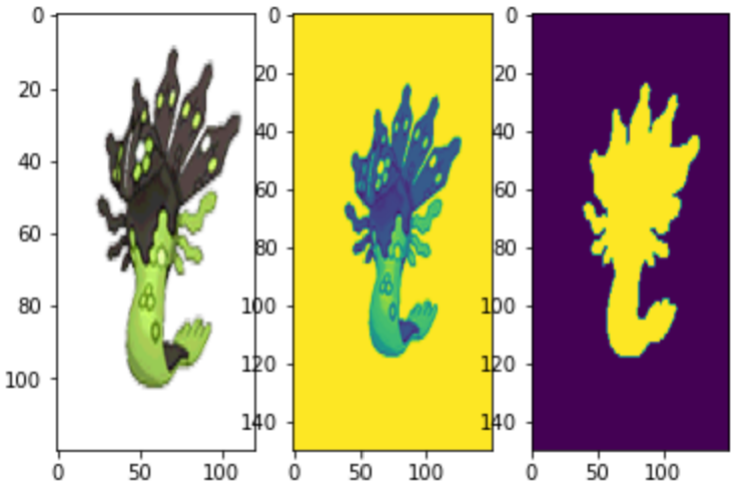

Figure 3: Image preprocessing for Zernikes distance

image must first be segmented. Colors are not important for the outlines, so the image is colored black and white. The following figure shows the steps from the input image to the image for which the Zernike moments are calculated. The purple and yellow areas are black and white.

The Zernike moments are then calculated for the black, white segmented image. This sophisticated computational operation is performed by the Mahotas Library.

6 Implementation of CBIR processes

With the help of discriptors, different features are calculated for each image. Like the color, the pixel intensity or different shapes. These values can be stored to greatly reduce the search time. The indexes used here are separated by the discriptors, but could also be combined in more sophisticated approaches. Each index consists of a key (one-to-one Pokemon-name) and the associated calculated values.

7 Documentation

The code has been uploaded to GIthub and is available at this link: Github Repo

8 Presentation

To further improve the functionality, we have created a website. Currently, the site works by selecting a pokémon from the main page and then clicking on run the code and in the next page it will show you information and measurements of the pokémon image. https://pokemonchecker.site

References

[1] GeeksforGeeks. *Zernike Distance*. 2022. URL: https://www.geeksforgeeks.org/mahotas-zernike-moments/?ref=lbp.

[2] Gui Gui Le Gui. *Pokémon type classifier using their colors*. 2016. URL: https://guiguilegui.wordpress.com/2016/05/23/pokemon-type-classifier-using-their-colors/.

[3] Opengenus. *Manhattan Distance*. 2022. URL: https://iq.opengenus.org/manhattan-distance.

[4] New York Post. *Pokémon Charizard card sold for a whopping* $420,000 at auction.. 2022. URL: https://nypost.com/2022/03/28/pokemon-charizard-card-sold-for-whopping-420k-at-auction/.

[5] Rosebrock. *Indexing with Zernike Moments*. 2021. URL: https://pyimagesearch.com/2014/04/07/building-pokedex-python-indexing-sprites-using-shape-descriptors-step-3-6/.

[6] Wikipedia. *Hellinger Distance*. 2022. URL: https://en.wikipedia.org/wiki/Hellinger_distance.