REVERSE IMAGE SEARCH ENGINE

[1]ANIRUDH. S, [1]VINAY. M, [1]SACHIN S. SRIVATHSA, [1]R. ASHWIN

[1]III Year B.Tech. Students

Department of Computer Science and Engineering

SRM Institute of Science and Technology, Ramapuram Campus, Chennai, India

Abstract : Reverse Image search engines are those special kind of search engines where the user don't need to input any keyword to find pictures. Instead, you must put a picture and the engine finds the images like one you entered. Thus, you can get to know everything you wish to, just with the help of one picture.

Google images is a widely used website to search images. Introduced 15 years ago, Google images have the largest database than most of the other sites. In June 2011, Google Images introduced the feature of reverse image search. The mechanism of reverse photo search here is via uploading an image from your computer or pasting the link of the image in the search bar itself, it works equally well both ways, or you can simply drag and drop the image in search bar. Google images use algorithms based on various attributes like shape, size, colour and resolution to get the similar pictures.

The goal of this project is to create a basic image search engine. This search engine is localized and based on images that are indexed by the user. For this project, I have indexed photos and then calculated the colour histogram of these photos.

Later the user has used animage from one of the 5 locations, calculate the histogram and check for similar images, the goal being getting the images from the same location, in the top 5 result images.

Index Terms – Images, search engine

I. PROBLEM IDENTIFICATION

Existing System:

With Current Image Search technology even though the objective to search images is the same, its technology covers a variety of matters, and is roughly classified as follows. Image search Of digitized images, although trials to extract in conformity with a required condition had started since the latter half of the 20th century, the upper limit of computers' calculation speed and data volume to be accumulated was strict initially. In recent years, due to the enhancement of technology in the overall sector of computers, data volume and search speed have reached the limit capable of practical applications, and research of methods as to how efficient and correct search can be carried out have flourished. Common object recognition Of image search technologies, this is the one to learn characteristics from multiple targeting graphics and recognize the same category as its object. For example, if a box type thing with four tires is recognized as a "Vehicle," recognition is carried out for a thing with similar characteristics as a "Vehicle."

Drawbacks of Existing System:

The image search engine of firms has been characterized by technology called object recognition (if it is comprised of graphics composed of ○△□, it is a technology capable of recognizing each independently), not all but individual objects can independently be recognized and judge similarly. Loaded graphics for targeting searches comprehensively digitize the vectors at geometric character for each minimum unit object and output search results by computing those from both the parts and the whole and determining the similarity of the individual or overall. While it is a technology that has been developed mainly for the purpose to automatically recognize specific objects of figures and characters from digitized photographs or frames, this technology is also effective in search of line drawing figures such as trademarks and technological development is being implemented. While photographs have depth, shading or the like, since trademarks essentially are flat surfaces, decent outcomes can be brought about without particular tunings.
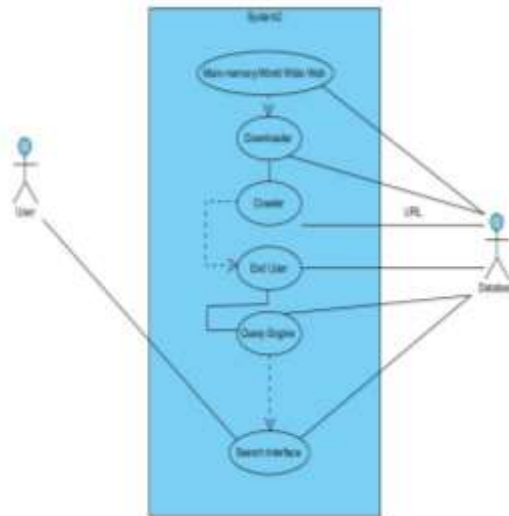
II. ARCHITECTURE DIAGRAM

Fig. 1 Architecture Diagram

III. MODULES

The project is divided into 4 modules as follows:

Descriptor – We have used 'color' as the main image descriptor.

Indexing – The cPickle module that we have used for indexing. "Pickling" is the process whereby a Python object hierarchy is converted into a byte stream, and "unpickling" is the inverse operation, whereby a byte stream is converted back into an object hierarchy.

Similarity Matrix - The similarity of 2 images is calculated by using a distance metric. In this case, the chi-squared distance was used.

Searching - We apply the descriptor on the query image and then ask your distance metric to rank how similar your images are in your index to your query images. Sort your results via similarity and then examine them.

To run the program successfully, the following packages will have to be installed for Python: a) Numpy b) cPickle c) OpenCV d) Glob e) Tkinter.

Tkinter:

The Tkinter module ("Tk interface") is the standard Python interface to the Tk GUI toolkit. Both Tk and Tkinter are available on most Unix platforms, as well as on Windows systems. This Python function will parse the commands and the arguments and convert them into a form that makes them look as if they had come from a Tk script instead of a Python script.

OpenCV:

OpenCV (Open Source Computer Vision Library) is released under a BSD license and hence it's free for both academic and commercial use. It has C++, C, Python and Java interfaces and supports Windows, Linux, Mac OS, iOS and Android. OpenCV was designed for computational efficiency and with a strong focus on real-time applications.

Cpickle:

The pickle module implements an algorithm for turning an arbitrary Python object into a series of bytes. This process is also called serializing" the object. The byte stream representing the object can then be transmitted or stored, and later reconstructed to create a new object with the same characteristics.

IV. TECHNICAL REQUIRMENTS

Software Requirements:

Python: Python is an interpreted high-level programming language for general-purpose programming. Created by Guido van Rossum and first released in 1991, Python has a design philosophy that emphasizes code readability, notably using significant whitespace. It provides constructs that enable clear programming on both small and large scales.

VMware: VMware, Inc. is a subsidiary of Dell Technologies that provides cloud computing and platform virtualization software and services. It was the first commercially successful company to virtualize the x86 architecture.VMware's desktop software runs on Microsoft Windows, Linux, and macOS, while its enterprise software hypervisor for servers, VMware ESXi, is a bare-metal hypervisor that runs directly on server hardware without requiring an additional underlying operating system.

V. CONCLUSION

Reverse image search is one of those handy innovations that's often hard to come up with specific uses for. Sure, you can use it to track down the origin of a photo, but it's also useful for hunting down product names, apartments, and even recipes. Reverse Image search engines are those special kind of search engines where you don't need to input any keyword to find pictures. Instead, you must put a picture and the engine finds the images like one you entered. Thus, you can get to know everything you wish to, just with the help of one picture.

VI. REFERENCES

1. https://www.pyimagesearch.com/2014/12/01/complete-guide-building-image-search-engine-python-opencv/
2. https://github.com/jayneeshw/Image_Search_Python/blob/master/README.md
3. http://www.pyimagesearch.com/
4. http://opencv.org/
5. http://www.lfd.uci.edu/~gohlke/pythonlibs/