

## A Content Based Image Retrieval Model for E-Commerce

Gibson Kimutai, Prof. Wilson Cheruiyot, Dr. Calvins Otieno

Student, School of Computing and Information Technology, JKUAT, Kenya

Associate Professor, School of Computing and Information Technology, JKUAT, Kenya

Lecturer, School of Computing and Information Technology, JKUAT, Kenya

### Abstract:

In the last decade, large database of images have grown rapidly. This trend is expected to continue in to the future. Retrieval and querying of these image in efficient way is a challenge in order to access the visual content from large database. Content Based Image Retrieval (CBIR) provides the solution for efficient retrieval of image from these huge image database. Many research efforts have been directed to this area with color feature being the mostly used feature because of its ease of extraction. Although many research efforts have been directed to this area, precision of majority of the developed models are still at less than 80%. This is a challenge as it leads to unsatisfying search results. *This paper proposes a Content Based Image Retrieval model for E-Commerce.*

**Keywords:** *Image retrieval, text-based image retrieval, content-based image retrieval, E-Commerce, E-Bay, shape, performance evaluation, precision, recall*

### 1.Introduction

In the present day world of mass internet penetration and the advent of the e-commerce era the number of products being bought and sold online has increased exponentially in the past few years (Bengio, 2011). In 2012, Business to Consumer (B2C) E-commerce sales grew by 21.1% to top \$1 trillion for the first time (Bengio, 2011). This is expected to grow steadily at the rate of around 20% and is estimated to hit \$2.5 trillion by 2018 (Bengio, 2011). Given the explosive growth in the number of products being sold online and the relative heterogeneity in the categories these products could be allotted to, it has become physically impossible and infeasible to manually tag these products (Bengio, 2011). Besides not everyone will tag the same images with the same tags. This leads to discrepancy in the kinds of tags allotted to the products (Bengio, 2011). Search engines looking for products based on customers query heavily rely on these tags allotted to each image to return accurate and meaningful results to customer's queries but mainly only the product images are available which is impossible for the search engine to make sense of (Bengio, 2011). Besides the discrepancy in tagging leads to a lot of useful search results to get excluded (Bengio, 2011).

Electronic commerce, or E-commerce, is defined to be the process of businesses trading with other businesses and the formulation of internal processes using electronic links (J. Ohene-Djan, 2008).

Popular internet commerce websites such as Amazon, Ebay, Snapdeal and FlipKart contains tremendous amounts of product related image (Kore, Tidke, & Chandre, 2015). Images in E-commerce websites are accompanied by tags, links, producer and Consumer information, annotations and comments as shown in figure 2.1 (Kore et al., 2015)

### 2.0 Methodology

The model is made for image retrieval in online business. The online shopping model is used for enhancing the system clients understanding. We have designed a model which the user can understand and use it easily for a broad e-commerce application, such as image search, and item data gathering subsystem. This model may give a direct contact between images in the database and our online shopping, on which purchasers are automatically and found directly. Simultaneously, it can be utilized to give exactness of online advertising to companies. Figure 1, shows the proposed model's components.

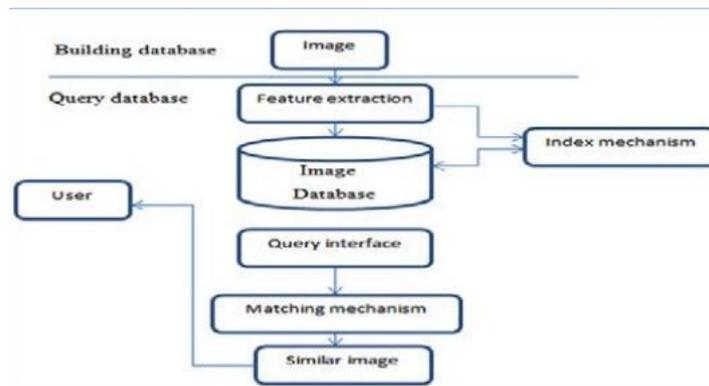


Figure 1: The components of similar image retrieval Approach

In this paper, the implementation is based on some programming languages and techniques, which are explained as follows: Laravel PHP framework MVC is used for the web application structure that is supported by Microsoft that executes the model-view-controller (MVC) design. The programming depends on Laravel PHP framework and it permits software designers to fabricate a web application as a composition of three parts: Model, View, and Controller. The MVC model characterizes web applications with 3 rationale layers:

- I. Model (business layer)
- II. View (display layer)
- III. Controller (input control)

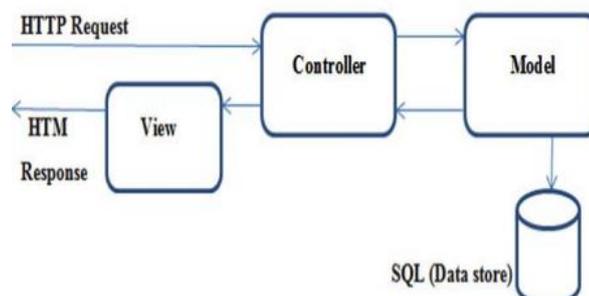


Figure 2: MVC structure query

Figure 2 , displays how MVC works together. MYSQL Server is used to implement a database in this project because it is secure and it can process many users at the same time (Ashraf, Bashir, Irtaza, & Mahmood, 2015). Search image in this system will depend on Laravel framework and also it can be used on visual programs either, but because this work has been created for images, that is why visual Laravel; it has more compatibility for such things, because of security and flexibly. The system coding consists of four class for each searching images inside the database for comparing Red, Green, and Blue (RGB). If the image which has been searched for was existed in the database or had the most similarities there by 90% or more, then the result will be shown it takes the top on the searching results. This paper uses a method which is called process image (fileinfo[ ]files, fileinfo target):

- I. FileInfo[] files: This parameter is the list of all of the images that have been sent to the database.
- II. FileInfo target: This parameter is the image which is compared with the list of the images in the database i.e. File Info[] files. Thus, it will retrieve all those images in the database which resembles 90% of the image which has been used in the search engine. There is a class which is called Comparable Image to calculate the matching by passing each one of them through the calculation.

## 2.1 Dataset used

To implement the Developed content based image retrieval model, Ebay database is used. The reason for choosing Ebay over other E-Commerce sites like Amazon and Jumia is that it has an Application

Programming Interface where developers are encouraged to experiment on the search and results. Many authors including (Cai, Gao, Yu, Huang, & Cai, 2017) have reported that E-Bay is the most influential E-commerce thus our model is worth testing with a world class site. E-bay have a wide range of variety of products but for our experiment we choose the following categories

**Table 1:Dataset**

| S/no | Category    | Items  |
|------|-------------|--|
| 1    | Electronics | Smartphones, Laptops, Cameras, Televisions   |
| 2    | Fashion     | Dresses, Trousers, Shirts, Watches, Handbags |
| 3    | Motors      | Cars and Car Electronics                     |

### 3.0 Results and Discussion

This application uses image processing to aid shopping. The developed system is allowing the user to submit an image and will then return similar items using image retrieval techniques that exist in the database. The main results in this project are represented by retrieved image displaying accuracy. The retrieval results are not only one image but could be a group of images displayed by their similarities to the image that exists in the database. The retrieval results find similar images in E-Bay site and show prices of those items and directs users on how to purchase them.

The steps in the implementation of the model are:

#### Step 1-Authentication

#### Step 1 - Browse query image

In this step the users chooses input image to retrieve the similar images from the database. after choosing image the user loads image so that the query system check for similar images in the database.

#### Step 2 - Select the Image for Retrieval

The type of feature used for retrieval depends on the type of images within the collection. For example, if searching an image collection made up of flowers, it is preferable to use a global image feature, such as a color histogram that captures the color content of the entire scene. However, if the goal is to find specific objects within the image collections, then local image features extracted around object key points are a better choice

The following images are going to help getting an idea of how to approach the problem.



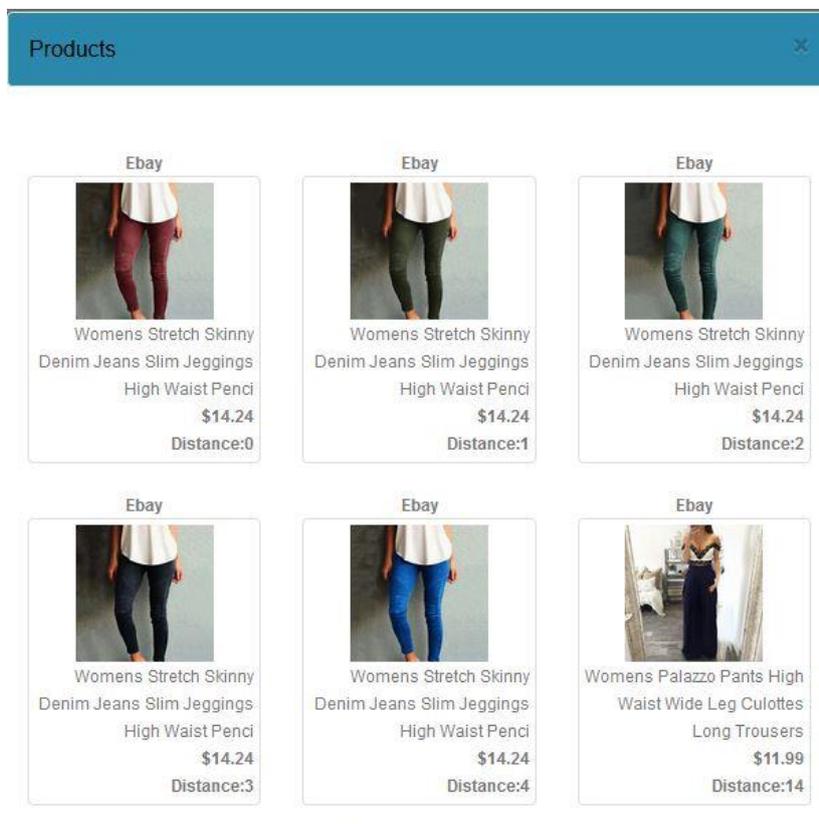
### Step 3 - load the Images in the database

Now the Features are created, the entire flower images set can be loaded for search. The loading procedure extracts features from each image using the custom extractor function from step 1. The extracted features are encoded into a visual word histogram and added into the image load.

Because the loading step processes thousands of images, the rest of this example uses a saved load to save time. A load may be created locally by running the step given above. Consider enabling parallel computing to reduce processing time.

### Step 4 - Search for Similar Images

The final step is to search for relevant images. Search images option returns the image IDs



| Product  | Price   | Distance    |
|--|---------|-------------|
| Womens Stretch Skinny Denim Jeans Slim Jeggings High Waist Penci | \$14.24 | Distance:0  |
| Womens Stretch Skinny Denim Jeans Slim Jeggings High Waist Penci | \$14.24 | Distance:1  |
| Womens Stretch Skinny Denim Jeans Slim Jeggings High Waist Penci | \$14.24 | Distance:2  |
| Womens Stretch Skinny Denim Jeans Slim Jeggings High Waist Penci | \$14.24 | Distance:3  |
| Womens Stretch Skinny Denim Jeans Slim Jeggings High Waist Penci | \$14.24 | Distance:4  |
| Womens Palazzo Pants High Waist Wide Leg Culottes Long Trousers  | \$11.99 | Distance:14 |

On the other hand, this work has made the system convenient to the user and it doesn't need to waste too much time for searching products. This is realized by the fact that is only needed to upload any item, and then the results will display similar images in a fast way as the system a fast connection to the database. The system has been developed based on the experience gained through our investigation to other system similar methods to this system. It is important to mention some drawbacks of other papers that differ this paper from. Which are; slow image search in e-commerce, lack of retrieving identical or desired results in many cases, retrieving single examples when a product is searched for and inaccurate. This study has endeavored to overcome these problems or at least mitigate those in (Hu, Yin, Han, & Yu, 2014) and (Cai et al., 2017). In this paper, all those problems have been tried to be solved with an improved speed. In result, users' satisfaction can be achieved meanwhile users meet their searched products easily. As well as, this approach leads in showing identical items in the database and suggest them to the users. Explained in the following features:

1. Search speed and efficient: This paper has an active retrieval scheme. Since the image database has a large number of images that the images retrieval need at the same time. Search accuracy: In this paper, the process of image retrieval in the database is Produced and shown exactly.
2. Improved data security and accuracy.

#### 4.0 Conclusion and Future Research

A Content Based Image Retrieval Model For E-Commerce has been proposed in this paper. The model has been implemented in a Laravel PHP Framework environment and MYSQL database. The developed was accessed and it produced a better experience to the users than when using searching by text. Future Research is directed on models that can integrate all E-Commerce sites and allow user to search for commodities simultaneously.

#### Acknowledgement

The first author wishes to thank Prof. Wilson Cheruiyot and Dr. Calvins Otieno (Second and Third authors respectively) for all their inputs in designing the model and in the writing of the Masters thesis as a whole. A special thanks to the VU Vereniging for its support through the legacy of the late Ms. Grietje Wille.

#### 5.0 References

- [1] Ashraf, R., Bashir, K., Irtaza, A., & Mahmood, M. T. (2015). Content Based Image Retrieval Using Embedded Neural Networks with Bandletized Regions. *Entropy*, 17(6), 3552–3580. <https://doi.org/10.3390/e17063552>
- [2] Bengio, Y. (2011). Deep Learning of Representations for Unsupervised and Transfer Learning. In *Proceedings of the 2011 International Conference on Unsupervised and Transfer Learning Workshop - Volume 27* (pp. 17–37). Washington, USA: JMLR.org. Retrieved from <http://dl.acm.org/citation.cfm?id=3045796.3045800>
- [3] Cai, Z., Gao, W., Yu, Z., Huang, J., & Cai, Z. (2017). Feature extraction with triplet convolutional neural network for content-based image retrieval. In *2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA)* (pp. 337–342). <https://doi.org/10.1109/ICIEA.2017.8282867>
- [4] Hu, Y., Yin, H., Han, D., & Yu, F. (2014). The Application of Similar Image Retrieval in Electronic Commerce [Research article]. <https://doi.org/10.1155/2014/579401>
- [5] J. Ohene-Djan. (2008). *Electronic commerce*.
- [6] Kore, V. S., Tidke, B. A., & Chandre, P. (2015). Survey of Image Retrieval Techniques and Algorithms for Image-rich Information Networks. *International Journal of Computer Applications*, 112(6), 39–42.