" SKETCH BASED IMAGE RETRIEVAL USING CNN"

Muhammad Omer, Sana Mohib, Shahira Malik and Dr. Maria Waqas

Computer & Info System Engineering
NED University of Engineering and Technology
Karachi, Pakistan
(omerhameed98, malikshahira,Sanamohib80) @yahoo.com, mariaw@neduet.edu.pk

Abstract:-

Educational games are a vital source of learning for children in the modern education system. They induce the motivation for learning in children by their interactive manner and also encourage self learning in young children. With the advancement of web application interfaces, the application of sketches has increased to quite great extents where we are using sketches for interaction. These kind of web applications are also used in educational gaming. Hence this motivated us to develop such a web application known as "The sketch based image retrieval system" for educational gaming purposes. We propose such a system which will take the input in the form of sketch from the user of the application and then by using machine learning algorithm find and display the natural image of the sketch drawn or uploaded. Our application will also classify the sketch taken as input.

Keywords: Sketch-based image retrieval, Deep learning, Image synthesis, CNN.

I. Introduction:-

To provide a more interacting and engaging environment in order to encourage the self learning in the young children. We have developed a web based application that aims to provide educational gaming facility and make the user capable enough to classify and recognize objects.

The sketch based image retrieval using sketches can be viable and basic in our everyday life, for example, Medical analysis, advanced library, web crawlers, wrongdoing anticipation, photograph sharing locales, topographical data, instructive gaming, picture search and detecting remote frameworks.

II. Methodology

The user is allowed to give input in one of the two ways, Either he/she can draw the sketch on the sketch pad present in our application or he/she can upload a sketch which is already drawn. The system then process the sketch and displays the top matching results of natural images Elaborating the retrieval of the natural image, the uploaded or drawn sketch is first available in base 64 format which is then converted in png format in backend .The png image is passed as a query against our dataset of sketch and natural images and the system extracts the features and retrieves the corresponding natural image in json array and that array is displayed as an output in front end. In this way the outcome will be displayed on the screen of the user in the form of closest natural image of the sketch given as an input to the system. In this particular section we will reference the work related to our system of sketch based image retrieval, search and also recognition. We will also present the contribution made by our project in a very summarized manner.

Computer vision is a very broad category and under it comes the problems which are related to the recognition of objects. In the recent times the capability of guessing and classifying the sketches which are hand drawn are of great interest. A system called the "MindFinder" which is discussed and described in [Cao et al. 2010] and [Wang et al. 2010], does the operation of image search based the rough doodle which is drawn and given as input to the system. Another system for image search is described in [Sun et al. 2013] lets the user of the system to give the information about the color and edge in a sketch. Annotated sketches in the form of blue print for the formation of composited scenes is used in [Chen et al. 2009] by the help of the pictures which are present online. In [Eitz et al. 2012b], a method is presented which returns objects which are 3 dimensional based on sketches in the form of queries. A study is described in which results are presented about how the sketches are sketched by artists in [Cole et al. 2008].

A tool for the retrieval of images of sketch based input which can be used on the mobile device as it has very minor consumption of memory is presented in a very detailed manner in [Tseng et al. 2012]. In gaming context sketches are considered as described in [Ribeiro and Igarashi 2012].

Primarily our project of "Sketch based image retrieval" is based on [Eitz et al. 2012a], as described above it lets the user of the system to roughly sketch an object an as a result provides guesses of the category of the object in a ranked fashion.

A. Project Flow

The diagram presented below defines the work flow of our system in a very concise manner that how the front end which is Angular and backend which is Flask are working and in what flow the project is operating and displaying the required output.

The user first draws the sketch or uploads it though the front end then the encoding and decoding of the image is performed as discussed in the Methodology section above and then by extracting the features of the sketch given as input, the top four matches are displayed as an outcome on the screen of the user.

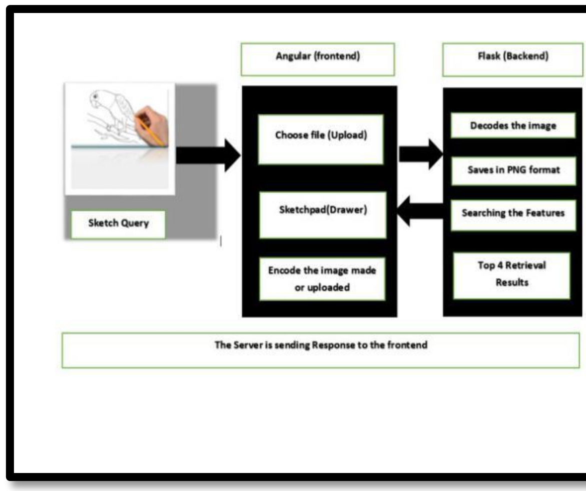


Figure 1. Hand Drawn Image

B. Use Case Diagram

The functionality of the sketch based image retrieval system can also be demonstrated with the help of a use case diagram

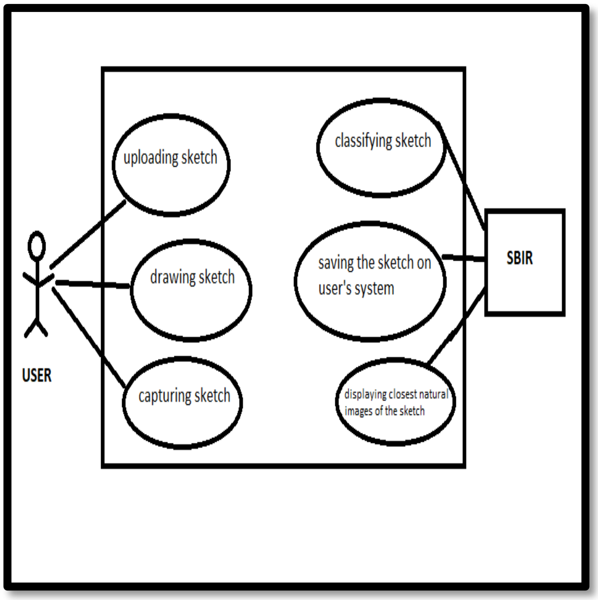


Fig 2.Use Case Diagram

III. Creating The Sketchy Database

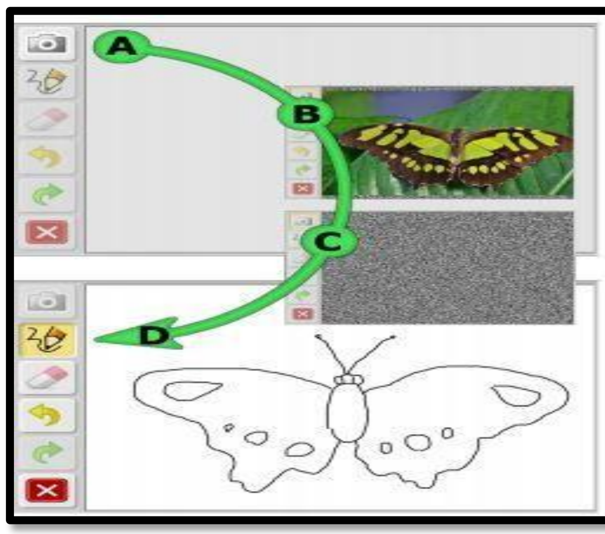
We have effectively prepared a model which traverses 125 classifications and comprises of 12,500 photographs and 75,471 human sketches.

A. Dataset Gathering

We have mainly used the database called "Sketchy". It is a very vast collection of sketch and natural photo pairs. This dataset basically consists of sketches which are sampled from one hundred and twenty five categories and contains seventy five thousand and four hundred seventy one sketched of twelve thousand and 500 objects. In the sketchy dataset there is a very fine-grained association between sketches and natural photographs and this dataset can be used for the training of cross domain convolutional networks, in which there is a common space of features and natural photographs and sketches are embedded in that space. This dataset clearly outperforms the deep features as well as the handcrafted ones which have been trained for classification of sketch.

We utilize our information base as a benchmark for fine-grained recovery and show that our educated portrayal fundamentally beats both handmade highlights just as profound highlights prepared for sketch or photograph order. Past picture recovery, we accept the Sketchy information base opens up new open doors for sketch and picture comprehension and amalgamation.

Figure 3: A range of "sketchability" for three ImageNet classes: pony, apple, and bunny.



igure4: Sketch arrangement interface.(a) Pressing a catch reveals (b) a photo for 2 seconds followed by (c) a clatter spread for one second. (d) The part by then uses pencil, eraser, and fix instruments to make their sketch.

B. Preparation of data and Pre-training:

a) Pre-training:

firstly train each sub network for sketch and image characterization. The organizations freely learn loads fitting for every area with no underlying imperatives for regular installing. We start with AlexNet or GoogLeNet prepared on

ImageNet. The sketch network is calibrated to perceive the 250 classes .We partition the dataset into a train/test split of 18k/2k and after tweaking accomplish 77.29% precision with AlexNet and 80.85% exactness with GoogLeNet.

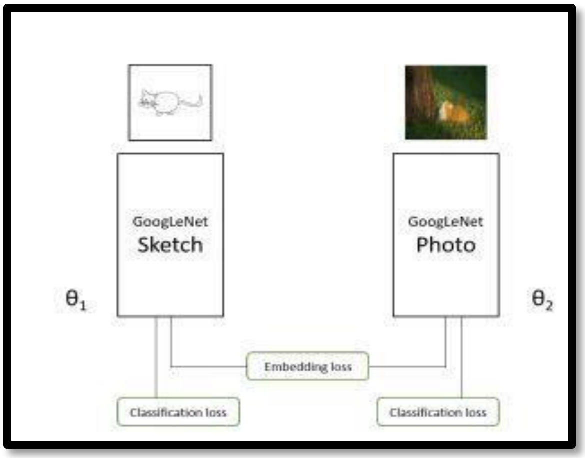


Figure 5: GoogleNet Sketch & Photo

Network is trained using Caffe. Its gives a total toolbox to preparing, testing, tweaking, and sending models, with very much recorded models for these undertakings. Thusly, it's an ideal beginning stage for scientists and different designers hoping to hop into cutting edge AI. Simultaneously, it's possible the quickest accessible usage of these calculations, making it promptly helpful for modern organization.

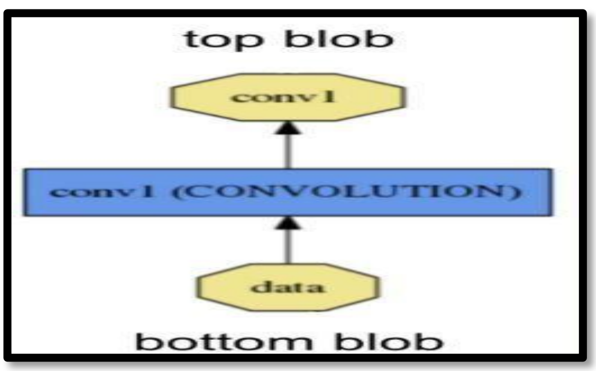


Figure 6: Caffe Blob Diagram

Flask the way toward planning a web application less difficult. lets us focus around what the clients are mentioning and what kind of reaction to give back. HTTP is the show for destinations. The web uses it to relate and talk with PCs and laborers. Let me give you an instance of how you use it reliably. Right when you type the name of a site in the area bar of your program and you hit enter. What happens is that a HTTP request has been sent to a laborer. We will form code that will manage the laborer side planning. Our code will get requests. It will understand what those requesting are overseeing and what they are asking. It will in like manner comprehend what response to send to the customer. To do this we will use Flask.

b) Fine-grained training data:-

The Sketchy data base gives fine-grained correspondence among draws & photos that can be used as specific sets for setting up our cross-space CNNs. We hold out 10% of the data for testing. The primary 'jittering' we use is reflecting. Positive sketch-photo sets are pondered together the grounds that we would lean toward not to be reflect invariant. In the wake of mirroring, the 90% of data used for planning gives in excess of 100,000 positive sets (22,500 pictures with in any occasion 5 draws each).

c) Sketch Normalization:-

We reliably focus layouts from the data base with the objective that the academic depiction isn't fragile to the by and large zone and size of a sketch. This normalization makes our benchmark harder in light of the fact that it breaks the spatial correspondence among depicts and photos. However, in a down to earth sketch-based picture recuperation circumstance we expect the customer should be invariant to region and scale and rather needs to facilitate stance, & diverse attributes. We release the balanced sketch-photo data, notwithstanding, as it would be more appropriate planning data for a couple of circumstances (for instance making sense of how to convey photos like portrayals).

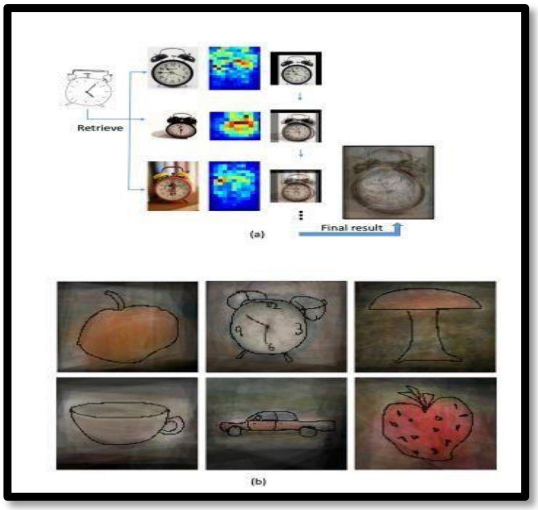


Figure 7: In (a) we adjust top recovery results (b) normal pictures.

IV. RESULT AND DISCUSSION

The framework represents to a novice sketcher running an application, the sketcher deliberately draws various articles by adding gradual line strokes to their current outlines, and watches changes in the brought outcomes back. This attracting strategy encourages sketchers to refine their inquiries, additionally to comprehend which strokes are significant for recovering wanted photograph pictures.

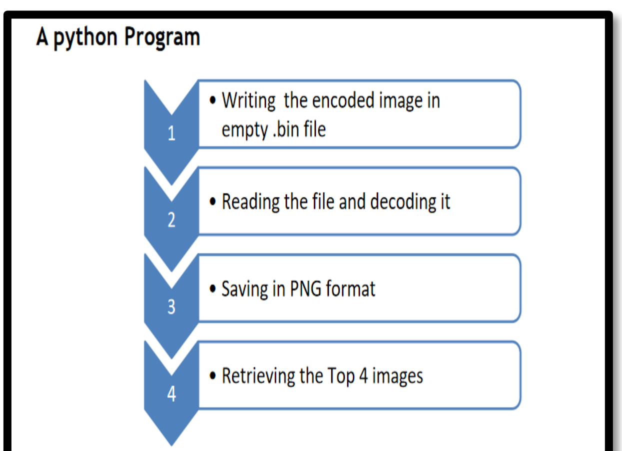


Figure 8: Conversion of user implemented image

When the user draw a bad hand drawn image or choose image from the PC, the image is encoded in base64 format so the backend is not able to draw the image .What we had first we create a .bin files and write the encoded bytes in it. After this process ,we read those bytes and store in variable. Then we create another empty file with a PNG format and decode those bytes which were read earlier and writes in PNG file which is just created

A. Limitations:-

Some of the challenges faced in the implementation of the sketch based image retrieval system are stated below.

This system requires comparison across two domains which are natural photographs and sketches. It is a challenge as understanding them individually is very hard as they contain distinct appearances.

Another challenge faced is that sketch-to-image comparison is difficult because the end users can not draw the sketch perfectly. The users only draw the salient structures and objects very poorly so there is a distortion in the shapes and the scales.

Integrating the sketch pad for directly drawing the sketch through our application was also not less than a challenge for us

.

B. Beneficiaries of the project:-

The Image search could be performed potentially by only feeding a rough sketch of the desired object which is in the mind of the user of the application to the system. In the area of education young children could learn to draw objects in a computer game that automatically evaluates the category of the sketched object.

V. Acknowledgments

The "Badly Drawn Bunnies" caption was enlivened by an article by Rebecca Boyle talking about "How Do Humans Sketch Objects?" [Eitz et al. 2012a].

Reference:-

J Canny. 1986. A Computational Approach to Edge Detection. *IEEE Anal. Mach. Intell.* 8, 6 (June 1986), 679–698. DOI: https://doi.org/1 1986.4767851

Xiaochun Cao, Hua Zhang, Si Liu, Xiaojie Guo, and Liang Lin. 2013. Symmetry-Aware Flip Invariant Sketch Histogram Shape Descript *International Conference on Computer Vision (ICCV).*

Bryan Catanzaro. 2013. kmeans. https://github.com/bryancatanzaro/k

N. Dalal and B. Triggs. 2005. Histograms of oriented gradients for huma *2005 IEEE Computer Society Conference on Computer Vision and Patt (CVPR'05)*, Vol. 1. 886–893 vol. 1. DOI: https://doi.org/10.1109/CVPI

Mathias Eitz, James Hays, and Marc Alexa. 2012. How Do Humans S *ACM Trans. Graph. (Proc. SIGGRAPH)* 31, 4 (2012), 44:1–44:10.

Mathias Eitz, Kristian Hildebrand, Tamy Boubekeur, and Marc Alexa Based Image Retrieval: Benchmark and Bag-of-Features Descriptors *and Computer Graphics* 17, 11 (2011), 1624–1636. DOI: https://doi.org/ 2010.266

Mathias Eitz, Ronald Richter, Tamy Boubekeur, Kristian Hildebrand, a 2012. Sketch-based Shape Retrieval. *ACM Trans. Graph.* 31, 4, Article 10 pages. DOI: https://doi.org/10.1145/2185520.2185527

Mathias Eitz, Ronald Richter, Kristian Hildebrand, Tamy Boubekeur, a 2011. Photosketcher: interactive sketch-based image synthesis. *I Graphics and Applications* (2011).

Rui Hu, Mark Barnard, and John Collomosse. 2010. Gradient field sketch based retrieval and localization. In *Image Processing (ICIP), International Conference on.* IEEE, 1025–1028. http://ieeexplore.iee all.jsp?arnumber=5649331

Rui Hu and John Collomosse. 2013. A Performance Evaluation of Gradi Descriptor for Sketch Based Image Retrieval. *Comput. Vis. Image* (July 2013), 790–806. DOI: https://doi.org/10.1016/j.cviu.2013.02.005

Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonath Girshick, Sergio Guadarrama, and Trevor Darrell. 2014. Caffe: Architecture for Fast Feature Embedding. *arXiv preprint arXiv:140*

Qiang Li, Yahong Han, and Jianwu Dang. 2015. Sketch4Image: a novel sketch-based image retrieval based on product quantization wit uals. *Multimedia Tools and Applications* (2015). DOI: https://d s11042-015-2645-y

Yi Li, Tim Hospedales, Yi-Zhe Song, and Shaogang Gong. 2014. Intra-c based image retrieval by matching deformable part models. In *Pro British Machine Vision Conference.* BMVA Press. DOI: https://doi.or 115