

Enriching Image Retrieval System Through CNN for Sketches And Images



Shreya Kukdeja, A. G. Phakatkar

Abstract: Image retrieval is being a one of the widely research areas in the current paradigm of the software industry. This is mainly due to as a proverb "image can speak many words", which is because of its diversity of the contained objects and their pattern. Many search engines including Google, is providing the option of searching images by giving query image, most of the time this searching is done by the averaging the features of the images which yields considerable low precision. Convolution neural network and Recurrent neural network are widely used mechanisms to handle the image processing techniques in image retrieval. On using of CNN and RNN many systems are yielding the low accuracy because of the features that they are considering. And again, only finger counting systems are dealing with the image retrieval mechanism for the input of the object rather than the whole image. So as a tiny step towards this, this research article proposes a model of image retrieval using the input as image sketch and images using the histogram features and Region of interest based on the position, volume, color and orientation through the interactive CNN model. The image search is performed using the CNN through K means Clustering and Haar wavelets.

Index Terms: CNN, Haar Wavelet, Histogram features, Region of Interest.

I. INTRODUCTION

Analysis of images has been a difficult task for a computer as it cannot extract the semantics and the objects in the image efficiently. For a human being, the recognition is near instantaneous, the moment we see an image we are able to recognize what it contains and the various objects in the image, which is almost impossible for a computer that calculates in 1's and 0's, any aspect of the image would be the same for the computer, it will treat it as data without any understanding of what it contains.

This is highly difficult in a scenario where an image is required and needs to be retrieved by the system. Due to the fact that the meaning of any image is very subjective and changes according to a particular individual. As every individual understands the same image differently, it becomes extremely difficult to retrieve the image as per the given limitations and features.

Due to large advancements in the technology and the proliferation of images everywhere, there has been an unprecedented increase in the number of images. Every Smartphone nowadays has a camera and the users can click any number of photographs or images in a day. This is adding to the ever-increasing number of images in the digital domain. The digital image is different from the images taken with on photographic film. The Digital images are captured with the help of a light sensor that captures the light entering the sensor at that instant.

The Digital image is composed of numerous pixels, pixels are nothing but small dots composed of colors arranged in a grid-like structure. The digital images are essentially a scan of the light that passes through the sensor and the image is formed with the help of pixels that capture and emulate the exact same light and color levels from the object that was photographed. Most of the digital images are stored in a compressed format, this is due to that fact that some of the values of pixels can be very similar to one another and can help reduce the storage utilized.

The image retrieval process cannot be done as the way humans determine the objects in an image. It is very different and complicated for the computer, therefore, the computer needs to extract the relevant information about the image, such as shape, texture, and color, these are known as the image features. these features give an understanding of the objects contained in the image. But, some of the images do not constitute the majority of their area with the object of interest and this needs to be addressed by cropping the object out of the image. This serves a dual purpose, firstly, the object in the cropped or segmented image now acquires the maximum area in the image, and secondly, the reduced size is a boon for image retrieval as fewer resources would be required to process the image.

There are various techniques designed for the purpose of image retrieval as it is one of the most influential mechanisms in the field of computers. The first one is the Text-Based Image Retrieval technique; this technique utilizes the metadata present in the images to help retrieve the images through the keywords. This technique is not the most accurate ones from the lot. The second one is a Content-Based Image Retrieval technique, as the name suggests, this technique utilizes the content in the images such as the textures, shapes colors etc. this method is based on the indexing done by a computer vision system and is highly helpful in segregating large databases. The third one is a Sketch-Based Image Retrieval technique; this technique requires the user to make a sketch of the image or the object in an image that is supposed to be retrieved.

Manuscript published on November 30, 2019.

* Correspondence Author

Shreya Kukdeja*, Department of Computer Engineering, Pune Institute of Computer Technology, Pune, India.

Prof. A. G. Phakatkar, Department of Computer Engineering, Pune Institute of Computer Technology, Pune, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

And lastly, Semantic Based Image Retrieval which utilizes the semantics in the image to retrieve the relevant image. This technique is very thorough as it analyses the whole image and assigns semantic labels to the image that can help retrieve the image efficiently.

RNN or Recurrent Neural Networks are the type of networks that are an offshoot of the main branch of Neural Networks, it is quite powerful and is generally used for the purpose of predicting the next word in a sentence. This use case is possible entirely due to a unique feature in this type of neural network, which is not shared by any other neural networks, hidden layer. This hidden layer or state is responsible for remembering the previous information therefore, this allows the RNN to predict the next word or item.

Convolutional Neural Networks are computational algorithms derived from the set of Artificial Neural Networks. These networks are defined as a set of algorithms designed to emulate the inner workings of the human brain. As we all know the human brain is an organic computational workhorse. The human brain has enabled us to achieve some impressive feats, humans as a species have conquered most of the elements present on our planet. Therefore, to harness the power of that process has inspired the researchers to develop an algorithm based on the human brain.

The Artificial Neural Network shares the same basic unit of computation from the human brain; the neuron. The neuron is the most basic component and the smallest computational unit in the brain, it is capable of firing an electrical impulse when a stimulus that is received is more than the threshold values. The neuron is interconnected with a large number of neurons that work in tandem to provide a neuron network that is capable of learning.

The Convolutional Neural Network is derivative of the Artificial Neural Network, which has the same layered system with numerous neurons arranged together, with only one key difference between the two, the Convolutional Neural Networks have a convolution in their algorithm as the name suggests, but the Artificial Neural network doesn't. As the neurons have a specific threshold that is supposed to be exceeded, if not then the neuron won't activate and hence won't fire the signal. The Convolutional Neural Networks work in various different layers of neurons, these layers are interconnected to each other. The connections are made in such a way that the output of one layer is given as input to the next layer. This ensures that the information or the data is filtered through the network, in a layer by layer fashion. This filtering is what provides the unprecedented strength for the classification purposes, this makes the Convolutional networks really apt for classification tasks. This is the reason why Convolutional Neural Networks are heavily utilized for Image processing tasks. The convolution in the neural network ensures that the images are classified correctly every time it is passed through the layers of the Convolutional network, which act as filters. The Convolutional networks, therefore, provide extensive robustness coupled with very high accuracy and efficiency.

This research paper dedicates section 2 for analysis of past work as literature survey, section 3 deeply elaborates the proposed technique and whereas section 4 evaluates the performance of the system and finally section 5 concludes the paper with traces of future enhancement.

II. LITERATURE SURVEY

Q. Zhang [1] states that most of the images being circulated on the internet are in compressed form, as this allows for conservation of resources such as bandwidth and storage. As CBIR (Content Based Image Retrieval) requires the images to be stored in binary form to assist in image retrieval through its features. Therefore, the authors have implemented a deep-network based image coding scheme to achieve greater compression and better retrieval. The authors in this paper have not measured the impact of the reduction of the number of feature and the alleviating the precision of the features. The compression applied in [1] reduces the image quality and its subsequent retrieval, which is not experienced by the image in the methodology proposed in this paper. E. Pinho explains that there has been a lot of advancements in the area of medical information retrieval, due to increased interest amongst the researchers. This is due to the fact that there has been an increase in the amount of digital imaging being utilized in the medical field. This has pressed for a system that can help integrate these into archives. Therefore, the authors have deployed a system that achieves multimodal information retrieval for the medical images [2]. The system is also capable of classification of the medical images. The only drawback this paper includes is that the authors have not performed a complete analysis of the solution in terms of Retrieval. D. Saravanan elaborates on the extensive usage of images in communication nowadays. It has been extensively utilized to convey a greater understanding through visual means in various fields, such as Healthcare, Media and Education. As the image retrieval is an important concept in these types of scenarios, the authors have proposed a methodology for the retrieval of images based on features, such as shape, color, and texture, in addition to the implementation of Data Mining Techniques. the main drawback of the proposed technique is the increased time complexity that is observed. The content-based retrieval in [3] is very inferior to the Convolutional Neural Network used in this paper, which accelerates the retrieval process. G. Schaefer introduces the irregularities that are observed when performing image retrieval with the help of Content-Based Image Retrieval (CBIR). The Content-Based Image Retrieval specifically works with the pixel data in the images, as most of the images that are circulated on the internet are compressed images, therefore, the author in this paper has presented an innovative technique for the retrieval of compressed images with a modified CBIR approach[4]. The proposed methodology is computationally extensive, in comparison with the traditional techniques.

J. Lei [5] explores the techniques for image retrieval based on sketches. The author states that there are various implementations of Sketch-Based retrieval techniques, but the systems primarily utilize handmade sketches for retrieval purposes. Therefore, the authors have presented a technique for the Sketch-Based Image Retrieval (SBIR) with the addition of a Deep Learning model that is trained to identify distinguishing features. the main drawback of this technique is that the deep learning algorithm based on neural networks require a huge number of images to train.

The proposed system in [5] utilizes a pre-trained VGG19 framework for the retrieval purposes, unlike our technique that trains the system with a training dataset in real-time. L. Xie [6] states that a better image retrieval technique can be utilized for effective and efficient retrieval of images. The authors propose that this could be done with the help of amalgamating image classification with various other retrieval algorithms to form a methodology named as ONE (Online Nearest-Neighbor Estimation). The authors state that the inclusion of various retrieval algorithms makes sure that the system will be highly accurate. The only drawback of this technique is that the inclusion of a number of algorithms has made the proposed technique to be highly computationally expensive.

A. Razavian explains that for the purpose of efficient visual instance retrieval, there is a need for a pipeline-based Convolutional Neural Network for the image representation. This technique puts more weight on the local features in the image which is inclined towards the geometric invariance present in the images [7]. This is the reason for utilizing a multiscale scheme that utilizes the Convolutional Neural Network efficiently. But the authors of this paper haven't investigated the use of domain adaptation for increasing the performance of the system. The system outlined in [7] has a greater space and time complexity in comparison with the model presented in this paper, due to the fact that in [7] each image is uploaded with in 4 different scales for the image retrieval. Y. Kalantidis elaborates on the proposed system for the utilization of deep Convolutional features in a cross-dimensional weighting for the purpose of image retrieval. The system has been utilized with a heavy emphasis on the Convolutional network, that uses non-parametric weighting schemes which are elementary to increase performance. The authors have tried to predict the performance of the system to gauge its efficiency. The one feature this technique lacks is the learning of the weights by the system to facilitate easier and faster retrieval. [8]

Y. Uchida introduces a system that utilizes binary features in an image through the application of a Fisher Vector Representation to retrieve an image. The closed form Fischer vectors of the binary features for the estimation of the Bernoulli mixture [9]. The researchers also included a fast approximation technique to make the calculations of the Fischer Vectors faster. The authors have implemented this technique and confirmed the efficient of the Fischer Vectors of the binary features. One drawback in this technique was that the authors did not utilize the Fischer Vector for the classification of the images. The improvement of accuracy in [9] is not as significant as the one outline in this paper as Convolutional Neural Networks are highly accurate and considerably faster than the Fischer vectors. A. Alzu'bi explores the technique of utilizing the Convolutional Neural Networks for the purpose of performing Content-Based Image Retrieval tasks with a compact and bilinear feature. The authors have utilized this technique to retrieve images without the semantic, meta-data but the Convolutional Neural Networks for parallelly extracting the image features [10]. The architectures proposed by the authors need to be calibrated to ensure that the feature extraction is on point and poor calibration would lead to low dimensional representations. L. Zheng [11] states that the usage of Convolutional Neural Networks for the purpose of image retrieval has been gaining interest lately. The authors have proposed a technique for the Convolutional Neural Networks

based image retrieval techniques to improve the retrieval performance of those systems. The researchers improve the performance by adding average polling to the activation maps and by utilizing larger images which in turn increases accuracy. The discriminative power of the system remains the same and is the only drawback in this technique. [11] Discusses the Feature Extraction accuracy of CNNs which is achieved by providing the highest resolution of the image, this is a problem as it increases the space complexity as images with higher resolution take up more space, which is not a problem with the methodology proposed in this paper.

E. Mohedano explains that to improve the image retrieval of the system, there needs to be a strategy developed. The authors in this paper have presented a system for efficient retrieval of images by introducing Bag of Words and Convolutional neural networks for performing the search [12]. The proposed system is highly aggressive and is also very extensible. The technique proposed by the authors has a very large amount of time complexity, which has been outweighed by the accuracy of the system. Y. Li proposes a technique for the retrieval of images based on Convolutional neural networks in conjunction with the bag of words technique. The presented system has been named as a Multi-layer Order less Fusion [13]. The Convolutional neural networks in this technique are utilized to achieve a multi-scale and multi-category fusion in the layers. The fusion is carried out by the Bag of Words framework for increasing the performance of the system as a whole. The authors have not utilized the reciprocal and region proposal for the post processing in this paper.

T. Yu introduces the concept of the Fuzzy Objects Matching, which detects the approximate locations in the image for discerning the object of interest in the image. To achieve efficient and effective results, the authors factorized the feature matrix with a set of fuzzy objects. This is done to ensure that the performance and efficiency of the system are not affected. The proposed algorithm has been deployed in a real-world scenario with impressive results, the proposed technique outperformed the traditional technique by a large margin. [14] H. Azizpour explores the realm of Convolutional Neural Networks that are trained on the Image Net. These images are quickly becoming the standard for image retrieval. Nowadays, Convolutional Neural Networks have been utilized extensively for the purpose of image retrieval. The authors in this paper study various implementations of the Convolutional Neural networks for the purpose of image retrieval which has become an effective representation of the purpose of visual recognition. The researchers successfully correlated the facts with the help of extensively studying 16 visual recognition tasks.

III. PROPOSED METHODOLOGY

The proposed system for image retrieval is depicted in figure 1. The steps that involve in the retrieval process are elaborated below.

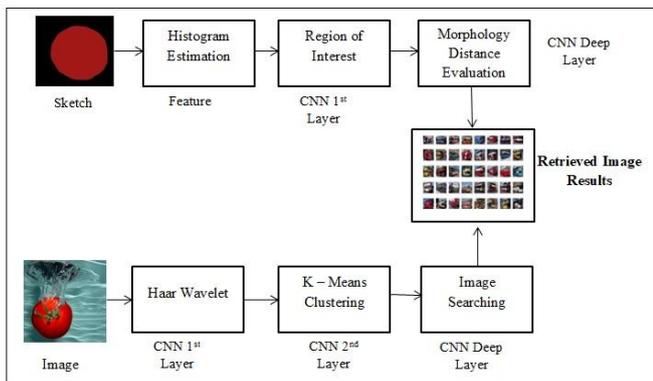


Figure 1: Overview of the Proposed System

1. Retrieval Through image Sketch- This is the first step in the proposed system that is responsible for storing the images from the dataset in the root of the device which is further segmented into subfolders, which contain the respective images. As the proposed model receives the root folder of the dataset, then each image is read in the object of BufferedImage to extract the important features from them.

[A] Feature Extraction –This step creates the histogram for the images. A Histogram is an essential concept that segregated the images into segments based on the intensity of the pixels. This histogram evaluation can be depicted clearly in algorithm 1 given below.

The obtained histogram array is applied on the image pixels to get the histogram image for all the images in the input image dataset, then it is stored in the same location as of the images. The histogram array is written in a file which is subsequently used to search or query the sketched object provided by the user.

ALGORITHM 1: Histogram Estimation

```

// Input: DB image DIMG
// Output: Histogram Array HG [ ]
// Histogram Estimation(DIMG)
1: Start
2: HG=NULL
3: for i = 0 to size of Width of DIMG
4: for j=0 to size of Height of DIMG
5: PSIGN = DIMG(i,j) RGB
6: SRGB[3]=NULL
7: SRGB[0]→R= ( PSIGN >> 16 & HD)/255
8: SRGB[1]→G= ( PSIGN >> 8 & HD)/255
9: SRGB[2]→B= ( PSIGN >> 0 & HD )/255
10: normalize SRGB
11: MIN=getMIN(SRGB)
12: HG[min]= HG[min]+ 1.0
13: End for
14: End for
15: for i = 0 to size of HG
16: HG[i] = HG[i] / (Width of DIMG * Height of DIMG)
17: End for
18: return HG
    
```

[B] Region of Interest Estimation - CNN First Layer-A BufferedImage object is formed in this step of the proposed system based on the sketch provided by the user for the purpose of image retrieval. Then this sketch object is matched with all of the stored histogram arrays of the dataset images. Firstly, the region of interest for the drawn sketch is

evaluated. In this process, the sketch image is assumed to have a black background. All the non-black pixels are counted as N. Then a mass center of the drawn sketch is estimated using equation 1 and 2.

$$Cx = \frac{\sum_{i=0}^w i}{N} \quad \text{---(1)}$$

$$Cy = \frac{\sum_{j=0}^h j}{N} \quad \text{---(2)}$$

where,

C_x - Mass Center x

C_y - Mass Center y

w,h- Width and height of the image

Then the measured mass center is used to identify the elongation of the region of interest, this is done by multiplying the angular distances. The normalized elongations provide the region of interest that is used to search the stored histograms of the dataset images.

[C] Morphology Distance Evaluation - CNN Deep Layer-A Euclidean distance is evaluated for the four parameters which are subsequently used to identify the objects in the image. The four parameters utilized for searching purposes are position, volume, color, and orientation. The Euclidean distance is being evaluated using the equation no 2.

$$ED = \sqrt{(x1 - x2)^2 + (y1 - y2)^2} \text{----- (3)}$$

Where,

E_D - Euclidean distance

x1,x2,y1,and y2 are the pixel positions

The Evaluated distances are further used to find the product function combined with the weights of the decided parameters results in the morphological distance. The morphological Distance of the object is calculated as depicted in equation 4.

$$\text{Dist} = E_D * \text{weight}[0] + E_D * \text{weight}[1] + E_D * \text{weight}[2] + E_D * \text{weight}[3] \text{---(4)}$$

Where,

Dist- Object Distance

weight- Array of the parameters

The images with the distance less than that of the threshold are considered as the output for the given input sketch object.

2. Retrieval Through image - In this part of the proposed model initially all the images are stored in a single root folder and that is fed to the system to learn by indexing the images.

[A] Haar Wavelet - CNN First Layer- As the system receives the dataset images, it starts indexing them by resizing the images into a size of 256 X 256. Then the resized images are sent to the Haar wavelet to estimate the index of the RGB color channel. The obtained Haar indices are added into a list to cluster them efficiently.

[B] K- Means clustering-CNN Second Layer - Here in this step of the proposed model the indices of the dataset images are clustered based on the following steps.

[i] Distance Estimation - Here the distances of the Haar index are estimated using equation 5.

$$E_D = (\sqrt{\sum R} + \sqrt{\sum G} + \sqrt{\sum B}) / 3 \text{---(5)}$$



[ii] *Centroids Estimation* - Here the centroids are estimated based on the biggest distances in the range in comparison to the previous distance.

[iii] *Cluster formation* - The clusters are formed based on the maximum iteration of 200 for the assigned centroids.

Once the images are clustered based on the distances of the Haar features of the images, then these clusters are labeled with the distance of the centroid image to be stored in an object of the indexer.

[C] *Image Searching - CNN Deep Layer*- Once the query image is fed to the system, then a distance is evaluated for the query image based on the Haar feature. Then this distance is matched with the distances of the centroids of the cluster. The maximum matched distance of a centroid cluster obtained is then provided as the output.

IV. RESULT AND DISCUSSIONS

The proposed methodology of the Image retrieval system through input sketches and images is developed in Windows machine. Which is equipped with a Core i5 Pentium processor along with 6GB primary memory. Proposed system uses the Java programming language for the development purpose by using Netbeans 8.0 as the Standard IDE. Proposed model uses the generic image dataset obtained through the UCI public dataset repository. The model is put under the hammer to prove its authenticity by conducting some experiments as described below.

Dataset collection - The proposed model is developed for the image retrieval system, Where the images are retrieved for both the input object through the form of a sketch or by an input image. The model is developed for the all kinds of images in generic form, So the generic datasets are collected through the UCI dataset repository and store in a folder on the hard drive. The proposed model of image retrieval is developed to retrieve the matched images for the input sketch and images. Hence, retrieval step is classified in two parts and they are elaborated as below.

For the Evaluation purposed proposed model considers the precision and recall as the standard measuring parameters. Precision always tells us the performance of the system. Precision can be described as the number of the relevant images are retrieved from the database for the number of the relevant and Irrelevant images are retrieved from the database.

Whereas Recalls are generally described as the relevant images retrieved over the relevant images identified. And this can be more clearly stated as the ratio of the number of relevant images are retrieved to the sum relevant and not relevant images are retrieved. Precision can be more effectively explained as below.

A = The number of relevant images extracted for the given number of query images.

B= The number of irrelevant images extracted for the given number of query images.

C = The number of relevant images is not extracted for the given number of query images.

So, precision can be given as,

$$\text{Precision} = (A / (A + B)) * 100$$

$$\text{Recall} = (A / (A + C)) * 100$$

When precision and recall are estimated the obtained results are tabulated in the table 1 and table 2 for the given input of the sketch and the images.

Table 1: Precision and Recall Data for the input Query sketch

No of Runs	Dataset Size	Classes	Relevant Images Retrieved (A)	Irrelevant Images Retrieved (B)	Relevant Images not Retrieved (C)	Precision = (A/(A+B))*100	Recall= (A/(A+C))*100
1	1000	Bus	84	3	8	96.55172414	91.30434783
2	2000	Dinosaur	45	3	3	93.75	93.75
3	3000	Flower	34	1	1	97.14285714	97.14285714
4	4000	Horse	87	3	2	96.66666667	97.75280899
5	5000	Mountain	22	0	1	100	95.65217391
6	6000	Food	97	4	7	96.03960396	93.26923077
7	7000	monument	55	3	5	94.82758621	91.66666667
8	8000	Butterfly	24	0	2	100	92.30769231
9	9000	Nature	78	1	5	98.73417722	93.97590361
10	10000	People	154	3	4	98.08917197	97.46835443



Figure 2: Precision and Recall for input Query Sketch

On comparing the data of the table 1 which yields the average precision of 97.18 % and recall of 93.2% for the input of the query sketch image. Whereas the table 2 yields the average precision of 97.5 % and Recall of 96.1% for the input of the query image.

When the precision of the proposed model for image retrieval system is compared with that of the [16], which is based on the deep feature learning mechanism. Proposed system's precision of 97.18 % is bit higher than that of the [16], which is about 96.5%.

Table 2: Precision and Recall Data for the input Query Image

Dataset Size	Classes	Relevant Images Retrieved (A)	Irrelevant Images Retrieved (B)	Relevant Images not Retrieved (C)	Precision = (A/(A+B))*100	Recall= (A/(A+C))*100
1000	Bus	102	3	7	97.14285714	93.57798165
2000	Dinosaur	117	4	4	96.69421488	96.69421488
3000	Flower	147	4	2	97.35099338	98.65771812
4000	Horse	158	5	4	96.93251534	97.5308642
5000	Mountain	98	3	3	97.02970297	97.02970297
6000	Food	87	2	2	97.75280899	97.75280899
7000	monument	115	2	4	98.29059829	96.63865546
8000	Butterfly	101	3	4	97.11538462	96.19047619
9000	Nature	88	2	4	97.77777778	95.65217391
10000	People	187	2	7	98.94179894	96.39175258

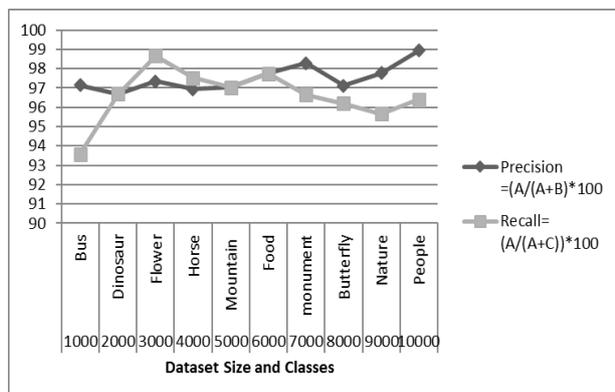


Figure 3: Precision and Recall for input Query Image

This slight increment in precision can be seen both in the retrieval process based on the input of query sketch and image. [16] mainly works on the bag of word model of the SIFT features, so this BoW model may not tag all of the objects in the image. Whereas the proposed model attained a slight increase in the percentage of the precision due to using of Histogram estimation and Haar wavelets which are extracting all the small details of the image pattern to retrieve the images in a best possible way.

V. CONCLUSION AND FUTURESCOPE

The research article mainly concentrates to attain the excellence in image retrieval scheme based on the input image object in the form of sketches and the image itself. To achieve this proposed model excellently blends the Histogram images as the Region of interest by using the deep learning mechanism of the Convolution neural network. On the other hand, proposed model deals with the image as the input, where Haar wavelet features are used to generate the image indices. These indices are used to estimate the distances to cluster them using the K-means technique which is blended with CNN. The evaluation of the results clearly shown that the proposed model attained an average precision of around 97.18 % and recall of 96.1%. Which is just slightly higher than the deep learning model of SIFT image features are stated. In the feature this mechanism of image retrieval can be boosted with the deployment of distributed paradigm to deal with the high definition images.

REFERENCES

1. Q. Zhang, D. Liu, and H. Li, "Deep Network-Based Image Coding for Simultaneous Compression and Retrieval", IEEE International Conference on Image Processing (ICIP), 2017.
2. E. Pinho and C. Costa, "Extensible Architecture for Multimodal Information Retrieval in Medical Imaging Archives", 12th International Conference on Signal-Image Technology & Internet-Based Systems, 2016.
3. D. Saravanan, S. Vijaya Lakshmi and Dennis Joseph, "Image Retrieval by image feature using Data Mining Technique", International Conference on Inventive Systems and Control, 2017.
4. Gerald Schaefer, "Fast Compressed Domain JPEG Image Retrieval", International Conference on Vision, Image and Signal Processing, 2017.
5. J. Lei, K. Zheng, H. Zhang, X. Cao, N. Ling, and Y. Hou, "Sketch-Based Image Retrieval Via Image-Aided Cross-Domain Learning", IEEE International Conference on Image Processing (ICIP), 2017.
6. L. Xie, R. Hong, B. Zhang, and Q. Tian, "Image classification and retrieval are one", in ACM International Conference on Multimedia Retrieval, 2015.
7. A. Razavian, J. Sullivan, and S. Carlsson, "Visual instance retrieval with deep Convolutional networks", ITE Transactions on Media Technology and Applications, vol. 4, no. 3, pp. 251–258, 2016.

8. Y. Kalantidis, C. Mellina, and S. Osindero, "Cross-dimensional weighting for aggregated deep Convolutional features", in European Conference on Computer Vision, 2016.
9. Y. Uchida, S. Sakazawa, and S. Satoh, "Image retrieval with fisher vectors of binary features", ITE Transactions on Media Technology and Applications, vol. 4, no. 4, pp. 326–336, 2016.
10. A. Alzu'bi, A. Amira, and N. Ramzan, "Content-based image retrieval with compact deep Convolutional features", Neurocomputing, vol. 249, pp. 95–105, 2017.
11. L. Zheng, Y. Zhao, S. Wang, J. Wang, and Q. Tian, "Good practice in CNN feature transfer", 2016.
12. E. Mohamedano, K. McGuinness, N. E. O'Connor, A. Salvador, F. Marques, and X. Gir'oi Nieto, "Bags of local Convolutional features for scalable instance search", in ACM International Conference on Multimedia Retrieval, 2016.
13. Y. Li, X. Kong, L. Zheng, and Q. Tian, "Exploiting hierarchical activations of neural network for image retrieval", in ACM International Conference on Multimedia, 2016.
14. T. Yu, Y. Wu, S. D. Bhattacharjee, and J. Yuan, "Efficient object instance search using fuzzy objects matching", AAAI Conference on Artificial Intelligence, 2017.
15. H. Azizpour, A. Sharif Razavian, J. Sullivan, A. Maki, and S. Carlsson, "From generic to specific deep representations for visual recognition", IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015.
16. Yue Lv, Wengang Zhou, Qi Tian, Shaoyan Sun and Houqiang Li, "Retrieval Oriented Deep Feature Learning with Complementary Supervision Mining", IEEE TRANSACTIONS ON IMAGE PROCESSING, 2018.

AUTHORS PROFILE



Shreya Kukdeja Bachelor of Computer Engineering, Image Processing and Machine learning.



Prof. A. G. Phakatkar Masters of Computer Engineering, Data mining, Image processing, Text mining and Database, LMISTE.