International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 08 | Aug - 2021 ISSN: 2582-3930

RETRIEVAL OF IMAGE BASED CONTENT USING PRETRAINED DEEP CONVOLUTIONAL NEURAL NETWORK BASED FEATURES EXTRACTION TECHNIQUES.

Shalaka R. Kangale (M.E.II year), Sipna College of Engineering & Technology

Dr. Ajay A. Gurjar, Prof. Sipna College of Engineering & Technology

ABSTRACT

Multimedia content analysis is applied in different real-world computer vision applications, and digital images constitute a major part of multimedia data. In last few years, the complexity of multimedia contents, especially the images, has grown exponentially, and on daily basis, more than millions of images are uploaded at different archives such as Twitter, Facebook, and Instagram. To search for a relevant image from an archive is a challenging research problem for computer vision research community. The content based image retrieval aims to find the similar images from a large scale dataset against a query image. Generally, the similarity between the representative features of the query image and dataset images is used to rank the images for retrieval. In early days, various hand designed feature descriptors have been investigated based on the visual cues such as color, texture, shape, etc. that represent the images. However, the deep learning has emerged as a dominating alternative of hand-designed feature engineering from a decade. It learns the features automatically from the data. This paper presented the system of content based image retrieval (CBIR) using features extracted from pre-trained Deep Convolutional Neural Network. The proposed system is implemented and analyzed using MATLAB software.

Keywords: Content Based Image Retrieval, Semantic Gap, Machine Learning, Deep Learning.

INTRODUCTION

Recent technology development in different fields has made sizably voluminous digital image databases practical. Well organized database and efficient browsing, storing, and retrieval algorithms are very paramount in such systems. Image retrieval techniques were developed to avail these components. Image Retrieval was originated from Information Retrieval, which has been very active research topic since 1940s. Content Predicated Image Retrieval (CBIR) researches endeavor to automate such intricate process of retrieving images that are homogeneous to the reference image or descriptions given. The shared and stored multimedia data are growing, and to search or to retrieve a relevant image from an archive is a challenging research problem. The fundamental need of any image retrieval model is to search and arrange the images that are in a visual semantic relationship with the query given by the user [2].

Content Based Image Retrieval (CBIR) is the procedure of automatically retrieving images by the extraction of their low-level visual features, like color, texture, shape properties or any other features being derived from the image itself. The performance of a CBIR system mainly depends on these selected features. Thus, it can be said that through navigation, browsing, query-by- example etc, we can calculate the similarity between the low-level image contents which can be used for the retrieval of relevant images. The most challenging issue associated with CBIR systems is reducing the semantic gap. It is the information lost by representing an image in terms of its features i.e., from high level semantics to low level features. This gap exists between the visual information captured by the imaging device and the visual information perceived by the human vision system (HVS) and it can be reduced either by embedding domain specific knowledge or by using some machine learning technique to develop intelligent systems that can be trained to act like HVS [1] [2].

However, the improper selection of features can decrease the performance of image retrieval model. The image feature vector can be used as an input for machine learning algorithms through training and test models and it can improve the performance of CBIR [1, 2]. There has been a significant growth in machine learning research but mainly deep learning has already demonstrated its potential in large-scale visual recognition. The main reasons behind its success are the availability of large annotated data sets, and the GPUs computational power and affordability. Deep learning is a subset of machine learning which uses a hierarchical level of artificial neural networks to carry out the process of machine learning. The term Deep Neural Network (DNN) refers to describe any network that has more than three layers of non-linear information stages in its architecture and Deep Learning (DL) is a collection of algorithms for learning in Deep Neural Networks, used to model high-level abstractions in data. Thus, deep learning techniques gives a direct way to get feature representations by allowing the system (deep network) to learn complex features from raw images without using hand crafted features. The recent revolutions in computer vision and image recognition, thanks to the deep learning breakthrough, make the deep learning seems a potential bridge to this gap for retrieving images, because it has the ability to process raw data and build the internal feature representation of it through its multiple nonlinear layers of abstraction to provide eventually a high conceptual representation of the image. In other words, the deep learning has the capability to learn the image semantic representation through its training phase. Deep learning has been successfully applied to many problems e.g., computer vision and pattern recognition, computer games, robots and self-driving cars, voice recognition and generation, music composition and natural language processing [1][4][5][6]. Due the success of deep leaning, and the importance of feature extraction in CBIR systems, in this work we propose a pre-trained Deep CNN for learning feature representation in CBIR. The proposed system learns how to extract relevant features from a given images database and then applies this information in image retrieval process.

LITERATURE SURVEY

Research work presented by various authors related to CBIR and various techniques related to feature extraction described is as given below.

Ibrahim Abood, Zainab et.al. [3] proposes content based image retrieval (CBIR) using four feature extraction techniques. The four techniques are colored histogram features technique, properties features technique, gray level co-occurrence matrix (GLCM) statistical features technique and hybrid technique. From the results, it is concluded that, for the database images used in this work, the CBIR using hybrid technique is better for image retrieval because it has a higher match performance (100%) for each type of similarity measure so; it is the best one for image retrieval. Sirisha Kopparthi et.al. [4] proposed experiments which are carried out in two datasets such as UC Merced Land Use Dataset By using a pre-trained model that is trained on millions of images and is fine-tuned for the retrieval task. Heba Abdel-Nabi et.al. [6] proposed approach overcomes these difficulties with the aid of the most fast growing technology, namely Deep Learning. The experimental results demonstrate the effectiveness of the proposed scheme in terms of the number of relevant retrieved images of the query results, and the mean average precision, while keeping low computational complexity since it uses an already trained deep convolutional model called AlexNet.

M. Alrahhal et.al. [7] proposed a new CBIR system using Local Neighbor Pattern (LNP) with supervised machine learning techniques. Performance analysis shows that LNP gives better performance regarding the average recall than LBP, LDP, and LTrP. P. Kaur et.al [8] proposed a new system in which features are extracted using Gabor filtering which are further optimized using lion optimization. The proposed method is tested in terms of various parameters that show improved results are achieved using Lion optimization as compared to cuckoo search optimization.

Deepa Dubey et.al. [9] clarifies for the most part about the determination of the picture highlights like shading, surface, and edge for the substance based picture recovery framework which utilizes the intelligent hereditary calculation. Since, Two-dimensional entropy utilizes both the dark estimation of a pixel and the nearby normal dim estimation of it, and along these lines gives better results. Ying Liu et.al. [10] proposes an image database retrieval algorithm based on the framework of transfer learning and feature fusion. Last, the semantic feature extracted from the CNN is fused with traditional low-level visual feature to improve the retrieval accuracy further.

Palepu Pavani et.al. [11] proposed an enhanced relevance-feedback method to support the user query based on the representative image selection and weight ranking of the images retrieved. From these experiments, the proposed learning method has enabled users to improve their search results based on the performance of CBIR system. Yogita D. Shinde et. al. [12] presented Self Mutated Hybrid Wavelet transform (SMHWT) which is used to form by using same component transform. Results show that the proposed Self Mutated Hybrid Wavelet Transform containing Sine transform as a component gives better performance improvement across all tried variations of SMHWTs.

Rajeev Srivastava et.al. [13] proposed a fast and effective CBIR system which uses supervised learning-based image management and retrieval techniques. This work is evaluated and compared with the conventional model of the CBIR system on two benchmark databases, and it is found that the proposed work is significantly encouraging in terms of retrieval accuracy and response time for the same set of used features. Nehal M. Varma et.al. [16] proposes a mixture of hybrid features for based image retrieval in which color features are obtained using color histogram method which will ultimately give the color gradient.

Most of researcher performed CBIR using various techniques which majorly classified using machine learning with manual feature extraction and automated feature extraction on different standard benchmark dataset. In manual feature extraction, features used to extract from images is not suitable for each dataset available, so it affects the system retrieval accuracy. But in case of automated feature extraction, it extracts low level to high level features that's why it is suitable for any kind of datasets. Hence, we proposed our architecture using automated feature extraction using pre-trained model.

PROPOSED WORK

The entire working process of the presented method is shown in Fig. 1. The presented model consists a series of processes which are discussed below.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 08 | Aug - 2021 ISSN: 2582-3930

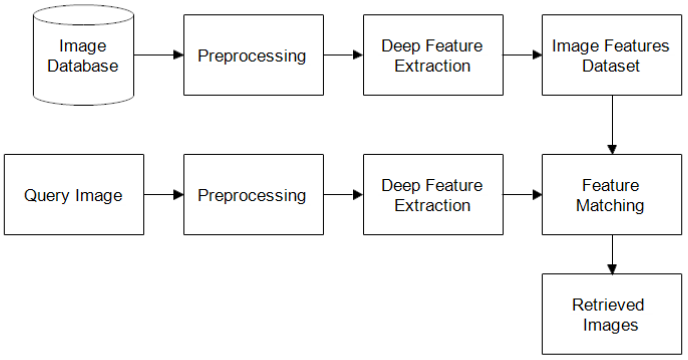


Fig. 1: Proposed methodology of system

1. Image Database / Query Image

In proposed work, standard benchmark image database, Caltech-101 is used for CBIR. (http://www.vision.caltech.edu/Image_Datasets/Caltech101/). The Caltech 101 data set consists of a total of 9,146 images, split between 101 different object categories, as well as an additional background/clutter category. Each object category contains between 40 and 800 images. Common and popular categories such as faces tend to have a larger number of images than others. Each image is about 300x200 pixels. Images of oriented objects such as airplanes and motorcycles were mirrored to be left to right aligned and vertically oriented structures such as buildings were rotated to be off axis.

2. Pre-processing

Pre-processing is the name used for operations on images at the lowest level of abstraction. The aim of the pre-processing is an improvement of the image that suppresses unwilling distortions or enhances some image features, which is important for future processing of the images. This step focuses on image feature processing. Process is performed. Main operation performed is image resizing. The images used by dataset and query image need to resize as per size of image given in the pre-trained deep CNN model.

3. Deep Feature Extraction

In proposed system, image features are analyzed based on the pre-trained deep convolutional neural network (resnet-101). Feature extraction is the easiest and fastest way to use the representational power of pre-trained deep networks. First we load a pre-trained network which is trained on more than a million images and can classify images into several object categories. As a result, the model has learned rich feature representations for a wide range of images. The network constructs a hierarchical representation of input images. Deeper layers contain higher-level features, constructed using the lower-level features of earlier layers. To get the feature representations of the images, use activations on the global pooling layer at the end of the network. The global pooling layer pools the input features over all spatial locations, giving final features.

ResNet-101 is a convolutional neural network that is 101 layers deep. A residual neural network (ResNet) is an artificial neural network (ANN) of a kind that builds on constructs known from pyramidal cells in the cerebral cortex. Residual neural networks do this by utilizing skip connections, or shortcuts to jump over some layers. Typical ResNet models are implemented with double- or triple- layer skips that contain nonlinearities (ReLU) and batch normalization in between. An additional weight matrix may be used to learn the skip weights; these models are known as HighwayNets. Models with several parallel skips are referred to as DenseNets.

Mostly in order to solve a complex problem, we stack some additional layers in the Deep Neural Networks which results in improved accuracy and performance. The intuition behind adding more layers is that these layers progressively learn more complex features.

For example, in case of recognising images, the first layer may learn to detect edges, the second

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 08 | Aug - 2021 ISSN: 2582-3930

layer may learn to identify textures and similarly the third layer can learn to detect objects and so on. But it has been found that there is a maximum threshold for depth with the traditional Convolutional neural network model.

Here is a plot shown in fig 2 that describes error% on training and testing data for a 20 layer Network and 56 layers Network.

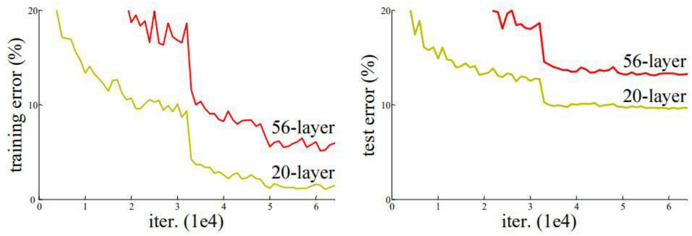
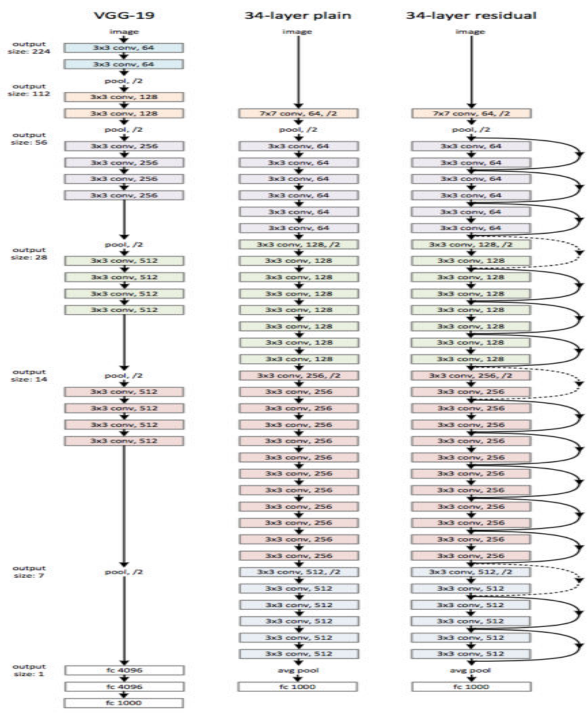


Fig 2: Training and Testing Error for N-layers

We can see that error% for 56-layer is more than a 20-layer network in both cases of training data as well as testing data. This suggests that with adding more layers on top of a network, its performance degrades.



This could be blamed on the optimization function, initialization of the network and more importantly vanishing gradient problem. We might be thinking that it could be a result of overfitting too, but here the error% of the 56-layer network is worst on both training as well as testing data which does not happen when the model is overfitting.

ResNet Architecture

ResNet network uses a 34-layer plain network architecture inspired by VGG-19 in which then the shortcut connection is added. These shortcut connections then convert the architecture into the residual network as shown in the figure below:

Fig 3: ResNet 101 Architecture

4. Feature Matching

The matching of features is carried out between the feature database and query image feature using K-nearest neighbor (KNN) having distance calculation. The image features that get matched are displayed in an ascending order and thus the number of similar images were retrieved as per given query.

RESULTS AND DISCUSSION

The proposed work is implemented on Intel CORE processor i5, 8GB RAM Laptop configuration and operating system is windows 10. MATLAB R2018b software was used to write the programming code in this we used Image processing and deep learning toolbox and the images is taken from Caltech-101 standard benchmark dataset for CBIR.

1. Dataset Feature Extraction

In our experiment, we have used following set of images given in table 1 as shown below. For feature extraction process of dataset or query image, we used automated feature extraction based on Resnet-

101 pre-trained deep convolutional neural network in which it consists of total 317 layers including input, feature, classification and output layer shown in fig 4 and 5. We have used 'pool5' feature layer to extract the features from images.

Table 1: List of Images used in Caltech-101 Dataset

Sr. No.	Image Category	Number of Samples
1	Airplanes	60
2	Ant	42
3	Barrel	47
4	Bonsai	40
5	Buddha	50
6	Butterfly	50
7	Cannon	43
8	Chandelier	50
9	Dollar bill	52
10	Elephant	42
11	Flamingo	50
12	Headphone	40
13	Lotus	50
14	Revolver	35
15	Sunflower	60

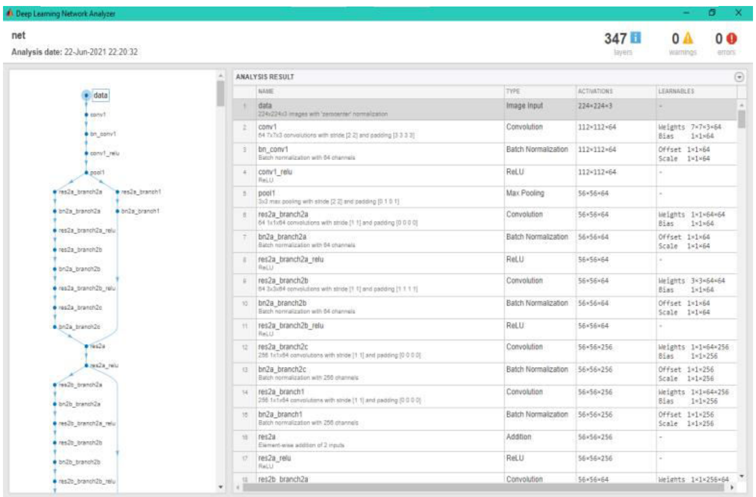


Fig: 4: Initial layers of Resnet-101 architecture

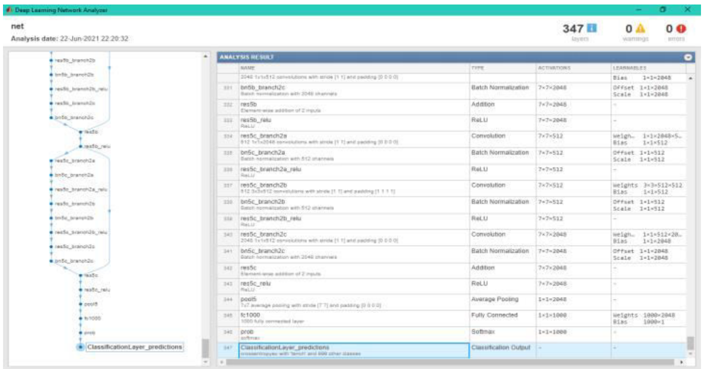


Fig: 5: Final layers of Resnet-101 architecture

Fig 6 showed the samples images of dataset with total 2048 features of whole dataset images having dimension of 711*2048 in fig 7. Total evaluation time required for project to evaluate dataset features is shown in fig 8 given below.

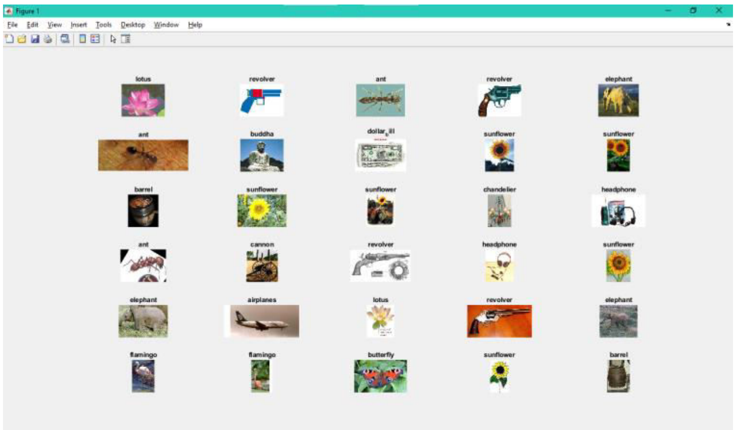


Fig: 6: Sample Images of Caltech-101 Dataset

```
Name            Size          Bytes   Class    Attributes

featuresTrain   711x2048      5824512 single
```

Fig: 7: Feature dimensions of used dataset images

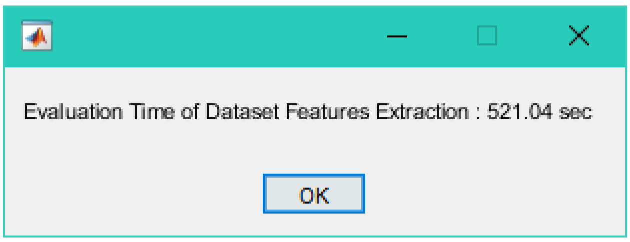


Fig: 8: Evaluation time required for dataset image feature extraction

2. Image Retrieval based on Deep Features

In our experimentation, we have tested an image from each category for retrieval. Some tested query and their equivalent retrieval images are displayed in following figures for Butterfly, Sunflower and Chandeliers is shown in following figures.

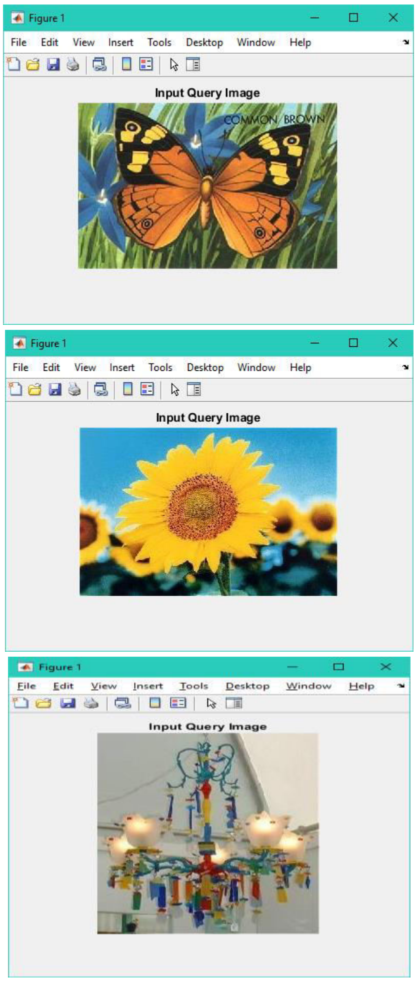


Fig: 9: Input Test Query Image – butterfly, sunflower and chandelier

Name	Size	Bytes	Class	Attributes
featuresTest	1x2048	8192	single	

Fig: 10: Feature dimension of Test Query Image

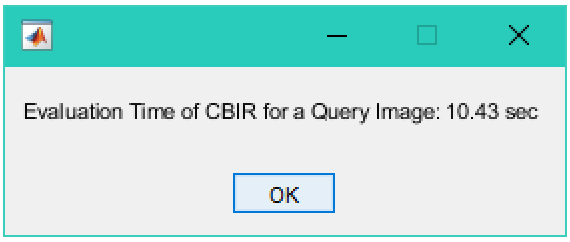


Fig: 11: Evaluation time required for CBIR of Test Query Image

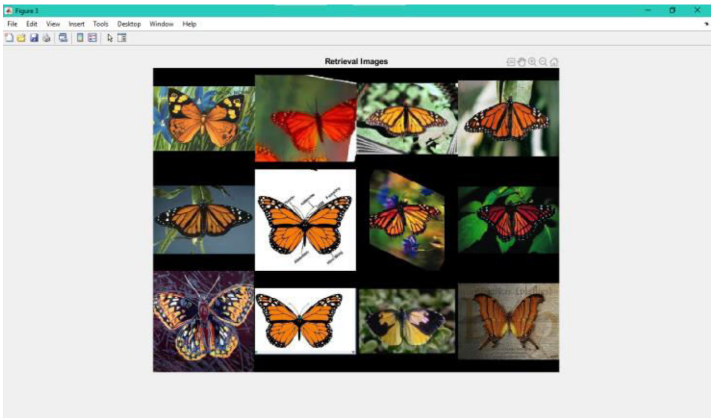


Fig: 12: Top 12 retrieval images w.r.t. Test Query Image - butterfly

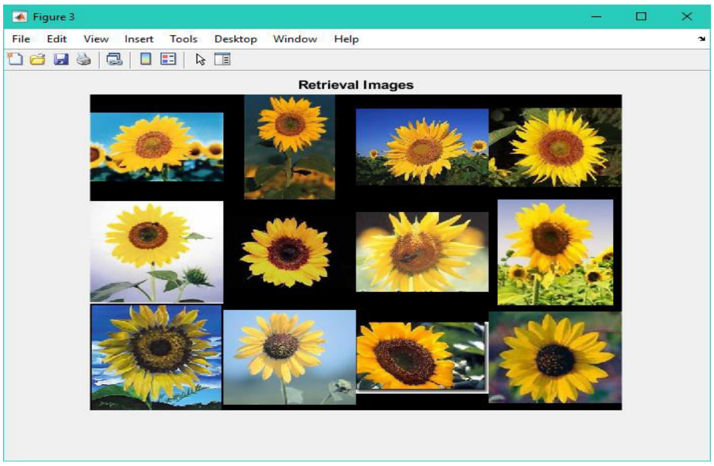


Fig: 13: Top 12 retrieval images w.r.t. Test Query Image – sunflower

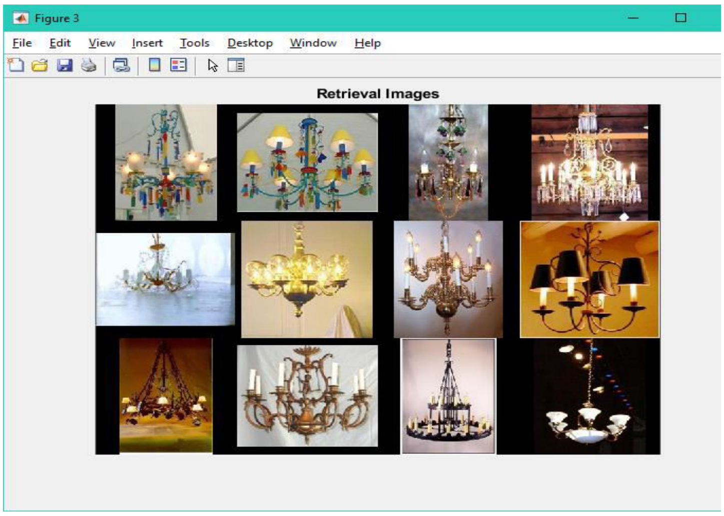


Fig: 14: Top 12 retrieval images w.r.t. Test Query Image – chandelier

CONCLUSION AND FUTURE SCOPE

Content based image retrieval is a technique which uses the visual contents to search images from large scale image databases. The proposed retrieval system uses integrated features from pre-trained deep neural network. On initializing the pre-trained model for new images will increases the performance and accuracy of the image retrieval framework. With these pre-trained models and also with assistance of transfer learning it is conceivable to retrieve new images with a superior performance. On initializing the pre-trained model trained on ImageNet for new images will increases the performance and accuracy of the image retrieval framework. With these pre-trained models and also with assistance of transfer learning it is conceivable to recover new images with a superior performance. By registering the similarity of features from feature layers the similar images are recovered. More precision rate is for the retrieval for Euclidian distance measures.

Further implementation in this project can include all category of dataset images. Also by using GPU, we can reduce the time of training of dataset features. Same model can be applied for any kind of dataset regarding to content based image retrieval (CBIR).

REFERENCES

[1] Dubey, Shiv. (2020), “A Decade Survey of Content Based Image Retrieval using Deep Learning”.

[2] Latif, Afshan & Rasheed, Aqsa & Sajid, Umer & Jameel, Ahmed & Ali, Nouman & Ratyal, Naeem Iqbal & Zafar, Bushra & Dar, Saadat & Sajid, Muhammad & Khalil, Tehmina. (2019), “Content-Based Image Retrieval and Feature Extraction: A Comprehensive Review”. Mathematical Problems in Engineering. 2019. 10.1155/2019/9658350.

[3] Ibrahim Abood, Zainab & Muhsin, Israa & T., Nabeel. (2013), “Content-based Image Retrieval (CBIR) using Hybrid Technique”. International Journal of Computer Applications. 83. 17-24. 10.5120/14500-2307.

[4] Kopparthi, Sirisha & Nynalasetti, Konadala Kameswara Rao. (2020), “Content based Image Retrieval using Deep Learning Technique with Distance Measures”. Science, Technology & Human Values. 9. 251-261.

[5] Ouhda, Mohamed & Elasnaoui, Khalid & Ouanan, Mohammed & Aksasse, B.. (2019), “Content-Based Image Retrieval Using Convolutional Neural Networks”. 10.1007/978-3-319-91337-7_41.

[6] H. Abdel-Nabi, G. Al-Naymat and A. Awajan, "Content Based Image Retrieval Approach using Deep Learning," 2019 2nd

International Conference on new Trends in Computing Sciences (ICTCS), Amman, Jordan, 2019, pp. 1-8, doi: 10.1109/ICTCS.2019.8923042.

[7] M. Alrahhal and K. P. Supreethi, "Content-Based Image Retrieval using Local Patterns and Supervised Machine Learning Techniques," 2019 Amity International Conference on Artificial Intelligence (AICAI), Dubai, United Arab Emirates, 2019, pp. 118-124, doi: 10.1109/AICAI.2019.8701255.

[8] Palwinder Kaur, Dr. Rajesh Kumar Singh, “Content Based Image Retrieval Using Machine Learning And Soft Computing Techniques”, International Journal Of Scientific & Technology Research Volume 9, Issue 01, January 2020, pp. 1327-1332

[9] Deepa Dubey, Vivek Suryawanshi, “Color, Texture and Shape Hybrid Features Based Image Retrieval System”, International Journal of Computer Science and Information Technologies, Vol. 7 (4) , 2016, 1795-1800

[10] Liu, Ying & Peng, Yanan & Lim, Kengpang & Ling, Nam. (2019). A novel image retrieval algorithm based on transfer learning and fusion features. World Wide Web. 22. 10.1007/s11280-018-0585-y.

[11] Palepu Pavani and T. Sashi Prabha, “Content Based Image Retrieval Using Machine Learning Approach”, FICTA 2013, Advances in Intelligent Systems and Computing 247, pp 173-179. DOI: 10.1007/978-3-319-02931-3_21

[12] D.Shinde, Yogita & Thepade, Sudeep. (2015). Content based Image Retrieval using Sectorisation of Self Mutated Hybrid Wavelet Transforms. International Journal of Computer Applications. 120. 16-19. 10.5120/21408-4419.

[13] Singh, Vibhav & Srivastava, Rajeev & Pathak, Yadunath & Tiwari, Shailendra & Kaur, Kuldeep. (2019). Content-based image retrieval based on supervised learning and statistical-based moments. Modern Physics Letters B. 33. 1950213. 10.1142/S0217984919502130.

[14] Poorani, M. (2013). Integrated Feature Extraction for Image Retrieval. 2. 28-35.

[15] Praveen Yadav, Mrs. Mamta Yadav, “A Hybrid Technique For Content Based Image Retrieval Using Dwt And Modified K-Means”, International Journal of Computer Science and Mobile Computing, IJCSMC, Vol. 5, Issue. 7, July 2016, pg.545 – 552

[16] N. M. Varma and A. Riyaz, "Content Retrieval using Hybrid Feature extraction from Query Image," 2018 International Conference on Information, Communication, Engineering and Technology (ICICET), Pune, India, 2018, pp. 1-4, doi: 10.1109/ICICET.2018.8533700.