Image Retrieval Using Deep Learning

Shraddha Mane[1st], Kritika Sharma[2nd], Jyoti Ware[3nd], Prof.. Zarina Shaikh[4th]
Student[1st,,2nd], Alard Charitable Trust's Alard College Of Engineering And Management
Professor[3rd], Alard Charitable Trust's Alard College Of Engineering And Management

Abstract: - During previous number of years, the planet Wide Web (WWW) has become a particularly well-liked data source. To with success utilize the huge amount of knowledge that the net provides, we wish a good thanks to explore it. Image knowledge is far additional voluminous than matter knowledge, and visual data can't be indexed by ancient methods Developed for compartment allocation matter data. Therefore, Content-Based Image Retrieval has received a wonderful deal of interest among the analysis community. A CBIR system operates on the visible options at low-level of a user's input image that makes it difficult for the users to plot the input and additionally does not supply adequate retrieval results. In CBIR system, the study of the helpful illustration of options and appropriate similarity metrics is very necessary for improving the performance of retrieval task. Linguistics gap has been the most issue that happens between image pixels at low level and linguistics at high-level understood by humans. Among varied ways, machine learning (ML) has been explored as a feasible thanks to cut back the linguistics gap. Galvanized by the present success of deep learning ways for pc vision applications, during this paper, we tend to aim to confront AN advance deep learning methodology, referred to as Convolutional Neural Network (CNN), for learning feature representations and similarity measures. During this paper, we tend to Explored the applications of CNNs towards determination classification and retrieval issues. For retrieval of comparable pictures, we tend to in agreement on victimization transfer learning to apply the Google Net deep design to our downside. Extracting the last-but-one absolutely connected layer from the retraining of Google Net CNN model served because the feature vectors for each image, computing Euclidian distances between these feature vectors which of our question image to come the highest matches within the dataset.

Keywords: Deep Learning, Convolutional Neural Network, Transfer Learning,

I. INTRODUCTION

In the days of Internet boom where social networks and reasonable smartphones are capable of taking high-quality photos and videos, users have automatic access to several images across the Web. Given these circumstances, the necessity to search, filter and organize the images is a lot more crucial. In the case of small collections, it is possible to search for the specified pictures or duplicates manually. This becomes impractical if the quantity of items increases. To deal with this fast growth there is a need to develop the image retrieval systems that will operate extensively. The main manage and enquire database of images in a precise manner.

CBIR is the strategy of automatically indexing pictures by the extraction of visible features at low-level, like shape, color and texture and these indexed features are entirely responsible for the retrieval of images. In typical CBIR systems the visible information of the pictures in the database of images is separated and illustrated by multidimensional vectors of features. The vectors of features derived from the pictures present in the database then form a database of features. To fetch similar pictures, the query image is provided by users to retrieval system. Image retrieval system then modifies these query images into a representative model of feature vectors. The resemblance between the query picture's feature vector and the vectors of pictures in the database is then studied, and retrieval is executed with the help of an indexing strategy. The indexing procedure specifies an economical manner to find out similar pictures in the image database.

CNNs are a specific type of ANN for handling data that features a grid-like topology like, image data, which is a 2D grid of pixels. CNNs are merely ANNs that involves the use of convolution instead of conventional matrix multiplication operation in a minimum of one in all their layers. Convolution supports three essential concepts that can facilitate in improving a ML system: parameter sharing, equivariant representations, and sparse interactions. CNNs are eminent for their potential to learn shapes, textures, and colors, making this problem suitable for the application of neural networks.

II. LITERATURE SURVEY

[1] "Rethinking the inception architecture for computer vision"

Author: C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2818–2826.

Description: Convolutional networks are at the core of most state-of-the-art computer vision solutions for a wide variety of tasks. Since 2014 very deep convolutional networks started to become mainstream, yielding substantial gains in various benchmarks. Although increased model size and computational cost tend to translate to immediate quality gains for most tasks (as long as enough labeled data is

provided for training), computational efficiency and low parameter count are still enabling factors for various use cases such as mobile vision and big-data scenarios. Here authors are exploring ways to scale up networks in ways that aim at utilizing the added computation as efficiently as possible by suitably factorized convolutions and aggressive regularization. We benchmark our methods on the ILSVRC 2012 classification challenge validation set demonstrate substantial gains over the state of the art: 21.2% top-1 and 5.6% top-5 error for single frame evaluation using a network with a computational cost of 5 billion multiply-adds per inference and with using less than 25 million parameters. With an ensemble of 4 models and multi-crop evaluation, we report 3.5% top-5 error and 17.3% top-1 error on the validation set and 3.6% top-5 error on the official test set.

[2] "Going deeper with convolutions"

Author: C. Szegedy,W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.

Description: We propose a deep convolutional neural network architecture codenamed Inception that achieves the new state of the art for classification and detection in the ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14). The main hallmark of this architecture is the improved utilization of the computing resources inside the network. By a carefully crafted design, we increased the depth and width of the network while keeping the computational budget constant. To optimize quality, the architectural decisions were based on the Hebbian principle and the intuition of multi-scale processing. One particular incarnation used in our submission for ILSVRC14 is called GoogLeNet, a 22 layers deep network, the quality of which is assessed in the context of classification and detection**.**

[3]"Image Based Search Engine Using Deep Learning"

Author: Surbhi Jain, Joydip Dhar, in Proceedings of 2017 Tenth International Conference on Contemporary Computing (IC3), 10-12 August 2017, Noida, India.

Description: During previous couple of years, the World Wide Web (WWW) has become an extremely well-liked information source. To successfully utilize the vast quantity of information that the web provides, we want an effective way to explore it. Image data is much more voluminous than textual data, and visual information cannot be indexed by traditional strategies developed for indexing textual information. Therefore, Content- Based Image Retrieval (CBIR) has received an excellent deal of interest within the research community. A CBIR system operates on the visible features at low-level of a user's input image that makes it troublesome for the users to devise the input and additionally doesn't offer adequate retrieval results. In CBIR system, the study of the useful representation of features and appropriate similarity metrics is extremely necessary for improving the performance of retrieval task. Semantic gap has been the main issue which occurs between image pixels at low level and semantics at high-level interpreted by humans. Among varied methods, machine learning (ML) has been explored as a feasible way to reduce the semantic gap. Inspired by the current success of deep learning methods for computer vision applications, in this paper, we aim to confront an advance deep learning method, known as Convolutional Neural Network (CNN), for studying feature representations and similarity measures. In this paper, we explored the applications of CNNs towards solving classification and retrieval problems. For retrieval of similar images, we agreed on using transfer learning to apply the Google Net deep architecture to our problem. Extracting the last-but-one fully connected layer from the retraining of Google Net CNN model served as the feature vectors for each image, computing Euclidean distances between these feature vectors and that of our query image to return the closest matches in the dataset.

[4] "A Novel Technique For Content Based Image Retrieval Using Color, Texture And Edge Features"

Author : Manpreet Kaur, Neelofar Sohi

Description: The massive growth of digital technology along with use of internet has increased the use of audio-visual data such as images and videos in many domains like digital museums, commercial use, crime prevention, medical images, remote sensing and so on. With increasing volume of digital data, search and retrieval of relevant images from large datasets in accurate and efficient way is a challenging problem. CBIR combines the contents or features of image like color, texture, edges rather than keywords, labels related to an image. This paper presents systematic literature review of various image retrieval techniques presenting the basic concepts and available methods with their research gaps. In this study, retrieval techniques based on features like HSV, Color Moment, HSV and Color Moment, Gabor Wavelet and Wavelet Transform, Edge Gradient are studied and implemented. An approach is proposed for retrieval based on combination of color, texture and edge features of image. Performance evaluation of studied image retrieval techniques and proposed technique is done using parameters like Sensitivity, Specificity, Retrieval score, Error rate and Accuracy. Experimental results of performance evaluation demonstrate that proposed technique outperforms other techniques**.**

[5] "Implementation of Deep Learning Algorithm with Perceptron using Tensor Flow Library"

Author: Arshiya Begum, Farheen Fatima and Asfia Sabahath, in International Conference on Communication and Signal Processing, April 4-6, 2019, India.

Description: In recent years, Deep Learning, Machine Learning, and Artificial Intelligence are highly focused concepts of data science. Deep learning has achieved success in the field of Computer Vision, Speech and Audio Processing, and Natural Language Processing. It has the strong learning ability that can improve utilization of datasets for the feature extraction compared to traditional Machine Learning Algorithm. Perceptron is the essential building block for creating a deep Neural Network. The perceptron model is the more general computational model. It analyzes the unsupervised data, making it a valuable tool for data analytics. A key task of this paper is to develop and analyze learning algorithm. It begins with deep learning with perceptron and how to apply it using Tensor Flow to solve various issues. The main part of this paper is to make perceptron learning algorithm well behaved with non-separable training datasets. This type of algorithm is suitable for Machine Learning, Deep Learning, Pattern Recognition, and Connectionist Expert System.

III. PROPOSED SYSTEM

Our approach incorporates three modules. The first module is the removal of background from the image to eliminate undesirable features and extraction of desirable features to represent useful information in the image. The second module is the use of transfer learning to retrain the pre-trained deep CNN model to study field specific features representations. Finally, the third module retrieves pictures similar to the query picture based on the features matching from the trained model using Euclidean distance as the similarity metric.

Advantages:

- Using small amount of data massive data can be generated at runtime.
- More than one related outcomes occur by only one search.
- Accuracy should be increased.
- Efficiency is increased

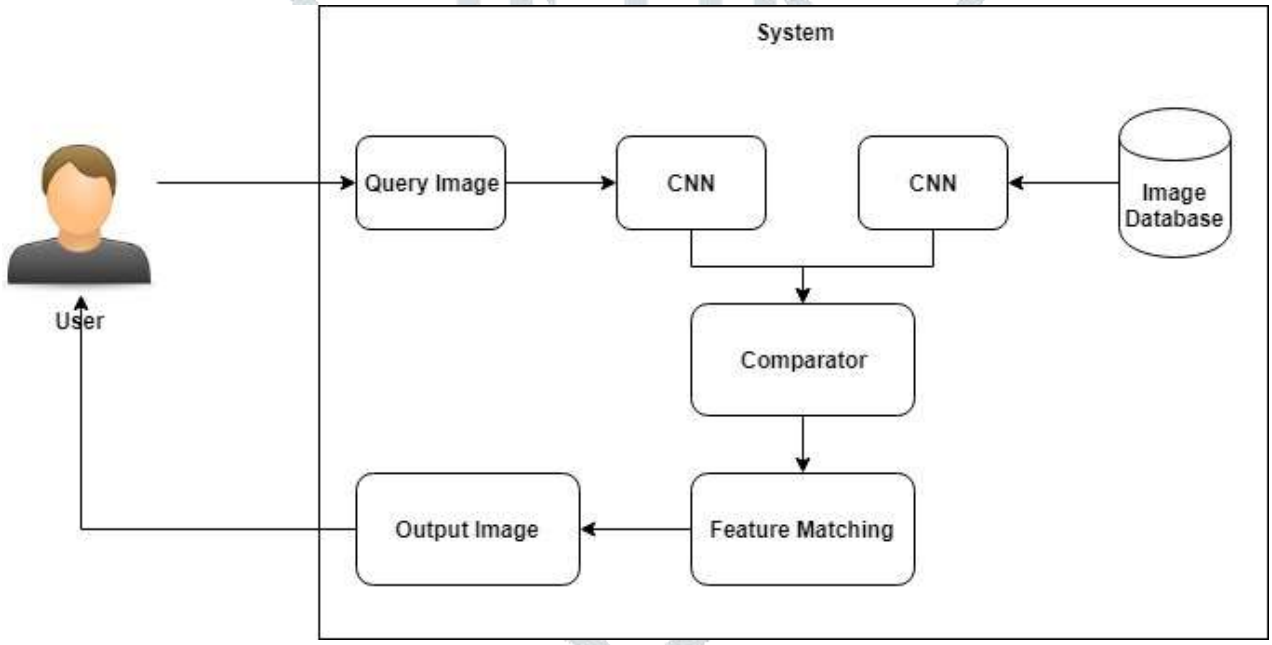

Figure:1 System Architecture Diagram

IV. RESULTS:

The following graph represents the difference between Existing and Proposed System's time taken in ms and the accuracy of VGG-19K1, VGG-19K2, VGG-19DG, VGG-19BR, VGG-19FS, VGG-19BO models in % for the result of Image based Search.

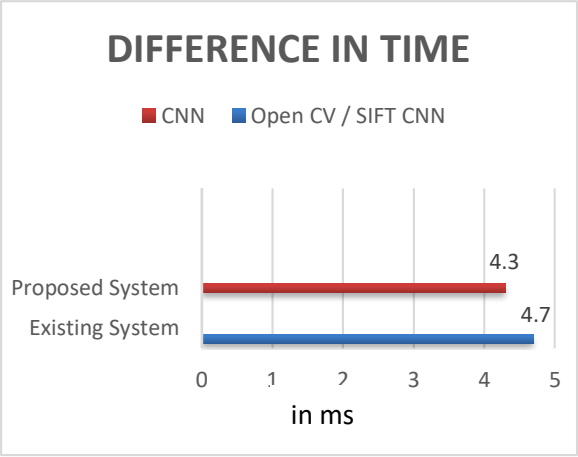

Fig 2. Time-Comparison Graph

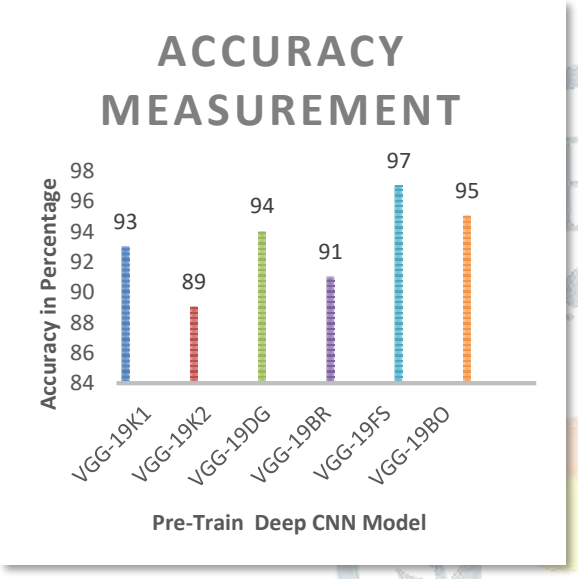

Fig 3. Accuracy Graph of Overall System

V. CONCLUSION

This paper presents a technique to retrieve similar images using transfer learning. We achieved top-1 accuracy as 97% and 89% accuracy for top-10 predictions in correctly classifying historical images in test data. The accuracy obtained for classification task on Fish and Dogs datasets for top-1 and top-3 predictions are 97% and 95% respectively. To retrieve images similar to query image, Euclidean distance was used as similarity metric on the last-but-one fully connected layer of the find-tuned VGG-19K1 model.

VI. FUTURE WORK

More vast and varied datasets can be added inside the system. Different pre- trained libraries can be added and used in the system.

VII. REFERENCES

[1] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2818–2826.

[2] A. Iliukovich-Strakovskaia, A. Dral, and E. Dral, "Using pre-trained models for fine-grained image classification in fashion field," 2016.

[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[4] C. Szegedy,W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.

[5] K. Lin, H.-F. Yang, J.-H. Hsiao, and C.-S. Chen, "Deep learning of binary hash codes for fast image retrieval," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015, pp. 27–35.

[6] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," Pattern recognition, vol. 40, no. 1, pp. 262–282, 2007.

[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[8] C. Szegedy,W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.

[9] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2818– 2826.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105.

[11] A. Babenko, A. Slesarev, A. Chigorin, and V. Lempitsky, "Neural codes for image retrieval," in European conference on computer vision. Springer, 2014, pp. 584–599.

[12] R. Xia, Y. Pan, H. Lai, C. Liu, and S. Yan, "Supervised hashing for image retrieval via image representation learning." in AAAI, vol. 1, 2014, p. 2.

[13] Surbhi Jain, Joydip Dhar, " Image Based Search Engine Using Deep Learning", Proceedings of 2017 Tenth International Conference on Contemporary Computing (IC3), 10-12 August 2017, Noida, India.