International Journal of Interdisciplinary Innovative Research &Development (IJIIRD)
ISSN: 2456-236X
Vol. **05** Special Issue **01 | 2020**

Rotation Invariant for Image Retrieval and Deep Learning

[1]Payal Bhole, [2]Saurabhi Desai, [3]Mansi Jadhav, [4]Mrs. D. N. HIRE

[1,2,3,4]Dept of E&TC Dr. D. Y. Patil Institute of Engineering, Management and Research,
D. Y. Patil Educational Complex, Sector 29, Nigidi Pradhikaran, Akurdi, Pune 411044.

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 (CBIR) 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 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 Convolution Neural Network (CNN), for learning feature representations and similarity measures. During this paper, we tend to explore 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, Convolution Neural Network, Transfer Learning,

1. INTRODUCTION

In the days of Internet boom where social networks and reasonable smart phones 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 involve 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.

2. LITERATURE SURVEY

Paper 1: Using pre-trained models for fine-grained image classification in fashion field. Description: Fashion image classification offers online stores a fast and effective way to manage the volume of their products. It could be used in product categorization by their brands, types, styles, etc. It could be also used to make fashion ensembles of clothes and even to create product recommendations. The results of apparel image classification are also useful for such purposes as product filtering and search. A wide range of different models have been successfully applied to coarse level image classification, but only a few of them could be applied to fashion images due to the nature of the data: such images could belong to the one high-level topic such as 't-shirt', 'skirt' or 'sport shoes', but at the lower level be fine-grained (e.g. 'high heels', 'medium heels'). Each apparel shop can have its own fine-grained hierarchy leading to difficulties to acquire a big dataset. This work shows necessity of pre-trained model usage to achieve a good predictive quality. In addition, the article contains a description and an experimental study of an approach to stack features from pre trained model and features learned directly from an image pixel representation. This approach significantly outperforms baseline methods in fine-grained image classification and a classification based only on a pre-trained model.

Paper 2: Going deeper with convolutions.
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 Google Net, a 22 layers deep network, the quality of which is assessed in the context of classification and detection.
Paper 3: Deep learning of binary hash codes for fast image retrieval.
Description: Approximate nearest neighbor search is an efficient strategy for large-scale image retrieval. Encouraged by the recent advances in convolutional neural networks (CNNs), we propose an effective deep learning framework to generate binary hash codes for fast image retrieval. Our idea is that when the data labels are available, binary codes can be learned by employing a hidden layer for representing the latent concepts that dominate the class labels. The utilization of the CNN also allows for learning image representations. Unlike other supervised methods that require pair-wised inputs for binary code learning, our method learns hash codes and image representations in a point-wised manner, making it suitable for large-scale datasets. Experimental results show that our method outperforms several state-ofthe- art hashing algorithms on the CIFAR-10 and MNIST datasets. We further demonstrate its scalability and efficacy on a large-scale dataset of 1 million clothing images.
Paper 4: Rethinking the inception architecture for computer vision.
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.

3. PROPOSE SYSTEM

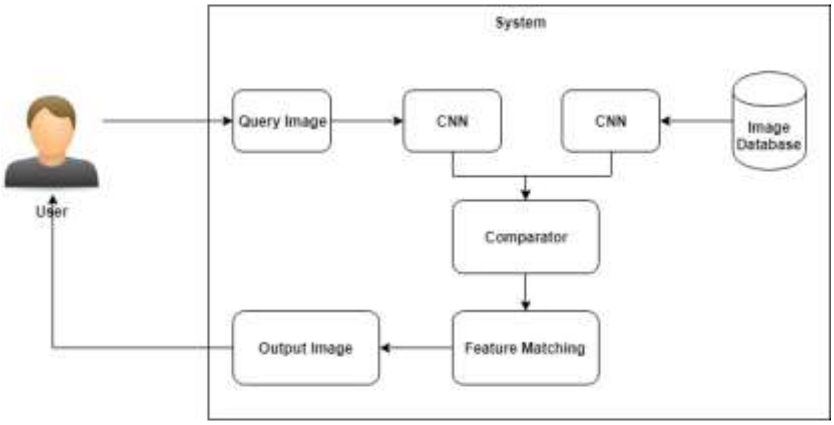
Our approach incorporates three modules. The first module is the removal of background from the image to eliminate undesirable features and extraction of desirable features to represent useful information in the image. The second module is the use of transfer learning to retrain the pre-trained deep 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.

3.1 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

3.2 Methodology:

Basically, the input query image will be given to the CNN model where it will be compared to the images that are already present in the database and provides all the related images (all rotation) at the output side .Type of CNN module used entirely depends upon the kind of features that are to be extracted from the input query image. Tensor-flow software will be used for the purpose of image feature extraction. Further these extracted features are compared with the images in the database and we obtain all the related images as an output



4. CONCLUSION

Using this technique we can retrieve similar images with all angular rotation as that of the query input image from the database that already exists.

5. 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.