International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

A Deep Learning Framework for Content Based Image Retrieval using Enhanced InceptionV3

[1]Vignesh Thati, [2]Rohan Yadav Kothapally, [3]Sri Phani Kumar, [4]V Satheesh Kumar

20311A0571 Dept. CSE SNIST,HYD 20311A0571@sreenidhi.edu.in
20311A0577 Dept. CSE SNIST,HYD 20311A0577@sreenidhi.edu.in
20311A0561 Dept. CSE SNIST,HYD 20311A0561@sreenidhi.edu.in
Asst. Professor Dept. CSE SNIST,HYD satteesh3@gmail.com

Abstract - The invent of cloud computing made multimedia applications very popular. Particularly image content is widely used by different computer vision applications in the world. In this regard, Content Based Image Retrieval (CBIR) plays crucial role in accessing most relevant images. The existing heuristics based approaches for CBIR have certain limitations. Artificial Intelligence (AI) enabled approaches came into existence that exploit learning based approaches leveraging performance of CBIR. However, there is need for improvement in deep learning models towards optimizing CBIR system that retrieves images to satisfy user intent. In this study, we proposed a deep learning architecture that is trained to realize an effective CBIR system using a modified InceptionV3 model. We proposed an algorithm known as Learning Based Image Retrieval (LBIR) which takes user query image as input and retrieves most relevant images that reflect user intention. Our testing findings showed that, with the highest accuracy of 96%, the suggested CBIR system outperforms several of the current ones.

Keywords – Content Based Image Retrieval, Artificial Intelligence, Deep Learning, Modified Inception V3

1. INTRODUCTION

With the availability of cloud computing environments and the supportive ecosystem with distributed programming frameworks, many computer vision applications in the real world are dealing with multimedia objects and processing them with ease. With respect to images, there are number of applications that involve in storing and retrieval of images in diversified domains. In this context, retrieving images based on the given input image became very popular. This phenomenon is also known as Content Based Image Retrieval (CBIR).

The research related to CBIR systems reveal traditional heuristics based approaches and AI enabled methodologies in retrieving images efficiently [3], [4]. With the emergence of deep learning models, processing image content became easier. Moreover, deep learning models like Convolutional Neural Network (CNN) is found to have ability to deal with images [3]. There are many CNN variants as found in the literature. However, as one size does not fit all, CNN cannot be used directly with all kinds of applications.

In order to improve efficiency in CBIR system, in this paper, we modified a pre-trained deep learning model and CNN variant known as Inception V3. The modified InceptionV3 is exploited in our proposed framework to achieve efficient image retrieval in query by example fashion. Literature review has revealed significant progress in usage of deep learning models.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Ghosh et al. [6] examined how deep learning affects picture segmentation, classifying key methods and providing an intuitive explanation of each one's contribution. With deep learning, image matching—which is essential for visual applications. With insights and comparisons, this survey navigates both ancient and current methodologies [7]. Zhang et al. [13] examined the use of deep learning to visual location recognition, describing techniques, resources, and future perspectives. Ozturk et al. [15] presented OCAM, a unique triplet-learning technique that overcomes obstacles and outperforms current methods for medical image retrieval.

From the literature, it is revealed that the existing deep learning models need to be optimized towards improving efficiency of CBIR. Motivated by this fact, we enhanced InceptionV3 model in such a way that its performance is improved in analyzing images and retrieve images based on the given input image more efficiently. Our contributions in this paper are as follows.

1. We proposed a deep learning framework based on modified InceptionV3 model which is trained to realize an efficient CBIR system.
2. We proposed an algorithm known as Learning Based Image Retrieval (LBIR) which takes user query image as input and retrieves most relevant images that reflect user intent.

The rest of the paper is structured as follows. Section 2 reviews prior works on CBIR focusing more on deep learning techniques. Section 3 presents the proposed deep learning framework along with modified InceptionV3 algorithm. Section 4 describes observations associated with the empirical study made in this paper. Section 5 concludes our research and provides scope of the research possible in future.

2. RELATED WORK

This section presents review of literature pertaining to image retrieval with heuristic approaches and deep learning methods. Lu et al. [1] improved binary codes and get beyond current deep hashing restrictions, Deep Fuzzy Hashing Network (DFHN) integrates fuzzy logic with DNN. Shen et al. [2] with encryption make it possible to model and share data securely. Privacy is the first priority in a blockchain-based system that is being developed for medical picture retrieval. Gu et al. [3] for image retrieval, the Clustering-driven Unsupervised Deep Hashing (CUDH) technique concurrently learns binary codes and discriminative clusters. Cai et al. [4] enabled hypothesis testing and the application of domain knowledge, refinement tools improve user confidence and utility in machine learning (ML)-based image retrieval for medical decision-making. Hossain et al. [5] discussed prominent datasets, assessment criteria, and deep learning-based picture captioning systems, evaluating their effectiveness, drawbacks, and merits.

Ghosh et al. [6] examined how deep learning affects picture segmentation, classifying key methods and providing an intuitive explanation of each one's contribution. Ma et al. [7] with deep learning, image matching—which is essential for visual applications—develops. With insights and comparisons, this survey navigates both ancient and current methodologies. Liu et al. [8] examined more than 300 submissions and suggests future paths. In computer vision, generic object recognition is revolutionized by deep learning. Sumbul et al. Zhang et al. [9] examined the use of deep learning to visual location recognition, describing techniques, resources, and future perspectives. Ozturk et al. [10] presented OCAM, a unique triplet-learning technique that overcomes

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

obstacles and outperforms current methods for medical image retrieval.

3. PROPOSED CBIR SYSTEM

This section provides details about the proposed CBIR system which is based on a deep learning framework that exploits modified InceptionV3 model for efficient image retrieval when query is given by example.

3.1 Problem Definition

When an image is given as input, developing a CBIR system based on deep learning to retrieve images efficiently reflecting the user intent is the problem considered.

3.2 Our Framework

We proposed a deep learning based framework meant for improving efficiency in retrieval of images provided an image as input in query by example fashion. The proposed framework is based on modified Inception V3 model which is found efficient in retrieving features and analyzing them from input query image and also images in the training dataset. The input query image is taken by the user and most relevant images are retrieved by the system. The system has both training process and testing process. In the training process the system uses modified InceptionV3 model to train with all the training images in the dataset. Once the training is completed, the model is used in the online phase to get the input query image and extract its features for further processing. System has a matching module which makes use of distance based similarity metric to identify images that match with the query image while retrieving resultant images. It makes use of the features database which has all features of training images. The system is efficient because the training samples will be improved over a period of time. Thus, the knowledge of the system gets increased with every set of new training samples. Since InceptionV3 is a pre-trained model which gets retrained as and when new samples are available, it gains knowledge incrementally. Figure 1 shows the architecture of the proposed CBIR system.

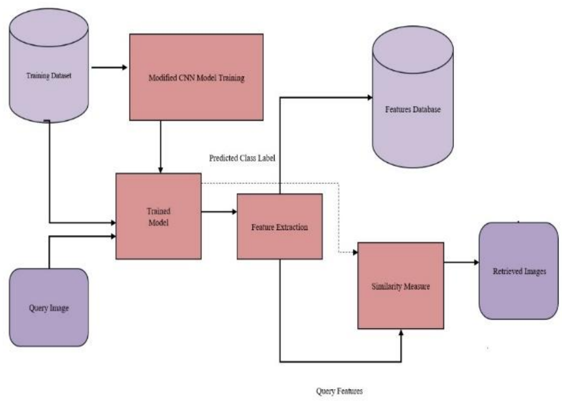


Figure 1: The proposed framework for efficient learning based CBIR

3.3 Enhancing InceptionV3 Model

We enhanced inceptionV3 model to enable it for efficient retrieval of images given a query image. We analyzed inceptionV3 architecture and it is improved further towards more depth in analyzing image content. InceptionV3 is a CNN variant which is widely used in computer vision applications. It is also used in some of the CBIR systems and found to be efficient in retrieval of images. In this paper, we modified it by adding certain layers as described below to improve its performance further. The overview of InceptionV3 architecture is found in Figure 2.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

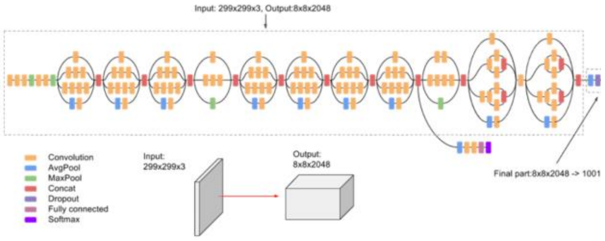


Figure 2: Architecture of Inception V3 model

The InceptionV3 model is based on CNN architecture which is pre-trained with ImageNet dataset. Compared with other CNN variants, InceptionV3 model is found to be more efficient and faster due to its internal architecture and modus operandi. The characteristics of InceptionV3 model made it more practical in solving many real-world problems as part of computer vision applications. The network exploits modular system and helps in progressively analogizing the features and optimizing the features. The modified InceptionV3 model supports transfer learning in order to learn from a greater number of training samples in future. The transfer learning process improves stability of the model besides its generalization to get rid of negative influence by image pixels in the eventual classification of images. Since InceptionV3 pre-trained with ImageNet dataset, it has existing learning weights that are reused in future with retraining of the model. The convolutional layers associated with Inception V3 model are frozen before the fully connected and soft max layers. By adjusting network parameters continuously, the softmax layer and fully connected layer are retrained to obtain deep level features in order to improve performance of CBIR system.

3.4 Proposed Algorithm

We proposed an algorithm known as Learning Based Image Retrieval (LBIR) which takes user query image as input and retrieves most relevant images that reflect user intention

Algorithm: Learning Based Image Retrieval (LBIR)

Inputs

Image dataset D, input query image q

Output

CBIR results R, Performance Statistics P

1. Begin
2. D'←Pre-Process(D)
3. (T1, T2)←SplitData(D')
4. Configure modified InceptionV3 model
5. Train InceptionV3 model with T1
6. Save model
7. Load saved model
8. Take the query image q
9. F←ProcessQuery(q or any random sample from T2)
10. R←Matching(F, features of T1)
11. P←Evaluation(R, ground truth)
12. Display R
13. Display P
14. End

Algorithm 1: Learning Based Image Retrieval (LBIR)

As presented in Algorithm 1, it takes dataset and query image as input and provide CBIR results and also performance statistics. It has provision for pre-processing of images by improving quality and splitting data into training (T1) and testing (T2). The modified IncpetionV3 is configured and trained with T1. The trained model is used to extract features from query image and then there is matching for image retrieval.

3.5 Dataset Details and Evaluation Procedure

Dataset used in the experiments has 25,000 images consisting of 6 categories as presented in Table 1. Each image in the dataset is of size 150x150.

Category	Description
0	Buildings
1	Forest

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

2	Glacier
3	Mountain
4	Sea
5	Street

Table 1: Image categories

The data set is divided into 14000 training samples, 3000 test samples and 7000 validation samples. The proposed system is evaluated using the following performance metrics.

$$\text{Precision (p)} = \frac{TP}{TP+FP} \quad (1)$$

$$\text{Recall (r)} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{F1-score} = 2 * \frac{(p*r)}{(p+r)} \quad (3)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

The measures used for performance evaluation result in a value that lies between 0 and 1. These metrics are widely used in machine learning research.

4. EXPERIMENTAL RESULTS

This section presents results of experiments with a prototype application to evaluate the proposed deep learning framework and its underlying algorithm. Experiments are made with number of randomly selected input images and the retrieved results are observed for their correctness. The proposed deep learning model known as modified InceptionV3 is evaluated and compared with existing deep learning models. This section provides results of some of the queries besides presenting performance statistics in terms of accuracy, F1-score, precision and recall. The performance of enhanced InceptionV3 model is compared against deep learning models like CNN and InceptionV3.

Query Image

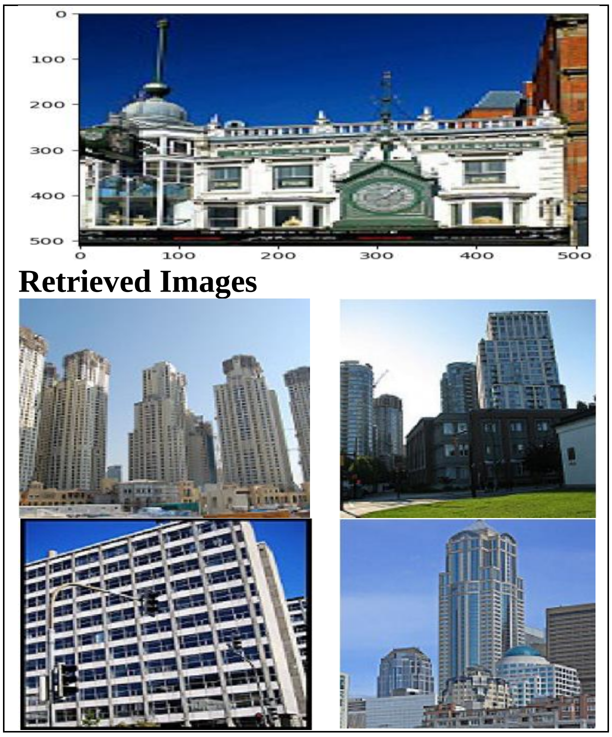
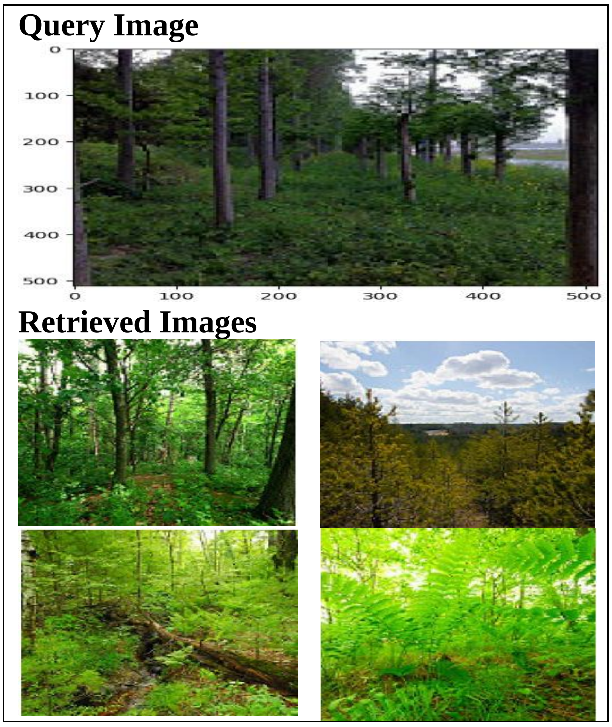


Figure 3: Query image and resultant images of experiment one

As presented in Figure 3, the given input query image is processed by the proposed CBIR system and the resultant images are found relevant to given input image as part of experiment one.



International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Figure 4: Query image and resultant images of experiment two

As presented in Figure 4, the given input query image is processed by the proposed CBIR system and the resultant images are found relevant to given input image as part of experiment two.

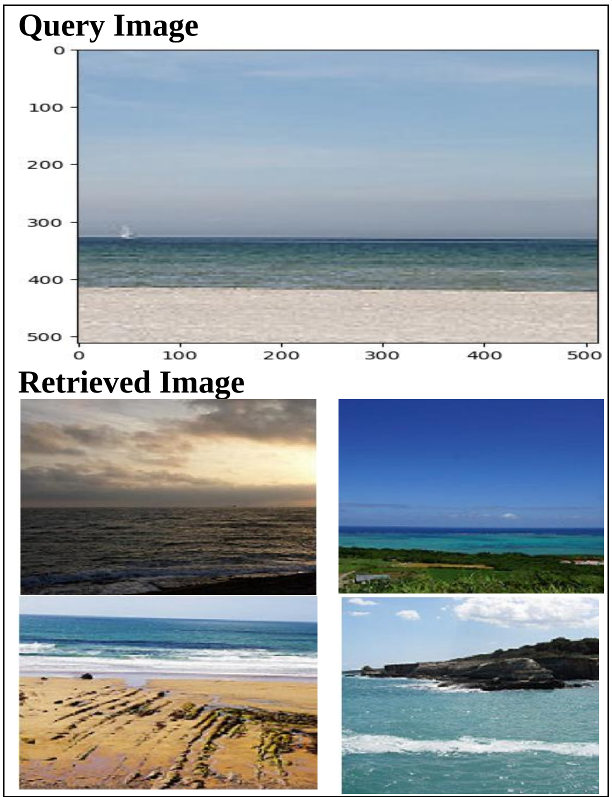


Figure 5: Query image and resultant images of experiment three

As presented in Figure 5, the given input query image is processed by the proposed CBIR system and the resultant images are found relevant to given input image as part of experiment three.

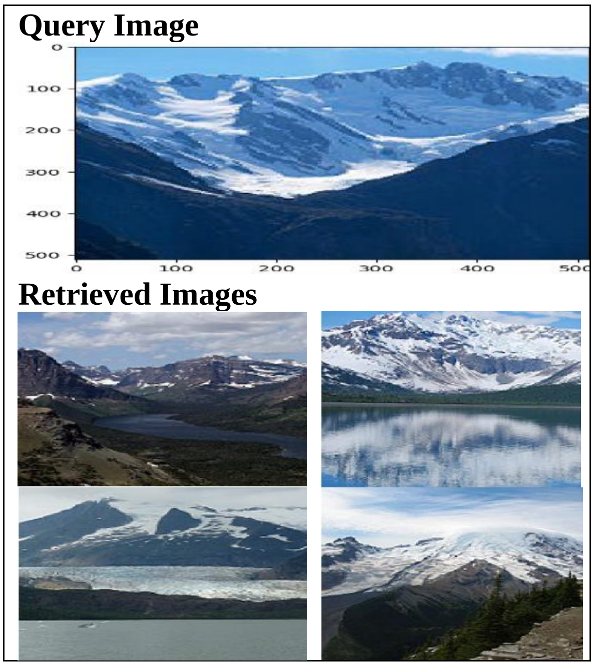


Figure 6: Query image and resultant images of experiment four

As presented in Figure 6, the given input query image is processed by the proposed CBIR system and the resultant images are found relevant to given input image as part of experiment four.

Model	Precision	Recall	F1-Score	Accuracy
CNN	0.89	0.92	0.9	0.93
InceptionV3	0.94	0.89	0.91	0.92
Modified InceptionV3	0.97	0.91	0.94	0.96

Table 2: Performance comparison among model

As presented in table two the performance of modified InceptionV3 is compared against baseline models like CNN and InceptionV3.

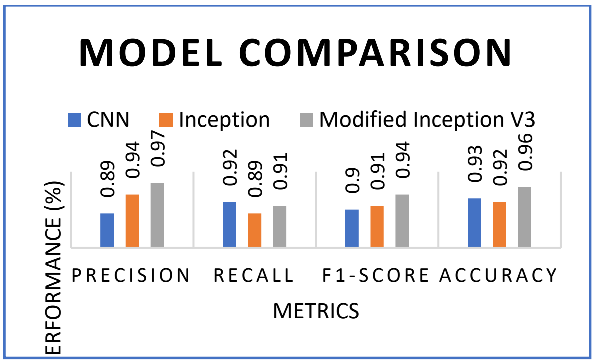


Figure 7: Performance of different models in retrieval of images

The enhanced InceptionV3 model is compared against existing deep learning models to evaluate its performance as shown in Figure 7. Each model showed different level of performance due to their internal architecture and modus operandi. With respect to precision CNN model achieved 89%, InceptionV3 94% and modified InceptionV3 97% precision. The baseline CNN model achieved 92% recall, InceptionV3 89% and modified InceptionV3 91% recall. with respect to F1-score the baseline CNN model could achieve 90%, InceptionV3 91% and modified InceptionV3 94% F1-score. The baseline CNN model achieved 93% accuracy, InceptionV3 model 92% and modified InceptionV3 model 96% accuracy. The highest accuracy is achieved by the modified InceptionV3 model which is used in the proposed CBIR system.

6. CONCLUSION AND FUTURE WORK

In this paper, we proposed a deep learning framework based on modified InceptionV3 model which is trained to realize an efficient CBIR system. The proposed framework is based on supervised learning process where the inception V3 model which has been modified is trained with training samples and used for efficient retrieval of images in query by example fashion. The model is found efficient in retrieval of images provided any given query image. We proposed an algorithm known as Learning Based Image Retrieval (LBIR) which takes user query image as input and retrieves most relevant images that reflect user intention. Our experimental results revealed that the proposed CBIR system is better than many existing ones with highest accuracy 96%. In future, we intend to improve our framework with GAN architecture and hybrid deep learning models towards leveraging CBIR system further.

REFERENCES

[1] Lu, Huimin; Zhang, Ming; Xu, Xing; Li, Yujie and Shen, Heng Tao (2020). *Deep Fuzzy Hashing Network for Efficient Image Retrieval. IEEE Transactions on Fuzzy Systems, 1–1.* http://doi:10.1109/tfuzz.2020.2984991

[2] Shen, Meng; Deng, Yawen; Zhu, Liehuang; Du, Xiaojiang and Guizani, Nadra (2019). *Privacy-Preserving Image Retrieval for Medical IoT Systems: A Blockchain-Based Approach. IEEE Network, 33(5), 27–33.* http://doi:10.1109/MNET.001.1800503

[3] Gu, Yifan; Wang, Shidong; Zhang, Haofeng; Yao, Yazhou; Yang, Wankou and Liu, Li (2019). *Clustering-driven unsupervised deep hashing for image retrieval. Neurocomputing, 368, 114–123.* http://doi:10.1016/j.neucom.2019.08.050

[4] Cai, Carrie J.; Stumpe, Martin C.; Terry, Michael; Reif, Emily; Hegde, Narayan; Hipp, Jason; Kim, Been; Smilkov, Daniel; Wattenberg, Martin; Viegas, Fernanda and Corrado, Greg S. (2019). *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19 -*

Human-Centered Tools for Coping with Imperfect Algorithms During Medical Decision-Making. 1–14. http://doi:10.1145/3290605.3300234

[5] Hossain, MD. Zakir; Sohel, Ferdous; Shiratuddin, Mohd Fairuz and Laga, Hamid (2019). *A Comprehensive Survey of Deep Learning for Image Captioning. ACM Computing Surveys, 51(6), 1–36.* http://doi:10.1145/3295748

[6] Ghosh, Swarnendu; Das, Nibaran; Das, Ishita and Maulik, Ujjwal (2019). *Understanding Deep Learning Techniques for Image Segmentation. ACM Computing Surveys, 52(4), 1–35.* http://doi:10.1145/3329784

[7] Ma, Jiayi; Jiang, Xingyu; Fan, Aoxiang; Jiang, Junjun and Yan, Junchi (2020). *Image Matching from Handcrafted to Deep Features: A Survey. International Journal of Computer Vision.* http://doi:10.1007/s11263-020-01359-2

[8] Liu, Li; Ouyang, Wanli; Wang, Xiaogang; Fieguth, Paul; Chen, Jie; Liu, Xinwang and Pietikäinen, Matti (2019). *Deep Learning for Generic Object Detection: A Survey. International Journal of Computer Vision.* http://doi:10.1007/s11263-019-01247-4

[9] Zhang, Xiwu; Wang, Lei and Su, Yan (2020). *Visual Place Recognition: A Survey from Deep Learning Perspective. Pattern Recognition, 107760–*. http://doi:10.1016/j.patcog.2020.107760

[10] Şaban Öztürk, Emin Çelik, Tolga Çukur. (2023). Content-based medical image retrieval with opponent class adaptive margin loss. *Elsevier*. 637, pp.1-10. https://doi.org/10.1016/j.ins.2023.118938

[11] Image-Net Mini Dataset. Retrieved from http://cs231n.stanford.edu/tiny-imagenet-200.zip