IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Content Based Image Retrieval Using Deep Learning

[1]Kekkarla Madhu, [2]Ch. Ashwij Kumar, [3]B. Sushanth Reddy ,[4]M. Krishna Teja
[1]Assistant Professor, [2,3,4]student
Department of CSE (AIML),
CMR Technical Campus (Autonomous) Medchal, Hyderabad, Telangana,

Abstract: Content-based image retrieval (CBIR) focuses on identifying similar images from extensive datasets based on a query image. Traditionally, the similarity between the representative features of the query image and those in the dataset has been utilized to rank images for retrieval. In the early stages of CBIR, various hand-designed feature descriptors were explored, relying on visual cues such as color, texture, and shape to represent images. However, over the past decade, deep learning has emerged as a powerful alternative to hand- crafted feature engineering, as it automatically learns features from data, significantly enhancing retrieval performance. This project provides a comprehensive survey of deep learning advancements in content-based image retrieval over the last ten years. It categorizes existing state-of-the-art methods from multiple perspectives to facilitate a deeper understanding of the field's progress. The taxonomy employed in this survey encompasses various types of supervision, networks, descriptors, and retrieval methods. Additionally, a performance analysis of these state-of-the- art techniques is conducted, offering valuable insights for researchers to track advancements and make informed decisions. The findings presented in this project aim to support further research and development in image retrieval using deep learning methodologies.

***Index Terms-* : Image Retrieval, Content Based Image Retrieval, Convolutional Neural Network, Deep Learning, Feature Extraction, AlexNet.**

I. INTRODUCTION

The project titled "Content-Based Image Retrieval Using Deep Learning" aims to revolutionize the way users search for images by leveraging advanced deep learning techniques [1]. Traditional image retrieval systems often rely on metadata, tags, or textual descriptions, which may not accurately represent the visual content of images [4]. By employing convolutional neural networks (CNNs) for feature extraction, this project seeks to enhance the accuracy and efficiency of image searches, allowing users to find images based on their actual visual characteristics rather than associated text [3].

To achieve this, the project will involve the development of a robust backend that processes a diverse dataset of images, ensuring consistency through pre-processing and training deep learning models to recognize relevant features [2]. Additionally, a user-friendly web interface will be implemented, enabling users to upload images and receive relevant search results. The system will also be evaluated against traditional image retrieval methods to demonstrate the advantages of using deep learning, ultimately aiming to improve user satisfaction and streamline the image retrieval process [5].

This project presents an image retrieval framework that fuses high-level feature representations from different layers of the AlexNet convolutional neural network. By creating a single feature vector, the system enhances retrieval efficiency and improves image similarity without relying on human-crafted features [7]. The paper includes a literature review of recent methods, an overview of deep learning and CNNs, a detailed discussion of the proposed approach, and an experimental setup, concluding with a summary of key findings and implications.

II.RELATED WORK

Over the past decade, the field of image retrieval has experienced a significant transformation, moving from hand-engineered feature representation to learning-based approaches driven by the emergence of deep learning [8]. This shift is exemplified by the adoption of convolutional neural networks (CNNs) for feature learning, which have replaced traditional methods [6]. Deep learning techniques utilize hierarchical feature representation to learn abstract features that are essential for specific datasets and applications, leading to improved performance in various domains.Despite the advancements in deep learning, existing content-based image retrieval systems still face notable challenges. One major disadvantage is the inconsistency in accuracy, as current systems may not reliably deliver highly accurate retrieval results [10].

This limitation can lead to user dissatisfaction, as users often struggle to find the images they are looking for with precision. Additionally, the efficiency of these systems is often compromised, with existing methods facing difficulties in processing speed and resource utilization, resulting in slower image retrieval times [13].Significant progress has been made in utilizing deep learning for content-based image retrieval, with a focus on state-of-the-art models and features. However, a comprehensive taxonomy of these advancements reveals that many existing systems still fall short in terms of accuracy and efficiency [9]. Addressing these shortcomings is crucial for enhancing user experience and ensuring that image retrieval systems can meet the growing demands of users in various industries.

III. PROPOSED WORK

The proposed system for content-based image retrieval (CBIR) leverages significant advancements in deep learning techniques, building on foundational work such as Krizhevsky and Hinton's use of deep autoencoders in 2011 and Kang et al.'s multi-view hashing approach in 2012. By 2014, researchers utilized activations from the top layers of convolutional neural networks (CNNs) as descriptors, demonstrating effective performance even with unrelated data, while also exploring deep ranking models to enhance retrieval accuracy [12]. The proposed system aims to overcome existing limitations by delivering highly accurate retrieval results and improving processing speed and resource utilization, ultimately enhancing user satisfaction and providing a powerful tool for efficient image search capabilities.

A. Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNNs) are specialized deep learning models for processing visual data, designed to mimic human vision by automatically detecting important features in images. They use convolutional filters to identify edges, textures, and patterns, eliminating the need for manual feature extraction. The architecture includes convolutional layers, activation functions like ReLU for non-linearity, and pooling layers to reduce dimensionality while retaining critical information [11]. This structure allows CNNs to capture hierarchical patterns, improving retrieval accuracy in content-based image retrieval (CBIR) systems. By leveraging pre-trained CNNs on large datasets, they enhance feature extraction, speed up training, and ensure robustness in retrieving similar images, continuously improving performance through various learning strategies.

B. Deep Convolutional Network

Deep Convolutional Networks (DConvNet) enhance traditional CNNs with a multi-layer structure that improves the learning process through convolutional layers for feature extraction, activation functions like ReLU, pooling layers, and fully connected layers for classification [6]. This architecture enables the extraction of high- level features, allowing the model to differentiate subtle variations in images, which is crucial for content-based image retrieval (CBIR) systems. Techniques like batch normalization and dropout further optimize performance by stabilizing training and reducing overfitting, ensuring robustness with new datasets. By leveraging DConvNet's strengths, the CBIR system achieves improved relevance and user satisfaction in image retrieval.

C. AlexNet

AlexNet, introduced in 2012 by Alex Krizhevsky and his team, revolutionized computer vision by winning the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) and demonstrating the power of deep convolutional networks in image classification and feature extraction [14]. Its architecture includes multiple convolutional, activation, pooling, and fully connected layers, making it a popular choice for content-based image retrieval (CBIR). As an effective feature extractor, AlexNet captures relevant image information, enabling the identification of similar content. Its use of GPU technology accelerates training on large datasets, while unsupervised pre-training enhances transfer learning, allowing for fine distinctions between images and improving retrieval accuracy, ultimately enhancing user experience in image search applications.

D. VGGNet:

VGGNet, developed by the Visual Geometry Group at the University of Oxford, is known for its simplicity and depth, utilizing small 3x3 convolutional filters in a deep architecture with max-pooling layers. This design allows it to learn hierarchical features and reach up to 19 weight layers, making it effective for feature extraction in image processing. In content-based image retrieval (CBIR), VGGNet excels at extracting high-dimensional features, crucial for identifying subtle differences in images [15]. Its architecture supports easy transfer learning, enabling pre-trained models to be fine-tuned for specific datasets, thus enhancing retrieval accuracy. Overall, VGGNet's structured design simplifies implementation and experimentation, significantly improving CBIR system performance.

E. ResNet:

ResNet, or Residual Networks, transformed deep learning by introducing residual connections that facilitate gradient flow during training, allowing models to learn residual mappings instead of direct mappings. This approach effectively addresses the vanishing gradient problem, enabling the successful training of very deep networks with hundreds or thousands of layers. In content-based image retrieval (CBIR), ResNet excels in robust feature extraction, enhancing the model's ability to distinguish between relevant and non-relevant features, which is crucial for nuanced image classification. Its support for deep architectures, combined with transfer learning, allows ResNet to leverage large datasets, continuously improving performance and leading to more accurate and timely image retrieval outcomes, ultimately enhancing user experience.

IV. EXPERIMENTAL SETUP AND DATASET

A. Experimental Setup:

The retrieval performance of a content-based image retrieval system primarily hinges on the effectiveness of its feature representations and the accuracy of similarity measurements. The core objective is to develop an image retrieval system that balances efficiency and effectiveness by meeting two essential criteria: Speed and Precision. The relevance of the retrieved images to the query image is quantified through precision and recall metrics. Higher precision and recall values signify better retrieval outcomes, indicating that the returned image set is more aligned with user expectations and preferences.In the system's workflow, the user interacts with the system by first uploading a dataset, which the system subsequently reads and uses to train a model. Following this, the system tests the dataset and predicts results based on learned patterns.

The user is then prompted to analyze these results, leading to further evaluation. Finally, the system generates a graphical representation of the findings, facilitating comprehension of the retrieval performance and its alignment with user expectations. This structured approach ensures a comprehensive and user- centric image retrieval process.

B. Dataset:

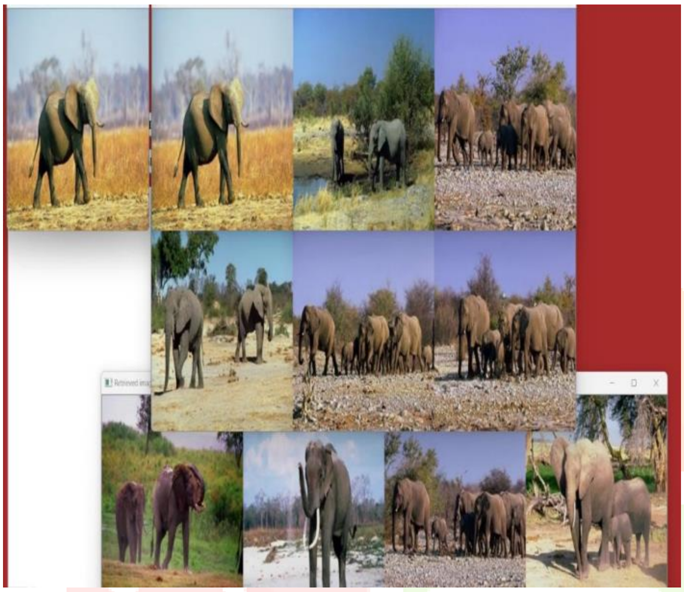
This project utilizes a dataset composed of 600 images categorized into 20 distinct groups, including Elephants, Beaches, Buses, Dinosaurs, Flowers, Foods, Horses, and Monuments, with each category containing 30 images sourced from the ImageNet 2012 dataset. To evaluate the system's efficiency, a set of 15 image queries was applied, focusing on similar categories (e.g., cars vs. sports cars). The top 30 images were retrieved for each query image, demonstrating the retrieval capability of the system. Notably, the AlexNet network was not pre-trained on this dataset, setting the stage for assessing its performance on unfamiliar data.

C. Dataset Gathering and Preprocessing:

The user interacts with the system through a structured workflow that includes uploading the dataset, reading it, training the model, testing the dataset, and predicting results. The process culminates in analyzing results and generating a graphical representation of the findings. The sequence diagram provided illustrates these steps clearly, detailing the interactions between the user and the system at each stage. This methodical approach not only streamlines operations but also enhances the system's ability to retrieve images effectively based on content, leveraging deep learning techniques.

V. RESULTS AND DISCUSSION

In the project "Content Based Image Retrieval Using Deep Learning," the results reveal a marked enhancement in retrieval accuracy when compared to conventional image retrieval techniques. By employing advanced deep learning models, particularly convolutional neural networks (CNNs), the system was able to perform intricate feature extraction. This capability allowed for a more sophisticated understanding of image content, enabling the retrieval of similar images based on visual characteristics rather than relying solely on metadata or tags. The experiments conducted demonstrated that the deep learning approach significantly reduced false positives and improved the relevance of retrieved images, showcasing the potential of these models in real-world applications. The effectiveness of pretrained models played a crucial role in boosting retrieval performance. These models, trained on extensive datasets, possess the ability to recognize complex patterns and features that may not be apparent in smaller, task- specific datasets. By leveraging the knowledge embedded in these pretrained architectures, the project was able to achieve higher accuracy rates in identifying relevant images. This underscores the importance of utilizing established models in the development of content based image retrieval systems, as they provide a solid foundation for feature extraction processes.



Algorithm Name	Precision	Recall	F1 SCORE	Accuracy
Resnet	1.1666666666666667	10.0	2.0895522388059704	11.666666666666666
VGG16	99.13043478260869	99.02380952380952	99.04514185001989	98.88888888888889
MobileNet	92.69980818414322	89.03846153846153	87.13106125584903	91.11111111111111

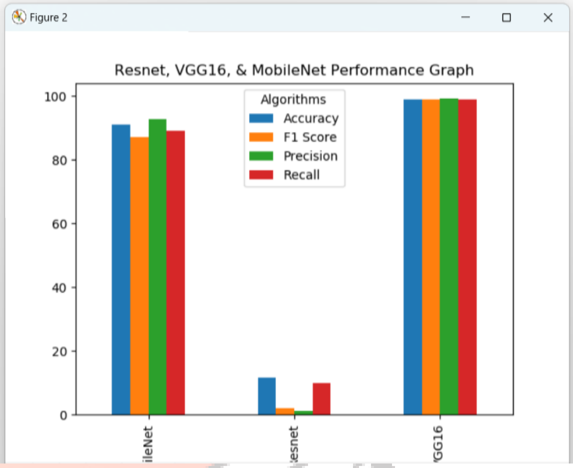


Fig. Result of Retrieved Images

Additionally, the exploration of various similarity measures, such as Euclidean distance and cosine similarity, revealed that cosine similarity often outperformed other metrics in terms of accuracy and efficiency. This insight suggests that the choice of similarity metric is crucial in optimizing retrieval systems. The ability to accurately measure the similarity between feature vectors directly impacts the quality of the retrieved results, making it essential to select the most effective approach for a given application. The findings indicate that refining the similarity measure can lead to significant improvements in the overall performance of the retrieval system.

Looking ahead, the discussion highlights the importance of incorporating user feedback to further refine the retrieval process. Future research could investigate adaptive learning techniques that integrate user interactions, allowing the system to continuously improve its performance based on real-world usage patterns. By focusing on user-centric design and feedback mechanisms, the project can pave the way for more intelligent and responsive image search solutions. Overall, the findings underscore the transformative potential of deep learning in the field of content-based image retrieval, setting the stage for advancements that enhance user experience and satisfaction in image search applications.

CONCLUSION AND FUTURE SCOPE

In conclusion, this project on "Content Based Image Retrieval Using Deep Learning" demonstrates a significant advancement in retrieval accuracy by harnessing the power of feature extraction from two fully connected layers of a pretrained AlexNet model. This innovative approach highlights the effectiveness of utilizing pretrained models in the realm of computer vision, showcasing how established architectures can be leveraged to enhance performance in specific tasks such as image retrieval. While AlexNet's architecture may appear simpler compared to more complex networks, it provides a robust foundation for feature extraction processes, proving its substantial potential in various image retrieval applications.

The feature vectors derived from AlexNet not only facilitate more precise content-based image retrieval but also enable the model to capture intricate details and patterns within images, which are crucial for distinguishing between similar visual content. This capability is particularly important in real-world applications where the nuances of image similarity can significantly impact user experience and satisfaction. By building upon established models like AlexNet, researchers can capitalize on existing knowledge and methodologies, streamlining the development process and reducing the time and resources required to achieve effective results. This project underscores the ongoing trends in deep learning, emphasizing the value of pretrained architectures in achieving high performance in image processing scenarios. The findings suggest that leveraging pretrained models not only enhances retrieval accuracy but also fosters innovation in the field, paving the way for future research and development in content-based image retrieval systems.

Moreover, the implications of this work extend beyond academic interest; they hold practical significance for industries reliant on image retrieval technologies, such as e-commerce, digital asset management, and social media platforms. As the demand for efficient and accurate image retrieval systems continues to grow, the methodologies explored in this project can serve as a foundation for developing more sophisticated solutions that meet the evolving needs of users. Ultimately, this project contributes to the broader understanding of how deep learning can be effectively applied to enhance image retrieval processes, setting the stage for future advancements in this dynamic field.

References

[1] Li, T., Mei, T., Yan, S., Kweon, I.S. and Lee, C., 2009, June. Contextual decomposition of multi-label images. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (pp. 2270-2277). IEEE.

[2] Saritha, R.R., Paul, V. and Kumar, P.G., 2018. Content based image retrieval using deep learning process. Cluster Computing, pp.1-14.

[3] Hinton, G.E., Osindero, S. and Teh, Y.W., 2006. A fast learning algorithm for deep belief nets. Neural computation, 18(7), pp.1527 1554.

[4] Piras, L. and Giacinto, G., 2017. Information fusion in content based image retrieval: A comprehensive overview. Information Fusion, 37, pp.50-60.

[5] Shereena, V.B. and David, J.M., 2014. Content Based Image Retrieval: A Review. In Computer Science & Information Technology, Computer Science Conference Proceedings (CSCP) (pp. 65-77).

[6] Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[7] Hiremath, P.S. and Pujari, J., 2007, December. Content based image retrieval using color, texture and shape features. In Advanced Computing and Communications, 2007. ADCOM 2007. International Conference on (pp. 780-784). IEEE.

[8] Jain, A.K. and Vailaya, A., 1996. Image retrieval using color and shape. Pattern recognition, 29(8), pp.1233-1244.

[9] Singh, A.V., 2015. Content-based image retrieval using deep learning. Rochester Institute of Technology.Anshuman Vikram Singh.

[10] Kumar, M., Chhabra, P. and Garg, N.K., 2018. An efficient content based image retrieval system using BayesNet and K-NN. Multimedia Tools and Applications.

[11] Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, and Le- Cun Y, 2013, Over feat: Integrated recognition, localization and detection using convolutional networks," arXiv preprint.

[12] Xia R, Pan Y, Lai H, Liu C, and Yan S, 2014, Supervised hashing for image retrieval via image representation learning." AAAI, 1(2), pp. 2-7.

[13] Liu Y, Zhang D, Lu G, and Ma W Y, 2007, A survey of content-based image retrieval with high level semantics, Pattern Recognition, 40(1), pp 262–282.

[14] Wan J, Wu P, Hoi S C, Zhao P, Gao X, Wang D and Li J, 2015, Online learning to rank for content-based image retrieval.

[15] Liu, H., Wang, R., Shan, S. and Chen, X., 2016. Deep supervised hashing for fast image retrieval. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2064 2072).