ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A DYNAMIC IMAGE RETRIEVAL FRAMEWORK BASED ON FUSION-BASED FEATURE EXTRACTION USING DEEP LEARNING

[ab]ALLA TALIB MOHSIN[1], [ac]MOHD SHAFRY BIN MOHD RAHIM[2]
[a]NORHAIDA BTE MOHD SUAIB[3], [d]ALI MUTTALEB HASAN[4], [b]SARAH ABDULKAREEM BAQER[5], [e]SAFA RIYADH WAHEED[6]
[a]Faculty of Computing, University Technology of Malaysia Johor Bahru, MALAYSIA.
[b]Faculty of education and sport sciences, University of Kufa, Iraq.
[c]Faculty of Computer & Information Technology, Sohar University, Sohar, Oman
[d]Ministry of Finance, General Authority for Customs
[e]Open College of Education, Ministry of Education, Najaf, Iraq

E-mail: [1]talab_alaa@yahoo.com,[2]shafry@utm.my[3], haida@utm.my,[4] alimatlab65@yahoo.com, [5]almosawisara200@gmail.com, [6]safa_albdeary@hotmail.com

ABSTRACT

Images are integral to human communication, and with the rapid growth of multimedia data, finding relevant images in large archives has become a significant challenge. Content-Based Image Retrieval (CBIR) offers a solution by retrieving visually similar images based on content rather than textual annotations. Despite its potential, CBIR systems face critical challenges, including irrelevant region detection, sensitivity to variations in brightness and size, and the absence of predefined class information in datasets. To address these issues, this study proposes a CBIR framework that integrates low- and high-level features such as texture, shape, and color for robust image representation. The framework employs wavelet-based Local Ternary Pattern (LTP) for texture extraction and incorporates a dynamic weight allocation mechanism, which adapts to statistical metrics like mean and variance to enhance retrieval accuracy. Comprehensive evaluations of the Corel-1k and Corel-10k datasets demonstrate the method's effectiveness in retrieving visually similar images with high precision. The proposed approach surpasses existing techniques, including CBIR-ANR, OMCBIR, and CNN-QCSO, in terms of precision, memory efficiency, and visual quality. The result of this work for the feature types included the dataset in the Retrieval Performance on the Corel 1K Dataset of fused features, proposing a precision of 88.1, recall of 80.5, and MAP of 88.3, and the Retrieval Performance on the Corel 10K Dataset of fused features proposed the precision of 83.7, recall of 76.2, and MAP of 83.9. This study establishes a promising direction for developing efficient CBIR systems capable of handling large-scale image datasets while improving retrieval performance.

Keywords: *Image Retrieval; CBIR; Features extraction; CNN; LTP.*

1. INTRODUCTION

Images have always played a crucial part in human communication since ancient times. Images enhance the user-friendliness and clarity of the communication medium. Technology has advanced to the point that more people are using cellophanes, cameras, and the Internet. With the increasing amount of shared and stored multimedia data, finding a relevant image in vast archives has become a difficult research problem. Such databases produce excellent results in commerce, academics, fashion, journalism, hospitals, historical research, government, surveillance, engineering, architecture, design, and crime prevention. Furthermore, the feature fusion substantially raises both memory and computational demands. Overwhelming computational burden may lead to delayed picture retrieval, particularly for extensive image databases or applications that need real-time processing. Optimized indexing methods or approximation search approaches may enhance retrieval time, but they may compromise accuracy or need more implementation complexity. The majority of feature extraction approaches exhibit high sensitivity to noise, particularly in areas with homogeneous characteristics, and face difficulties in complicated circumstances where subtle distinctions are crucial. Both noise robustness and the ability to capture small texture details are necessary to overcome different brightness and scaling levels. In the realm of large-scale, intricate, and high-performance image retrieval applications, deep learning CBIR emerges as the preferred option. When provided with a substantial dataset, deep learning models will far surpass conventional statistical approaches by effectively capturing overarching, conceptual characteristics. The deep models provide exceptional performance in terms of computation and accuracy. Combined with an appropriate framework, it can be used further to improve the retrieval performance of the whole model.

2. RESEARCH QUESTIONS

a. How can a feature-level fusion approach be developed to enhance the retrieval performance and robustness of CBIR systems capable of handling heterogeneous datasets?
b. How dynamic features of weighting based on statistical measures affect the accuracy and robustness of the CBIR system?
c. How to integrate handcrafted features and deep features for improving CBIR system?

3. RESEARCH OBJECTIVES

The performance and accuracy of an image retrieval system, this study focuses on the feature vector extraction process that reflects the visual image contents. The primary goals of the research were:

A. The feature-level fusion technique that integrates multiple visual features to improve retrieval accuracy in heterogeneous image datasets.
B. The effect of dynamic feature weighting is based on statistical mean and variance on the accuracy and robustness of CBIR systems.
C. Combine the dynamic weighted features with deep learning-based features for improving the CBIR system.

4. LITERATURE REVIEW

The process of merging features from several sources might result in data with a large number of dimensions. To handle the issue of dimensionality and computing complexity, it may be necessary to use techniques such as PCA or t-SNE. To prevent any one sort of feature from overpowering the fusion process, it is necessary to normalize or standardize features that have diverse scales and units. Features from various modalities may possess diverse levels of significance for retrieval. To achieve successful fusion, it is necessary to calculate the suitable weights for each feature type, taking into consideration their impact on the total retrieval performance. Developing a successful approach for merging characteristics and making judgements on retrieval may be difficult and often involve trial and error and expertise in the field.

Table 1: A critical review of feature extraction techniques in CBIR

Feature Category	Feature Type	Feature Technique	Advantage	Disadvantage
Global Features	Color Feature	Color Histogram	simple to use and fast computation	lost spatial information and no color similarity
		Color Correlogram	provide spatial information	slow computation and high dimensionality
		Color Co-occurrence Matrix	effective texture representation, captures the color variations, sensitive to spatial information, directional sensitivity	high computational complexity, sensitivity to noise, difficulty in capturing high-level semantics, inadequate for complex or natural images, limited rotation and scale invariance, redundancy in features
	Texture Feature	Tamura Feature	generates the histogram for texture features	represent the specific types of texture

		Steerable pyramid	rotation invariant	more computation and storage
		Wavelet Transform	less retrieval time	poor performance
		Gabor Wavelet Transform	achieves the highest retrieval results	computationally intensive
	Shape Features	Boundary-Based Features	boundary-based features, invariance to geometric and photometric transformations, easy to implement	insensitive to large variations in scale and rotation, high computational cost, not optimal for small objects
		Region-Based Features	robust to noise and small variations, useful for segmentation-based analysis,	dependency on accurate segmentation, high computational complexity, inflexibility with variations in object size and orientation
Local Features	Gradients based methods	Scale-invariant feature transform (SIFT)	scale-invariant, rotation-invariant, and lighting-invariant.	computationally intensive, unsuitable for real-time applications.
		Gradient location and orientation histogram	improved dispersion of the characteristics, enhanced precision in spatial representation, invariant to scale and rotation, making it resistant to alterations in image size and orientation.	noise-sensitive, computationally difficult, demanding in terms of processing time and memory, and not well-suited for real-time applications.
		Speed-Up Features (SURF)	invariant to scale and rotation, well-suited for real-time applications, less affected by picture noise and blurring.	lacking complete rotation invariance
	Intensity based methods	Local binary pattern (LBP)	computational simplicity and efficiency, rotation invariance, illumination invariance, easy to implement, low dimensionality	sensitivity to noise, not scale invariant

5. IMAGE DATASET

The suggested image retrieval approach is validated using three standard datasets. The Corel-1000 database has 1000 photos categorized into ten groups: African tribes, Beaches, Buildings, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains, and Foods (see to Figure 1) [1]. Each category consists of 100 photos, each with dimensions of either 384 × 256 or 256 × 384. Each picture in this database is captured to query. The photos are categorized into idea groups, with 90 images allocated for training and 10 images for testing in each group.

The retrieval accuracy has been improved throughout the Corel-1k, Corel-10k, Corel1k scale, and Corel-1k illumination datasets. The Corel-10k collection has 100 distinct categories, with a wide range of subjects like sunsets, beaches, flowers, buildings, cars, horses, mountains, fish, cuisine, doors, and more. It consists of a total of 10,000 photos. Every category consists of 100 photographs with dimensions of either 192×128 or 128×192 in the JPEG format.

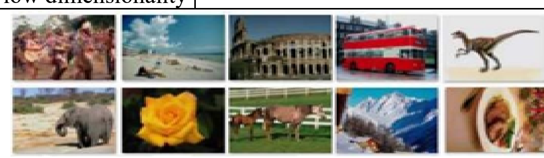


Figure 1. Sample Images From Each Class Of Corel-1k Dataset [1].

The assessments are performed using Corel-1k and Corel-10k datasets, each consisting of 1000 photos from 10 distinct categories and 10,000 images from 100 distinct categories, respectively. The effectiveness of the suggested approach is additionally confirmed by testing on the Corel-1k Scale and Corel-1k Illumination datasets. The Corel-lk illumination and Corel-lk scale datasets are created by modifying the channels of the first 20 images from each class in the Corel-1k dataset. Factors of -60, -30, 0, 30, and 60 are added to the channels for the Corel-lk illumination dataset. For the Corel-lk scale dataset, the first 20 images from each class in the Corel1k dataset are scaled at different scales of 0.5, 0.75, 1, 1.25, and 1.5 respectively. The dataset was queried with each picture, and the top 20 photos with the smallest distance were obtained. Every picture in the collection is used as a query image, and the most comparable photos are obtained.

6. PROPOSED FEATURE EXTRACTION BASED CNN

The proposed scheme combines diverse features to create a comprehensive feature representation of images. The feature vectors extracted from LTP, DWT, CH, and EH were combined into a single, unified feature vector before performing similarity matching. This allows the system to work with a rich representation of the image's content and improves retrieval performance by leveraging the strengths of each feature type. By considering multiple aspects of the image (texture, color, shape, and frequency content) [2], the system can more accurately find relevant images. The algorithm does a comparison between the features of the input picture and the features of the photos in the database. The most comparable photos to the input image are fetched from the database in the final step. The user can modify the quantity of photos that are obtained. The images that have the highest degree of similarity with the supplied picture will be shown at the top of the list. Figure 2 shows an image that is obtained from the database and retrieved using the suggested approach.

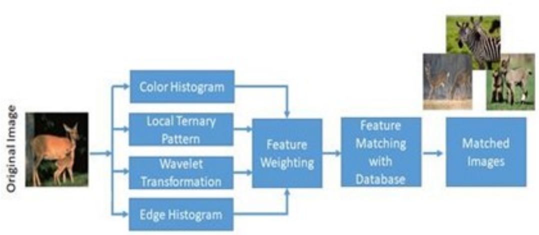


Figure 2. Flowchart Of The Proposed CBIR Framework

By fusing different feature benefits from the complementary strengths of these features, leading to several advantages, such as improved retrieval accuracy, increased robustness and enhanced versatility. The system becomes more robust to variations in lighting, noise, image scale, and rotation, as different features provide resilience to different types of distortions.

This approach effectively fuses different kinds of features —LTP for texture, DWT for multi-scale frequency analysis, CH for color distribution, and EH for shape and structure detection. By combining these features, the technique captures multiple visual aspects of images. Its distinctiveness comes from its ability to dynamically adjust feature relevance on a per-query basis, providing more accurate and context sensitive retrieval results. The feature fusion in CBIR systems results in a powerful multi-feature approach that improves retrieval effectiveness and robustness. However, different feature types may yield scores on different scales, and improper normalization can lead to biased retrieval. Therefore, the system must be optimized for different use cases, balancing performance and computational efficiency. The major strength of this technique is dynamic weight assignment based on query relevance. This technique dynamically assigns higher weights to features for the database image class that are more relevant for a given query instead of treating all features equally for every image class and query. For example, if color is a defining characteristic for one query, the Color Histogram will receive a higher weight. If texture is more important for another query, LTP or DWT will dominate. This adaptivity ensures that the CBIR system tailors its retrieval strategy to the specific needs of the query, unlike static systems where the importance of each feature is pre-set. Overall, the effectiveness of this CBIR system lies in its comprehensive multifeatured fusion strategy, the incorporation of statistical measures for dynamic weighting, and its robust similarity calculation methods, all tailored to enhance retrieval performance and adaptability to diverse datasets. These elements collectively contribute to improved accuracy and efficiency in the image retrieval process. Here's a breakdown for each subroutine.

7. MULTI-FEATURE FUSION APPROACH

The proposed system utilizes a fusion of four distinct feature extraction methods: Local Ternary Patterns LTP for texture, DWT for multi-scale frequency analysis, CH for color distribution, and EH for shape and structure detection. This multi-faceted approach allows for a more comprehensive representation of images, addressing various visual aspects that are crucial for accurate retrieval. The flow diagram of this module is given in Figure 3.

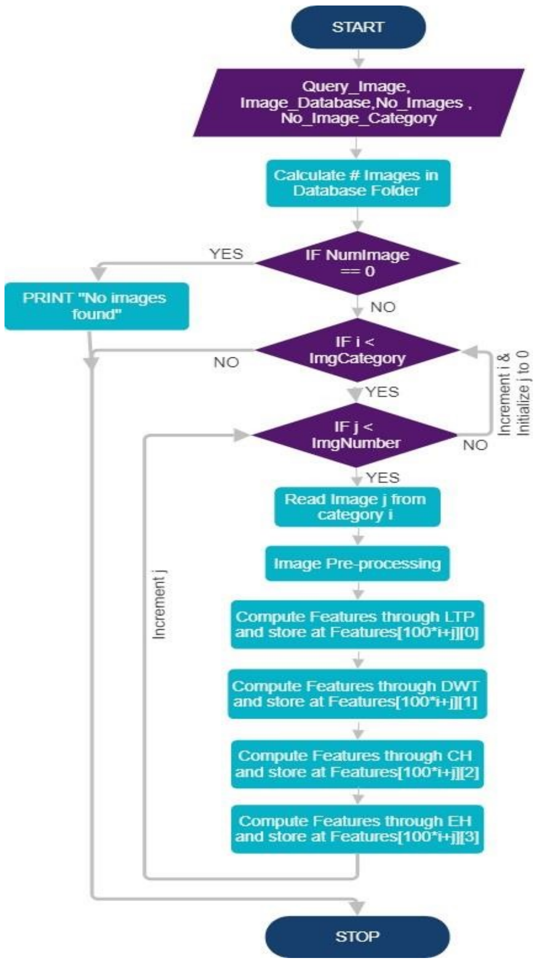


Figure 3. Flow Diagram Of Feature Extraction Subroutine

8. CENTROID OF VISUAL FEATURES COMPUTATION

The centroid of visual features is a single feature vector representing the "center" of a set of feature vectors within a class. It captures the central characteristics of the visual features across the class. A feature centroid represents the average feature vector for all images in a particular class or category. It enables the system to treat the entire class as a single entity by encapsulating its essential features [3], simplifying comparisons. The class centroid can quickly indicate the query's similarity to that class. The centroid is less sensitive to outliers and minor variations in feature vectors compared to comparing each individual image. Thus, it increases accuracy when categorizing a query image or retrieving similar images. The flow diagram of this module is given in Figure 4.

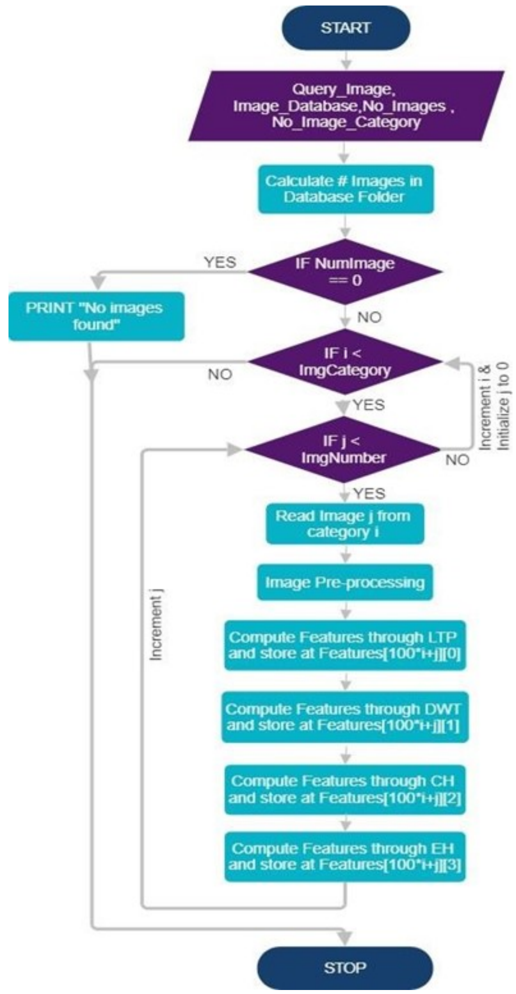


Figure 4. Flow Diagram Of Calculating Centroid For Each Class.

This is a structured way to understand the feature distribution within each class, enhancing the retrieval process by establishing a clear reference point (centroid) for comparison. The centroid acts as a prototype or representative for a class. By computing the centroid, you get a single feature vector that summarizes the characteristics of all the images in that class. This reduces the need to compare a query image with every image in the class, improving retrieval efficiency. The centroid provides an overall description of the class's characteristics, ensuring that the retrieved images reflect the general visual properties of the class, which might be missed by individual image comparisons. Some individual images may have noise or outliers that deviate significantly from the main visual characteristics of their class. The centroid, being the average of all images in a class, can mitigate the effect of such outliers, leading to more robust retrieval results.

9. MEAN AND VARIANCE OF IMAGE CLASSES

The use of mean and variance for each class in combination with normalized Euclidean distances is a unique aspect that provides additional context to the feature distances. Contributes in centroid-based CBIR system:

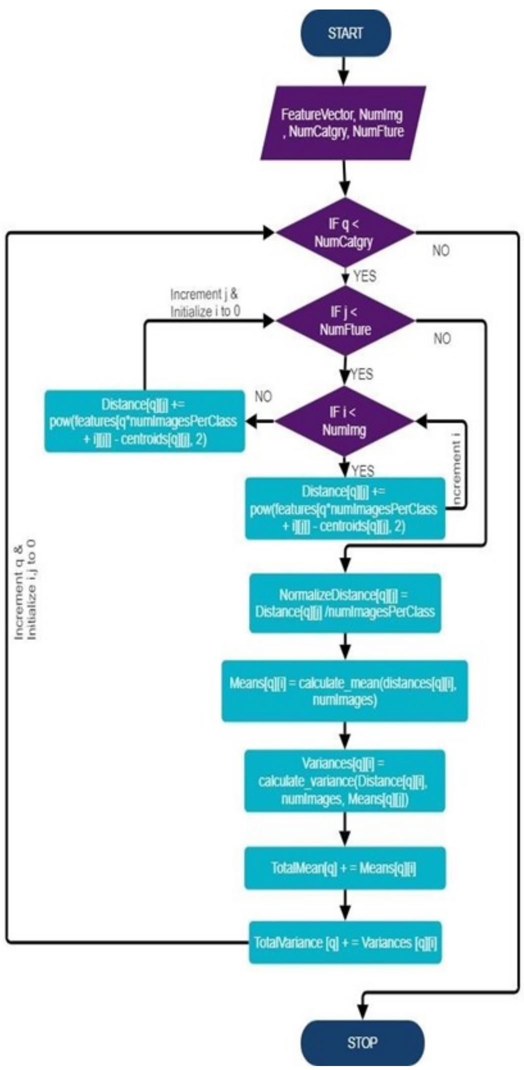


Figure 5. Flow Diagram Of Calculating Mean And Variance For Each Class By Evaluating

The statistical characteristics of the feature distributions, the system can better assess the relevance and similarity of images. Incorporating mean and variance calculations for each class and using them with normalized Euclidean distances adds an additional layer of statistical insight to image retrieval. The mean provides the central tendency of features within each class, while the variance measures how much the features of the images in a class deviate from the mean. This helps in determining not only the proximity (using normalized Euclidean distance) but also the consistency of the feature distributions [4]. Figure 5 and Figure 6. Show flow diagrams that combine mean, variance, and normalized Euclidean distance that contributes in centroid-based CBIR system.

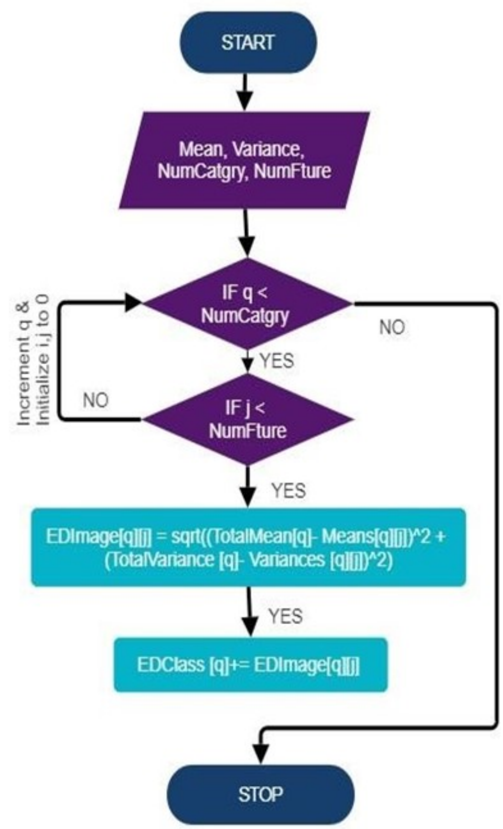


Figure 6. Flow Diagram Of Calculating Normalized Euclidean Distance Of Each Class

10. DYNAMIC FEATURE WEIGHTING

By calculating weights for each feature vector belonging to a specific class (ensuring the sum of all weights equals 1) [5], the system dynamically adjusts the influence of different features based on their statistical properties. This weighting mechanism enhances the model's adaptability to various types of queries and improves the robustness of the retrieval process. The flow diagram of this module is given in Figure 7.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

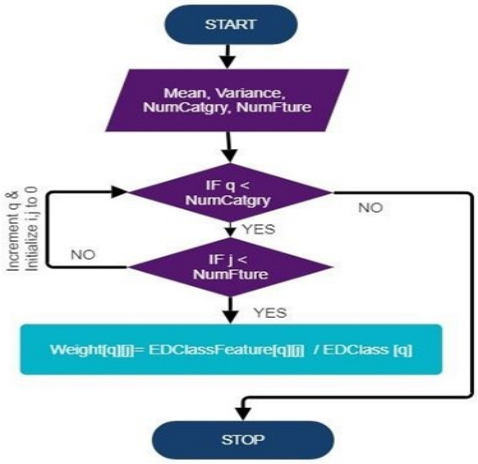


Figure 7. Flow Diagram Of Calculating Dynamic Weight For Each Feature Of The Classes

This approach adjusts each feature dimension by its variance, so features with less variability are given more weight. Lower Euclidean distance suggests that the feature vectors in the class are closer to each other, which can imply a more defined class structure. Higher Euclidean distance may suggest that the class has more variability, potentially indicating a broader or less cohesive grouping. This analysis is beneficial in CBIR as it provides both the central tendency (centroid) and the dispersion (variance-adjusted distance) of feature vectors, enhancing the understanding of class similarity and structure.

11. IMAGE RANKING AND RETRIEVAL

The method includes a systematic approach to sorting the computed distances in ascending order, allowing for efficient extraction of the top 20 results. This step ensures that users receive the most relevant images promptly, enhancing user experience. The flow diagram of this module is given in Figure 8. The combination of these methodologies makes the CBIR system robust against variations in image quality, illumination, and noise, which are common in heterogeneous datasets. This adaptability is crucial for real-world applications where image conditions can vary significantly. Dynamic weighting enhances retrieval performance by allowing the system to adaptively emphasize or de-emphasize certain features based on the query, class statistics, or relevance feedback. This tailored approach to feature weighting enables the CBIR system to deliver more accurate, relevant, and robust search results, ultimately improving the user experience by providing images that better match the query's intent.

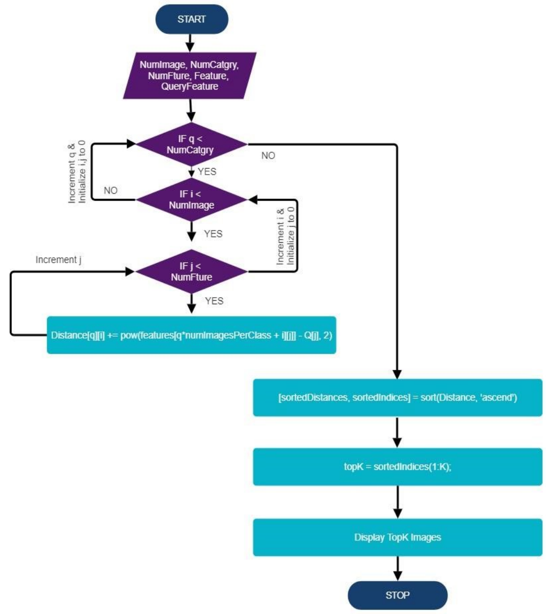


Figure 8. Flow Diagram Of Image Ranking And Retrieval

12. RESULTS AND DISCUSSION

A comparison of the proposed technique has also been done over Corel1k and Corel-10k. Given each query image, the top 20 matching images have been retrieved from the dataset using minimum distance. Each image of the dataset is used as a query image and most similar images are retrieved [6]. The distance computation between dataset images and query images is according to the histogram intersection distance metric. If the given query and the images retrieved relate to a similar class then the retrieval is considered to be accurate otherwise it will be regarded as incorrect. The performance assessment of the proposed technique has been done using recall and precision. The proposed framework's retrieving performance of top 24 bus images from the Corel-1K datasets was compared with three four earlier and state-of-the-art methods: CBIR-ANR (Content-Based Image Retrieval using Adaptive Neighborhood Relevance) [6], OMCBIR (Optimized Multi-Feature CBIR) [7], and CNN-QCSO (Convolutional Neural Network with Quantum-Chaotic Swarm Optimization) [8]. Visual results from Figure 9 clearly illustrate the eminence of the proposed approach over others in the form of its retrieval results. The proposed framework demonstrates high precision, retrieving images with strong similarities in color, texture, and structure to bus images within the top 24. Fusion of LTP, DWT, color histogram, and edge histogram effectively captures both global and local image characteristics,

Journal of Theoretical and Applied Information Technology
31st May 2025. Vol.103. No.10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

yielding results that accurately reflect bus image attributes

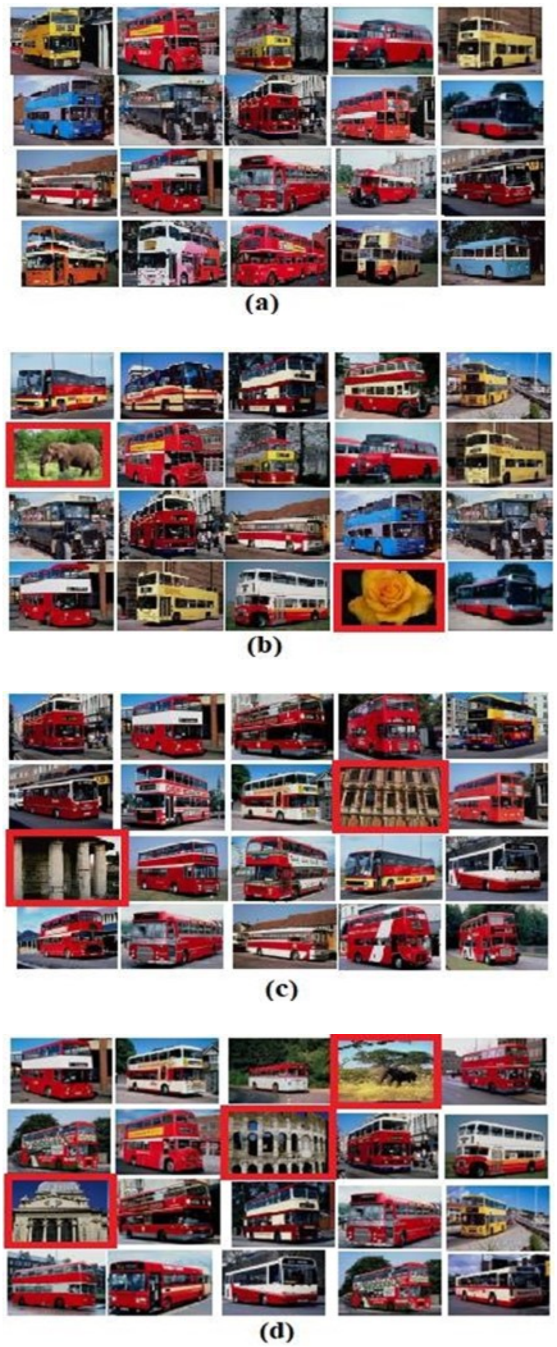


Figure 9. Visual Comparison For The Retrieval Of Top 24 Bus Images From Corel-1k Dataset (A) Proposed Framework (B) CBIR-ANR (C) OMCBIR (D), And CNNQCSO.

In analyzing a chart that compares image retrieval performance for different numbers of retrieved images across proposed framework, CBIR-ANR [6], OMCBIR [7], and CNN-QCSO [8] from the Corel-1K dataset. As the number of retrieved images increases (e.g., from top-10 to top-20), most retrieval systems show a decline in precision. This happens because the initial top results often contain the most relevant images, while results become more varied as more images are retrieved. The proposed framework maintains a higher precision with increasing numbers of images than the other methods, it suggests better generalization in retrieving relevant images.

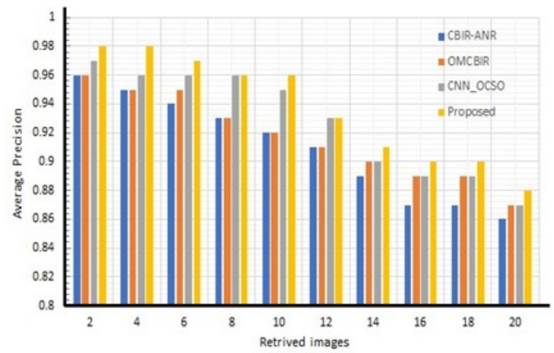


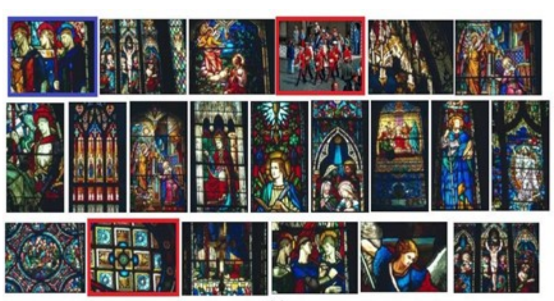
Figure 10. Comparison Of Existing Techniques With The Proposed Technique Using Number Of Retrieved Images And Precision Over Corel-1k Dataset

As the proposed framework consistently performs better than CBIR-ANR, OMCBIR, and CNN-QCSO across all retrieval counts, this would demonstrate its robustness in extracting and matching features that better align with the dataset's structure. CNN-QCSO and proposed framework often closely matched. CNN-QCSO image features effectively in complex datasets but may require fine-tuning for domain specific image sets like Corel-1K. Furthermore, it also suggests that the proposed system utilizes feature extraction techniques that are more aligned with this type of content. A strong result for the proposed framework in terms of precision stability and performance consistency across varying retrieval counts, would indicate a significant improvement over previous methods, suggesting a more effective approach to feature extraction and similarity measurement tailored to the Corel-1K dataset's characteristics. The Corel-10K dataset is a large image collection containing 10,000 images across various categories including landscapes, animals, vehicles, buildings, and more abstract or pattern-based images, covering a broad spectrum of visual features (color, texture, shape, and composition). This diversity increases the difficulty of image retrieval, as images from different classes may share similar visual characteristics (e.g., colors or textures) [9], leading

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to potential overlap. Thus, it is a more extensive and challenging dataset compared to Corel-1K, as it includes a broader array of categories with greater intra-class variability, making it ideal for evaluating the scalability and robustness of image retrieval algorithms. Visual results from Figure 10 clearly illustrate the eminence of the proposed approach over others in the form of its retrieval results from Corel-10K dataset. The proposed framework effectively retrieves semantically relevant images with high precision (0.9) due to its multi-feature fusion approach where each feature type contributes to a more nuanced understanding of image characteristics, which enhances semantic matching [19]. The dynamic weighting within the framework allows it to consistently identify and retrieve the most relevant images. By leveraging comprehensive feature information, it minimizes false positives in the top 20 images.

CBIR-ANR uses an adaptive neighborhood relevance approach, which performs well on smaller datasets but is challenged by the diversity and scale of Corel10K. Relevance: It can retrieve semantically similar images within a neighborhood relevance model; however, it tends to retrieve visually similar images rather than semantically relevant ones across larger categories. Figure 11 shows five mismatched images out of 20 semantic images from Corel-10k dataset [20]. Two images out of 5 mismatched images were similar but non-target classes in the top 20. The fixed neighborhood approach of CBIR-ANR becomes less adaptable for categories with high intra-class variability or overlapping features, leading to reduced precision. As a result, CBIR-ANR retrieve more false positives compared to the proposed framework, especially in categories with subtle distinctions from Corel-10K dataset.


(a)


(b)


(c)


(d)

Figure 11. Visual Comparison For The Retrieval Of Top 20 Semantic Images From Corel-10k Dataset (A) Proposed Framework (B) CBIR-ANR (C) OMCBIR (D), And CNN-QCSO

OMCBIR optimizes the use of multiple features for retrieval, focusing on color and edge information, which helps improve retrieval accuracy for distinct and visually unique categories. OMCBIR retrieves relevant images for visually distinct categories [21] (e.g., vibrant colors or unique textures), it may struggle with categories where color or edge features alone are insufficient for capturing the semantic content. Thus, while optimized for certain features, it may miss finer semantic details. OMCBIR's feature optimization reduces adaptability, as it doesn't dynamically adjust to individual category traits.

CNN-QCSO leverages CNNs for deep feature extraction, which excels in capturing intricate patterns and semantic meaning in images, combined with QCSO for search optimization. This framework retrieves semantically rich images with high relevance, thanks to the CNN's feature

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

extraction capabilities [22]. CNN-QCSO often captures deeper contextual details, making it highly effective for complex categories in Corel-10K. Proposed Framework yields best balances precision and high relevance on both Corel-10K: Dynamic weighting and multi-feature fusion help it retrieve relevant bus images with minimal false positives. CBIR-ANR: Performs reasonably well for smaller datasets but lacks adaptability for large-scale datasets, resulting in lower precision on Corel-10K. The proposed framework provides a holistic representation of the representation of the holistic image, leading to higher precision and relevance in retrieval results. This is particularly beneficial for large and diverse datasets like Corel-10K, where single-feature approaches and basic fusion methods often fall short [23]. The proposed framework still maintains relevance but may encounter slight decreases in precision due to dataset size and diversity. However, its dynamic weighting mechanism allows it to adjust feature importance, making it resilient to the dataset's scale, leading to a high number of correctly retrieved images in the top 24 semantic images.

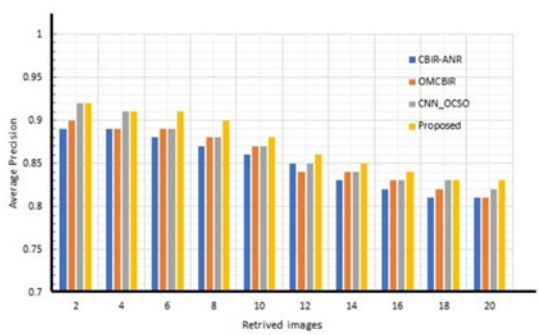


Figure 12. Comparison Of Existing Techniques With The Proposed Technique Using Number Of Retrieved Images And Precision Over Corel-10k Dataset

13. PROPOSED MODEL

The proposed CBIR model. The proposed framework employs a hybrid approach to image retrieval, integrating features extracted using pretrained CNNs based features with dynamically weighted handcrafted features. This integration aims to leverage the complementary strengths of both feature types to enhance retrieval accuracy and robustness across diverse datasets. First, the feature extraction process begins with pretrained CNNs, specifically VGG-16, Mobile Net, and ResNet-50. Second, handcrafted features are used to complement the deep features obtained from CNNs.

The proposed CBIR framework consists of three main components:

1. **Feature Extraction Using CNNs**: Deep features are extracted from VGG-16, MobileNet, and ResNet-50 pretrained models.
2. **Weighted Handcrafted Features**: handcrafted features such as Local Binary Patterns (LBP), and color quantization are fused and weighted based on their contribution to retrieval performance.
3. **Feature Integration and Retrieval**: Deep features and weighted handcrafted features are combined using a feature fusion strategy, and similar measures are applied to retrieve relevant images.

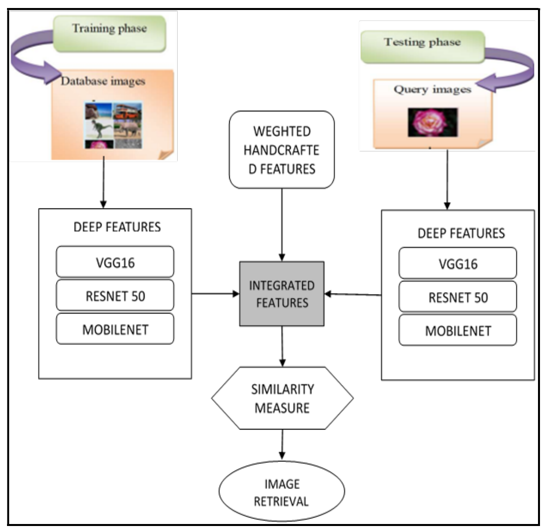


Figure 13: Proposed Deep Learning Based CBIR Model

A. Feature Extraction with Pretrained CNNs

In modern CBIR systems, pre-trained CNN models such as ResNet-50, VGG-16, and VGG-19 have proven invaluable for extracting high-level features from images. These models, trained on large-scale datasets like ImageNet, possess a rich repository of hierarchical visual knowledge, enabling them to recognize complex patterns and semantic relationships.

B. Feature Integration

The weight of handcrafted features are concatenated with the CNN features to create a unified feature vector that encapsulates both low-level and high-level information. The deep features and weighted handcrafted features are combined using a feature fusion strategy, and similar measures are applied to retrieve relevant images as follows:

$$F_{integ} = \alpha \cdot F_{deep} + \beta . F_{handcrafted}$$

Where F_{integ} is the final integrated feature vector,

F_{deep} deep feature vector extracted from CNNs.

$F_{handcrafted}$ denote the weighted handcrafted

features. α and β is the fusion weights determine

experimentally.

C. *Similarity Measurement and retrieval*

The similarity measurement and retrieval process is carried out using Euclidean Distance (ED) as the primary metric for assessing the similarity between feature vectors. The Euclidean Distance quantifies the geometric distance between two vectors in the feature space, providing a straightforward and effective means of comparing images. A smaller Euclidean Distance indicates a higher similarity between the query image and the retrieved images, while a larger distance suggests a greater dissimilarity. The similarity measure can be computed as follows:

$$\text{Similarity } (F_1, F_2) = \frac{F_1 \cdot F_2}{||F_1|| \cdot ||F_2||}$$

F_1 *and* F_2 represent the fused feature vectors of the query image and a database image. $||F_1|| \cdot ||F_2||$ denote the norms of the respective feature vectors.

D. *Evaluation*

To evaluate the proposed model, Precision, recall and mean average precision (mAP) have been used which can be computed as follows:

$$Precision@K = \frac{Nuber\ of\ relevant\ images\ in\ Top - K}{K}$$

$$Recall@K = \frac{Nuber\ of\ relevant\ images\ in\ Top - K}{Total\ Number\ of\ relevant\ images}$$

$$mAP = \frac{1}{Q} \sum_{Q=1}^{Q} \frac{\sum_{k=1}^{K} P_k . rel_k}{Total\ \ relevant\ images\ for\ query\ q}$$

Where Q is the total number of queries, P_k is the precision at rank k, rel_k is relevant indicator for image at rank (1 if relevant, 0 otherwise)

14. RESULTS AND DISCUSSION

The evaluation of the proposed CBIR framework is conducted using standard metrics including precision, recall, F1score, **mAP** and retrieval time to assess classification and retrieval performance. Two datasets were used which include Corel 1K and Corel 10K. Table 5.1 shows the comparison results of the proposed model compared to the baseline methods.

Table 2: Retrieval Performance on Corel 1K Dataset

Feature Type	**Precision@10 (%)**	**Recall@10 (%)**	**mAP (%)**
Handcrafted Features Only	72.5	65.2	70.3
VGG-16	78.4	72.1	79
MobileNet	76.9	70.5	77.5
ResNet-50	81	74.6	81.2
Chen (2022)	83.2	76	84.5
Fused Features (Proposed)	**88.1**	**80.5**	**88.3**

Table 3: Retrieval Performance on Corel 10K Dataset

Feature Type	**Precision@10 (%)**	**Recall@10 (%)**	**mAP (%)**
Handcrafted Features Only	62.1	58.4	65
VGG-16	72.3	67.8	74.5
MobileNet	70.8	66	72.2
ResNet-50	76.2	70.9	76.5
He (2020)	74.5	68.8	76.2
Fused Features (Proposed)	**83.7**	**76.2**	**83.9**

15. CONCLUSION

In this paper, we proposed an advanced CBIR framework that combines multiple visual features, including LTP, DWT, color histogram, and edge histogram, to enhance image retrieval performance. The fusion of these features enables a comprehensive representation of image content, capturing both global and local characteristics. A dynamic feature weighting mechanism was developed to adaptively emphasize significant

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

features based on statistical properties, improving retrieval robustness and accuracy. The framework was evaluated on Corel-1K and Corel-10K datasets, demonstrating superior precision and recall compared to state-of-the-art methods such as CBIR-ANR, OMCBIR, and CNN-QCSO. The novelty of this works the image representation that includes several feature types That efficacy of image retrieval can be improved by having access to several feature types and the merged using CBIR. The goal of this study is to build image retrieval, based on LTP and statistical weighting of parameter, which integrates the image information that is classified as color or texture features. This will improve the efficacy of the similarity approaches and CBIR system. Moreover, an automated feature vector format for images may be introduced by this technique. Ultimately, the efficacy of the image representation and retrieval technique may be improved by having access to the integrated various feature types as they become accessible. It is thought that the new system is capable of retrieving images that are comparable to the image query. The developed algorithms seem to surpass several current state-of-the-art CBIR techniques and achieve exceptional results in terms of retrieval accuracy.

REFERENCES:

[1] Pradhan, J., Pal, A. K. & Banka, H. 2022. A Cbir System Based On Saliency Driven Local Image Features And Multi Orientation Texture Features. Journal Of Visual Communication And Image Representation, 83, 103396.

[2] Sukhia, K. N., Ali, S. S., Riaz, M. M., Ghafoor, A. & Amin, B. 2021. Content-Based Image Retrieval Using Angles Across Scales. Ieee Geoscience And Remote Sensing Letters, 19, 1-5.

[3] Swapna, M. & Hegde, N. Noise Removal Filtering Methods For Mammogram Breast Images. International Conference On Communications And Cyber Physical Engineering 2018, 2023. Springer, 1033-1046.

[4] Taheri, F., Rahbar, K. & Salimi, P. 2023. Effective Features In Content-Based Image Retrieval From A Combination Of Low-Level Features And Deep Boltzmann Machine. Multimedia Tools And Applications, 82, 37959-37982.

[5] Yadav, S. P. & Yadav, S. 2020. Image Fusion Using Hybrid Methods In Multimodality Medical Images. Medical & Biological Engineering & Computing, 58, 669-687.

[6] Vieira, G. S., Fonseca, A. U. & Soares, F. 2023. Cbir-Anr: A Contentbased Image Retrieval With Accuracy Noise Reduction. Software Impacts, 15, 100486.

[7] Zhang, X., Bai, C. & Kpalma, K. 2023. Omcbir: Offline Mobile Contentbased Image Retrieval With Lightweight Cnn Optimization. Displays, 76, 102355.

[8] Kumar, S. N. & Kumar, C. S. 2023. Fusion Of Cnn-Qcso For Content Based Image Retrieval. Journal Of Advances In Information Technology, 14, 668673.

[9] Mehidi, I. Pre-Processing Techniques: A Review For Retinal Image Segmentation. 2024 2nd International Conference On Electrical Engineering And Automatic Control (Iceeac), 2024. Ieee, 1-6.

[10] Adnan, M. M., Rahim, M. S. M., Al-Jawaheri, K., Ali, M. H., Waheed, S. R., & Radie, A. H. (2020, September). A Survey And Analysis On Image Annotation. In 2020 3rd International Conference On Engineering Technology And Its Applications (Iiceta) (Pp. 203-208). Ieee.

[11] Singh, B., Kaur, E. K. & Choudhary, C. A Machine Learning Model For Content-Based Image Retrieval. 2023 2nd International Conference For Innovation In Technology (Inocon), 2023. Ieee, 1-6.

[12] S. R. Waheed, M. S. M. Rahim, N. M. Suaib, And A. Salim, "Cnn Deep Learning-Based Image To Vector Depiction," Multimedia Tools And Applications, Pp. 1-20, 2023.

[13] Zhang, X., Bai, C. & Kpalma, K. 2023. Omcbir: Offline Mobile Contentbased Image Retrieval With Lightweight Cnn Optimization. Displays, 76, 102355.

[14] Zhang, Y., Qian, X., Tan, X., Han, J. & Tang, Y. 2016. Sketch-Based Image Retrieval By Salient Contour Reinforcement. Ieee Transactions On Multimedia, 18, 1604-1615.