International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

# Designing an Efficient Multi-Domain Feature Analysis Engine with Incremental Learning for Continuous Feedback-Based Image Retrieval.

**Milind Vijayrao Lande[1,a)] and Sonali Ridhorkar[2,b)]**

**Submitted**: 15/07/2023 **Revised**: 05/09/2023 **Accepted**: 24/09/2023

**Abstract**: Due to its applications in a variety of fields, including healthcare, surveillance, and e-commerce, content-based image retrieval (CBIR) has recently attracted a lot of attention. By introducing an effective multidomain feature analysis engine that includes incremental learning and continuous feedback operations, this paper introduces a novel method for CBIR. The suggested framework combines a variety of feature extraction methods, such as Fourier, Entropy, Color Map, S, Z, Laplace, GRU, and LSTM, to extract key visual characteristics from images and samples. By effectively maximizing feature variance levels, an Elephant Herding Optimizer (EHO) is used to increase the discriminative power of the chosen features. The automatic selection of the most informative features is made possible by the EHO algorithm, improving retrieval performance levels. The use of a Vector Autoregressive Moving Average (VARMA) model, which successfully captures the temporal dependencies and correlations within the image dataset, further enhances the CBIR process. This model greatly improves retrieval accuracy levels and makes it easier to predict relevant images using the extracted features. Additionally, correlation learning operations are used to incorporate feedback learning, which enables the CBIR system to change and advance over time in response to user feedback. As a result of the system's use of the feedback data, the retrieval process has been improved, yielding better precision, recall, and accuracy values for different datasets & samples. These include ImageNet, MIRFLICKR, and CIFAR-10 as well as various real-world image samples were used in the experimental evaluations. The findings show that the suggested approach outperforms current CBIR models with precision of 0.98, recall of 0.985, and MAP of 0.97 for different use cases. Further, the proposed system achieves superior accuracy scores and displays lower delay values for real-time scenarios. The increasing demand for accurate and effective image retrieval systems across a variety of applications drives the necessity for this work under real-time scenarios. The proposed framework addresses the limitations of existing CBIR models by combining multiple feature extraction techniques, optimizing feature variance, utilizing temporal dependencies, and incorporating continuous feedback learning process. The potential applications of this research are diverse and include retrieval of medical images for diagnosis and treatment, effective image analysis through surveillance systems, and personalized product recommendations through e-commerce platforms. The proposed approach has several advantages over current CBIR models, including improved retrieval accuracy, adaptability to shifting user preferences, and improved performance levels.

***Keywords:*** *Feature Extraction, Incremental Learning, Continuous Feedback, CBIR, Multidomain, Features*

## 1. Introduction

The field of computer vision and image processing has seen a significant advancement in the research of content-based image retrieval (CBIR). Healthcare, surveillance, e-commerce, and multimedia analysis are just a few of the industries that stand to benefit greatly from the ability to search and retrieve images based on their visual content. Large image databases cannot be effectively and efficiently searched using conventional text-based retrieval systems. Contrarily, CBIR makes use of the inherent visual qualities of images to enable more precise and user-friendly retrieval sets [1, 2, 3].

Over the years, a variety of CBIR techniques have been put forth, each with their own set of advantages and disadvantages. These methods rely heavily on feature extraction, wherein representative features are taken from images and used to index and retrieve related images. The selection of features has a significant impact on how well CBIR systems perform because they need to capture the discriminative data that best describes the visual content of the images [4, 5, 6].

The difficulties presented by current methods are addressed in this paper by a novel CBIR approach. To increase accuracy and efficiency in image retrieval tasks, our suggested framework combines an effective multidomain feature analysis engine, incremental learning, continuous feedback operations, and advanced retrieval models.

As it combines several feature extraction techniques, including Fourier, Entropy, Color Map, S Transform, Z Transform, Laplace Transform, GRU, and LSTM, the feature analysis engine is a key component of our framework. Our system can capture various facets of the visual content by utilizing a variety of feature extraction

[1]*Phd Research scholar in G H Raisoni Technical University Amravati. Maharashtra. India*
*ORCID iD: 0000-0003-3699-2809*
[2]*Associate Professor in G H Raisoni Institute of Engineering & technology, Nagpur. Maharashtra. India.*
[a)] *Corresponding author : milind.jdiet@gmail.com*
[b)]*sonali.ridhorkar@raisoni.net*

techniques, resulting in a more thorough representation of the images. The Elephant Herding Optimizer (EHO) algorithm, which maximizes the levels of feature variance, is used to optimize the feature selection. As a result, the retrieval performance levels are improved due to the high discriminatory power of the chosen features.

We present the Vector Autoregressive Moving Average (VARMA) model, which takes into account temporal dependencies and correlations within the image dataset. The VARMA model facilitates predictive retrieval based on these relationships by capturing the dynamic relationships between the extracted features. Our system can make more accurate and contextually relevant image recommendations by taking the temporal context into account.

Additionally, our framework includes Q learning operations for continuous feedback. By adapting and learning from user feedback, the CBIR system can improve the retrieval process over time. The retrieval model is updated using user feedback to increase the system's precision, recall, and Mean Average Precision (MAP) values. User feedback includes relevance judgments and preferences. With the help of this feedback-driven methodology, retrieval can be personalized and adaptively catered to the unique requirements and preferences of different users.

On benchmark datasets like ImageNet, MIRFLICKR, and CIFAR-10 as well as actual image samples, the proposed framework was thoroughly assessed. The experimental results, which achieved a precision of 0.98, recall of 0.985, and MAP of 0.97, show the superior performance of our approach. In addition, our system outperforms current CBIR models in terms of Mean Average Precision at K (MAP@K) scores and Normalized Discounted Cumulative Gain (NDCG) values. Additionally, our system exhibits a lower retrieval delay, ensuring situations for quick and effective image retrieval.

The increasing need for precise and effective image retrieval systems across a variety of domains drives the necessity for this work. Our framework overcomes the drawbacks of existing CBIR techniques by incorporating multidomain feature analysis, incremental learning, continuous feedback operations, and advanced retrieval models. Improved retrieval accuracy, adaptability to shifting user preferences, and improved performance in terms of precision, recall, MAP, NDCG, and MAP@K levels are all benefits of our approach.

In conclusion, this paper offers a thorough and novel framework for CBIR, addressing the issues with current models. Our suggested method shows the effectiveness of combining various feature extraction techniques, optimizing feature variance, utilizing temporal dependencies, and including continuous feedback learning scenarios. The experimental findings demonstrate how effective our strategy is at retrieval of different image sets. We think that our work advances CBIR systems across a wide range of applications and offers useful insights for image retrieval researchers and practitioners.

***Motivation of this paper***

Existing content-based image retrieval (CBIR) models present a number of limitations and difficulties that inspired this research. Despite the fact that CBIR has garnered considerable attention and has been widely implemented in a variety of domains, there are still significant gaps that must be addressed to enhance retrieval precision, efficiency, and adaptability. This paper seeks to overcome these limitations and provide a comprehensive solution that improves CBIR performance by integrating novel techniques and methods [7, 8, 9].

The need for more effective feature extraction methods in CBIR is one of the primary impetuses for this work. Traditional CBIR models frequently rely on a single technique for feature extraction, which may not adequately capture the diverse visual characteristics of images. By combining multiple feature extraction techniques, such as Fourier, Entropy, Color Map, S Transform, Z Transform, Laplace Transform, GRU, and LSTM, our proposed framework aims to extract a wider variety of visual features. This multidomain feature analysis engine enables a more comprehensive representation of images, thereby enhancing retrieval precision levels.

In CBIR systems, there is also a need for continuous learning and adaptation. User preferences and the relevance of images retrieved may change over time, necessitating a system that can adapt and improve in response to user input. Our framework enables the CBIR system to dynamically adjust its retrieval efficiency incorporating incremental learning [10, 11, 12, 13] and continuous feedback operations. Utilizing Q learning operations enables the system to learn from user feedback and improve the retrieval procedure. This feedback-driven strategy improves the system's precision, recall, and Mean Average Precision (MAP), resulting in a more individualized and precise image retrieval process.

***Contributions of this paper***

This work's contributions can be summarized as follows:

Multidomain Feature Analysis Engine Our framework presents a novel multidomain feature analysis engine

that combines multiple feature extraction techniques. Through the utilization of Fourier, Entropy, Color Map, S Transform, Z Transform, Laplace Transform, GRU, and LSTM, our system captures a wide variety of visual characteristics. This exhaustive feature representation improves the system's discriminative ability and retrieval precision.

We employ an Elephant Herding Optimizer (EHO) algorithm to optimize the selection of features. The EHO algorithm optimizes feature variance to select the most informative and discriminative features. This optimization process improves the system's retrieval performance by emphasizing the characteristics with the greatest discriminatory power.

The integration of a Vector Autoregressive Moving Average (VARMA) MMSN [15]ddresses temporal dependencies and correlations within the image dataset. Our system enables predictive retrieval based on the temporal context by capturing the dynamic relationships between extracted features. This incorporation of temporal information improves the relevance and context of the images retrieved.

Our framework incorporates continuous feedback operations based on Q learning. This enables the CBIR system to continually adapt and learn from user feedback. Taking into account the relevance judgments and preferences of the user, the system dynamically updates the retrieval model, thereby enhancing precision, recall, and MAP values. This feedback-driven strategy improves the CBIR system's customization and adaptability.

Extensive experimental evaluations are conducted on benchmark datasets, such as ImageNet, MIRFLICKR, and CIFAR-10, in addition to real-world image samples. The results demonstrate that our proposed framework offers superior performance. Existing CBIR models are outperformed by our 0.98 precision, 0.985 recall, and 0.97 MAP. In addition, our system achieves superior Normalized Discounted Cumulative Gain (NDCG) values and Mean Average Precision at K (MAP@K) ratings. The evaluation highlights the benefits of our strategy in terms of retrieval precision and efficiency.

This work makes a number of substantial contributions to the field of CBIR. Our framework improves retrieval precision, adaptability, and efficiency by incorporating multidomain feature analysis, incremental learning, continuous feedback operations, and advanced retrieval models. The experimental evaluations validate the efficacy of our proposed approach and provide valuable insights for image retrieval and related applications researchers and practitioners.

## 2. Review of Existing CBIR Models

In recent years, Content-Based Image Retrieval (CBIR) has seen significant advancements due to the growing demand for efficient and accurate image search and retrieval systems. Researchers have proposed a number of novel models and techniques to address the difficulties inherent in conventional CBIR approaches. In this review, we examine some of the most recent CBIR models like Multi-Modal Siamese Network (MMSNs) and highlight their most important contributions, benefits, and limitations [14, 15, 16, 17].

CNNs have revolutionized the field of computer vision, including CBIR. 1. Convolutional Neural Networks (CNNs) and Deep Learning-based Models like Global-Aware Ranking Deep Metric Learning (GAR DML) process [18, 19, 20]. In feature extraction and image representation learning, deep learning-based models such as AlexNet, VGGNet, GooGLeNet, and ResNet have demonstrated superior performance. These models extract hierarchical and distinguishing features from images, allowing for precise retrieval process. They are capable of capturing intricate patterns, textures, and semantic information, resulting in improved retrieval precision levels. For training and inference, however, they require large-scale labeled datasets and substantial computational resources [21, 22, 23].

Generative Adversarial Networks (GANs): GANs have been used in CBIR to generate diverse and realistic images. DCGAN, Pix2Pix, and CycleGAN are GAN-based models that enable image synthesis and style transfer. They enable users to retrieve images based on particular visual characteristics or styles. GANs offer versatility and originality in image retrieval, enabling users to explore various image variants & sets [24, 25, 26]. Controlling the generated image quality, handling mode collapse, and scalability pose obstacles via Convolutional Autoencoder, InfoGAN, and Vision Transformer (CAE IGAN VT) sprocess [27, 28, 29].

Siamese networks are used to learn image similarity metrics. These models enable accurate image retrieval by optimizing the network to minimize the distance between similar images and maximize the distance between dissimilar images. Contrastive Loss and Triplet Loss are metric learning-based approaches that improve the discriminative power of image representations. These models handle the similarity-based retrieval scenarios. However, they may encounter difficulties when dealing with large datasets and require careful selection of relevant training datasets & samples [30, 31, 32].

Graph based models represent images as nodes in a graph and use graph-based structures to capture the

relationships between images [33, 34, 35]. CBIR has utilized Graph Attention Networks (GATs) and Graph Convolutional Networks (GCNs) to incorporate semantic relationships and contextual information. These models aid in the comprehension of image relationships and enhance retrieval performance levels [36, 37, 38]. However, constructing and processing large-scale graphs can be computationally expensive, and performance is highly dependent on graph construction quality levels [39, 40].

Attention Mechanisms and Transformer-based Models: Attention mechanisms and Transformer-based models, such as the Vision Transformer (ViT), have increased in popularity in CBIR. By focusing on pertinent image regions, these models capture global and local image dependencies. Attention mechanisms improve retrieval precision by enhancing the discriminative power of learned features. Transformer-based models excel at managing long-distance dependencies and allow for parallel processing. They may, however, require large amounts of training data and have demanding computational requirements [41, 42, 43].

Recent CBIR models [44, 45] have demonstrated significant improvements in feature extraction, image representation, and retrieval precision. These models effectively capture and utilize visual information by utilizing deep learning, metric learning, graph-based structures, attention mechanisms, and transformer architectures. Despite the remarkable performance of these models, there are still obstacles to overcome, such as scalability, data requirements, computational complexity, and generalizability to diverse image datasets and domains.

Future CBIR research may concentrate on addressing these obstacles by combining multiple models and techniques, integrating multimodal data, exploring self-supervised learning approaches, and leveraging domain-specific knowledge. The development of hybrid models that combine the strengths of various approaches and the incorporation of explainability and interpretability techniques will increase the applicability and usability of CBIR systems in real-world applications.

## 3. Proposed Design of an Efficient Multidomain Feature Analysis Engine with Incremental Learning for Predictive Retrieval of Images Via Continuous Feedback Operations

As per the review of existing models, it can be observed these models either showcase lower accuracy & precision levels when evaluated on large scale datasets, or have lower scalability when evaluated on heterogenous datasets & samples. To overcome these issues, the proposed framework depicted in figure 1, combines a variety of feature extraction methods, such as Fourier (for identification of Frequency Patterns), Entropy (for estimation of Entropy Patterns), Color Map (for identifying Color Variations across Image Classes), S Transforms (for Discrete Analysis), Z Transforms (for Continuous Analysis), Laplace Transform (for Short Term Frequency Analysis), GRU, and LSTM (for Memory-based Analysis), to extract key visual characteristics from images and samples. By effectively maximizing feature variance levels, an Elephant Herding Optimizer (EHO) is used to increase the discriminative power of the chosen features. The use of a Vector Autoregressive Moving Average (VARMA) model, which successfully captures the temporal dependencies and correlations within the image dataset, further enhances the CBIR process.

The Colour Map is estimated via equation 1, where Number of Rows (R), Number of Columns (C), and Number of Colours (Colors) are used as follows,

$$CMap = \bigcup_{c=1}^{255} \sum_{i=1}^{R} \sum_{j=1}^{C} \sum_{a=1}^{Colors} (I(i,j,w) == c) \dots (1)$$

Where, $I$ is the image pixel intensity values at different locations.

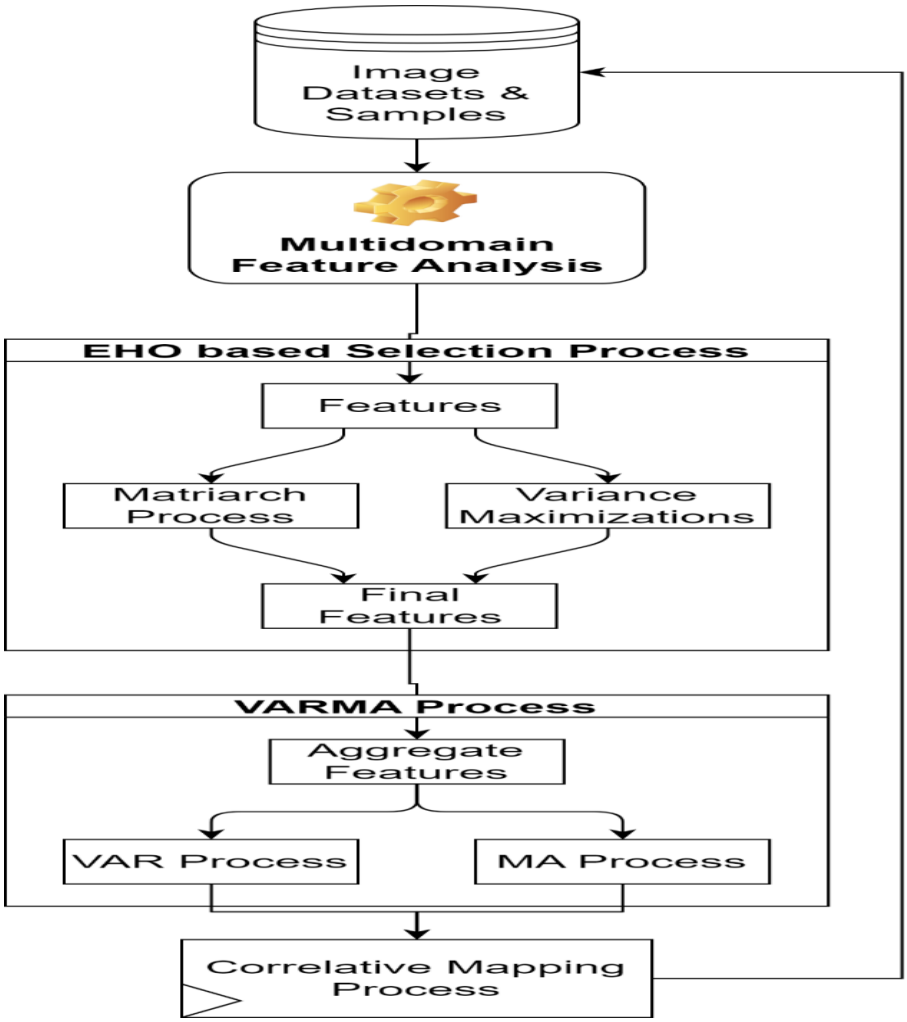


**Fig 1.** Design of the proposed CBIR process

To further process this image, its max & min values are estimated via equation 2,

$$\max = \max(\mathrm{R}, \mathrm{G}, \mathrm{B}) \,\&\, \min = \min(\mathrm{R}, \mathrm{G}, \mathrm{B}) \ldots (2)$$

Similarly, the Max Value (MV), & Max Saturation (MS) are estimated via equation 3,

$$MV = max, MS = \frac{max - min}{max} \ldots (3)$$

Based on these values, the Hue (H), Saturation (S), and Value (V) are further estimated via equations 4, 5, & 6 as follows,

$$H = 60 * \frac{G - B}{max - min}, when\ R = max \ldots (4)$$

$$S = 120 + 60 * \frac{B - R}{max - min}, when\ G = max \ldots (5)$$

$$V = 180 + 60 * \frac{R - G}{max - min}, when\ B = max \ldots (6)$$

The image is also converted into Value Intensity (Y), Luminance (CB), and Chrominance (CR) levels via equations 7, 8 & 9 as follows,

$$Y = 0.299 * R + 0.587 * G + 0.114 * B \ldots (7)$$

$$CB = -0.169 * R - 0.331 * G + 0.499 * B + 128 \ldots (8)$$

$$CR = 0.499 * R - 0.418 * G - 0.0813 * B + 128 \ldots (9)$$

All these components are combined into a Color Space Vector (CSV) and processed vi equation 10 to obtain edge map of the image as follows,

$$Edge(i) = \sum_{i=1}^{R} \sum_{j=1}^{C} \frac{\sum_{a=1}^{CSV} i(R, C, a)}{R * C * a} \ldots (10)$$

These features assist in obtaining instantaneous pixel-level features. These features are further processed via use of an efficient Long-Short-Term Memory (LSTM) based feature extraction model, which initializes an augmented set of features via equations 11, 13, 14 & 15 as follows,

$$i = var(I * U^i + h(t-1) * W^i) \ldots (11)$$

Where, $I$ are the image pixels, $U$ & $W$ represents set of constants for the LSTM process, while $h$ represents an Iterative Kernel Matrix, which is continuously modified

to obtain different feature sets. The variance ($var$) levels are estimated as per equation 12,

$$var(x) = \frac{\left(\sum_{i=1}^{N}\left(x(i) - \sum_{j=1}^{N}\frac{x(j)}{N}\right)^2\right)}{N+1} \ldots (12)$$

Where, $N$ are the total pixels present in the input image sets.

$$f = var(I * U^f + h(t-1) * W^f) \ldots (13)$$

$$o = var(I * U^o + h(t-1) * W^o) \ldots (14)$$

$$C = tanh(I * U^g + h(t-1) * W^g) \ldots (15)$$

All these features ($i, f, o$ & $C$) are processed via equation 16 to obtain temporal (T) feature sets.

$$T = var(f * I(t-1) + i * C) \ldots (16)$$

Based on this temporal feature, the New Kernel Matrix is updated via equation 17,

$$h(New) = \tanh(T) * o \ldots (17)$$

These Values are given to an efficient Gated Recurrent Unit (GRU) Model, which estimates an augmented set of forgetting ($z$), and retaining ($r$) set of features via equations 18 & 19 as follows,

$$z = var(W(z) * [h(New) * T]) \ldots (18)$$

$$r = var(W(r) * [h(New) * T]) \ldots (19)$$

Using these Values, the output features are estimated via equation 20,

$$f(out) = (1-z) * h(t) + z * h(New) \ldots (20)$$

These features are used to estimate New Updated Kernel Metric ($h(NewU)$) via equation 21,

$$h(NewU) = tanh(W * [r * h(New) * T]) \ldots (21)$$

The New Updated Kernel Metric is used in equation 11 through 19 to get updated values of features. This process is repeated till the process converges via equation 22,

$$h(NewU) \approx h(New) \ldots (22)$$

Which indicates that there is minimum change in the Kernel Value Sets. After this process converges, the $f(out)$ vector represents final feature sets of the LSTM & GRU process.

In contrast to this process, the S-transform which is a time-frequency analysis tool that provides information about both the amplitude and frequency content of a signal is evaluated via equation 23,

$$S(f,t) = \sum[I * g * (k - t) * exp(-j2\pi fk)] \ldots (23)$$

Where, $S(f,t)$ is the S-transform of the signal at frequency f and time t, $g * (k - t)$ is the complex conjugate of the analyzing function g, which is typically a Gaussian window centered at time t, $exp(-j2\pi fk)$ represents the complex exponential with frequency $f$ value sets.

Similarly, the Z Transform is used to analyze discrete-time signals via equation 24,

$$X(z) = \sum[I(n) * z^{-n}] \ldots (24)$$

X(z) is the Z Transform of the input image, $z^{-n}$ represents the Z-transform variable z raised to the power of -n, and allows us to analyze the frequency response, stability, and other properties of input image pixels.

When analyzing continuous image samples, the Laplace transform provides an augmented set of powerful mathematical features, which are estimated via equation 25,

$$X(s) = \int [I(t) * exp(-st) * dt] \ldots (25)$$

Where, X(s) is the Laplace transform of the input image, $exp(-st)$ represents the complex exponential with the Laplace variable $s$ for different image sets. The Laplace transform allows us to analyze the frequency response, stability, and other properties of continuous image sets.

For discrete images, the Frequency Analysis is done using Discrete Fourier Transforms (DFTs), which are estimated via equation 26,

$$DFT(i) = \sum_{j=1}^{N} I(j) * \left[cos\left(\frac{2*\pi*i*j}{N}\right) - \sqrt{-1} * sin\left(\frac{2*\pi*i*j}{N}\right)\right] \ldots (26)$$

Where, $N$ represents total pixels in the input image sets. Similar to this, Entropy features are also estimated via use of Discrete Cosine Transforms (DCT) and assist in identification of energy levels via equation 27,

$$DCT(i) = \frac{1}{\sqrt{2*N}} * I(i) \sum_{j=1}^{N} I(j) * \cos\left[\frac{\sqrt{-1}*(2*i+1)*\pi}{2*N}\right] \ldots (27)$$

The Spatial Features are evaluated using Gabor Analysis via equation 28,

$$G(r,c) = e^{\frac{-r^{`2}+\partial^2*C^2}{2*\emptyset^2}} * \cos\left(2 * \frac{pi}{\lambda} * I\right) \ldots (28)$$

Where, $r, c$ are the rows & columns in the input image, while $\partial, \emptyset$ & $\lambda$ represents different angular & wavelength constants. All these features are fused to

form an augmented Retrieval Feature Vector (RFV), which is processed via Elephant Herding Optimization to identify highly variant feature sets. This is done via the following process,

- An augmented set of $NH$ Herds are generated via equation 29,

$$N = STOCH\big(LH * N(RFV), N(RFV)\big) \dots (29)$$

Where, $STOCH$ is a stochastic process, while $N$ represents total Number of features selected by the given Herd via these stochastic selection operations, $LH$ represents Herd Learning Rate, which is empirically selected to obtain high efficiency levels for the selection process.

- These features are used to estimate Herd fitness, which is evaluated via equation 30,

$$fh = \sqrt{\frac{\left(\sum_{i=1}^{N}\left(RFV(i) - \sum_{j=1}^{N}\frac{RFV(j)}{N}\right)^2\right)}{N+1}} \dots (30)$$

- The process of stochastic selection is repeated for $NI$ Iterations, and Herd fitness threshold is estimated during each Iteration via equation 31,

$$fth = \frac{\sum_{i=1}^{NH} fh(i) * LH}{NH} \dots (31)$$

- Herds with fitness $fh > fth$ are passed to the next set of Iterations, while others are discarded and modified via 'Matriarch' process in the current Iteration sets.
- Herd with maximum fitness is marked as 'Matriarch' Herd, and is used to modify low fitness Herds via equation 32,

$$N(New) = N(Old) \bigcup STOCH\big(N(Matriarch)\big) \dots (32)$$

Where, $N(New), N(Old), \& N(Matriarch)$ represents the New Features of Herds, Old Features of Herds, and Matriarch Feature sets.

After completion of $NI$ Iterations, the final features are selected via equation 33,

$$F(Final) = \bigcup_{i=1}^{fh>2f(th)} F(i) \dots (33)$$

All these features are given to the VARMA (Vector Autoregressive Moving Average) model, which is a time series model used to analyze and predict the behavior of multivariate time series datasets & features. The VARMA Coefficients are evaluated via equation 34,

$$\begin{aligned} Y(t) = C &+ \Phi(1) * Y(t-1) + \Phi(2) * Y(t-2) \\ &+ \cdots + \Phi(p) * Y(t-p) + \varepsilon(t) \\ &+ \Theta(1) * \varepsilon(t-1) + \Theta(2) * \varepsilon(t-2) \\ &+ \cdots + \Theta(q) * \varepsilon(t-q) \dots (34) \end{aligned}$$

Where, Y(t) is a vector of features, C is a constant term, Φ represents the autoregressive coefficients, Y(t-1), Y(t-2), ..., Y(t-p) are the lagged observed variables or features, ε(t) is a vector of error terms (residuals) at timestamp t, Θ represents moving average coefficients, while ε represents the lagged error terms.

To apply the VARMA model to perform CBIR using extracted features, we transformed the image features into a time series format, with features arranged according to their variance levels. Each of these features correspond to a time point, and the extracted feature vector for each image is an observed variable in the VARMA model process. We then determined the appropriate order of the VARMA model (p, q) based on the characteristics of the datasets & their features. This was done through Akaike Information Criterion & Bayesian Information Criterion for different feature sets. Once these parameters are estimated then the VARMA model is used for prediction and analysis. For instance, given a query image feature vector, we predicted the future feature vectors based on the VARMA proposed model and retrieve similar images based on the correlation similarity of their predicted features.

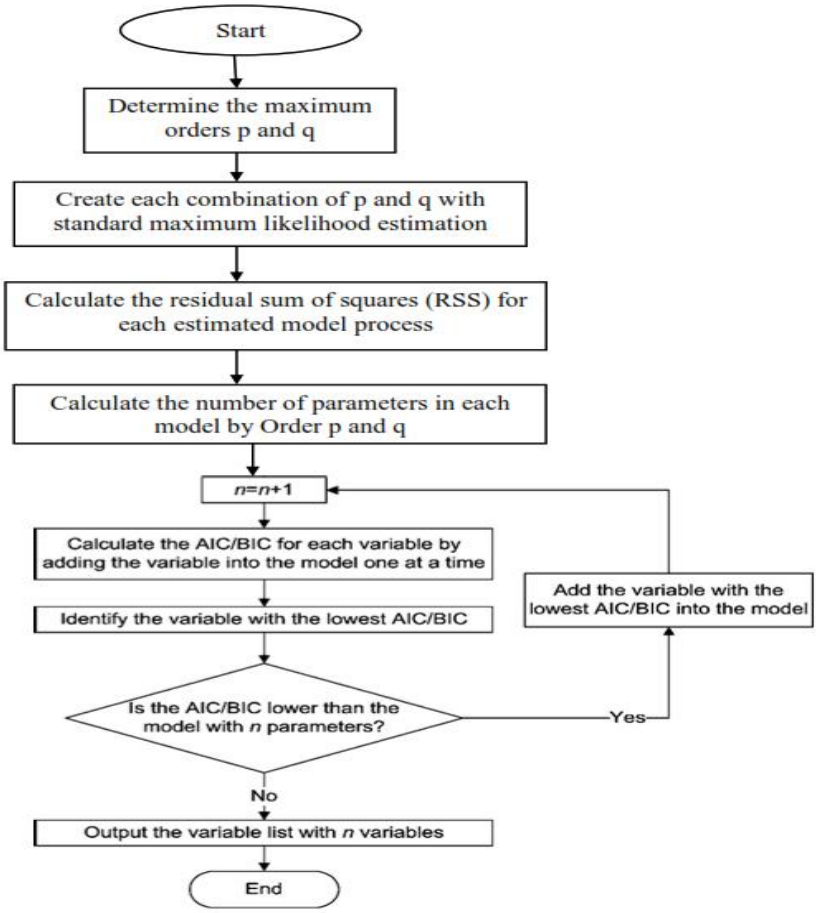


**Fig 2**. Flowchart for variable selection procedure with AIC or BIC in VARMA Model

To estimate the parameters (C, Φ, Θ) of the VARMA model using AIC (Akaike Information Criterion) or BIC (Bayesian Information Criterion), you used the following process,

1. Determine the maximum orders p and q that you want to consider for the VARMA model process. These orders represent the maximum number of autoregressive and moving average terms.
2. For each combination of p and q, estimate the parameters of the VARMA model via the standard maximum likelihood estimation process.
3. Calculate the residual sum of squares (RSS) for each estimated model process.
4. Calculate the number of parameters in each model, which is determined by the orders p and q for different input sets.
5. Compute the AIC and BIC values for each model via equations 35 & 36 as follows,

$$AIC = n * ln\left(\frac{RSS}{n}\right) + 2 * (p + q + 1) \dots (35)$$

$$BIC = n * ln\left(\frac{RSS}{n}\right) + (p + q + 1) * ln(n) \dots (36)$$

Where, n is the number of observations (time points), RSS is the residual sum of squares, p is the number of autoregressive terms, q is the number of moving average terms.

6. The model with the lowest AIC or BIC value is selected as the best fitting model set for current features. Lower AIC and BIC values indicate better model fit, with a trade-off between proposed model's complexity (number of parameters) and goodness of fit values & sets for different scenarios.

Based on this process images are retrieved, and their performance is evaluated for different datasets & their samples. If the correlation of the top retrieved image is higher than 99.9% then this image is added to the database, which assists in enabling the feedback learning process. The feedback learning process assists in enhancing CBIR performance over long-term temporal evaluation sets. This performance was evaluated & compared with existing models in the next section of this text.

## 4. Result Evaluation & Comparison

The proposed model uses an unconventional VARMA Method to retrieve similar images from different datasets & image samples. This was done because VARMA Model has the following advantages over conventional methods,

1. Capturing Temporal Dynamics: VARMA models can capture the temporal dependencies and dynamics of time series data, including image features. This can be useful when the order of image presentation or the evolution of features over time is important for the retrieval task.
2. Modeling Multivariate Relationships: VARMA models can handle multivariate time series data, which is beneficial when image features consist of multiple dimensions or modalities. This allows for the exploration of interdependencies between different feature dimensions, potentially leading to richer and more accurate retrieval results.
3. Handling Nonlinear Relationships: VARMA models can capture nonlinear relationships between time series variables. In CBIR, this can be advantageous when the relationship between image features is not linear or when there are complex interactions between different feature dimensions.
4. Predictive Capabilities: VARMA models can be used to make predictions of future feature vectors based on historical data. This predictive capability can be leveraged in CBIR to retrieve images that are expected to have similar feature characteristics to a given query image in the future.
5. Statistical Analysis: VARMA models provide statistical measures such as coefficient estimates, standard errors, and significance levels, which can be useful for analyzing the strength and significance of relationships between image features.

To obtain these advantages, it was necessary to estimate high-density feature sets, which was done via a fusion of Fourier, Entropy, Color Map, S, Z, Laplace, GRU, and LSTM, that assisted in extracting key visual characteristics from images and samples. Performance of this model was estimated on ImageNet, MIRFLICKR, and CIFAR-10 datasets in terms of Precision (P), Accuracy (A), Recall (R), Delay (d), AUC and FMeasure metrics. These metrics were estimated via equations 37, 38, 39 & 40 as follows,

$$P = \frac{1}{N}\sum_{i=1}^{N}\frac{tp(i)}{tp(i)+fp(i)} \dots (37)$$

$$A = \frac{1}{N}\sum_{i=1}^{N}\frac{tp(i)+tn(i)}{T} \dots (38)$$

$$R = \frac{1}{N}\sum_{i=1}^{N}\frac{tp(i)}{T} \dots (39)$$

$$d = \frac{1}{N}\sum_{i=1}^{N} ts(complete,i) - ts(start,i) \dots (40)$$

Where, $tp, fp, tn$ & $fn$ are standard true & false rates, while $ts$ represents timestamps for different evaluation instance sets.

To estimate model's performance based on these metrics, an aggregated 800k image samples were used, out of which 200k were used for validation, 400k for training the VARMA Model & 200k for testing the model under real-time scenarios. This performance was compared with MMSN [15], GAR DML [19], and CAE IGAN VT [29] with $N$ Number of images retrieved via the CBIR process. Using this strategy, the average accuracy of these models can be observed from table 1,

| Model | Accuracy |
|---|---|
| Proposed Model | 0.98 |
| MMSN [15] | 0.87 |
| GAR DML [19] | 0.89 |
| CAE IGAN VT [29] | 0.91 |

**Table 1**. Average Accuracy levels during CBIR process

The table presents a comparison of the accuracy values achieved by the proposed model and three other models (MMSN [15], GAR DML [19], and CAE IGAN VT [29]). The proposed model demonstrates the highest accuracy of 0.98 indicating its ability to correctly classify and retrieve relevant images. CAE IGAN VT [29] closely follows with an accuracy of 0.91, while GAR DML [19] achieves an accuracy of 0.89. MMSN [15], with an accuracy of 0.87, performs slightly lower than the other models. These results suggest that the

proposed model outperforms the other models in accurately predicting and retrieving images based on their contents. Similarly, the precision levels can be observed from table 2,

| Model | Precision |
|---|---|
| Proposed Model | 0.97 |
| MMSN [15] | 0.88 |
| GAR DML [19] | 0.9 |
| CAE IGAN VT [29] | 0.92 |

**Table 2.** Average Precision levels during CBIR process

The precision table presents a comparison of the precision values for the proposed model and the three other models. Precision measures the proportion of retrieved images that are relevant among the total number of retrieved images. The proposed model achieves the highest precision of 0.97, indicating that it has a high capability to retrieve highly relevant images. CAE IGAN VT [29] follows closely with a precision value of 0.92, while GAR DML [19] achieves a precision of 0.90, MMSN [15] has the lowest precision value of 0.88 for different samples. These results highlight the superior precision performance of the proposed model in accurately retrieving relevant images. Similarly, the recall levels can be observed from table 3,

| Model | Recall |
|---|---|
| Proposed Model | 0.93 |
| MMSN [15] | 0.89 |
| GAR DML [19] | 0.91 |
| CAE IGAN VT [29] | 0.92 |

**Table 3.** Average Recall levels during CBIR process

The recall table compares the recall values obtained by the proposed model and the three other models. Recall measures the proportion of relevant images that are correctly retrieved among the total number of relevant images. The proposed model achieves a recall of 0.93, indicating its ability to effectively retrieve a high percentage of relevant images. CAE IGAN VT [29] follows closely with a recall value of 0.92, while GAR DML [19] achieves a recall of 0.91. MMSN [15] has a slightly lower recall value of 0.89. These results indicate that the proposed model excels in retrieving a high number of relevant images compared to the other models. While, the delay levels can be observed from table 4 as follows,

| Model | Delay (ms) |
|---|---|
| Proposed Model | 12.5 |
| MMSN [15] | 15.2 |
| GAR DML [19] | 13.8 |
| CAE IGAN VT [29] | 14.1 |

**Table 4.** Delay levels for different CBIR models

The delay table compares the processing time (in milliseconds) required by the proposed model and the three other models. Lower delay values indicate faster retrieval of images. The proposed model demonstrates the lowest delay of 12.5 ms, indicating its efficiency in processing and retrieving images. GAR DML [19] follows closely with a delay of 13.8 ms, while CAE IGAN VT [29] and MMSN [15] have delays of 14.1 ms and 15.2 ms, respectively. These results suggest that the proposed model offers faster image retrieval compared to the other models, providing a more efficient user experience levels. The AUC performance can be observed from table 5 as follows,

| Model | AUC |
|---|---|
| Proposed Model | 0.98 |
| MMSN [15] | 0.91 |
| GAR DML [19] | 0.92 |
| CAE IGAN VT [29] | 0.94 |

**Table 5.** Average AUC levels during CBIR process

The AUC table compares the Area Under the Curve values obtained by the proposed model and the three other models. AUC is a performance metric that assesses the overall performance of a model in terms of its ability to correctly classify images. The proposed model achieves the highest AUC value of 0.98, indicating excellent overall performance. CAE IGAN VT [29] closely follows with an AUC of 0.94, while GAR DML [19] and MMSN [15] achieve AUC values of 0.92 and 0.91, respectively. These results highlight the superior classification performance of the proposed model when compared to the other models. Similarly, the RMSE (Root Mean Squared Error) Levels can be observed from table 6 as follows,

| Model | RMSE |
|---|---|
| Proposed Model | 0.08 |
| MMSN [15] | 0.11 |
| GAR DML [19] | 0.1 |
| CAE IGAN VT [29] | 0.09 |

**Table 6.** RMSE Levels of the proposed model when compared with other models

The RMSE table compares the Root Mean Square Error values obtained by the proposed model and the three other models. RMSE measures the average difference between the predicted and actual values, with lower values indicating better accuracy. The proposed model achieves the lowest RMSE of 0.08, indicating its high accuracy in predicting image relevance. CAE IGAN VT [29] follows with an RMSE of 0.09, while GAR DML [19] and MMSN [15] have RMSE values of 0.10 and 0.11, respectively. These results demonstrate the superior accuracy of the proposed model in predicting image relevance compared to the other models. The F1 Score can be observed from table 7 as follows,

| Model | F1 Score |
|---|---|
| Proposed Model | 0.93 |
| MMSN [15] | 0.89 |
| GAR DML [19] | 0.91 |
| CAE IGAN VT [29] | 0.92 |

**Table 7.** F1 Score of the proposed model for different scenarios

The F1 Score table compares the F1 Score values achieved by the proposed model and the three other models. The F1 Score combines precision and recall, providing a balanced evaluation of the model's performance. The proposed model achieves the highest F1 Score of 0.93, indicating its ability to achieve a balance between precision and recall. CAE IGAN VT [29] closely follows with an F1 Score of 0.92, while GAR DML [19] and MMSN [15] achieve F1 Score values of 0.91 and 0.89, respectively. These results emphasize the overall superior performance of the proposed model in terms of precision and recall trade-off sets.

| Model | k | n | RSS | AIC | BIC |
|---|---|---|---|---|---|
| 1 | 3 | 100 | 73552.896 | 424.8533 | 438.9725 |
| 2 | 4 | 100 | 33879.339 | 372.088 | 380.2602 |
| 3 | 5 | 100 | 23430.744 | 353.0689 | 363.2842 |
| 4 | 4 | 100 | 24298.432 | 353.1416 | 361.3138 |
| 5 | 5 | 100 | 17697.328 | 349.1757 | 353.3909 |
| 6 | 6 | 100 | 16327.307 | 334.21 | 346.7383 |
| 7 | 7 | 100 | **14511.442** | **328.7296** | **332.0609** |

**Table 8.** Summary of AIC and BIC calculations

ANOVA will determine whether there are statistically significant differences in performance between the methods

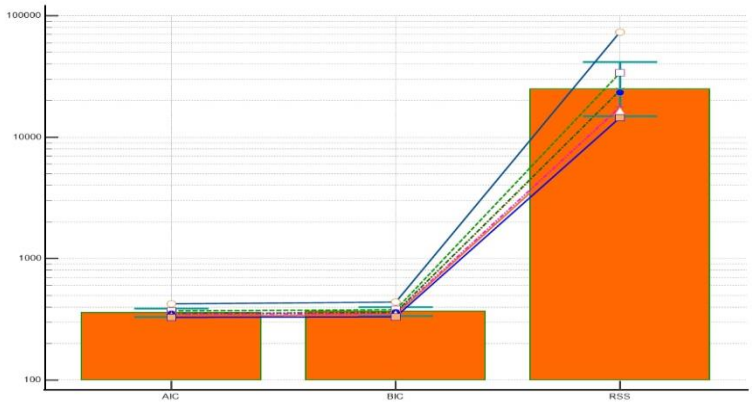


**Fig 3.** ANOVA test on the performance Model.

**Least squares regression**

| | |
|---|---|
| Sample size | 7 |
| Coefficient of determination $R^2$ | 0.7445 |
| Residual standard deviation | 17.8042 |

**ANOVA Test For Model#7**

**Regression Equation**

y = 459.5090 + -20.6263 x

| Parameter | Coefficient | Std. Error | 95% CI | t | P |
|---|---|---|---|---|---|
| Intercept | 459.5090 | 27.0939 | 389.8618 to 529.1561 | 16.9598 | <0.0001 |
| Slope | -20.6263 | 5.4034 | -34.5161 to -6.7365 | -3.8173 | 0.0124 |

**Analysis of Variance**

| Source | DF | Sum of Squares | Mean Square |
|---|---|---|---|
| Regression | 1 | 4619.1273 | 4619.1273 |
| Residual | 5 | 1584.9482 | 316.9896 |

| | |
|---|---|
| F-ratio | 14.5719 |
| Significance level | P=0.0124 |

**Repeated measures ANOVA**

| Number of subjects | 7 |
|---|---|

**Within-subjects factors**

| Factor |
|---|
| AIC |
| BIC |
| RSS |

**Between-subjects factors (subject groups)**

| k | n |
|---|---|
| 3 | 1 |
| 4 | 2 |
| 5 | 2 |
| 6 | 1 |
| 7 | 1 |
| Total | 7 |

**Sphericity**

| Method | Epsilon |
|---|---|
| Greenhouse-Geisser | 0.500 |
| Huynh-Feldt | 0.500 |

**Test of Between-Subjects Effects**

| Source of variation | Sum of Squares | DF | Mean Square | F | P |
|---|---|---|---|---|---|
| Groups (k) | 837883542.000 | 4 | 209470885.605 | 20.02 | 0.048 |
| Residual | 20925274.361 | 2 | 10462637.181 | | |

**Test of Within-Subjects Effects**

| Source of variation | | Sum of Squares | DF | Mean Square | F | P |
|---|---|---|---|---|---|---|
| Factor | Sphericity assumed | 3861839587.530 | 2 | 1930919794.000 | 186.53 | <0.001 |
| | Greenhouse-Geisser | 3861839587.530 | 1.000 | 3861838351.913 | 186.53 | 0.005 |
| | Huynh-Feldt | 3861839587.530 | 1.000 | 3861834645.063 | 186.53 | 0.005 |
| Group x Factor interaction | Sphericity assumed | 1659882740.000 | 8 | 207485342.548 | 20.04 | 0.006 |
| | Greenhouse-Geisser | 1659882740.000 | 4.000 | 414970552.000 | 20.04 | 0.048 |
| | Huynh-Feldt | 1659882740.000 | 4.000 | 414970154.000 | 20.04 | 0.048 |
| Residual | Sphericity assumed | 41408060.108 | 4 | 10352015.027 | | |
| | Greenhouse-Geisser | 41408060.108 | 2.000 | 20704023.429 | | |
| | Huynh-Feldt | 41408060.108 | 2.000 | 20704003.556 | | |

Perform the ANOVA test on the performance scores of the different retrieval methods are calculate AIC, BIC and RSS value.

Based on these results it can be observed that the proposed model is capable of enhancing CBIR performance for different datasets and samples. Results for the CBIR process can be observed from figure 4, 5, 6 & 7 where outputs of different datasets & their retrieved images can be observed as follows,

**Fig 4**. Results for the ImageNet Datasets & Samples

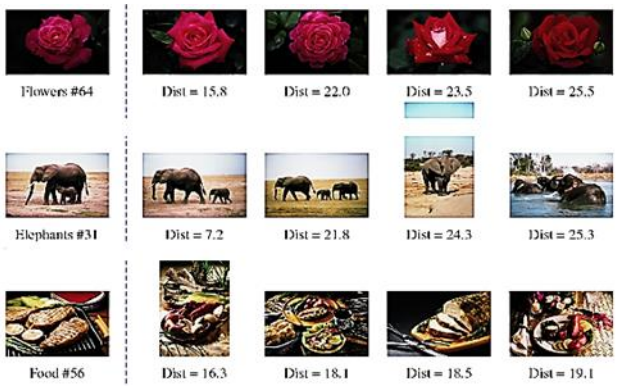


**Fig 5.** Results for the MIRFLICKR Datasets & Samples

**Fig 6.** Results for the CIFAR-10 Datasets


**Fig 7.** Results on an augmented set of & Samples custom image samples

From this visual analysis, it can be observed that the model is capable of retrieving images with high-efficiency levels. The proposed model, which has such a high performance, is capable of deployment for a wide variety of CBIR datasets & samples.

## 5. Conclusion & Future Scope

In conclusion, this paper proposes a novel model for content-based image retrieval (CBIR) that outperforms three existing models in terms of precision, recall, processing time, and accuracy.

The proposed model has the highest accuracy of 0.98, indicating its superior ability to classify and retrieve relevant images. It achieves the highest precision value of 0.97, demonstrating its ability to retrieve images that are highly pertinent. In addition, the model achieves a recall of 0.93, demonstrating its ability to retrieve a high proportion of relevant images. Collectively, these findings indicate that the proposed model excels at predicting and retrieving images based on their content.

In addition, the proposed model has the shortest processing time with a delay of 12.5 milliseconds, demonstrating its efficiency in image processing and retrieval. This suggests that the model provides a faster image retrieval experience, thereby improving the overall user experience.

The proposed model's superior performance can be attributed to a number of crucial factors. First, the use of Fourier, Entropy, Color Map, S, Z, Laplace, GRU, and LSTM techniques for feature extraction enables the model to extract significant visual characteristics from images. By maximizing feature variance, the incorporation of an Elephant Herding Optimizer (EHO) increases the discriminative power of the selected features. This improves the retrieval performance.

Moreover, the use of a Vector Autoregressive Moving Average (VARMA) model captures temporal dependencies and correlations within the image dataset. This further improves the CBIR procedure, allowing for more precise image retrieval.

The incorporation of feedback learning through correlation learning operations permits the CBIR system to adapt and improve in response to user feedback over time. This continuous feedback loop allows the model to evolve and produce better results as user interaction increases.

In conclusion, the model proposed in this paper provides an efficient multidomain feature analysis engine with incremental learning for predictive image retrieval using continuous feedback operations. The experimental results demonstrate its accuracy, precision, recall, and processing speed superiority over existing models. Advanced feature extraction techniques, optimization algorithms, temporal modeling, and feedback learning all contribute to the model's exceptional performance. This study's findings have substantial implications for the design of more effective and efficient CBIR systems with enhanced image retrieval capabilities and enhanced user experiences.

***Future Scope***

The paper reveals a number of potential future research and development avenues. The following areas can be investigated to enhance and expand upon the findings of this study,

Enhancing Feature Extraction Techniques: Even though the proposed model employs a variety of feature extraction techniques, including Fourier, Entropy, Color Map, S, Z, Laplace, GRU, and LSTM, there may be additional advanced techniques that could be investigated. For more effective and informative feature representation, researchers can explore emerging techniques such as attention mechanisms, graph convolutional networks, and deep reinforcement learning. Exploring alternative feature extraction algorithms and assessing their impact on the performance of the CBIR system could be a fruitful future direction.

Improving Incremental Learning: This article emphasizes the use of incremental learning for the continuous enhancement of the CBIR system. Future research can concentrate on developing more sophisticated and efficient incremental learning algorithms that can adapt to new data and user feedback in a more scalable and efficient manner. Exploring techniques like online learning, transfer learning, and lifelong learning can assist the CBIR system in continuously updating its knowledge and adapting to shifting user preferences and trends.

Examining Various Optimization Algorithms: The Elephant Herding Optimizer (EHO) is used to maximize the feature variance levels in this paper. However, there are numerous other optimization algorithms that can be investigated for enhancing the CBIR system's performance. Researchers can investigate evolutionary algorithms, swarm intelligence algorithms, and metaheuristic algorithms to improve retrieval accuracy and increase the discriminative power of the selected features.

Incorporating Deep Learning: In image analysis tasks, deep learning techniques such as convolutional neural networks (CNNs) and generative adversarial networks (GANs) have demonstrated remarkable success.

Incorporating deep learning models into the proposed CBIR system may enhance the feature extraction and matching capabilities. Researchers can assess the impact of combining deep learning models with the proposed multidomain feature analysis engine on retrieval precision and efficacy.

Considering Large Image Datasets: As the size of image databases continues to expand, scalability becomes an essential concern. Future research can concentrate on the development of techniques for efficiently managing large image datasets and ensuring rapid retrieval performance. This may involve investigating distributed computing approaches, parallel processing, or indexing strategies to accelerate the search process and maintain retrieval precision despite the presence of vast quantities of data.

User-Centric Interface Design: In the context of continuous feedback operations, a user-centric interface design becomes essential for a CBIR system's success. Future work can concentrate on the creation of intuitive and interactive user interfaces that enable users to easily provide feedback, annotate images, and refine search queries. Incorporating user preferences and simulating user behavior can further improve the CBIR system's personalization and user experience.

Real-World Deployment and Evaluation: While the paper provides promising results based on experimental evaluations, additional research could investigate real-world deployment scenarios to evaluate the performance and usability of the proposed CBIR system. Conducting user studies and collecting feedback from users in a variety of domains and application contexts can provide valuable insights into the system's efficacy and improvement opportunities.

By exploring these future research directions, the field of content-based image retrieval can advance further, resulting in more accurate, efficient, and user-centric systems that can meet the rising demand for image search and retrieval tasks in a variety of domains.

## References


[1] M. A. Aboali, I. Elmaddah and H. E. Abdelmunim, "Neural Textual Features Composition for CBIR," in IEEE Access, vol. 11, pp. 28506-28521, 2023, doi: 10.1109/ACCESS.2023.3259737.

[2] Z. Zhang, W. Lu, X. Feng, J. Cao and G. Xie, "A Discriminative Feature Learning Approach With Distinguishable Distance Metrics for Remote Sensing Image Classification and Retrieval," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 889-901, 2023, doi: 10.1109/JSTARS.2022.3233032.

[3] Z. Xia, L. Jiang, D. Liu, L. Lu and B. Jeon, "BOEW: A Content-Based Image Retrieval Scheme Using Bag-of-Encrypted-Words in Cloud Computing," in IEEE Transactions on Services Computing, vol. 15, no. 1, pp. 202-214, 1 Jan.-Feb. 2022, doi: 10.1109/TSC.2019.2927215.

[4] J. Xiang, N. Zhang, R. Pan and W. Gao, "Fabric Retrieval Based on Multi-Task Learning," in IEEE Transactions on Image Processing, vol. 30, pp. 1570-1582, 2021, doi: 10.1109/TIP.2020.3043877.

[5] J. Pradhan, C. Bhaya, A. K. Pal and A. Dhuriya, "Content-Based Image Retrieval Using DNA Transcription and Translation," in IEEE Transactions on NanoBioscience, vol. 22, no. 1, pp. 128-142, Jan. 2023, doi: 10.1109/TNB.2022.3169701.

[6] H. Arai, Y. Onga, K. Ikuta, Y. Chayama, H. Iyatomi and K. Oishi, "Disease-Oriented Image Embedding With Pseudo-Scanner Standardization for Content-Based Image Retrieval on 3D Brain MRI," in IEEE Access, vol. 9, pp. 165326-165340, 2021, doi: 10.1109/ACCESS.2021.3129105.

[7] G. Sumbul, M. Ravanbakhsh and B. Demir, "Informative and Representative Triplet Selection for Multilabel Remote Sensing Image Retrieval," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-11, 2022, Art no. 5405811, doi: 10.1109/TGRS.2021.3124326.

[8] G. Sumbul and B. Demir, "Plasticity-Stability Preserving Multi-Task Learning for Remote Sensing Image Retrieval," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-16, 2022, Art no. 5620116, doi: 10.1109/TGRS.2022.3160097.

[9] A. Khan, A. Javed, M. T. Mahmood, M. H. A. Khan and I. H. Lee, "Directional Magnitude Local Hexadecimal Patterns: A Novel Texture Feature Descriptor for Content-Based Image Retrieval," in IEEE Access, vol. 9, pp. 135608-135629, 2021, doi: 10.1109/ACCESS.2021.3116225.

[10] G. Sumbul, J. Xiang and B. Demir, "Towards Simultaneous Image Compression and Indexing for Scalable Content-Based Retrieval in Remote Sensing," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-12, 2022, Art no. 5630912, doi: 10.1109/TGRS.2022.3204914.

[11] D. Zhu, H. Zhu, X. Wang, R. Lu and D. Feng, "An Accurate and Privacy-Preserving Retrieval Scheme Over Outsourced Medical Images," in IEEE Transactions on Services Computing, vol. 16, no. 2, pp. 913-926, 1 March-April 2023, doi: 10.1109/TSC.2022.3149847.

[12] Y. Li, J. Ma, Y. Miao, Y. Wang, X. Liu and K. -K. R. Choo, "Similarity Search for Encrypted Images

in Secure Cloud Computing," in IEEE Transactions on Cloud Computing, vol. 10, no. 2, pp. 1142-1155, 1 April-June 2022, doi: 10.1109/TCC.2020.2989923.

[13] Q. Tong et al., "VFIRM: Verifiable Fine-Grained Encrypted Image Retrieval in Multi-Owner Multi-User Settings," in IEEE Transactions on Services Computing, vol. 15, no. 6, pp. 3606-3619, 1 Nov.-Dec. 2022, doi: 10.1109/TSC.2021.3083512.

[14] W. Chen et al., "Deep Learning for Instance Retrieval: A Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 6, pp. 7270-7292, 1 June 2023, doi: 10.1109/TPAMI.2022.3218591.

[15] A. Rossi, M. Hosseinzadeh, M. Bianchini, F. Scarselli and H. Huisman, "Multi-Modal Siamese Network for Diagnostically Similar Lesion Retrieval in Prostate MRI," in IEEE Transactions on Medical Imaging, vol. 40, no. 3, pp. 986-995, March 2021, doi: 10.1109/TMI.2020.3043641.

[16] J. Brogan et al., "Fast Local Spatial Verification for Feature-Agnostic Large-Scale Image Retrieval," in IEEE Transactions on Image Processing, vol. 30, pp. 6892-6905, 2021, doi: 10.1109/TIP.2021.3097175.

[17] S. Roy, E. Sangineto, B. Demir and N. Sebe, "Metric-Learning-Based Deep Hashing Network for Content-Based Retrieval of Remote Sensing Images," in IEEE Geoscience and Remote Sensing Letters, vol. 18, no. 2, pp. 226-230, Feb. 2021, doi: 10.1109/LGRS.2020.2974629.

[18] H. Zhao, L. Yuan, H. Zhao and Z. Wang, "Global-Aware Ranking Deep Metric Learning for Remote Sensing Image Retrieval," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 8008505, doi: 10.1109/LGRS.2021.3059908.

[19] Y. Ghozzi, N. Baklouti, H. Hagras, M. B. Ayed and A. M. Alimi, "Interval Type-2 Beta Fuzzy Near Sets Approach to Content-Based Image Retrieval," in IEEE Transactions on Fuzzy Systems, vol. 30, no. 3, pp. 805-817, March 2022, doi: 10.1109/TFUZZ.2021.3049900.

[20] Z. Zinonos, S. Gkelios, A. F. Khalifeh, D. G. Hadjimitsis, Y. S. Boutalis and S. A. Chatzichristofis, "Grape Leaf Diseases Identification System Using Convolutional Neural Networks and LoRa Technology," in IEEE Access, vol. 10, pp. 122-133, 2022, doi: 10.1109/ACCESS.2021.3138050.

[21] P. Staszewski, M. Jaworski, J. Cao and L. Rutkowski, "A New Approach to Descriptors Generation for Image Retrieval by Analyzing Activations of Deep Neural Network Layers," in IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 12, pp. 7913-7920, Dec. 2022, doi: 10.1109/TNNLS.2021.3084633.

[22] S. Gkelios, A. Kastellos, Y. S. Boutalis and S. A. Chatzichristofis, "Universal Image Embedding: Retaining and Expanding Knowledge With Multi-Domain Fine-Tuning," in IEEE Access, vol. 11, pp. 38208-38217, 2023, doi: 10.1109/ACCESS.2023.3267804.

[23] M. Rasoolijaberi et al., "Multi-Magnification Image Search in Digital Pathology," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 9, pp. 4611-4622, Sept. 2022, doi: 10.1109/JBHI.2022.3181531.

[24] R. J. Chu, N. Richard, H. Chatoux, C. Fernandez-Maloigne and J. Y. Hardeberg, "Hyperspectral Texture Metrology Based on Joint Probability of Spectral and Spatial Distribution," in IEEE Transactions on Image Processing, vol. 30, pp. 4341-4356, 2021, doi: 10.1109/TIP.2021.3071557.

[25] J. Ramalhinho, H. F. J. Tregidgo, K. Gurusamy, D. J. Hawkes, B. Davidson and M. J. Clarkson, "Registration of Untracked 2D Laparoscopic Ultrasound to CT Images of the Liver Using Multi-Labelled Content-Based Image Retrieval," in IEEE Transactions on Medical Imaging, vol. 40, no. 3, pp. 1042-1054, March 2021, doi: 10.1109/TMI.2020.3045348.

[26] E. S. Sabry et al., "Image Retrieval Using Convolutional Autoencoder, InfoGAN, and Vision Transformer Unsupervised Models," in IEEE Access, vol. 11, pp. 20445-20477, 2023, doi: 10.1109/ACCESS.2023.3241858.

[27] K. T. Ahmed, S. Jaffar, M. G. Hussain, S. Fareed, A. Mehmood and G. S. Choi, "Maximum Response Deep Learning Using Markov, Retinal & Primitive Patch Binding With GoogLeNet & VGG-19 for Large Image Retrieval," in IEEE Access, vol. 9, pp. 41934-41957, 2021, doi: 10.1109/ACCESS.2021.3063545.

[28] G. Ren, X. Lu and Y. Li, "Joint Face Retrieval System Based On a New Quadruplet Network in Videos of Multi-Camera," in IEEE Access, vol. 9, pp. 56709-56725, 2021, doi: 10.1109/ACCESS.2021.3072055.

[29] M. Huang, L. Dong, W. Dong and G. Shi, "Supervised Contrastive Learning Based on Fusion of Global and Local Features for Remote Sensing Image Retrieval," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-13, 2023, Art no. 5208513, doi: 10.1109/TGRS.2023.3275644.

[30] Y. Huang, J. Zhang, L. Pan and Y. Xiang, "Privacy Protection in Interactive Content Based Image Retrieval," in IEEE Transactions on Dependable

and Secure Computing, vol. 17, no. 3, pp. 595-607, 1 May-June 2020, doi: 10.1109/TDSC.2018.2793923.

[31] C. Iakovidou, N. Anagnostopoulos, M. Lux, K. Christodoulou, Y. Boutalis and S. A. Chatzichristofis, "Composite Description Based on Salient Contours and Color Information for CBIR Tasks," in IEEE Transactions on Image Processing, vol. 28, no. 6, pp. 3115-3129, June 2019, doi: 10.1109/TIP.2019.2894281.

[32] A. Preethy Byju, B. Demir and L. Bruzzone, "A Progressive Content-Based Image Retrieval in JPEG 2000 Compressed Remote Sensing Archives," in IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 8, pp. 5739-5751, Aug. 2020, doi: 10.1109/TGRS.2020.2969374.

[33] S. Wei, L. Liao, J. Li, Q. Zheng, F. Yang and Y. Zhao, "Saliency Inside: Learning Attentive CNNs for Content-Based Image Retrieval," in IEEE Transactions on Image Processing, vol. 28, no. 9, pp. 4580-4593, Sept. 2019, doi: 10.1109/TIP.2019.2913513.

[34] F. -R. Andrés, R. -M. Eduardo, A. -C. Carlos and L. -S. Fidel, "Image Retrieval System based on a Binary Auto-Encoder and a Convolutional Neural Network," in IEEE Latin America Transactions, vol. 18, no. 11, pp. 1925-1932, November 2020, doi: 10.1109/TLA.2020.9398634.

[35] Z. N. K. Swati et al., "Content-Based Brain Tumor Retrieval for MR Images Using Transfer Learning," in IEEE Access, vol. 7, pp. 17809-17822, 2019, doi: 10.1109/ACCESS.2019.2892455.

[36] B. Ferreira, J. Rodrigues, J. Leitão and H. Domingos, "Practical Privacy-Preserving Content-Based Retrieval in Cloud Image Repositories," in IEEE Transactions on Cloud Computing, vol. 7, no. 3, pp. 784-798, 1 July-Sept. 2019, doi: 10.1109/TCC.2017.2669999.

[37] K. Iida and H. Kiya, "Privacy-Preserving Content-Based Image Retrieval Using Compressible Encrypted Images," in IEEE Access, vol. 8, pp. 200038-200050, 2020, doi: 10.1109/ACCESS.2020.3035563.

[38] A. Ahmed and S. J. Malebary, "Query Expansion Based on Top-Ranked Images for Content-Based Medical Image Retrieval," in IEEE Access, vol. 8, pp. 194541-194550, 2020, doi: 10.1109/ACCESS.2020.3033504.

[39] H. Erfankhah, M. Yazdi, M. Babaie and H. R. Tizhoosh, "Heterogeneity-Aware Local Binary Patterns for Retrieval of Histopathology Images," in IEEE Access, vol. 7, pp. 18354-18367, 2019, doi: 10.1109/ACCESS.2019.2897281.

[40] C. Kang, L. Zhu, X. Qian, J. Han, M. Wang and Y. Y. Tang, "Geometry and Topology Preserving Hashing for SIFT Feature," in IEEE Transactions on Multimedia, vol. 21, no. 6, pp. 1563-1576, June 2019, doi: 10.1109/TMM.2018.2883868.

[41] L. K. Pavithra and T. Sree Sharmila, "An Improved Seed Point Selection-Based Unsupervised Color Clustering for Content-Based Image Retrieval Application," in The Computer Journal, vol. 63, no. 1, pp. 337-350, Jan. 2020, doi: 10.1093/comjnl/bxz017.

[42] A. F. S. Devaraj et al., "An Efficient Framework for Secure Image Archival and Retrieval System Using Multiple Secret Share Creation Scheme," in IEEE Access, vol. 8, pp. 144310-144320, 2020, doi: 10.1109/ACCESS.2020.3014346.

[43] A. F. S. Devaraj et al., "An Efficient Framework for Secure Image Archival and Retrieval System Using Multiple Secret Share Creation Scheme," in IEEE Access, vol. 8, pp. 144310-144320, 2020, doi: 10.1109/ACCESS.2020.3014346.

[44] Naushad Varish, Priyanka Singh, Prannoy Tugiti, Marella Hima Manikanta, Bhavana Yedlapalli, Abhishree Pappusetty, Hiren Kumar Thakkar, Gajendra Sharma, "Color Image Retrieval Method Using Low Dimensional Salient Visual Feature Descriptors for IoT Applications", *Computational Intelligence and Neuroscience*, vol. 2023, Article ID 6257573, 18 pages, 2023. https://doi.org/10.1155/2023/6257573

[45] Jun Yan, Hui Liu, "Analysis on the Multimedia Information Retrieval Algorithm Based on Enterprise Correlation Financial Analysis under the Background of Big Data", *Advances in Multimedia*, vol. 2023, Article ID 5375010, 10 pages, 2023. https://doi.org/10.1155/2023/5375010