Iraqi Journal for Computer Science and Mathematics

Journal Homepage: http://journal.esj.edu.iq/index.php/IJCM
e-ISSN: 2788-7421 p-ISSN: 2958-0544

An Overview of Content-Based Image Retrieval Methods And Techniques

M.H.Hadid[1] *, Qasim Mohammed Hussein[1], Z.T.Al-Qaysi [1], M.A.Ahmed [1], Mahmood M. Salih [1]

[1]Tikrit University, College of Computer Science and Mathematics, Computer Science Department, Iraq

*Corresponding Author: M.H.Hadid
DOI: https://doi.org/ 10.52866/ijcsm.2023.02.03.006
Received ; April 2023; Accepted May 2023 ; Available online July 2023

ABSTRACT: With the development of Internet technology and the popularity of digital devices, Content-Based Image Retrieval (CBIR) has been quickly developed and applied in various fields related to computer vision and artificial intelligence. Currently, it is possible to retrieve related images effectively and efficiently from a large-scale database with an input image. In the past ten years, great efforts have been made for new theories and models of CBIR, and many effective CBIR algorithms have been established. Content-based image retrieval helps to discover identical images in a big dataset that match a query image. The query image's representative feature similarities to the dataset images typically assist in ranking the images for retrieval. There are various past studies on different handicraft feature descriptors according to the visual features that describe the images: color, texture, and shape. However, deep learning has been the dominant alternative to manually planned feature engineering; it automatically takes the features from the data. The current work reviews recent advancements in content-based image retrieval. For a deeper understanding of the advancement, the explanation of current state-of-the-art approaches from various vantage points is also conducted. This review employs a taxonomy encompassing various retrieval networks, classification types, and descriptors and this study will help researchers make more progress in image retrieval.

Keywords: Image Retrieval, Content-Based Image Retrieval (CBIR), Handcraft, Traditional Learning, Deep Learning.

1. INTRODUCTION

Numerous devices for image capture have emerged in the digital environment of the present day. Using image processing techniques, it is now easier than ever to store a large number of images. Since the images and databases keep growing and getting bigger daily, we need new, fast, and effective ways to find images[1]. Investigating image retrieval has grown more intriguing, and getting the required images from a big and diverse library is difficult. Image retrieval technology is grouped into: Text-based Image Retrieval (TBIR) and Content-based Image Retrieval (CBIR) [2].Text-based Image Retrieval (TIBR) requires retrieving images by manually matching keywords provided for each image [3]. However, manual processing proved impracticable due to the following factors: first, the difference in the labels of the images by diverse annotators, because every user uniquely interpreted the images, making it challenging to obtain the correct ones through keyword searches. Second, a typographical error when giving keywords to images can potentially result in the retrieval of unwanted images. Thirdly, annotating images in big datasets is time-consuming and subjective [4]. CBIR is a method for retrieving images according to their visual information including color, texture, and shape [5, 6]. The primary objective of CBIR is to increase the efficiency of image indexing and retrieval, thereby decreasing the requirement for human participation in the indexing process [7]. Any performance of the image retrieval method relies on how similar are the computations to the images. In an ideal world, the way to figure out how similar two images are should be different, reliable, and quick [8].

This paper will focus on image retrieval based on content. The classification of current state-of-the-art methods from various angles also begins for a deeper understanding—It Will aid in advancing image retrieval research by discussing state-of-the-art methods. This review employs a taxonomy encompassing various networks, categories, descriptors, and classification types. The study described will help researchers make more progress in the field of image retrieval. The remaining sections are organized as follows. In Section 2, we explain content-based image retrieval. The third section contains a literature review on image retrieval. It is divided into four sections, the first focusing on handicraft, the second on deep learning, the

Corresponding author: ziadoontareq@tu.edu.iq

third on combining handicraft and deep learning, and the fourth on combining deep learning and traditional learning (classification, dimensionality reduction). Section 3 next comprises image retrieval techniques. In Section 4, the paper's conclusion is presented approaches. Below Figure (1) shows the taxonomy that this review uses for the categorization of the current image retrieval methods.

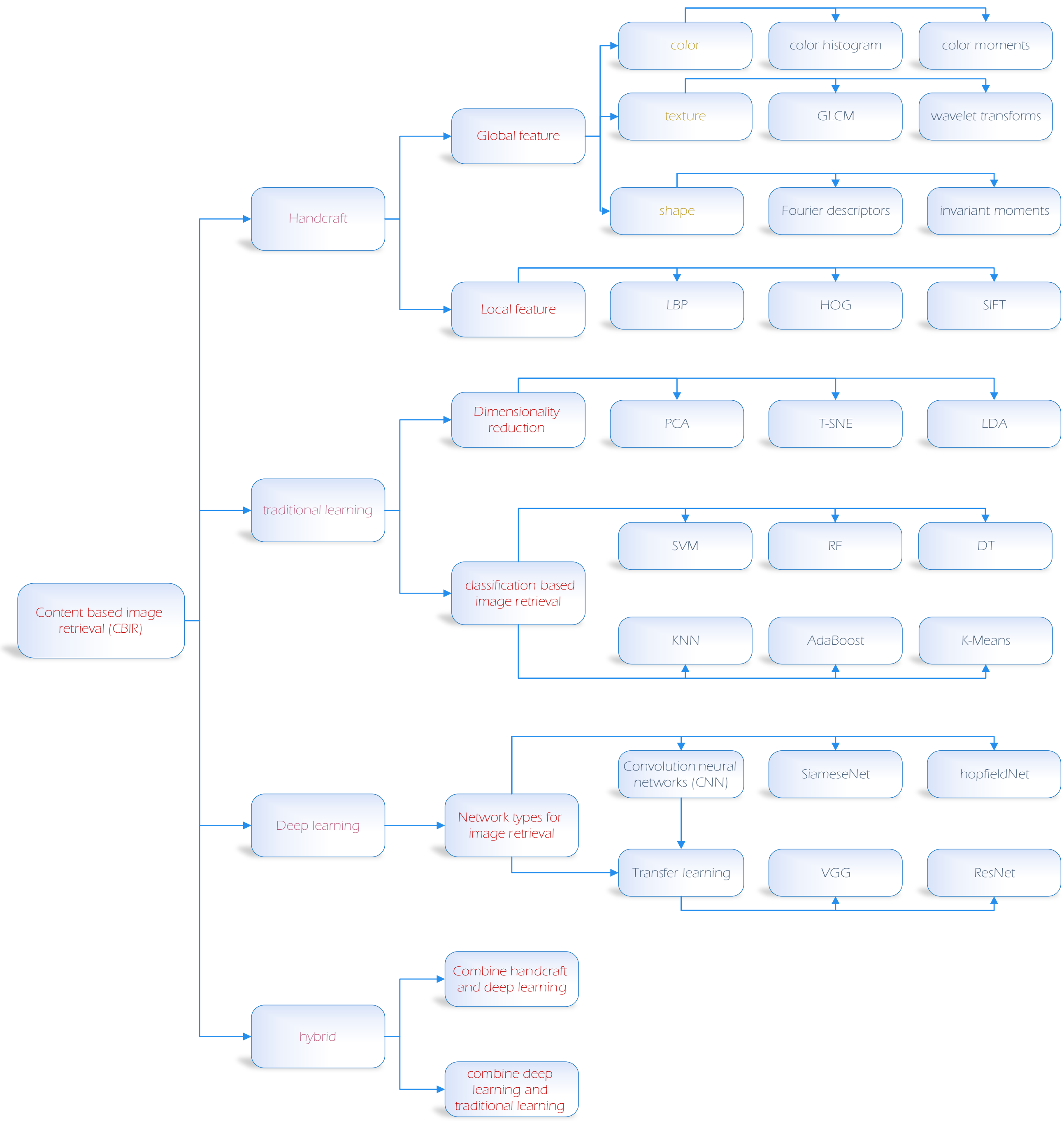

FIGURE 1. Taxonomy that this review uses for the categorization of the current image retrieval methods

2. CONTENT BASED IMAGE RETRIEVAL (CBIR)

CBIR is a significant and difficult problem applicable to m human endeavor industries [9]. CBIR is a method for searching an image collection using a query image. CBIR differs from other image retrieval systems because it accepts images as input rather than tags or text [10]. Numerous image retrieval techniques concentrate on low-level image characteristics such as textures, shapes, and colors [11]. Every CBIR system is typically divided into offline and online phases [12]. In the offline phases, features from large image collections (trains the system) create a database of features. This phase is typically lengthy, based on training image number utilized for training the system. During the online phases, similar features are extracted from the query image generating a distance metric between the query image's features and the database images' features for determining the similarity degree. The user is then shown images with a high similarity or a low level of distance [13]. Below Figure (2) shows the general of CBIR system

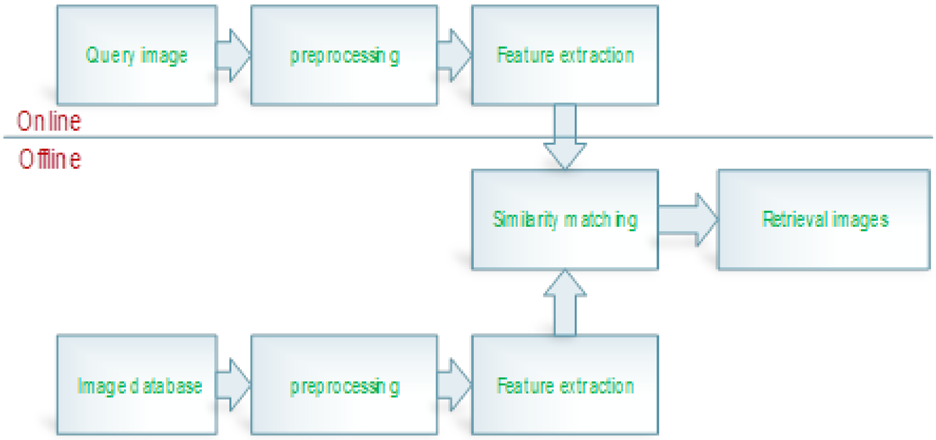

FIGURE 2. The general of CBIR system

3. EVOLUTION OF LITERATURE FOR CONTENT BASEB IMAGE RETRIEVAL

This section reviewed the literature review of three-year of image retrieval models (2020, 2021, and 2022).

3.1 HANDCRAFT

Mistry, Yogita D. in (2020) [14] proposed Combining color and texture descriptors including the binary Gabor pattern (BGP). In addition, Gabor wavelet texture features, and fuzzy histogram-based descriptors are also involved. It also includes segmentation-based fractal texture analysis (SFTA) and edge histogram descriptor (EHD). We used a Laplacian score to minimize the dimensionality of feature vectors. Using the Wang database, the retrieval rate is found using ten classes of 1,000 images.
Vimina, E. 1., & Divya, M. O. in (2020) [15] propose exploiting the RGB color space. MMLBP is a unique compact feature-generating approach for capturing multichannel texture data. Finally, a multi-channel texture-color descriptor for image representation for color image retrieval. The suggested algorithm has been used on benchmarked datasets Wang's 1 K, Corel 5 K, Corel 10 K, Coloured Brodatz Texture, and Zubud.
Ghahremani, Mona, et al. in (2021) [16] They suggested that the image be normalized and the median filter be used for noise reduction. Using the SLIC superpixel the color channel is altered. They are combining HOG and LBP methods for optimization purposes. These ideas are used to determine the image pattern and content-based image retrieval.
Hatibaruah, Rakcinpha, et al. in (2021) [17] they proposed a new low-dimensional feature descriptor referred to as local bit plane-based differences and adder pattern (LBPDAP). In each analyzed bit plane, the LBPDAP takes advantage of the center-neighboring pixel difference relationships and the neighboring pixel mutual difference connections. Using an adder, the collected dissimilarity information is joined to produce joint dissimilarity information. To create the proposed local patterns, the relationship between the center pixel and the joint dissimilarity information is encoded.
Kumar, N. R., & Kumar, Y. R. in (2021) [18] proposed the image retrieval process, integrating the Tamura texture features and Wavelet transform features. Calculating the Hausdorf similarity distance between each images and forming a matrix. On a dataset of Medical images.

Varish, Naushad. in (2022) [19] proposed the use of the probability histogram model; the color moments are retrieved from the image by the pre-processed HSV color components. The texture moments are obtained by determining the interrelationship between DCT blocks, and from these, GLCM-based statistics are produced, which provide crucial textual information about the image. Eventually, we calculated a unified feature descriptor through combining color, shape, texture, and feature moments. The suggested algorithm has been applied to the Corel-1K, OT-8, and GHIM-10K images dataset.

3.2 DEEP LEARNING

Pawar, Aashay. in (2020) [20] proposed utilizing Image-based search engines: The extension of the Content-Based Image Retrieval (CBIR) and evaluation of a Deep Learning strategy using an Autoencoder, a Convolutional Neural Network trained on a handwritten MNIST dataset. Autoencoder successfully showed quicker and more effective for Content-Based Image Retrieval.

Allegretti, Stefano, et al. in (2021) [21] they suggested a new content-based image retrieval system for dermoscopic image analysis that utilizes extracted features from Convolutional Neural Networks to collect related images. The designed method includes ResNet-50 initially trained to divided the dermoscopic images; the extraction portion of the feature which is then segregated, and an embedding network on top. The embedding obtains a different demonstration enabling the examination of image similarities

Kokilambal, S. in (2021) [22] proposed a novel CBIR model using a ResNet 50 model-based feature extractor. In addition, the Adadelta optimizer is used to effectively tune the ResNet-50 model's hyperparameters to improve performance of the retrieval. Also, Euclidean distance functions as a similarity metric for the identification of the images in the database that are highly similar to the Query image on the Corel benchmark dataset.

Staszewski, Paweł, et al. in (2021) [9] suggested a unique approach for the CBIR system according to the neural codes (neural activations) of deep neural networks utilizing the VGG16, ResNet50 neural network. The descriptors comprise information derived from the activation of convolutional and fully connected network layers, typically employed for classification tasks. Because of this technique, the images from the dataset are semantically identical to the query images and other features, including texture, color distribution, and image background. The suggested algorithm has been applied to the ImageNet1M images dataset.

Zhang, Kai, et al. in (2022) [23] they proposed Combining content-based image retrieval with convolutional Siamese neural networks developed CBIR-CSNN models. This CBIR-CSNN performs well when distinguishing CT images. And they used different distance functions.

Eswaran, A., & Varshini, E. in (2022) [24] proposed employing an image search engine system in several pre-trained CNN models for image feature extractions and aimed at to make very small changes to the chosen ResNet-50 model for the custom dataset in order to make it work as well as possible. The matching of feature resemblance was achieved by picking a similarity measure using Approximate Nearest Neighbor (Annoy) Indexing, with time and precision adjusted.

3.3 COMBINING HANDCRAFT AND DEEP LEARNING

Devulapalli, Sudheer, et al. in (2021) [25] proposed built a feature vector for Image Retrieval by combining deep learning and handcraft features. They act as a pre-trained feature extractors and fewer feature dimensions than other deep learning models, so Googlenet was used. Using the Gabor filter, custom-built features are retrieved and combined. Also, for the calculation of the Euclidean distance between the query image and the input image, the resultant feature vectors are utilized.

Pathak, D., & Raju, U. S. N. in (2021) [26] suggested a framework for efficient image retrieval by fusing high-level features of upgraded DarkNet-53, GN-Inception-Darknet-53 (GN-Inception-Darknet-53), with handcrafted features taken from both RGB and HSI color space. DDBTC is a low-level characteristic of the RGB color space. On RGB color images, HOG is used to extract shape features. DSP and GLCM are utilized to extract local texture features. The proposed method has applied on the the three color texture image datasets VisTex, Color Brodatz and Stex, and the three natural image datasets Corel-1 K, Corel-10 K and Corel-5 K.

Yelchuri, Rajesh, et al. in (2022) [27] proposed deep and hand-crafted features for constructing new texture image retrieval systems by many experiments conducted to extract deep features by many popular Convolutional Neural Network architectures (AlexNet, InceptionV3, VGG16, and ResNet50) to predict the class membership of the query image to every output class. Also, the similarities of these images and every images in the database are calculated by a adjusted distance matrix in the hand-crafted space wavelet features by normalized city-block

Karthik, T. S., et al. in (2022) [28] they propose (EOCBIR-HFSN) approach using handcraft features based on local binary patterns (LBP) and SqueezeNet-based deep features. In addition, the grasshopper optimization process is used to fine-tune the SqueezeNet model's hyperparameters (GOA). Test this method using the COREL-10K dataset.

3.4 COMBINE DEEP LEARNING AND TRADITIONAL LEARNING

Kumar, R. B., & Marikkannu, P. in (2020) [29] proposed from the image database, features of images including textures, shapes, standard deviations, and means are extracted. The K-means approach for clustering extracted features is provided. From the cluster, the proposed PSO-ANN classifier receives features as input. Through this enhanced classifier, we retrieve the related images to the query ones.

Jha, J., & Bhaduaria, S. S. in (2020) [30] propose Use RGB Histogram (fast) color features and Histogram of Oriented Gradient texture data; the method extracts handcrafted characteristics. We utilized unsupervised machine learning, k-means clustering, to create a perfect cluster of images with minimal outliers. This approach has sped up the retrieval process by omitting the entire search for similarity matches in the dataset. The recommended modification has been made to the Dataset of Historical Monuments.

Satish, B., & Supreethi, K. P. in (2021) [31] propose the gathered data is preprocessed using normalization, and HOG is used to extract features. Then, the suggested ICNN is employed to classify query images, leading to the classification of relevant images. The suggested ICNN methodology is simple, fast, and order-independent. The new approach performs better than the existing algorithms on high-dimensional training setting regarding the learning scaling behavior, for developing the ICNN technique classification and retrieval outcomes.

Desai, Padmashree, et al. in (2021) [32] they propose CNN is used for feature extraction in Content-Based Image Retrieval systems, while SVM is used for classification. The application of SVM decreased the time necessary to get the results. SVM will calculate the distance between the query image's features and the complete dataset's features. Using Corel 1 K dataset.

Karthik, K., & Kamath, S. S. in (2021) [33] They suggested an Model for retrieval of medical images utilizing a classification-based content-based image retrieval technique. The approach consists of a CNN-based model for categorizing medical images; the findings are used to facilitate the retrieval of related images. It was observed that the model produced effective retrieval outcomes.

Khan, Umer Ali, et al. in (2021) [34] proposed A CBIR technique according to the genetic algorithm (GA), the hybrid features descriptor, and the support vector machine (SVM) classifier for image retrieval in a multi-class scenario. Precisely, we utilized the first three color moments, Daubechies Wavelet, Haar Wavelet, and Bi-Orthogonal wavelets, for the extraction of feature, refined the features by GA, and trained the multi-class SVM by one-against-all strategies: on 4 standard datasets—WANG, CIFAR-10, Oxford Flower, and kvasir.

Reena, M. R., & Ameer, P. M. in (2022) [35] proposed Deep learning features collected from the activations of the feature layer in a pre-trained network named ResNet-101, followed by dimensionality reduction with t-Distributed Stochastic Neighbor Embedding (t-SNE), are used to pick discriminant features for the image retrieval system. To obtain similar images from a database, Euclidean distance is utilized for the calculation of the similarity.

4. IMAGE RETRIEVAL METHODS

This paper presents CBIR methods that extract information or features from the images.

4.1 HANDCRAFT

The content of an image is recognizable by visual components such as texture, color, and shape in the human visual system. So, this paper presents CBIR methods that extract information in global and local manners.

4.1.1 GLOBAL IMAGE FEATURES HANDCRAFT

In content-based image retrieval tasks, color, texture, shape, and spatial information are the most frequently employed features that describe the image as a whole [36].

4.1.1.1 COLOR FEATURE

The color feature is a fundamental visual feature used to retrieve images based on their color similarity. It is relatively resistant to background problems and independent of image size, orientation, and scale variations. Humans recognize images mainly based on their colors [37, 38]. Color features are the most intuitive and most dominant low-level image features which are very stable and robust in comparison with other image features such as texture and shape. Since these features are not sensitive to rotation, translation, and scale changes, they could apply to CBIR systems. What is more, the color feature calculation cost is relatively lower than other features [39].

A. COLOR HISTOGRAM

The color descriptor has been frequently utilized in CBIR for image retrieval due to its stable and robustness. Image translation, rotation, and resizing have no effect on it. The color histogram is the most popular and straightforward way. The color histogram describes the probabilities of the three color channels' intensity. The Histogram of an image is calculated by counting the number of each color pixel in the image, with each color pixel included in separate bins [40]. The histogram is a graph that shows the number of color values falling in a number of resolution ranges or bins. From the image histogram, a set of features could be extracted which are named color histogram features [39].

B. COLOR CO-OCCURRENCE MATRIX

The color co-occurrence matrix (CCM) is a typical technique for recording color variations in an image, which provides color characteristics. It calculates the chance of each pixel sharing the same color with its neighboring pixel. Each pixel in an image corresponds to four adjacent pixels, allowing each image to be represented by four image motifs of the scan pattern, which may then be assembled into four two-dimensional matrices [37]. Color texture analysis using color Co-occurrence Matrix (CCM) method is based on the hypothesis that the use of color features in the visible spectrum provides additional image characteristics over the traditional grey-level representation [41].

4.1.1.2 TEXTURE FEATURE

Textures are visible patterns with similar unachievable features with a single color or intensity. Texture provides essential information on the structural organization of surfaces and their relation to their surroundings. The human eye recognizes the many texture features, such as regularity, directionality, smoothness, and coarseness [38, 42]. Another important element in visual perception is "Texture". Texture also can be used to separate regions of interest in an image. Texture is also one of the most used low-level visual features that refer to innate surface properties of an object and their relationship to the surrounding environment and it also contains important information about the structural arrangement of surfaces and their relationship with the surroundings. Texture can be defined as, "a region in the image has a constant texture if a set of local statistics or other local properties of the picture are constant, gradually differing or approximately periodic" [43].

A. GRAY-LEVEL CO-OCCURRENCE MATRIX (GLCM)

GLCM) known as the Gray-Level Spatial Dependence Matrix (GLSD Matrix) is a statistical approach for analyzing texture determining the spatial link between pixels [44]. Gray-level co-occurrence matrix (GLCM) shows the spatial relationship between the pixels of a grayscale image. GLCM determines how many times two pixels with fixed values are adjacent. Two pixels are adjacent, if they are situated side by side horizontally, vertically or in any other direction [45].

B. WAVELET TRANSFORM

Using a multi-resolution method, the wavelet transform [46] incorporates texture analyses and categorizations. In wavelet transform, an image is deconstructed based on shifted and stretched wavelet processes. Applying the filter bank to all columns of the image and rows of the resultant coefficient facilitate a discrete wavelet transform in two dimensions [37]. Wavelet transform could be applied to images as 2-dimensional signals. To refract an image into k level, first, the transform is applied on all rows up to k level while columns of the image are kept unchanged. Then this task is applied to columns while keeping rows unchanged. In this manner frequency components of the image are obtained up to k level. These components are LL which is an approximation of the image and HL, LH, and

HH which are horizontal, vertical, and diagonal frequency details. These frequency components in various levels let us to better analyze the original image or signal [39].

4.1.1.3 SHAPE FEATURE

Important for recognizing and distinguishing real-world items are shaped features (SF). They are humans' most prominent visual cues for checking/matching similarity. Two types of SF exist region-based (RB) and contour-based (CB). RB pulls characteristics from the entire object, whereas CB calculates SF from the object's boundary [47]. The extraction of shape features is performed primarily to capture the shape properties (such as moment, region and boundary, etc.) of the image items. Such extraction facilitates the storage, transmission, comparison and identification of shape. It is crucial that the features of shape are robust to rotation, translation and scaling. Studies have been carried out to figure out the best way to discover the features of shape to allow the storage, transmission, or recognition of shape. In this regard, no mathematical transformation is involved in the shape features selected. For each pixel, the colored image has three values, and for extracting the shape features, the RGB color image is converted to grayscale image [48].

A. FOYRIER DESCRIPTORS

This technique requires applying Fourier transform to the image's shape boundaries. The coefficients of the Fourier transform are typically referred to as the shape's Fourier descriptors (FD)descriptors of the shape. They are durable and straightforward to derive. The noise does not affect Fourier descriptors [47]. Fourier transform is a classic technique in the field of conventional digital image processing that converts images to be processed from the spatial domain to the frequency domain. And then, the images further processed by various filters are converted back to the spatial domain by inverse Fourier transform. One of the most commonly used filters is the low-pass filter which attenuates the high-frequency content of the image in the frequency domain and preserves low-frequency content that contains most of the information, which can suppress noise and smooth (or blur) the images. In recent works, some improved variants of the Fourier transform and various filters are widely used in various sub-fields of image processing [49].

4.1.2 LOCAL IMAGE FEATURES HANDCRAFT

They are typically obtained by splitting the image into segments or calculating specific essential locations including corners, edges and blobs. Local characteristics are invariant concerning scale, rotation and translation [37]. It has also been demonstrated that local feature descriptors are more effective at enhancing the performance of the CBIR system. The LBP, HOG, scale-invariant feature transform (SIFT), and speeded-up robust feature (SURF) methods are commonly employed today for the extraction of the local image features.

A. LOCAL BINARY PATTREN (LBP)

The LBP code is the binary texture analysis approach version comparing the value of the central pixel to its surrounding pixels. LBP is resilient due to its invariant to monotonic grayscale modifications. LBP is also computationally uncomplicated. It is unable to retain global spatial information [50]. LBP was used mainly in texture classification but its computational ease leads it to be further used in medical imaging, image classification, object tracking, and facial expression recognition. This method uses a small window of an image. Each of the N neighboring pixels surrounding the center pixel is compared to the center pixel and a binary value (0 or 1) is assigned based on this intensity difference. The final result is obtained after multiplying these bits with specific weights. The center pixel is replaced with this value, the binary pattern value for that center pixel. Thus by replacing each center pixel with its binary pattern value a local binary map of the image is generated in its gray level. A histogram of this local binary map is calculated to create the feature vector [51].

B. HISTOGRAM OF ORIENTED GRADIENT (HOG)

Locally normalized feature descriptor that outperforms other well-known feature descriptors, including wavelet. HOG describes a local item's shape and appearance based on the direction of the object's edge or the distribution of local intensity gradients, even if specific information about corresponding angles or edge placements is not known [36]. The basic idea of the HOG algorithm is to use the gradient information of each pixel to extract discriminating features. HOG features are normally extracted from various window sizes in the image [52]. The process of computing HOG starts by dividing the image into cells and grouping the cells in the blocks. The gradient magnitude and gradient orientation are computed for all pixels within the block [53].

C. SCALE INVARIANT FEATURE TRANSFORM (SIFT)

The SIFT identifies extrema in scale space and extracts invariants of location, scale, and rotation. The SURF is an effective method of the SIFT using the factor of the Hessian matrix to detect feature points and hastens processes with integral graphs [50]. The SIFT detector extracts a number of attributes from an image in such a way that is reliable with changes in the lighting impacts and perspectives alongside other imagining viewpoints. The SIFT descriptor will distinguish nearby elements of an image [54]. SIFT was developed by David Lowe in 2004 as a continuation of his previous work on invariant feature detection, and it presents a method for detecting distinctive invariant features from images that can later be used to perform reliable matching between different views of an object or scene. Two key concepts are used in this definition: distinctive invariant features and reliable matching. The reason SIFT gives better and more reliable matching than those obtained from its predecessor descriptors lies in the cascade filtering approach that is used to detect the features which in return transforms the image data into scale-invariant coordinates relative to the local features. Lowe formulated SIFT algorithm into four major computational stages: a)Scale-Space Extrema Detection b)Keypoint Localization c)Orientation Assignment d)Keypoint Descriptor [55].

4.2 TRADITIONAL LEARNING

CBIR can be implemented using machine learning algorithms which can be used to alleviate any ambiguity in results. Can be used to enhance the learning process to abridge semantic gaps between the user and CBIR and abate the imbalance of classification problems, imbalance, and insufficient training set issue.

4.2.1 DIMENSIONALITY REDUCTION

Dimensionality reduction is the elimination of redundant information on high-dimensional properties. Typically, two methods reduce the dimension of a feature: feature extraction and mapping. Feature extraction is selecting a subset of high-dimensional features, whereas feature mapping is mapping high-dimensional features to low-dimensional features via a function [56]. The essential aspect of the CBIR system is the effective extraction of image features [57].

A. PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA is one of the most used data dimension reduction algorithms. It is an orthogonal feature transformation technique that transforms correlated sampled variables into uncorrelated linear sampled variables. These new variables are primary components because they resemble the original variables (PC). Features are linearly integrated to generate multiple characteristics, facilitating the computation of maximum variance [58].

B. T-DISTRIBUTED STOCHASTIC NEIGHBOR EMBEDDING (T-SNE)

t-SNE (t-distributed Stochastic Neighbor Embedding) is recognized as one of the most effective dimensionality reduction and data visualization technologies [59].It is a probabilistic strategy for visualizing high-dimensional data in 2D or 3D space, seeking to retain the local structure of the data as much as possible in the low-dimensional area. Despite the initial introduction of data visualization, the cluster high-dimensional data at any Euclidean distance has been expanded [60].

C. LINEAR DISCRIMINANT ANALYSIS (LDA)

It is a supervised, linear FEA. Nevertheless, other writers assert that LDA can also function as a linear classifier. LDA shows a new feature space to project data to maximize the separability of classes. It takes from a dataset's independent features best differentiate the types (dependent features). Therefore, the number of components produced is less than the number of classes [61].

4.3 IMAGE CLASSIFICATION

A. RANDOM FOREST (RF)

Random forest [62] is a simple, controlled classification method that functions fast on massive datasets, analyses tens to thousands of input instances without variable elimination, and recognizes essential classification characteristics. Regarding outliers and noise, it is reasonably robust. In contrast, each tree constructs itself differently from other trees in the same forest and frequently chooses its features randomly. After making the forest, each decision tree in the forest determines which class (for classification) this input data belong to, then the model chooses the most specific type (majority voting) [63].

B. SUPPORT VECTOR MACHINE (SVM)

It's a supervised machine learning technique that conducts categorization using labeled training data. The Support Vector Machine (SVM) is an efficient classifier [64, 65]. The low-level feature and the desired outcome play a crucial role in the data training process. In machine learning, the returned query image feature is used as input to determine the query image's image categorization. Examine the design of a CBIR system for non-texture image databases with multi-class SVM classifier integration. It is noted that the dimension and distance of the query image's feature vector relative to the database are dramatically reduced, resulting in more similarity between classes [38].

C. K-NEAREST NEIGHBORS (KNN)

One of the simple algorithms is the nearest neighbor algorithm, where the data are distributed in the hyperdimensional space, and the computation takes place depending on the "K" nearest neighbors by simply copying the labels of the weight without any specific effects on computation for prediction. Even though it is computationally intensive, it is very simple to implement [66]. The closest neighbor has the most similar data [67].

D. K-MEANS CLUSTERING

K-means is one of the most straightforward unsupervised learning techniques for solving the well-known clustering problem. A simple method is used to divide a given data set into a specified set clusters. The central concept establishes k centers, one for each cluster. These centers are located ingeniously as various locations produce varied outcomes. Therefore, arranging them as far apart as feasible would be preferable [68].

4.4 DEEP LEARNING

Deep learning is a subclass of machine learning consisting of a family of machine learning algorithms that aim to model high-output data by employing deep structures constructed with several indirect conversions [69].The application of image retrieval has dominated the majority of deep learning experiments [70].It has been demonstrated that deep learning gives superior performance in data processing due to improved image accuracy. Among the deep learning algorithms, convolutional neural networks are suitable for image [71].

4.4.1 NETWORK TYPES FOR IMAGE RETEIEVAL

4.4.1.1 CONVOLUTIONAL NEURAL NETWORKS (CNN)

In CBIR, convolution neural networks are crucial permitting the extraction of particular traits utilized for the creation of an effective description [72]. Similarity assessments based on diverse information, including shape, text, color, and texture, as well as characteristics gathered inside convolutional neural networks (CNNs), are utilized to bridge the semantic gap between imaging features and high-level visual concepts [73]. There are three basic types of layers, each consisting of three layers: convolution layer (Conv layer) (feature extraction), pooling layer (dimensionality reduction), and fully connected layer [74]. The Conv layers are comprised of a sequence of convolution kernels, each extracting various features from the input images, beginning with simple features including edges and forms and progressing to more complicated and specific features [75]. The activation layer is a non-linear operation of CNN models learning the non-linear representation of the output volumes of the previous Conv or FC layer by non-linear activation functions: Sigmoid, Tanh, or rectified linear activation function (ReLU) [76, 77]. Using non-linear numerical operations such as average-pooling, sum-pooling, or max-pooling, the POOL layer can reduce the spatial dimensions of the derived feature maps and the number of network parameters. A max-pooling function, for instance, separates input data into non-overlapping rectangles and stores the maximum value for each region [78]. A layer that is completely connected resembles a flattened vector and acts as a bridge between the two-dimensional convolutional and the one-dimensional Softmax layers. The latter receives information from the completely connected layer and computes the possibilities of every class by normalized exponential functions, and returns the classification result for the course with the highest possibility [79, 80]. Below Figure (3) shows the Illustration of a typical architecture of CNN.

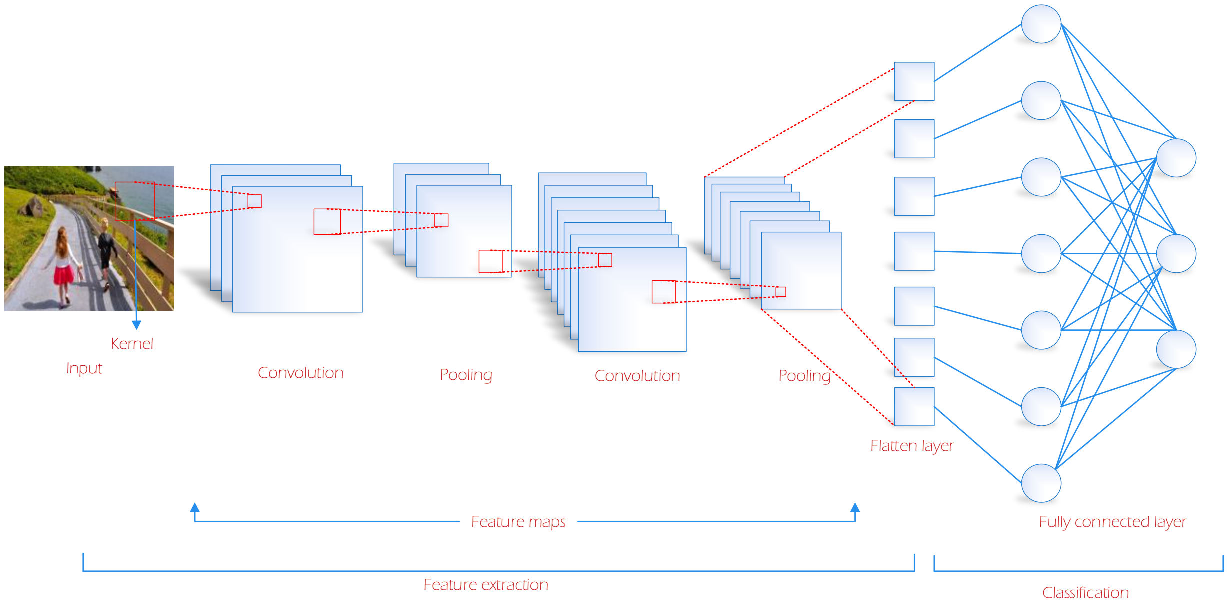

FIGURE 3. Illustration of a typical architecture of CNN.

4.4.1.1.1 TRANSFER LERNING (CNN PRETRAINED MODEL)

A. VGG-16

VGG (Visual Geometry Group) model with 16 network layers. The dataset they used is the ILSVRC 2012 dataset. The model exceeds the previous generation of models, which won the ILSVRC-2012 and ILSVRC-2013 competitions [57]. The Visual Geometry Group network (VGG-16) serves as a highly accurate feature extractor. The input image to the VGG-16 network is of fixed size, i.e., 3×224×224. It is passed through a stack of various convolutional layers of different receptive fields. The stride rate for convolutional layers and pooling layers remains the same throughout the VGG-16 network which is 3×3 with stride 1 in the convolutional layer and 2×2 with stride 2 in the pooling layer. The first two convolutional layers have 64 and 128 filters, respectively. The rest of the convolutional layers include 256, 512, and 512 filters, respectively. Border pixels are padded before each convolutional operation, which can preserve the feature maps size the same as the input. The VGG-16 ended with three fully connected layers. The first two FC layers consist of 4096 neurons while the final FC layer compresses the features to 1000 dimensions [81].

B. ResNet

The second pre-trained CNN model is ResNet, a convolutional neural network that consists of several layers deep. Both this model and the previous model are trained on more than a million images from the ImageNet database [82]. Deep residual network or ResNet. This architecture is formed to defeat quandaries in deep learning training because deep learning training, in general, takes quite a lot of time and is limited to a certain number of layers. The explication of the intricacy

introduced by ResNet is to apply to skip connection or shortcut. The advantage of the ResNets model compared to other architectural models is that the performance of this model does not decrease even though the architecture is getting deeper. Besides, computation calculations are made lighter, and the ability to train networks is better [83]. This approach is predicated on the notion that the next layer must learn new and distinct information from the prior input. [65].

4.4.1.2 HOPFIELD NETWORK

A Hopfield net is a recurrent neural network with a synaptic connection pattern such that the activity dynamics are governed by an underlying Lyapunov function. State of the system evolves to a (local) minimum of the Lyapunov function from any initial state [84].

4.4.1.3 SIAMESE NETWORK

A Twin/Siamese network is a form of neural network that uses the distance between image pair feature [8]. It consists of two indistinguishable subnetworks with a similar configuration and shared characteristics. These networks achieved widespread popularity for tasks involving the identification of similarities between two data points with their corresponding identical subnetworks, with a third module combining their outputs to provide the last results [85] .

5. CONCLUSION

The main focus of this study is to present an overview of different techniques that are applied in different research models. it is summarized that image features representation is done by the use of low-level visual features such as color, texture, and shape called handcraft and high-level visual features such as human interpretation. CBIR and image representation are with traditional machine learning approaches that have shown good results in various domains. Optimization of feature representation in terms of feature dimensions can provide a strong framework for the learning of classification-based models and it will not face the problems like overfitting, also several research in CBIR, utilize deep neural networks and they have shown good results on many datasets. this work is a recent review of the methods of image retrieval. where this article focuses primarily on content-based image retrieval in the years 2020, 2021, and 2022. A complete taxonomy is offered in various retrieval methodologies, network types, descriptor types, and classifications types. It is clear that CBIR has made great progress in theory, technology, and application. However, there are still many challenges, especially with the emergence of big data and the utilization of deep learning techniques, which will play an important role in artificial intelligence in the future for our daily life.

Funding

None.

ACKNOWLEDGEMENT

None.

CONFLICTS OF INTEREST

None.

REFERENCES

[1] O. Mohamed, E. A. Khalid, O. Mohammed, and A. Brahim, "Content-based image retrieval using convolutional neural networks," in Lecture Notes in Real-Time Intelligent Systems, 2019, pp. 463-476.

[2] J. Yao, Y. Deng, Y. Yu, and C. Sun, "A fast image retrieval method with convolutional neural networks," in 2017 36th Chinese Control Conference (CCC), 2017, pp. 11110-11115.

[3] J. Pradhan, A. Ajad, A. K. Pal, and H. Banka, "Multi-level colored directional motif histograms for content-based image retrieval," The Visual Computer, vol. 36, pp. 1847-1868, 2020.

[4] M. Alshehri, "A content-based image retrieval method using neural network-based prediction technique," Arabian journal for science and engineering, vol. 45, pp. 2957-2973, 2020.

[5] V. Tyagi, "Content-based image retrieval," Springer Nature, 2017.

[6] R. Ashraf, M. Ahmed, S. Jabbar, S. Khalid, A. Ahmad, S. Din, et al., "Content based image retrieval by using color descriptor and discrete wavelet transform," Journal of medical systems, vol. 42, pp. 1-12, 2018.

[7] S. A. Wadhai and S. S. Kawathekar, "Techniques of content based image retrieval: a review," IOSR J Comput Eng (IOSR-JCE), pp. 75-79, 2019.

[8] S. R. Dubey, "A decade survey of content based image retrieval using deep learning," IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, pp. 2687-2704, 2021.

[9] 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," IEEE Transactions on neural networks and learning systems, vol. 33, pp. 7913-7920, 2021.

[10] F. Mustafic, I. Prazina, and V. Ljubovic, "A new method for improving content-based image retrieval using deep learning," in 2019 XXVII international conference on information, communication and automation technologies (ICAT), 2019, pp. 1-4.

[11] A. J. Trappey, C. V. Trappey, and S. Shih, "An intelligent content-based image retrieval methodology using transfer learning for digital IP protection," Advanced Engineering Informatics, vol. 48, p. 101291, 2021.

[12] X. Li, J. Yang, and J. Ma, "Recent developments of content-based image retrieval (CBIR)," Neurocomputing, vol. 452, pp. 675-689, 2021.

[13] A. Qayyum, S. M. Anwar, M. Awais, and M. Majid, "Medical image retrieval using deep convolutional neural network," Neurocomputing, vol. 266, pp. 8-20, 2017.

[14] Y. D. Mistry, "Textural and color descriptor fusion for efficient content-based image retrieval algorithm," Iran Journal of Computer Science, vol. 3, pp. 169-183, 2020.

[15] E. Vimina and M. Divya, "Maximal multi-channel local binary pattern with colour information for CBIR," Multimedia Tools and Applications, vol. 79, pp. 25357-25377, 2020.

[16] M. Ghahremani, H. Ghadiri, and M. Hamghalam, "Local features integration for content-based image retrieval based on color, texture, and shape," Multimedia Tools and Applications, vol. 80, pp. 28245-28263, 2021.

[17] R. Hatibaruah, V. K. Nath, and D. Hazarika, "Computed tomography image retrieval via combination of two local bit plane-based dissimilarities using an adder," International Journal of Wavelets, Multiresolution and Information Processing, vol. 19, p. 2050058, 2021.

[18] N. R. Kumar and Y. R. Kumar, "Fusion-based Medical Image Retrieval using Tamura and Wavelet features," in 2021 4th International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE), 2021, pp. 609-614.

[19] N. Varish, "A modified similarity measurement for image retrieval scheme using fusion of color, texture and shape moments," Multimedia Tools and Applications, vol. 81, pp. 20373-20405, 2022.

[20] A. Pawar, "Evaluation of autoencoder for CBIR system in deep learning," in 2020 IEEE 17th India Council International Conference (INDICON), 2020, pp. 1-4.

[21] S. Allegretti, F. Bolelli, F. Pollastri, S. Longhitano, G. Pellacani, and C. Grana, "Supporting skin lesion diagnosis with content-based image retrieval," in 2020 25th International Conference on Pattern Recognition (ICPR), 2021, pp. 8053-8060.

[22] S. Kokilambal, "Intelligent content based image retrieval model using adadelta optimized residual network," in 2021 International Conference on System, Computation, Automation and Networking (ICSCAN), 2021, pp. 1-5.

[23] K. Zhang, S. Qi, J. Cai, D. Zhao, T. Yu, Y. Yue, et al., "Content-based image retrieval with a Convolutional Siamese Neural Network: Distinguishing lung cancer and tuberculosis in CT images," Computers in biology and medicine, vol. 140, p. 105096, 2022.

[24] A. Eswaran and E. Varshini, "Reverse Image Search Engine for Garment Industry," in 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), 2022, pp. 414-418.

[25] S. Devulapalli, A. Potti, R. Krishnan, and M. S. Khan, "Experimental evaluation of unsupervised image retrieval application using hybrid feature extraction by integrating deep learning and handcrafted techniques," Materials Today: Proceedings, 2021.

[26] D. Pathak and U. Raju, "Content-based image retrieval using feature-fusion of GroupNormalized-Inception-Darknet-53 features and handcraft features," Optik, vol. 246, p. 167754, 2021.

[27] R. Yelchuri, J. K. Dash, P. Singh, A. Mahapatro, and S. Panigrahi, "Exploiting deep and hand-crafted features for texture image retrieval using class membership," Pattern Recognition Letters, vol. 160, pp. 163-171, 2022.

[28] T. Karthik, R. Krishna, T. R. Rao, V. Manoranjithem, S. Kalaiarasi, and B. Jegajothi, "Evolutionary Optimization Algorithm on Content based Image Retrieval System using Handcrafted features with Squeeze Networks," in 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), 2022, pp. 1425-1431.

[29] R. B. Kumar and P. Marikkannu, "An efficient content based image retrieval using an optimized neural network for medical application," Multimedia Tools and Applications, vol. 79, pp. 22277-22292, 2020.

[30] J. Jha and S. S. Bhaduaria, "A novel approach for retrieval of historical monuments images using visual contents and unsupervised machine learning," International Journal, vol. 9, 2020.

[31] B. Satish and K. Supreethi, "An independent condensed nearest neighbor classification technique for medical image retrieval," Journal of Ambient Intelligence and Humanized Computing, pp. 1-11, 2021.

[32] P. Desai, J. Pujari, C. Sujatha, A. Kamble, and A. Kambli, "Hybrid Approach for Content-Based Image Retrieval using VGG16 Layered Architecture and SVM: An Application of Deep Learning," SN Computer Science, vol. 2, pp. 1-9, 2021.

[33] K. Karthik and S. S. Kamath, "A deep neural network model for content-based medical image retrieval with multi-view classification," The Visual Computer, vol. 37, pp. 1837-1850, 2021.

[34] U. A. Khan, A. Javed, and R. Ashraf, "An effective hybrid framework for content based image retrieval (CBIR)," Multimedia Tools and Applications, vol. 80, pp. 26911-26937, 2021.

[35] M. R. Reena and P. Ameer, "A content-based image retrieval system for the diagnosis of lymphoma using blood micrographs: An incorporation of deep learning with a traditional learning approach," Computers in Biology and Medicine, vol. 145, p. 105463, 2022.

[36] I. M. Hameed, S. H. Abdulhussain, and B. M. Mahmmod, "Content-based image retrieval: A review of recent trends," Cogent Engineering, vol. 8, p. 1927469, 2021.

[37] J. M. Patel and N. C. Gamit, "A review on feature extraction techniques in content based image retrieval," in 2016 international conference on wireless communications, signal processing and networking (WiSPNET), 2016, pp. 2259-2263.

[38] N. K. Rout, M. Atulkar, and M. K. Ahirwal, "A review on content-based image retrieval system: Present trends and future challenges," International Journal of Computational Vision and Robotics, vol. 11, pp. 461-485, 2021.

[39] A. Rashno and E. Rashno, "Content-based image retrieval system with most relevant features among wavelet and color features," arXiv preprint arXiv:1902.02059, 2019.

[40] C. Irawan, W. Listyaningsih, C. A. Sari, and E. H. Rachmawanto, "CBIR for herbs root using color histogram and GLCM based on K-nearest neighbor," in 2018 International Seminar on Application for Technology of Information and Communication, 2018, pp. 509-514.

[41] Y. Hendrawan, M. Fauzi, N. Khoirunnisa, M. Andreane, P. Hartianti, T. Halim, et al., "Development of colour co-occurrence matrix (CCM) texture analysis for biosensing," in IOP Conference Series: Earth and Environmental Science, 2019, p. 012022.

[42] L. Yang, M. Gong, and V. K. Asari, "Diagram image retrieval and analysis: Challenges and opportunities," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 180-181.

[43] M. A. M. Shukran, M. N. Abdullah, and M. S. F. M. Yunus, "New approach on the techniques of content-based image retrieval (CBIR) using color, texture and shape features," Journal of Materials Science and Chemical Engineering, vol. 9, pp. 51-57, 2021.

[44] M. J. J. Ghrabat, G. Ma, I. Y. Maolood, S. S. Alresheedi, and Z. A. Abduljabbar, "An effective image retrieval based on optimized genetic algorithm utilized a novel SVM-based convolutional neural network classifier," Human-centric Computing and Information Sciences, vol. 9, pp. 1-29, 2019.

[45] M. F. Sadique and S. Haque, "Content-based image retrieval using color layout descriptor, gray-level co-occurrence matrix and k-nearest neighbors," International Journal of Information Technology and Computer Science, vol. 12, pp. 19-25, 2020.

[46] Z. Al-Qaysi, A. Al-Saegh, A. F. Hussein, and M. Ahmed, "Wavelet-based Hybrid Learning Framework for Motor Imagery Classification," 2022.

[47] A. O. Salau and S. Jain, "Feature extraction: a survey of the types, techniques, applications," in 2019 international conference on signal processing and communication (ICSC), 2019, pp. 158-164.

[48] M. K. Alsmadi, "Content-based image retrieval using color, shape and texture descriptors and features," Arabian Journal for Science and Engineering, vol. 45, pp. 3317-3330, 2020.

[49] Z. Zhou, X. Wang, C. Li, M. Zeng, and Z. Li, "Adaptive deep feature aggregation using Fourier transform and low-pass filtering for robust object retrieval," Journal of Visual Communication and Image Representation, vol. 72, p. 102860, 2020.

[50] S. Wang, K. Han, and J. Jin, "Review of image low-level feature extraction methods for content-based image retrieval," Sensor Review, 2019.

[51] A. K. Bhunia, A. Bhattacharyya, P. Banerjee, P. P. Roy, and S. Murala, "A novel feature descriptor for image retrieval by combining modified color histogram and diagonally symmetric co-occurrence texture pattern," Pattern Analysis and Applications, vol. 23, pp. 703-723, 2020.

[52] S. Ghaffari, P. Soleimani, K. F. Li, and D. W. Capson, "Analysis and comparison of FPGA-based histogram of oriented gradients implementations," IEEE Access, vol. 8, pp. 79920-79934, 2020.

[53] S. Abouzahir, M. Sadik, and E. Sabir, "Bag-of-visual-words-augmented Histogram of Oriented Gradients for efficient weed detection," Biosystems Engineering, vol. 202, pp. 179-194, 2021.

[54] S. Gupta, M. Kumar, and A. Garg, "Improved object recognition results using SIFT and ORB feature detector," Multimedia Tools and Applications, vol. 78, pp. 34157-34171, 2019.

[55] N. U. R. Malik, A. G. Airij, S. A. Memon, Y. N. Panhwar, S. A. Abu-Bakar, and M. A. El-Khoreby, "Performance comparison between SURF and SIFT for content-based image retrieval," in 2019 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), 2019, pp. 214-218.

[56] J. Qin, J. Chen, X. Xiang, Y. Tan, W. Ma, and J. Wang, "A privacy-preserving image retrieval method based on deep learning and adaptive weighted fusion," Journal of Real-Time Image Processing, vol. 17, pp. 161-173, 2020.

[57] R. Bibi, Z. Mehmood, R. M. Yousaf, T. Saba, M. Sardaraz, and A. Rehman, "Query-by-visual-search: multimodal framework for content-based image retrieval," Journal of Ambient Intelligence and Humanized Computing, vol. 11, pp. 5629-5648, 2020.

[58] P. Ray, S. S. Reddy, and T. Banerjee, "Various dimension reduction techniques for high dimensional data analysis: a review," Artificial Intelligence Review, vol. 54, pp. 3473-3515, 2021.

[59] M. Kimura, "Generalized t-SNE Through the Lens of Information Geometry," IEEE Access, vol. 9, pp. 129619-129625, 2021.

[60] G. L. Potratz, S. W. A. Canchumuni, J. D. B. Castro, J. Potratz, and M. A. C. Pacheco, "Automatic lithofacies classification with t-SNE and K-nearest neighbors algorithm," Anuário Do Instituto De Geociências, vol. 44, 2021.

[61] F. Anowar, S. Sadaoui, and B. Selim, "Conceptual and empirical comparison of dimensionality reduction algorithms (pca, kpca, lda, mds, svd, lle, isomap, le, ica, t-sne)," Computer Science Review, vol. 40, p. 100378, 2021.

[62] M. M. Salih, Z. Al-Qaysi, M. L. Shuwandy, M. Ahmed, K. F. Hasan, and Y. R. Muhsen, "A new extension of fuzzy decision by opinion score method based on Fermatean fuzzy: A benchmarking COVID-19 machine learning methods," Journal of Intelligent & Fuzzy Systems, pp. 1-11, 2022.

[63] O. S. Kareem, A. Al-Sulaifanie, D. A. Hasan, and D. M. Ahmed, "Segmenting and Classifiying the Brain Tumor from MRI Medical Images Based on Machine Learning Algorithms: A Review," Asian Journal of Research in Computer Science, vol. 10, pp. 50-61, 2021.

[64] M. AHMED, M. D. SALMAN, R. ADEL, Z. A.-Q. ALSHARIDA, and M. M. HAMMOOD, "AN INTELLIGENT ATTENDANCE SYSTEM BASED ON CONVOLUTIONAL NEURAL NETWORKS FOR REAL-TIME STUDENT FACE IDENTIFICATIONS," Journal of Engineering Science and Technology, vol. 17, pp. 3326-3341, 2022.

[65] S. M. Samuri, T. V. Nova, B. Rahmatullah, S. L. Wang, and Z. T. Al-Qaysi, "Classification model for breast cancer mammograms," IIUM Engineering Journal, vol. 23, pp. 187-199, 2022.

[66] T. Vijayan, M. Sangeetha, A. Kumaravel, and B. Karthik, "Feature selection for simple color histogram filter based on retinal fundus images for diabetic retinopathy recognition," IETE Journal of Research, vol. 69, pp. 987-994, 2023.

[67] Y. Chen, X. Zeng, X. Chen, and W. Guo, "A survey on automatic image annotation," Applied Intelligence, vol. 50, pp. 3412-3428, 2020.

[68] B. Mahesh, "Machine learning algorithms-a review," International Journal of Science and Research (IJSR).[Internet], vol. 9, pp. 381-386, 2020.

[69] S. Kumar, "A comprehensive study on deep learning approach for CBIR," in 2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT), 2021, pp. 560-564.

[70] I. Bansal and S. Dhar, "A Framework and Techniques for Image-based Search Application with an E-commerce

Domain," in 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), 2021, pp. 1102-1109.

[71] A. Sezavar, H. Farsi, and S. Mohamadzadeh, "Content-based image retrieval by combining convolutional neural networks and sparse representation," Multimedia Tools and Applications, vol. 78, pp. 20895-20912, 2019.

[72] 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," IEEE Transactions on neural networks and learning systems, 2021.

[73] K. Kobayashi, R. Hataya, Y. Kurose, M. Miyake, M. Takahashi, A. Nakagawa, et al., "Decomposing normal and abnormal features of medical images for content-based image retrieval of glioma imaging," Medical image analysis, vol. 74, p. 102227, 2021.

[74] Z. Rian, V. Christanti, and J. Hendryli, "Content-based image retrieval using convolutional neural networks," in 2019 IEEE International Conference on Signals and Systems (ICSigSys), 2019, pp. 1-7.

[75] Y. Liu, H. Pu, and D.-W. Sun, "Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices," Trends in Food Science & Technology, vol. 113, pp. 193-204, 2021.

[76] M. Paoletti, J. Haut, J. Plaza, and A. Plaza, "Deep learning classifiers for hyperspectral imaging: A review," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 158, pp. 279-317, 2019.

[77] M. Ahmed, Z. Al-qaysi, M. L. Shuwandy, M. M. Salih, and M. H. Ali, "Automatic COVID-19 pneumonia diagnosis from x-ray lung image: A Deep Feature and Machine Learning Solution," in Journal of Physics: Conference Series, 2021, p. 012099.

[78] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on Convolutional Neural Networks (CNN) in vegetation remote sensing," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 173, pp. 24-49, 2021.

[79] S. Sony, K. Dunphy, A. Sadhu, and M. Capretz, "A systematic review of convolutional neural network-based structural condition assessment techniques," Engineering Structures, vol. 226, p. 111347, 2021.

[80] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, et al., "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," Journal of big Data, vol. 8, pp. 1-74, 2021.

[81] F. Liu, Y. Wang, F.-C. Wang, Y.-Z. Zhang, and J. Lin, "Intelligent and secure content-based image retrieval for mobile users," IEEE Access, vol. 7, pp. 119209-119222, 2019.

[82] A. Ahmed, "Pre-trained CNNs models for content based image retrieval," International Journal of Advanced Computer Science and Applications, vol. 12, 2021.

[83] D. Sarwinda, R. H. Paradisa, A. Bustamam, and P. Anggia, "Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer," Procedia Computer Science, vol. 179, pp. 423-431, 2021.

[84] C. Bauckhage, R. Sanchez, and R. Sifa, "Problem solving with hopfield networks and adiabatic quantum computing," in 2020 International Joint Conference on Neural Networks (IJCNN), 2020, pp. 1-6.

[85] R. Kapoor, V. Malhotra, S. Yasser, and V. Wadhwa, "A Comparative Analysis of Capsule Networks in Content based Image Retrieval tasks: CapsCBIR," in 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 279-283.