June 2021

A Review on Recent Advances in Content-Based Image Retrieval used in Image Search Engine

Smita V. Bhoir Ms.
Department of Computer Engineering, K.J. Somaiya College of Engineering, Ramrao Adik Institute of Technology, Mumbai, Maharashtra, India, smitapatilbe@gmail.com

Sunita Patil
K.J. Somaiya Institute of Engineering and Information Technology, Mumbai, Maharashtra, India, spatil@somaiya.edu

A Review on Recent Advances in Content-Based Image Retrieval used in Image Search Engine

Abstract

Since the advent of visual data on the Web, there has been a more significant increase in image search activity. Lack of knowledge of visual content can lead to inconsistencies in methods that employ text retrieval. Searching an image and getting a relevant image is a challenging research issue for the computer vision community. The value of recent research on Content-Based Image Retrieval (CBIR) has gone up significantly in the last decade because it has focused on discovering relevant images. The first is the problem is due to the intention gap and the second is due to the semantic gap. A good deal of work has been done on CBIR, image classification, and interpretation of images in the last two decades. The paper presents a systematic overview of recent research in the field of CBIR paper. Also, the images have a lot of features that make them stand out from others, the comprehensive analysis of various feature extraction and image processing techniques such as color, texture, shape, low-level feature extraction, and recent machine and deep learning techniques is given in this paper. The prospective research directions are explored in-depth and finally concluded in order to stoke interest in further research in this field.

Keywords: Image Search Engine; Web Crawler; Content-Based Image Retrieval; Feature Extraction; Machine Learning; Deep Learning;

1 Introduction

The World Wide Web is a digital medium of information in which information is explored around the world by as many people as possible. The search engine (Hernández, Rivero, and Ruiz, 2019) is a tool used by internet users to search for the results of their needs and return pages or images from searchable locations (e.g. website, a searchable image, or video). Now here come web crawlers or web spiders, which scan the internet looking for web-accessible information and copy the results into a centralized location where they can be easily analyzed by the search engine (Vishakha, 2019)(Vagač *et al.*, 2016). Web spider-based applications are making their way into more uses in various search engines, scraping data, and cyber security, so the greater the daily value and frequency of their use, the more spiders become (Samuel, 2019). Billions of people are expected to post and view digital images with the advancement of digital devices and the expansion of the internet. Since having equal access to digital images and the Internet expands the ways that we can search images, there are numerous new applications that we didn't know about. The algorithm searches for visual documents in an unstructured document collection in which it is easy to search for images of all types of text or images which match the search terms (Hanjalic, 2012). While image searching has been thoroughly investigated since the early 1990s, other techniques have yet to truly flourish due to a lack of user attention, but these other scalability issues have still opened the door for advances in computer vision and multimedia. Without going through the hassle of uploading and reindexing each piece of visual content, such as a piece of artwork, a traditional image search engine will only look at surrounding metadata, like titles and tags (Hwang, Lee, and Ha, 2020). Instead of attempting to return an accurate image for a given query, the algorithm that does most of the keyword expansion looks for textual information and returns relevant images(Hwang, Lee, and Ha, 2020)(Guinness, Cutrell, and Morris, 2018). This approach has shown to be more effective in the past. More specifically, content-based visual retrieval is difficult due to the inherent absence of comprehension and the semantic gap. This difference intends to make a point about the difficulties faced by the user. A useful question expands into an accurate picture or sketch illustration that helps to represent the expected results or impacts the reader can see. To obtain, the index, you should always run the file through a keyword and map file browser first, as an O(N*log(N) approach (O is the number of random locations that each element must be queried) works as opposed to an O(N**2) approach that takes the logarithm of the number of locations to compute each variation(Yang *et al.*, 2019). It develops at a high level due to the ambiguity in articulating the low-level visual processes concerning the higher-level cognitive functions. A lot of attempts have been made to close the gaps on both sides of the education and business sectors to expand these paraphrases. The main goal of this paper is to trace recent developments in Content-Based Image Retrieval while searching image data. These recent developments in CBIR modeling, especially those associated with machine learning and deep learning, are explained in detail with current concepts for computer vision. Additionally, future work in this area may include the expansion of studies that promote more research.

The organization of the paper is in this manner, and now follow is the expansion: In Section 2 Image search engine and the current issues in its working is given. In Section 3 CBIR system and its workflow along with feature extraction methods required are discussed. SECTION 4 discusses color-based feature extraction, SECTION 5 covers texture-based feature extraction, SECTION 6 mentions CBIR-shape feature extraction and SECTION 7 shows additional techniques for CBIR low-level feature extraction. SECTION 8 addresses the existence of extensive use of basic Machine Learning techniques and discusses results from using them, and SECTION 9 highlights how CBIR-based learning has been applied successfully in the world of commerce and the conclusion of the study is given. Finally, SECTION 10 ends with future recommendations in CBIR.

2 Image Search Engine

Multimedia Information Retrieval (MIR) has emerged as an even more significant technology in the exponential rate of technology growth, which gives rise to significant amounts of multimedia content with ubiquitous access to both the images and videos. The increased complexity in Multimedia Information Retrieval and satisfying semantically rich, contextual, and individual needs have demanded level of quality is becoming a priority(Hanjalic, 2012). This concept can be quite challenging for people with poor vision, there are many difficulties for individuals who are trying to comprehend web-how-manage images. Some of how images are made more usable are through the use of alternative text in HTML markup. This method is often called the "alt" alternative text used in the "img" tag of HTML which is defined as the "alt" attribute. Conversely, however, this technique has only been applied by a small percentage of people (Guinness, Cutrell, and Morris, 2018). This estimate based on past studies shows that about half of the images on highly trafficked websites are also described using ALT attributes. Although alt text image information is required for search engine indexing, it may also be manipulated for purposes other than describing an image, thus doing so will often yield incorrect descriptions, which may either call an image by its filename. The browser renders an image to the specified dimensions in its original context while a visual description might not be available in another context. The same image appears differently in various parts of the page contexts, resulting in an inconsistent browsing experience for people who rely on screen reader assistive technology. The primary challenges in content-based visual retrieval are the lack of an accurate representation of intention the fault lies in the lack of identifying intent in the search results. The intention gap refers to the difficulty that a user suffers to precisely express the expected visual content by a query at hand and the semantic gap originates from the difficulty in describing high-level semantic concept with the low-level visual feature(Amelio, 2019)(Celik and Bilge, 2017). It will be important to narrow those gaps if extensive efforts are to be put in place. The extensive literature survey conducted has shown that there is a need for a more efficient system for obtaining information about individual users' specific data from the web. CBIR techniques can assist in expanding your search results.

3 **Content-based Image Retrieval (CBIR)**

Content-based Image retrieval examines visual contents in search results, instead of specific visual features of the image. To ensure that CBIR includes a query image as input, it computes visual content and visually similar images proximity in the feature vector is used to identify query-to-feature images. At the pixel level, some attributes control how the output looks (such as color, shape, and texture) but sorting is performed by the feature values that are found in the CBIR(Zhang, Islam and Lu, 2012). The implementations of the above models in these areas have seen the incorporation of CBIR and feature extraction methods into numerous other fields such as medical imaging, meteorology, surveillance, and video analytics to significantly increase their use. The description of the fundamental concepts and mechanisms of image searching (Tian, 2013) (Zhang and Lu, 2004)(Zhang, Islam and Lu, 2012) is provided in Figure 1.

The major requirement of any image retrieval system is to be able to quickly and effectively locate images that do not have human involvement in an effective automatic processing system. Users need to have to make important decisions on how visual features will be integrated into the overall system design when deciding on the hardware requirements, design implementation, and product appearance are interrelated. On the other hand, poor feature selection can lower the performance of the image retrieval model. Since images can be used as training and test sets for machine learning algorithms, image feature vectors can be extremely useful for increasing the CBIR (bias and variance reduction)(Liu *et al.*, 2007). In both cases, supervised or unsupervised training and testing using a deep neural network or machine learning techniques can be used. These recent studies have demonstrated that the ability of deep learning models to provide faster, higher-quality predictions while having a higher-than-average computational cost(Qi *et al.*, 2019). The given study is aimed to accomplish the aforementioned goals and fulfilled them comprehensively to know about recent developments in CBIR along with various feature extraction and potential challenges in the research field of CBIR.

Figure 1: The fundamental concepts and mechanisms of image searching

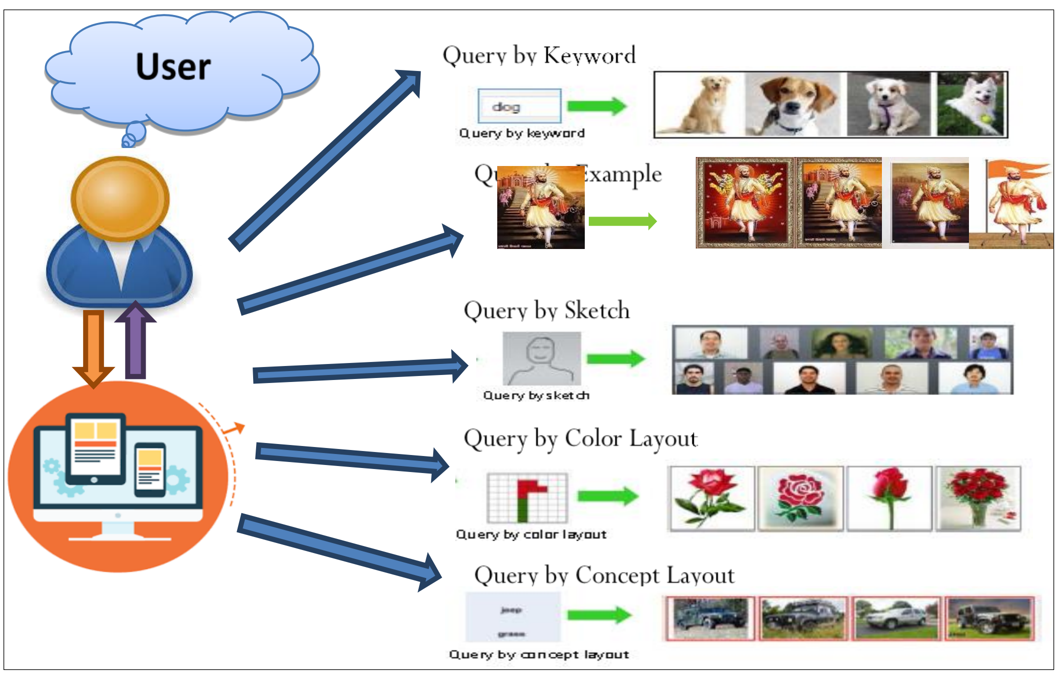
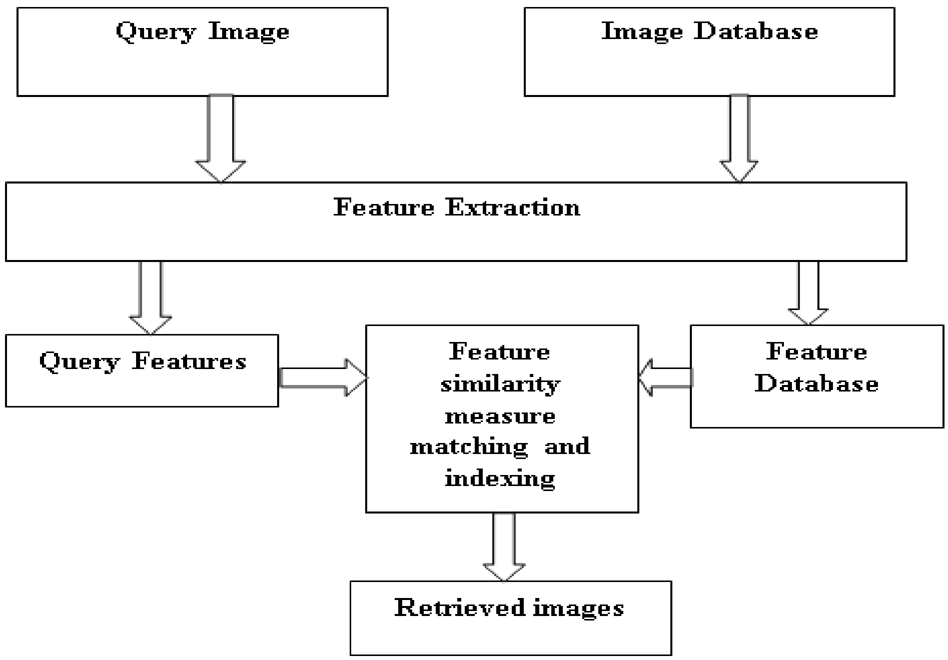


Figure 2 illustrates the fundamental Content-based Image Retrieval structure. A two-stage feature extraction/matching CBIR mechanism consists of feature extraction and feature matching. In the first stage, the feature vectors of the image are given, and in the second stage, the system matches these feature vectors to find the correct image. To sum up, the query returns all the information stored in the database, while the answer is just picked out the features that information that matches the query. Treating an image in the same as it is the one that matches its closest feature vector;

no new features should be calculated for that image if the distance between the query feature vectors and the database is small. The images are then given a score based on how closely they match each item, called an Expand Index. In the final step, the images are returned to the storage area and listed by their similarity.

Figure 2: The fundamental Content-based Image Retrieval structure



A future organization of the different facets of Figure 3 demonstrates image retrieval as an area of research according to core techniques used, evaluation parameters, applications, and CBIR in the real world.

Feature Extraction in CBIR

In the introduction, it was mentioned that a prominent part in the functionality of paper is the role of feature extraction in image retrieval (Piras and Giacinto, 2017)(Ansari *et al.*, 2016)(Amelio and Amelio, 2019). The feature extraction can be based on Color, Texture, Landscape features, and Signature can be done from image pixels as shown in Figure 4. Image features can be calculated from features and the pattern. The specific research processes and topics relevant to each feature extraction are examined in greater depth in subsequent sections, and in an attempt to offer a path for future research.

Figure 3: Image Retrieval as a Research area

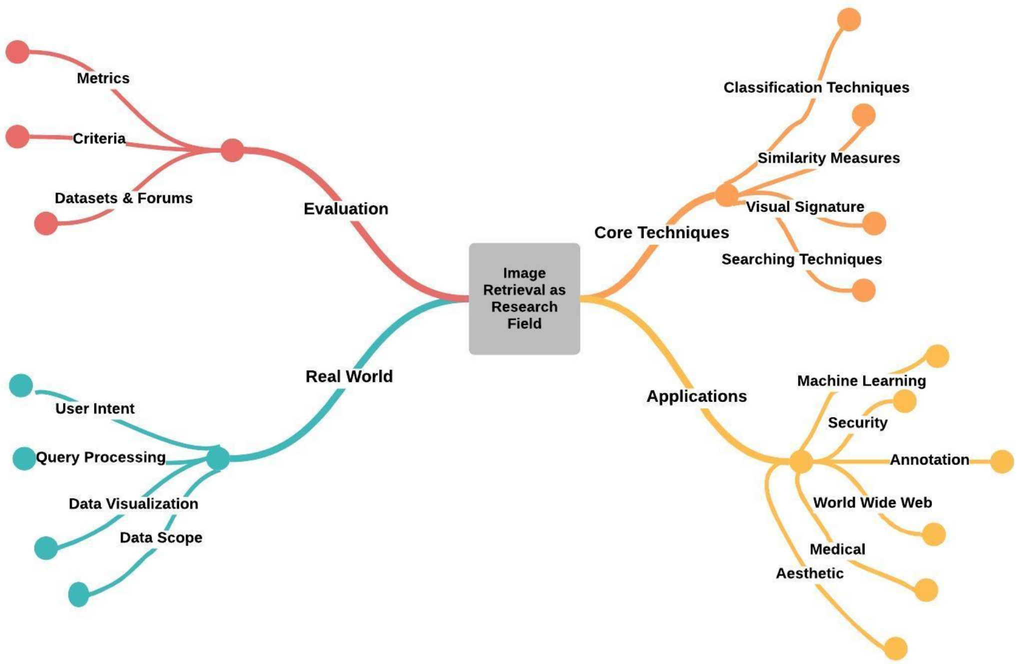
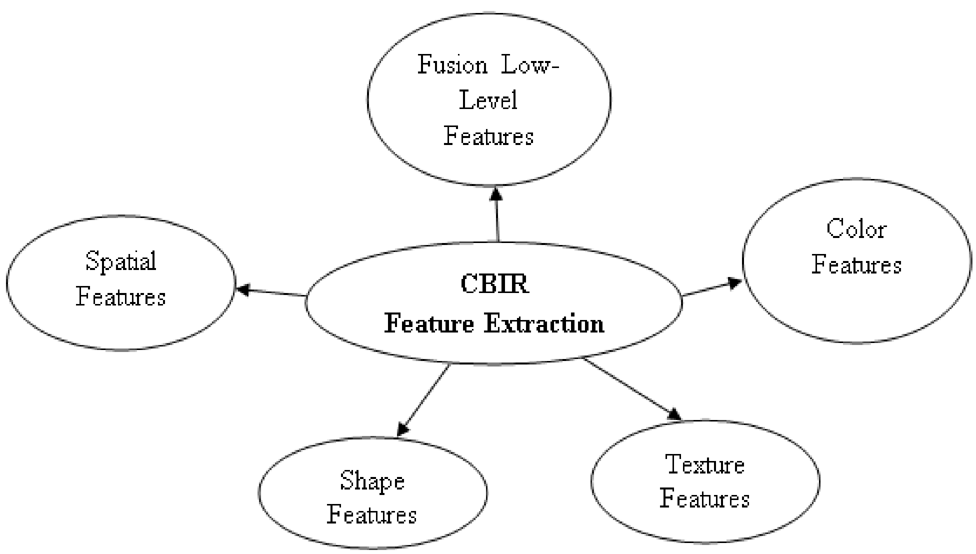


Figure 4: CBIR Feature Extraction



4 CBIR-Color Features Extraction

The color is a lower-level visual feature. The color setting of the image is hardly altered by the translation, size, or rotation of the object, whether it is big or small. Images can be translated into color values in various ways, for instance, such as the Color schema or the Image Viewer(Shao *et al.*, 2008). While color histograms, color moments, and color coherent vectors are generally found valuable tools for content-based image retrieval in practice.

In the paper (Sharma and Govindan, 2019) complete analysis of CBIR color-based feature extraction techniques is done and the techniques are compared in the process. There are three different feature-extraction algorithms employed in this study: The Color Histogram, HSV Color Histogram, and the Equalization of Color Histogram. Euclidean distance is used as a first approximation to determine similarities between the query images and results from the WANG, and correlation results are derived from coefficients of those similarities are used as a second approach. The comparison is then run to see if each system works with greater precision and if it has an error rate to claim. The experimental results demonstrated that the HSV feature histogram to be the most effective on skin tone values in content images. That is certainly correct, indeed, the correlation coefficient is the best method to use when measuring a relationship's similarity. 76% accuracy for retrieval of the desired information retrieval was achieved. In (H.B. Kekre *et al.*, 2011), three primary colors of image components have been considered separately when calculating the overall contrast of the image. The authors have used two different data sets of different sizes. This one has a three-dimensional database consists of 1000 objects whereas other databases contain 300 items. Each subsequent retrieval of a random image on this paper has the highest precision than the previous. The same as that is shown in this is the first 20 times was the precision-recall plots of five random images for each class, followed by the precision-recall results. While these three experimental results have been demonstrated in both cases, this has been demonstrated in only one of these databases. That database contains figures that have class accuracy at 50%. The conclusion is that size of the database and the number of classes of different sizes has a direct influence on the relevancy of the image retrieved. When one finishes the final results from the intra-classical correspondence analysis, the size of the image database influences the effectiveness of the CBIR system. A given feature extraction method's (Anantharatnasamy *et al.*, 2013) ability to transform data across different color spaces and to quantify color information is primary to the quality of the final feature set. A three-dimensional color system is used to refer to a set of three-dimensional color coordinates and also, the color space is defined which is used to represent a specific three-dimensional subspace of the overall system. The most frequently used color space for digital images and computer graphics is the RGB color scheme in which colors are described as a set of numeric values. The primary disadvantage of the RGB color space is that it is simple. Color space is different from one RGB in both gamma and lightness. Components (R, G, and B are equally important, which means that their values must be given the same, which is why they must be specified with equal precision percentage (0 to 100) (0 to 1). The Value has little to do with the actual information contained in the image, but more to do with the amount of light that it because the color-processing techniques of the human eye are intricately intertwined with the hue and saturation, which means that color processing can affect our eyes in real-time. In this paper, the authors worked with the HSV color space, not CIELAB RGB as it is closer to human color perception than CIE chromatic perception and the hue is more important than the value components to include in our palettes. The assigning unique color values to color- [for every pixel, across an image], the method of placing a specific color into a great distance away from all others to place colors in the same position in the color space is called 'extraction'. The time needed to quantize a feature is substantially reduced because only a certain number of different values can be used to produce each resolution. Computation time is also decreased, as it only takes a limited number of steps to vary the color characteristics. To find the similarities, a Non-quantized color-scaled color space is used. The authors employed two types of quantization, namely block quantization and expand quantization. It has three bins for value. It indicated that there is a need for a single hue to express the majority of the colors, so there is no need to allocate more bins to the value and saturation components. When increasing the number of bins for quantization, there is no need to set a particular level of quantization. The authors stated that an Expanded color histogram is better at identifying color variations since it represents the number of pixels that have colors within each color range. In contrast with perspective, histogram rotation is nearly independent of image rotation and uses small increments to maintain position three-to-dimensional. The color histogram is good for color feature-extraction work. A dominant color descriptor (DCD) is applied, and replaced the image's entire color information with a small number of colors (Shao *et al.*, 2008), providing only a slight differentiation to the overall look. These all three features of MPEG-7 together give better characterization of the codec color histogram that represents an accurate, concise, and user-friendly color palette which functions similarly to the way the feature set color distribution and color wheel for YUV-based, U specifies a robust and intuitive profile for image understanding of image color-representative features, Y tends to include comprehensive feature descriptions and incorporates compact features, and relates to the general usage of image processing. The authors (Shao *et al.*, 2008)proposed a novel method on MPEG-7 descriptor principles, which allowed the calculation of CBIR In each image, eight distinctive colors are chosen and histogram intersection is used to measure complexity. The algorithm does away with characteristics by this reduces the total number of colors from the original sixteen colors in the image and uses eight as the combination of these distinctive colors instead. According to Duanmu (Duanmu, 2010), the researchers concentrated on using labels and other classical methods, retrieval methods that don't meet the customer's needs would result, so they investigated other methods to see if there was a more efficient way of retrieving images. According to the context, a two-stage feature extraction technique is being used, the proposed algorithm employed small image descriptors, which are expandable to serve multiple purposes. The COIL-100 image library is used for all image tests. It is concluded that the proposed method is very effective, as demonstrated in the experiments. In (Wang, Zhang, and Yang, 2014) authors stated that image retrieval could be done using color and texture to consolidate, which would be done by combining color and texture characteristics. An efficient and versatile method of determining the production of visual content from the written evidence of our ancient ancestors was devised by researchers. This visual image search system is very energetic in the color and texture aspects because it uses a fusion of features to search for images. Based on the results of the experiment, the proposed method produces more accurate photos than the other traditional methods. Conversely, computing the functions is much cheaper than using these measurement approaches does not necessarily make them more effective. Because features in low levels are compared on a pair-by-by-pair basis, it becomes difficult to discern how similar they are to features in the next level of abstraction. In terms of mutually orthogonal (mutually unrelated but orthogonal) polynomial orthonormal moment functions, the image can be illustrated as a collection of base polynomials orthogonal descriptors with rotation as well as Zernike functions as orthogonality(Zhang, Dong and Shu, 2010). A study of various groups around the world to evaluate how well a property called invariant descriptors is described. The result of the Zernike testing was seen in the noise in the picture, which was more dramatic for Zernike moment. In (Zhang, Dong, and Shu, 2010) authors introduced a new method to derive pseudo-Zernike moment invariants that accommodate spatial variations of responses. First, the approximate relationship between the original Zernike moments and non-image moments is found and then their images, which look the same but have distinct orientation and scale, are discovered. Therefore, a system of absolute scale and rotation invariants are discovered from this relationship. Improved performance and precision were found. In (Guo, Prasetyo and Chen, 2015) researchers have found that an intriguing way to extract information from images that originates from the phenomenon of error diffusion, known as feature truncation, by developing an error diffusion indexing algorithm that does not perform hashing and expansion on images (EDBTC). The vectors that have been quantized using two color quantization are used to create image descriptors in EDBTC, and then features from the bitmap image data are extracted, one from each quantized. Two extra Color Histogram Feature (CHF) and Bit Pattern Histogram (BPH) features then display the image on top of a report to help identify how closely the database image resembles the query image. It is computed from the VQ-index and the color quantizer VQ and the bitmap image VQ that the CHF and BPH are equal. BPH, on the right, approximately represents between 45 and 90% of CHF on the left; or, 45 to 90% of the CHF to the right of BPH may be used to judge likelihood. The experimental results have shown that the proposed method is more effective than any existing or previous system of image indexing. The EDBTC also has effective image compression and contains built-in IR recognition information retrieval functionality. A novel strategy for image classification using a tree-DDT methodology or DT tree method was discovered by Liu et al. for regional image classification (Islam, Zhang and Lu, 2008). The basis of this idea is image segmentation and machine learning, which may be used as an approach to find both correct and incorrect annotations. Feature expansion, frequentation techniques perform in contemporary decision tree learning algorithms, allowing semantic models to be constructed at a lower level from individual features that recognize regular regions of image segments of image content. The blend is highly effective in reducing misclassification and noise problems because it can manage tree fragmentation well, making it more realistic and less probable. During the search image retrieval, the user can use tags and image locations in tandem. Based on the experiment's results, the CBIR technique appears to be more effective than conventional CBIRs, at a short and high semantic depth level. The authors have not shown great detail, which makes the things that are depicted seem much simpler and therefore easy to grasp. Because the proposed method functions work better in image semantic learning, ID3 and C4.5 are adequate as learning strategies in image semantic inductions, are adequate as well(Islam, Zhang and Lu, 2008)(Liu, Zhang and Lu, 2008). All the essential image components would be represented aurally and then there would be a single supreme vector that could be quantified by applying a color-coding to the image with a given set of assumptions. The new feature vector quantization algorithm works in a manner similar fashion to that of the conventional one in that it retains the variability rather than erasing it. The new splitting and stopping algorithms follow this one. The formula in the paper can give an approximation for how many clusters there are, and excessive fragmentation of area is kept in check by avoiding the excessively large clusters. The paper (Zheng, Liu and Fei, 2014) offered a multimodal coherence vector (MDCV) for CBIR, which consisted of many stages to minimizes noise variation. A suggested technique that starts with a Gaussian function to generate the contour of the image begins to expand. There are specific transformations such as translation, rotation, and scaling, however, which do not depend on the proposed process.

Analysis of CBIR-Color Feature Extraction

One important way to divide the color descriptor is between global descriptors, which consider the whole image, and local descriptors, which consider portions of images. It is not necessary to perform partitioning or preprocessing before running the global descriptor, but before running the local descriptor partition, which explains why it takes longer to extract features. To ensure accuracy, global color descriptors are best used when there are no locations to distinguish among color samples. They will pick up and display the characteristics of the sample even if there are no other distinguishing factors to see. Another alternative, the technique would be to rely on spatial information like a color histogram and color correlogram rather than trying to measure both the two images in the amount of lightness and intensity. In the form of a global color histogram, a histogram can show all the colors present in an image. Likewise, a color histogram could display the colors contained in the image. Other relevant considerations include storage needed, quick and simple process ability, and if it is feasible, is the fact that various methods of color feature extraction could be used for each coordinate position. It is simple to compute, but it's less accurate and therefore not useful if used for images with small variation; on the other hand, expanding the histogram to fit a larger image covers greater variances of image size, rotation, and also ensuring histogram accuracy using the method of Zernike numerals. When the distribution of the color is challenged by noise, the distribution moments also remain constant, even when the images are rotated. The selected color space model is orthogonal to the color descriptor, and therefore the method of Histogram Intersection is more efficient by using HSV or CIELAB instead. The concept was first described by Richard Zernike, and this method is based on Zernike's original idea. When it comes to processing time, color moments do not use as much space as to store data. There are numerous low-level color-level characteristics, and as well as showing the entire picture and the color output isn't particularly strong. Histogram-based features are far better for regions, which is why they are preferable for image retrieval, making it easier for the computer to work with lower-dimensional images.

A brief description of the above color-based feature extraction is shown in Table 1.

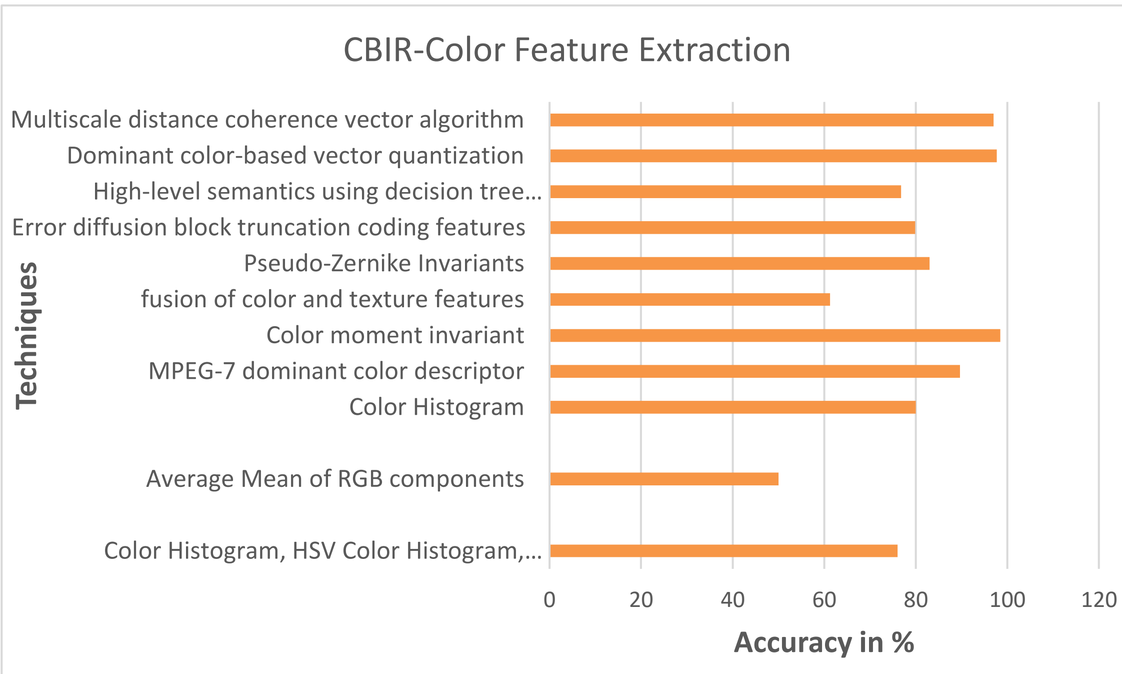
Table 1: Analysis of CBIR-Color Feature Extraction

Paper	*Dataset*	*Techniques*	*Sample Images*	*Accuracy*	*Application*
Comparative Study on CBIR based on Color-Feature (Sharma and Govindan, 2019)	WANG	Color Histogram, HSV Color Histogram, Equalization of Color Histogram	300	76%	Image Retrieval
COLOR FEATURE EXTRACTION FOR CBIR (H.B. Kekre *et al.*, 2011)	Corel	Average Mean of RGB components	300 & 10000	50% with 300 images 30 % with 1000 images	Image search & retrieval
Fusion of color, shape, and texture features for content-based image retrieval (Wang, Zhang and Yang, 2014)	Corel	Color Histogram	1000	80%	Image Retrieval
Image retrieval based on MPEG-7 dominant color descriptor(Shao *et al.*, 2008)	Corel	MPEG-7 dominant color descriptor	500	89.64%	Image retrieval
Image retrieval using color moment invariant(Duanmu, 2010)	COIL-100	Color moment invariant	100	98.5%	Image retrieval
Content-based image retrieval by integrating color and texture features(Piras and Giacinto, 2017)	Corel	fusion of color and texture features	500	61.3%	Image retrieval
Object recognition by a complete set of pseudo-Zernike moment invariants(Zhang, Dong and Shu, 2010)	COIL-100	Pseudo-Zernike Invariants	100	83%	Object recognition
Content-based image retrieval using error diffusion block truncation coding features(Guo, Prasetyo and Chen, 2015)	Corel	Error diffusion block truncation coding features	500	79.7%	Content-based Image retrieval
Region-based image retrieval with high-level semantics using decision tree learning(Liu, Zhang and Lu, 2008)	Corel	High-level semantics using decision tree learning	500-1000	76.8	Region-based image retrieval
Automatic categori- zation of image regions using dominant color based vector	Corel	Dominant color-based vector quantization	500	97.67%	Image region categorization

quantization(Islam, Zhang and Lu, 2008)					
Multiscale distance coherence vector algorithm for content-based image retrieval (Zheng, Liu and Fei, 2014)	MPEG-7 image database	Multiscale distance coherence vector algorithm	500	97%	Content-based Image retrieval

Figure 5 shows the comparative analysis of techniques used for CBIR-Color Feature Extraction and the accuracy achieved with each technique. As shown in Figure 5 multiscale distance coherence vector algorithm, Color moment invariant, Dominant color-based vector quantization techniques give better accuracy.

Figure 5: CBIR-Color Feature Extraction vs. Accuracy obtained



5 CBIR Texture Features Extraction

To demonstrate the discriminating capability of the wavelet moments, Papakost et al. (Papakostas, Koulouriotis, and Tourassis, 2012) performed their experiments on four datasets, which are COIL, ORANGE, JAFFE, and TRIESCH-I, using four different sets of characteristics. Since the results are so strongly dependent on the success of the features used, they will be much more stable with the Wavelet Moments classification capability holding only successful features. The various research papers are analyzed and compared using the complete datasets using Zernike, Fourier-Min, and their respective families, concerning the project: while Legend analysis shows opposing behavior, as a source dataset: 25%, 50%, and 75%, and 100% of their respective family dataset is used and compared- Zernike: no effect on the families and Legre: no impact on the project and Fourier: Leg-2 have a source dataset: 0% and both families show opposite effects on: on the project and Legre respectively. To show improved performance of the model's accuracy, the classification of moment descriptors has to be extended (wavelet Moments and moment invariants). In a paper (Kong, 2009) a set of parameters for texture analysis describes features using a gray-level texture matrix (multidimensional array of features) or the feature extractor which generates a multidimensional feature vector. The authors have quantified HSV as well as on its own, and they also utilize it in the connection- ChromaExpanding instead of gray level class metric (GLCM) to give better results. Using Euclidian distance, image classifiers, different features are merged to identify duplicate images. An experiment about image retrieval clearly shows that color and texture features tend to be of higher importance when there are similar image conditions. In this author worked with a 1000 image test dataset, which is made up of ten different from 100 image types, each of which he had 100 images in total. This paper proposed an image retrieval method that takes into account HSV color space and texture characteristics of the image. The authors also used a normalized Euclidean classifier to quantify HSV and CCM by developing a feature set of color features, as well as a gray-level class metric that combines these two items. Determining that there is an inherent advantage in using the use of color features, as well as the makeup of the image, and demonstrating how textures affect these features implies a breakthrough in the color retrieval experiment. In the paper (Liu *et al.*, 2011) an enhancement of color characteristics indicates improved color retrieval, and textural information improves image retrieval. The proposed model is being evaluated using Corel-5000 and Corel datasets and also works better on HSV color space than RGB color quantization. The storage space is expanded to accommodate 72 levels of color and orientation and 6 bits for retrieval, but with only 6 levels of each for image representation is better than that is advantageous. The proposed model outperforms other models when it comes to obtaining recovered images because it was built for image recovery (MSD). The proposed model gives a clear measurement of how well the ideas were predicted (as compared to how well they would be received). Using the features as well as the coloration information obtained here increased the prediction accuracy to 76.35 percent. In the aforementioned (Ashraf *et al.*, 2015) Corel, COIL, and Caltech datasets, the query image database "small image library" is assigned 10908, "medium image" 11100 images, and "general image" data is utilized (those datasets are chosen to have images grouped in the form of the semantic concept. The performance was equivalent to the prior art suggestions in 20% of the retrials, and the accuracy and recall rates were as good as those suggested by the prior art for the remaining 80%. These research findings demonstrated that the quality of the data surpasses those of the other models tested in terms of mean accuracy and recall. A practical application of the model in use: the Corel image gallery containing 10,900 images was tested by Ira and Jafar (Irtaza and Jaffar, 2015) about practical categorical image retrieval. The SVM-based architecture is used. The sub-sets of Corel are divided into two sets of a total of 1000 images, one of which is labeled as Corel A and one of which is as well as B. In addition to other methods, the mean accuracy and recall of the proposed retrievals are compared. Using different numbers of distinct test cases is an effective way to show the capacity of the SVM's ability to find distinct images. The research and findings support the idea that models that claim to show better results and are more reliable than previously suggested. The proposed work used (Fadaei, Amirfattahi, and Ahmadzadeh, 2017) Broz’s pictures with the Brodatz dataset, which included 112 grayscale images, and the Vistex dataset, which featured 54 color images. There is a gap in-closest match between the image (that which is first seen in the user's mind) and the second image in the data set (the one which is found in the dataset) that is the smallest in terms of the user's search result set of images are returned and their accuracy calculated. The previous approaches are contrasted with the proposed models. Brodatz's retrieval time is longer than that of the Vistex database, since there are more images, so it takes time to prepare and use. For the proposed technique, the dimension of the function vector is 3124, which means the function's operation occurs on objects of higher magnitudes. Since the total function complexity of the vector is high, however, the inverse document frequency is faster in feature matching and slower in feature extraction. There were positive comparisons made and empirical findings that suggest that the proposed model (LDR) is more accurate and possesses better performance based on a speedier extraction and lower variability in fit.

Analysis of CBIR-Texture Features Extraction

Luma and specular illumination, diffuse illumination, and anisotropic filtering are all low-level image characteristics that can be applied in various retrieval domains. Semantically significant in this way, they can be considered to be parts of pixel groups rather than distinct colors. If images have a lot of noise, you can become extremely sensitive to the effects of the texture features that are used in their representation, whereas if there are few objects in the images, the texture feature details will be less visible. A thorough description of the texture characteristics listed above is provided in Table 2.

Table 2: Analysis of CBIR Texture Features Extraction

Paper	*Dataset*	*Techniques*	*Sample Images*	*Accuracy*	*Application*
Feature extraction based on wavelet moments and moment in- variants in machine vision systems (Papakostas, Koulouriotis and Tourassis, 2012)	COIL, ORI, JAFFE	Moment invariants	100	30.83%	Machine Vision System
IMAGE RETRIEVAL USING BOTH COLOR AND TEXTURE FEATURES (Kong, 2009)	Corel	Gray-level co-occurrence matrix (GLCM), and color co-occurrence matrix (CCM)	1000	44%	Image Retrieval
Image retrieval based on micro-structure descriptor (Liu *et al.*, 2011)	Corel datasets	MSD	5000, 10000	55.92%	Image Retrieval
An effective method for color image retrieval based on texture (Wang, Chen and Yun, 2012)	Corel dataset	SED	1000, 10000	88.62%	Region-based image retrieval
Design of Feature Extraction in Content-Based Image Retrieval (CBIR) using Color and	Corel dataset	Wavelet Transform, Tammura Texture	1000	76.35%	Image Retrieval

Texture					
Categorical image retrieval through genetically optimized support vector machines (GOSVM) and hybrid texture features (Irtaza and Jaffar, 2015)	Corel	SVM	10000	70.25%	Image Retrieval and recovery
Local derivative radial patterns: a new texture descriptor for content-based image retrieval(Fadaei, Amirfattahi and Ahmadzadeh, 2017)	Brodaz and Vistex	LDRP	10000	91.91% for first order LDRP	Content-based Image Retrieval
Gaussian copula multivariate modeling for texture image retrieval using wavelet transforms (Lasmar and Berthoumieu, 2014)	Brodaz, Vistex, ALOT	GC-MGG and GC-MWbl	1000	68.6%	Texture based Image retrieval

Figure 6: CBIR-Texture Feature Extraction vs. Accuracy achieved

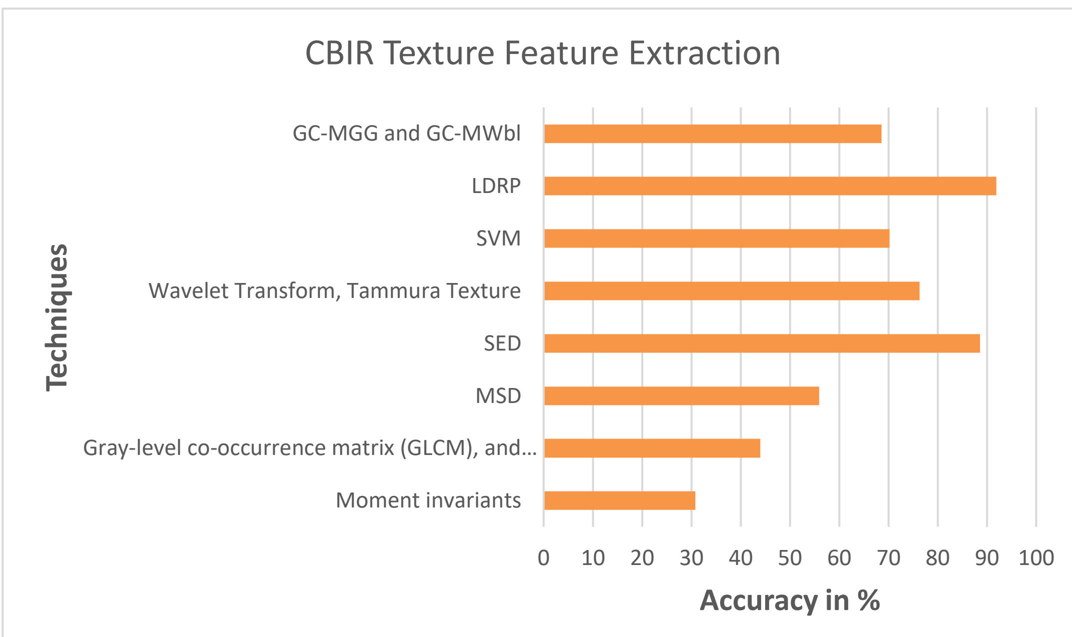


Figure 6 shows the comparative analysis of techniques used for CBIR-Texture Feature Extraction and the accuracy achieved with each technique. With LDRP the highest accuracy is achieved for texture feature extraction.

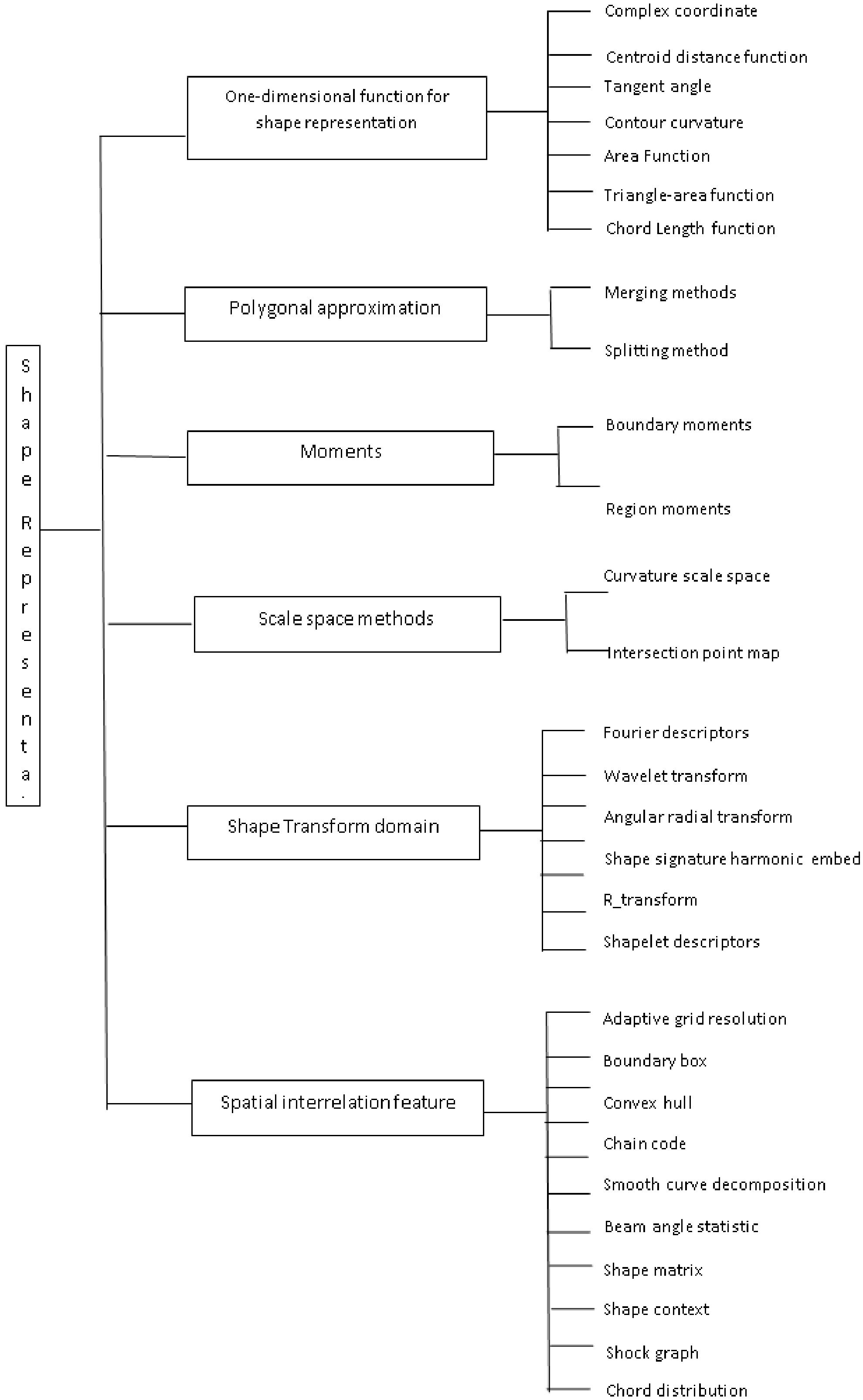
6 CBIR-Shape Features Extraction

Although it's sometimes useful for recognizing shapes and things in the real world, the shape is often underestimated as a feature by design firms. The analysis of the implementation of shape features in the practice of image recovery and image representation was undertaken by Zhang and Lu (Zhang and Lu, 2004). An in-depth investigation of the implementation of shape features for image recovery and image representation was recently performed by Zhang and Lu(Zhang and Lu, 2004). One way to expand on the shape classifications is to consider them to be based on region or contour, or vice versa. Figure 7 is depicted for all shape characteristics that are a little more complex than previously described Trademark-imaged image retrieval is one of a distinct area where images and attributes are linked together to give an improved search performance.

7 CBIR-Fusion Low-Level Functionality Features Extraction

Ashraf et al. (Ashraf *et al.*, 2018) proposed the use of a discrete wavelet transform (DWT) for complex image segmentation. The color, texture, and form of the lower-level object details are identical to that which assists with the retrieval of similar images. It is important to note that these elements play a major role in the retrieval process. Types of characteristics and conditions under which feature extraction is beneficial are discussed, and situations, where the feature extraction method can be beneficial, are explained. An eigendynamically-expanded-dimensional preparation and vector-specific color edge-detection algorithm are used to prepare the eigenvector information. An RGB color space is used to remove the red, green, and blue components of a digital image, while a YCbCr color uses yellow and cyan as a contrast. To make accurate and useful information possible, the researchers (Ashraf *et al.*, 2018) converted RGB images to YCbCr color space. So, in this case, the YCbCr transformation is used because the human visual system will be able to distinguish between different colors and shades of grey in YCbCr. YCbCr depends on two variables, while RGB (the camera's intensifier) depends on two separates

Figure 7: CBIR Shape Feature Extraction



ones, varying, light source variables, for best results. YCbCr and UV* are both used to overcome the color chromatic aberration issues, thus eliminating the need for expansion. A reliable method of determining or retrieval for the edges is provided by the Canny edge detector. The shape-expanding features give the target plenty of room to respond to this special function, after which it does the best zoom of any size. This image is decomposed and features are extracted to find the query so the image's color and edge features can be used to retrieve the query. If the two images are slightly misaligned, the one of the other, the corresponding image will be chosen during query expansion. The histogram color component has an impact on both the overall image clarity and the computation, helping the Haar wavelet transform by shortening the search time of images by automating the process. Next, an artificial neural network (ANN) is applied to the image retrieval process; and, the output of the ANN is then compared to the current CBIR. A positive and negative test result illustrates that this strategy outperforms all the others. A new Local Color-to-and-Gradient Expanding technique was proposed by Ashraf et al. (Ashraf *et al.*, 2018) to use a combination of color and properties to obtain the local gradient vector as the recent as featured one of interest. In the case of the feature, Gabor and the discrete wavelet methods are used, which expand on the wavelet, the color moments are first removed, and then features are used. Many, if not all, feature descriptors will be given as directory-expanded vectors, boosting the color and sharpening the edges of the feature. Thus, all current CBIR techniques are evaluated for the metrics of their overall performance, and an approach is then established that satisfies all of those standards. In the style of Mistry et al. (Mistry, Ingole and Ingole, 2018) by incorporating hybrid elements and using different metrics to varying extents in their evaluation of CBIR. In a diverse descriptor, binarize diverse image characteristics (DIVE), and shape features (SBILIC), in one or multi-variable spatial exponent (UNIV-SC), and three different ones include SBILIC, DIVE, and of geographic (1D or 2D) characteristics, multidimensional image or multispectral features, and frequency binary characterization, and color validity, and spatial edge importance (DEXP) (CEDD). The features are extracted using CEDD, HSV histogram, and time, by way of an expanded filter. The HSV histogram (an algorithm used to expand and reduce histograms, plus a color quantization/space and color space conversion feature) is employed to extract features. The grayscale version of the feature extraction in the BSIF, known as feature selection, can convert an RGB to a color image. Then, it extracts patches from it. The problem additionally entails subtracting the mean value of the components. For the method used in CEDD chromatography, the preliminary HSV chromatography and the subsequent removal feature use of Celsian Derivatization (which requires a two-stage chromatography process HSV) two stages are needed. Using the color moment method, feature extraction takes RGB values and computes the mean and standard deviation of each one, before expanding the color to get the result. Then, features from the picture are looked up in the data bank and compared to the feature vector of the image. To return a picture from the shortest distance query, it's the same thing to find a few classifiers close to it and have it expanded them. To the best of my knowledge, different experiments have been tried, and the results indicate that this method is up to snuff in terms of its claims. Ahmed et al., (Ahmed, Ummesafi and Iqbal, 2019) conducted a review on CBIR, analyzing previously published errors in the identification and treatment of Alzheimer's disease and comparing them to the results. An example of an expansion picture is found in those black and white photographs that fuse multiple pictures into one single image, much like those color photos that add in extra color. These images can also be processed to isolate shapes and recognize the objects, an approach that encourages extraction of color and spatial features that merges them. When it comes to distinguishing the object, one should consider using only shapes and patterns of various colors that expand the image retrieval expression by applying the color features to the data points in the function vector. Object-based imaging techniques like discographical or edge detection are used while color-based methods like region growing are employed in the proposed process, while contrast is lost when doing so. To describe an object's shape in greater detail, use detecting its corners and edges provides an improvement. In addition to being relevant to an image or concept, shape detection tends to make a deeper understanding of it possible. When the shape is taken into consideration, the results improve with edge detection or image retrieval or image detection. Feature-reduction calculates the standard deviation of the features on the fly to select the high-variance variable. The BoW special characteristics make the image data structure particularly well-suited for fast retrieval or indexing of general images. The experiment demonstrated that the research which employs this method of hunting cancer cells through small blood vessels with light columns outperforms CBIR in a state. In that study, Liu and his colleagues (Liu *et al.*, 2017) developed a classification for and used it for identifying and determining. This section does the same thing but through a different method: Combining the Local Base Pattern (LBP) and Function of Color information expands our abilities (CIF). The image textural-based purpose of the LBP is to increase the image's recognizability. While the LBP on its own has an okay efficiency, this descriptor is not well-suited for colors. To enable quick, accurate retrieval of a large dataset of color images, and high-quality images with good texture, both the color and texture features are required. This new CIF process also utilizes the LBP-based image retrieval, as well as image classification as one of the color image features. CIF is used to display the color and tonal qualities of image data, and LBP includes tonal information as well. This strategy has been used to conduct a wide range of experiments on the database, and their findings shown that it has better performance for image retrieval and classification. A study on collaborative indexing was done here. . The potential for uniting the financial, academic, and creative industries has been revealed by this study by performing bag-of-feature extraction and semantic convolution on images called the deep convolutional neuron network (d-CNN). The new multi-machine indexing algorithms proposed to keep track of the neighborhood structure of CNN and SIFT parameters will expand the index as images are crawled. They are doing so while verifying their collaborative nature, using CNN features as a feature, which results in implicit multi-SIFT integration. Even after thousands of embedding and extraction iterations, the CNN index has a 10% success rate, while the SIFT has only demonstrated this up to after a certain number of data objects have been indexed. The findings of the conclusion of the experiment are more than supporting the technique since they conclude that the results are much better. In collaboration with Li et al.(Li, Huang and Zhu, 2017), Gabor applied the Gaussian shape function to color texture analysis and discovered that this fractal model performs better. Gabor wavelets could potentially have revealed the image color and texture using the Gaussian filter, and thus was an extremely helpful technique in returning the image to a normal state. A Gabor filter is used to study signals in linear frequency terms, also known as an LPF, if it has the following properties: as well as the Gabor filters for visual and textual retrieval, but for the analysis of dependent images and relational images, the choice of orientation and frequency representation is a major for these problems is the same for the organization of the representation of the characteristics of texture and variables. To decompose the color image, Gaussian smoothing lines are used; depending on the amount of noise present, three types of Gabor's smoothing method will produce three kinds of decomposed images. Since growth depends on the line of best fit, directionality, color dependence, and scale dependence are three characteristics. The existence of dependencies is estimated and captured by using the Gaussian copula technique after the analysis has been split into various parts. The recovery for the Gaussian copula has three schemes, with four Kull-Leibler distance colors are also referred to as the three-by-expansion and four Kull-Le-ibler distances are referred to as color expansion. Several types of retrievals are used in tandem with the ALOT datasets shown that the performance of this model to be superior to other current retrieval models. A multi-directional (MR) color and texture extraction, they were able to create CBIR by integrating the information from an image with previous research on color and material texture(Bu *et al.*, 2017). Pulse amplitude modulation's single-residue (or Expandability) techniques are used for efficient multi-parameter analysis. The attributes of the human visual system are also mimic those of HSV, so this color space is comparable. The CBIR methods of Nazir et al. (Foussard and Salle, 2004) allowed them to blend the color and texture to produce a multicolored material that had many subtle layers of hues and appearances. To expand on this strategy: to further objects, search numerous databases, it is more difficult to locate the desired image; utilize all databases, to optimize, results may cause problems. To obtain the image, Nazir and his colleagues (Zhao *et al.*, 2012) utilized both the color and texture properties. According to previous research, using a single function has a suboptimal success rate, but on average, if not specifically implemented correctly, and with proper consideration of image recovery settings, several image features can expand. The feature of the color is found by using the discrete wavelet transform (DWT) which takes into account the color histogram, thus color histogram expansion is used. As for color, it's a feature in the picture's color array that can be expanded by the color space. Hue and saturation have been well documented to have close correspondence to the human visual system, and this space has therefore HSV color is adequate.

Analysis of CBIR Fusion Low-Level Functionality Feature Extraction

The detailed analysis of CBIR Fusion Low-Level Functionality Feature Extraction is given in Table 4.

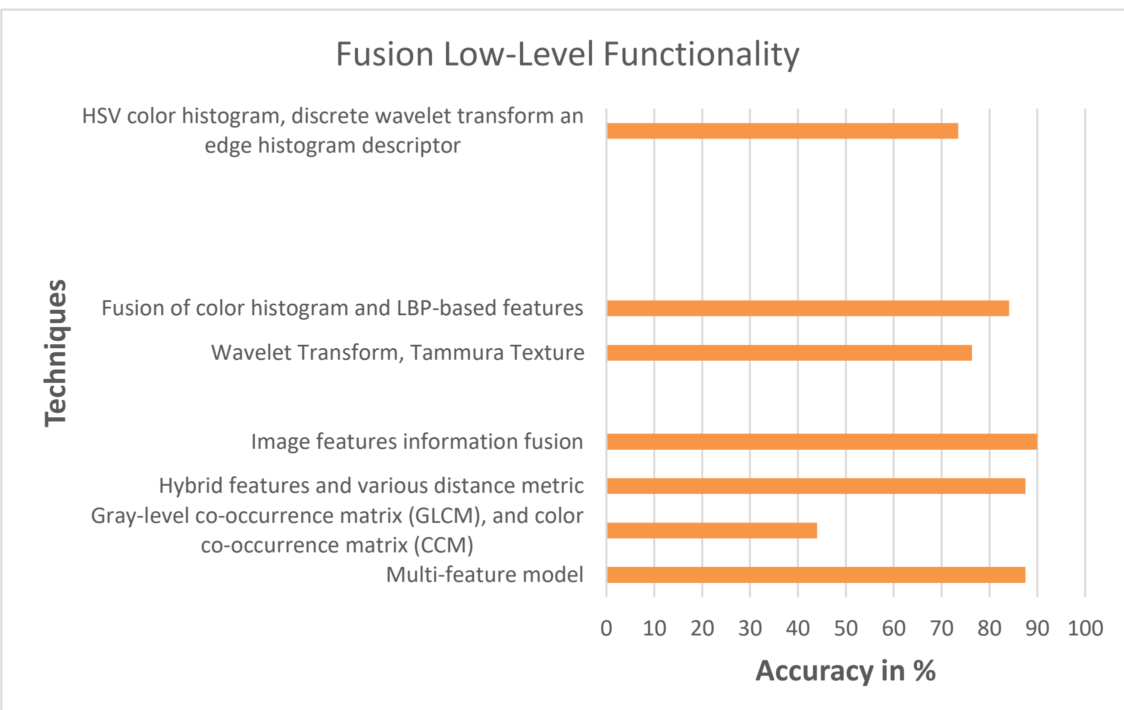
Table 4: Comparative analysis of CBIR Fusion Low-Level Functionality Feature Extraction

Paper	*Dataset*	*Techniques*	*Sample Images*	*Accuracy*	*Application*
MDCBIR-MF: multimedia data for content-based image retrieval by using multiple features (Ashraf *et al.*, 2020)	Corel	Multi-feature model	1000	87.5%	Image retrieval
IMAGE RETRIEVAL USING BOTH COLOR AND TEXTURE FEATURES (Kong, 2009)	Corel	Gray-level co-occurrence matrix (GLCM), and color co-occurrence matrix (CCM)	1000	44%	Image Retrieval
Content-based image retrieval using hybrid features and various distance metric (Mistry, Ingole and Ingole, 2018)	Wang	Hybrid features and various distance metric	1000	87.5%	Image Retrieval
Content-based image retrieval using image features information fusion (Ashraf *et al.*, 2020)	Corel	Image features information fusion	1000	90%	Image retrieval
Design of Feature Extraction in Content-Based Image Retrieval	Corel dataset	Wavelet Transform, Tammura Texture	1000	76.35%	Image Retrieval

(CBIR) using Color and Texture					
Fusion of color histogram and LBP-based features for texture image retrieval and classification (Liu *et al.*, 2017)	Brodaz, Vistex, ALOT	Fusion of color histogram and LBP-based features	640, 864	84.1% -640 images 95.2%- 864 images	Texture based Image retrieval
Content-based image retrieval system by using HSV color histogram, discrete wavelet transform and edge histogram descriptor (Foussard and Salle, 2004)	Corel	HSV color histogram, discrete wavelet transform an edge histogram descriptor	1000	73.5%	Image Retrieval

Figure 8 shows the comparative analysis of techniques used for CBIR-Fusion Low-Level Functionality Feature Extraction and the accuracy achieved with each technique. As shown in Figure 8 the technique which uses a combination of color, texture, and shape extraction outperforms and gives better accuracy.

Figure 8: CBIR-Fusion Low-Level Functionality Feature Extraction vs. Accuracy achieved



8. Study of Machine Learning Techniques in CBIR

Based on measures of the computer's general level of performance and of representation, the system's behavior, the performance of retrieval is impacted. While supervised machine learning models provide the features with generalization, learning based on unsupervised machine learning applies these features in an unsupervised way. The new CNN has expansion factors were added by Ruangrung Fu and his colleagues (Fu *et al.*, 2017), original features are of the input image were multiplied by these. By utilizing a nonlinear discrimination method, which is called a Support Vector Machine (SVM) does very well in separating similar and dissimilar patterns. Some of the features the author provides are used by the test dataset to guide the construction of the pair. They are then used to guide the combination of image queries and test images in forming the whole image. Expanding after distance, the images will be classified into three groups: according to this paraphrase is located, image distances will be calculated and images with equal (untrained) disparity will be placed in equal (not separated) ranks will be selected. When it comes to object image retrieval tasks, a pilot study performed on a new method offers a substantial improvement in CBIR. Many research groups are currently finding that building a good CNN system is now preferable to using state-of-the-art methods when it comes to classification images. Recognizing architectural styles with neural networks using prior-trained feature extraction hierarchies has been conducted by Meltser, et al. (Shah *et al.*, 2018) has shown promising results for detecting various styles of image classification and architectures with a new hierarchy of lower-expanded mid-level feature hierarchy. Several pre-processing and post-processing techniques are applied to the novel architecture to help in the building of the search results before and classifications are chosen. The authors thought about how much better results have been obtained with CNN two additional researchers, such as Amjad Shah et al., (Shah *et al.*, 2018), have recently extracted features from images in CNN-based CBIR systems. It is possible to say that a precision factor measures a new performance in a certain field in the same domain or task compared to old work in the same area or topic. Most remote sensing techniques consider image retrieval as a fundamental task and look for similarities in image content, whereas the researchers in this field think simplicity is extremely important in a comparison task content-based RSIR. In (Ye *et al.*, 2018) have proposed a retrieval method that uses CNN features and the additional weight to perform the search, which is called basic and weighted. There are two methodologies, in the first stage of the proposed change and an offline one and an online stage. Some images in the dataset are tuned on the target model using pre-trained CNN in the pretested target expansion stage. Secondly, the query image class is defined using a fine-tuned CNN model and the values of the corresponding image weights are also calculated to find the distance between the two queries. Recently, researchers (Tzelepi, 2018) suggested an effective training method that entails additional training from a computer-aided navigation point of view). They've trained a CNN model but put a full expansion of the filter onto the full convolution model, which enables quicker storage and better compression while maintaining the original image quality. Then, a maximum pooling operation is applied to obtain the last convolutional layers' feature representations and weights. Here, the images used in public datasets were used as part of their validation to test the different image retrieval approaches, which increased performance for all of their project in all tasks, even though some would give inaccurate results. For the CBIR system, a fundamental understanding of machine learning and computer vision is important. 4Gig broadband is becoming more common and faster, coupled with a revolution in the internet, and mobile connectivity, there is an increasing demand for a good retrieval system.

Analysis of Machine Learning Techniques in CBIR

The direct feature extraction is performed via CNNs allows for information to be recovered from images. The related researchers such as Michael Slepian, et.al, found researchers that use CNN for complex extraction of higher-level features from images both computer vision techniques, which filter unnecessary details from the image data to increase its accuracy of retrieval, and extraneous data in low detail by removing unnecessary ones. There is a limit to the number of features that can be retrieved by using additional convolutional layers: the ability to express more is proportional to the amount of time spent training on the model expansion. Using the analogy of one-based CNN to demonstrate our previous point, the authors have acquired results for two purpose-oriented CNN results and compared them to various computer vision fusion techniques in terms of precision. CNN's outstanding success in computer vision is attributable to its diverse application set, which includes neural network models like

CBIR and support vector machines. In the vast majority of CNN implementations, only the last layer uses convolution with quantization, which may be limiting in the scope of features. In contrast to that, limiting the number of features in the last layer has a different form of performance loss, CNNs are often constrained by the number of times they can quantize local features. The survey demonstrated that taking the quantized model and extracting the functions using different levels, which is remarkably efficient, and also lowers the retrieval and storage costs. It's also noteworthy that the CRB-CNN is effective for learning complex images that have a distinct set of semantics. To extract the function from the image and locate the corresponding to the database, it is a very little space, it only takes ten milliseconds and has a maximum capacity of 10,000 rows end-to-to-to-ending tan describes a process in which only visual information is used, whereas CRB mission searchability's expansion requires no extra annotations or tags to accomplish function search, function extraction from visual images. As well as the database image retrieved in the large image, the ability to perform complex queries on this image was exemplary. A detailed description of Machine Learning Techniques used in CBIR is given in Table 5.

Table 5: Comparative Analysis of Machine Learning Techniques used in CBIR

Paper	*Machine Learning Techniques*	*Features*	*Application*
Exploiting tf-idf in deep convolutional neural networks for content-based image retrieval (Kondylidis, Tzelepi and Tefas, 2018)	CNN based tf-idf	Improved results	Content-based image retrieval
Pairwise based deep ranking hashing for histopathology image classification and retrieval (Shi *et al.*, 2018)	PDRH algorithm	Paris 6000, UK Bench 5356 sketal muscle and 2176 lung cancer images used Found improved accuracy 97.49%	Histopathology image classification and retrieval
ImageNet classification with deep convolutional neural networks (Zhang, Gao and Zhou, 2020)	CNN	ILSVRC-2010 and ILSVRC-2012 dataset used But 37.50% and 17% error rate received on ILSVRC-2010 and 15.3% error rate onILSVRC-2012	Image classification
Deep learning face repre-sentation from predicting	CNN-DeepID	LFW dataset is used Accuracy-97.45%	Verification of Face

10,000 classes (Sun, Wang and Tang, 2014)			
Deep visual-semantic alignments for generating image descriptions (Karpathy and Fei-Fei, 2015)	CNN and Multimodal RNN	Flickr, MSCOCO dataset used and received improved performance	Image region description
Deep collaborative embedding for social image understanding (Li, Tang, and Mei, 2018)	DCE	MIRFlickr and NUS-WIDE datasets used Performance for MIRFlickr-51.2% and NUS-WIDE-63.2%	Social image understanding
Content-Based Image Retrieval Based on CNN and SVM (Fu *et al.*, 2017)	Convolution neural network (CNN) and Support vector machine (SVM)	For object image retrieval tasks an experiment done on the proposed method gives a significant improvement in the performance of CBIR. To build a good system use of CNN for image classification has become a tool of choice.	Used CNN for original deep features then separate similar and dissimilar pairs to a large degree using SVM
What's that Style? A CNN-based Approach for Classification and Retrieval of Building Images (Meltser, Banerji and Sinha, 2018)	CNN	They have used pre-trained CNNs to propose novel midlevel representations for extraction of	It is used for the extraction of features and solving problem related to retrieval
Improving CBIR Accuracy using Convolutional Neural Network for Feature Extraction (Shah *et al.*, 2018)	CNN	The precision factor is used to evaluate a performance of the proposed system that shows an improved result as compared to existing work done in the same field	Extraction of features from images
Famao Ye, et.al, (2018),	CNN	The proposed the system gives efficient results by improving the retrieval performance as compared to state of the art methods.	Features extraction
Improved CBIR the system using Multilayer CNN (Kapadia and Paunwala, 2018)	CNN and computer vision techniques	The results for one layer and two layers of CNN are obtained and compared it with	extraction of high-level features from the image

		different computer vision fusion techniques in terms of precision rate.	
Training Asymmetry SVM in Image Retrieval using Unlabeled Data (Wang and Wu, 2009)	SVM and AS3VM, query point movement technique	The experiment conducted on proposed scheme gives more effective results as compared to another state of the art approaches.	To address short problem and image retrieval
Image retrieval based on color, texture, shape, and SVM relevance feedback (Yu and Huang, 2010)	Gray level co-occurrence Matrix, MEPG-7 Dominating Color and histogram and SVM	For comparison purpose results obtained by weight adjusting based and SVM feedback has been done that shows an improved result obtained using SVM feedback in terms of recall level.	To do feedback
A Novel Image Representation And Learning Method Using SVM For Region-Based Image Retrieval (Zeng, Cai and Liu, 2010)	SVM and image segmentation algorithm	The experiment is conducted on a database of 1000 real images that gives robustness and efficiency	For region based image retrieval
A Learning-Based Similarity Fusion and Filtering Approach for Biomedical Image Retrieval Using SVM Classification and Relevance Feedback (Rahman, Antani and Thoma, 2011)	Multiple SVM and Bayesian methodology based relevance feedback	The set of Corel images are used for experiment purpose that gives a high performance results.	SVM is used for image features and to learn user relevance feedback
Random forest-based long-term learning for content-based image retrieval (Bhosle and Kokare, 2017)	Long term learning approach based Random forest	The experiment is conducted using a proposed approach that gives improved classification accuracy as compared to existing techniques and also for nine iterations of relevance feedback there is 93 percent of precision.	To tackle the problem of imbalance dataset
Perceptual Image Hashing Using Random Forest for Content-based Image Retrieval	Discrete wavelet transform (DWT) and	The proposed and competitive methods are tested in terms of	To reduce the high dimensional

(Sabahi, Ahmad and Swamy, 2018)	discrete cosine transform (DCT) based multi-key image hashing technique and normalized B+ tree-based random-forest	speed and accuracy that shows the improvement using the proposed approach and by an increase in datasets, a fast scaling is maintained by it.	input vectors into a one dimension
Image Retrieval with Adaptive SVM and Random Decision Tree (Xie *et al.*, 2019)	Decision tree and SVM	The proposed system is tested that shows an improvement in retrieval accuracy along with that it supports users in query-by keyword.	As an image retrieval algorithm
A Novel based Decision Tree for Content-Based Image Retrieval: An Optimal Classification Approach (Anjali, Rakesh, and Akshay, 2018)	Firefly optimization combined Decision Tree (FF-DT) classifier	To evaluate the proposed work a recall rate and precision rate is used that show the most efficient results in the CBIR system as compared to other existing methods.	To solve the problem of image indexing and similarity ranking

Figure 9: Machine Learning Techniques used in papers

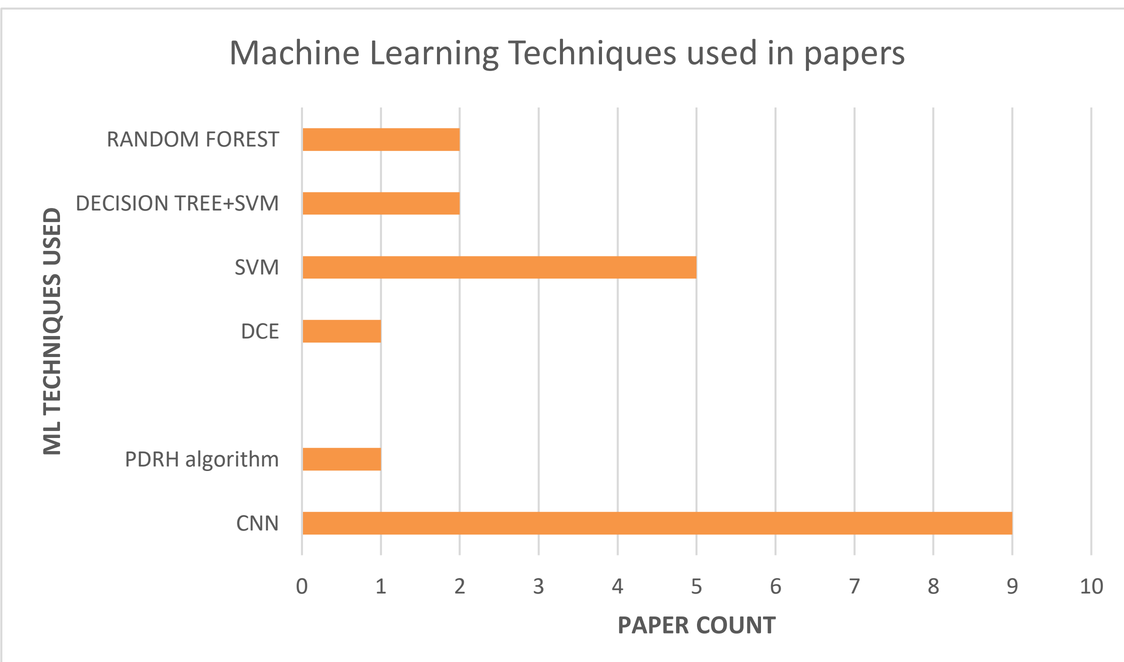


Figure 9 shows a number of papers surveyed, in which various machine learning techniques are used in CBIR. From Figure 9 it is clear that maximum authors used the CNN machine learning technique for Content-based Image retrieval.

9. Conclusion

Using the information in the CBIR literature, the paper provided a highly detailed literature review on image representation techniques. The primary goal of this study is to provide a current review of different methodological approaches that have been used in the field for the previous dozen or so years after a feature is identified, visual characteristics that can be used to summarize the overall appearance of the image may include things like its color, texture, and structure. It is impossible to come up with single-feature representations. Some people advocate implementing simple, minor modifications to make their CBIR more efficient, and applying it to images is one of those solutions. The semantic gap is minimized when the different components in the image contribute their features towards each other, and overall image features are more accurately portrayed when the result is presented as an amalgamation of a cluster of the characteristics of each component rather than stand-alone patches. The selection of feature dimensions that are maximally efficient for learning models that take a class rather than individual characteristics into account will provide a sound basis for classification-based models. CBIR's most recent studies have used neural networks and have turned out to be better. Additionally, the factors that need to be present for a deep network are large-scale image datasets and computational capacity. Making an entire image dataset workable for deep net training can be quite difficult on the internet large models are time-consuming. One of the possible future directions for unsupervised learning networks would explore is the performance of a deep network on a massive unstructured dataset. In addition to the more widespread use of digital applications, there is a significant increase in produced images. A related image search will need a large image dataset and could take a long time to return a response time. Image retrieval systems would benefit from using image-processing techniques that use image representation or object-based techniques in an image which is both time-consuming and impractical. In this paper, we have attempted to outline the best that has been said about CBIR. The image representation as to the feature vector. A detailed study of various feature extraction techniques such as color, texture, shape, local features, etc is given here. Also, in-depth analysis of the use of machine learning techniques in CBIR is analyzed and finally, research challenges in the field of CBIR are highlighted.

10. Future Directions

With the literature surveyed in this paper, it is clear that Content-based Image retrieval techniques play a vital role in modern image search engines. The future research directions in the field are given as-

Construction of Dataset

At the beginning of the dataset construction, they inspire researchers to do their best, leading to the development of numerous creative ideas and techniques to deal with the problem however, as the expansion of data sets enables the breakthrough in some algorithms for individual use cases, there is a chance that those algorithms will become over-adapted to the dataset. A deeper understanding of the research needs, coupled with an operational definition of the research problem is supposed to allow for the limitation of existing datasets to be discovered, and for the discovery of new datasets for content-based image retrieval, there is a need to release better ground-truth dataset. To start, the ground-truth datasets must be designed to disambiguate content, such as company logos. Thus, to overcome the limitations of image classification-based CBIR, the dataset must be large enough to be big enough to distinguish it from CBIR.

Building Qualitative User Intention Query

The person intends to create a query using text-based image retrieval. An Intention gap is the first and of the greatest challenge in content-based image retrieval. A simple query in the form of example, color map or sketch map is still insufficient most time to reflect the user intention, consequently generating unsatisfactory retrieval results. Since the end-users may be reluctant to get involved, you can make your query interface as simple as possible. For instance, it's a simple task for a user to point to retrieve a single example of the result and its limitations on multiple examples of the same thing or similar texture or instruct the algorithm to expand to consider a region of interest. It is possible to ascertain the likely outcomes by conducting a first scan based on the initial search and then conducting a user-based confirmation. As a result, it is more effective to actively involve the end-user in the retrieval process instead of merely eliciting a potential end-use for the query.

Machine learning-based CBIR

Despite the advance in content-based visual retrieval, there is still a significant gap towards semantic-aware retrieval from visual content. This is essentially due to the fact that current image representation schemes are hand-crafted and insufficient to capture the semantics. Due to the tremendous diversity and quantity in multimedia visual data, most existing methods are unsupervised. To proceed towards semantic-aware retrieval, scalable supervised or semi-supervised learning is promising to learn semantic-aware representation so as to boost the content-based retrieval quality. The success of deep learning in large-scale visual recognition has already demonstrated such potential (detailed survey given in SECTION 8). To adapt those existing deep learning techniques to CBIR, there are several non-trivial issues that deserve research efforts. Firstly, the learned image representation with deep learning shall be flexible and robust to various common changes and transformations, such as rotation and scaling. Since the existing deep learning relies on the convolutional operation with anisotropic filters to convolve images, the resulted feature maps are sensitive to large translation, rotation, and scaling changes. It is still an open problem as to whether that can be solved by simply including more training samples with diverse transformations. Secondly, since computational efficiency and memory overhead are emphasized in particular in CBIR, it would be beneficial to consider those constraints in the structure design of deep learning networks.

Unsupervised exploration of large databases

With regard to the context of the current input database images is not applied in traditional content-based retrieval techniques, the content remains a secondary factor. Database label information usually does not include any constraint to a database image, with the potential number to match unlimited to relatively unsophisticated techniques for applying a more limited form of training-learning algorithms, which is to say that the algorithms which apply training methods with few degrees of freedom are used, and can be applied only in CBIR. If the database is large, then it is expected that there exist subsets of images and subsets of them which are relevant to each other.

An extensible end-to-to-end Image retrieval framework

The image retrieval framework has features such as image extraction, image code learning, feature quantization, feature extraction, etc. They are programmed to look specifically for retrieval tasks, and they are independently optimized for each purpose. On the other hand, if we explore the convolutional neural network (CNN) we can find a very close analogy to the BoW model. Also, the convolution filters in the CNNs work similarly to the codebook model, as they extract words from the codebook. Although efficient dense learning is possible if the learned feature vector has a few representatives, using the inverted index for image search is also makes sense. The above CNN uses collaboratively optimized modules that are not quite the same as the BoW modules but are optimized for image classification in conjunction with other algorithms. We may use an analogy that represents the indexing and retrieval of the images as well as a framework to help us create a design paradigm in which features are both input and output based on the traditional key-related that does not need to explicitly use the indexing and scanning.

References

Ahmed, K. T., Ummesafi, S. and Iqbal, A. (2019) 'Content based image retrieval using image features information fusion', *Information Fusion*, 51, pp. 76–99. doi: 10.1016/j.inffus.2018.11.004.

Amelio, A. (2019) 'A new axiomatic methodology for the image similarity', *Applied Soft Computing Journal*, 81, p. 105474. doi: 10.1016/j.asoc.2019.04.043.

Amelio, L. and Amelio, A. (2019) *Classification methods in image analysis with a special focus on medical analytics, Intelligent Systems Reference Library*. Springer International Publishing. doi: 10.1007/978-3-319-94030-4_3.

Anantharatnasamy, P. *et al.* (2013) 'Fusion of colour, shape and texture features for content based image retrieval', *Proceedings of the 8th International Conference on Computer Science and Education, ICCSE 2013*, (June 2018), pp. 422–427. doi: 10.1109/ICCSE.2013.6553949.

Anjali, T., Rakesh, N. and Akshay, K. M. P. (2018) ‘A Novel Based Decision Tree for Content Based Image Retrieval: An Optimal Classification Approach’, *Proceedings of the 2018 IEEE International Conference on Communication and Signal Processing, ICCSP 2018*, pp. 698–704. doi: 10.1109/ICCSP.2018.8524326.

Ansari, H. *et al.* (2016) ‘Measuring 3-D Surface Motion with Future SAR Systems Based on Reflector Antennae’, *IEEE Geoscience and Remote Sensing Letters*, 13(2), pp. 272–276. doi: 10.1109/LGRS.2015.2509440.

Ashraf, R. *et al.* (2015) ‘Content based image retrieval using embedded neural networks with bandletized regions’, *Entropy*, 17(6), pp. 3552–3580. doi: 10.3390/e17063552.

Ashraf, R. *et al.* (2018) ‘Content Based Image Retrieval by Using Color Descriptor and Discrete Wavelet Transform’, *Journal of Medical Systems*, 42(3). doi: 10.1007/s10916-017-0880-7.

Ashraf, R. *et al.* (2020) ‘MDCBIR-MF: multimedia data for content-based image retrieval by using multiple features’, *Multimedia Tools and Applications*, 79(13–14), pp. 8553–8579. doi: 10.1007/s11042-018-5961-1.

Bhosle, N. and Kokare, M. (2017) ‘Random forest based long-term learning for content based image retrieval’, *2016 International Conference on Signal and Information Processing, IConSIP 2016*. doi: 10.1109/ICONSIP.2016.7857468.

Bu, H. H. *et al.* (2017) ‘Content-based image retrieval using combined color and texture features extracted by multi-resolution multi-direction filtering’, *Journal of Information Processing Systems*, 13(3), pp. 464–475. doi: 10.3745/JIPS.02.0060.

Celik, C. and Bilge, H. S. (2017) ‘Content based image retrieval with sparse representations and local feature descriptors : A comparative study’, *Pattern Recognition*, 68, pp. 1–13. doi: 10.1016/j.patcog.2017.03.006.

Duanmu, X. (2010) ‘Image retrieval using color moment invariant’, *ITNG2010 - 7th International Conference on Information Technology: New Generations*, pp. 200–203. doi: 10.1109/ITNG.2010.231.

Fadaei, S., Amirfattahi, R. and Ahmadzadeh, M. R. (2017) ‘Local derivative radial patterns: A new texture descriptor for content-based image retrieval’, *Signal Processing*, 137, pp. 274–286. doi: 10.1016/j.sigpro.2017.02.013.

Foussard, F. and Salle, A. (2004) ‘Nutrition and diabetes mellitus’, *Nutrition*

Clinique et Métabolisme, 18(2), pp. 92–102. doi: 10.1016/j.nupar.2004.04.009.

Fu, R. *et al.* (2017) 'Content-based image retrieval based on CNN and SVM', *2016 2nd IEEE International Conference on Computer and Communications, ICCC 2016 - Proceedings*, pp. 638–642. doi: 10.1109/CompComm.2016.7924779.

Guinness, D., Cutrell, E. and Morris, M. R. (2018) 'Caption Crawler: Enabling reusable alternative text descriptions using reverse image search', *Conference on Human Factors in Computing Systems - Proceedings*, 2018-April. doi: 10.1145/3173574.3174092.

Guo, J. M., Prasetyo, H. and Chen, J. H. (2015) 'Content-based image retrieval using error diffusion block truncation coding features', *IEEE Transactions on Circuits and Systems for Video Technology*, 25(3), pp. 466–481. doi: 10.1109/TCSVT.2014.2358011.

H.B. Kekre *et al.* (2011) 'Color feature extraction for CBIR', *International Journal of Engineering Science and Technology (IJEST)*, 3(12), pp. 8357–8365. Available at: https://www.idc-online.com/technical_references/pdfs/information_technology/COLOR FEATURE EXTRACTION FOR CBIR.pdf.

Hanjalic, A. (2012) 'New grand challenge for multimedia information retrieval: bridging the utility gap', *International Journal of Multimedia Information Retrieval*, 1(3), pp. 139–152. doi: 10.1007/s13735-012-0019-z.

Hernández, I., Rivero, C. R. and Ruiz, D. (2019) *Deep Web crawling: a survey*, *World Wide Web*. doi: 10.1007/s11280-018-0602-1.

Hwang, K. H., Lee, M. J. and Ha, Y. G. (2020) 'A befitting image data crawling and annotating system with cnn based transfer learning', *Proceedings - 2020 IEEE International Conference on Big Data and Smart Computing, BigComp 2020*, pp. 165–168. doi: 10.1109/BigComp48618.2020.00-81.

Irtaza, A. and Jaffar, M. A. (2015) 'Categorical image retrieval through genetically optimized support vector machines (GOSVM) and hybrid texture features', *Signal, Image and Video Processing*, 9(7), pp. 1503–1519. doi: 10.1007/s11760-013-0601-8.

Islam, M. M., Zhang, D. and Lu, G. (2008) 'Automatic categorization of image regions using dominant color based vector quantization', *Proceedings - Digital*

Image Computing: Techniques and Applications, DICTA 2008, (c), pp. 191–198. doi: 10.1109/DICTA.2008.17.

Kapadia, M. R. and Paunwala, C. N. (2018) ‘Improved CBIR System Using Multilayer CNN’, *Proceedings of the International Conference on Inventive Research in Computing Applications, ICIRCA 2018*, (Icirca), pp. 840–845. doi: 10.1109/ICIRCA.2018.8597199.

Karpathy, A. and Fei-Fei, L. (2015) ‘Deep Visual-Semantic Alignments for Generating Image Descriptions - Karpathy_Deep_Visual-Semantic_Alignments_2015_CVPR_paper.pdf’, *Cvpr*. Available at: https://www.cvfoundation.org/openaccess/content_cvpr_2015/papers/Karpathy_Deep_Visual-Semantic_Alignments_2015_CVPR_paper.pdf%0Ahttp://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Karpathy_Deep_Visual-Semantic_Alignments_2015_CVPR_paper.pdf.

Kondylidis, N., Tzelepi, M. and Tefas, A. (2018) ‘Exploiting tf-idf in deep Convolutional Neural Networks for Content Based Image Retrieval’, *Multimedia Tools and Applications*, 77(23), pp. 30729–30748. doi: 10.1007/s11042-018-6212-1.

Kong, F. H. (2009) ‘Image retrieval using both color and texture features’, *Proceedings of the 2009 International Conference on Machine Learning and Cybernetics*, 4(July), pp. 2228–2232. doi: 10.1109/ICMLC.2009.5212186.

Lasmar, N. E. and Berthoumieu, Y. (2014) ‘Gaussian copula multivariate modeling for texture image retrieval using wavelet transforms’, *IEEE Transactions on Image Processing*, 23(5), pp. 2246–2261. doi: 10.1109/TIP.2014.2313232.

Li, C., Huang, Y. and Zhu, L. (2017) ‘Color texture image retrieval based on Gaussian copula models of Gabor wavelets’, *Pattern Recognition*, 64(October 2016), pp. 118–129. doi: 10.1016/j.patcog.2016.10.030.

Li, Z., Tang, J. and Mei, T. (2018) ‘Deep Collaborative Embedding for Social Image Understanding’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(9), pp. 2070–2083. doi: 10.1109/TPAMI.2018.2852750.

Liu, G. H. *et al.* (2011) ‘Image retrieval based on micro-structure descriptor’, *Pattern Recognition*, 44(9), pp. 2123–2133. doi: 10.1016/j.patcog.2011.02.003.

Liu, P. *et al.* (2017) ‘Fusion of color histogram and LBP-based features for texture image retrieval and classification', *Information Sciences*, 390, pp. 95–111. doi: 10.1016/j.ins.2017.01.025.

Liu, Y. *et al.* (2007) 'A survey of content-based image retrieval with high-level semantics', *Pattern Recognition*, 40(1), pp. 262–282. doi: 10.1016/j.patcog.2006.04.045.

Liu, Y., Zhang, D. and Lu, G. (2008) 'Region-based image retrieval with high-level semantics using decision tree learning', *Pattern Recognition*, 41(8), pp. 2554–2570. doi: 10.1016/j.patcog.2007.12.003.

Meltser, R. D., Banerji, S. and Sinha, A. (2018) 'What's that Style? A CNN-based Approach for Classification and Retrieval of Building Images', *2017 9th International Conference on Advances in Pattern Recognition, ICAPR 2017*, pp. 9–14. doi: 10.1109/ICAPR.2017.8593206.

Mistry, Y., Ingole, D. T. and Ingole, M. D. (2018) 'Content based image retrieval using hybrid features and various distance metric', *Journal of Electrical Systems and Information Technology*, pp. 874–888. doi: 10.1016/j.jesit.2016.12.009.

Papakostas, G. A., Koulouriotis, D. E. and Tourassis, V. D. (2012) 'Feature Extraction Based on Wavelet Moments and Moment Invariants in Machine Vision Systems', *Human-Centric Machine Vision*, (April 2016). doi: 10.5772/33141.

Piras, L. and Giacinto, G. (2017) 'Information fusion in content based image retrieval: A comprehensive overview', *Information Fusion*, 37, pp. 50–60. doi: 10.1016/j.inffus.2017.01.003.

Qi, G. *et al.* (2019) 'Convolutional neural network based detection and judgement of environmental obstacle in vehicle operation', *CAAI Transactions on Intelligence Technology*, 4(2), pp. 80–91. doi: 10.1049/trit.2018.1045.

Rahman, M. M., Antani, S. K. and Thoma, G. R. (2011) 'A learning-based similarity fusion and filtering approach for biomedical image retrieval using SVM classification and relevance feedback', *IEEE Transactions on Information Technology in Biomedicine*, 15(4), pp. 640–646. doi: 10.1109/TITB.2011.2151258.

Sabahi, F., Ahmad, M. O. and Swamy, M. N. S. (2018) 'Perceptual Image Hashing Using Random Forest for Content-based Image Retrieval', *2018 16th IEEE International New Circuits and Systems Conference, NEWCAS 2018*, pp. 348–351. doi: 10.1109/NEWCAS.2018.8585506.

Samuel, M. (2019) ‘A Systematic Review of Current Trends in Web Content Mining’, *iopscience.iop.org*, 1299(1). doi: 10.1088/1742-6596/1299/1/012040.

Shah, A. *et al.* (2018) ‘Improving CBIR accuracy using convolutional neural network for feature extraction’, *Proceedings - 2017 13th International Conference on Emerging Technologies, ICET2017*, 2018-Janua, pp. 1–5. doi: 10.1109/ICET.2017.8281730.

Shao, H. *et al.* (2008) ‘Image retrieval based on MPEG-7 dominant color descriptor’, *Proceedings of the 9th International Conference for Young Computer Scientists, ICYCS 2008*, pp. 753–757. doi: 10.1109/ICYCS.2008.89.

Sharma, H. and Govindan, K. (2019) *Advances in Computing and Intelligent Systems, International Conference on Advancements in Computing & Management (ICACM-2019)*.

Shi, X. *et al.* (2018) ‘Pairwise based deep ranking hashing for histopathology image classification and retrieval’, *Pattern Recognition*, 81, pp. 14–22. doi: 10.1016/j.patcog.2018.03.015.

Sun, Y., Wang, X. and Tang, X. (2014) ‘Deep learning face representation from predicting 10,000 classes’, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1891–1898. doi: 10.1109/CVPR.2014.244.

Tian, D. P. (2013) ‘A review on image feature extraction and representation techniques’, *International Journal of Multimedia and Ubiquitous Engineering*, 8(4), pp. 385–395.

Tzelepi, M. (2018) ‘Fully unsupervised optimization of CNN features towards Content Based Image Retrieval’, *2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*, pp. 1–5.

Vagač, M. *et al.* (2016) ‘Crawling images with web browser support’, *2015 IEEE 13th International Scientific Conference on Informatics, INFORMATICS 2015 - Proceedings*, pp. 286–289. doi: 10.1109/Informatics.2015.7377848.

Vishakha (2019) ‘Issues and Challenges with the Web Crawlers’, *International Journal of Science and Research (IJSR)*, 8(3), pp. 1791–1795. Available at: https://www.ijsr.net/archive/v8i3/ART20196562.pdf.

Wang, W. S. and Wu, J. J. (2009) ‘Training asymmetry SVM in image retrieval using unlabeled data', *Proceedings of the 2009 2nd International Congress on Image and Signal Processing, CISP'09*. doi: 10.1109/CISP.2009.5303323.

Wang, X. Y., Chen, Z. F. and Yun, J. J. (2012) 'An effective method for color image retrieval based on texture', *Computer Standards and Interfaces*, 34(1), pp. 31–35. doi: 10.1016/j.csi.2011.05.001.

Wang, X. Y., Zhang, B. B. and Yang, H. Y. (2014) 'Content-based image retrieval by integrating color and texture features', *Multimedia Tools and Applications*, 68(3), pp. 545–569. doi: 10.1007/s11042-012-1055-7.

Xie, X. *et al.* (2019) 'Image retrieval with adaptive SVM and random decision tree', *Proceedings of the International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), I-SMAC 2018*, (1), pp. 784–787. doi: 10.1109/I-SMAC.2018.8653699.

Yang, S. *et al.* (no date) 'Skeletonnet: A hybrid network with a skeleton-embedding process for multi-view image representation learning', *ieeexplore.ieee.org*. Available at: https://ieeexplore.ieee.org/abstract/document/8695120/?casa_token=-R4g95Cgv-QAAAAA:nLS0FDSYbIpNk_nxheCLimt6jQLQKinOa4g3eWxE6sAKwnuSYimPQr_GCnbvaxKX9quzGn6Kxhr5DLY (Accessed: 29 April 2021).

Ye, F. *et al.* (2018) 'Remote sensing image retrieval using convolutional neural network features and weighted distance', *IEEE Geoscience and Remote Sensing Letters*, 15(10), pp. 1535–1539. doi: 10.1109/LGRS.2018.2847303.

Yu, X. and Huang, X. (no date) 'm = L a ; yixi m b k k f (x) = La ; y ; k (xi , x) + b signee) f (x) Ii Ilmll ' Y ; = -1', pp. 779–781.

Zeng, Z., Cai, S. and Liu, S. (2010) 'A novel image representation and learning method using SVM for region-based image retrieval', *Proceedings of the 2010 5th IEEE Conference on Industrial Electronics and Applications, ICIEA 2010*, pp. 1622–1626. doi: 10.1109/ICIEA.2010.5514756.

Zhang, D., Islam, M. M. and Lu, G. (2012) 'A review on automatic image annotation techniques', *Pattern Recognition*, 45(1), pp. 346–362. doi: 10.1016/j.patcog.2011.05.013.

Zhang, D. and Lu, G. (2004) 'Review of shape representation and description techniques', *Pattern Recognition*, 37(1), pp. 1–19. doi:

10.1016/j.patcog.2003.07.008.

Zhang, H., Dong, Z. and Shu, H. (2010) 'Object recognition by a complete set of pseudo-Zernike moment invariants', *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, (1), pp. 930–933. doi: 10.1109/ICASSP.2010.5495286.

Zhang, Y., Gao, J. and Zhou, H. (2020) 'Breeds Classification with Deep Convolutional Neural Network', *ACM International Conference Proceeding Series*, pp. 145–151. doi: 10.1145/3383972.3383975.

Zhao, Z. Q. *et al.* (2012) 'Cooperative sparse representation in two opposite directions for semi-supervised image annotation', *IEEE Transactions on Image Processing*, 21(9), pp. 4218–4231. doi: 10.1109/TIP.2012.2197631.

Zheng, J., Liu, X. and Fei, Y. (2014) 'Multiscale distance coherence vector algorithm for content-based image retrieval', *Scientific World Journal*, 2014. doi: 10.1155/2014/615973.