

Improvement CBIR Performance of Region-based Segmentation on DCT Images

Suhendro Y. Irianto
Department of Informatics,
Institute of Informatics and
Business Darmajaya
93 PagarAlam Street, Bandar
Lampung , Indonesia

Sri Karnila
Department of Information
System, Institute of Informatics
and Business Darmajaya
93 PagarAlam Street, Bandar
Lampung, Indonesia

Dona Yulawati
Department of Information
System, Institute of Informatics
and Business Darmajaya
93 PagarAlam Street, Bandar
Lampung, Indonesia

ABSTRACT

In this paper, we tried to discover the JPEG image which is presently an worldwide normal for pictures on the internet. Furthermost of the images presently on the storage media used regularly or on the net are in JPEG format. This JPEG format has some benefits compared to others, one of the benefit of JPEG is the size compared to others non JPEG data or information, so JPEG has an significant role in saving capacity much quality of the image. We used more than 50,000 natural images collected from the internet and other sources. Even though JPEG image has some advantages compared to others format, some researches still continue to find methos to improve CBIR performance on images, particularly on natural images. This is a reason why our research was carried out, in this work we applied region-based segmentation to improve the performance of image searching, both accurately and effectivity. We applied histogram based segmentation on natural images before deploying content based image retrieval. The evidence seems to indicate that the split and merge segmentation on JPEG image for image retrieval demonstrates a higher precision than on RGB images. Even though the precision is not radically different, the Split and Merge approach can be used as an alternative technique to improve the effectiveness of image retrieval, particularly for DCT based images. Statistically, it also can be concluded that if the number of regions generated during segmentation is high, the precision tends to be higher. In the near future this trend could be considered for larger database with greater varieties of image category in order to get more accurate results.

General Terms

Content Based Image Retrieval

Keywords

Keywords: CBIR, region segmentation, DCT, JPEG

1. INTRODUCTION

The works tried to discover the JPEG image which is presently used as an worldwide ordinary for images in the internet. Most of the images currently on the storage media which we use frequently or in the internet are in JPEG format [1]. This JPEG format has some benefits compared to other images format[2], one of the merit is the dimensions compared to others data or information, so JPEG has a significant role in saving stowage capacity lacking much quality of the image [1][3]. JPEG is a recognized standard from ISO / ITU-T formed in the first 1990s. There are some models that exist in JPEG created by, namely the baseline, lossless, progressive and hierarchical models. The baseline

model is the furthermost widespread and only supports lossy coding, and few provide lossless coding. In the baseline model the image will be converted into 8x8 blocks where each block is changed using discrete cosines transform technology or often called DCT. Then the modified block coefficient is quantized with the same quantizer, after being scanned in zigzag way and coded using Huffman coding. The baseline mode for the DC coefficients of all blocks is coded separately from 63 AC coefficients. The DC coefficient is coded differently using first order prediction using the difference formula = $DC_i - DC_{i-1}$, where DC_i and DC_{i-1} are coefficients of present 8x8 DC pixel chunk and pixel coefficient of the earlier 8x8 DC chunk. Lossless mode is very different algorithm compared to baseline mode, lossless mode uses prediction scheme based on 3 adjacent pixels.

Previous studies tried all aspects while still using entropy coding before extraction was carried out. Since the compressed file consists of variable length code which cannot be recognized without entropy decoding. Until a fairly good technique can be developed, the direct approach to the compressed domain can be said to be premature because feature extraction in the transformation of the domain coefficient is limited in some cases such as shape and color. According to [3],[4] stated that another problem is the successful development of algorithms for image processing in the pixel domain which cannot be used in JPEG images.

This research needs to be carried out by considering more and more data in form of digital images (unmoving images) and moving pictures. Unmoving image data and moving picture involve a stowage area that is plentiful larger than text. As a significance of this development, a technique can accelerate the method of indexing, retrieving, and transferring digital images is needed as well as to reduce space needed in storage, so that growing data in form of still and moving images do not decrease rapidly and presentation of hardware and image data allocation in the internet. A previous study [5] using more than 200 queries reserved from 500 training images (face image training) in his study, showed results of the matching process produced were quite good at around 77%. Algorithm used in this study is DC extraction. By using the algorithm implemented on JPEG images, then the effectiveness of this algorithm can be calculated by an average of more than 77%. To improve his research, this work added one process, namely segmentation before matching is expected to improve the results of the matching process.

In [6] splitting and merging stages implemented using the watershed transform and the self-organizing map (SOM). It is established a segmentation algorithm named "Region

Competition” in [7]. This associations universal optimization approaches (snakes/balloons and region growing) to assurance optimum global conjunction. According to [8] who is demarcated the J-image by limited windows in a quantized class-map. The high and low values in J-images parallel to probable limitations and centre of areas.

In supervised segmentation, a pixel classifier is qualified for top panel of color space with a model of entity colours. The image is segmented by conveying the pixel to one of predefined classes. The mutual methods for supervised segmentation, as well as extreme probability, decision tree, nearest neighbour and neural networks, are evaluated [8]. Supervised segmentation is deployed for the segmentation of video in [4]. The segmentation of image frames is formed into a hierarchy by three classifiers, k-nearest neighbour, naïve Bayes, and support vector machine. In [9], image segmentation is achieved by a controlled pixel arrangement procedure. The regulation of lowest distance result is used to allocate individually pixel to a exact class in a colour texture space.

2. RELATED WORKS

In splitting and merging segmentation, primarily the full image is measured as a region. If the array of features inside this section is bigger than a determined rate, then the area is split into four quadrants and individually quadrant is verified in similar means till every quadrangular section shaped in this means comprises pixels with series of features in specified rate. Split and merge algorithms twitch at approximately transitional level of quadtree (block of size $2l \times 2l$, where $l < n$) and squared Individually block for further splitting into four square sub-blocks and any two together blocks for merger. At the end over, a check is made for merging any two together sections [10].

Image segmentation is dividing one object with others in an image or among objects with a background confined in an image. With this segmentation method, individually object in the image can be reserved exclusively so that it can be used as input for other processes. Two main methods in image segmentation, one is constructed on the edge (edge-based) and the other based on the region (region-based). Segmentation is based on edge of dividing the image based on discontinuity between sub-regions. The segmentation process carried out in this study is edge detection based on the Prewitt method and Sobel method [10]. In [11] splitting and merging stages are implemented using the watershed transform and the self-organizing map (SOM). On the other hand [6] introduced a segmentation algorithm called “Region Competition” in [12]. This associations comprehensive optimization approaches (snakes/balloons and region growing) to assurance optimum global merging. According to [13] definite the J-image using limited windows in a quantized class-map. The high and low standards in J-images resemble to probable margins and midpoints of areas.

In supervised segmentation, a pixel classifier is accomplished for top partition of colour interplanetary with a trial of entity colours. The image is segmented by conveying the pixel to one of predefined classes. The mutual methods for supervised segmentation, such as determined probability, decision tree, nearest neighbour and neural networks, are evaluated. In [14] supervised segmentation is laboured for the segmentation of video shots in [8]. The segmentation of image frames is formed into a hierarchy by three classifiers, k-nearest neighbor, naïve Bayes, and support vector machine. In [15], image segmentation is completed by a supervised pixel

arrangement process. The statute of least distance result is used to dispense individually pixel to a exact class in a color texture space.

In splitting and merge segmentation, firstly the full image is measured as one section. If the series of features in this section is larger than a determined value, then the section is split into four quadrants and individually quadrant is verified in similar way till each square section formed in this way comprises pixels with series of features in the given value. Split and merge algorithms start at some intermediate level of the quadtree (block of size $2l \times 2l$, where $l < n$) and patterned each block for further splitting into four square sub-blocks and any two adjacent blocks for merging. At the end again a check is made for merging any two adjacent regions [16]. Lastly, split and merge (SaM) segmentation is applied to calculate the regions used to build the indexing keys for image retrieval.

2.1 Content Based Image Retrieval

Recently, the cost of providing large amounts of internal and peripheral storage has developed quite economy, manufacture it more monetary to store art groups, digital images, medical images, folders of family photographs, and all-purpose assortments of images. The size of an image database can be huge containing hundreds to millions of images [17]. Some image databases have been created to display the working of specific content base image retrieval systems, such as [3], [5], [18], and [19].

WebSeek is a model image and video search engine which gathers images and videos from the Web and catalogues. Blobworld is a scheme for content based image retrieval operating on the principle of discovery comprehensible sections that match to objects. The PicToSeek scheme uses color disparity, saturation, evolution forte, related and greyness as the indexing structures to offer content based entree to images. Whilst [20] stated that C-BIRD is a system that uses an illumination invariant method to retrieve colour images over a color canal stabilization phase. Introduce by [18] MARS as considered as multimedia analysis and retrieval system) which is a system that permits image queries with colour, texture, shape and outline features.

Image content includes[9]: (i). Perceptual properties like colors, texture, shape, and longitudinal relations, (ii) Semantic primitives matching to concepts, like objects, characters, and passages, (iii) Imprints, reactions and meanings associated with combinations of perceptual features. The method used to determine the likeness amongst query images and image in the database depends on which image distance measure or measures are employed. There are four (Squire et al. 1999) major classes of measures: color similarity, texture likeness, shape likeness, and object and relationship likeness.

Content based image retrieval has a wide-ranging space for application, and several measures used. [17] used a list of keywords as due to individually image is explained by keywords (comment words) which describe content of image, and this is termed query-by-text (QBT). QBT is applied when users need to easily prompt their quest requests. In additional requests, enquiries may be delivered through provided that instance images, and this category of image retrieval is consequently called query-by-example (QBE). Most QBE retrieval is “content-based image retrieval”[7].

Content-based image retrieval is the retrieval of relevant images from an image database based on automatically

derivative features. The necessity for this kind of method has improved tremendously in various application areas as well as biomedicine, crime anticipation, the martial, commerce, values, education, entertainment, and web image cataloguing and penetrating. For this purpose, content-based image retrieval has been broadly considered. Though, space limits do not tolerate this to current a comprehensive examination. Instead, importance is located on some of literatures that is utmost connected to the research proposed.

A wide range of indexing and retrieval approaches have been used in image retrieval systems. Some systems use keywords and full-text descriptions to index images. Other systems use features such as colour histogram, colour layout, limited texture, wavelet coefficients, and shape to index images. An example of a recent system is SimPLicity,[21]. Furthermost image retrieval systems sustenance one or more of the subsequent possibilities. Moreover arbitrary glancing, search by example, search by sketch, search by text (with keyword or language), and navigation with tailored image classes. A useful objective of an image retrieval system is to retrieve the images in rank directive, where the rank is determined from the significance to the enquiry. The global retrieval effectiveness of a system can be measured only if the actual relevancies are known. In broad outline, an information retrieval system assessment test-bed contains of a group of N images, a set of M standard enquiries, a set of ground-truth significance scores for the benchmark queries, and a set of assessment

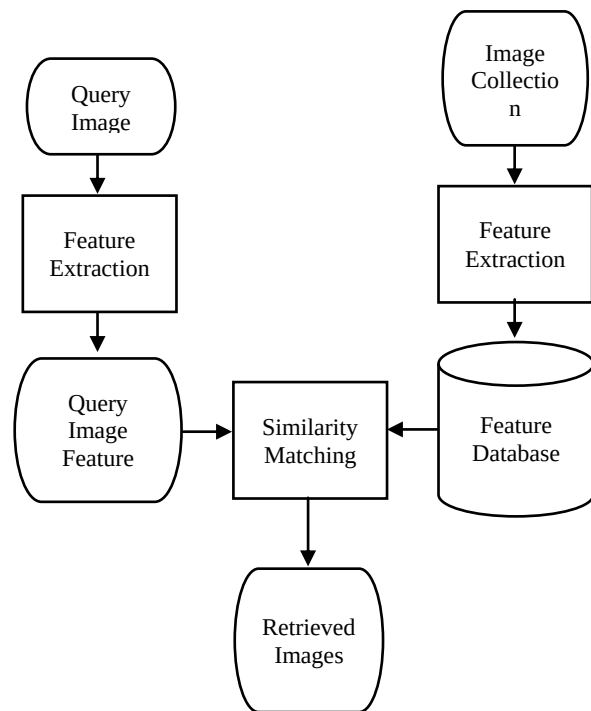


Figure 1. The CBIR system

2.2 Precision and recall

It is not useful to retrieve all the appropriate images in a huge image collection. In various researches therefore, the assessments are premeditated based on a few images retrieved throughout a query by computing precision and recall.

$$precision = \frac{c}{b}$$

$$recall = \frac{c}{a}$$

Where a is the total amount of appropriate images in the group, b is the number of images retrieved and c is the number of relevant images retrieved.

These are standard procedures in image retrieval, which give a virtuous signal of system recital. However, neither precision nor recall value unaccompanied contains sufficient information. We can constantly make recall 1, merely by retrieving all images. Correspondingly, precision can be kept high by retrieving only a few images. Thus, precision and recall should either be used together or the number of images retrieved should be specified. Precision and recall are often averaged, but it is important to know the basis on which this is done [22].

2.3 Retrieval efficiency

According to [5] which defined retrieval efficiency as in the equation (1). If the amount of images retrieved is subordinate than or identical to the amount of relevant images, this rate is the precision; else it is the recall of a query. This explanation can be ambiguous due to it mixtures two average measures,[23].

$$Retrieval\ efficiency = \begin{cases} \frac{N_r}{N} & , \text{ if } T > N_r \\ \frac{N_r}{T_r} & , \text{ otherwise} \end{cases} \quad (1)$$

Where

N_r = Amount of related images retrieved

T = Total amount of images retrieved

T_r = Total amount of relevant images

2.4 JPEG

JPEG is the standard method used in compression for photographic images. The term JPEG takes its name from the Joint Photographic Experts Group committee that issued the standard from the format in 1992, which was later recognized as ISO / IEC 10918-1 in 1994. This JPEG format regulates how a digital image is made properly through a series of byte compression method. so the jpeg file size becomes smaller, which is about 10% of the bmp file format without compression. JPEG is made for lossy compression images. Generally JPEG-format files use extensions (.jpeg, .jpg, .jpe, .jfif, .jif). In addition, JPEG is also able to provide colors with a depth of 24 Bits, corresponding to 16 million colors. The image file compression standard formed by the Joint Photographic Experts Group produces very large compressions but with the result of distortions in images that are almost always invisible. is an image format, very useful for creating high-quality photographic images in very small file sizes. This graphic file format has been accepted by the Telecommunication Standardization Sector or ITU-T and the International Organization for Standardization or ISO [11].

3. MATERIALS AND METHODS

In this work, 50,000 images were employed, which consisted of seven classes including bike, building, car, cat, flower, mountain, and sky. In order to reduce algorithm complexity, this deployed preliminary processing by resizing images as 2 KB to 10 KB for all in the database. Examples of images group in the database can be observed on fig.3.



Figure 3. A sample of the ten image categories in the database

The work was carried out by adopting the split and merge algorithm introduced by [15] and integrated with CBIR method, the proposed algorithm is presented at figure 5.

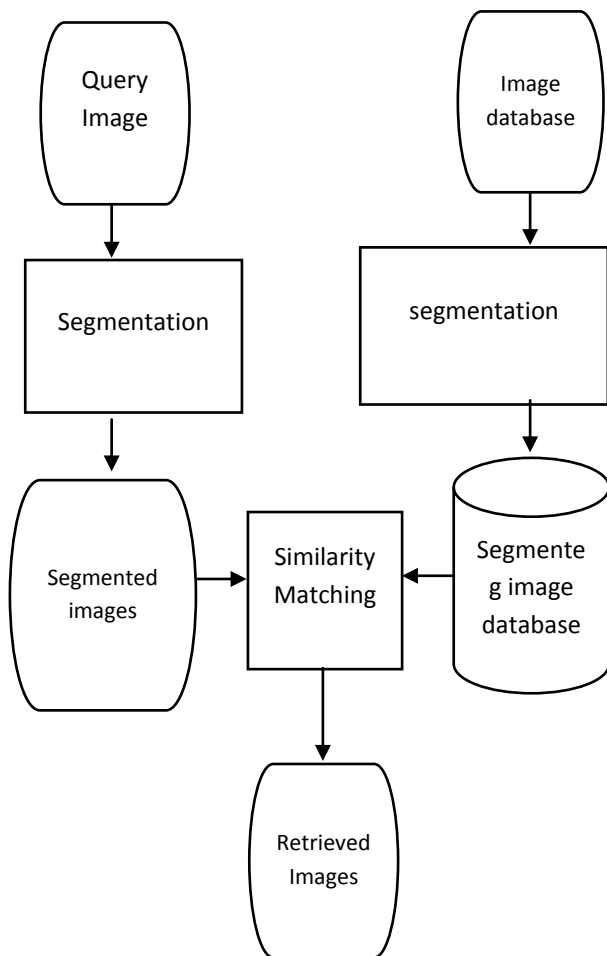


Figure 2. The proposed system (hybrid of segmentation and CBIR)

Figure 2 and 3, explained of CBIR method modification with region growing segmentation method. The proposed this modification is to accelerate and improve accuracy of image retrieval since images be segmented first before applying CBIR. It is carried out CBIR merely of matching on region interest rather than others. This research seem very simple , but shows significant result improvement compared to DCT based CBIR in term of precision and recall. Table 1, demonstrates the outcomes of the work with 10 images

category og groups. Splitting and merging algorithms can be made extra effective by recursively splitting the image into minor areas till all individual areas are articulate, then recursively merge these regions to produce larger comprehensible areas. First, to split the image, start by seeing the whole image as one area. (i) If the whole area is coherent or has appropriate similarity, leave it unchanged. (ii). If the region is not sufficiently comprehensible, split it into four quadrants and recursively apply these steps to each new region. Most split and merge segmentation algorithms are applied in the pixel domain, and very rarely in the compressed domain. For this reason, in the current chapter a method of image segmentation in the compressed domain is proposed, specifically in the JPEG DCT coefficient domain. Furthermore, the proposed method is used to segment the DC image, and then calculate a transform distance to build indexing keys for image retrieval. The work presents a new technique for deriving indexing keys on DCT segmented image in order to reduce complexity and the time consumed for indexing.

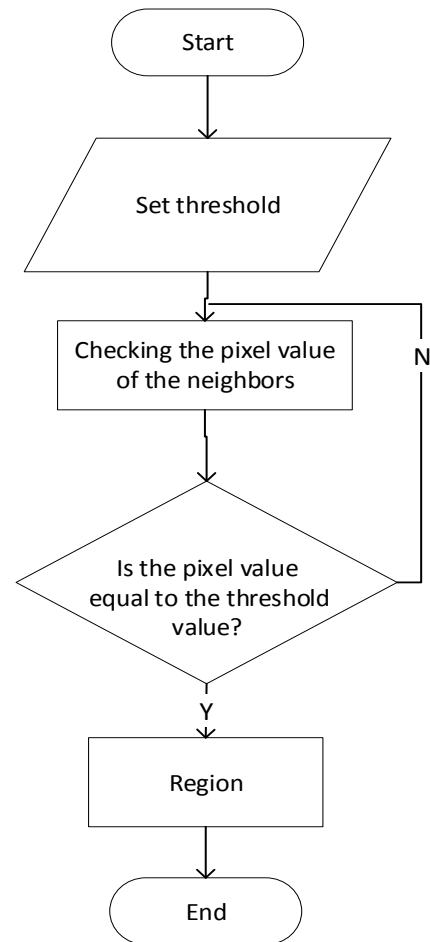


Figure 4. Split and Merge segmentation algorithm

4. RESULTS AND DISCUSSION

In the results the cat class achieved a good precision of 89% while the mount class achieved a precision of 22% which is very significantly lower. This work also evaluated image retrieval using all the DCT coefficients and the result was quite different. The average precision achieved by split and merge was then 85% and 51% for RGB/DCT images, for further illustrate can be examine on figure 5. The main advantages of using segmentation on DCT images for image

retrieval is that the proposed technique might well save time and reduce computing costs. The merit of using a segmented image for image retrieval is that we do not need to compare one image to another using the full image but only by using the region that is the most important part (object) of the image, which means saving time and reducing computing costs. It can be said that using more regions in a segmented image leads to a higher precision in image retrieved. The result of Split and merge and DCT based methods is presented by precision Table 1.

Table 1. The average of precision of 10 group images with 50,000 images as a database

IMAGE	S&M*	DCT*
cat	0,66	0,33
flower	0,76	0,34
sky	0,78	0,76
bike	0,78	0,37
building	0,79	0,37
car	0,87	0,67
bear	0,87	0,69
mount	0,89	0,44
model	0,95	0,39
texture	0,98	0,72
Average	0,833	0,508

*average of precision with 20 queries

Table 1 explained average of precision of 10 image groups from cat, flower, and texture. And every group was carried out as many as 20 queries. Therefore, for this research, we applied 200 queries. In addition to the experimental results mentioned, figure 5 shows a rough trend that precision increases as the number of areas growths. The number of regions is calculated based on average of ten image queries made on each class. The largest number of regions is 343 for the cat class, with a precision of 98% whilst the lowest is 212 for the cat class with a precision of 66%. An application of split and merge retrieval to RGB and DCT images. Efficiency of image search, in this research, 200 `each query 20 most similar images will be showed which will then be calculated precision:

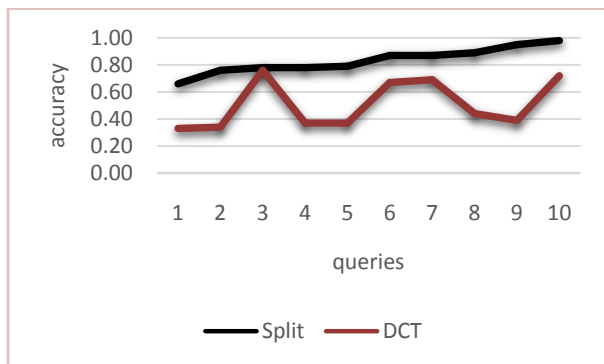


Figure 5. Average of precision if 10 image groups

Figure 5 explains the precision of image retrieval where magenta line represent retrieval on DCT images. Whilst Split and Merge segmentation retrieval represented with blue

line. When the results was analysed and split and merge retrieval method demonstrates higher precision compare to DCT retrieval. This condition due to that split and merge segmentation method was deployed first as previous process before CBIR applied. Split and merge segmentation was carried out in order to make matching development in CBIR more effective. This cause matching process only match object selected without others.

5. CONCLUSION AND FUTURE WORKS

The evidence seems to indicate that the split and merge segmentation on DCT image for image retrieval demonstrates a higher precision than on RGB images Even though the precision is not radically different, the Split and Merge approach can be used as an alternative technique to improve the effectiveness of image retrieval, particularly for DCT based images. Statistically, it also can be concluded that if the number of regions generated during segmentation is high, the precision tends to be higher. In the near future this trend could be considered for larger database with greater varieties of image category in order to get more accurate results.

6. ACKNOWLEDGMENTS

Our thanks to the research centre of Darmajaya Informatics and Business Institute, Indonesia who have contributed finishing the work from the beginning.

7. REFERENCES

- [1] Y. Dong, "Image Classification in JPEG Compression Domain for Malaria Infection Detection," 2022.
- [2] M. Image and T. Ccl, "No Title."
- [3] A. Kumar, "Automatic Feature Weight Determination using Indexing and Pseudo Relevance Feedback for Multi-feature Content Based Image Retrieval," pp. 1–9, 2018.
- [4] I. M. Hameed, S. H. Abdulhussain, B. M. Mahmmod, I. M. Hameed, S. H. Abdulhussain, and B. M. Mahmmod, "Content-based image retrieval: A review of recent trends COMPUTER SCIENCE | REVIEW ARTICLE Content-based image retrieval: A review of recent trends," *Cogent Eng.*, vol. 8, no. 1, 2021, doi: 10.1080/23311916.2021.1927469.
- [5] N. P. Kumar, "Content Based Image Retrieval Using Shape , Color and Texture," vol. 7, no. 5, pp. 1151–1158, 2018, doi: 10.21275/ART20182457.
- [6] M. Lachaize *et al.*, "Evidential split-and-merge : Application to object-based image analysis To cite this version : HAL Id : hal-01918408 Evidential Split-and-Merge : Application to Object-Based," 2018.
- [7] R. Aslanzadeh, "An Efficient Evolutionary Based Method For Image Segmentation," pp. 1–17.
- [8] C. Ipn, "segmentation in fluorescence microscopy images," 2021, doi: 10.1016/j.bspc.2019.101575.Split.
- [9] O. A. Adegbola *et al.*, "Modified one-class support vector machine for content-based image retrieval with relevance feedback Modified one-class support vector machine for content-based image retrieval with relevance feedback," *Cogent Eng.*, vol. 5, no. 1, pp. 1–20, 2018, doi: 10.1080/23311916.2018.1541702.

- [10] I. O. F. S. Segmentation, "Markku Hautakoski Helsinki University of Technology Institute of Photogrammetry and Remote Sensing Helsinki, Finland," pp. 32–39.
- [11] T. Zhang, G. Lin, W. Liu, J. Cai, and A. Kot, "Splitting vs . Merging : Mining Object Regions with Discrepancy and Intersection Loss for Weakly Supervised Semantic Segmentation," pp. 1–17.
- [12] G. R. Jothilakshmi, P. Sharmila, and A. Raaza, "Mammogram Segmentation using Region based Method with Split and Merge Technique," vol. 9, no. October, 2016, doi: 10.17485/ijst/2016/v9i40/99589.
- [13] C. T. Rueden *et al.*, "ImageJ2: ImageJ for the next generation of scientific image data," pp. 1–26, 2017, doi: 10.1186/s12859-017-1934-z.
- [14] D. Adlakha, D. Adlakha, and R. Tanwar, "Analytical Comparison between Sobel and Prewitt Edge Detection Techniques," vol. 7, no. 1, pp. 1482–1485, 2016.
- [15] L. Gatos, "Split-and-Merge Segmentation of Aerial Photographs," vol. 6, 1988.
- [16] S. K. Katiyar, "Comparative analysis of common edge detection techniques in context of object extraction," vol. 50, no. 11, pp. 68–79.
- [17] B. Shikha, P. Gitanjali, and D. P. Kumar, "An Extreme Learning Machine-Relevance Feedback Framework for Enhancing the Accuracy of a Hybrid Image Retrieval System," pp. 15–27, 2020, doi: 10.9781/ijimai.2020.01.002.
- [18] K. Porkaew, S. Mehrotra, M. Ortega, and K. Chakrabarti, "Similarity Search Using Multiple Examples in," pp. 68–75, 1999.
- [19] T. Gevers and A. W. M. Smeulders, "PicToSeek : Combining Color and Shape Invariant Features for Image Retrieval," no. January 2000, 2014.
- [20] Z. Li, O. R. Za, B. Yan, and C. Va, "C-BIRD : Content-Based Image Retrieval from Digital Libraries Using Illumination Invariance and Recognition Kernel C-BIRD: Content-Based Image Retrieval from Digital Libraries Using Illumination Invariance and Recognition Kernel," no. March 1998, 2015, doi: 10.1109/DEXA.1998.707425.
- [21] [21] E. R. Andersson, R. Sandberg, and U. Lendahl, "Notch signaling: simplicity in design , versatility in function," no. September, 2011, doi: 10.1242/dev.063610.
- [22] S. Blue, "C OMBINING S TRUCTURE , C OLOR AND T EXTURE FOR I MAGE R ETRIEVAL: A P ERFORMANCE E VALUATION Qasim Iqbal and J . K . Aggarwal Computer and Vision Research Center Department of Electrical and Computer Engineering The University of Texas at Austin Austin , Texas 78712 , USA," vol. 2, pp. 438–443, 2002.
- [23] S. Eickeler, "High quality face recognition in JPEG compressed images," no. February 1999, 2014, doi: 10.1109/ICIP.1999.821721.