# Colour-Based Extraction Methods for the Classification of Breast Milk (ASI)

**Rogayah[*1], Waliya Rahmawanti[2], Nur Azizah[3]**
[1]Fakultas Teknologi Industri Universitas Gunadarma, [2]Fakultas Ilmu Komputer Universitas Gunadarma, [3]Fakultas Sains & Technology University Of Raharja
Email: [*1]**rogayah@staff.gunadarma.ac.id**, [2]waliya@staff.gunadarma.ac.id, [3]nur.azizah@raharja.info

### Abstract

*The development of cellular devices makes accessing information in the form of text or images more easier. In line with the growing field of computer vision, various processes in image/image processing continue to increase. Image processing can be done by increasing image quality (image enhancement) and image recovery (image restoration). Feature extraction is divided into three types, namely feature form extraction, texture feature extraction, and color feature extraction. The application of color-based feature extraction methods has been widely used by researchers in the process of classification of various objects. This paper aims to review the technology that can be applied to image processing in a CBIR system with the object of breast milk so that it can measure the quality of breast milk based on its color.*

*Keywords — Breast Milk, CBIR (Content Based Information Retrieval), Color Feature Extraction, Image Processing.*

## 1. INTRODUCTION

The development of mobile devices that is very rapid makes accessing information in the form of text and images easier. The availability of supporting devices such as cameras and scanners makes anyone able to capture images well thanks to the higher quality of camera resolution. In line with the growing field of computer vision today, various processes in image / image processing continue to increase. Image processing is every form of signal processing where input is an image, such as a photo or video frame, while the output of image processing can be an image or a number of characteristics or parameters related to the image. Most processing techniques involve or treat photos as two-dimensional dimensions and apply standard signal processing techniques to them, usually this refers to digital image processing, but can also be used for optical and analog image processing. Image acquisition or the resulting image input in the first place is referred to as imaging.

Image processing can be done by increasing image quality (image enhancement) and image recovery (image restoration). Improving image quality aims to improve certain features of an image. The term extraction feature or feature extraction is a feature taking from a form which later the values obtained will be analyzed for the next process. Feature extraction is divided into three types, namely feature form extraction, texture feature extraction, and color feature extraction. The application of color-based feature extraction methods has been widely used by researchers in the process of classification of various objects. The research conducted by Setyaningtyas in 2019, using the YUV Color Moment color feature extraction method, extracting texture features using Haralick, and feature selection using Information Gain. The classification used is K-Nearest Neighbor (KNN). The highest accuracy obtained is 94.26% obtained from the combination features of the two selected feature extraction methods. From

these results, it can be concluded that the application of a combination of feature extraction methods namely color and texture, as well as feature selection methods greatly influence the food image classification process.

Content-based information retrieval (CBIR) is a computer-based system that can be used to store, find and display the document images sought in a digital image document database based on image processing, one way that can be done for image processing is to use methods color feature extraction. Other research objects that can be developed can be either in the form of water or liquid objects such as breast milk or commonly called ASI. Mother's Milk (ASI) is the only perfect and best food for babies because it contains nutritional elements needed by babies to achieve optimal growth and development. There is no food that is better than exclusive breastfeeding for newborns until the first 6 months of life [4]. According to the Ministry of Health on ministerial decree Number: 450 / MENKES / SK / VI / 2004 concerning exclusive breastfeeding in Indonesia, namely: the first, establish exclusive breastfeeding in Indonesia for 6 months and continue until the age of 2 years with appropriate supplementary food. Secondly, health workers inform all new mothers to give exclusive breastfeeding by referring to the 10 steps of successful breastfeeding.

Nutritional content in food intake for the body must be balanced, consisting of carbohydrates, proteins, vitamins, minerals and other substances that are needed, as well as the nutritional content of breast milk consists of elements of water, colostrum, carbohydrates, proteins, fats, carnitine, vitamins and enough minerals. Table 1 shows a comparison of the nutritional levels of collostrum, foremilk, and hind milk.

**Table 1.** Comparison of the nutritional levels of colostrum, foremilk, and hindmilk in breast milk.

| Composition | Collostrum | Foremilk | Hindmilk |
|---|---|---|---|
| Energy | 57.0 | 63.0 | 65.0 |
| Laktosa (gr/100 ml) | 5 | 6.7 | 7.0 |
| Lemak (gr/100 ml) | 2.9 | 3.6 | 3.8 |
| Protein (gr/100 ml) | 1.195 | 0.965 | 1.324 |
| Mineral (gr/100 ml) | 0.3 | 0.3 | 0.2 |
| Immunoglobulin: | | | |
| Ig A (mg/100 ml) | 335.9 | - | 119.6 |
| Ig G (mg/100 ml) | 5.9 | - | 2.9 |
| Ig M (mg/100 ml) | 17.1 | - | 2.9 |
| Lisosin | 14.2-16.4 | - | 24.3-27.5 |
| Laktoferin | 420-520 | - | 250-270 |

The quality of nutritional content is a very important part to be known by nursing mothers to support the growth of the baby.

**Figure 1.** Breast Milk Color

The color in breast milk is not always the same as described in Figure 1. that colostrum is yellowish and thicker because it contains high protein and vitamin A, besides colostrum also contains immune substances which are important to protect babies from infectious and bacterial diseases. Fat-soluble vitamins are higher than mature breast milk. Breast milk released at the first time the milking / pumping process will look "thinner" than the milk released in the next minutes called foremilk (because it is rich in protein) the color tends to be bluish clear white. While the milk that comes out a few minutes later will look thicker, also called hindmilk (rich in fat) in thick white.

## 2. STATE OF THE ART

### A. Image

Image (image) is one of the multimedia components that plays an important role as a form of visual information. The image is literally interpreted as an image in a two-dimensional field (dwimatra) which is composed of many pixels, the smallest part of the image.

### B. Digital Image

Digital imagery is produced from analog image recording to storage in the form of image files using a scanner, digital camera, camcorders or transformed by digitizing the analog image. Analog imagery is the result of reflecting light from images received by optical systems such as the human eye and analog cameras. Digital imagery is a matrix that contains real and complex values written in bits on rows and columns. Digital images can be written in the form of a function f (x, y), which is the light intensity of the matrix M lines and N columns. The image points at position (x, y) are picture elements or pixels.

### C. Extraction of Color Features

Feature extraction is the stage of extracting the features / information from an object in an image that you want to recognize / distinguish from other objects. The extracted features are then used as input parameters to distinguish between objects from one another at the identification / classification stage. To distinguish an object with a certain color can use the value of hue which is a representation of visible light (red, orange, yellow, green, blue, purple). The value of hue can be combined with the value of saturation and value which is the brightness of a color. To get these three values, it is necessary to convert the color image space that was originally RGB (Red, Green, Blue) to HSV (Hue, Saturation, Value).

### D. Color Feature Extraction Using the BQMP Method

The pixel value in the image is a random vector, this image color distribution will be equivalent to the probability distribution of this random vector. With probability theorems, random vectors can be characterized by the moment of statistics. Application of the Binary Quaternion-Moment-Preserving (BQMP) threshold technique proposed for color feature extraction problems. The BQMP threshold technique can divide the data set into two subsets

and, each has a representation, i.e., each. If we change the data point with and with, the first 3 four-digit moments are stored in the resultant and. It is known that the higher the sequence of moments, the less important the moment is. Thus, keeping the moment until the third moment will not interfere with the image distribution. The termination process is for a fixed number of clusters and the number of cluster variables is different from each other. The process for clusters still ends when the specific number of pixel clusters has been extracted. While the process for the number of cluster variables ends when the number of pixel clusters is extracted until enough to represent the image. This variance threshold is predetermined based on the results and previous experiments.

### E. Color Feature Extraction Using the Histogram Method

The histogram calculates the number of pixels of each type and can be made quickly by reading each pixel of the image only once and adding the appropriate histogram tray. The relative color histogram does not change in translation, rotation on the imaging axis, small off-axis rotation, scale changes and partial occlusion. To extract the color histogram, at first, the RGB color space is converted to the HSV color space. Color histograms are calculated based on scalable colors. In Srivastava's study, the color space was uniformly quantified to 16 hue levels, 4 saturation levels and values giving a total of 256 bins. To decrease this number and make the application scalable, the histogram is encoded using Haar transform. The use of subset coefficients in Haar representation is 64 bins. Global color histograms are made for all images. For the complexity of image computing, the number of homogeneous blocks that cannot be divided again is calculated for each level.

### F. Color Feature Extraction Using the Color Histogram Method

Color Histogram is the most widely used technique for extracting image color features. This represents an image from a different perspective. This represents the frequency distribution of the color bin in an image. This calculates the same pixels and saves them. There are two types of color histograms, namely. global color histogram and local color histogram. The color histogram is proposed as a global color descriptor that analyzes each statistical color frequency in an image. This is used to solve problems such as changes in translation, rotation and point of view. The local color histogram focuses on the individual parts of an image. The local color histogram considers the spatial distribution of pixels lost in the global color histogram. The color histogram is easy to calculate and insensitive to small variations in images so it is very important in indexing and retrieving image databases. Apart from these advantages, he faced two major weaknesses. First, overall spatial information is not taken into account. Second is that the histogram is not strong and unique because two different images with the same color distribution lead to the same histogram, the same display image with different lighting makes a different histogram.

### G. Extraction of Color Features Using the Gabor Method

New approach to multi-texture images, segmentation based on the formation of effective textures feature vectors with gray space scales. Vector texture features obtained from the integration of Gabor single filters with local binary patterns (LBP). The research conducted by Wicaksono proposed a method for extracting color and texture features using the Gabor-LBP filter applied for research data on image segmentation taken from the Berkeley Segmentation Dataset (BSDS) secondary data. BSDS is an RGB image that will be converted to L * a * b * color space, so that it has three values L, a and b. From these values represent the grayscale L color space value to extract the texture feature while the values a and b represent the feature color. The method proposed to obtain a single Gabor filter becomes efficient for a small number of texture classes, but for three or more, it is not possible to distinguish the difference in texture. When integrated with LBP image segmentation it gets better. Integrating Gabor - LBP filter extraction and the colors of components a * and b * into image features that are used as image attributes in grouping using Fuzzy C-Means (FCM).

## 3. RESEARCH METHOD

Content-based information retrieval (CBIR) is a computer-based system that can be used to store, find and display the document images sought in a digital image document database based on image processing, one way that can be done for image processing is to use methods color feature extraction.

Several studies have been carried out for image processing based on color feature extraction, including the research of Kekre in 2011, presented an algorithm to search for and take pictures as outlined in this study. Use the average of each component (color averaging) of the Red, Green and Blue images separately and two databases of different sizes. One database consists of 300 images while the other database has 1000 images in it. This paper shows the first 20 draws from a particular class, precision cross-recall plots from five randomly selected images for the same and finally cross-recall average grade precision. These three results have been shown for both databases. The average precision and memory of each class, which has a cross point of 50% for the database of 300 images which is a good performance and 30% for a database of 1000 images. The results also concluded that the size of the database and the number of different classes affected the relevance of the images taken. To take the desired picture, the user must provide a request image. The system then performs certain features of the extraction procedure and represents it in the form of vector features. The similarity of distance between feature vectors from the request image and people from the image in the database is then calculated and retrieval is done with the help of an indexing scheme. The indexing scheme provides an efficient way to search image databases. The retrieval system has recently included user relevance feedback to modify the retrieval process to produce search results that are more semantically and semantically meaningful. The final retrieval results from the CBIR system are based on how well the similarities between the classes they have and the size of the image database.

Srivastava researchers and friends in 2015 conducted a study on CBIR which aims to take relevant images from large image databases, by measuring the similarity between request images and database images. It is based on three low-level basic features, namely color, texture and shape, and color is the most important among them. The methods (Color Histogram, Intersection Histogram, and Color Correlogram) are used and selected according to the application they are using. Methods that consider spatial information from images are increasingly popular, because they are more beneficial in a variety of other perspectives. In addition, the method must be strong against noise, scaling, rotation, flipping, and efficient in storage space and time complexity parameters.

Researchers Wicaksono and friends in 2015 conducted experiments and tests on colors that contain a lot of information and vision humans can see thousands of color combinations and intensities compared to gray scale or black and white (binary). BSDS image segmentation using FCM by applying color feature and texture extraction using Gabor filter-LBP then for the use of color features compared to the grayscale level increased the accuracy of 16.54% for Gabor filter texture from 76.83% to 93.37% and increased to 14 , 57% of 78.85% of LBP filters became 93.42%. While the LBP texture feature can help improve image segmentation accuracy by 2.02%, even though the small color space scale is from 76.83% to 78.85% and 0.05% in the L * a * b * color space of 93.37% to 93.42%.

## 4. RESULTS AND DISCUSSION

Patil and Dusan in 2012, conducting research using color feature extraction techniques to consider the color distribution of images. For the color feature extraction process clustering is used, where the number of clusters is fixed and the number of cluster variables is formed. The proposed technique maintains the image color distribution and reduces the distortion that occurs during the feature extraction process by using a binary moment preserving (BQMP)

technique. The pixel value in the image is considered a random vector, this image color distribution will be equivalent to the probability distribution of this random vector. With probability theorems, random vectors can be characterized by the moment of statistics. With this system, efficient relevance feedback is developed to improve the efficiency of content-based shooting systems. Experimental results show an increase in the proposed system of previous binning techniques and ACE. This research attempts to develop methods and solutions to problems and shortcomings that have been carried out by researchers. This research is expected to be implemented on objects in the form of breast milk. The stages of research carried out are as follows:

1. Taking Image / Image
2. Conducting Image Acquisition
3. Image processing, consisting of: Pre-approval and Extraction Feature
4. Classification

As done by Patil and Dusan, this study uses the BQMP method, the following figure 2 shows how the system works.
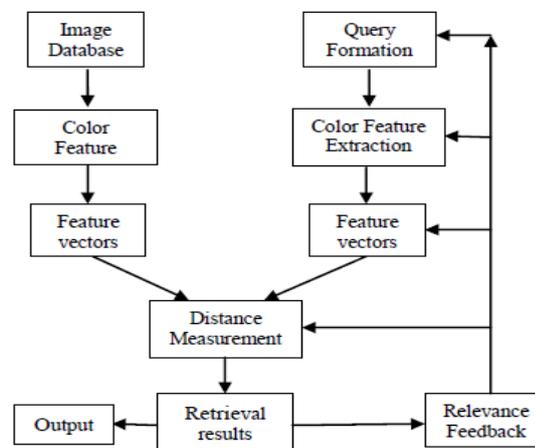


**Figure. 2.** Working Model System

In the proposed CBIR system, the relevance feedback process is carried out as follows:

a. User requests to the CBIR system by sending a request image.
b. The system will extract color features by considering the color distribution of the image. Using the extraction strategy with fixed cardinality and variable cardinality, different sets of extracted colors are generated and this is represented by a histogram.
c. After feature extraction, all features extracted will form in feature vectors.
d. The query image is then compared with the image database by calculating the distance inequality between them.
e. Rank all relevant images and sort them in ascending order based on the distance difference.
f. Show the resulting image with the highest rating.
g. If the user does not meet the retrieval results displayed by the system, the user can provide relevant image feedback to the system.
h. As before, the system processes the feature extraction module.
i. The system then ranks the image and displays the image that is most relevant to the user.

## 5. CONCLUSION

The researcher refers to the research conducted by Patil and Dusan in their research regarding the use of color feature extraction methods using the clustering method, in which the method has successfully compared the ACE method with the proposed method. The development of existing methods with feature extraction uses a classification method by researchers to look for differences between the two in breast milk objects (ASI).

## 6. REFERENCES

[1]     Amalia, Ruhana dkk. (2016) "Pengaruh Waktu dan Suhu Penyimpanan Terhadap Kadar Asam Amino Taurin pada ASI". Indonesian Journal of Human Nutrition, vol. 3, no.2, pp. 60 – 68.

[2]     Boltz, S. (2009). Image Segmentation Using Statistical Regian Merging. IEEE Trans. [online] 12[2], pp. 1452-1458.

[3]     Kekre, H.B. (2011). "Color Reature Extraction For CBIR". International Journal of Engineering Science and Technology (IJEST), vol 3 no. 12. pp. 8358- 8365.

[4]     Lawrance, RA. (1980). Breas- Feeding, a guide for the Medical Profession. St.Louis, Toronto, London:  The Mosby Company.

[5]     Patil, Sandip.S, Dusane, Atul P. (2012). "Use of Color Feature Extraction Technique based on Color Distribution and Relevance Feedback for Content based Image Retrieval". International Journal of Computer Applications. vol. 52, pp. 9-11.

[6]     R, Munir. (2004). Pengolahan Citra Digital dengan Pendekatan Algoritmik. Bandung, Indonesia: Informatika.

[7]     Srivastava, Divya, Wadhvani, Rajesh, Gyanchandani, Manasi. (2015). "A Review: Color Feature Extraction Methods for Content Bases Image Retrieval". IJCEM (International Journal of Computational Engineering & Management), vol. 18 Issue 3. Pp.9-13.

[8]     Wahidun, Sipayung. (2014). "Perancangan Citra Watermaking Pada Citra Digital Menggunakan Metode Discrete Cosine Transform (Dct")". Pelita Informatika Budi Darma, vol. 7,  no. 52, pp. 104-107.

[9]     Yanuar, Wicaksono, dkk. (2015). "Color and Texture Feature Extraction Using Gabor Filter – Local Binary Patterns for Image Segmentation with Fuzzy C-Means". Journal of Intelligent Systems, vol. 1, no. 1. p 15-21.