Medicine Recognition Using Intrinsic Geometric and Texture Property from Pill Image

By

Tanjina Piash Proma

ID: 1320810

August, 2017

Supervised By:

Dr. M. Ashraful Amin

Head, Department of Computer Science & Engineering
School of Engineering & Computer Science

Senior Project Report

Bachelor of Science

Computer Science

Independent University, Bangladesh

Department of Computer Science & Engineering

School of Engineering & Computer Science

Independent University, Bangladesh

Abstract

Prescription pills mostly comes in air-tight plastic bottles or labeled Ziploc bags where the identity of the pills are unobtainable once they are removed from their specified labels, bottles or bags. It causes a great dilemma for common people outside of the expertise field to know the pills identity without the labels. And the task becomes even harder for the people of visual impairment syndrome. People of hazy or blurred vision or the elderly people finds it way too challenging just to identify the pills. And various pills of various shapes, size, texture, color comes with a diverse set of medicinal components which creates confusion among pills of same color and shape to identify based on a specific texture. Even if they configure the shape of the pill, the texts imprinted on the pill remains unknown to them. And all these causes as in most cases, the prescription pills do not come with blister or alu alu packaging where the identity of the pill is available, making it further complicated to specify the medicine. Recognizing a pill from a given dataset using computer vision techniques requires multiple steps. In this paper, the splitting processes of a dataset according to the shape and the texts embossed on the pills, will be described. To find the shape information we used fundamental geometric properties such as: eccentricity, extent and narrowness of pill which can be extracted from image using carefully selected image processing techniques. And for finding the texts segmentation and text recognizing techniques has been used. Reference values of discriminative parameters are determined using 'RxIMAGE', National Library of Medicine, USA database. The overall shape discrimination accuracy of the proposed system is 93.75%, detecting the number of color is around 95.6% and text recognition accuracy of the printed texts is 81.32%.

List of Publications

[1] Md. Zakir Hossan, Tanjina Piash Proma, M. Ashraful Amin “Medicine Recognition Using Intrinsic Geometric Property from Pill Image” The 14th Pacific Rim International Conference on Artificial Intelligence 2016 (PRICAI 2016)

Acknowledgements

I would like to thank my Supervisor Dr. Md. Ashraful Amin, Head of CSE Dept. of Independent University, Bangladesh (IUB) for supporting, providing me with all the necessary facilities for the research. I am extremely thankful and indebted to him for sharing expertise, and sincere and valuable guidance and encouragement extended to me.

I would also like to thank Md. Zakir Hossain, co-author of the papers published on international conferences, for his assistance, guidance, help and support.

I take this opportunity to express gratitude to all of the Department faculty members for their help and support. I also thank my parents for the unceasing encouragement, support and attention. I am also grateful to my partner who supported me throughout this venture and encouraged me to do better work.

Table of Contents

List of Figures

1 Introduction

For Proper medication, accurate identification of pills is crucial. Without the consultancy of the doctors as consuming pill is dangerous, in that way random consumption of pills without distinguishing them can unfortunately be severely threatening. Thus, the consumption of pills requires special attention for ensuring the safety of consumer's health condition. And for the case of identification, the better technique is to concentrating on the key features of a pill like texture, to sort out the initial assessment. After that the medicinal properties of the pill should be reviewed. As the identification process is highly dependable on vision and to make it precise, the perspective of classifying the texture has to be reasonable as well. In this paper, the process of extracting the features to identify the pills based on the texture will be discussed.

1.1 Background and Context

Prescription pills are available in blister, alu-alu or container (bottle) pack, Ziploc pack. When a medical pill is out of its pack, it's almost impossible task to recognize it without the label of the pack. It's also possible that in any way the label can be damaged. In many cases someone has to take two or more pill at a time, so it is hard to find the correct pill from several cases where the label is damaged or all of them are out of designated containers. Older people will face even more difficulty identifying pills out of its container.

For a visually impaired person this problem is even worse. As the printed labels are no help to them. Though sometimes they can identify some pills by touching the embossed imprints on the surface of a pill. However, this works with some tablet form of pills, as in the capsule form of pill it is not possible to emboss any text. Moreover, sometimes the pills may not even have any imprint on them. However, sometimes there may be some imprint on the surface of a capsule but that is not sensible by touching. All of these issues may leads to wrong medication which may cause an unwanted serious health hazard.

So for wrong medication, lack of identification of pill is one of the primary reason. A common problem in the emergency room is often patients come with medication errors that is either they have taken the wrong medication or the dose of medication is wrong. It can cause something disasters consequences. After an earnest try of doctors and pharmacists about correct medication mistakes may happen due to pill identification mistakes.

At the current time there exist thousands of different types of medical pills as well as their number of varieties in size, shape, color and texture are enormous. Recognition of any individual pill becomes worse when there is no standard for identification of medical pills either local or international arena. Moreover, there is no standardization or guideline of pill shape, size, color or texture with respect to the type of medication it is. Though it is possible to identify some pills from the printed label (or text) which is available on the surface, however only trained pharmacists can do this. Most of the time people doesn't have any idea what is in the printing or many of the cases there isn't any text or logo available at all. It is truly impractical to search and find information about pills by using catalog or any book. Though most of the company provides adequate documentation, but it is very tough and tedious process to link an arbitrary pill and the literature.

1.2 Scope and Objectives

In the context of computer vision we need a set of discernible veasual features, by those we can localize or identify some object from the pill image. Usually if a problem is difficult for human eye and brain to discriminate it also is difficult for computers to do so. In this case the problem is, in a pill image there is hardly any discriminating visual cue other than the geometric characteristics. It becomes harder when plenty of pills look alike or share exact same features values. For a medical pill the usual discriminating features are size, shape, color, textural pattern, text or logo printed on the surface. However, the warning is classical optical character recognition (OCR) technology doing its best to find and interpret text from image and it is far beyond to extract text from pills in image, because most of the time text or other labels are embossed on the surface of pill (in tablet form of pills). In the other case (capsule form of pills) the text printed in such a way that the rate of failure is very high with traditional OCR. Above this in most of the cases there is no text or discriminatory sign on the pill at all.

There are different ways the shape of an object can be recognized. It can be done by comparing each pixel value of the object in the image with another image stored in the computing machine that we call templet matching. It is also a favorable way that, after careful analysis find some metrics, extract their value and finally make a decision by comparing those value with some values from empirical study.

1.3 Achievements

In this paper we propose a method to provide an initial sorting of pill images based on proper measurement of their intrinsic geometric parameters values, their number of colors i.e. if the pill is of a specific color or it is a capsule of two colors and finally the text printed on the surface of that pill.

2 Prior Works on Pill Image Recognition

The amount of work on automated medicine recognition is insufficient as it is relatively a new concept. Young-Beom Lee et al. [8] developed an application that is able to automatically identify illicit drugs. M. Hagedoorn [2], and, R. C. Veltkamp and M. Hagedoorn [3] in their work provide a detailed survey about shape recognition techniques or algorithms. Such as tree pruning, generalized Hough transform, Fourier descriptor, statistics, wavelet transform, deformable templets, curvature scale space, relaxation labeling, neural network. A. Hartl et al. [4], and A. Hartl and C. Arth [5] in their work tried to recognize medical pills using mobile device. However, their proposed method limited to find shape and color of the pills.

The usual approaches to find features in image for object recognition are the use of, Scale Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF), Pyramid Histogram Of visual Words (PHOW) [6,7, 9-11].

From a set of reference images first interest points are extracted and stored in a dataset. In a new image each features of the recognized object is compared with the stored features individually by calculating the Euclidean distance between their feature vectors. A subset of interest point that best goes with the object and its orientation, location, scale are kept. SIFT [6] feature descriptor is scale and rotation invariant and also partially affine distortion and illumination changes invariant.

The other popular local feature descriptor is SURF [7]. The standard implementation of SURF is several times faster than SIFT. SURF follows the same principles and steps as SIFT but the details of each steps are different. SURF is somewhat inspired by SIFT. Using the multi-resolution pyramid technique, image is transformed into coordinate's space where each points are SURF descriptor. SURF detector is also scale invariant.

Above mentioned feature descriptors performs excellent with objects that contains variations within the object under consideration. However, medical pill doesn't contain enough texture or corner features thus SIFT or SURF cannot extract enough discriminating features to separate pills from each other. The most dominant features for pills are size, shape and color and finally the text printed on them.

International Drug Enforcement Administration like U.S. Drug Enforcement Administration's Office of Forensic Sciences (DEA) has been publishing Microgram Bulletin and Microgram Journal to assist and serve forensic scientists for detection and analyses of drug-related substances. Food and Drug Administration (FDA) regulation requires that every prescription pill or capsule sold in the market must have unique-look for easy identification in terms of size, shape, colour and imprint. Among those features, imprint is an indented or printed mark on a pill, tablet or capsule. Imprints can be a symbol, text, set of digits, or any combination of them. Drug makers use imprints to identify the chemical substance and their quantity in each pill and put special imprint on the pill for advertisement purposes. When a new illicit psychoactive substance is first detected in the Figure 1. Example images of illicit drug pills or tablets. Market, its information is recorded in the law enforcement databases. This information includes chemical and physical description, where the physical description includes size, colour, and imprints of illicit drug pills.

Law enforcement units would like to automatically extract the information about the drug pill (i.e., type of pill, manufacturing location and the manufacturer) from the imprints by matching them with known patterns in their databases. Unlike legal drug pills that can be identified from the information provided by pharmaceutical companies and enrolled in FDA database, matching imprints with the patterns of previously seized pills is the most effective way to identify illicit drug pills. Therefore, it is important to develop an image based matching tool to automatically identify illicit drug pills based on their imprint, size, shape, colour, etc. There are a few web sites that provide keyword-based legal drug pill identification tools. The keywords are based on the size, shape and colour of the pill (e.g., round, diamond, red, etc.), but do not utilize the imprint. For retrieving information from the surface of the pill images, text recognition process is highly recommended. Even though, OCR technologies is mostly restricted to finding texts that are printed on clean backgrounds, and cannot identify texts printed on shaded or textured backgrounds.

There have been only a few studies on automatic pill identification. Gerents et al used aspect ratio, circularity, moment invariants and gradients. Only 75 duplicate images taken from three different pills were used for the image retrieval experiment out of a drug tablet database containing several images. To develop a successful automatic pill matching system, it is important to generalize the variations in the appearance of the pills, due for instance to changes in viewpoint, illumination or occlusion. We utilize the edge information to characterize the imprint patterns on the drug pill images. Edge detection is one of the most useful image processing methods to extract object features. It serves to simplify image analysis by drastically reducing the amount of data to be processed, while at the same time preserving useful structural information about object shape and boundary. Given the processed edge image, invariant moment features proposed by Hu and grid intensity methods can be used to generate feature vectors from the edge image. With the extracted feature vectors, we use the L2-norm to compute the similarity between two pill images. The proposed method is evaluated on a database with thousands of pill images. To the best of our knowledge, this is the first national effort in developing an automatic pill identification based on imprints and texture using a large scale real pill image database.

3 Proposed Method

As the only features that are available for discrimination are shape, size and color the proposed method intends to extract values of intrinsic geometric property of medical pills from image by using image processing techniques for a set of training image and then determine a threshold of those parameters and use it to differentiate pills. To determine the values of the intrinsic geometric features we need a data set. For this experiment we used NLM's (U.S. National Library of Medicine) RxIMAGE dataset from their NLM Pill Image Recognition Project [1].

The available features of pill are mainly their shape, size, color and significant engraved or imprinted text on its surface. Even though there has been works on feature finding in an image for object recognition using Scale Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF), Pyramid Histogram of visual Words (PHOW) [6, 7, 9-11]. These feature descriptors are useful for object with variations. And medicinal pills doesn't have that many variant features to use the above mentioned algorithms to apply. Hagedoorn [2], and Veltkamp and Hagedoorn [3] in their work provide a detailed survey about shape recognition techniques or algorithms. Such as tree pruning, Hough transform, Fourier descriptor, statistics, wavelength transform, deformable templets, curvature scale space, relaxation labeling, and neural network.

The proposed method is to use the prime features like shape, size and the printed text. This method intends to extract values of intrinsic geometric property of medical pills from image by using image processing techniques for a set of training image, then determine a threshold of those parameters and use it to differentiate pills. After that using the text segmentation technique and text detection method to find out the text regions and the containing texts printed on the pills.

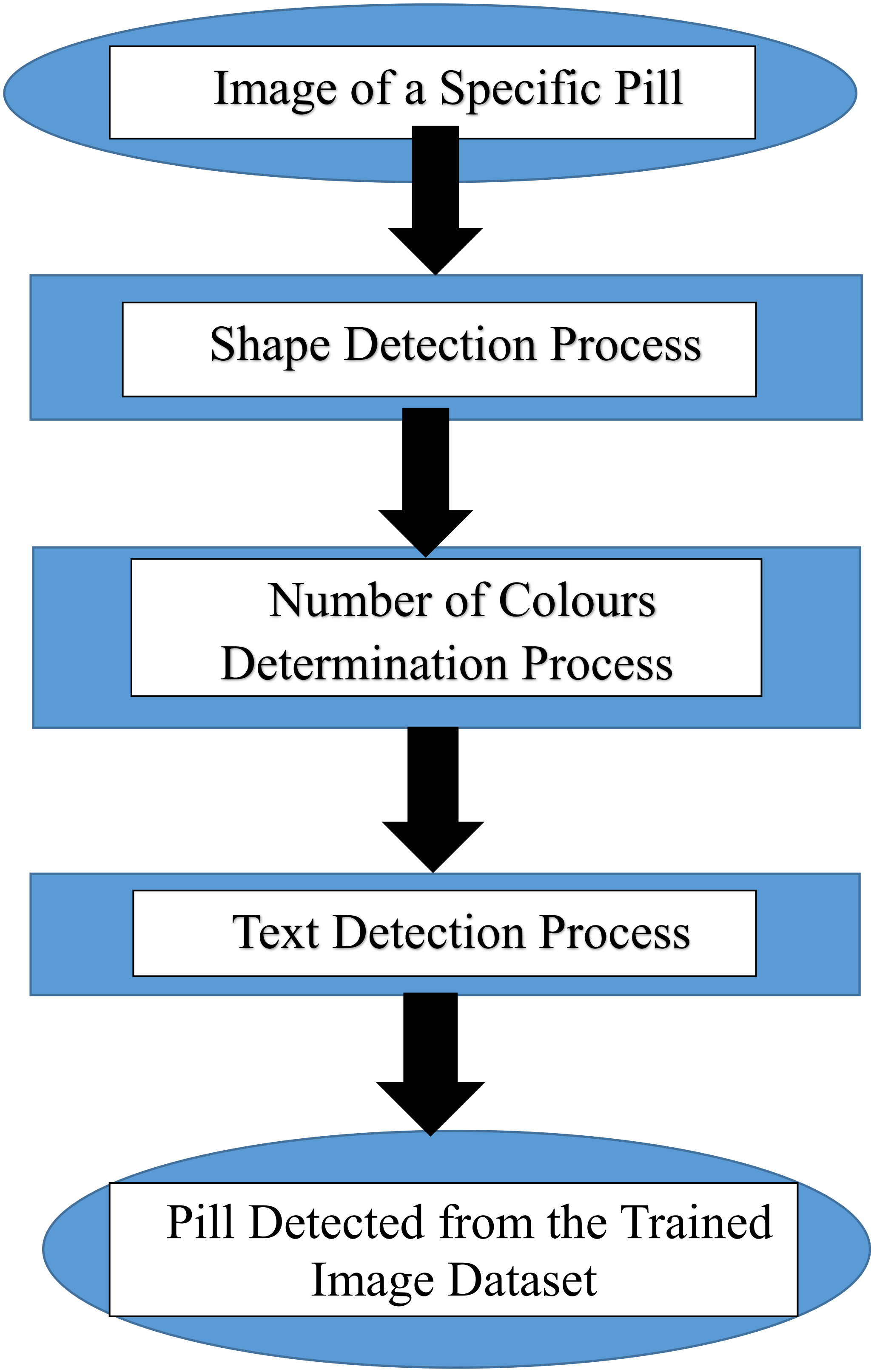


Figure 1. Overview of the Proposed Method

4 The NLM Dataset

On January 2016, NLM open a challenge under a federal notice “Pill Image Recognition Challenge [1]”. NLM’s purpose of this Challenge is to find a set of algorithm and software that can rank an input pill according to the similarity to images of unknown prescription pills to known prescription pill images in the NLM RxIMAGE dataset. There is two image-set one for consumer quality image and other for reference image and a ground truth table.

The RxIMAGE dataset contains 5000 images of 1000 different pills in the consumer quality image-set. To mimic the quality of image in consumer level the photos were taken with digital cameras built in to mobile devices.

There are 2000 (one for the back and other for the front) images of 1000 different pills in the reference image-set in the RxIMAGE dataset. This were taken with very high resolution digital camera and under proper lighting condition.

One CSV format ground truth table is also provided with the dataset. There are two column headers named ref_images (name of pills from reference image set) and cons_images (name of pills from consumer image set). A separate row is allocated for each reference, consumer-quality image pair in the table.

Two-column is provided in table with column headers ref_images and cons_images. In the table for every row, the first column named after a reference image, and the second column named after a consumer-quality image corresponding to the reference image. A separate row is there in the table for each (reference image, consumer-quality image) pair, even when match found of multiple reference and consumer-quality images are all photos of the same pill. In Fig 1 we are providing sample images form the RxIMAGE dataset.

Figure 2. Sample images from the reference set of the NLM RxIMAGE dataset

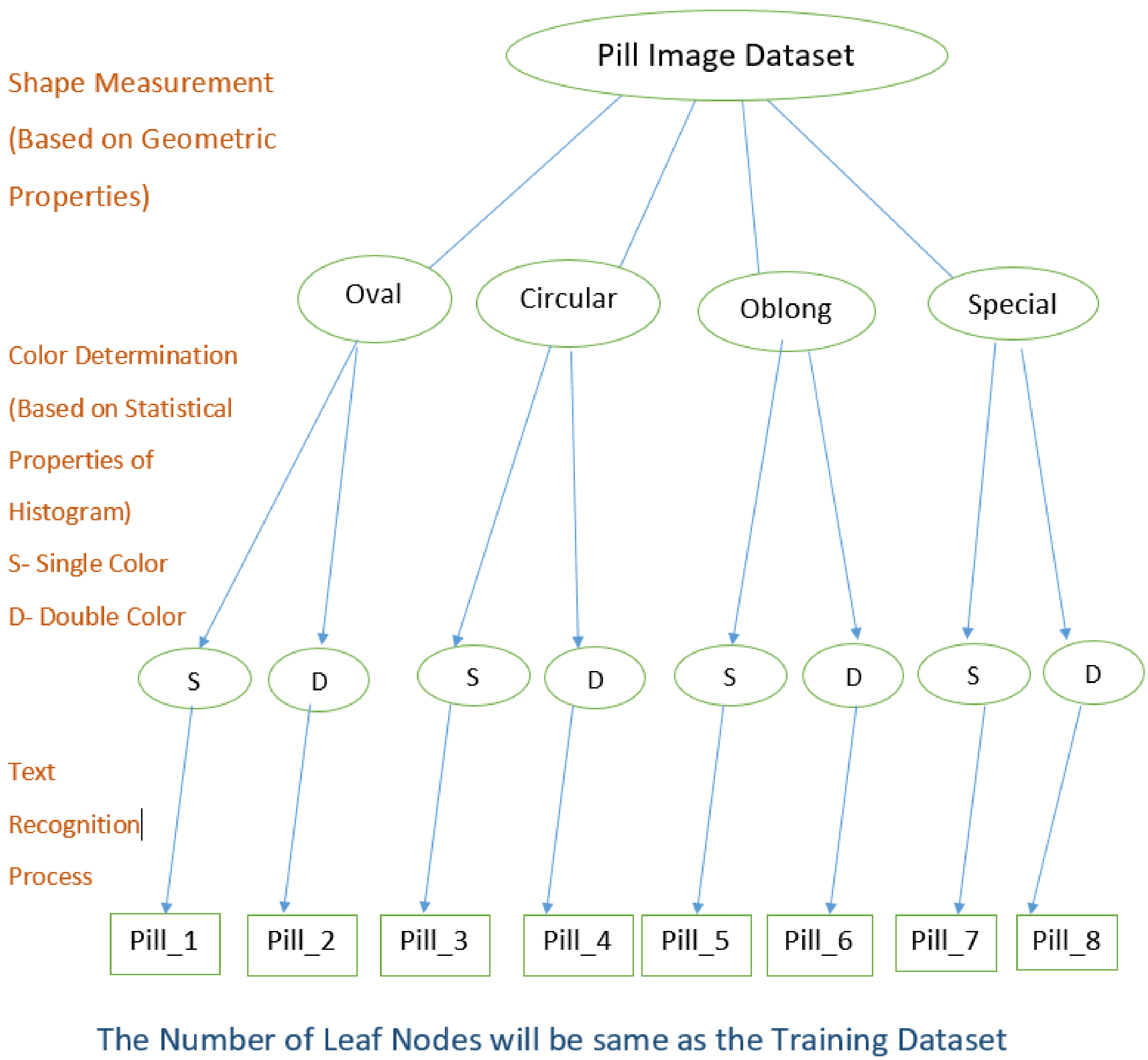


Figure 3. Classification of the Pill Image Dataset

5 Image Processing Steps for Shape Detection

Ones the images are collected next task is to measure the geometric properties of the main object (pill) in an image. However, before we can do that series of image processing steps are required to acquire a shape of the pill in the image. Consider the image in Fig 4. It is not possible to measure the intrinsic parameters values.

Figure 4. Color pill image from the reference database

Converting RGB image to Grayscale image: Every reference image in the database is a 1600 X 2400 X 3 matrix. That is each image is 1600 X 2400 pixel and in RGB format. RGB is additive colour model where Red, Green and Blue colour channel. For farther processing the colour image is converted into a 1600 X 2400 8-bit grayscale image. Fig. 5 illustrates the grayscale form of sample RGB image.

Figure 5. 8-bit Gray pill image for the image from Figure 4

Edge Detection: Ones the images are converted into a 8-bit gray image Canny edge detection algorithm is applied on the gray image to acquire edge in the image**.** To detect boundary most of the time we just need the lower level details. To find the shape most of the information can be extracted from its boundary. After application of canny edge detector, the Fig. 6 is produced.

Figure 6. Detected Edges of the pill image for the image from Figure 5

Closing Morphological Operation: To segment out the pill from image it is necessary to find the boundary of pill. Morphological closing operation is used with disk shape structuring element to connect any edge pixel that is missed in the Canny edge detector during edge filtering process. Which makes a robust continuous boundary (Fig. 7).

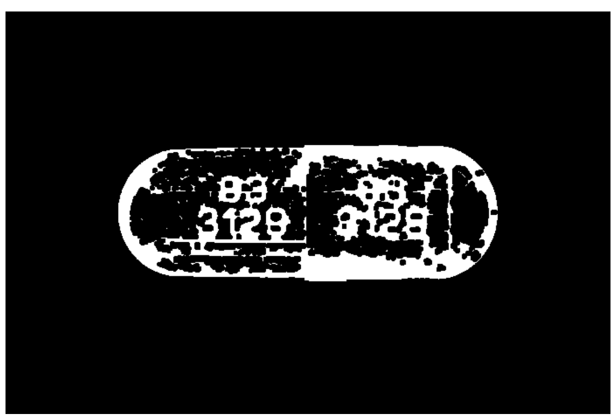


Figure 7. Edge image of Fig 6 after application of closing morphological operator

Fill Morphological Operation: Note that the image in Fig 8 contains many smaller components, however we are interested in a single object representing the pill. Thus we apply fill morphological operator to finally acquire a single object. Fig 8 shows the final product.

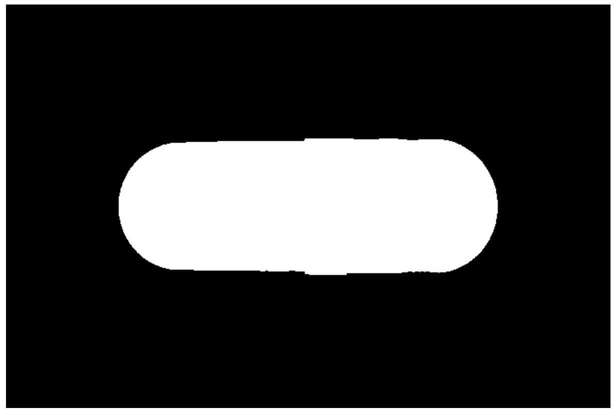


Figure 8. Image after application of fill morphological operation on the image of Fig 7

Figure 9. Segmented pill region from pill image acquired from Fig 8

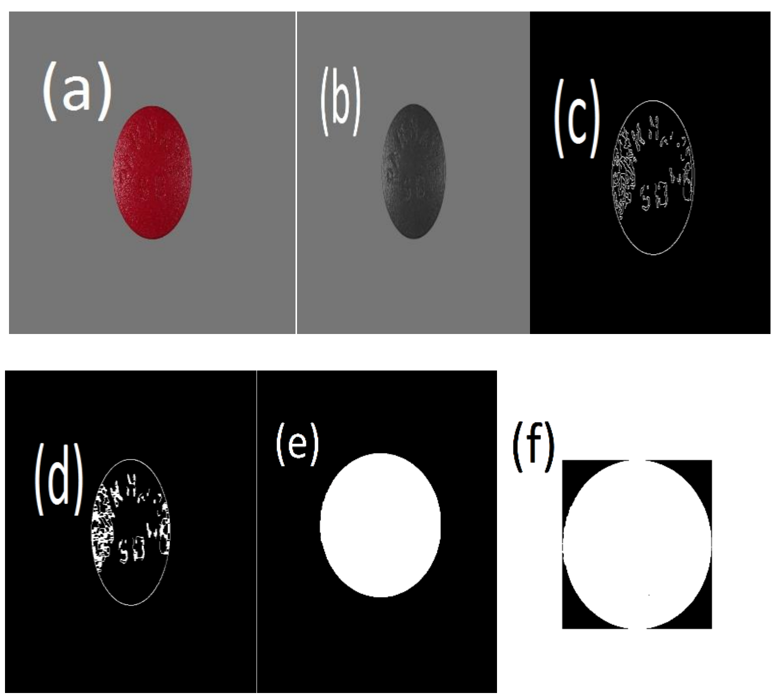


Figure 10. Shape detection for circular pill

Detecting the Pill in the Image: Now apply boundary enclose on the whole image the pill will be detected in the image. Fig 9 shows the final cropped image containing the pill only. The final image dimensions are the dimensions of the pill itself.

5.1 Metric for Shape Measurement

Now that we have the pill image in more suitable form for measuring geometric parameter values. Depending upon the types of measurements used in image processing or analysis it is possible to split and categorize any dataset in various way. One can categorize the measurement types based on scale. Sometimes the largest scale measurement referred as form that reflects geometric dimensions of the object. Roundness, which can express the radius of curvature of the object corners in the next smaller scale measurement. The types of measurement can also be categorized based on the assumption level and the degree to which the results are calculated. In this case linear results such as area, perimeter can be calculated from the pixel map of the image. By these, other results such as spherical equivalent volume and the circular equivalent diameter are calculated.

Finally, using several relations and ratios of the mentioned factors, metrics such as the aspect ratio, circularity, convexity, solidity, spherical equivalent volume, eccentricity, extent, elongation, convex hull area etc. may be calculated.

To separate the RxIMAGE dataset we used Eccentricity, Extent and Narrowness (using aspect ratio) to separate the dataset into four categories. These are, Circular, Oval, Oblong, and Special. Please refer to Fig 2, where we have provided samples of each type.

Eccentricity: The ratio of the distance between two focal of the ellipse and its major axis length is known as eccentricity. Fig. 8 illustrates, how eccentricity be a measure of shape. Its value is between 0 and 1. The shape with 0 (zero) eccentricity is actually a circle and eccentricity 1 (means) it’s a line or line segment. It’s a rotation and scale invariant property.

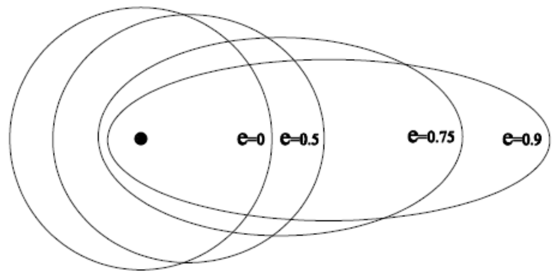


Figure 11. A perfect circle has an eccentricity of 0, while an oval or ellipse has 1

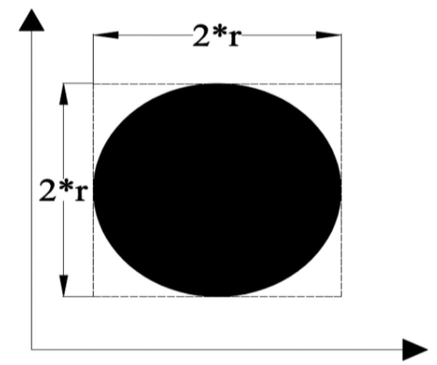


Figure 12. Dimension of bounding box for a circle.

Extent: It is the ratio of the area of the region to the total area of its bounding box (bounding box: the smallest rectangle containing the region) Fig. 11 illustrate and Eq. 1 gives the formula to measure extent of an image object. Though it scale invariant, however not rotation invariant. Thus we have to rotate the image according to its major axis to measure the correct value of extent. In theory if the extent is 0.7854 it's a circular shaped object.

$$Extent = \frac{Object\ Area}{Bounding\ Box\ Area} \quad (1)$$

Narrowness: It is defined as the absolute difference of aspect ratio and inverse of aspect ratio. The ratio between major axis and minor axis length is the measure aspect ratio (Eq. 2). If the narrowness is 0 (zero) the shale is circular. A higher value of narrowness (Eq. 3) will means the object is narrower.

$$aspectRatio = \frac{majorAxisLength}{minorAxisLength} \quad (2)$$

$$narrowness = abs\left(aspectRatio - \frac{1}{aspectRatio}\right) \quad (3)$$

6 Image Processing Steps for Finding number of Colors

Another most distinctive feature of pills are their color. With their particular color the pills can be identified more easily. The pills have various colors and shades. Some are of one individual color, others are of two colors. Basically the capsules have two colors which is another significant feature just to inform the blind people that the medicine is a capsule containing 2 colors helping them to identify the pill better.

The pills are of various colors. The capsules mostly consists of two colors. Whereas the tablets consists only one color. For determining how many colors are there in the pill region, the classification method that has been used is Decision Tree. Decision Tree Classifier is a simple and widely used classification technique. It applies a straightforward idea to solve the classification problem. Decision Tree Classifier poses a series of carefully crafted questions about the attributes of the test record. Each time it receive an answer, a follow-up question is asked until a conclusion about the class label of the record is reached.

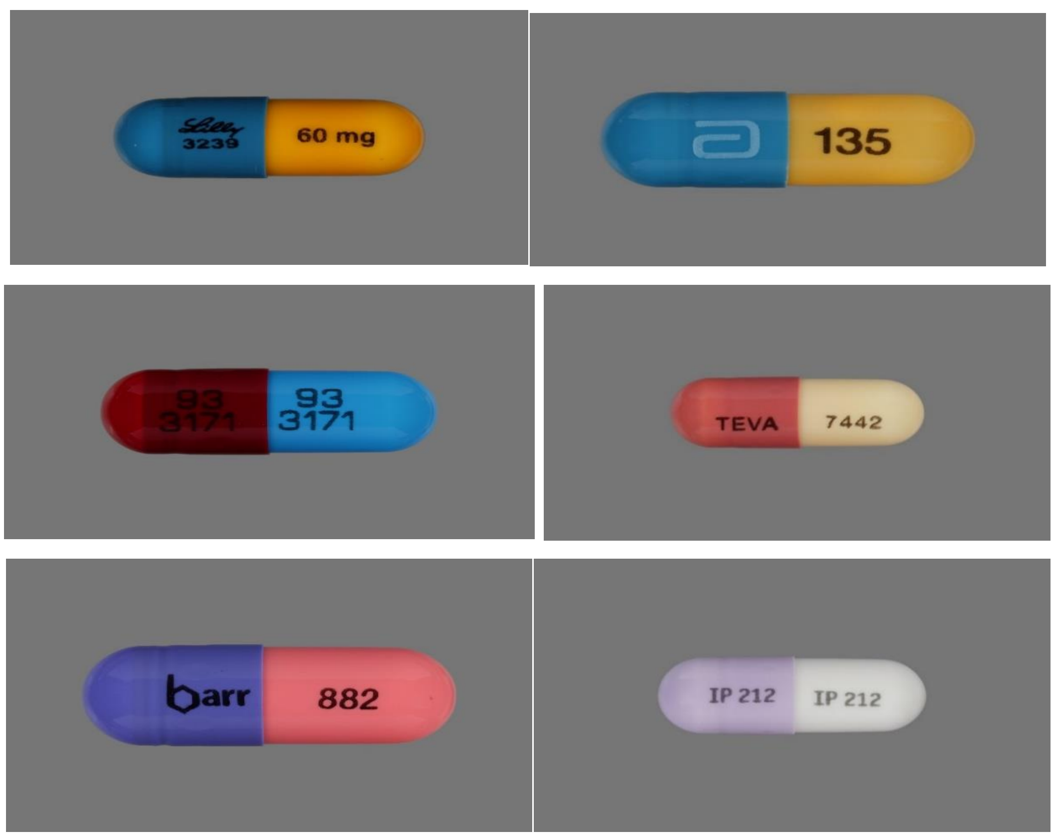


Figure 13. Two colored pills

Figure 14. Single colored pills

6.1 Segmenting the pill image

The first fact that have been considered on the Color based Classification is the number of colors of the pill. As shown above the pills have at most two colors. So the first classification would be on the number of colors present. The proposed method is to first use clustering algorithm to find out the color numbers. The RGB images first go through the morphological operations to segment the pills from the background. The images are segmented close to their shape and based on that three images are extracted from the original image. The extracted images are, morphological segment image (the one extracting the shape of the pill from the background) ;Color segment Image (the RGB image with only the pill in it with a bit of a background present due to the shape variance with the bounding boxes) and gray segment image (the grayscale version of the RGB image). These segmented images are substantially advantageous for finding both the shape and color of the pill images.

Now that the segmented images have been computed, the next step is to section the pill region using the morphological images and the grayscale images.

6.2 Generating Histogram for finding the number of Peak present

Now that the pill region is known, the next step is to calculate the histogram of the pill region. As the pill region only consists of the pixel values of the pill, the histogram calculated only shows the histogram of the pixels present in the pill region. The histogram is now filtered with predefined 2-D filters for fine tuning it. Later on the edge is calculated from it gives a clear picture of the number of peaks present in the histogram. With the number of peaks available on the filtered histogram, the number of colors existing in the pill is obtained.

Later on some statistical measurements are calculated from the filtered picture for probability distribution. Measures like Kurtosis and Skewness helps to find out the symmetry of the distributed data.

Skewness measures the relative length of the two peaks of the histogram and **Kurtosis** measures the combined size of the two peaks of the histograms. As the histogram can only give a general idea of the shape, kurtosis and skewness gives more precise evaluation. The departure of the horizontal symmetry presents the value and direction of skew and kurtosis presents how tall and sharp the central peak is. The values of kurtosis and skewness represents a huge amount of information regarding the histogram. The shade of the pill can be known by these values. For that, we need to work on the bimodal distribution, a continuous probability distribution with two different modes.

Sarle's bimodality coefficient b is

B = (γ^2 +1)/ κ

Where γ is the skewness and κ is the kurtosis. The kurtosis is here defined to be the standardized fourth moment around the mean. The value of b lies between 0 and 1. The logic behind this coefficient is that a bimodal distribution will have very low kurtosis, an asymmetric character, or both all of which increase this coefficient.

The formula for a finite sample is

$$b = \frac{g^2 + 1}{k + \frac{3(n-1)^2}{(n-2)(n-3)}}$$

Where n is the number of items in the sample, g is the sample skewness and k is the sample excess kurtosis.

The value of b for the uniform distribution is 5/9. This is also its value for the exponential distribution. Values greater than 5/9 may indicate a bimodal or multimodal distribution. The maximum value (1.0) is reached only by a Bernoulli distribution with only two distinct values or the sum of two different Dirac delta functions (a bi-delta distribution).

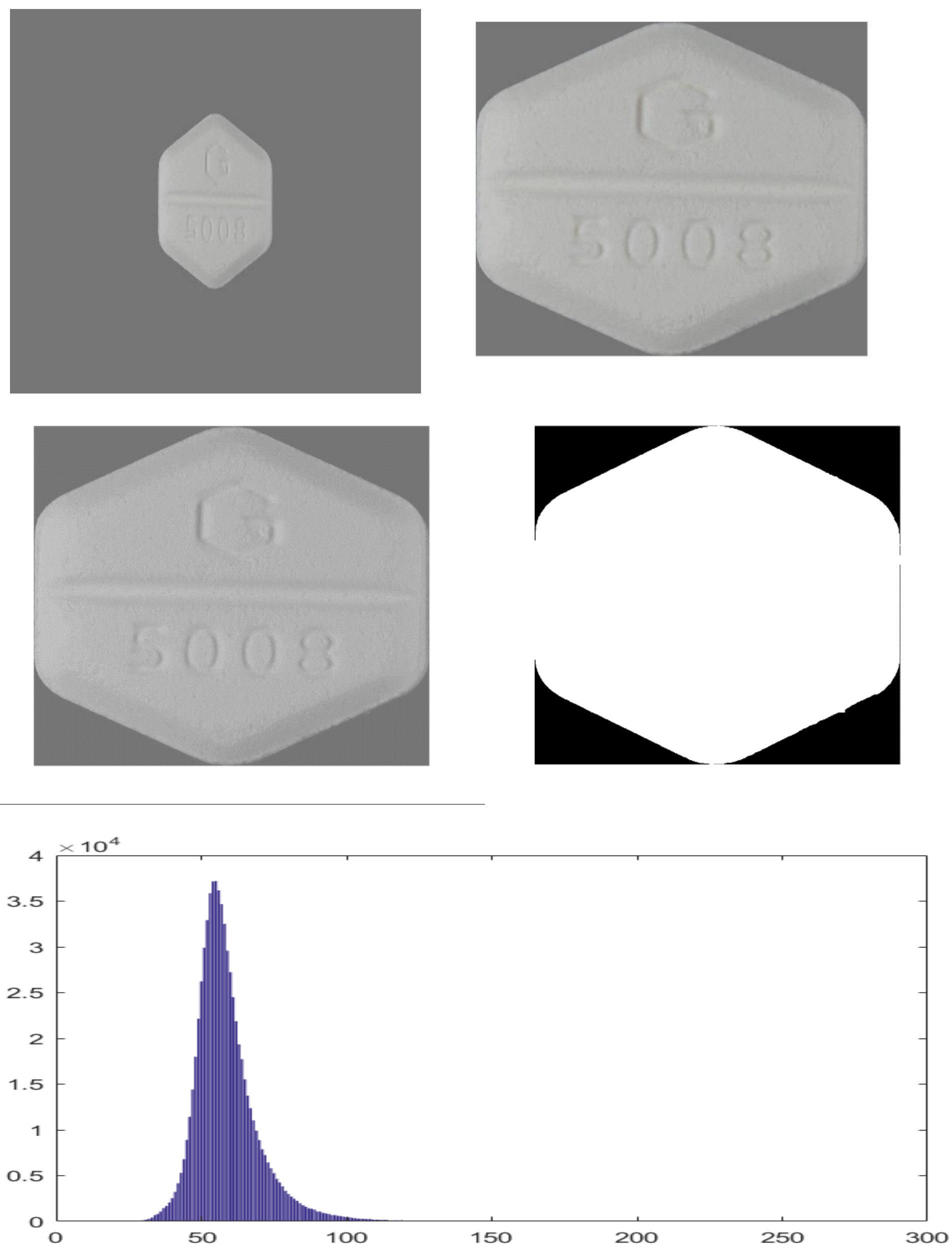


Figure 15. Detecting Number of colors with a single colored pill

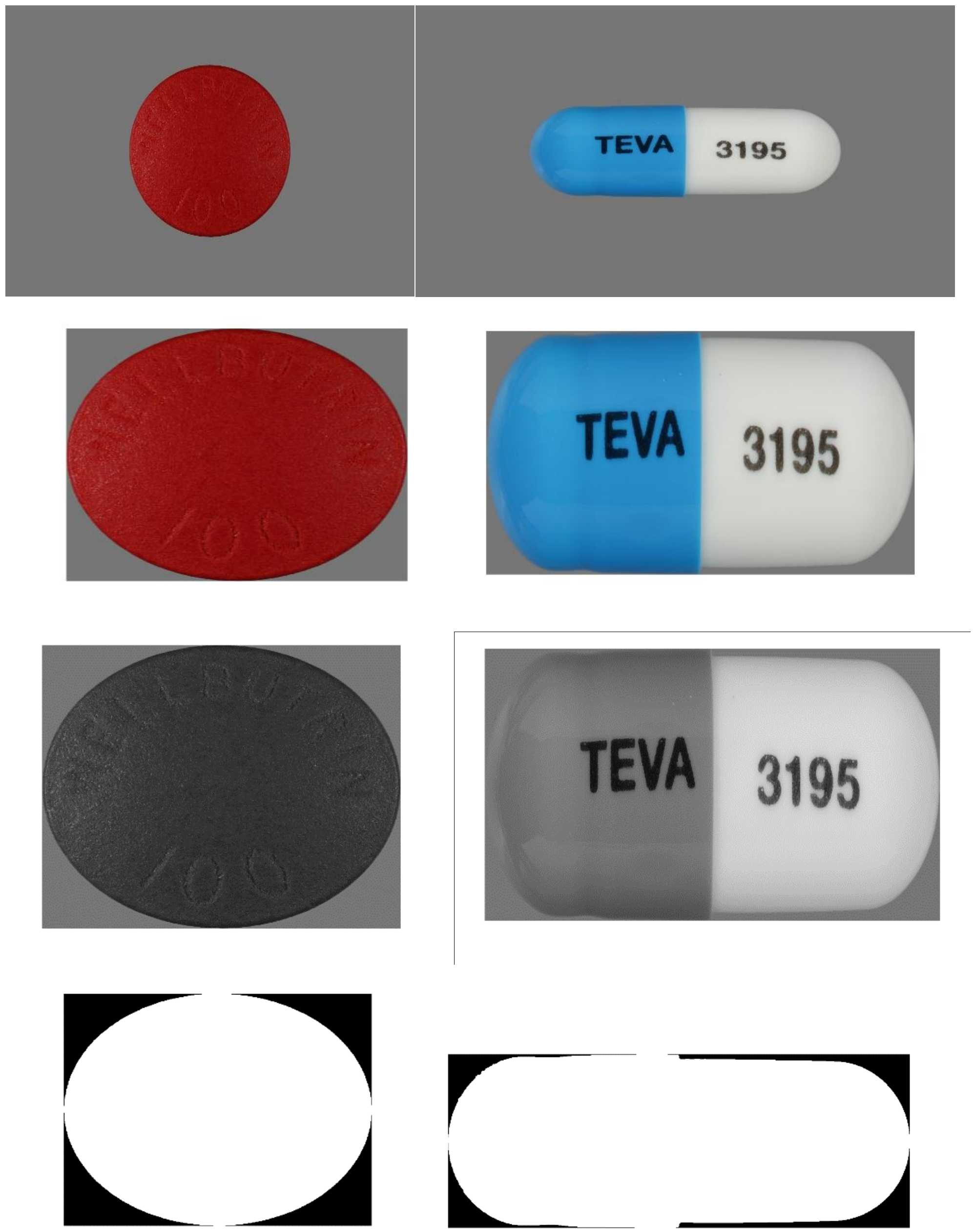


Figure 16. Segmenting the Pill image and detecting the pill region of Single colored pill (Left) and two colored pill (Right)

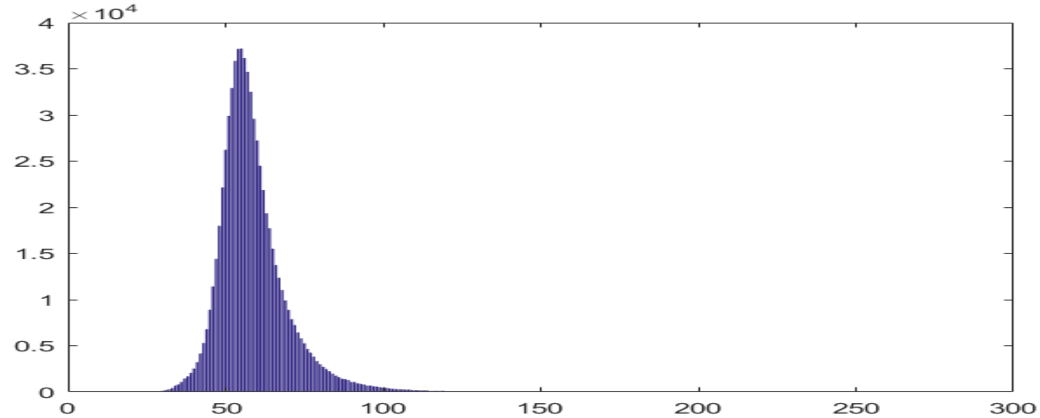


Figure 17. Histogram of the pill region of a single colored pill

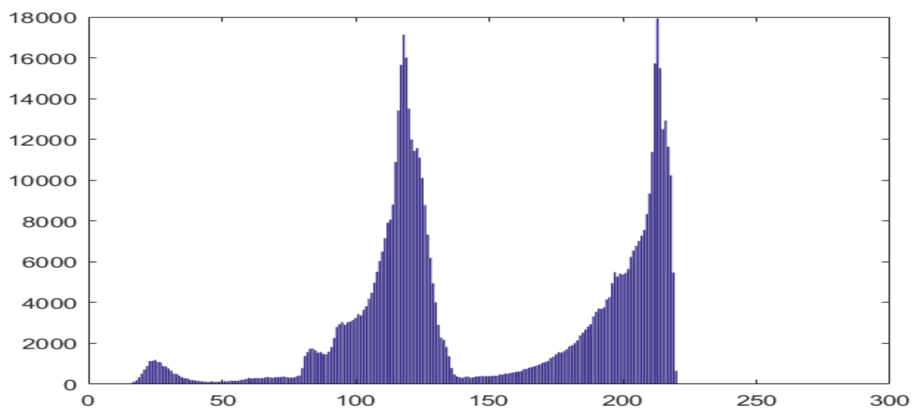


Figure 18. Histogram of the pill region of a two colored pill

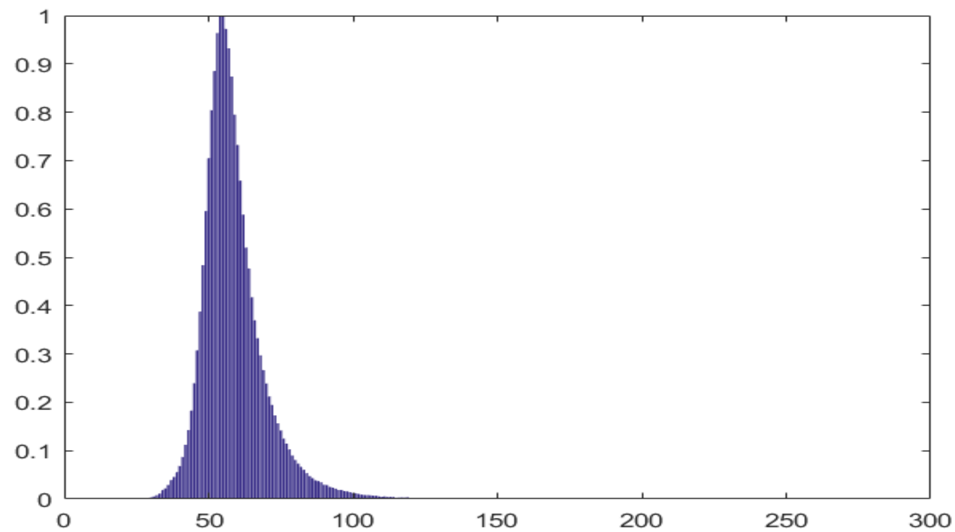


Figure 19. Filtered Histogram of the pill region of a single colored pill

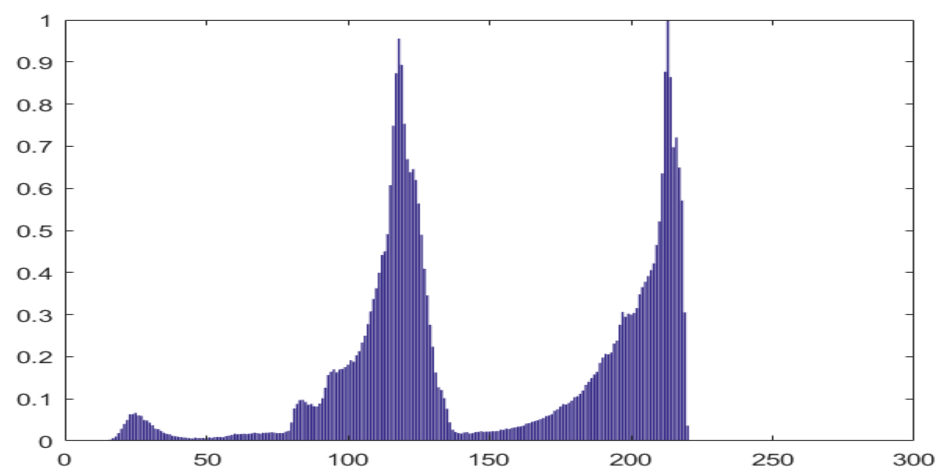


Figure 20. Filtered Histogram of the pill region of a two colored pill

There are times when the histogram looks a bit confusing by showing 3 peaks. But the dataset does not contain any pill of three colors. It happens because of the different shades of the colors present.

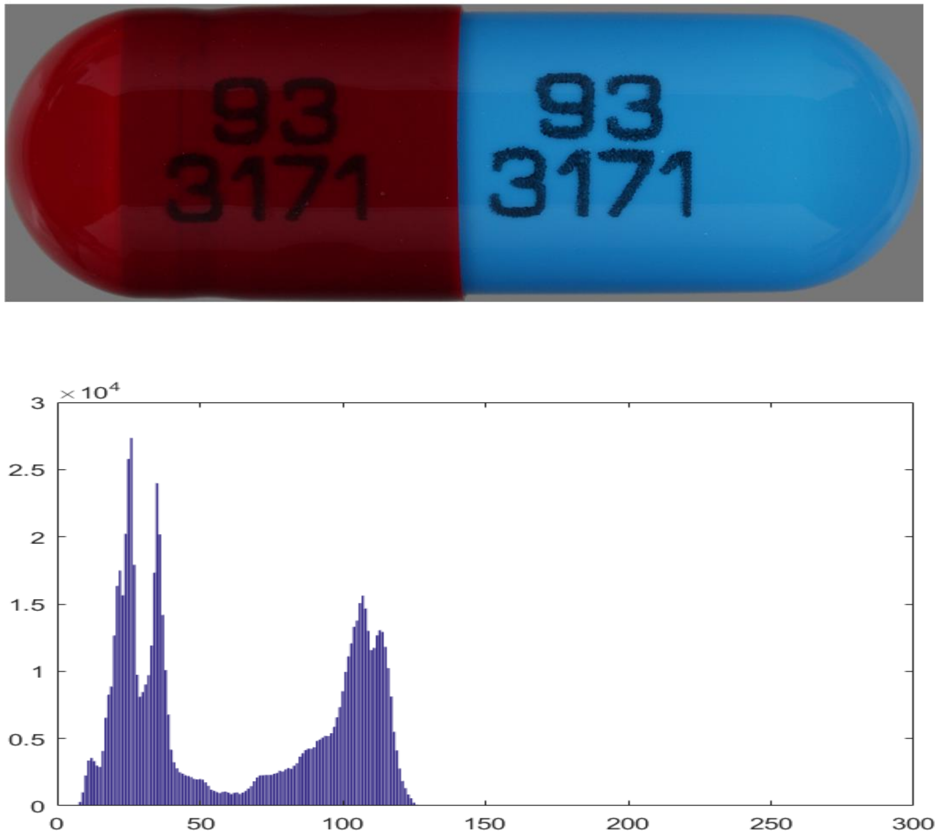


Figure 21. Detecting Number of colors for shaded Pills

In Figure 20, the histogram shows three peaks as the red colored side has two shades of it present on the pill. The left side is a bit lighter and the text part is darker, making the presence of two different shades of red and thus making three peaks in the histogram. There are some images where this kind of shade difference occurred.

7 Image Processing Steps for text recognition

All Pills have a texture which as in itself is a significant feature. The ones with a notable features like colored texts, engraved texts or some engraved lines makes the pills more easier to recognize and also helps the human eye to easily detect the pills. But for the blind people the colored texts are impossible to recognize. But if the texts on the pills can be recognized or readable, can be of great help for the blind ones. For finding the text on the pill, there are a series of image processing steps required. At first, the segmentation of the text from the image is done followed by the optical character recognition (OCR).

7.1 Image Processing steps for printed texts on pill surface

7.1.1 Detect Possible Text Regions

For detecting the text region a feature detector has been used. Maximally Stable Extremal Regions (MSER) is a feature detector which works well for finding text

Figure 22. Original Image

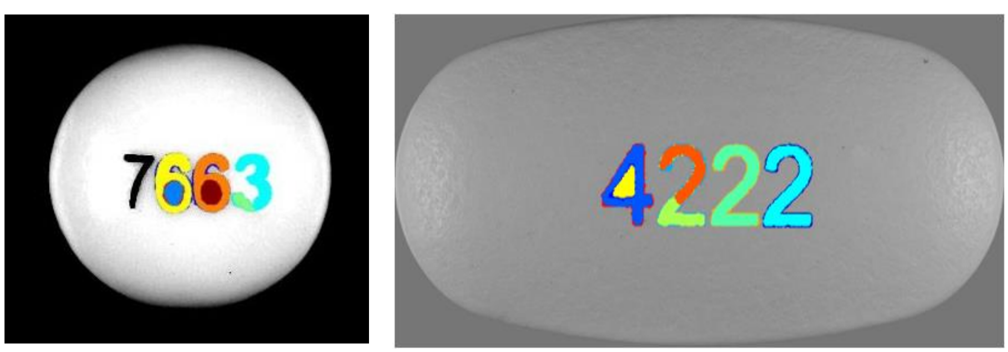


Figure 23. MSER region

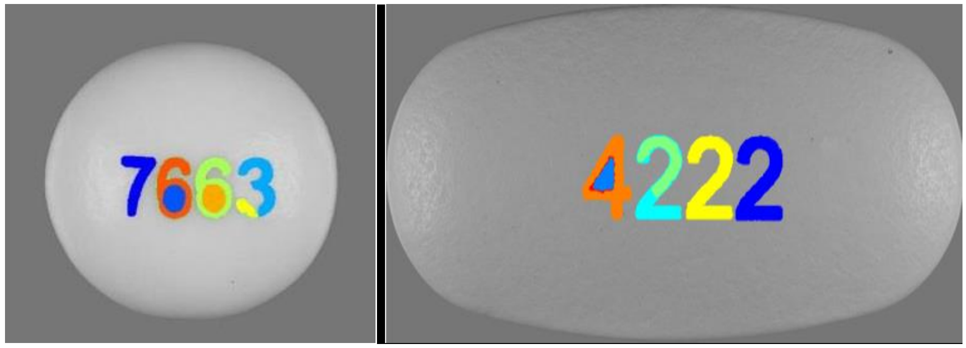


Figure 24. After removing Non-Text Regions Based on Basic Geometric Properties

regions [13] as the texts are always high contrast and stable color compared to the surrounding. It works on the Grayscale image and gives an output area of consistent color and high contrast points. (Figure 24).

7.1.2 Remove Non-Text Regions Based On fundamental Geometric Properties

The previous process grabs out the text regions along with some other stable regions like lines, curves or other mark and mistakes them as text. To remove them the non-text region has been filtered based on their geometric properties used in the previous shape detection method. This is the rule based approach. Another approach is to use Machine learning. There are also solutions which includes both [16].

Among the several geometric properties, there are some which are better for detecting text and non-text region [14, 15]. The geometric properties used here are eccentricity, solidity, Euler number, aspect ratio, extent. (Figure 25).

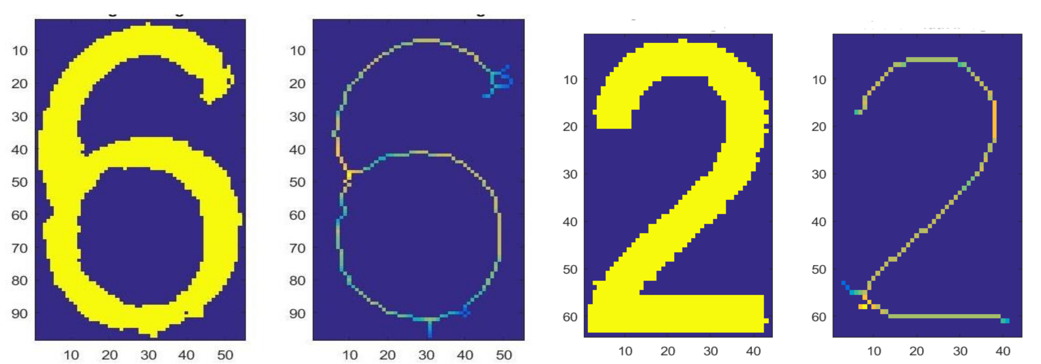


Figure 25 Region Image & Mark Width Image

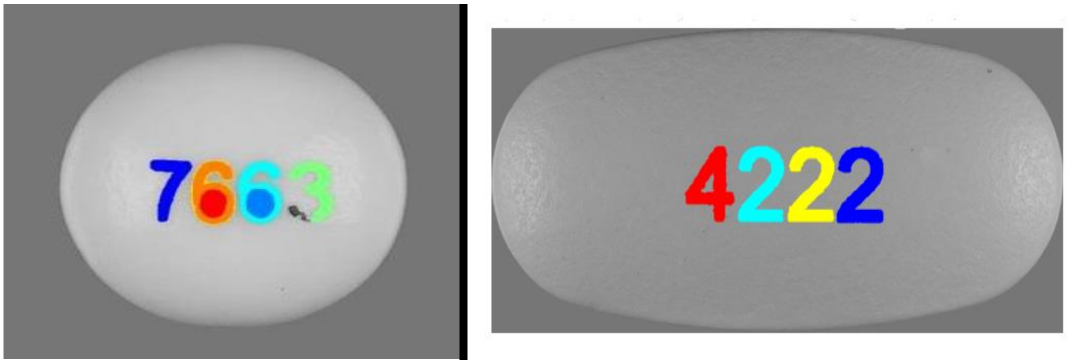


Figure 26. After removing non-text regions based on mark width variation

7.1.3 Remove Non-Text Regions Based On Mark Width Variation

Another technique of removing the non-text region is to check the width of the mark. This is the width of the lines, curves of the character. The texts in the text region have same mark width and the non-text regions have different width causing more width variation. After the calculation it is noticeable that the mark width of a text region is same meaning the region contains text. Then a threshold has been applied for removing non-text region for varying font styles. (Figure 26).

7.1.4 Merge Text Regions for Final Detection

Till now the individual text characters has been detected. Now we need to merge the characters for finding a meaningful text which is easily understandable. For using the optical character recognition, the characters should be merged to form a text line.

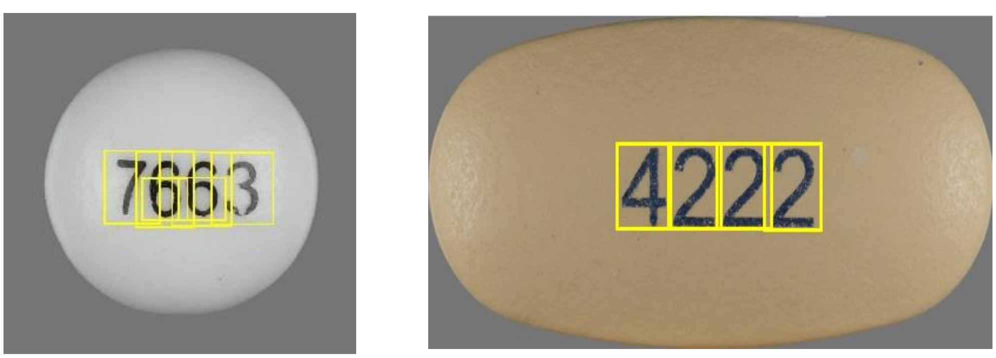


Figure 27. Merged bounding boxes text

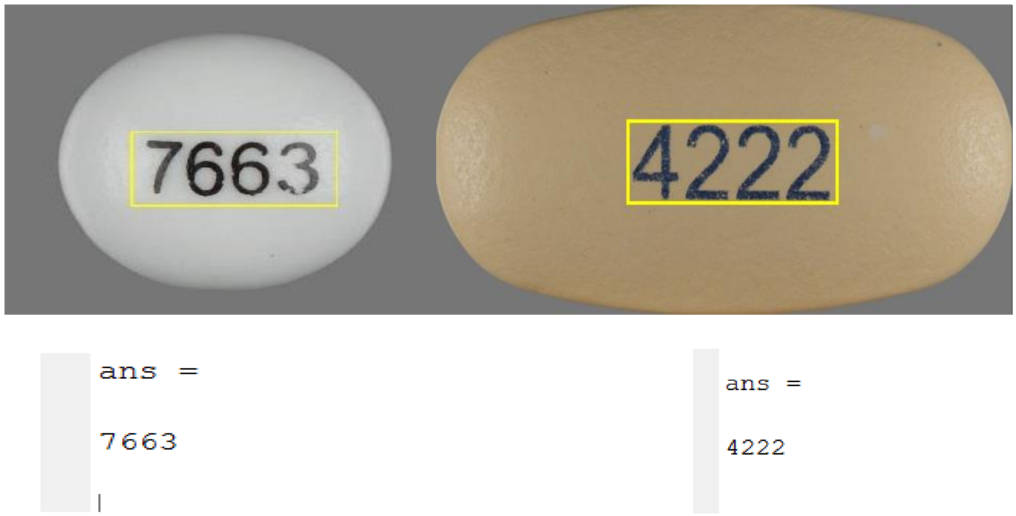


Figure 28. Detected Text

For that the text regions of the characters should be merged and find the connected regions and form bounding boxes. The connected regions should be found by stretching the bounding boxes that causes overlapping of the same text region forming a string of bounding boxes. After merging these bounding boxes single bounding box for individual text lines are formed. (Figure 27 & 28).

7.1.5 Recognize Detected Text Using OCR

Now that the text region is detected, OCR function can be used for recognizing the texts. It gives back the text on the surface of the pills. (Figure 8). Even though the OCR systems does not work well with the printed, shaded backgrounds, due to the preprocessing of the image, it becomes easier to recognize the texts.

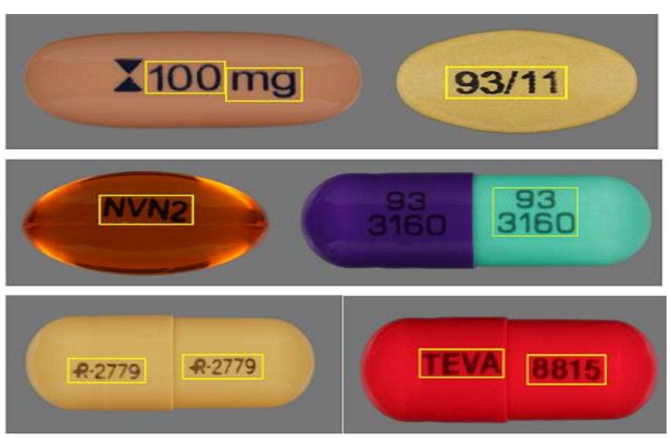


Figure 29. Detected Text from other pills

7.2 Image Processing Steps for Engraved Text on pill surface

The dataset that has been worked on has 2000 images out of which almost 55% of the pills have engraved texts on them. The texts are sometimes too hard to recognize. Due to light settings during the image capture, the text inscribed on the pills get lightened up and the whole image flushes out the possible text region and creates more noise. Due to those noises the preprocessing steps for the text detection becomes far tougher than it already is. So the text detection process includes feature extraction to reduce the noises as much as possible.

Figure 30. Sample images of engraved pills

7.2.1 Preprocessing the image

The Initial step of detecting text from the engraved pill is to preprocess the image as such that the engraved region looks like a text for the OCR. For that reason, the image has been reduced to the size of the pill region to make it easier for the system to work on.

❖ Segmenting Pill region from the background of the image

First of all, the morphologically segmented images of the pills are taken under consideration. Morphology is a broad set of image processing operations that process images based on shapes. Structuring Elements. An essential part of the morphological dilation and erosion operations is the structuring element used to probe the input image.

Morphological image processing is a collection of non-linear operations related to the shape or morphology of features in an image. Morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images. Morphological operations can also be applied to greyscale images such that their light transfer functions are unknown and therefore their absolute pixel values are of no or minor interest.

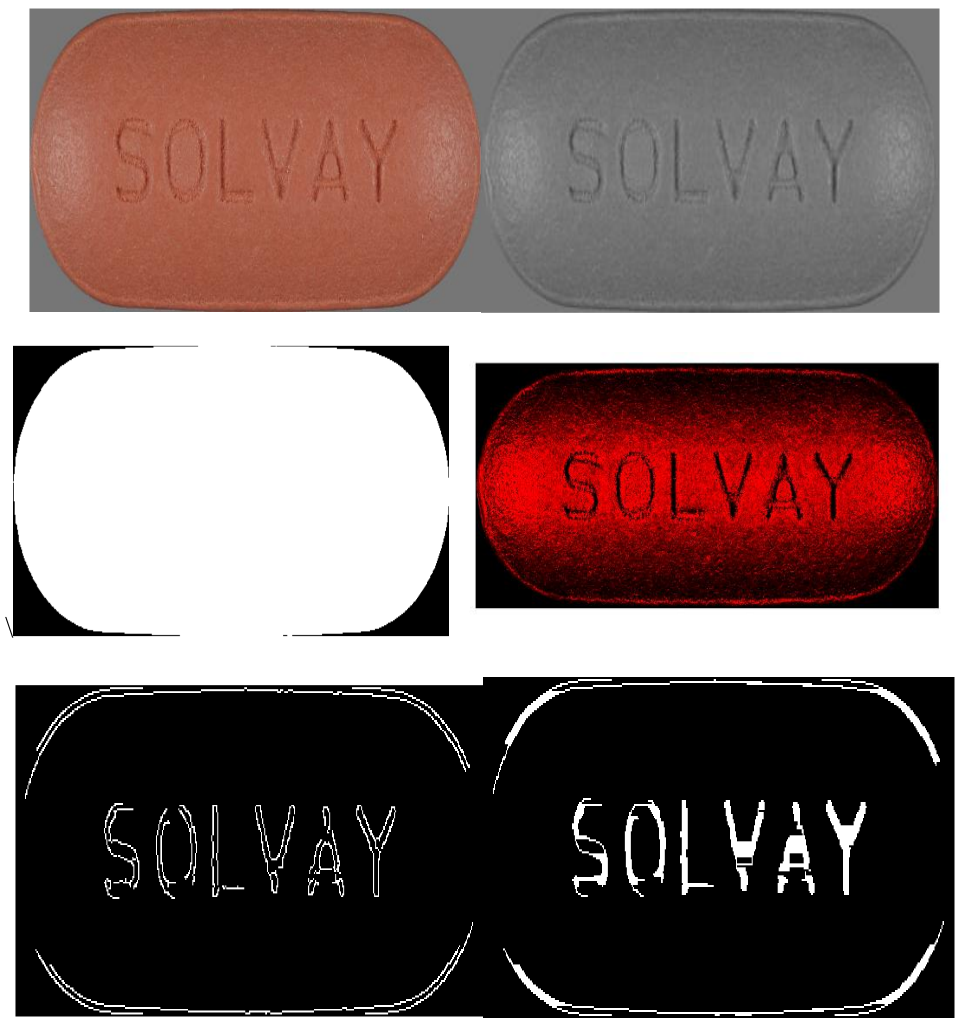


Figure 31. Text detection of engraved pill images

❖ Adjusting Contrast level of the segmented image

The images that has been obtained after segmenting the pill regions are adjusted in a contarst scale so that the engraved texts can be easily recognised. Contrast adjustment adjusts the contrast of an image by linearly scaling the pixel values between upper and lower limits. Pixel values that are above or below this range are saturated to the upper or lower limit value, respectively.

❖ Edge detection of the adjusted contrast image

Now that we got the adjusted contrast image, we have to send the images for edge detection process. Image result for edge detection.Edge detection is an image processing technique for finding the boundaries of objects within images. It works by detecting discontinuities in brightness. Edge detection is used for image segmentation and data extraction in areas such as image processing, computer vision, and machine vision.

❖ Refining the edge image for text detection

After the edge detection, we need to fine tune the image further by resizing the image, morphologically filling and providing structural element in it. A morphological operation on a binary image creates a new binary image in which the pixel has a non-zero value only if the test is successful at that location in the input image.

The structuring element is a small binary image, i.e. a small matrix of pixels, each with a value of zero or one:

1. The matrix dimensions specify the size of the structuring element.

2. The pattern of ones and zeros specifies the shape of the structuring element.

3. An origin of the structuring element is usually one of its pixels, although generally the origin can be outside the structuring element.

❖ Text detection using OCR

Now that the image is fine tuned, they are sent to OCR for text recognition. it is generally an "offline" process, which analyses a static document. Instead of merely using the shapes of glyphs and words, this technique is able to capture motions, such as the order in which segments are drawn, the direction, and the pattern of putting the pen down and lifting it. This additional information can make the end-to-end process more accurate.

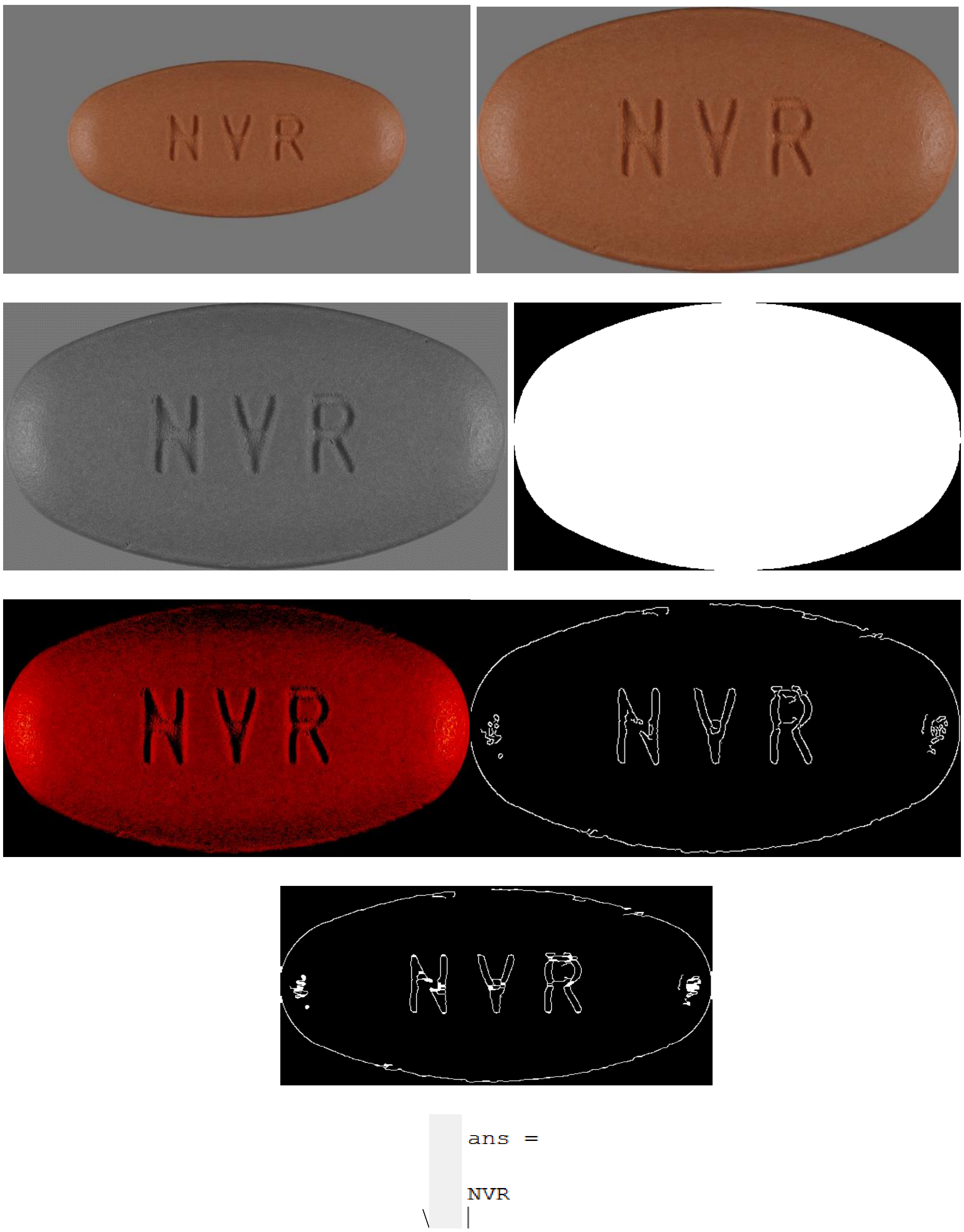


Figure 32. Text detection of engraved pill images

There are cases when the preprocessing is not done perfectly and the text recognition system fails to recognize the texts.

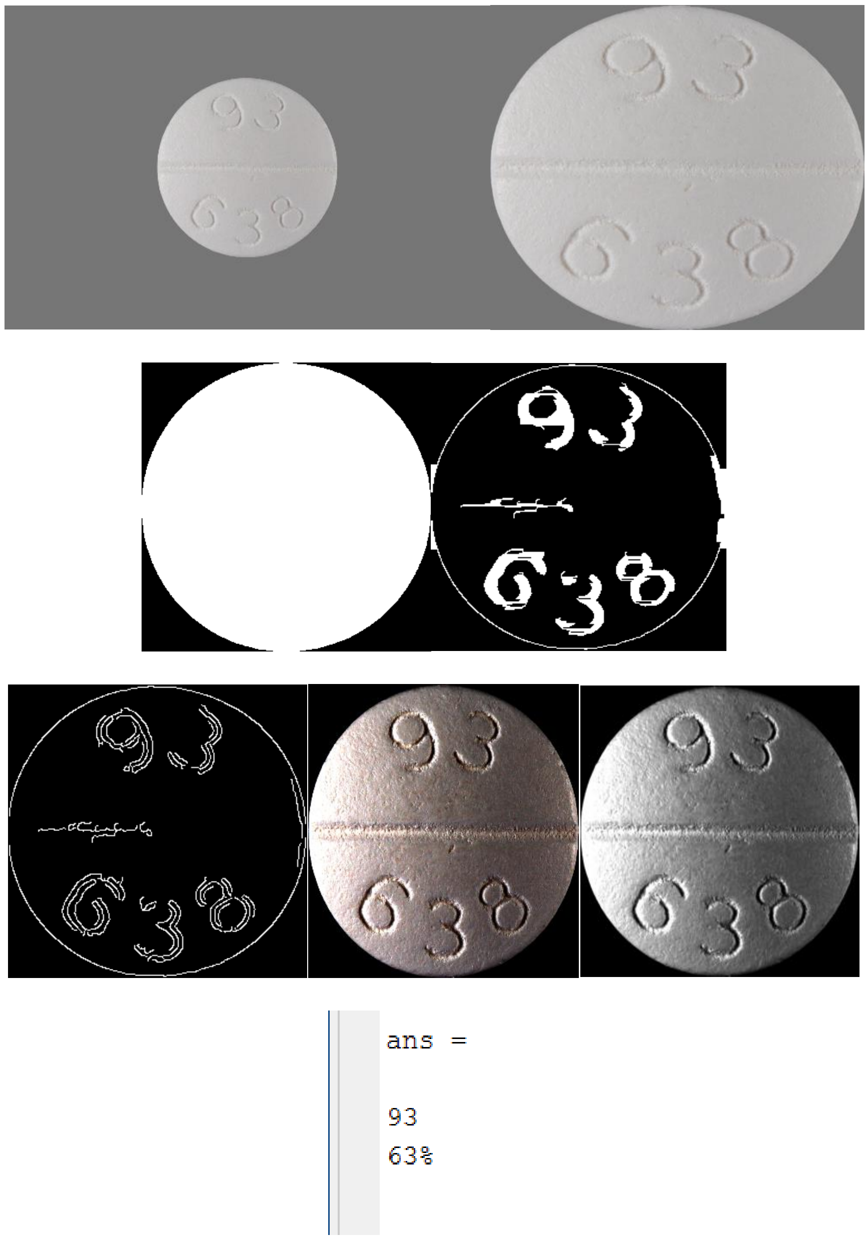


Figure 33. Failed cases of text detection

Figure 34. Failed cases of text detection

7.3 Experimental Results and Discussion

We did all our experiments using the reference image-set from the RxIMAGE dataset. Table I provides distribution of data according to class.

TABLE I. Distribution of training and testing image

Type	Training Set	Testing Set
Circular	814	90
Oval	188	20
Oblong	722	80
Special	78	8
Sub-total	1802	198
Total	2000	

For train dataset the values for Eccentricity, Extent and Narrowness and empirically measured the minimum and maximum values for all three features so that four shapes can be differentiated from each other.

The maximum and minimum values for three features is given in Table II. As during the image processing steps, information lose occurs and shapes does not appear as strict and structured as they should be, that is why value for each of the features is a range not a single floating point number. The shapes of the pills are not exactly the geometric shape they are supposed to be.

TABLE II. Shapes vs minimum and maximum value of the features

		Circular	**Oval**	**Oblong**	**Special**
Eccentricity	Min	0.0107	0.7212	0.7637	0.0298
	Max	0.2104	0.8938	0.9496	0.8522
Extent	Min	0.7713	0.7598	0.8003	0.6050
	Max	0.7935	0.8013	0.9732	0.9439
Narrowness	Min	0.00011471	0.7509	0.9034	0.0008874
	Max	0.0453	1.7818	2.8779	1.3883

Based on the minimum and maximum values of all three feature values of Table II test dataset was tested. Value of Table II can group the test dataset into four shape category with 93.75% accuracy. Detailed result of this testing is given in Table III.

TABLE III. Detailed test results for all shapes of the test dataset

	In Test Image Set	**Correctly Identified**	**Accuracy**
Circular	90	90	100%
Oval	20	20	100%
Oblong	80	80	100%
Special	8	6	75%
Total	198	196	99%
Average Accuracy			93.75%

Here note that, except the special shaped pills all others are correctly grouped according to their shapes. The two pills that are not correctly identified are classified as Ovals. In Fig 10 we are providing the pill images that are not correctly identified. It can be said that this particular morph is from an oval shaped object roughly we may call it oval. Intrinsic value suggest that it is an oval however visually they are not.

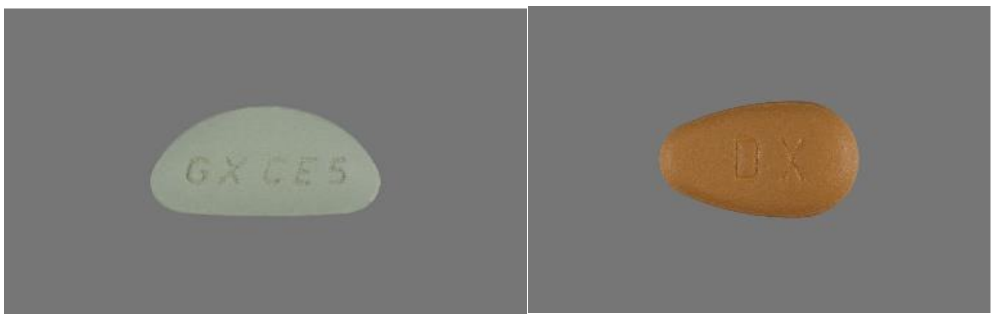


Figure 35. Two pill images that were classified as oval shaped image instead of being classified as special type

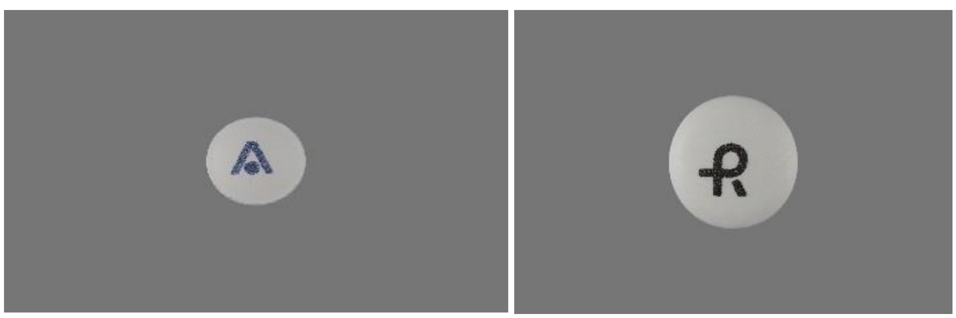


Figure 36. Two pill images that has been misread as "A" & "R" respectively.

From the RxIMAGE dataset around 61% of the pills have engraved text on its surface. And our challenge was to find the text of the remaining 39% pill's imprinted text. By using the proposed method, we have checked 780 images if they provide the correct text or not. And the accuracy of the method is around 81% for the imprinted pills.

8 Conclusion

In this paper we have proposed a fundamental geometric feature based approach to discriminate pill images into four categories: Circular shaped, oval shaped, oblong shaped and special shaped. The proposed method can discriminate them with 93.75% accuracy. Moreover, the pills were classified based on the number of colors in them with 95.3% accuracy. Then we figured out the text imprinted on the pills surface with 81% accuracy by following text segmentation and text detection method. And finally, the engraved texts on the pills were detected with 77% accuracy rate based on their variety of size and contrast level.

However, it might appear in mind that size and color should be more dominant feature to discriminate prescription pills from their images. However, taking size as a feature to discriminate pills from images is a difficult one, as size is not scale invariant and photo is taken from two different height, so pills will have different size in the image. The goal of this work is to provide an initial screening of the pill images into smaller groups based on shape and text and then next we are working on how to farther split the dataset according to the color of the pill to farther refine the search and make a mobile application on the whole system. It will be a complete system providing assistance in medicinal pill recognition.

References

[2] G. Pill Image Recognition Challenge by National Library of Medicine (NLM) http://pir.nlm.nih.gov/challenge/.

[3] M. Hagedoorn, “Pattern Matching Using Similarity Measures”, PhD thesis, Universiteit Utrecht, 2000.

[4] R. C. Veltkamp, and M. Hagedoorn, “State of the Art in Shape Matching”, Technical Report, Utrecht, 1999.

[5] A. Hartl, C. Arth, and D. Schmalstieg “Instant Medical Pill Recognition on Mobile Phone”, IASTED International Conference on Computer Vision, 2011.

[6] A. Hartl “Computer-Vision based Pharmaceutical Pill Recognition on Mobile Phones”, CESCG, 2010.

[7] D. G. Lowe. Distinctive Image Features from Scale- Invariant Keypoints. IJCV, 2004.

[8] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-Up Robust Features (SURF). CVIU, 2008.

[9] Y. B. Lee, U. Park, and A. K. Jain, “PILL-ID: Matching and Retrieval of Drug Pill Imprint Images”, ICPR, 2010.

[10] S. Saha, A. Mahmudy, A. A. Ali, M. A. Amin, “Classifying Digital X-Ray Images into Different Human Body Parts”, ICIEV, 2016.

[11] M. M. Rahman, B. Poon, M. A. Amin, H. Yan, “Recognizing Bangladeshi Currency for Visually Impaired”, ICMLC, 2014.

[12] E. Hossain, S. M. S. Alam, A. A. Ali, M A. Amin, “Fish Activity Tracking and Species Identification in Underwater Video” ICIEV, 2016.

[13] M A. Amin, and M. K. Mohammed, “Overview of the ImageCLEF 2015 medical clustering task”, ImageCLEF2015, 2015.

[14] Chen, Huizhong, et al. "Robust Text Detection in Natural Images with Edge-Enhanced Maximally Stable Extremal Regions." Image Processing (ICIP), 2011 18th IEEE International Conference on. IEEE, 2011.

[15] Gonzalez, Alvaro, et al. "Text location in complex images." Pattern Recognition (ICPR), 2012 21st International Conference on. IEEE, 2012.

[16] Li, Yao, and Huchuan Lu. "Scene text detection via stroke width." Pattern Recognition (ICPR), 2012 21st International Conference on. IEEE, 2012.

[17] Neumann, Lukas, and Jiri Matas. "Real-time scene text localization and recognition." Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012.

Appendix 1: Conference Papers

The paper published on International Conferences are attached below.

Paper 1 is published on The 14th Pacific Rim International Conference on Artificial Intelligence 2016 (PRICAI 2016) titled "Medicine Recognition Using Intrinsic Geometric Property from Pill Image".

Medicine Recognition Using Intrinsic Geometric Property from Pill Image

Md. Zakir Hossan, Tanjina Piash Proma, M. Ashraful Amin

Computer Vision and Cybernetics Group, Department of Computer Science and Engineering, Independent University, Bangladesh, Bashundhora R/A, Dhaka, Bangladesh

zshujon@gmail.com, tanjinaproma@gmail.com, aminmdashraful@iub.edu.bd

Abstract. It is often the case that prescription pills do not come with blister or alu alu packaging where the identity of the pill is available, rather it comes in air-tight plastic bottles or labeled Ziploc bags. Problem with such bottle or pack is that, if the label is removed then it becomes difficult to tell what the pill is. Moreover, there is the issue of visually impaired people having difficulty identifying pills outside the pack. There are such many scenarios where it is good to have an automated pill recognition system. Due to the large variety of size, shape, color, texture it is a difficult task for human to tell about the identity of any individual medical pill. To localize a pill from a given dataset using computer vision techniques requires multiple steps. This paper will describe how to split a dataset according to the shape of pill. To find the shape information we used intrinsic geometric properties such as: eccentricity, extent and narrowness of pill which can be extracted from image using carefully selected image processing techniques. Reference values of discriminative parameters are determined using 'RxIMAGE', National Library of Medicine, USA database. The overall shape discrimination accuracy of the proposed system is 93.75%.

Keywords: Medical Imaging, Pill Image, Eccentricity, Extent, Narrowness

1 Introduction

Prescription pills are available in blister, alu-alu or container (bottle) pack and Ziploc pack. When a medical pill is out of its pack, it is almost an impossible task to recognize. It is also possible that in any way the label can be damaged. In many cases someone has to take two or more pill at a time, so it is hard to find the correct pill from several where the label is damaged or all of them are out of designated containers. Older people will face even more difficulty identifying pills out of its container.

For a visually impaired person this problem is even worse. As the printed labels are of no help to them. Though sometimes they can identify some pills by touching the embossed imprints on the surface of a pill. However, this works with some tablet form of pills, as in the capsule form of pill it is not possible to emboss any text. Moreover, sometimes the pills may not even have any imprint on them. However, sometimes here may be some imprint on the surface of a capsule but that is not sensible by touching. All of these issues may leads to wrong medication. Which may cause an unwanted serious health hazard.

adfa, p. 1, 2011.

Automated recognition of medicine is relatively a new concept though there exists a few work on this issue. Lee et al. [8] developed an application that is able to automatically identify illicit drugs. Hartl et al. [4], and Hartl [5] in their work tried to recognize medical pills using mobile device. However, their proposed method is limited to find shape and color of the pills.

In this paper we propose a method to provide an initial sorting of pill images based on proper measurement of their intrinsic geometric parameters values.

2 Proposed Method

The usual approaches to find features in image for object recognition are the use of, Scale Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF), Pyramid Histogram of visual Words (PHOW) [6,7, 9-11].

Above mentioned feature descriptors perform excellent with objects that contains variations within the object under consideration. However, medical pill does not contain enough texture or corner features thus SIFT or SURF cannot extract enough discriminating features to separate pills from each other. The most dominant features for pills are size, shape and color. Hagedoorn [2], and Veltkamp and Hagedoorn [3] in their work provide a detailed survey about shape recognition techniques or algorithms. Such as tree pruning, Hough transform, Fourier descriptor, statistics, wavelength transform, deformable templets, curvature scale space, relaxation labeling, and neural network.

Because only features available for discrimination are shape, size and color; the proposed method intends to extract values of intrinsic geometric property of medical pills from image by using image processing techniques for a set of training image and then determine a threshold of those parameters and use it on differentiate pills.

2.1 The NLM Dataset

It is common practice to use standard datasets [1, 12] for developing and testing a new idea. On January 2016, NLM open a challenge under a federal notice “Pill Image Recognition Challenge [1]”. NLM’s purpose of this Challenge is to find a set of algorithm and software that can rank an input pill according to the similarity to images of unknown prescription pills to known prescription pill images in the NLM RxIMAGE dataset. There is two image-set one for consumer quality image and other for reference image and a ground truth table.

The RxIMAGE dataset contains 5000 images of 1000 different pills in the consumer quality image-set. To mimic the quality of image in consumer level the photos were taken with digital cameras built in to mobile devices. Fig. 1 shows some sample images from the database.

Fig. 1. Sample images from the reference set of the NLM RxIMAGE dataset

2.2 Image Processing Steps

Ones the images are collected next task is to measure the geometric properties of the main object (pill) in an image. However, before we can do so, series of image processing steps are required to acquire the shape of the pill in the image. Consider the image in Fig 2(a). It is not possible to measure the intrinsic parameters values directly.

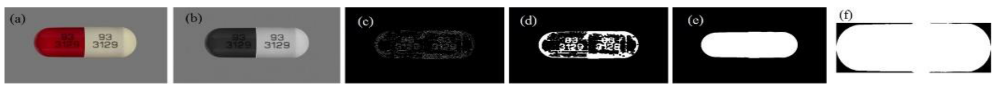


Fig. 2. (a) 24-bit Color pill image from the reference database, (b) 8-bit Gray pill image for the image from color image, (c) Detected Edges of the pill image for the Gray image, (d) Edge image after application of closing morphological operator, (e) Image after application of fill morphological operation, (f) Segmented pill region from pill image.

Converting RGB image to Grayscale image: Every reference image in the database is a 1600 X 2400 X 3 matrix. For farther processing the color image is converted into a 1600 X 2400 8-bit grayscale image (Fig. 2(b)).

Edge Detection: Ones the images are converted into an 8-bit gray image Canny edge detection algorithm is applied on it to acquire edge in the image (Fig. 2(c)).

Closing Morphological Operation: To segment out the pill from image it is necessary to find the boundary of pill. Morphological closing operation is used with disk shape structuring element to connect any edge pixel that is missed in the Canny edge detector during edge filtering process (Fig. 2(d)).

Fill Morphological Operation: Note that the image in Fig. 2(d) contains many smaller components, however we are interested in a single object representing the pill. Thus we apply fill morphological operator to finally acquire a single object (Fig. 2(e)).

Detecting the Pill in the Image: Now apply boundary enclose on the whole image the pill will be detected in the image (Fig. 2(f)).

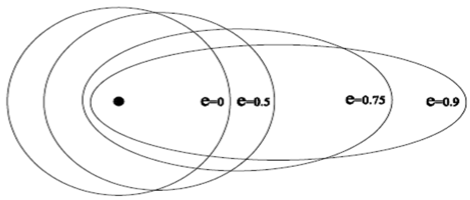
2.3 Metric for Shape Measurement

Now that we have the pill image in more suitable form for measuring geometric parameter values. Depending upon the types of measurements used in image processing or analysis it is possible to split and categorize any dataset in various way. One can categorize the measurement types based on scale. Roundness, which can express the radius of curvature of the object corners in the next smaller scale measurement. The types of measurement can also be categorized based on the assumption level and the degree to which the results are calculated. In this case linear results such as area, perimeter can be calculated from the pixel map of the image. By these other results such as spherical equivalent volume and the circular equivalent diameter are calculated.

Finally, using several relations and ratios of the mentioned factors, metrics such as the aspect ratio, circularity, convexity, solidity, spherical equivalent volume, eccentricity, extent, elongation, convex hull area etc. may be calculated.

To categorize the RxIMAGE dataset we used Eccentricity, Extent and Narrowness (using aspect ratio) to separate the dataset into four categories. These are, Circular, Oval, Oblong, and Special (Fig 1).

Eccentricity: The ratio of the distance between two focal of the ellipse and its major axis length is known as eccentricity. Fig. 3(left) illustrates, how eccentricity can be a measure of shape. Its value is between 0 and 1. The shape with 0 (zero) eccentricity is actually a circle and eccentricity 1 (means) it's a line or line segment. It is a rotation and scale invariant property.



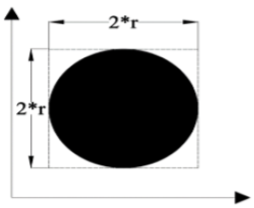


Fig. 3. (left) A perfect circle has an eccentricity of 0, while an oval or ellipse has 1, (right) Dimension of bounding box for a circle

Extent: It is the ratio of the area of the region to the total area of its bounding box (bounding box: the smallest rectangle containing the region) Fig. 3(right) illustrate and Eq. 1 gives the formula to measure extent of an image object. Though it scale invariant, however not rotation invariant. Thus we have to rotate the image according to its major axis to measure the correct value of extent. In theory if the extent is 0.7854 it's a circular shaped object.

$$Extent = \frac{Object\ Area}{Bounding\ Box\ Area} \quad (1)$$

Narrowness: It is defined as the absolute difference of aspect ratio and inverse of aspect ratio. The ratio between major axis and minor axis length is the measure aspect ratio (Eq. 2). If the narrowness is 0 (zero) the shale is circular. A higher value of narrowness (Eq. 3) will means the object is narrower.

$$aspectRatio = \frac{majorAxisLength}{minorAxisLength} \quad (2)$$

$$narrowness = abs\left(aspectRatio - \frac{1}{aspectRatio}\right) \quad (3)$$

3 Experimental Results and Discussion

We did all our experiments using the reference image-set from the RxIMAGE dataset. Table I provides distribution of data according to class.

TABLE I. Distribution of training and testing image

Type	Training Set	Testing Set
Circular	814	90
Oval	188	20
Oblong	722	80
Special	78	8
Sub-total	1802	198
Total	2000	

For all the images of the train dataset we had calculated the values for Eccentricity, Extent and Narrowness and empirically measured the minimum and maximum values for all three features such that we can differentiate four shapes from each other. The maximum and minimum values for three features is given in Table II. Here note that, it is possible to find in theory one precise value for each of the features, however in reality the value is rather a range not a single floating point number. We believe this is happening because, during the image processing steps, information lose occurs and shapes does not appear as strict and structured they should be. Moreover, the shapes of the pills are most of the time circle like, oval like, oblong like, triangle like, trapezoid like not exactly the geometric shape they are supposed to be.

TABLE II. Shapes vs minimum and maximum value of the features

		Circular	Oval	Oblong	Special
Eccentricity	Min	0.0107	0.7212	0.7637	0.0298
	Max	0.2104	0.8938	0.9496	0.8522
Extent	Min	0.7713	0.7598	0.8003	0.6050
	Max	0.7935	0.8013	0.9732	0.9439
Narrowness	Min	0.00011471	0.7509	0.9034	0.0008874
	Max	0.0453	1.7818	2.8779	1.3883

Based on the minimum and maximum values of all three feature values of Table II test dataset was tested. Value of Table II can group the test dataset into four shape category with 93.75% accuracy. Detailed result of this testing is given in Table III.

TABLE III. Detailed test results for all shapes of the test dataset

	In Test Image Set	Correctly Identified	Accuracy
Circular	90	90	100%
Oval	20	20	100%
Oblong	80	80	100%
Special	8	6	75%
Total	198	196	99%
Average Accuracy			93.75%

Fig. 4. Two pill images that were classified as oval shaped image instead of being classified as special type

Here note that, except the special shaped pills all others are correctly grouped according to their shapes. The two pills that are not correctly identified are classified as Ovals. In Fig 4 we are providing the pill images that are not correctly identified. It can be said that this particular morph is from an oval shaped object roughly we may call it oval. Intrinsic value suggest that it is an oval however visually they are not.

4 Conclusion

In this paper we have proposed an intrinsic geometric feature based approach to discriminate pill images into four categories: Circular shaped, oval shaped, oblong shaped and special shaped. The proposed method can discriminate them with 93.75% accuracy. However, it might appear in mind that size and color should be more dominant feature to discriminate prescription pills from their images. However, taking size as a feature to discriminate pills from images is a difficult one, as size is not scale invariant and photo is taken from two different height, so pills will have different size in the image. The goal of this work is to provide an initial screening of the pill images into smaller groups and then next we are working on how to farther split the dataset according to the color of the pill. Moreover, we intend to incorporate OCR techniques to farther refine the search. Such that we can determine the text or symbol inside the interior of a pill, and that will help us more to accurately identify what kind of medicinal pill is that.

References

1. Pill Image Recognition Challenge by National Library of Medicine (NLM) http://pir.nlm.nih.gov/challenge/.
2. M. Hagedoorn, “ Pattern Matching Using Similarity Measures”, PhD thesis, Universiteit Utrecht, 2000.
3. R. C. Veltkamp, and M. Hagedoorn, “State of the Art in Shape Matching”, Technical Report, Utrecht, 1999.
4. A. Hartl, C. Arth, and D. Schmalstieg “Instant Medical Pill Recognition on Mobile Phone”, IASTED International Conference on Computer Vision, 2011.
5. A. Hartl “Computer-Vision based Pharmaceutical Pill Recognition on Mobile Phones”, CESCG, 2010.
6. D. G. Lowe. Distinctive Image Features from Scale- Invariant Keypoints. IJCV, 2004.
7. H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-Up Robust Features (SURF). CVIU, 2008.
8. Y. B. Lee, U. Park, and A. K. Jain, “PILL-ID: Matching and Retrieval of Drug Pill Imprint Images”, ICPR, 2010.
9. S. Saha, A. Mahmudy, A. A. Ali, M. A. Amin, “Classifying Digital X-Ray Images into Different Human Body Parts”, ICIEV, 2016.
10. M. M. Rahman, B. Poon, M. A. Amin, H. Yan, “Recognizing Bangladeshi Currency for Visually Impaired”, ICMLC, 2014.
11. E. Hossain, S. M. S. Alam, A. A. Ali, M A. Amin, “Fish Activity Tracking and Species Identification in Underwater Video” ICIEV, 2016.
12. M A. Amin, and M. K. Mohammed, “Overview of the ImageCLEF 2015 medical clustering task”, ImageCLEF2015, 2015.