Medicine Recognition from Colors and Text

Tanjina Piash Proma
Computer Science & Engineering
Independent University Bangladesh
Dhaka, Bangladesh
+8801919471971
tanjinaproma@gmail.com

Md. Zakir Hossan
Computer Science & Engineering
Independent University Bangladesh
Dhaka, Bangladesh
+8801885570597
zshujon@gmail.com

M. Ashraful Amin
Computer Science & Engineering
Independent University Bangladesh
Dhaka, Bangladesh
+8801550153903
aminmdashraful@iub.edu.bd

ABSTRACT
People of hazy or blurred vision or the elderly people finds it way too challenging just to identify the pills if they are out of the box or packet. And various pills of various shapes, size, texture, color comes with a diverse set of medicinal components. It creates confusion among pills of same color and shape to identify based on a specific texture. For visually impaired people, even if they configure the shape of the pill, the color information and the texts imprinted on the pill remains unknown to them. In this paper, the splitting processes of a dataset according to the number of colors and the texts imprinted on the pills, will be described. Initially the color information were extracted by segmenting pill region from pill image and then some statistical measurements i.e. Kurtosis and skewness, are calculated for probability distributions generated from the image histograms. Thus figuring out the how many colors the pill surface consists of. For the text recognition, the probable text region is detected for an error free text detection. For high quality image data, the reference images from NLM RxIMAGE database has been utilized. The overall accuracy of the proposed system for number of color determination is 95.6% and text recognition accuracy is 81.32%.

CCS Concepts
• Computing methodologies→Artificial intelligence→ Computer vision→Computer vision problems→Matching

Keywords
Medical Imaging; Pill Image; Color determination; Kurtosis; Skewness; Text segmentation; Text recognition.

1. INTRODUCTION
The packaging of the prescription pills is done using blister, alu-alu or container (bottle) pack and Ziploc pack. Once a medical pill is out of its pack, the task to recognize the medical pill becomes impracticable. It is also possible that somehow the label can be damaged. In case of taking two or more pill at a time, it is hard to find the correct pill from several others where the label is damaged or all of them are out of assigned containers. For regular people this is a great hassle since it's not only an inconvenience but also a catalyst for serious health hazard.

The complications are severe for elderly people with incapacitated vision and further perilous for a visually impaired person, as they cannot recognize the color or the text printed on the pill. The only physical feature they can identify is the shape of the pill by touching it. The engraved ones are easier to identify by touch but the imprinted ones are unrecognizable since they are mostly printed on the smooth surface. And all these issues may lead to wrong medication which may cause an unwanted serious health hazard. If the text printed on the capsules and their number of colors could be recognized, then these severe risk factors can be avoided before any serious accident.

Image processing is an important technique used in automated recognition of human faces [1], recognize human from iris images [2], and help establish communication with hearing impaired with mass people [3]. People have also tried to use various image processing techniques to resolve the pill recognition problem. However, the amount of work on automated medicine recognition is insufficient as it is relatively a new concept. There are online websites who offer to identify the pill by giving input about the pill feature like color, shape and text [4, 5]. These search engines for pill identification are of great help for general mass. But none of them are helpful for visually impaired person since he cannot figure out these physical features. Lee et al. [6] developed an application that is able to automatically identify illicit drugs. Hartl et al. [7], and Hartl [8] 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. But the case of imprinted text has not been solved.

Human can easily differentiate shapes and similar methods are used in identification of plant species from leaf shape [9]. We also have worked on the shape detection of pill images with an accuracy of 93.75% [10]. Continuing with that detection criteria, the number of color and text detection has been resolved. The proposed method of this paper is to provide splitting algorithms for pill images based on their number of colors and then recognizing the text imprinted on them, continuing the previously measured shape using fundamental geometric property.

2. THE NLM DATASET
Implementing quality dataset [11, 12] for developing, training and testing of a system has always been a norm. On January 2016, NLM open a challenge under a federal notice "Pill Image Recognition Challenge [11]". 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

ICGSP '19, June 1–3, 2019, Hong Kong, Hong Kong

ACM ISBN 978-1-4503-7146-9/19/06…$15.00
https://doi.org/10.1145/3338472.3338484

to mobile devices. And the Reference Image dataset consists of high quality 2000 JPEG image files for 1000 pills. And for our work, we used the reference images. Figure 1 shows some sample images from the reference image dataset.

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

3. PROPOSED METHOD

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) [13, 14]. 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 [15], and Veltkamp and Hagedoorn [16] 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.

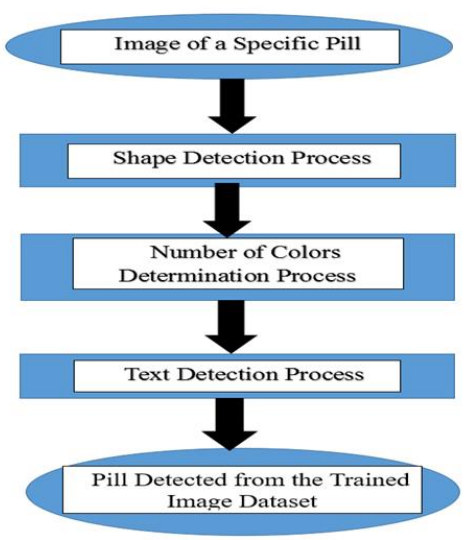

Figure 2. Overview of the proposed method.

The proposed method is to use the prime features like shape, number of colors and the printed text. Our previous work on shape measurement intended 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. This time, we have used the segmented pill image and calculated a statistical measurement from the histogram of the pill to find out how many colors are in the pill. Finally, we have used text segmentation technique and text detection method to find out the text regions and the containing texts printed on the pills. The proposed method uses decision tree classification method for classifying the images for the features like shape and color. After detecting the shape, number of color and texts imprinted on the pill, finally the pill is detected from the trained image dataset. Figure 2 shows the overview of the proposed method.

4. IMAGE PROCESSING STEPS FOR SHAPE MEASUREMENT

As mentioned previously, the procedures from shape detection method have further advanced the other feature detection criterion for this problem. The shape detection method was done by following several processes i.e. converting RGB image to Grayscale image, edge detection, closing morphological operation, fill morphological operation and finally detecting the pill in the image.

After completing these image processing, the geometric parameters of the processed images has been measured to classify the shape. Linear results such as area, perimeter have been calculated from the pixel map of the image. By these spherical equivalent volume and the circular equivalent diameter were calculated. Using several relations and ratios of the mentioned factors, metrics such as the aspect ratio, circularity, solidity, spherical equivalent volume, eccentricity, extent, elongation etc. have been 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 [10].

5. IMAGE PROCESSING STEPS FOR FINDING NUMBER OF COLORS

Another most distinctive feature of pills is their color. With their 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. It informs the blind people that the medicine is a capsule containing 2 colors helping them to identify the pill better.

5.1 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. Thus, 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. As illustrated in figure 3, 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. After the segmented images have been computed, the next step is to section the pill region using the morphological images and the grayscale images.

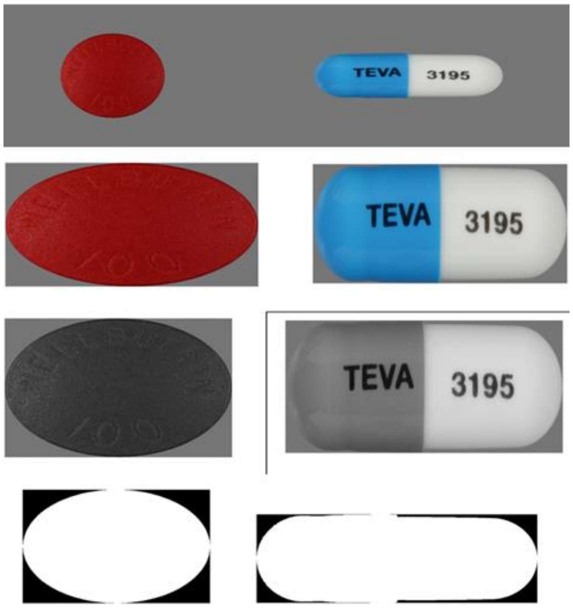

Figure 3. Segmentation of the pill image to find out the pill region. (Left) single colored pill, (Right) Two Colored Pill.

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 as illustrated in figure 4. With the number of peaks available on the filtered histogram, the number of colors existing in the pill is obtained.

5.2 Generating Histogram for finding the number of Peak present

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

$$B = (\gamma\text{^}2 + 1) / \kappa \qquad (1)$$

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

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

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).

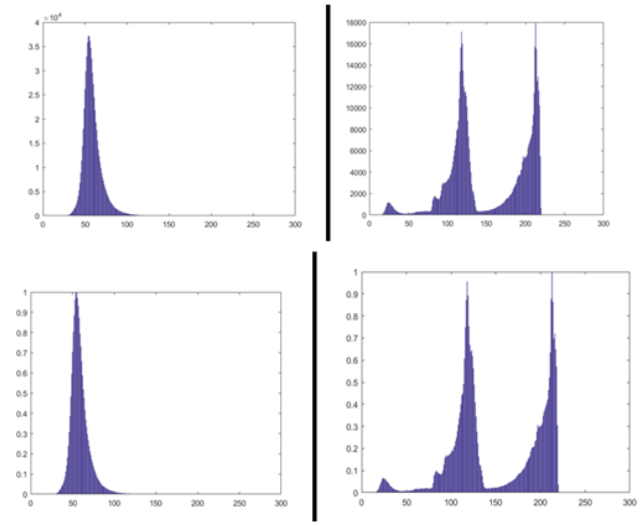

Figure 4. Histogram of (Left) single colored pill, (Right) Two Colored Pill.

6. IMAGE PROCESSING STEPS FOR TEXT RECOGNITION

All Pills have a texture which as in itself is a significant feature. The ones with some 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, it 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).

6.1 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 regions [17] 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 5 (Left)).

6.2 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 [18]. Among the several geometric properties, some are better for detecting text and non-text region [19, 20]. The geometric properties used here are eccentricity, solidity, Euler number, aspect ratio, extent. (Figure 5 (Right)).

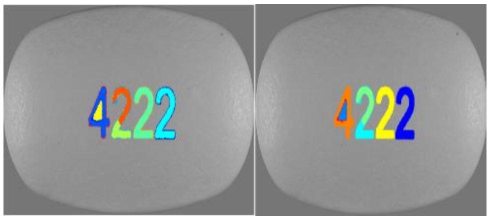

Figure 5. (Left) MSER region. (Right) After removing Non-Text Regions Based on Basic Geometric Properties (Right) Two Colored Pill.

6.3 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 6 & 7).

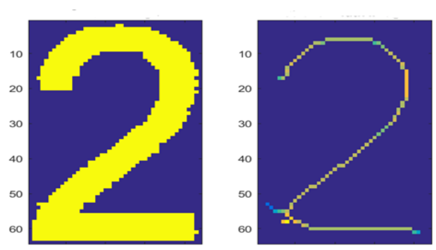

Figure 6. Region Image & Mark Width Image (Right)

Figure 7. After removing non-text regions based on mark width variation

6.4 Merge Text Regions for Final Detection

Till now the individual text characters has been detected. Now we need to merge the characters for finding meaningful text which is easily understandable. For using the optical character recognition, the characters should be merged to form a text line. 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 8).

6.5 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(right)).

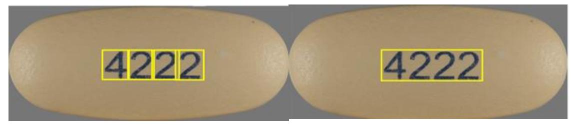

Figure 8. Bounding boxes text (left). Detected text (right)

7. EXPERIMENTAL RESULTS AND DISCUSSION

We did all our experiments using the reference image-set from the RxIMAGE dataset. Features are more prominent in the reference dataset.

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.

In Figure 9, 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.

There are cases when the texts printed on the pills can be misread due to being symbols or special font. The symbols which looks close to a text can lead to confusion as well as a different font style. In Figure 10 the symbols were misread as letter "A" and "R" even though they are not. Which is why the text recognition accuracy falls to 81.32% of the accuracy scale.

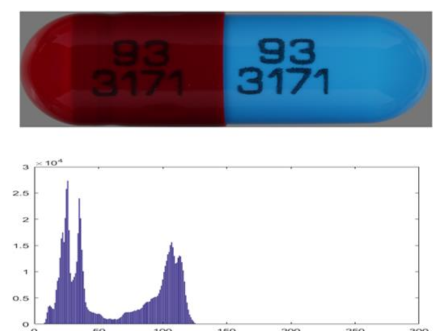

Figure 9. Detecting Number of colors for shaded Pills

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 81.32% for the imprinted pills. But for the Number of color detection, we used all the images from the dataset and found approximately a 5% inaccuracy for the images where pills had some lighting and contrast issues as mentioned above.

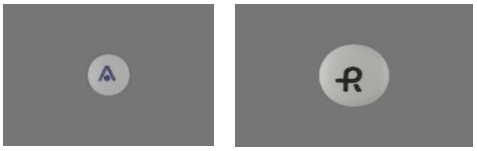

Figure 10. Two pill images that has been misread as "A" & "R" respectively.

8. CONCLUSION

In this paper, we have presented a fundamental feature-based approach to discriminate pill images. Then we figured out the number of color present on the pill with 95.6% accuracy and text imprinted on the pills surface with 81.32% accuracy by following text segmentation and text detection method. The goal of this work is to provide an initial screening of the pill images into smaller groups based on number of color and text and then next we are working on how to farther split the dataset according to the color of the pill. Moreover, we intend to work on the engraved pill text 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

9. ACKNOWLEDGEMENT

This project is supported by a grant from the information and Communication Technology Division (ICTD) of Bangladesh Government.

10. REFERENCES

[1] Iqbal Nouyed, Bruce Poon, M. Ashraful Amin, Hong Yan, "A study on the discriminating characteristics of Gabor phase-face and an improved method for face recognition", IJMLC, 2015.

[2] Iftakhar Hasan, M. Ashraful Amin, "Dual Iris Based Human Identification", Journal of Signal, Image and Video Processing, 2014.

[3] M. Ashraful Amin and Hong Yan, "Sign Language Finger Alphabet Recognition from Gabor-PCA Representation of Hand Gestures", ICMLC, 2007.

[4] Pill Identification Wizard from Drugs.com. https://www.drugs.com/imprints.php

[5] Pill Identifier (Pill Finder Wizard) - https://www.rxlist.com/pill-identification-tool/article.htm

[6] Young-Beom Lee, Unsang Park and Anil K. Jain, "PILL-ID: Matching and Retrieval of Drug Pill Imprint Images", ICPR, 2010.

[7] Andreas Hartl, Clemens Arth, and Dieter Schmalstieg, "Instant Medical Pill Recognition on Mobile Phone", IASTED International Conference on Computer Vision, 2011.

[8] Andreas Hartl, "Computer-Vision based Pharmaceutical Pill Recognition on Mobile Phones", CESCG, 2010.

[9] Javed Hosain and M. Ashraful Amin, "Leaf Identification Based Plant Classification", ICCIT, 2010.

[10] Md. Zakir Hossan, Tanjina Piash Proma and M. Ashraful Amin, "Medicine Recognition Using Intrinsic Geometric Property from Pill Image", PRICAI, 2016.

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

[12] M Ashraful Amin and Mahmood Kazi Mohammed, "Overview of the ImageCLEF 2015 medical clustering task", ImageCLEF2015, 2015.

[13] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", IJCV, 2004.

[14] Herbert Bay, Anderas Ess, Tinne Tuytelaars and Luc Van Gool, "Speeded-Up Robust Features (SURF)" CVIU, 2008.

[15] Michiel Hagedoorn, "Pattern Matching Using Similarity Measures", PhD thesis, Universiteit Utrecht, 2000.

[16] Remco C. Veltkamp and Michiel Hagedoorn, "State of the Art in Shape Matching", Technical Report, Utrecht, 1999.

[17] Huizhong Chen, Sam S. Tsai, George Schroth, David M. Chen, Radek Grzeszczuk, BerndGirod, "Robust Text Detection in Natural Images with Edge-Enhanced Maximally Stable Extremal Regions", ICIP, 2011.

[18] Lukas Neumann and Jiri Matas, "Real-time scene text localization and recognition", CVPR, 2012.

[19] Alvaro Gonzalez, Luis M. Bersaga, J. Javier Yebes, Sebastian Bronte, "Text location in complex images", ICPR, 2012.

[20] Yao Li and Huchuan Lu, "Scene text detection via stroke width", ICPR, 2012