March - April 2020
ISSN: 0193-4120 Page No. 4441 - 4449

An Image Processing Based Information Extraction Model for Preventing the Misuse of Drugs

Geun-Hwa Jeong[1], Koo-Rack Park[*2], Yun-Yeol Lee[3], Dong-Hyun Kim[4]
[1]M.S. Student, Department of Computer Engineering, Kongju National University, 31080, Rep. of Korea
[*2]Professor, Department of Computer Science & Engineering, Kongju National University, 31080, Rep. of Korea
[3]Professor, Department of Computer Science & Engineering, Kongju National University, 31080, Rep. of Korea
[4]Ph. D., Department of Computer Engineering, Kongju National University, 31080, Rep. of Korea
jgh910000@naver.com[1], ecgrpark@kongju.ac.kr[*2], alphaone@kongju.ac.kr[3], dhkim977@naver.com[4]

Article Info
Volume 83
Page Number: 4441 - 4449
Publication Issue:
March - April 2020

Abstract
Background/Objectives: As the medicine and pharmaceuticals are developed along with the rapid development of the society, human lives are extended to enter the aging society. Accordingly, new drugs are being manufactured and administered, but people in the informational vulnerability class can result in the misuse of drugs due to the absence of information. Therefore, drug information extraction model is required for convenience in verifying the information on drugs.

Methods/Statistical analysis: Non-prescription drugs are being sold worldwide, and the general public is granted with the accessibility and choice on the relevant drugs to increase the awareness on self-medication. People are attempting self-medication to treat their minor symptoms and illnesses, but adverse reactions on drugs from misuse are increasing annually. This is due to the great difficulty in identifying the information on drugs, and therefore in this study, drug images were acquired through sensor elements such as CCD or smartphone camera to identify the information on drugs, and drug shape, character string and color etc. There were analyzed to realize an image processing based drug information model for providing various information on the relevant drugs.

Findings: In this study, an image processing based medical information model was proposed to provide the information on drugs. In the proposed model, images were processed based on ROI setting, image binarization, and histogram equalization technique for quick processing on the inputted images. In addition, a drug information model was realized based on the technologies such as contour algorithm and color identification through RGB decision to extract the accurate shape of the drugs. The proposed model was tested to verify the accurate video image input of the drugs and print of the drug information. It is considered that when the relevant drug is searched through the proposed model in this study, existing inconveniences and difficulties will be reduced, and the existing method of adding various information to search the drug can be deviated to identify the drug information quickly through image acquisition processing.

Improvements/Applications: The proposed model can not only extract drug information, but also be applied to many fields through the recognition of various objects. For the direction of the study in the future, further studies should be continued on the algorithm strong against the changes in the value according to the shade and lighting in the process of acquiring the RGB value on image processing, and on the model on providing the information through the internet website.

Article History
***Article Received**: 24 July 2019*
***Revised**: 12 September 2019*
***Accepted**: 15 February 2020*
***Publication**: 26 March 2020*

***Keywords:** Medicine, ROI, Vision System, Binarization, Histogram Equalization.*

1. Introduction

From Year 2012, various countries around the world have been displaying non-prescription drugs referred to as OTC (Over The Counter) outside the displays of the pharmacy for sales to grant the general public with the accessibility and choice on the relevant drugs. In Korea, drugs for general drugs such as antipyretic drug or cold medicine are not only sold in pharmacies, but also in convenience stores to rapidly form the market for the general drugs sold for safety. Accordingly, the awareness is being increased on self-medication, where the patients are responsible for their health and disease issues. In other words, it signifies on supporting the self-medication of the people on minor symptoms or illnesses without wasting any time or money to enhance the choice and autonomy of the people [1]. However, the safety factors of monitoring on the relevant drugs can be overlooked due to the tags related to the relevant drugs [2]. Today, people can easily access through the internet to obtain information on drugs, but when the drugs critical to the human lives can even result in death from severe adverse reactions when the drugs are administered inappropriately, so information on drugs and symptoms are very important to reduce medical accidents. Accordingly, medical monitoring must be continued. Generally, physicians are provided with sufficient amount of information to select and decide the drugs for the treatment of the patients [3]. This drug monitoring faces great challenges to develop better healthcare system, and main tasks are web-based sales and information, globalization, extensive safety issues, continuous monitoring on the existing products, economic growth of public health and pharmaceutical industry, developing countries and rising countries, attitude and awareness on the benefits and damages of the drugs, and results and impacts [4]. The pattern of using drugs becomes a very important decision factor on safety of the drugs, and in countries that are capable of developing drugs, the use of drugs most widely used is more general [5]. Advertising prescription drugs directly to the consumers can be seen frequently in many countries, and through this information, patients can determine their treatment directly without the help of the physicians or pharmacists. As a result, illegal sales of drugs through internet is increasing, and it has great impact on over-prescription according to the physician's requirements [6,7]. Adverse reactions from drug misuse is increasing annually, and safety accidents and drug misuse by the socially underprivileged class such as the children, old and the infirm, and pregnant women including the general public are increasing to be exhibited as a social issue. Especially, the people vulnerable to understanding the information indicated on the drugs and acquiring the information are exposed to higher risks than general adults, and studies on assistive technology regarding the information service for the class vulnerable to information are in progress, but there are many difficulties in identifying the drug information for the pharmaceutical industry. Proper information on drugs is effective on recovery and maintenance of health, and appropriate drug information service is required. Currently, drugs are produced by classifying the mark, pattern and color, etc., and there are search services provided through this classification. However, the help from the professional pharmacist is required to identify the information on the drugs that are generally administered, or the information including the shape, color and formulation of the relevant drug must be entered for identification, so there is an inconvenience of input in the existing service, and it takes great amount of time to search the information. In this study, a model is proposed on acquiring the drug image through the sensor elements such as the CCD camera or smartphone camera to identify the drug information through image processing. By searching the relevant drug through the proposed model, it is expected that the existing inconveniences and difficulties will be reduced, and the existing method of entering various information for the search can be avoided to quickly identify the drug information through processing by image acquisition.

2. Related Works

2.1. ROI(Region of Interest)

There are three reasons for performing ROI analysis as follows [8]. Firstly, it enables simple data search; in complex design, convenient identification is possible when it is difficult to identify the activity pattern according to the condition in the overall map. Secondly, the number of statistical tests can be limited to several ROIs to control the errors. Thirdly, the test can be limited to the fields defined by function based on other information such as scan or option. This is especially useful in the complicated design with various options such as factorial design. The most general method for exploratory ROI analysis is developing small ROI in the activation cluster, and when subjected to large-scale cluster, it is useful to produce the ROI on the local maximum value to explore the various areas in the cluster for measurement, and it can be applied on exploring the activity patterns in various options. Analysis using robust statistics [9] can reduce the impact of singular value, so it can be utilized especially in this case.

2.2. Histogram Equalization

The gray value that one pixel of an inputted image has can be defined as histogram. When the inputted image is biased of being too bright or dark, the histogram of the brightness value can be distributed equally to adjust the brightness of the image. The purpose of histogram smoothing is to produce the histogram with constant distribution, and the bright image is darkened while the dark image is brightened to maintain the proper brightness value. In addition, it is creating the image with brightness level distributed evenly throughout the overall brightness scale, and the contrast of the brightness value close to the maximum value of the histogram is improved, but the contrast near the minimum point is reduced. The algorithm of histogram smoothing can be summarized as follows [10]. The following (Formula 1) calculates the accumulated histogram that is normalized, and (Formula 2) is for setting the transform function F for the histogram smoothing.

$$h_c(i) = h_c(i-1) + h(i) for\ i = 1, 2, \dots, \quad (1)$$

$$\mathrm{L} - 1 \text{ where } h_c(0) = h(0)$$

$$\mathrm{F(z)} = \mathrm{Ə}(L-1)h_c(z) for\ z = 0, 1, \dots, \quad (2)$$

$$L - 1$$

Here, L is the input image with the overall intensity levels, h is the normalized histogram, and h_c is the cumulative normalized histogram. Ə is the function rounded off into the closed integer. Afterwards, the inputted image is re-scanned, gray level z' is recorded of the image, and it is set as $z' = F(z)$.

Improving the gray value is a very important field on image processing for both the computer vision system and humans, which is used diversely in the pre-processing stage of voice recognition, medical image processing, text synthesis and image/video processing [11-14]. Also, among the histogram smoothing methods, the simple method of linear/non-linear gray level conversion function can be used [15], or the method of using the edge [16] and the method of using the connected component information [17] are used on complicated analysis of various image functions. This histogram smoothing redistributes the intensity distribution, and when the peak and valley are high in the histogram of the image, the peak and valley are maintained after the smoothing, but transferred. Therefore, each field is allocated with new intensity value in the histogram smoothing, and such value is based on the previous intensity level [18].

3. Proposed Model

3.1. Model Configuration

The following [Figure 1] is a configuration of the proposed model MID(Medication Information Detection). Probe is composed of the CCD camera, smartphone camera, lens and lighting. For the filming and input of a drug image, and the machine vision system for detecting the medical information is inputted of the image filmed through the probe to perform the pre-treatments such as binarization, ROI setting and histogram smoothing, etc. to recognize the image for detecting the medical information. The data similar to the result value is shown in a list to be provided to a user.

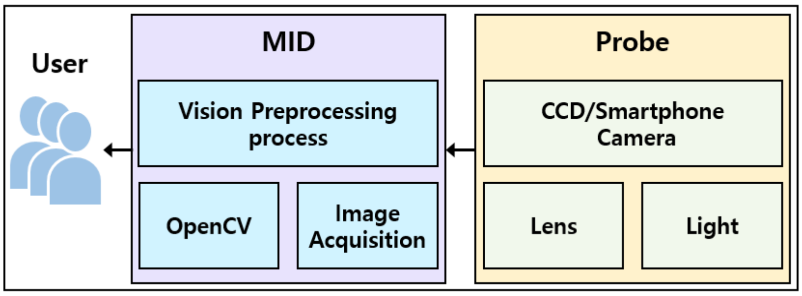


Figure 1. Proposed Model Configuration.

3.2. System Process

The following [Figure 2] is a process of the proposed model to be composed largely of 6 stages.

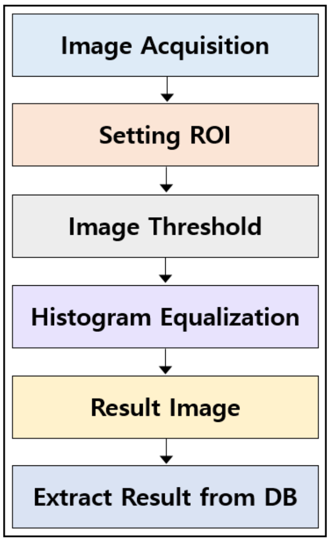


Figure 2. Proposed Model Process.

Firstly, a user acquired the image of the relevant drug through the probe. Secondly, the relevant work field is selected as the ROI setting stage. Thirdly, binarization is performed to acquire the threshold of the image, and fourthly, histogram smoothing is performed on the image completed with binarization. Fifthly, the image data completed with conversion is acquired, and sixthly, the image data completed with conversion is compared with the data stored tin the database to provide the result value to the user.

3.3. ROI Settings

The region performed with the work of setting the value on the location of the relevant tablet to acquire the information of the tablet in the image inputted by a user is referred to as ROI. The inputted image can be used with the Rect() instance to set the starting point of the relevant X, Y coordinates and finishing point of X, Y and thus to set the relevant region. When the region is set, there is an advantage of increasing the efficiency in work because the work is not performed in the irrelevant regions. To perform the work on converting the image acquired through ROI setting into Grayscale, cvtColor() instance is used to apply different RGB values in the image into the same value, and the values of 3 channels are all added to be divided by 3 to convert the relevant image into Grayscale. This work is required for the binarization. The following [Figure 3] (a) is the original image acquired through the probe, (b) is the image before converting to grayscale by the ROI setting, and (c) is the grayscale conversion image for the binarization work.

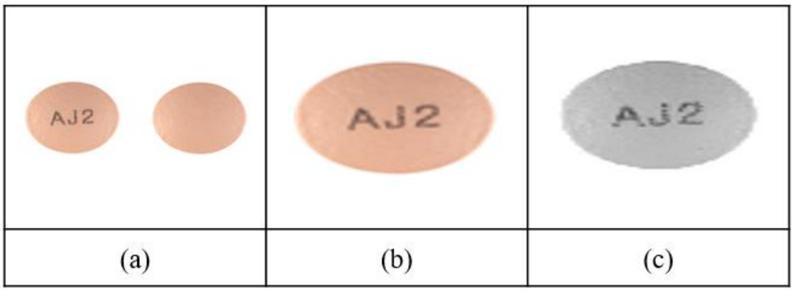


Figure 3. ROI and Grayscale Image.

3.4. Image Binarization

In [Figure 4], the image converted to grayscale (a) was binary-coded, and the threshold value image was extracted as shown in (b).

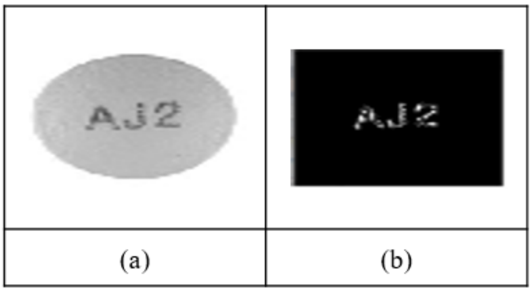


Figure 4. Threshold Acquisition Image.

The inputted image is applied with the binarization technique to detect the letters in the drug. Binarization is for converting the pixel value into 0 and 1 in the standard of the threshold value, and to perform the binarization, the value that becomes the standard of each pixel value in the image must be set. This standard value is the threshold value that specifies the black and white in the image. Threshold() instance was used to set the relevant image, threshold value and maximum value.

3.5. Histogram Equalization

The gray value for each pixel of the applied image can be defined as histogram, and previously, the contrast level is maximized and brightness is redistributed to improve the gray value in the image performed with the binarization. Through this work, the image brightness can be distributed evenly. In [Figure 5], (a) is the threshold extraction image, (b) is the image through histogram smoothing, and (c) is the result image.

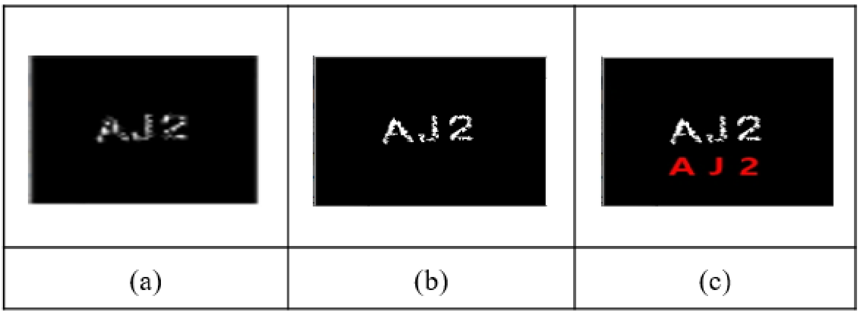


Figure 5. Histogram Equalization and Result Image.

3.6. Database Table

Database was applied with the MsSQL 1.50 version, and it was composed of 3 tables including the table for comparison with the data acquired after the image processing. The table stored with the information on the relevant drug when the list is provided to a user and the 'See Details' button is pressed, and the table for verifying the list of alternative drugs. The image of the acquired drug can be analyzed to compare with the stored data, the data related to the drug list can be provided to the user, and the list of alternative drugs can be verified. [Table 1] below shows the drug information table structure from the acquired image data.

March - April 2020
ISSN: 0193-4120 Page No. 4441 - 4449

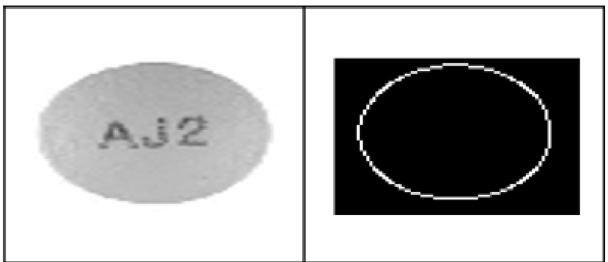
Table 1. Drug Information Table Structure.

No.	Field	Description	Data Type
1	Drug_Name(PK)	Drug Name	Varchar(255)
2	Color_R	Drug Color Red	int
3	Color_G	Drug Color Green	int
4	Color_B	Drug Color Blue	int
5	Drug_Shape	Drug Shape	Varchar(10)
6	Drug_Con	Drug Description	Varchar(10)

4. Test and Results

4.1. Contour and color Detection

Finally, the character string extracted from the image must be compared with the data stored randomly in the Database to list the relevant drugs into the list, but the method of searching the drugs only with the character string is not accurate. Accordingly, the color value through the RGB and the pattern of the relevant drug are used to perform the work for obtaining more accurate data. The image performed with Grayscale is applied with the contour algorithm with the boundary of the region with the equal color value or pixel value to acquire the contour of the image, and the extraction of the data related to the shape of the drug can be possible through this. To perform the contour extraction, the image performed with block and white image or binarization is used. In the [Figure 6] below, (a) is the image converted through grayscale, and (b) is the image detected with the contour line through the Contour.



(a) (b)

Figure 6. Contour Extraction.

The method of acquiring the RGB value can be classified into 3-channel and 1-channel, and the 3-channel refers to the RGB with color. In the 3-channel, the index of the image is used for channel access. 0 has the value of B, 1 with the value of G, and 2 has the value of R. On the initial image used, RGB value was assumed to be (220, 160,140), and the shape of the relevant drug is assumed to be ‘O’. The data of the relevant drugs were all stored randomly, and drug name, R, G, B, and drug shape were all set with the value to enter the data. The information obtained from the applied image was the value R(220,160,140), shape ‘O’, and character string of ‘AJ2’. The data that can be obtained through the shape and character string is the drug 1 and drug 3. But among the two drugs, the RGB value must be compared to distinguish which one is the drug in the inputted image. When the error range was assumed to be 10%, R was selected with the drugs in the range of value in 198~244, G in the range of value in 144~176, and B in the range of value in 126~154. In the case of drug 3, the R is 190 to be out of the error range, so the drug in the inputted image is delivered with the result value of drug 1. If there are several drugs within the error range, the list is organized in the order of smaller error on the relevant drugs to be provided to a user. The information value of the relevant drug can be saved to enable the user to be easily provided with the information on the listed drugs. [Figure 7] is the screen on the list of alternative drugs according to the extraction of the result value.

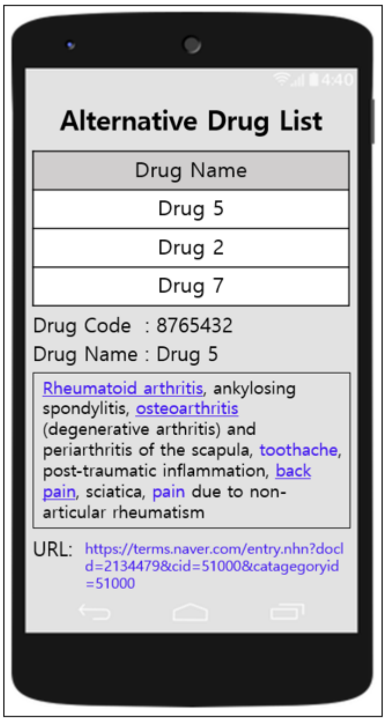


Figure 7. Alternative Drug List.

4.2. Drug Search

The following [Figure 8] is the image input screen and the screen showing the description on the drug.

When the image subjected for drug search is entered, the list of relevant drugs can be verified, and the alternate drug list can also be verified.

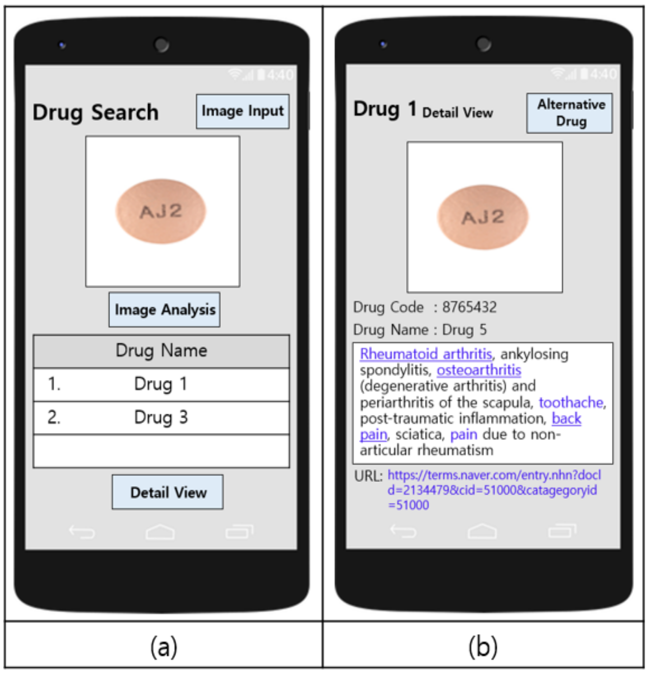


Figure 8. Result Screen.

Through this, the information can be acquired more quickly and accurately compared to the existing solution. Which can reduce the adverse reactions from the misuse of drug that are increasing annually.

5. Conclusion

Many countries in the world are selling non-prescription drugs to grant accessibility and choice to the general public on the relevant drugs. Accordingly, the awareness on self-medication is increasing to support the people of self-medication on the minor illnesses and symptoms to reduce the waste of money and time. It is expanding the autonomy and choice of the people, but the information on the relevant drugs can be easily searched through internet to neglect the consultation with the physicians. However, adverse reactions from the misuse of drugs are increasing annually, and safety accidents and drug misuse are increasing on the socially underprivileged class such as the children, old and the infirm, and pregnant women including the general public to be exhibited as a social issue. And there are many difficulties for the general public on identifying the information on drugs.

Therefore, proper information on the drugs can be effective on the recovery and maintenance of the health, and proper drug information services must be provided. Currently, drugs are produced by distinguishing the mark, pattern and color, etc., and search services are provided through these distinctions. However, the information on the drug can be identified only through the professional pharmacist or by entering the information including the shape, color and formulation of the relevant drug, so there is an inconvenience of input in the existing service, and it takes a lot of time to search the relevant drug. In this study, sensor elements such as the CCD camera or smartphone camera are used to acquire the drug image, and the drug shape, character string and color, etc. are analyzed to provide a user with the information,

March - April 2020
ISSN: 0193-4120 Page No. 4441 - 4449

drug code and name on the relevant drug through the information stored in the database. In addition, an alternative drug list on the relevant drug is provided to propose the model on reducing the misuse by the user and to identify the medical information based on image processing. By using the proposed model to search the relevant drug, it is expected to reduce the existing inconvenience and difficulties, and prevent the use of existing method on adding various information for the search to quickly identify the drug information through the image acquisition and processing method.

In the future, there should be further studies on the algorithm strong on the changes of the value according to the shadow and lighting in the process of acquiring the RGB value on image processing, and also on the search through the internet website.

References

[1] Brass, E. P. Changing the status of drugs from prescription to over-the-counter availability. The New England Journal of Medicine. 2001 Sep;345(11):810-816. DOI: 10.1056/NEJMra011080.

[2] Shaw, D., Graeme, L., Pierre, D., Elizabeth, W., & Kelvin, C. Pharmacovigilance of herbal medicine. Journal of Ethnopharmacology. 2012 Apr;140(3):513-518. DOI: 10.1016/j.jep.2012.01.051.

[3] Härmark, L., & Van Grootheest, A. C. Pharmacovigilance: methods, recent developments and future perspectives. European journal of clinical pharmacology. 2008 Jun;64(8):743-752. DOI: 10.1007/s00228-008-0475-9.

[4] Biswas P, Biswas A. Setting standards for proactive pharmacovigilance in India: The way forward. Indian Journal of Pharmacology. 2007; 39(3):124-128. DOI: 10.4103/0253-7613.33431.

[5] Bapna, J. S., Tripathi, C. D., Tekur U. Drug utilization patterns in the third world. Pharmaco economics. 1996;9(4):286-294. DOI: 10.2165/00019053-199609040-00002.

[6] Kane A, Lloyd J, Zaffran M, Simonsen L, Kane M. Transmission of hepatitis B, hepatitis C, and human immunodeficiency virus through safe injections in the developing world: mode-based regional estimates. Bulletin of the World Health Organization. 1999;77(10):801–807.

[7] de Vries C. S., Duggan C. A., Tromp T. F., de Jong-van den Berg L. T. Changing prescribing in the light of tolerability concerns. Drug Safety. 1999;21(3):153–160. DOI: 10.2165/00002018-199921030-00001.

[8] Poldrack, R. A. Region of interest analysis for fMRI. Social cognitive and affective neuroscience. 2007 Mar;2(1):67-70. DOI: 10.1093/scan/nsm006.

[9] Wager, T.D., Keller, M.C., Lacey, S.C., Jonides, J. Increased sensitivity in neuroimaging analyses using robust regression. Neuroimage. 2005 May;26(1):99-113. DOI: 10.1016/j.neuroimage.2005.01.011.

[10] S. H. Kwon, H. C. Jeong, S. T. Seo, I. K. Lee, C. S. Son. (2008). Histogram equalization-based thresholding. IEICE Transactions on Information and Systems. 2008;91(11):2751-2753. DOI: 10.1093/ietisy/e91-d.11.2751.

[11] A. Wahab, S. H. Chin, and E. C. Tan. Novel approach to automated fingerprint recognition. IEE Proceedings Vision, Image and Signal Processing. 1998 Jun;145(3):160-166. DOI: 10.1049/ip-vis:19981809.

[12] Diabetes Australia. Diabetes globally [Internet]. Canberra ACT: Diabetes Australia; 2012 [updated 2012 June 15; cited 2012 Nov 5]. Available from: http://www.diabetesaustralia.com.au/en/Understanding-Diabetes/Diabetes-Globally/ (website)

[13] Torre, A., Peinado, A. M., Segura, J. C., Perez-Cordoba, J. L., Benitez, M. C., Rubio, A. J. Histogram equalization of speech representation for robust speech recognition. IEEE Transactions on Speech and Audio Processing. 2005 Apr;13(3):355-366. DOI: 10.1109/TSA.2005.845805.

[14] Pizer, S. M. The medical image display and analysis group at the University of North Carolina: Reminiscences and philosophy. IEEE Transactions on Medical Imaging. 2003 ;22(1):2-10. DOI: 10.1109/TMI.2003.809707.

[15] Gonzalez, R. C., Woods, R. E. Digital image processing. 2nd ed. Reading, MA. Addison-

Wesley; 1992. p. 85-103.

[16] Boccignone, G. A multiscale contrast enhancement method. In Proceeding of International Conference on Image Processing. 1997 Oct;1:306-309. DOI: 10.1109/ICIP.1997.647767.

[17] Caselles, V., Lisani, J. L., Morel, J. M., Sapiro, G. Shape preserving local contrast enhancement. Proceedings of International Conference on Image Processing. 1997 Oct;1:314-317. DOI: 10.1109/ICIP.1997.647769.

[18] Bagade, S. S., Shandilya, V. K. Use of histogram equalization in image processing for image enhancement. International Journal of Software Engineering Research & Practices. 2011 Apr;1(2):6-10.