International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 3 March 2023, pp: 1599-1603 www.ijaem.net ISSN: 2395-5252

Enhancing mammographic image scan using contrast limited adaptive histogram equalization technique for machine learning purpose.

Eze Chikwado Emmanuel

Date of Submission: 20-03-2023 Date of Acceptance: 30-03-2023

ABSTRACT
Deep Learning models have proven useful in the analysis, classification and segmentation of medical images, as they have obtained higher accuracy levels compared to their human counterparts. Image preprocessing is fundamental to the development of successful machine learning models and is thus necessary to pay attention to methods used in achieving them. In this paper, mammography image scans from hospitals in Nigeria were preprocessed and enhanced using contrast limited adaptive histogram equalization technique. The image is primarily in the RGB format which is not best suited for histogram equalization since we would require it to be performed on the three channels of the image. Instead, the image is first converted to LAB format, where L holds the lightning value of the image, A and B saves the color information. Histogram Equalization is then performed on the L channel, after which the equalized L channel is merged back to the A and B channels. The image is then converted back to the RGB format. The result shows that the method improves the overall visibility of the images, and increases contrast details for even blurry image scans.

Key words: mammography, histogram equalization, cumulative density function (CDF)

I. INTRODUCTION

Mammograms are x-rays of the breast that show details of the internal structures of the breast. The doctor reading mammogram images will be looking for different types of breast changes, such as small white spots called calcifications, abnormal areas called masses, and other suspicious findings that could be signs of cancer[1]. Calcifications are tiny calcium deposits within the breast tissue. In the mammogram images, calcifications appear as white spots on mammograms. Calcification in the breast are of two types, macrocalcification and microcalcification. Macrocalcification are larger calcium deposits that are mostly due to aging of the breast arteries, old injuries or inflammation. Microcalcifications in breast tissue are one of the key indications appraised by the radiologist for identification of breast cancer in its early stage [1]. Microcalcifications in breast tissue are one of the key indications appraised by the radiologist for identification of breast cancer in its early stage [2].

Preprocessing of mammogram images is essential in the process of breast cancer examination as it could reduce the rate of false positives [2]. The most common contrast enhancement technique for preprocessing of mammogram images and medical images in general is the histogram equalization technique. A technique that has been developed where images are manipulated from its pixel intensity to create an image that is visually greater, called Image enhancement [3]. The purpose is to enhance images for humans visually by improving the interpretation of information contained in it, or also the result can be used as a high quality input for more image processing use [4]. As this technique improves human readability of image scans, the same is applicable to training of machine learning models for medical image classification. This is especially important as [5] noted that many of the decisions that affect a machine learning model's predictive behavior are made during data preprocessing.

II. RELATED LITERATURE

Several articles have been published on the use of histogram equalization for image enhancement. [6] Proposed to use some image processing methods as a data normalization method for machine learning. In addition to z-score normalization, histogram equalization was applied to training data and test data as a pre-processing

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 3 March 2023, pp: 1599-1603 www.ijaem.net ISSN: 2395-5252

method for machine learning. In [6]'s experiments, the proposed scheme is applied to a face-based authentication algorithm with SVM/random forest classifiers to confirm the effectiveness. For SVM classifiers, both z-score normalization and image enhancement worked well as a pre-processing method for improving the accuracy. In contrast, for random forest classifiers, image enhancement worked well, but z-score normalization did not improve accuracy. Histogram equalization was applied in the work done by [7] on corrosion areas and in dealing with low contrast present in shadow areas of an image. In [8], Image enhancement methods based on Histogram Equalization (HE) were studied. It presented an exhaustive review of these studies and suggested a direction for future developments of image enhancement methods. [9] Presented the results of theoretical, simulation and experimental studies focused on the improvement of the ultrasonic visualization of microcalcifications, methods for estimating changes in microcalcification which result from changes in aperture geometry or the presence of aberrator and the analysis of relative efficacy of spatial compounding and synthetic receive aperture geometries in the detection of microcalcifications.

III. METHODS.

a. Histogram Equalization.

The image histogram provides information about the intensity distribution of the pixels in the image. For example, images that are too light or too dark have a narrow histogram [10]. Histogram equalization ensures that the images are sharpened so that the hidden details are immediately obvious. This it achieved by spreading the image frequency histogram across the possible gray scale levels, typically between 0 to 255, from its initial clustered nature to a better even distribution across the range as illustrated in the figure 1 below.

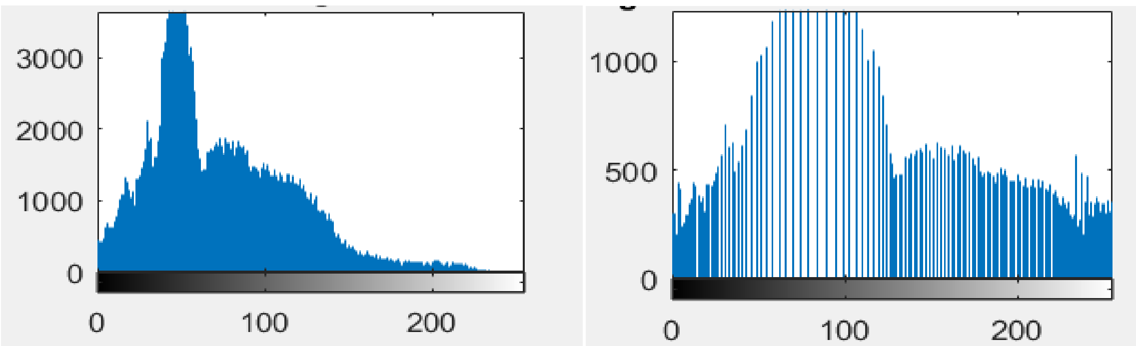


Figure 1: Histogram of an image before and after histogram equalization.

Mathematically, Histogram equalization is achieved thus:

let m be a given image represented as a r x c matrix of integer pixel intensities.

The pixel intensities ranges from 0 to L-1, where L is the possible number of intensity values, in this case 256.

Let p represent the normalized histogram of image m, then;

$P_n = \frac{\text{number of pixel with intensity } n}{\text{total number of pixels}} \quad n = 0,1,2,...,L-1$

The histogram equalized image g will be given then as

$g_{i,j} = \text{floor}((L-1)\sum_{n=0}^{i,j} Pn)$

The floor function rounds down to the nearest integer.

Hence, the transformed image g is obtained by transforming the pixels k of image m by the function

$T(k) = \text{floor}((L-1)\sum_{n=0}^{k} Pn)$

The motivation for this transformation comes from thinking of the intensities of m and g as continuous random variables X, Y on [0, L − 1] with Y defined by

$Y = T(X) = (L-1)\int_0^y Px(X)dx$

$P_x(X)$ is the probability density function of m

$\int_0^y Px(X)dx$ is the cumulative distributive function (cdf)

If for simplicity, we assume that T(X) is differentiable and invertible, it can then be shown that Y defined by T(X) is uniformly distributed on [0, L-1]. i.e $P_y(Y) = \frac{1}{L-1}$

Y = T(X)

Using change of variables formula

then, $Py(Y) = Px(X)\left|\frac{dX}{dY}\right|$ .... (1)

given that $Y = T(X) = (L-1)\int_0^y Px(X)dx$

taking the derivative of the equation above

$\frac{dY}{dX} = \frac{d(T(X))}{dY} = \frac{d((L-1)\int_0^y Px(X)dx)}{dY}$

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 3 March 2023, pp: 1599-1603 www.ijaem.net ISSN: 2395-5252

$\frac{dX}{dY}$ = (L-1) Px(X) .... (2)

If we substitute equation (2) into (1)

$P_y(Y) = Px(X) \left| \frac{1}{(L-1)\,Px(X)} \right|$

Since L >> 1, we can conveniently eliminate the absolute sign, and the Px(X) cancels out leaving us with

$P_y(Y) = \frac{1}{(L-1)}$

b. **Implementing Histogram Equalization using the openCV library in Python.**

The implementation of histogram equalization was achieved following the simple flow chart show below

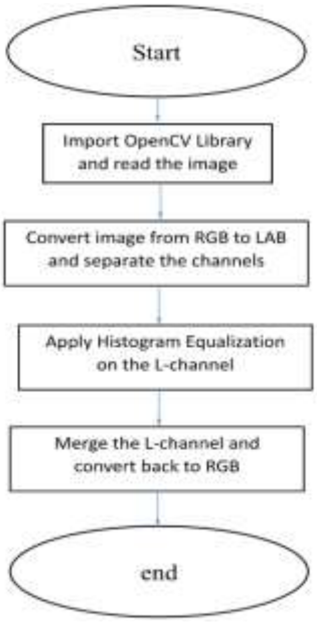
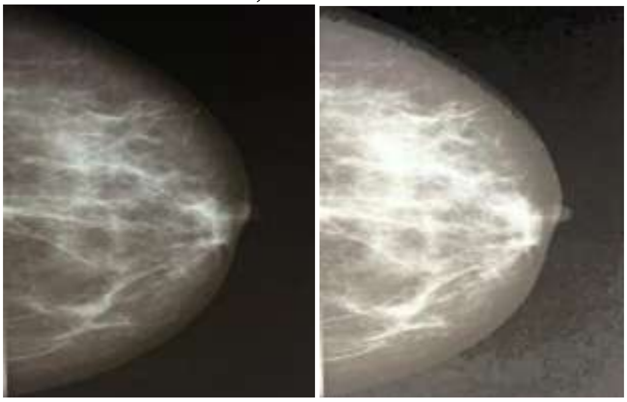


Figure 2: Histogram equalization flow chart

As is shown in figure 2, openCv2 library was first imported, the image was then read from the disk where it is located. The image is primarily in the RBG format which is not best suited for histogram equalization. Instead, the image is first converted to LAB format, where the L channel is the lightning channel of the image, A and B saves the color information of the image. Histogram Equalization is then performed on the L channel, after which the equalized L channel is merged back to the A and B channels. The image is then converted back to the RBG format.

IV. RESULTS AND ANALYSIS

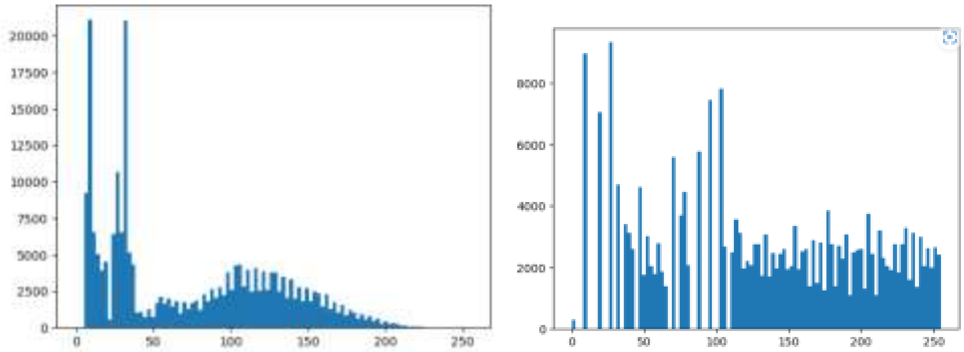
Histogram equalization was performed on a sample of mammogram scan, and the result is shown in figure 3.



(a) (b)

Figure 3: Histogram equalization applied on mammogram image scan, (a) is the original image while (b) is the histogram equalized image.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 3 March 2023, pp: 1599-1603 www.ijaem.net ISSN: 2395-5252

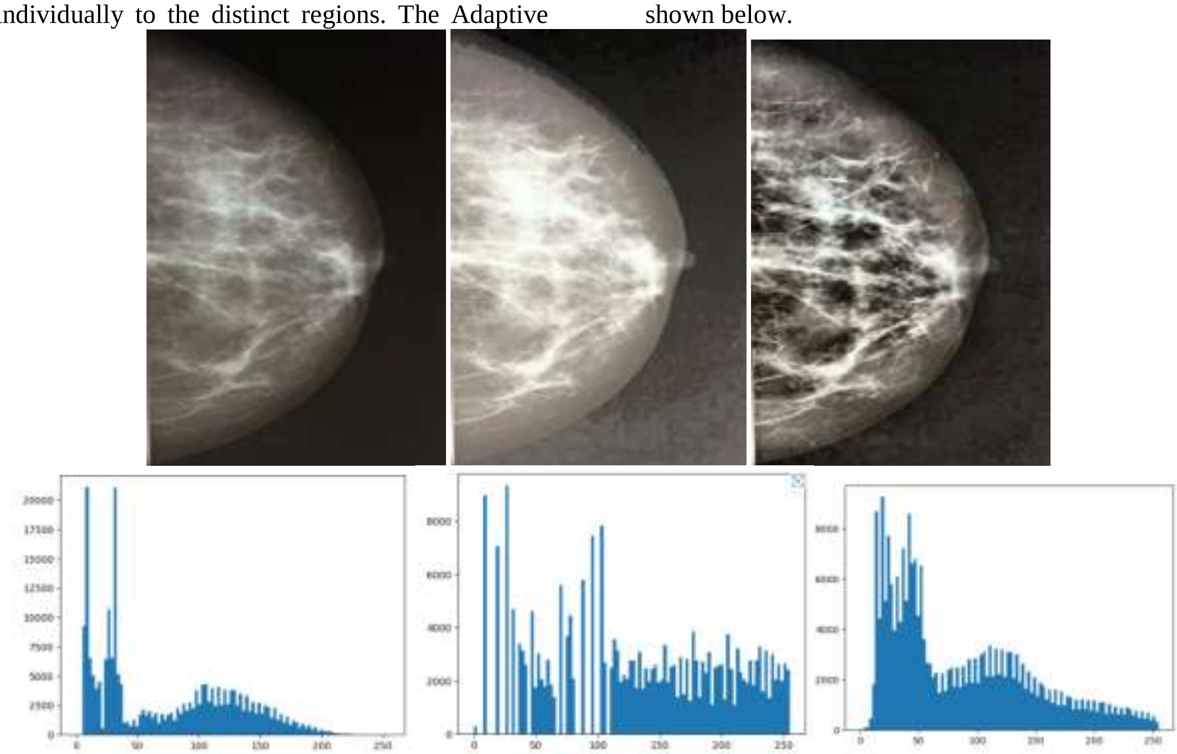


(a) (b)

Figure 4: Histogram distribution of the images in figure 3a and 3b respectively

Applying Contrast-Limited Adaptive Histogram Equalization(CLAHE)

The Histogram equalization technique fails when the input image has a large area with a low-intensity background. This causes the severe washing-out of images and thus effectively amplifies the noise in the image. To circumvent this, the technique of Contrast-Limited Adaptive Histogram Equalization is employed which is an improvement of Adaptive Histogram Equalization (AHE). In Adaptive Histogram Equalization, the image is divided into distinct regions, so that the histogram equalization technique is applied individually to the distinct regions. The Adaptive Histogram Equalization also failed as it has a tendency to over-amplify noise in relatively homogeneous regions of an image. Thus the contrast-limited adaptive histogram equalization solves all these problems by limiting the amplification. In CLAHE, the contrast amplification in the vicinity of a given pixel value is given by the slope of the transformation function. This is proportional to the slope of the neighborhood cumulative distribution function (CDF) and therefore to the value of the histogram at that pixel value. CLAHE was performed on the same image and the result is shown below.



(a) (b) (c)

Figure 5: Histogram equalization applied to the same mammogram image scan. (a) is the original image, (b) is histogram equalized while in (c), contrast limited adaptive histogram equalization was applied.

V. CONCLUSION

This paper presented a theoretical overview of contrast limited adaptive histogram equalization and applied the same in enhancing mammogram image scans. The algorithm improved the image quality of the scans and enhanced the lesions and calcifications to be seen more visibly even to the bare eye. The algorithm is targeted at preprocessing image scans for machine learning applications, thus, the result of the enhancement can be directed fed as input to convolutional neural network models for image classification applications.

REFERECES

1. "What does the doctors look for on a mammogram." American cancer society. https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/mammograms/what-does-the-doctor-look-for-on-a mammogram.html#:~:text=Some%20masses%20can%20be%20watched,it%20is%20to%20be%20cancer (accessed Jan. 26, 2023)
2. S. Tripathy, T. Swarnkar, "Unified Preprocessing and Enhancement Technique for Mammogram Images", Procedia Computer Science, Volume 167, 2020, Pages 285-292, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.223. (https://www.sciencedirect.com/science/article/pii/S1877050920306888)
3. H. Ibrahim and N. S. Pik Kong, "Brightness Preserving Dynamic Histogram Equalization for Image Contrast Enhancement," in IEEE Transactions on Consumer Electronics, vol. 53, no. 4, pp. 1752-1758, Nov. 2007, doi: 10.1109/TCE.2007.4429280..
4. B. Oktavianto and T. W. Purboyo, "A study of histogram equalization techniques for image enhancement" International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 2 (2018) pp. 1165-1170
5. C. V. Gonzalez Zelaya, "Towards Explaining the Effects of Data Preprocessing on Machine Learning," 2019 IEEE 35th International Conference on Data Engineering (ICDE), Macao, China, 2019, pp. 2086-2090, doi: 10.1109/ICDE.2019.00245.
6. A. Sada, Y. Kinoshita, S. Shiota and H. Kiya, "Histogram-Based Image Pre-processing for Machine Learning," 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE), Nara, Japan, 2018, pp. 272-275, doi: 10.1109/GCCE.2018.8574654.
7. J. Xiong, D. Yu, Q. Wang, L. Shu, J Cen, Q. Liang, H. Chen and B. Sun, "Application of histogram equalization for image enhancement in corrosion areas", Hindawi Shock and Vibration, 2021, article ID 8883571, page 13, https://doi.org/10.1155/2021/8883571
8. W. A. Mustafa, M. M. Abdul Kader, "A review of histogram equalization techniques in image enhancement application" IOP Conf. Series: Journal of Physics: Conf. Series 1019 (2018) 012026 doi :10.1088/1742-6596/1019/1/012026.
9. M. E Anderson, M. S. Soo, R. C. Bentl, and G. E. Trahey. The detection of breast microcalcifications with medical ultrasound. J Acoust Soc Am. 1997 Jan;101(1):29-39. doi: 10.1121/1.417973. PMID: 9000730.
10. T. KokKeong, and J. P. Oakley. "Enhancement of color images in poor visibility conditions." Image Processing, 2000. Proceedings. 2000 International Conference on. Vol. 2. IEEE, 2000.