



Histogram Specification for MRI Normalization

Styliani Tsihlaki

Hellenic Mediterranean University, Heraklion, Greece

Department of Electrical and Computer Engineering

e-mail: stsihlaki@gmail.com

1. INTRODUCTION

Magnetic Resonance Imaging (MRI) is a non-invasive method that is mainly based on the magnetic resonance of the Hydrogen nuclei, which are abundant in the human body [1]. MRI is widely used to study the human structure (e.g. the human brain, chest, leg, etc.) and to evaluate heterogeneities (e.g. tumors) [2], while it has electromagnetic RF type of radiation. Generally, human bodies are made of 75% water, while each molecule of the body's water has two hydrogen nucleus [3]. Basically, MRI depends on the magnetic properties and the number of hydrogen nuclei in the displayed area, unlike radiation imaging, in which the contrast depends on the varying attenuation of the structures depicted. Thus, the large number of water molecules in the human body often results in high resolution images [4]. Furthermore, in various pathological conditions, the distribution of water is altered, as well as in other small molecules that are rich in hydrogen [3][4]. During an MRI, running different sequences with various weightings, results to different contrasts that can be chosen for the displayed area. The three primary sequences are a) T1-weighted, b) T2-weighted and c) Proton Density (PD) weighting [1][4]. T1 anatomical scans usually have the best resolution and are useful for the location of anatomical structures, whereas T2 is useful for the detection of damaged brain areas. PD scan reveals the total densities of hydrogen/mm³. In contrast to Computed Tomography (CT), MRI intensities do not have a precise definition of the tissue, which allows images to have a wide range of intensity scaling, even if they originate from the same subject, or the same image sequence in different scanners [4]. MRI processing techniques, such as segmentation [5], are required to understand both normal ageing and disease progression, while image voltages are used as a key feature for almost all MRI processing algorithms.

In the image processing community, the technique of histogram matching, or else histogram specification, represents a way to modify image points, in order to align the histogram of an image with either a target histogram or a reference image histogram. Such methods can be a good starting point for standardizing a medical image. Histogram specification points to a group of image transformations that aims to get images that have a desired shape in their histograms¹. Over the years it has been widely used as a tool of digital image processing and as a technique of image enhancement [5]. One image enhancement parameter is the intensity, which is known as a continuous random variable (RV) defined by its function of the probability density (PDF). The transform function to be found in this setting must be such that the transformed RV has the defined PDF, while already given a known-distributed RV [6][7].

The required process to obtain a consistent histogram image is in accordance with the image optimization process, called histogram equalization. Histogram equalization is a tool used to process images, by adjusting the intensity distribution of the histogram, in order to change the contrast of the image [5]. Furthermore, the gray-levels of an image are distributed to the overall scale and each gray-level is given an equivalent number of pixels. This process results to a more balanced and better contrasting imaging, in terms of the human eye (Figure 1). Beyond the more pleasant display, equalized images can often make details more noticeable in the dark/light areas of the original images [5][6]². Higher performance in image enhancement can be achieved when attention is given to the human visual system (HVS) [5]. The image histogram is defined based on a particular HVS model, so that there is an equalized histogram to the subjectively perceived image [6]. Therefore, such methods may produce good contrast with a minimized effort and computational time. Overall, the purpose of this technique is to give the resulting image a linear cumulative distribution function (CDF), since a linear CDF refers to the cumulative sum of all the possibilities within its domain (Equation 1) and is associated with a uniform resulting histogram³.

¹ Source: https://en.wikipedia.org/wiki/Histogram_matching

² Source: https://en.wikipedia.org/wiki/Histogram_equalization

³ Source: http://www.sci.utah.edu/~acoste/uou/Image/project1/Arthur_COSTE_Project_1_report.html

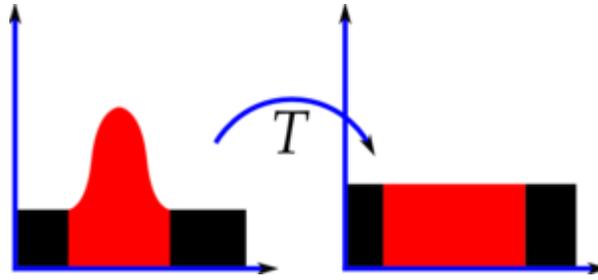


Figure 1. Histogram equalization ⁴

$$cdf(x) = \sum_{k=-\infty}^x P(k) \quad \text{Equation 1}$$

On the other hand, the normalization of an image's intensity is an important preprocessing step in MRI analysis [8][9]. There are discrepancies resulting from the widely varying intensity scales of MRIs, while errors in the output of algorithms or in the estimation process are introduced to tasks of image processing (e.g. segmentation). Thus, intensity normalization is of high importance [9]. Histogram transformations are used by a wide range of intensity normalization methods, which seek to detect a histogram-based one-to-one representation of the intensity, in order for MRI intensities to conform in a specific scale [10].

2. STATE-OF-THE-ART

In this section, a selection of histogram matching/specification and histogram-based normalization proposals and implementations from the digital image processing literature, were reviewed and discussed in detail.

The first approach is the one of Chen et al. [2], where they presented a histogram specification method that involved the automatic selection of reference frames. Their approach was specialized in glioma MRIs, in order to address the issue of inconsistency and efficiency in the data from multiple institutions. Since glioma is described as a high-risk/fatal type of brain tumor, the suggested approach of Chen et al. [2] is of high importance. Particularly, they aim to replace professionals (radiologists) in choosing the reference frames manually, with an automated selection of reference frames in the

⁴ Source: https://en.wikipedia.org/wiki/Histogram_equalization

histogram specification process. As we mentioned before, histogram specification is an important tool for image enhancements, as well as an extension of histogram equalization, which can ease the possible difficulties that may occur in histogram equalization. In this approach, they present and describe various concepts that they have used. Specifically, they set $r = \{r_{i,j}\}$ to be a *Height* \times *Width* discrete input digital image and $L = \{0, 1, 2, \dots, L - 1\}$ to represent the gray-levels of this image. They define the image histogram/gray-level probability density, as shown in Equation 2, where N represents the total number of the gray-level pixels and is defined as $N = \text{Height} \times \text{Width}$. CDF for the reference image z is also defined in Equation 3, while M represents the pixels that have a gray-level of L . In histogram specification, the transformation function y is retrieved (Equation 4) and the original image gray-level is translated. The aim is to match the histogram of the transformed image to the reference histogram, but in order to maintain the original image details, the transformation function y should increase in a monotonical way (Equation 5). In addition, Equation 6 presents the definition of the minimum value y , which allows for the mapped image to match a chosen histogram.

$$P_r(l) = \frac{N_l}{N}, \forall l \in L \quad \text{Equation 2}$$

$$v_z(y) = \sum_{i=0}^y P_z(i) = \sum_{i=0}^l \frac{M_i}{M}, \forall y \in L \quad \text{Equation 3}$$

$$v_z(y) = s_r(x) \quad \text{Equation 4}$$

$$y = v_z^{-1}[s_r(x)] \quad \text{Equation 5}$$

$$y = \underset{k}{\text{arg min}} |s_r(x) - v_z(k)| \quad \text{Equation 6}$$

Furthermore, the collection of reference frames was carried out by an automated grid search algorithm, known as HSASR (Algorithm 1). First, the search range is constrained through a rough scan of a collection of intra-glioma samples, and then the reference frame in the histogram is chosen based on a fine search through the samples extracted from the

previous scan. They performed glioma grading with the use of HSASR algorithm on two datasets (GliomaHPPH2018 and BraTS2017), while they also evaluated the efficiency of this method.

```

(i) Input: Mixed dataset  $F$ , data sample number  $N$ , number of each sample slice  $K$ 
(ii) Output: Optimal reference slice  $F_K$ 
(1) Step1: Coarse search
(2) Function HSBM ( $n, F1, l_1$ )//Histogram specification based on brain MRI
(3) {
(4)  $F1 = \text{Random}(\text{range}(0, N), F * 30\%)$ //Random selection of 30% of dataset
(5)   For each  $n \in F$  do
(6)     For each  $n1 \in F1$  do
(7)        $\text{cdf}_{F_{n1}} = \text{hist}_{F_{n1}} \cdot \text{cumsum}()$ 
(8)        $\text{cdf}_{F_{n2}} = \text{hist}_{F_n} \cdot \text{cumsum}()$ 
(9)       Calculate:  $\text{hist} = \arg\min_m |(\text{cdf}_{F_{n2}} / \text{shape}(F_{n2})) - ((\text{cdf}_{F_{n1}} / \text{shape}(F_{n1})) (m))|$ 
(10)    End for
(11)   Calculate:  $\text{Auc}(n_{k/2}) = \text{roc\_auc\_score}(\text{Label}_{\text{original}} - \text{Label}_{\text{predict}})$ 
(12) End for
(13) }
(14)  $F2 = \text{Sort max}(\text{Auc}_{F_{k/2}})$ 
(15)  $F2_C = \text{Max}(F2)$ 
(16) For each  $c \in F2$  do
(17)   if  $|F2_C - \text{Auc}(c)| \leq T$  then
(18)     Select  $c \rightarrow F3$ 
(19)   end
(20) End for
(21) Select the optimal set  $F3$ 
(22) Step 2 Fine search
(23) Initializes Pointers  $l_1, l_2$ 
(24) For each  $n \in F3$ //Search both ends
(25)    $l_1 = k/2$ 
(26)   For  $l_1 \leq k$  and exit (tumor $_{l_1++}$ );  $l_1++$  do
(27)     HSBM ( $n, F1, l_1$ )
(28)   End for
(29) End for
(30) For each  $n \in F3$ 
(31)    $l_2 = k/2$ 
(32)   For  $l_2 \leq k$  and exit (tumor $_{l_2++}$ );  $l_2++$  do
(33)     HSBM ( $n, F1, l_2$ )
(34)   End for
(35) End for
(36) Output: select the first optimal slice  $F_K$ 

```

Algorithm 1. Algorithm HSASR's pseudocode. Reprinted from [2].

The findings of the evaluation indicated that the method had some functional utility in the research of glioma grading and could increase the clarity of tumor intensity in the MRI. In conclusion, high efficiency was indicated in this approach, with a score of 0.978 in mean AUC, 94.13% in accuracy, 94.64% in sensitivity and 93% in specificity. Chen et al. [2] implied that their approach “*has a slight advantage over the result of a manually selected reference frame and it is about 15% higher on all indicators compared with those without HSASR algorithm. This method can also successfully minimize multicenter data variations and improve the efficiency of the prediction model*”.

Coltuc et al. [6] present experimental findings of their method, which includes the histogram specification of the discrete counterpart failure, while they also address several

image normalization and enhancement techniques. Their method consists of three main parts in the process of exact histogram specification, which aims to transform an image to obtain the exact desired histogram. The first part is the precise ordering of the image's pixels, as shown in Equation 7, where a $M \times N$ pixels string is defined by a clear ordering relation ($<$) with the pair (x_i, y_i) [7]. On the other hand, the second part includes the split of the ordered string in groups of pixels and the third one is the assignment of a gray-level to these pixels.

$$f(x_0, y_0) < f(x_1, y_1) < f(x_2, y_2) < \dots < f(x_{M-1, N-1}, y_{M-1, N-1}) \quad \text{Equation 7}$$

Exact histogram specification, as suggested by Coltuc et al. [6], can be addressed through the strict ordering of the image pixels and in many different ways. Their suggestion included the concept of imagining a family of neighborhoods around every pixel, in which the pixel tends to be lighter than another pixel, since its local mean is greater than the local mean of the other one. Thus, the actual gray-level ordering was optimized. They also refer to colored image histogram specification as not obvious [6], while they specify a precise ordering between colored pixels, through expressing the colored image in an RGB color space, and then proceed to process only the component of luminance. Overall, the main steps of the exact histogram specification approach involve (a) the RGB to hue-saturation-intensity (HSI) transition, (b) the ordering process, (c) the histogram specification implementation and (d) the transition back to RGB. Figure 2 presents the test images that were used to evaluate the aforementioned approach, in comparison to logarithmic and linear histogram specification presented in Figure 3. In conclusion, the proposed approach was considered to aid the tuning of histograms and in combination with the precise ordering that they proposed, the overall method was deemed interesting and important.



Figure 2. Original and perfectly equalized test images, respectively. Reprinted from [6]



Figure 3. Logarithmic and linear histogram specified test images, respectively. Reprinted from [6].

A histogram-based normalization approach was proposed and established by Sun et al. [8], aiming to decrease the intensity fluctuations in scanner sensitivity that are caused in response to changes in scanner efficiency. They performed an experiment with subjects between 20-30 years old from China, in which they aimed to the development of a young adult brain model. They scanned each subject twice, using various display parameters in two separate scanners and the brain MRI images extracted (two per subject) were a total of 22 images. The image with the lowest noise level was detected through noise assessment and was then classified as the reference image (high-quality), while the image with the higher noise level was also identified and classified as the low-quality image. The histogram of low-quality image was then normalized to the histogram of high-quality picture. The proposed normalization algorithm consists of two phases. First, they implemented intensity scaling (IS) for the reference image. Specifically, the reference image intensities were assigned a value between the high intensity region (HIR) and the low intensity region (LIR). In Equation 8, the prementioned application is presented, where f refers to the reference's gray-level value at specific points, while f' refers to the transformation of the related gray-level value.

$$f'(x, y, z) = \frac{f(x, y, z) - LIR}{HIR - LIR} \quad \text{Equation 8}$$

Proceeding to the second phase, they implemented histogram normalization (HN). Particularly, this process involved the low-quality image histogram as an input picture to be extended, in order to suit the histogram of the reference image, such that the range of intensity is among LIR and HIR for the image that was normalized (Equation 9).

$$g'(x, y, z) = \frac{HIR - LIR}{S_{max} - S_{min}} (g(x, y, z) - S_{min}) + LIR \quad \text{Equation 9}$$

In addition, three series of experiment were conducted, where they evaluated the proposed method in comparison to existing intensity normalization approaches and three main aspects were examined. The first aspect was if the comparison accuracy of the non-rigid registration was increased in relation to other methods of normalization. The second aspect checked whether the automated segmentation of brain tissue was enhanced through normalization, while the third one tested if normalization can aid creating different normalized MRI tissue bulks that correspond to high-quality MRI bulks of the same subject. Overall, the results indicated that Sun's et al. [8] approach has been shown to reduce the intensity variations from 7.5 to 2.5%. In conclusion, they suggest that, through their normalization approach, a better-quality brain model can be created [8], while the overall method can significantly boost the efficiency in image analysis.

3. METHODOLOGY

In this section, the methodologies followed and implemented are described, aiming to achieve MRI histogram specification. Both implementations revolved around three key concepts, but with different techniques followed in each one. These key concepts were: (a) histogram equalization (Figure 1), (b) histogram matching/specification, and (c) histogram comparison for the evaluation of efficiency of all the methods that were used.

The first methodology was implemented with the proposed approach of Coltuc et. al [6] followed as a reference. Specifically, the implementation of Coltuc's et. al [6] approach

was carried out by Stefano Di Martino⁵ and was reproduced and enhanced in this paper using Python 3.7 and all the necessary Python packages. In the first place, the target and reference images were imported. Target image was a brain MRI, while the reference image was an irrelevant one. These original images were processed through our implemented histogram equalization method, which essentially corresponds to a mapping transformation. First, we found the CDF of each image's histogram and performed a normalization to [0, 255] (Figure 4, Figure 5). Afterwards, the minimum histogram value was found, and the histogram equalization equation was applied [7], as shown in Equation 10. Furthermore, we used the masked array concept, which is provided by Numpy⁶, where all operations are conducted on non-masked elements. This process resulted to the creation of an informative table about each input pixel's output value, while we proceeded to the application of the transformation. The transformation results, as well as the resulted histogram equalized images, are presented in Figure 6 and Figure 7 respectively.

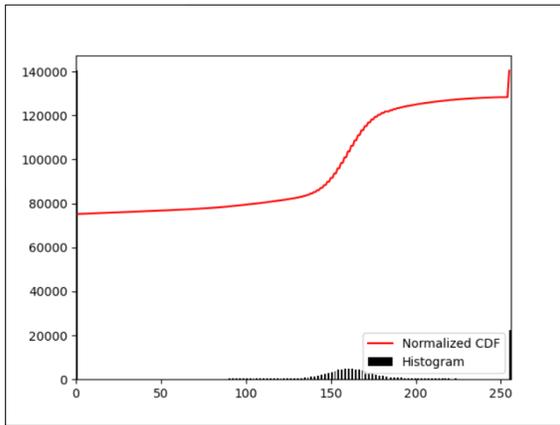


Figure 4. Histogram and CDF of histogram equalized target image

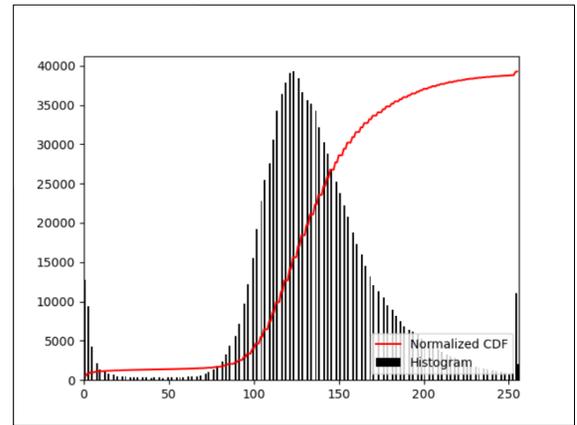


Figure 5. Histogram and CDF of histogram equalized reference image

$$h(v) = \text{round} \left(\frac{cdf(v) - cdf_{min}}{(M \times N) - cdf_{min}} \times (L - 1) \right) \quad \text{Equation 10}$$

⁵ GitHub account: <https://github.com/StefanoD>

⁶ Official website: <https://numpy.org/>

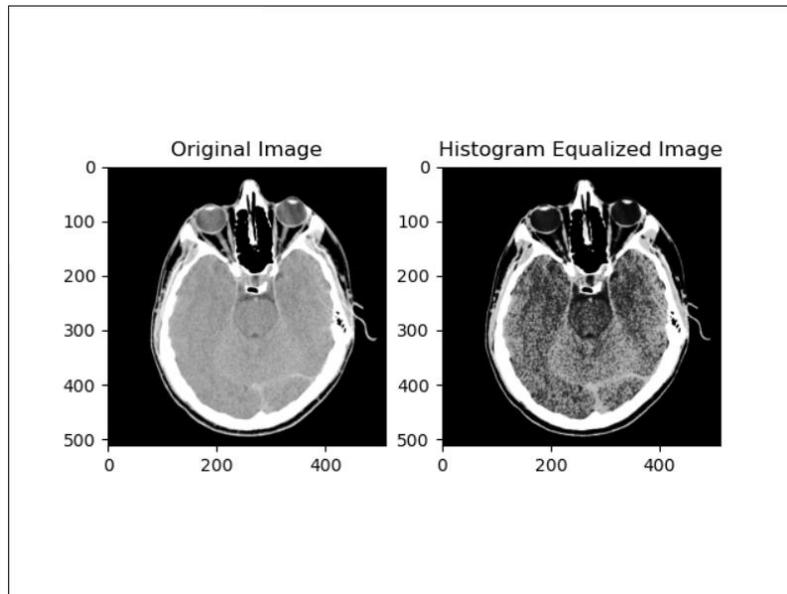


Figure 6. Histogram Equalization histogram and CDF Results

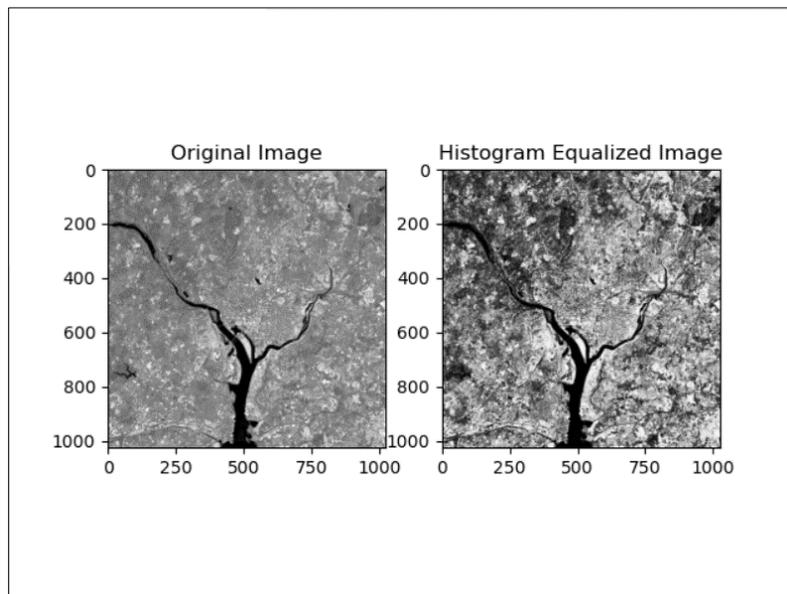


Figure 7. Histogram Equalization Resulted Image

In the next step, the histogram equalized images were analyzed through the exact histogram specification method, implemented by Stefano Di Martino, in contrast to traditional histogram matching algorithms, which only approximate a reference histogram. Exact histogram specification refers to the process of finding a transformation for a specific image, in order for its histogram to match the reference histogram exactly. This method extracts the reference image's histogram and matches the target image to the reference

histogram by using several kernels, which calculate the average of a neighborhood, in order to achieve the exact matching [6][11]. Thereby a pixel can not only be sorted after its value, but also after its average values, in more than one neighborhood. The result of the exact histogram specification of the reference histogram to the target image is presented in Figure 8, as well as the resulted image's histogram. Pseudo-color schemes were applied to these specific image plots, which is a technique that maps each of the gray-levels of a black and white image into a specific color and can enhance the contrast of the image and allow for an easier visualization.

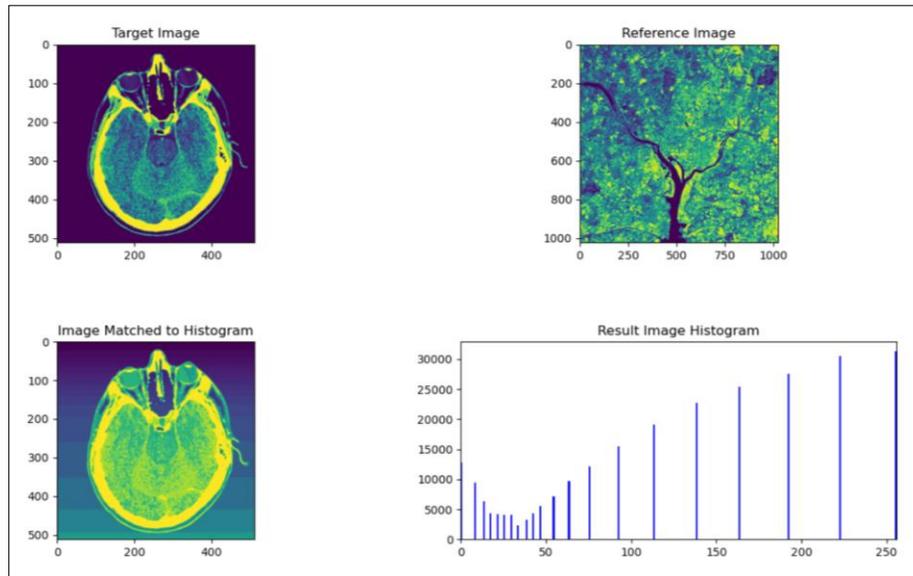


Figure 8. Exact Histogram Specification Results

The second implementation was carried out through MATLAB and Python, while it also includes the main three concepts (histogram equalization, histogram specification, histogram comparison), described above. Initially, the aforementioned target and reference images were imported (Figure 9), and two different MATLAB histogram equalization functions were applied to each image, in order to carry out the process of adjusting intensity values automatically. On the reference image, `histeq()` function⁷ was applied, which basically reinforces the contrast of the images. This is achieved through the transformation of the values in the intensity image or in the indexed image's color-map, such that the output image's histogram essentially relates to the given histogram. For histogram

⁷ Official website: <https://www.mathworks.com/help/images/histogram-equalization.html>

equalization on the target image, `adaphisteq()` function⁸ was applied to improve the contrast of the gray-scale image through the transformation of the values based-on Contrast Limited Adaptive Histogram Equalization (CLAHE)[12]. CLAHE refers to a variation of an image processing technique, called adaptive histogram equalization (AHE), and expunges the excessive noise amplification in relatively homogenous image parts⁹. The results of MATLAB's histogram equalization, on both target and reference images, are displayed in Figure 10. Furthermore, in Figure 11 and Figure 12, the computed CDFs and final histograms of the histogram equalized images are presented. For histogram specification, we used `imhistmatch()` function¹⁰ offered by MATLAB. This function adjusts the histogram of the target image aiming to match the histogram of the reference image. The output image extracted from the histogram matching function is shown in Figure 13.

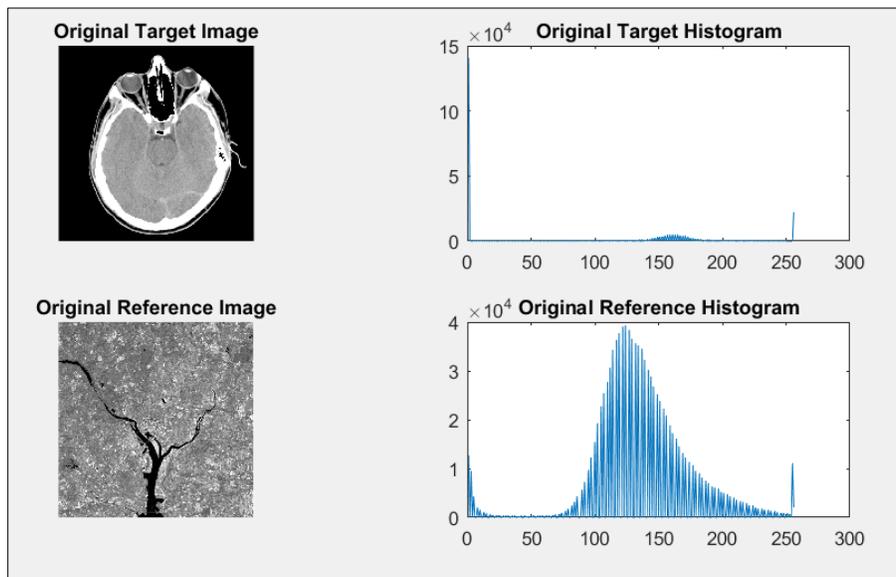


Figure 9. Target and reference images and their histograms

⁸ Official website: <https://www.mathworks.com/help/images/ref/adaphisteq.html>

⁹ Source: https://en.wikipedia.org/wiki/Adaptive_histogram_equalization

¹⁰ Official website: <https://www.mathworks.com/help/images/ref/imhistmatch.html>

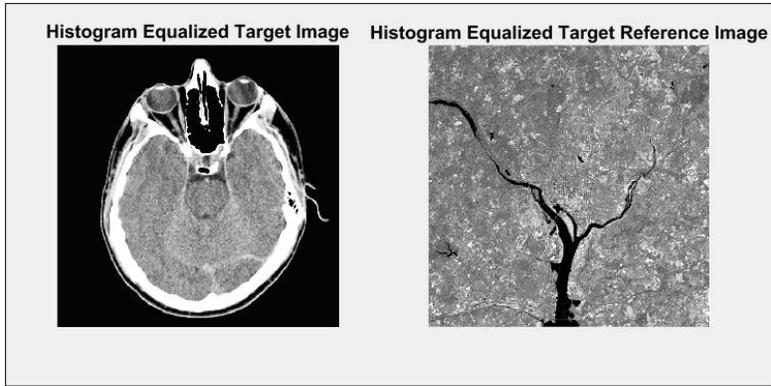


Figure 10. Histogram Equalized images

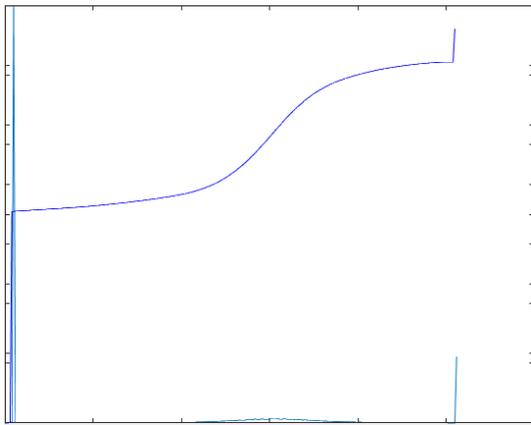


Figure 11. Histogram and CDF of histogram equalized target image

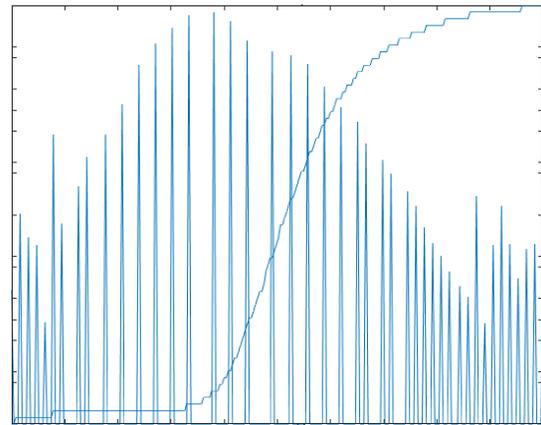


Figure 12. Histogram and CDF of histogram equalized reference image

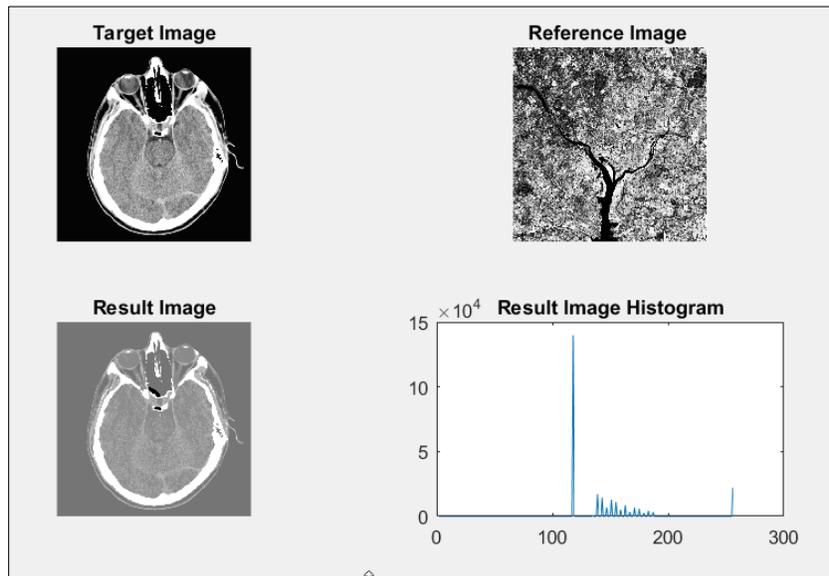


Figure 13. Histogram Matching Results

3.1. HISTOGRAM COMPARISON

An analysis of the methodologies' effectiveness was conducted and presented in this section. Specifically, we implemented a histogram comparison method, which involved

four different histogram distance-based functions. This histogram comparison method was used for both of our methodologies presented in Section 3. However, it was noted that the addition of Gaussian noise can possibly throw off these histogram comparison methods. In the first place, we initialized two dictionaries in our code. The first dictionary was created to store the histograms that would be extracted from the images in our dataset, with the filename set as the key and the histogram set as the value. The second dictionary was created to store the actual images, as well as to present the comparison results. Then, we extracted the histograms from the dataset, which contained all the images that will be a part of the comparison. Specifically, our dataset contained five images: 1) the original brain MRI, which was the target image, 2) an irrelevant image used as a reference for histogram specification, 3) the histogram equalized target MRI, 4) the histogram equalized reference image and 5) the histogram specified MRI. The histogram specified MRI was exported as a result of the histogram specification implementations, but due to the different techniques that were used, the output image was different in each methodology. As we mentioned above, in the first methodology we implemented exact histogram specification, while in the second one we implemented histogram matching. Following, the histogram comparison image (histogram specification output MRI) was then normalized and stored in the first dictionary. The above concludes the histogram computation process for each image in our dataset. The next step revolved around the comparison of the stored histograms. Specifically, four distance-based histogram comparison functions were used. The first function used was Correlation, which computes the correlation between two histograms, as described in Equation 11. The second one was Chi-Square function, which applies the Chi-Squared distance to the histograms based on Equation 12. The third function was Intersection, which calculates the intersection between two histograms, as seen in Equation 13, and finally was Bhattacharyya distance/Hellinger function, which is used to measure the overlap between two histograms and is described in Equation 14. For Correlation and Intersection, the score values are interpreted based on the higher value, which implies a more accurate match. On the other hand, Chi-Squared and Bhattacharyya distance methods are interpreted in the opposite way. In the final aspect of our histogram comparison implementation, we looped through the aforementioned methods, with the use of OpenCV

functions¹¹, and then we defined a third dictionary to store the comparison results, in which the filenames were set as the key, and the similarity score of these methods were set as a value.

$$d(H_1, H_2) = \frac{\sum_I (H_1(I) - \bar{H}_1)(H_2(I) - \bar{H}_2)}{\sqrt{\sum_I (H_1(I) - \bar{H}_1)^2 \sum_I (H_2(I) - \bar{H}_2)^2}} \quad \text{Equation 11}$$

$$d(H_1, H_2) = \sum_I \frac{(H_1(I) - H_2(I))^2}{H_1(I)} \quad \text{Equation 12}$$

$$d(H_1, H_2) = \sum_I \min(H_1(I), H_2(I)) \quad \text{Equation 13}$$

$$d(H_1, H_2) = \sqrt{1 - \frac{1}{\sqrt{H_1 H_2 N^2}} \sum_I \sqrt{H_1(I) \times H_2(I)}} \quad \text{Equation 14}$$

4. RESULTS

The results of the histogram comparison method described above (Section 3.1), which was based on four different histogram distance-based functions, are presented in this section. Both of our proposed implementations went through histogram comparison, aiming to analyze their efficiency in histogram specification. The histogram specification output image (HMres.tif) was compared to all the images in our dataset, based on histogram similarities. First, for the initial (Python) implementation, Correlation function revealed that the compared image's histogram scored 0.93 in similarity (Figure 14), in comparison to the histogram of the histogram equalized reference image (HErefer.tif). On the other hand, when compared to the equalized target image (HEtarget.tif) and the original target image (Fig1.tif), Correlation function indicated a score of 0.63. Finally, for the original reference image (Fig2.tif), Correlation score reached 0.47 in histogram similarity. As we mentioned before, in Chi-Squared function the lower the score the higher the similarity. Thus, in this case, Chi-Squared results revealed that the higher histogram similarity score

¹¹ Official website: <https://opencv.org/>

was reached with the histogram equalized reference image, with a score value of 0.45, while the second lower score was with the equalized target image with a value of 3.39. The lowest histogram similarities occurred when compared to the original images (target and reference), with score values of 1.72 and 3.51 respectively (Figure 15). Proceeding to Intersection, which is interpreted like Correlation function, the results once more indicated that the histogram equalized reference image was the most similar with a score value of 2.30, while the histogram of the original reference image was the least similar by reaching a value of 0.95 similarity (Figure 16). A slight variation was noted in the pattern of the distance-based functions' results, specifically in Intersection, where the original target image was placed higher in similarity than the histogram equalized target image. Finally, the implementation of Bhattacharyya/Hellinger, which is interpreted like Chi-Squared function, also confirmed that the highest histogram similarity was found in the histogram equalized reference image, while the results were respective to the ones presented by Correlation and Chi-Squared functions (Figure 17).

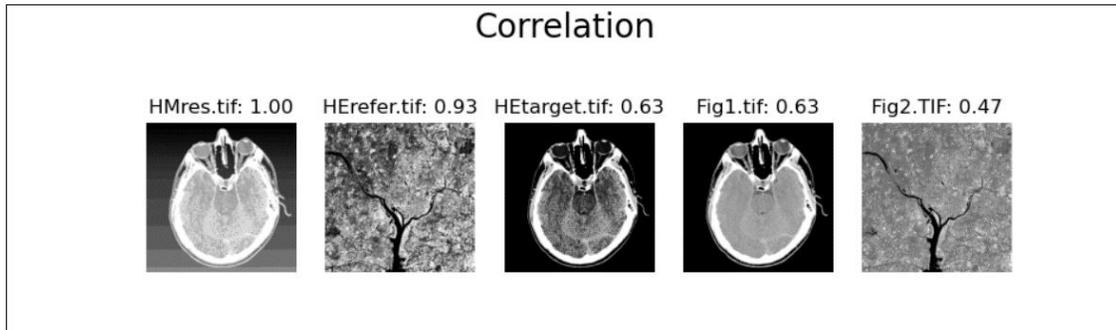


Figure 14. Scores of the Correlation function in implementation #1

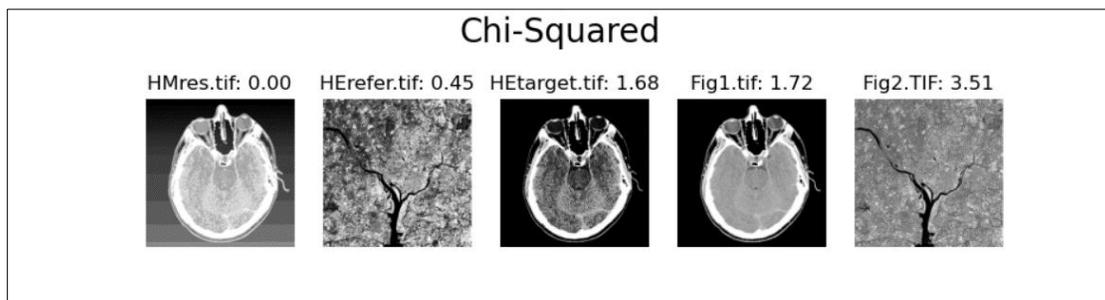


Figure 15. Scores of the Chi-Squared function in implementation #1

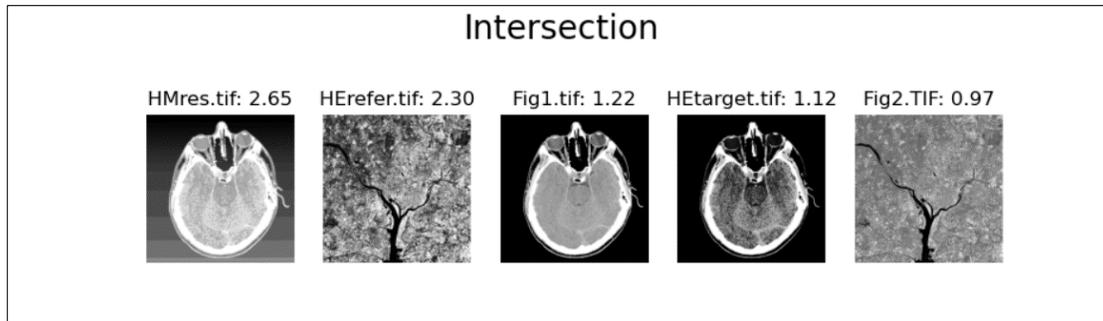


Figure 16. Scores of the Intersection function in implementation #1

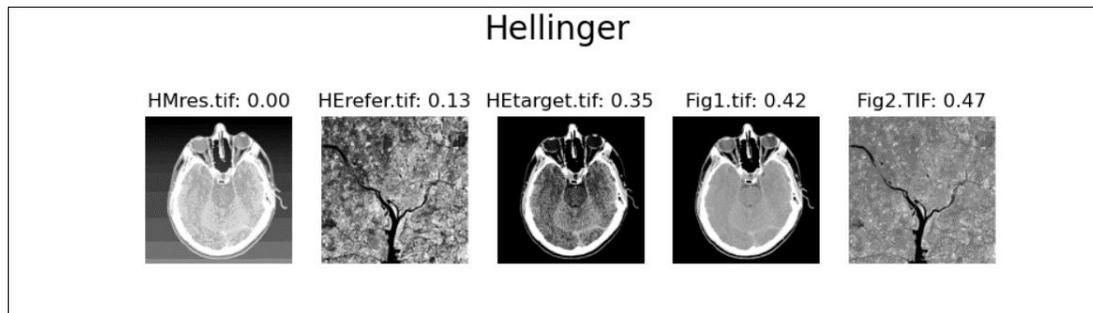


Figure 17. Scores of the Bhattacharyya/Hellinger distance function in implementation #1

Histogram comparison was conducted for the MATLAB implementation as well, where the extracted histogram matched image was a different one, due to the differences in the techniques used in each implementation. In terms of the histogram comparison of the images extracted and used in the second (MATLAB) implementation, Correlation function indicated that the histogram equalized reference image was the most similar to the histogram matched image (Figure 18). Overall, the same results were extracted through all four of the distance-based histogram comparison functions that were used. They all confirmed that the highest similarity was found in the histogram equalized reference image that was used as the reference in the histogram matching process (Figure 20 & Figure 21). Images that had lower similarity scores mostly retained the same similarity results, besides some slight variations in Chi-Squared function, which determined that the least similar image was the histogram equalized target image (Figure 19), instead of the original target image (Figure 18, Figure 20 & Figure 21).

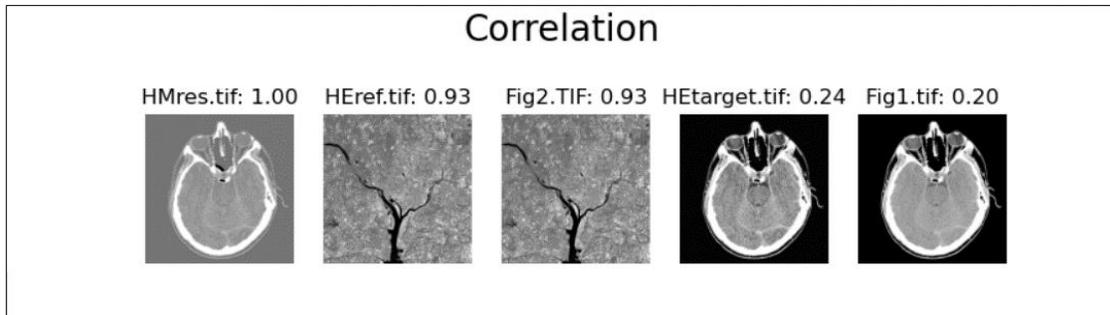


Figure 18. Scores of the Correlation function in implementation #2

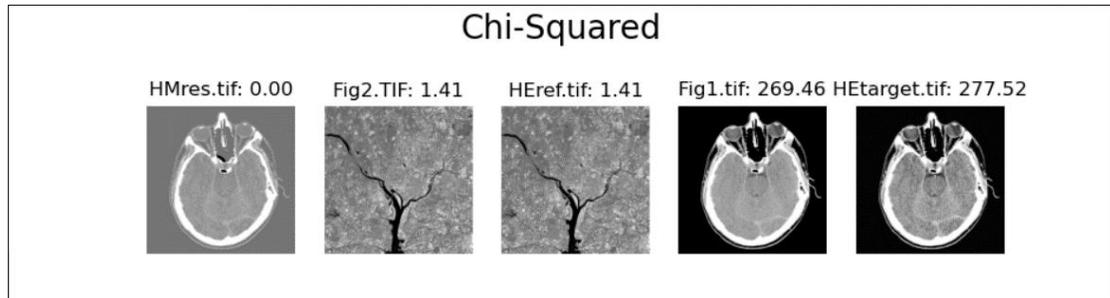


Figure 19. Scores of the Chi-Squared function in implementation #2

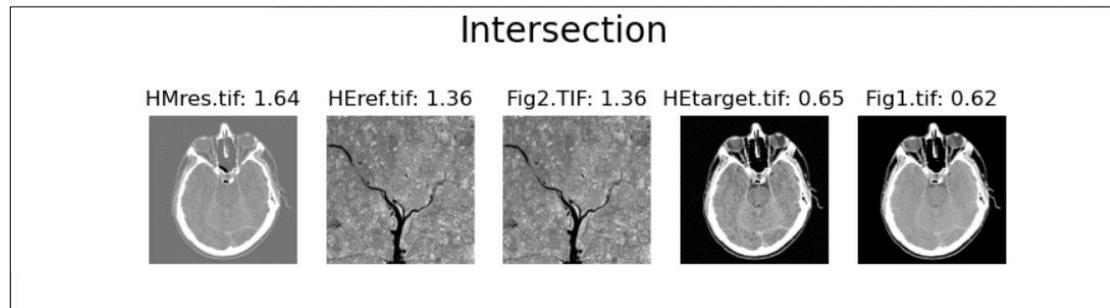


Figure 20. Scores of the Intersection function in implementation #2

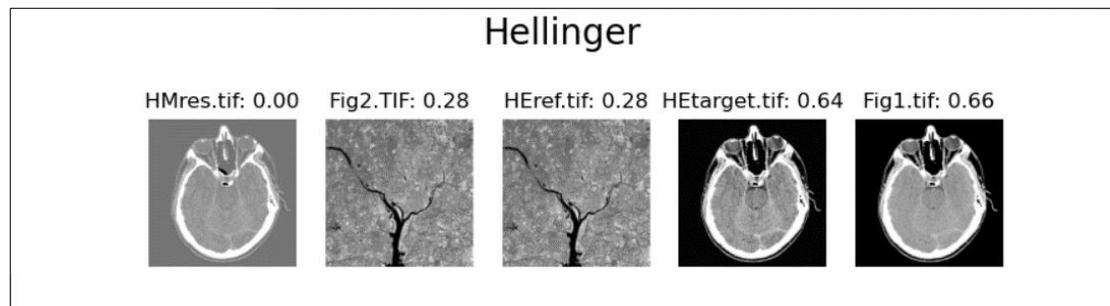


Figure 21. Scores of the Bhattacharyya/Hellinger distance function in implementation #2

5. DISCUSSION

The methodologies proposed in this paper consisted of three main components that were implemented in different ways. The first one, which refers to image's histogram equalization, was implemented to adjust the contrast of the image by using its histogram, while aiming to retain most of the information. Overall, the results of the histogram

equalization process on both methodologies revealed that the contrast was successfully adjusted, and the resulted images were clearly enhanced in terms of lighting conditions. Although histogram equalization was successfully implemented in both methodologies, there were differences in lighting and intensity between the output of MATLAB's default histogram equalization functions and the one we created in Python through mapping transformation. The second component describes histogram matching/specification. In Python implementation, an exact histogram specification method was implemented through the concept of an ordering relationship and involved a quite direct ordering of the pixels related to the image. After this ordering process, pixels were automatically divided into groups and were set to a specific gray-level. This ordering approach, established by Coltuc et al. [6] and implemented by Stefano Di Martino¹², complies with the natural ordering, which allows for the image's information to be mostly retained. Specifically, Coltuc et al. [6] stated that through their approach "*An ill-posed problem, exact histogram specification, is solved*". In MATLAB implementation, we used the default histogram matching function and generated a histogram-matched output image through relating target and reference histograms with the default number of bins (default = 64). Finally, a common histogram comparison approach, which involved four different distance functions, was implemented for both approaches to compare the histograms of the images contained in our dataset and measure their similarity. Similarity scores, in all four implemented functions, indicated a high similarity between the histograms of the compared image and the reference image, which was histogram equalized before histogram specification. On the other hand, the comparison revealed lower similarity scores, when the histogram of the compared image was contrasted to the histograms of the original images. In conclusion, the proposed approaches of histogram equalization, followed by (exact/not exact) histogram specification and ordering, and distance-based histogram comparison appears encouraging and relevant to the image processing world. Furthermore, it can be useful for medical image enhancement use cases (e.g. X-Rays enhancement and analysis), as well as for other image processing tasks, through some future modifications.

¹² GitHub account: <https://github.com/StefanoD>

6. REFERENCES

- [1] L. G. Nyúl, J. K. Udupa, and X. Zhang, “New variants of a method of MRI scale standardization,” *IEEE Trans. Med. Imaging*, vol. 19, no. 2, pp. 143–150, 2000.
- [2] X. Chen *et al.*, “Automatic Histogram Specification for Glioma Grading Using Multicenter Data,” *J. Healthc. Eng.*, vol. 2019, 2019.
- [3] D. S. Wishart, “Chapter 3: Small Molecules and Disease,” *PLoS Comput. Biol.*, vol. 8, no. 12, 2012.
- [4] M. Shah *et al.*, “Evaluating intensity normalization on MRIs of human brain with multiple sclerosis,” *Med. Image Anal.*, vol. 15, no. 2, pp. 267–282, 2011.
- [5] G. Thomas, D. Flores-Tapia, and S. Pistorius, “Histogram specification: A fast and flexible method to process digital images,” *IEEE Trans. Instrum. Meas.*, vol. 60, no. 5, pp. 1565–1578, 2011.
- [6] D. Coltuc, P. Bolon, and J. M. Chassery, “Exact histogram specification,” *IEEE Trans. Image Process.*, vol. 15, no. 5, pp. 1143–1152, 2006.
- [7] R. E. Woods and R. C. Gonzalez, *Digital image processing for visual prosthesis: Filtering implications*. 2011.
- [8] X. Sun *et al.*, “Histogram-based normalization technique on human brain magnetic resonance images from different acquisitions,” *Biomed. Eng. Online*, vol. 14, no. 1, pp. 1–17, 2015.
- [9] A. Brahim, J. Ramírez, J. M. Górriz, L. Khedher, and D. Salas-Gonzalez, “Erratum: Comparison between different intensity normalization methods in 123iioflupane imaging for the automatic detection of Parkinsonism (PLoS ONE 10:6 (e0130274) doi 10.1371/journal.pone.0130274),” *PLoS One*, vol. 10, no. 7, pp. 1–20, 2015.
- [10] S. Francisco, “PATCH BASED INTENSITY NORMALIZATION OF BRAIN MR IMAGES Sneathis Roy , Aaron Carass , and Jerry L . Prince Image Analysis and Communications Laboratory , Electrical and Computer Engineering , The Johns Hopkins University,” 2013.
- [11] S. W. Jung, “Two-dimensional histogram specification using two-dimensional cumulative distribution function,” *Electron. Lett.*, vol. 50, no. 12, pp. 872–874, 2014.
- [12] K. Zuiderveld, *Contrast Limited Adaptive Histogram Equalization*. Academic Press, Inc., 1994.