Automated Image Quality Assessment in Medical Images

Faranak Dayyani (500692664), Sarina Taki (500773183)

Abstract - This paper studies the various image quality metrics in order to obtain an automated image quality assessment tool for medical images. The three image quality metrics under study in this paper include noisiness of the image, contrast level as well as edge assessment. Noisiness refers to the amount of noise that exists in an image. Some noise levels whether high or low, could be difficult to detect manually which is why this tool measures it automatically. This could be specifically more challenging to do manually for medical images such as CT Scans and MRIs. Contrast level refers to the resolution of an image. Lastly, edge assessment can detect whether edges are fuzzy or sharp. These three metrics determine the image quality in this paper. Image quality is an important factor to take into account when performing image processing since it can help with design decisions [1].

I. INTRODUCTION

The purpose of this project is to measure image quality automatically using three different image quality metrics. Image quality represents the level of accuracy in imaging systems [1]. When dealing with medical images such as CT Scans and MRIs, they can have lots of variability in noise levels, contrast and edge strength [1]. One of the common sources of noise when taking medical images from patients is motion artifacts. The dataset used in this paper are Lung CTs and Brain MRIs.

Noisy-ness (N): Noise always exists in digital images and is added to the image during image acquisition and processing [2]. Image noise is defined as "random variation of brightness or color information in the images captured" [2]. In order to measure the noisy-ness of an image, multiple approaches were studied. From these approaches, peak signal-to-noise ratio (PSNR) was chosen for implementation. PSNR is the ratio between the maximum possible value of a signal and the distorting noise in an image [3]. PSNR is used to measure image quality for reconstructed images that have been previously compressed [4]. The unit for PSNR is in logarithmic decibel scale [3]. The mathematical representation of peak signal-to-noise ratio, PSNR can be seen below:

$$PSNR = 20log_{10}\left(\frac{MAX_f}{\sqrt{MSE}}\right) \qquad \text{Equation 1[3]}$$

In equation 1 above, MAX_f is the maximum intensity value of an image and MSE is the Mean Squared Error. The equation for Mean Squared Error can be seen below:

$$MSE = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}||f(i,j) - g(i,j)||^2 \qquad \text{Equation 2[3]}$$

In equation 2 above, MSE is the Mean Squared Error between two images, f(i,j) represents the matrix information for the original image, g(i,j) represents the matrix information for the noisy image. M and N represent the number of rows and columns in the image [3]. It can be described as "mean of the square of the differences in the pixel values between the corresponding pixels of the two images" [4].

Contrast (C): Contrast measurement is a well known quality for measuring Image Quality. Several methods for contrast measurement were studied and among Weber, Michelson and RMS that were most commonly used [5], RMS (root-mean-square) gave similar results as the other methods but had the advantage of simplicity. Implementation of root-mean-square (RMS) is simple and easy to understand. RMS is a statistical measure to help identify certain patterns in an image. The RMS of an image can be calculated using the equation below:

$$RMS = \sqrt{\frac{1}{MN}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}(I_{ij} - \bar{I})^2} \qquad \text{Equation 3 [6]}$$

In equation 3 above, M and N are size of the image. I_{ij} is the i-th & j-th elements of the 2D image. $\bar{I}$ is the average intensity of all pixel values in an image.

Edge Quality (E): The assessment for quality of edges in an image is an important topic in image processing as it helps determine the performance of edge detectors [7]. For performing edge quality assessment, various methods were taken into consideration such as measuring MSE (mean squared error) and SSIM (Structural similarity index for measuring image quality) [7]. Although all these methods were successful, they were either used previously in this paper or are too complex. The method that was chosen for measuring edge quality is an algorithm/function which was proven to be very successful after going through trial and error [8]. The input of this algorithm is a single image I with size MxN. The output is FM (Frequency Domain Image Blur Measure) [8]. The steps of this algorithm is explained below [8]:
Step 1: Computing the Fourier Transform representation (F) of image I
Step 2: Shifting the origin of F to centre (Fc)
Step 3: Obtaining the absolute value of Fc → AF=abs(Fc)
Step 4: Calculating the maximum value of the frequency component in F → M = max(AF)
Step 5: Setting threshold to → thres = M/1000 , M is value calculated in step 4
Step 6: Calculating the total number of pixels in F whose pixel values are greater than thres(T_H)
Step 7: Obtain FM from the equation below:

$$Image\ Quality\ Measure\ (FM) = \frac{T_H}{M \times N} \qquad \text{Equation 4[8]}$$

In equation 4 above, M and N are the dimensions of the input image I.

II. Methods

As it is mentioned earlier in the introduction, we measured PSNR, RMS and image sharpness of the image in order to justify the quality of noise, contrast and edge of the image as per slice or volume respectively [3], [10], [8].

In order to justify the quality of the noise of the image, Peak Signal-to-Noise Ratio (PSNR) of the image was measured in spatial domain [3]. Firstly, the image was converted to a double image since the type of most of the medical images is unsigned or signed integers [1]. It is worth to note that, double type images have the same pixel value as the unsigned or signed type of the image, but their pixels can be floating values as well [1]. Then after the image type was changed to double, the image was filtered by a built-in filter function in MATLAB called *imnlmfilt* which is a Non-Local-Means (NLM) filter. NLM is a low pass filter which removes noise by taking the average of the similar pixels/voxels based on the distance of their intensity [9]. Filtered image was used as a reference image to calculate Mean Square Error (MSE) of the input image by using equation 2 and the result of the MSE was used to measure the PSNR of the input image by using equation 1. According to National Instruments, 8-bit images with a PSNR value between 30dB to 50dB are not considered as noisy images, thus the filtered image and input image were converted to 8-bit signed integer before calculating the MSE and the PSNR of the images [3].

In order to justify the quality of the contrast of the image, Root Mean Square (RMS) of the image was measured in spatial domain [10]. Firstly, the image was converted to a double image, then the image was normalized to avoid getting RMS values out of the range of 0 to 1. At the end, the RMS of the normalized image was measured by using equation 3 and if the RMS value was within a specific interval, the image would have been considered as a high contrast image.

It is worth to note that, RMS value for a Lung CT image was low, which meant the image was considered as a dark image, however by observing the image and its histogram we concluded that the image was not a dark image but since Lung CT image had a large black area, the algorithm would have considered it a black image, therefore we came up with this idea to eliminate the pixels with zero intensity to avoid the large black area and as the result the accuracy of the algorithm improved.

In order to come up with a range for the RMS, which determines the boundary between high contrast image and dark or bright contrast image, RMS values of 78 images were measured and a box plot was plotted based on the obtained RMSs. According to the box plot, 25^{th} and 75^{th} percentiles of the box plot were assigned as the boundaries. Images with the RMS values between 0.46291 and 0.5537 were considered as high contrast images and images with the RMS values below the interval or above the interval were considered as dark or bright images, respectively. It is worth noting that the images were normalized since they had different ranges of pixel values which would lead to a different range of RMS values, therefore, it would be difficult to compare the RMS values of different images.

In order to justify the quality of the edge of the image, image sharpness was measured in the frequency domain [8]. Firstly, the image was converted to a double image, then the image was normalized to avoid getting values of image sharpness out of the range of 0 to 1. Fourier transform of the normalized image was computed to take the image in frequency domain by using Fast Fourier Transform (FFT) function of the MATLAB, then the frequency components were shifted to the center of the spectrum to obtain a frequency spectrum which contains one full contagious period [8]. After, the absolute value of the shifted spectrum was computed to obtain the magnitude frequency of the image, then the maximum value of the magnitude frequency was calculated which would be used to compute the threshold value by dividing it with 1000 [8]. Later, the number of pixels with intensity greater than the threshold value in the non-shifted frequency spectrum of the image was calculated [8]. At the end, the image sharpness of the image was measured by using equation 4 in which the total number of the pixels, which have the intensity greater than the threshold, was divided with the size of the input image [8].

In order to come up with a threshold for the image sharpness, which determines the boundary between strong edge image and weak edge image, image sharpness values of 78 images were measured and a box plot was plotted based on the obtained image sharpness values. According to the box plot, the median of the box plot was assigned as the threshold. Images with the image sharpness values greater than 0.014922 were considered as strong edge images and images with the image sharpness values less than 0.014922 were considered as weak edge images. It is worth noting that the images were normalized since they had different range of pixel values which would lead to a different range of image sharpness values, therefore, it would be difficult to compare the image sharpness values of different images.

III. Results

Table 1: This table shows the values of PSNR, RMS and image sharpness for both Lung CT slice and Brain MRI volume images.

	PSNR (dB)	RMS	Image Sharpness
Lung CT (slice)	48.8718	0.4372	0.0201
Brain MRI (volume)	50.6745	0.5386	0.0133

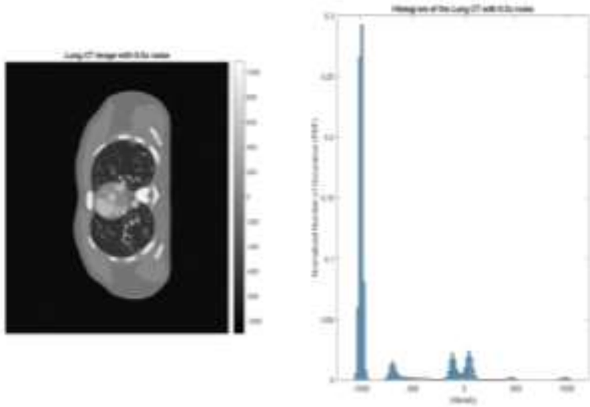

Figure 1: The image and histogram of a Lung CT slice with 0.5x noise for noise measurement

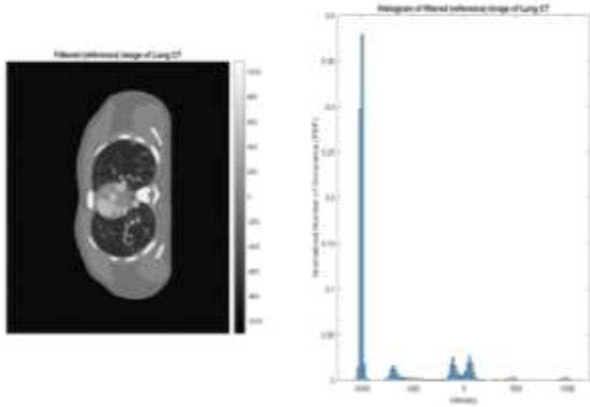

Figure 2: The image and histogram of the reference Lung CT slice for noise measurement

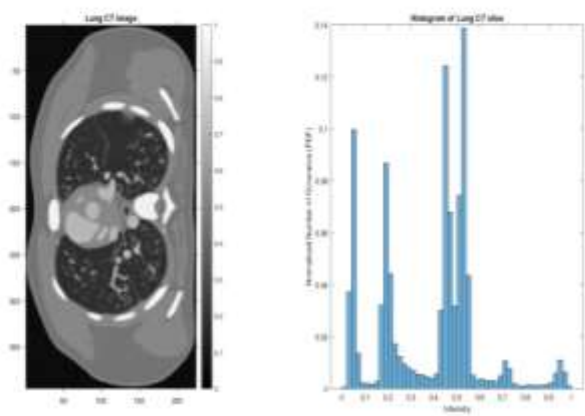

Figure 3: The image and histogram of the cropped Lung CT slice for contrast measurement

Figure 4: The magnitude frequency of the Lung CT slice for edge measurement

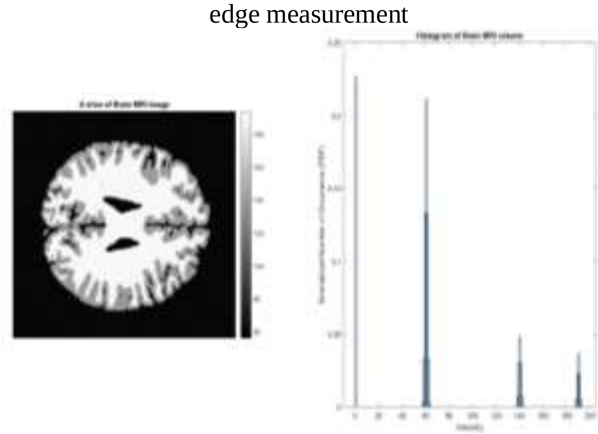

Figure 5: The image and histogram of a Brain MRI volume for noise and contrast measurement

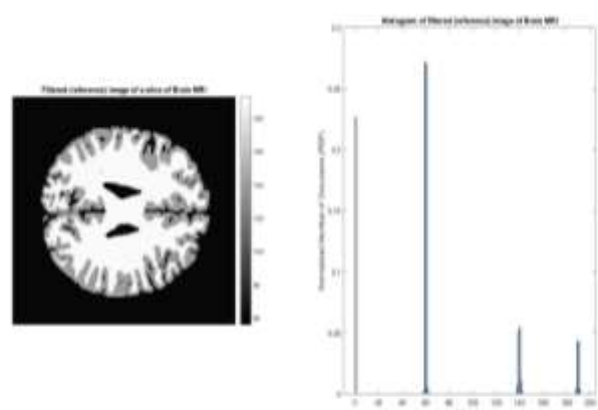

Figure 6: The image and histogram of the reference Brain MRI volume for noise measurement

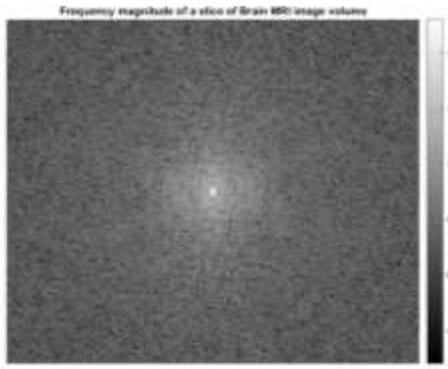

Figure 7: The magnitude frequency of the Brain MRI volume for edge measurement

IV. Discussion

Image quality metrics represent the level of accuracy in imaging systems [1]. Three different image quality metrics were used in order to evaluate Image quality in this project. These three metrics are noisy-ness (N), contrast (C), and edge quality (E). The dataset used were lung CT scans and Brain MRIs. One slice of a Lung CT and the volume of a Brain MRI were chosen for testing.

The PSNR method was used to measure the noisy-ness of our dataset. The range known to be of low amount of noise for PSNR values are between 30-50 dB. The values for 1 slice of a Lung CT and the entire volume of a Brain MRI, were 48.8718 and 50.6745, respectively. The value for Lung CT indicates that it is of good quality of noise, meaning it is not noisy. However, the Brain MRI is slightly above the 50dB point which gets categorized as noisy.

The RMS method was chosen to measure the contrast level of an image. After measuring the RMS value randomly for 78 images, a boundary region for a high contrast image was obtained. Images that have RMSs between 0.46291 and 0.5537 are categorized as high contrast. If the RMS is below 0.46291, it is considered a dark image and if the RMS is above 0.5537, it is categorized as a bright image. According to table 1 above, the Lung CT slice was of low contrast and has more dark regions than bright. This can be seen in figure 3. The histogram of figure 3 shows more intensity values on the left side of the graph which indicates a greater dark area in the image. However, the Brain MRI had a value of 0.5386 which means it is of high contrast. This is also evident in figure 5.

The edge quality of the images were measured by measuring the image sharpness of an image in the frequency domain [8]. The threshold value that determines strong and weak edge images is 0.014922. This means that any input image that has an image sharpness below 0.014922 is considered a weak edge image and if above 0.014922, is categorized as a strong edge image. By looking at the results in table 1, the Lung CT is of strong edge quality and the Brain MRI is of weak edge quality. These values make sense since the Brain MRI chosen for testing is known to be noisy and also has a weak edge quality. These two metrics go hand in hand when evaluating image quality.

V. Conclusion

In conclusion, the purpose of this experiment is to measure the image quality based on three image quality metrics such as image noisy-ness, contrast of the image and edge quality of the image. In order to measure each one of the image quality metrics, multiple approaches were studied which among all of them we used PSNR, RMS and FM to compute the noisy-ness, contrast and edge quality of an image, respectively. PSNR and RMS are statistical measurements which were used to compute the maximum signal to noise ratio and to identify certain patterns in an image, respectively [3] [5]. FM is an algorithm to compute edge quality by finding the total number of pixels with intensity values greater than a certain threshold in the magnitude frequency of the image [8]. According to the results, a slice of Lung CT image is considered as a not noisy image which is a dark contrast image and has strong edges, and a volume of Brain MRI image is considered as a noisy image which is a high contrast image and has weak edges.

References

[1] BME 872 Biomedical Image Analysis, "Project: Automated Image Quality Assessment in Medical Images", Ryerson University, Toronto, ON, rep., Winter 2020.

[2] A. Swain, "*Noise in Digital Image Processing*", 24-Nov-2019. [Online]. Available: https://medium.com/image-vision/noise-in-digital-image-processing-55357c9fab71

[3] National Instruments, "*Peak Signal-to-Noise Ratio as an Image Quality Metric*", 5-Mar-2019. [Online]. Available: https://www.ni.com/en-ca/innovations/white-papers/11/peak-signal-to-noise-ratio-as-an-image-quality-metric.html

[4] A. Robinson, "*How to Calculate PSNR*", 2-Mar-2019. [Online]. Available: https://sciencing.com/how-to-calculate-psnr-12751768.html

[5] A. Shaus, S. F. Golovin, B. Sober, E. Turkel, & E. Piasetzky, (2017). "Potential Contrast – A New Image Quality Measure", *Image Quality and System Performance XIV*, *14*, 52–58, 2017. [Online]. Available: https://pdfs.semanticscholar.org/e3f2/e337d4470545398cc6753a54c21debf9c37b.pdf

[6] R. Gholipour-Peyvandi, S.Z. Islami-Rad, R. Heshmati, S. Zaferanlouie, & M. Ghannadi-Maragheh, "Influence of gamma emitter source intensity against energy on the image RMS contrast". *The Egyptian Journal of Radiology and Nuclear Medicine*, *42*(0378-603X), 425–431. 2011. doi: https://doi.org/10.1016/j.ejrnm.2011.10.003

[7] D. Sadykova, & A. P. James, "*Quality Assessment Metrics for Edge Detection and Edge-aware Filtering: A Tutorial Review*", 1-Jan-2018. [Online]. Available: https://arxiv.org/pdf/1801.00454.pdf

[8] D. Kanjar, & V. Masilamani, "Image Sharpness Measure for Blurred Images in Frequency Domain". *Procedia Engineering*, *64*(18777058), 149–158. 2013. doi: 10.1016/j.proeng.2013.09.086

[9] S. Saladi, N. Amutha Prabha, "Analysis of denoising filters on MRI brain images". *Int. J. Imaging Syst. Technol*. 27: 201– 208. 17-Aug-2017 https://doi-org.ezproxy.lib.ryerson.ca/10.1002/ima.22225

[10] D. G. Pelli, & P. Bex, "Measuring contrast sensitivity". *Vision Research*, *90*(0042-6989), 10–14. 213. doi: https://doi.org/10.1016/j.visres.2013.04.015