Advances in Digital Imaging and Computer Vision

HISTOGRAM MATCHING/SPECIFICATION FOR MRI/CT NORMALIZATION

Ioannis Marinakis

HELLENIC MEDITERRANEAN UNIVERSITY | MTP212@EDU.HMU.GR

Introduction

Histogram matching and specification represent elegant point processing methods for aligning the histogram of an image with either a target histogram or with a histogram of a target image. Such methods can be an excellent starting point for image normalization. This task is particularly interesting for analyzing population data to develop diagnostic or predictive models in medical imaging, face recognition systems or affective computing applications based on images.

In this assignment we are going to conduct a small state-of-the-art literature review in histogram matching/specification for MRI/CT (medical imaging) normalization from published scientific papers and journals. Then we are going to try implement and reproduce the results of one of those advanced histogram matching/specification techniques. In the final step of this assignment, we will calculate the effectiveness of the technique we reproduced, in specific medical data regarding medical imaging (MRI/CT). Finally, we compare the results of this advanced technique with the standard histogram matching techniques that we were taught in Advances in Digital Imaging and Computer Vision Master's class.

State of the art review

For this small state-of-the-art review in histogram matching/specification for MRI/CT (medical imaging) normalization, we filtered the available web literature and we collected 5 very interesting articles that were published in major scientific journals. In this review we are going to explore the main concepts and methods proposed by the authors for histogram matching and specification, without getting into deep details. In these concepts, standard histogram matching techniques are used in combination with more advanced methods to achieve better results and enhance the quality of the medical images. Medical image data of good quality, can be very useful for medical experts and help them be more accurate in the diagnosis of a disease or pathology.

➢ MRI Scale Standardization

This paper [1], was published in IEEE TRANSACTIONS ON MEDICAL IMAGING journal from László G. Nyúl, Jayaram K. Udupa, and Xuan Zhang in year 2000. In their paper, authors state a major drawback of MRI scans which is the lack of a standard and quantifiable interpretation of image intensities. To be more specific:

"MR images taken for the same patient on the same scanner at different times may appear different from each other due to a variety of scanner-dependent variations and, therefore, the absolute intensity values do not have a fixed meaning."

That's a big drawback for MRI scans and its big obstacle to overcome because the data are inconsistent. The authors proposed a two-step method to standardize the MR images and make the extraction of quantitative information about healthy organs or about abnormalities, considerably simplified. Their method consists of 2 steps where in the first step the algorithm cut off the tails of the histogram of the MR image. This happens because the high-intensity tail corresponds to artifacts and outlier intensities and causes considerable inter and intra-patient/scanner variations. In the second step, a transformation is applied to the MR image using parameters that were extracted from the first step. The method is described in the article as it follows:

"In the first (training) step, a set of images of the same body region and protocol corresponding to a population of patients is given as input. The parameters of a histogram transformation are learned from these image data and a few additional input parameters are determined. These additional parameters may vary and are used to fine tune the transformation to the protocol, the body region, and possibly the application that will work on the transformed data. This step needs to be executed only once for a given protocol and for a given body region. In the second (transformation) step, the images are transformed using the parameters learned in the first step. This transformation is image dependent and needs to be done for each given image."

➢ Histogram Normalization based on intensity values

This paper [2] proposes a histogram normalization technique based on the histogram intensity values. It was published in year 2015 at BioMedical Engineering OnLine journal from Xiaofei Sun, Lin Shi, Yishan Luo et al. and the same problem of large intensity variations between scans is faced. They also expose some issues of the previous method that we described above and they comment on it as:

" But this method suffers from unreliable processing results. Because, in order for the existing histogram matching based on a joint histogram to achieve a more reliable implementation, it required a better prior knowledge based of the neighborhoods used to split up the image into K sub-images, which are corrected separately. However, this method relies on a non-rigid registration to match the

histogram, making it considerably slower than approaches which normalize an image without using additional preprocessing steps."

In their work they propose a new histogram normalization method on human brain magnetic resonance images, to reduce the intensity variation between MRIs obtained from different acquisitions. In their experiment, they scanned each subject twice on two different scanners using different imaging parameters. Their method and the algorithm steps are described as:

"The image with lower noise level was determined and treated as the high-quality reference image. Then the histogram of the low-quality image was normalized to the histogram of the high-quality image. The normalization algorithm includes two main steps:

1. *Intensity scaling (IS), where, for the high-quality reference image, the intensities of the image are first rescaled to a range between the low intensity region (LIR) value and the high intensity region (HIR) value;*

2. *Histogram normalization (HN), where the histogram of low-quality image as input image is stretched to match the histogram of the reference image, so that the intensity range in the normalized image will also lie between LIR and HIR."*

At the end of the article, the authors conclude their work with the following lines:

"The method can normalize scans which were acquired on different MRI units. We have validated that the method can greatly improve the image analysis performance. Furthermore, it is demonstrated that with the help of our normalization method, we can create a higher quality Chinese brain template "

➢ **Histogram equalization with tone mapping**

Histogram equalization is a well-known method and it's a special case of histogram matching. In histogram equalization, the specified histogram is uniformly distributed. This paper [3] proposes a weighted Histogram-Based tone mapping algorithm for CT images. It was published in the year 2018 at MDPI open-access journal from David Völgyes, Anne Catrine Trægde Martinsen, et al. A problem of histogram equalization methods, is either the spatially invariant weights that are used or the limited size of kernel due to the complexity of pairwise contribution calculation. In their work, they present a weighted histogram equalization-based

tone mapping algorithm that utilizes the Fast Fourier Transform for distance-dependent contribution calculation and distance-based weights.

The goal of this work is to achieve good local contract in the medical image. In their problem statement, authors explain the importance of good local contrast in medical imaging as:

"Good local contrast has a very important role in computed tomography (CT) images used for diagnosing pathologies. Unlike traditional photographs, CT images contain a measurement of material densities, and these are unaffected by irradiation. Traditional tone mapping assumes that the image is the product of illumination and reflectance."

Another problem the state is the use of a global histogram equalization would not yield good enough local contrast, while local methods might yield unwanted halos. That happens because the CT images have a huge dynamic range.

Considering all the above, they describe their main goal as:

"Our main goal is to develop an algorithm which is able to compress high dynamic range images into a low dynamic range while presenting as much local information as possible, preserving the main structures, not exhibiting strong halos, and more or less keeping pixel intensity ordering"

So, after stating all these known problems of high dynamic range images and local contracts, they built their approach based on the following ideas:

o Local neighborhood is important in order to determine a given pixel's intensity.
o Neighborhood should not have a strong cut-off; the weighted contribution of the whole image should be taken into account.
o The contribution is a decreasing function of the distance.
o The contribution can be calculated using a Fast Fourier transform (FFT).
o The intensity of the pixel is determined based on the local relative intensities in the source image.
o Locality and noise tolerance are equally important.

The extended proposed method, the theories, and mathematic formulas behind it are detailly described in their article. Finally, the authors conclude their paper with the following lines:

"Our proposed method yields good local contrast for CT images while maintaining a similar image structure to the reference CT image. This could contribute to improving the visualization of pathologies. The proposed method performs well in terms of structural similarity compared to popular tone mapping algorithms. The computation cost can be effectively reduced with approximations."

➢ Automatic Histogram Specification

This method [4] proposes histogram specification with an automatic selection of reference frames for magnetic resonance images to alleviate the problem of contrast inconsistencies among multicenter data. The selection of reference frames is automatically performed by an optimized grid search strategy with coarse and fine search. This research article is on a study for glioma, a fatal brain disease.

The main proposes of this study are:

1. *This paper proposes a histogram specification with automatic selection of reference frames (HSASR). This method can automatically select the suitable reference frame of histogram instead of radiologists during image enhancement, so as to enhance the consistency of brain tumors image contrast.*

2. *This paper proposes a set of image standardization algorithms to make the preprocessed multicenter data have better consistency, improve the adaptability of data, and improve the accuracy of the glioma prediction model. "*

At the preprocessing step of this method, histogram specification is conducted to The MRI data, among other preprocessing steps like format conversion and resampling. Then, they apply feature engineering to the data, in order to extract statistical and texture features from the image. The next step is to import those data into the model script of this study and make a prediction.

"*The model script is divided into two parts, namely, feature selection and classifier. Five selection methods and nine classifiers are combined to set up the classifier, and the training and testing sets are imported into the model for prediction.*"

Finally, the authors conclude their article with the following lines:

"*The core of histogram specification is to change the local brightness of the image according to the reference frame, but the reference frame of traditional histogram specification is usually manually selected by radiologists. This method not only increases the workload of radiologists, but also cannot guarantee the optimal reference frame.*"

The authors state the importance of histogram specification and the suitability of the reference frame. With this automatic histogram specification

method that they propose, by using calculating the features and using a series of sophisticated classifiers, this method is able to select the most suitable reference frame.

Methodology

For the purposes of this assignment, we are going to reproduce the results of this paper [3]. Their approach to this problem is really interesting and produces much better results than the standard histogram matching methodologies. In their solution, they are focusing on the enhancement of the local contracts which have a very important role in computed tomography (CT). We described the methodology and the main concept of this technique in our state-of-the-art review. Based on these methods and theories, a core algorithm was created and it can be summarized as:

1. read data → I0
2. reduce bit depth with dithering → I
3. generate W(r) weight array
4. loop over pixel values (z)
   – ID(x, y, z) = δ(I(x, y), z)
   ∗ use superpixels, if downscaling is required
   – convolve ID in x,y plane with W in order to get H
5. clip H peaks
6. redistribute clipped areas along the z-axis
7. determine local intensity from the local histogram and the original image
   – use bilinear interpolation, if superpixels were defined
8. convert the final result from float to integer with Floyd–Steinberg dithering

Based on this algorithm we are going to attempt to reproduce the results of this study and test it with our medical image data. One of the main reasons I selected to develop this method is because this algorithm doesn't require to select manually a reference frame for these medical image data, which in my opinion is very hard to select even for experienced experts.

Implementation

The source code for this implementation [3] is provided at GitHub, by the article authors. Those algorithms are implemented in Python3.6 using NumPy and SciPy. This algorithm can be also used as a library.

The calculations were performed on a notebook that is equipped with an Intel Core i7, 4710MQ @ 2.50GHz, and 24GB of DDR3 RAM. The operating system on this

machine was Ubuntu 18.04, 64bit and all the libraries and requirements to run this project were installed in an Anaconda environment. The DICOM files we used to extract the medical images were provided by the class professor. For the conversion of the DICOM (.npy) files to .png format, we used a Python script and the cv2 library. Plus, we wrote some extra scripts in Python with the Matplotlib library to visualize our results better. Octave is also used for the analysis of the effectiveness of both standard and the techniques we presented and to calculate the mean square error between source and target images. The octave scripts, original and processed images are contained in a folder attached with this assignment.

Results and Figures

In the next pages of this assignment, we present some of the results of the reproduced tone mapping - histogram equalization algorithm. The following figures visualize the original MR Images on the left side and the processed images are on the right side.

Then, by using Octave, we compare this proposed histogram equalization with the standard histogram equalization function of Octave, histeq. We also plot the histograms of the images and in the last step, we calculate the mean square error of each method calculating the effectiveness of the original and the target image.
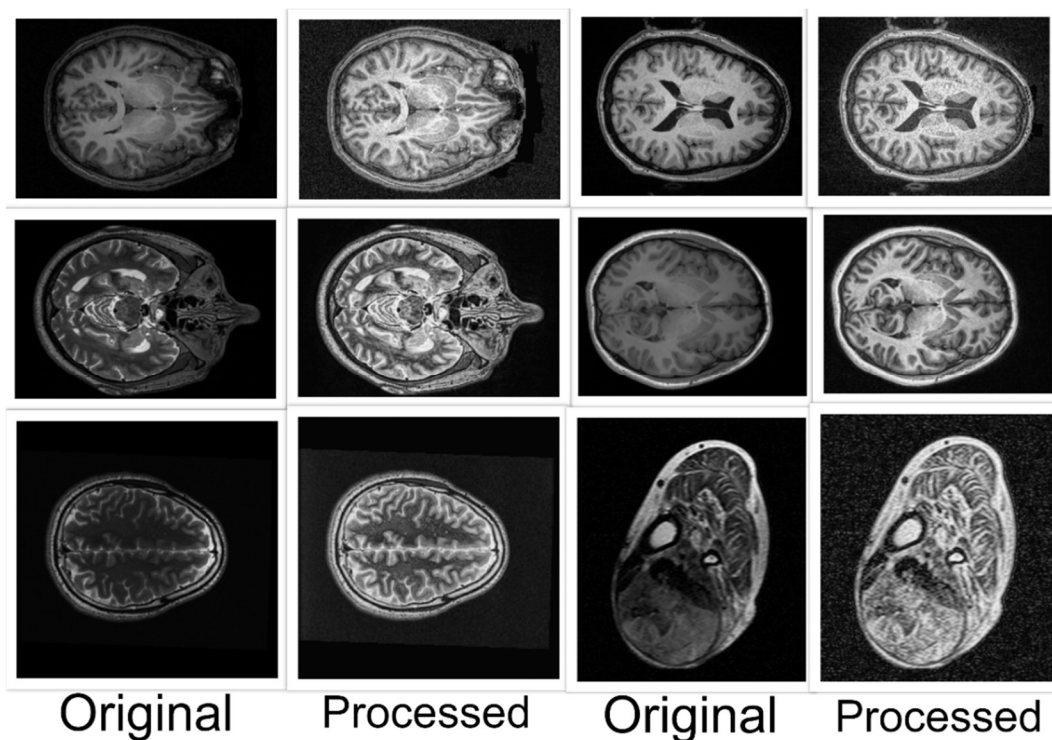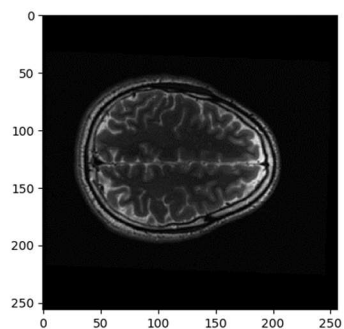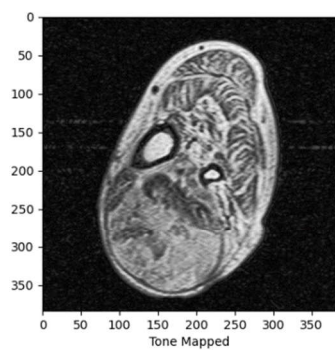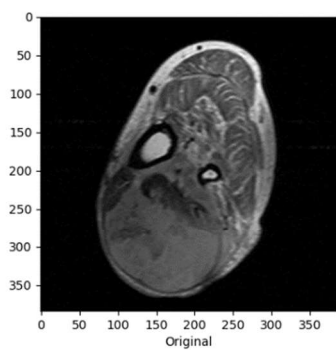
Tone Mapping Algorithm results (Original-Processed)

Original

Tone Mapped

Original

Tone Mapped

Original

Tone Mapped

Original

Tone Mapped

Original
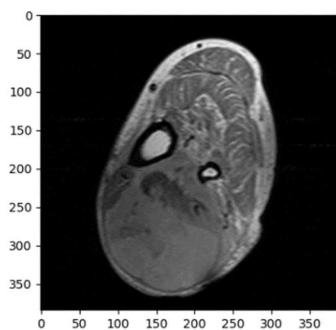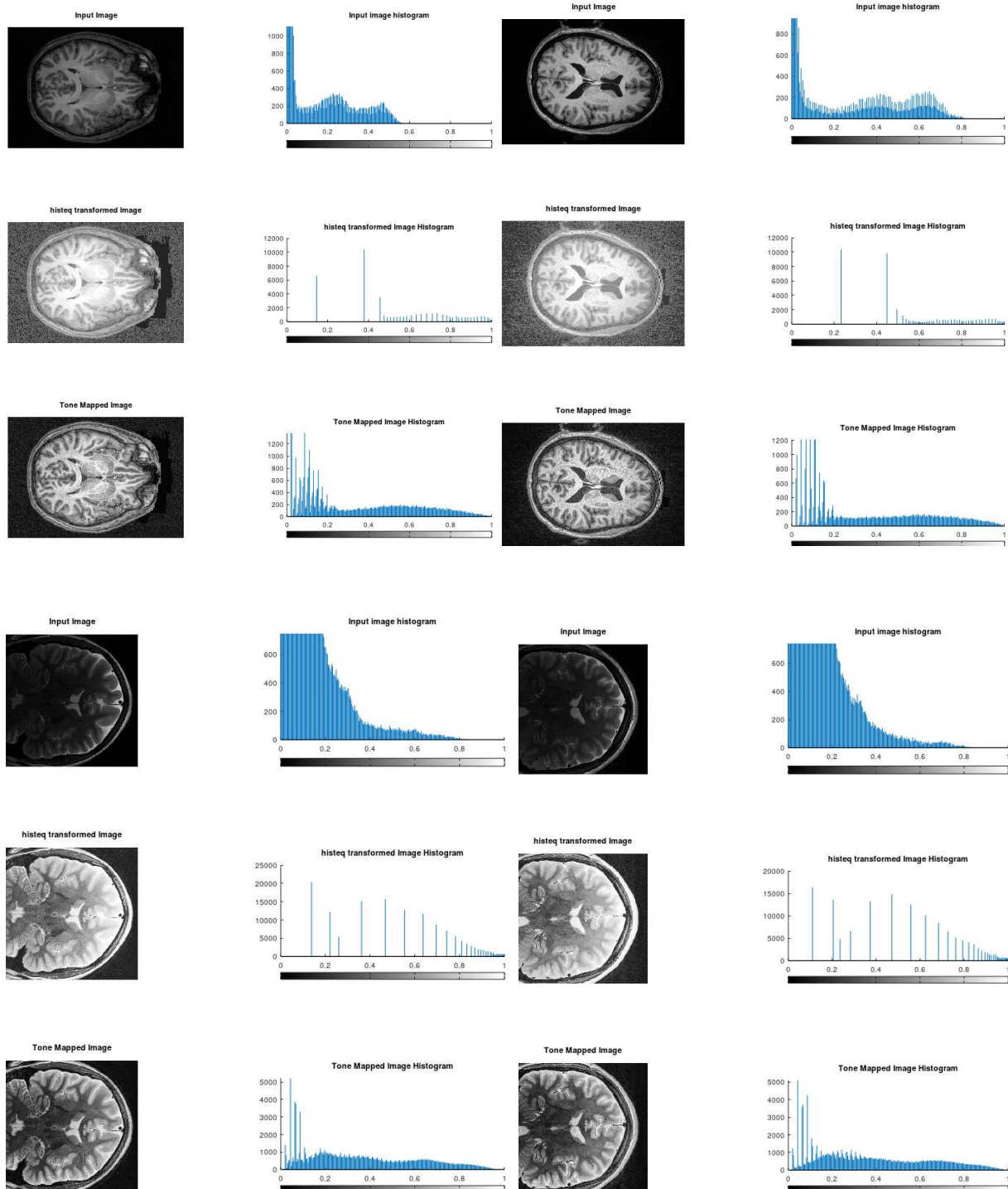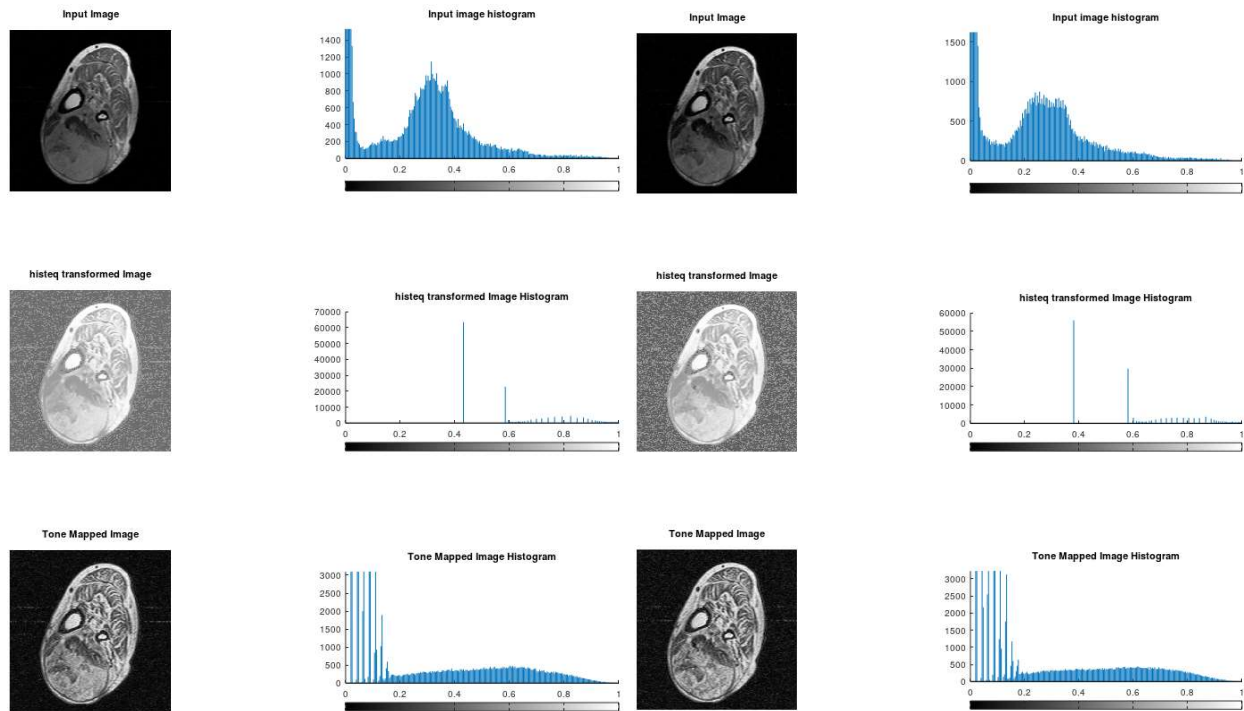
Tone Mapped

Original
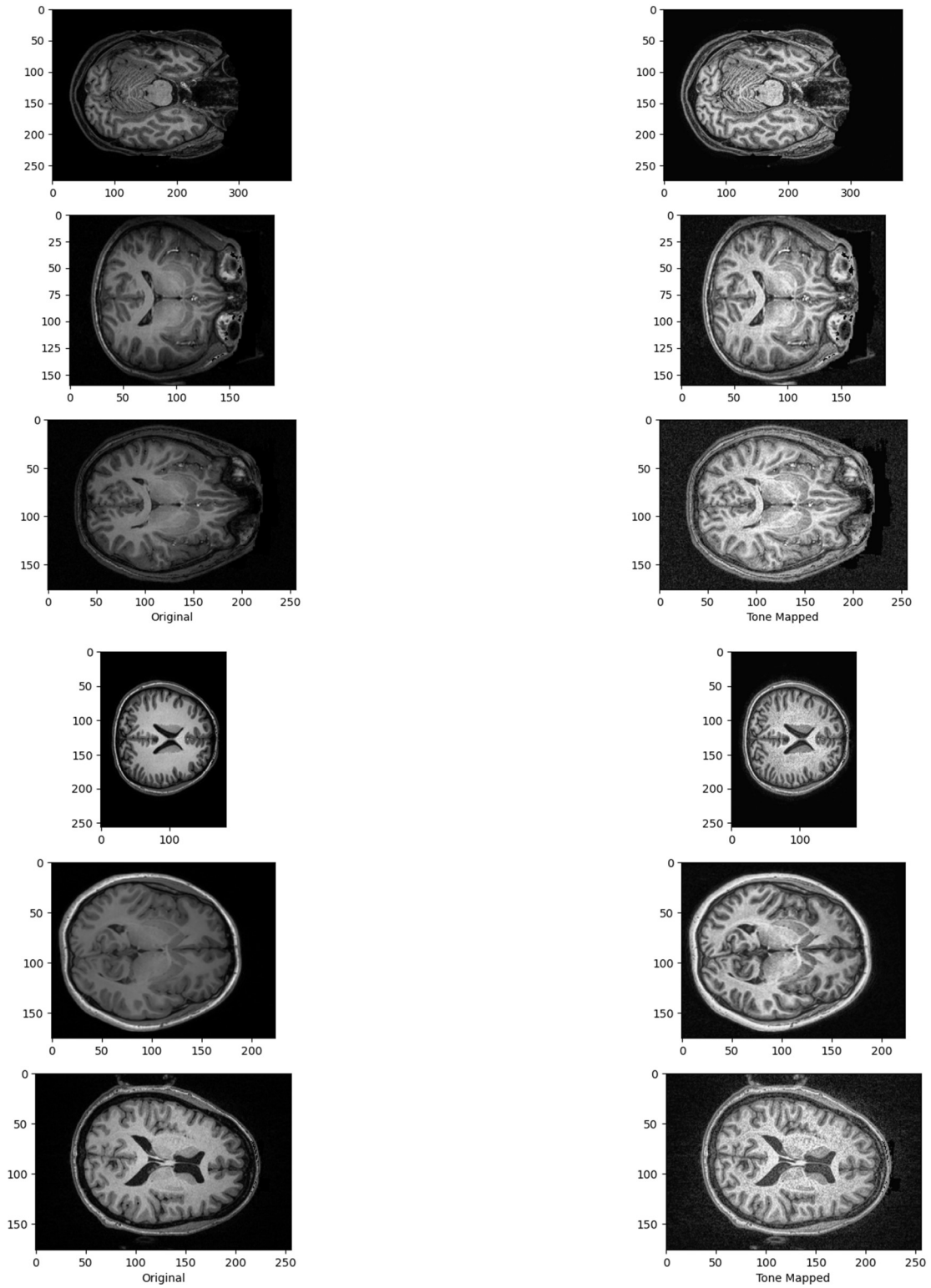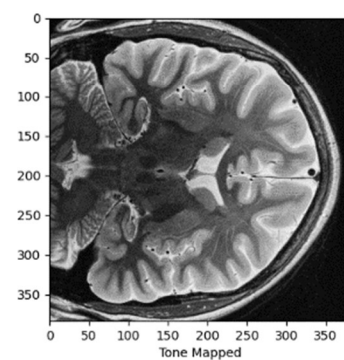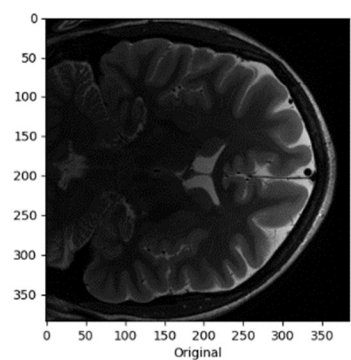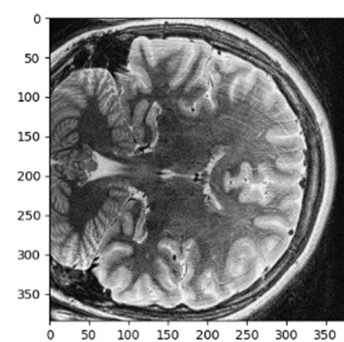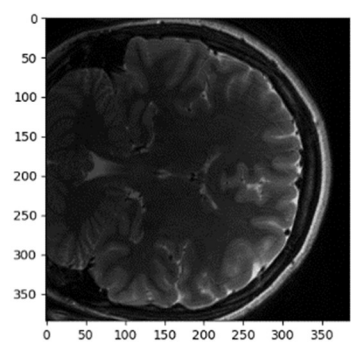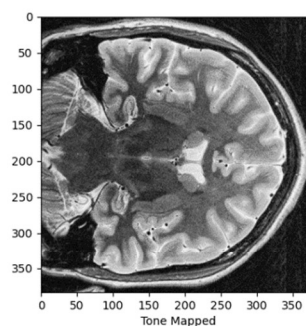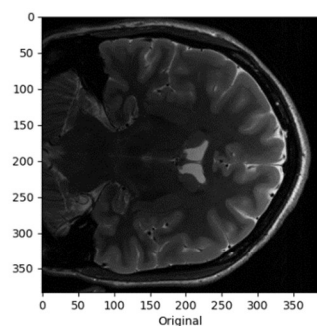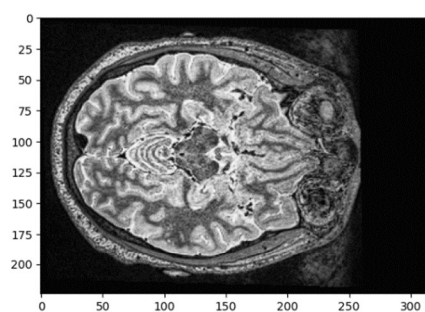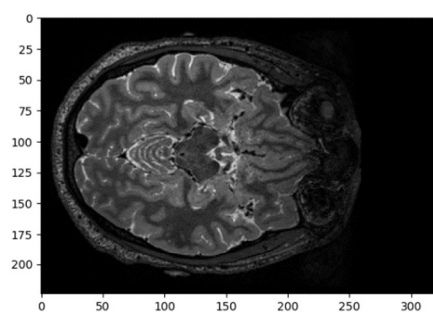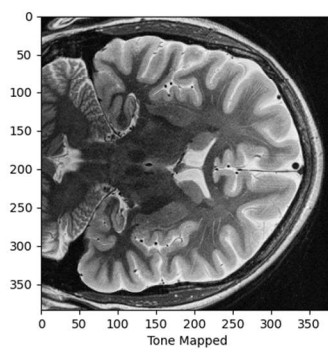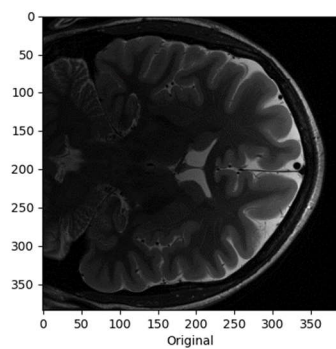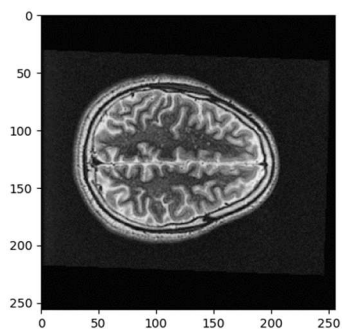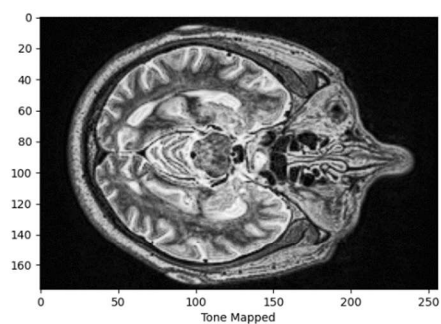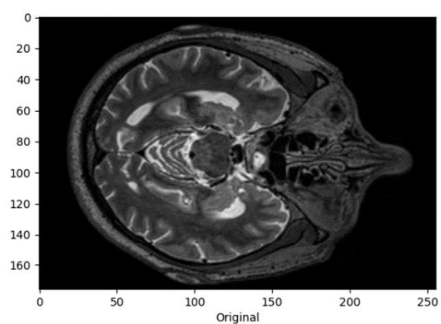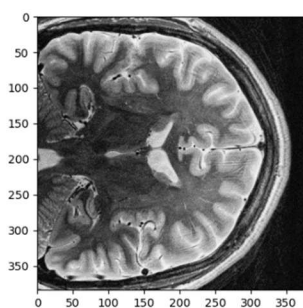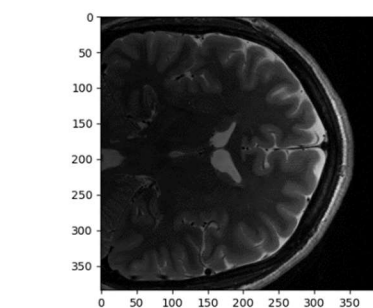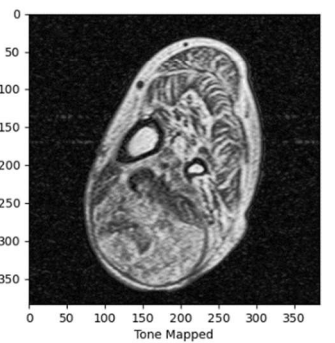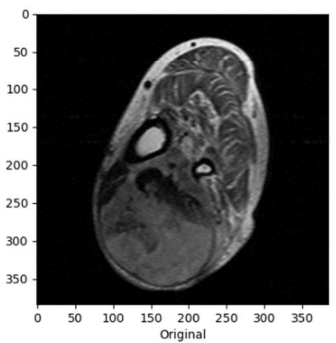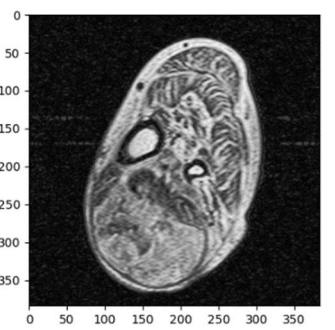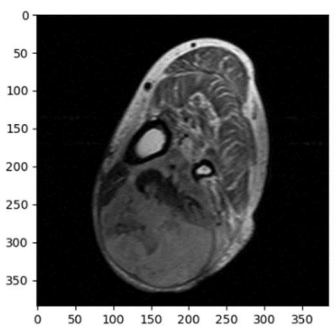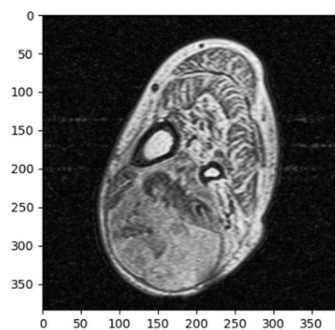
Tone Mapped

Histeq and Tone Mapping Algorithm Histograms

(Original-Histeq-Processed)

Analysis of the effectiveness and mean square error calculation

For the effectiveness analysis of the standard histeq and the proposed method, we selected randomly 12 different images from the set we were provided.

MSE	T1_1	T1_2	T1_3	T1_6	T2_3	T2_6	T2_4	T2_5	TE_	TE_	TE_	TE_9	Mean
Histeq	0.54	0.35	0.39	0.34	0.39	0.44	0.42	0.40	0.46	0.46	0.45	0.45	0.42
Proposed	**0.10**	0.15	0.19	0.12	0.26	**0.15**	0.25	0.25	**0.11**	0.12	0.15	0.15	**0.16**

The proposed method has 38% less mean square error and much better visual results.

Discussion

In this assignment we did a small state-of the-art review on using histogram matching/specification for MRI/CT normalization (population studies), and we also reproduced the results of one of those we reviewed. I believe the results we achieved are astonishing because by using some image processing techniques we could contribute to improving the visualization of pathologies. The output is much clearer than

input image and a lot of information that was barely visible even to experienced professionals, now is obvious.

We achieved that by combining several known image processing technics (simple like the ones we taught in class and some advanced) based on the histogram equalization and tone mapping in order to have better local contrast for CT images while maintaining a similar image structure to the reference CT image. The main difference between the proposed algorithm and traditional histogram equalization methods is the power-law-based distance-weighted contribution to the local histograms.

With this preprocessing method, by making the medical images (CT, MRI) clearer can be very useful for the medical experts and help with the visualization of pathologies and the diagnose. This method also enriches the quality of the medical image data that can be used for further research by scientists and achieve better results. The quality of these data is really important, especially when these data are used in unsupervised learning applications.

In contrast of everything we descripted and developed, in this paper [5], authors explain the failure of histogram matching methods and the propose a set of principles for image normalization and an explicit framework based on mixtures of distributions, where each fundamental distribution has a physical interpretation.

Concluding this assignment, I enriched my knowledge in histogram matching methods and at the same time we faced an important real word problem of medical imaging. I also tried to combine some advanced super-resolution algorithm (pretrained DNN's and GAN's models) in my implementation, to enhance the quality of the medical images and upscale them. Unfortunately, I wasn't able to implements these algorithms cause my machine doesn't support Nvidia's CUDA cores.

References

[1] L. G. Nyúl, J. K. Udupa and X. Zhang, "New Variants of a Method of MRI Scale Standardization," 2000. [Online]. Available: https://sci-hub.im/10.1109/42.836373.

[2] L. Shi and Y. Luo, et al. "Histogram-based normalization technique on human brain magnetic resonance images from different acquisitions," BioMedical Engineering OnLine, 2015. [Online]. Available: https://www.researchgate.net/publication/280536826_Histogram-based_normalization_technique_on_human_brain_magnetic_resonance_images_from_different_acquisitions.

[3] D. Völgyes, A. C. Trægde Martinsen, A. Stray-Pedersen, et al. "A Weighted Historgram-Based Tone Mapping Algorithm for CT Images," MDPI, 2018. [Online]. Available: https://www.mdpi.com/1999-4893/11/8/111/htm.

[4] X. Chen, W. Yaping, Z. Guohua, et al."Automatic Histogram Specification for Glioma Grading Using Multicenter Data," Journal of Healthcare Engineering, 2019. [Online]. Available: https://www.hindawi.com/journals/jhe/2019/9414937/.

[5] R. T. Shinoharaa, E. M. Sweeneyb, J. Goldsmith, et al."Statistical normalization techniques for magneticresonance imaging," ELSEVIER, 2014. [Online]. Available: https://www.researchgate.net/publication/267930907_Statistical_normalization_techniques_for_magnetic_resonance_imaging.