

Quick guide on radiology image pre-processing for deep learning applications in prostate cancer research

Samira Masoudi¹,^a Stephanie A. Harmon¹,^a Sherif Mehralivand¹,^a
Stephanie M. Walker¹,^a Harish Raviprakash¹,^b Ulas Bagci¹,^c
Peter L. Choyke,^a and Baris Turkbey^{a,*}

^aNational Cancer Institute, National Institutes of Health, Molecular Imaging Branch,
Bethesda, Maryland, United States

^bNational Institutes of Health, Department of Radiology and Imaging Sciences,
Bethesda, Maryland, United States

^cUniversity of Central Florida, Orlando, Florida, United States

Abstract

Purpose: Deep learning has achieved major breakthroughs during the past decade in almost every field. There are plenty of publicly available algorithms, each designed to address a different task of computer vision in general. However, most of these algorithms cannot be directly applied to images in the medical domain. Herein, we are focused on the required preprocessing steps that should be applied to medical images prior to deep neural networks.

Approach: To be able to employ the publicly available algorithms for clinical purposes, we must make a meaningful pixel/voxel representation from medical images which facilitates the learning process. Based on the ultimate goal expected from an algorithm (classification, detection, or segmentation), one may infer the required pre-processing steps that can ideally improve the performance of that algorithm. Required pre-processing steps for computed tomography (CT) and magnetic resonance (MR) images in their correct order are discussed in detail. We further supported our discussion by relevant experiments to investigate the efficiency of the listed preprocessing steps.

Results: Our experiments confirmed how using appropriate image pre-processing in the right order can improve the performance of deep neural networks in terms of better classification and segmentation.

Conclusions: This work investigates the appropriate pre-processing steps for CT and MR images of prostate cancer patients, supported by several experiments that can be useful for educating those new to the field (https://github.com/NIH-MIP/Radiology_Image_Preprocessing_for_Deep_Learning).

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JMI.8.1.010901](https://doi.org/10.1117/1.JMI.8.1.010901)]

Keywords: deep learning; medical images; image pre-processing; prostate cancer research.

Paper 20193VR received Jul. 23, 2020; accepted for publication Dec. 4, 2020; published online Jan. 6, 2021.

1 Introduction

In recent years, artificial intelligence (AI) has achieved substantial progress in medical imaging field where clinical decisions often rely on imaging data, e.g., radiology, pathology, dermatology, and ophthalmology.¹ Within radiology, AI has shown promising results in quantitative interpretation of certain radiological tasks such as classification (diagnosis), segmentation, and quantification (severity analysis). It is usually a time-intensive, error-prone, and non-reproducible procedure when a radiologist evaluates scans visually to report findings and make

*Address all correspondence to Baris Turkbey, Ismail.Turkbey@nih.gov

diagnostic decisions.² Alternatively, AI algorithms outperform the conventional qualitative approaches with faster pattern recognition, quantitative assessments, and improved reproducibility.¹⁻⁴ More specifically, the advent of deep neural networks along with the recent extensive computational capacity enabled AI to stand out by learning complex nonlinear relationships in complicated radiology problems. As a result, deep learning-based AI has met and even surpassed human-level performance in certain tasks. However, training these usually large models, requires massive amounts of data, which can be limited in medical imaging applications due to the concerns over data privacy as well as the paucity of annotation (labels) in supervised learning. With a continuing trend for developing universal data anonymization protocols in addition to open data sharing policies, larger clinical datasets have started to become available. Thus, training on massive dataset composed of different sources (institutions with different scanners, image qualities, and standards) is an ongoing potential pathway. Nonetheless, in cases with inevitable limited data, possible solutions are to use simpler designs or to employ transfer learning strategies based on giant datasets of clinical or natural images.⁵ Either scenario, training on a big dataset or adopting transfer learning, are only two other important reasons that resonate the demand for a series of preliminary pre-processing steps prior to training; however, the nature of radiology images intrinsically necessitates the preprocessing phase by itself.

Radiology images are acquired in a different way from natural pictures. Distinctive features of radiology images in each modality [i.e., computed tomography (CT) and magnetic resonance imaging (MRI)] are directly correlated with technical parameters used to generate these images. The detected (measured) signal within a scanner constitute the raw data that is reconstructed into an image in digital imaging and communications in medicine (DICOM) format as a standard in clinical medicine.⁶ DICOM files contain “metadata” in addition to the “pixel data,” which consists of image acquisition parameters, scanner information, and individual patient identification data. For clinical assessment, radiologists usually import this information as an image using the “picture archiving and communication system” (PACS).⁷ In some of their modality-specific and organ-specific workflows to fulfill a certain task, radiologists practice further adjustments to images often provided through a variety of third-party software programs embedded in PACS. These likely image modifications should be imitated as potential image pre-processing steps before training an AI system.

In the event that a massive dataset is built upon the data obtained from different patients, scanners, or multiple institutions, there are usually slight variations in the image quality, field of view, and resolution, which should be taken into consideration through a few pre-processing steps to create an integrated dataset.

In case of transfer learning, a model usually pre-trained on natural images is fine-tuned using radiology images. Natural images are obtained, stored, and loaded into memory with globally meaningful range of intensity values. However, based on the setting, clinical images are often acquired in multiple different ways with certain interpretation for their intensity values. Thus, to enable a robust transfer learning from the natural images into a dataset of radiology images, it is beneficial to apply several pre-processing steps.

Different pre-processing steps can also be viewed as a quality check for radiology images, which is beyond the embedded image quality filters at the scanner-level. To be more accurate, what is described in this paper is called post-processing according to medical physicists, and pre-processing according to image analysts. With that in mind, attaining a desired level of image quality in the training dataset can improve the succeeding quantification with deep learning.⁶

The nature of the pre-processing procedure strongly depends on the specific aim of the following processing algorithm and the image type. In practice, medical images can be provided in either formats of DICOM, analyze, Nifti, and Minc. Herein, our codes are based on the standard DICOM format;^{6,8} however, they can all be easily extended to any other format supported by the SimpleITK library.⁹ We also restrict our discussion to core pre-processing steps for anatomical MR and CT images with the aim of classification-, detection-, or segmentation-based applications using deep neural networks utilizing our institutional dataset of prostate cancer patients. Similar pre-processing methods can be extended to other imaging modalities or other medical research projects. Hence, we leave the choice of methods to users, with emphasis on the

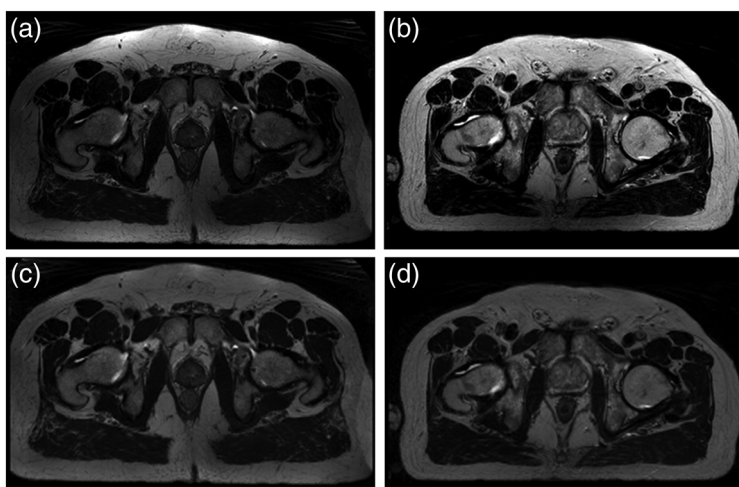


Fig. 1 (a) and (b) represent the MR images of the same anatomical slice in two different patients obtained within the same scanner, where inhomogeneity and intensity non-standardness are very well demonstrated. (c) and (d) show the same slices after bias field correction.

particular order that these methods should be applied to minimize the “uncertainties” coming from the data and enhance the performance of deep learning methods.¹⁰

1.1 What Do Medical Images Tell Us?

The clinical goal of medical imaging is to provide anatomical and functional information. The most common modalities include CT, x-ray imaging, MRI, ultrasound, and positron emission tomography (PET). Each modality undergoes certain acquisition conditions that are essential for quantitative assessment. CT scans are acquired by measuring x-ray attenuation through a rotating energy source that produces cross-sectional images. CT imaging provides anatomical information due to differences in tissue composition (i.e., density). The nature of CT acquisition results in a quantitative measurement of tissue density relative to water, known as Hounsfield unit (HU). Voxel HU values in CT images are largely considered reproducible with slight differences across different scanners and patients following the standard temperature and pressure specifications. While CT provides strong contrast of major anatomical landmarks, it is not the preferred modality in some clinical settings due to ionizing radiation and its limited soft-tissue contrast. Instead MR imaging provides excellent resolution and contrast of soft-tissue components. MR images represent the radio-frequency energy released after re-alignment of the protons in the presence of a strong magnetic field. Since MR acquisition results in voxel values obtained relative to each other, these images are subject to significant variations even when the same scanner is used to scan the same patient or the same organ (Fig. 1).

2 Methods

In this section, we go through the details of pre-processing steps for CT and MR images, respectively. Considering the deep networks’ task, the pre-processing may need to be applied at a certain level. For instance, data normalization may be performed using the image statistics at the slice-level (each single slice within an image), image-level, patient-level, scanner-level, institution-level, or an overall training data-level (Fig. 2). It is common to use the patient-level pre-processing in most studies, but there may be applications for which other methods are better suited.

2.1 Pre-Processing the CT Images

In case of CT images, potential sequence of pre-processing steps may be listed in order as denoising, interpolation, registration, organ windowing followed by normalization, and potentially zero-padding to improve the quality of training for deep learning algorithms.

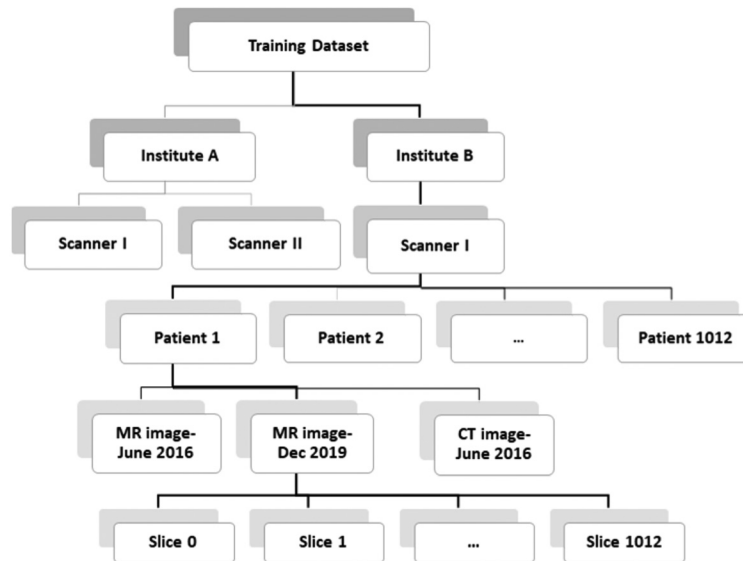


Fig. 2 Depending on the ultimate task of deep neural network, pre-processing is performed at a certain level of a dataset employing the respective statistics.

2.1.1 CT data denoising

There are several sources of disturbance in CT images, which mainly include beam hardening, patient movements, scanner malfunction, low resolution, intrinsic low dose radiation, and metal implants (Fig. 3). Each of these disturbances is addressed individually in the literature.^{11,12} Generally, images can be denoised within two domains: (i) spatial and (ii) frequency, which are comprehensively discussed in the literature.¹³

2.1.2 CT data interpolation

As it is usually the case for object detection and segmentation, there is a preference to have equal physical spacing for the input images. Maintaining the same resolution is desired to avoid center-specific or reconstruction-dependent findings. Images are usually interpolated in the xy plane and/or in z direction based on the desired physical spacing or ultimate number of the voxels. Function *ResampleImageFilter()* from the SimpleITK library performs this task with different interpolation methods. Cubic-spline and B-spline are generally better convolving functions to resample images as they perform close to an ideal low-pass filter.¹⁴ An inappropriate resampling step can negatively affect the subsequent registration and actual processing, such as degrading the resolution below the anticipated size of the detectable objects.¹⁵

2.1.3 Registration of CT Data

Image registration implies a spatial transformation to align an area of interest in a floating image with a reference image. Medical image registration in radiology is interpreted in two different ways: (i) slice-level registration in an image due to patient movements during the scan and (ii) image-level registration to have comparably aligned images in a training dataset. The first

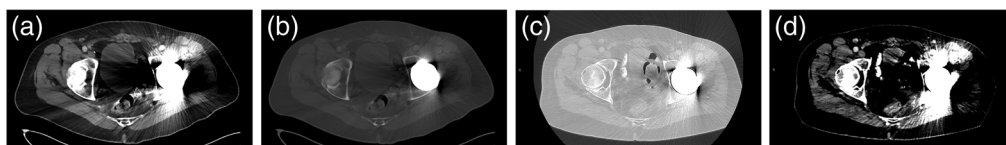


Fig. 3 Images show left hip prosthesis related diffuse streak artifacts obscuring the CT image quality after (a) soft tissue; (b) bone; (c) lung; and (d) liver windows.

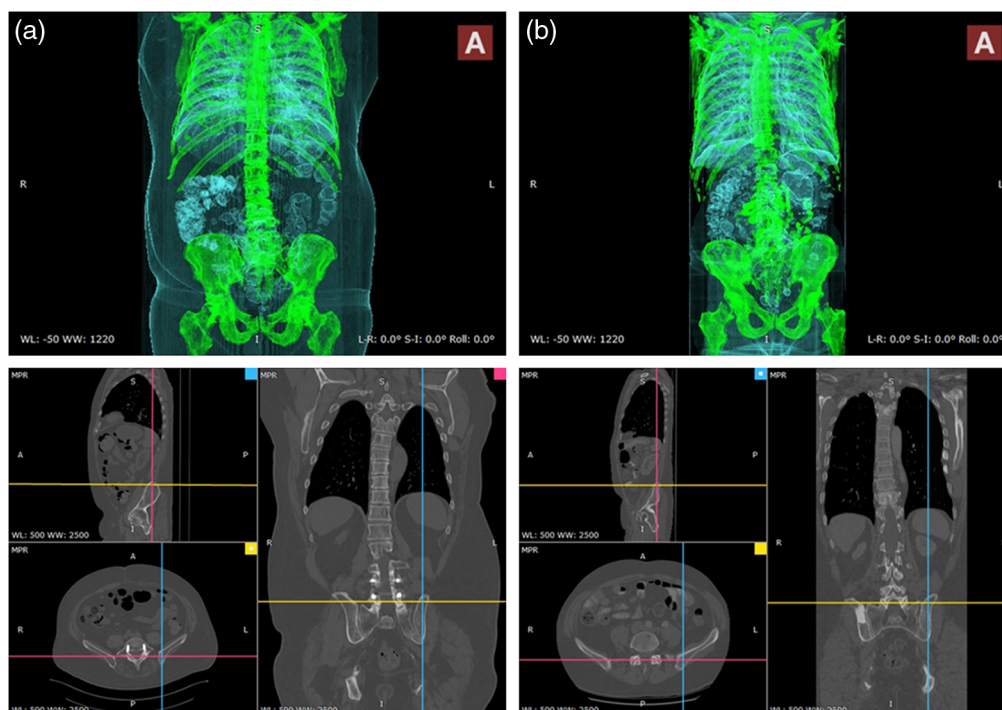


Fig. 4 Registered field of view in CT data of patients A and B (Appendix A1).

registration type may be required only in certain CT images that are the product of slow scanning, prior to denoising and interpolation. The latter type of image registration can ease the learning process of space-variant features for classification or segmentation using conventional machine learning tools. We emphasize that registration as pre-processing is not as critical for object recognition based on deep learning as it used to be for conventional methods. Deep learning algorithms can be trained to learn features invariant to affine transformation or even different viewpoints and illuminations.¹⁶ Yet, to facilitate training in complicated problems, registration may prove useful. In addition, there could be scenarios demanding registered images as the input for deep neural networks. For instance, the quality of synthetic CT generation is negatively affected by poorly registered pairs of MR and CT scans in training.¹⁷ Figure 4 shows 3D CT images of two patients captured within the same scanner where we care to have registered anatomical field of view as the input. Limiting the field of view for various purposes during the image acquisition, data curation, and through image pre-processing can improve the processing results in terms of fewer false positives in general. There exists a wide spectrum of methods from simple thresholding to segmentation for image registration in the literature which are beyond the scope of this work.^{18,19}

2.1.4 Windowing in CT data

As mentioned earlier, CT images are originally acquired in 12- or 16-bit DICOM format to present measures of tissue density in HU which are integer numbers within an interval of $(-1024 \text{ Hu}, 4000 \text{ Hu})$. These values may change to real numbers after potential denoising, interpolation, and registration. There is a limit to direct presentation of these values as an image to a radiologist. According to perception literature, the human eye can only distinguish 100 shades of gray simultaneously while it can discriminate several thousand colors.²⁰ Thus, for radiologists to be able to parse the complex images both faster and easier, one can incorporate color in clinical images or use image windowing to increase the contrast across a region of interest. In practice, the latter is commonly used where CT image values are clipped around a certain band of HU values for each organ, following unique standards from a texture-based dictionary described in Table 1. Obviously, computers are not limited by such a restriction, so windowing is not necessary if one is using the 16-bit images for training although they are not always supported by all

Table 1 Average Intensity Intervals in CT data that belong to different organs in four major areas of the body.²⁰

| Region | Organ | Intensity interval |
|---------------|----------------|--------------------|
| Head and neck | Brain | (40 ± 80/2) |
| | Subdural | (75 ± 430/2) |
| | Stroke | (36 ± 24/2) |
| | Temporal bones | (600 ± 280/2) |
| | Soft tissues | (375 ± 40/2) |
| Chest | Lungs | (-600 ± 1500/2) |
| | Mediastinum | (50 ± 350/2) |
| Abdomen | Soft tissue | (50 ± 400/2) |
| | Liver | (30 ± 150/2) |
| Spine | Soft tissue | (50 ± 250/2) |
| | Bone | (400 ± 1800/2) |

Table 2 Evaluating the effect of windowing and storage on binary bone lesion classification of CT images in terms of validation accuracy. Bolded values indicate the improved results.

| Windowing interval | Uint8 (%) | Uint16 (%) | Float32 (%) |
|--------------------|-----------|--------------|--------------|
| (400 ± 2000/2) | 76.50 | 77.34 | 78.17 |
| (0 ± 4000/2) | 75.63 | 77.38 | 76.90 |

deep learning libraries designed for 8-bit portable network graphics (PNG) or joint photographic experts group (JPEG) images.²¹ We believe storing images using 16-bit (either uint16 or float numbers) could improve the results. To demonstrate this principle, we designed an experiment to explore the effect of windowing and the format of image storage on the validation accuracy of a visual geometry group-16 (VGG-16) binary bone lesion classification (benign versus malignant). The pre-trained VGG-16 network on ImageNet was trained and evaluated using the lesion patches extracted from staging CT images of prostate cancer patients. Further details about this experiment are provided in Appendix B1. The obtained results are shown in Table 2 which vary over a limited interval of (76.99% ± 0.8%). One can observe that bone windowing certainly improved the accuracy while keeping data in float format led to the best comparative performance. Hence, if not limited by format, float and uint16 are generally preferred over the 8-bit images. In case of the 8-bit images, CT windowing is generally advised to avoid a non-reversible compression of visual details, where CT data are clipped to fit in an interval of $W_L \pm W_W/2$ (where W_L and W_W are the window level/center and window width, respectively, following Table 1). Otherwise, it is a good idea to perform CT windowing within a wide interval of (-1350 Hu, 2000 Hu) to remove the extreme values caused by metal artifacts as shown in Fig. 3(b).

2.1.5 Normalization of CT data

To stretch or squeeze the CT data intensity values so that they efficiently fit into the provided range of input images (8 or 16 bits) monotonically, we simply use a linear transformation,

enforcing two critical points (smallest and largest values) to be mapped to (0, 255) or (0, 65,535) respectively. Usually, we perform normalization at an institution-level or dataset level (utilizing the minimum and maximum pixel values among all the patients). A simple rule of thumb for implementing this is to use the windowing cut-off values of $W_L \pm W_W/2$ as the image extremes. Alternatively, one may normalize CT images at a patient-level, or even a slice-level for certain applications.

2.2 Pre-Processing the MR Images

While MR imaging is advantageous due to its superior soft-tissue contrast and non-ionizing radiation, these images are usually more challenging to be displayed or analyzed due to the lack of a standard image intensity scale. Additionally, MR imaging may be degraded by artifacts arising from different sources that should be considered prior to any processing. Inhomogeneity, motion and scanner-specific variations are among the major artifacts seen with the MRI. Figures 1(a) and 1(b) show how one similar anatomical slice in two different patients, obtained using the same scanner, have totally different intensity values. More importantly, Fig. 1 shows how the intensity values across the same organ within the same patient [for instance fat tissue in Fig 1(a)] vary in different locations. Potential steps to prepare MR images for further processing respectively include: denoising, bias field correction, registration, and standardization.²²

2.2.1 Denoising of MR data

Traditionally, denoising was an inevitable step to pre-process contaminated images. However, this phase is usually embedded within the current MR scanners, making acquired MR images to rarely suffer from direct noise distortion. In practice, additive Gaussian, Rician, and Speckle noise, respiratory and body movements and aliasing may still be major sources of contaminating noise in these images that should be addressed prior to any diagnosis. Motion can be avoided through several imaging protocols in this regard. There is also a long history of conventional methods proposed in the literature for motion artifact removal in MR images.^{23–35} Reliability of the derived diagnosis can be degraded by noisy image, which challenges both radiologists and automated computer-aided systems. To avoid false interpretations, it is critical to identify and exclude such poor images prior to any algorithmic analysis or inspection by the radiologists.

2.2.2 MR bias field correction

“Bias field” distortion is a low-frequency non-uniformity that is present in the MR data, causing the MR intensity values to vary across the images obtained from the same scanner, same patient, and even for the same tissue.³⁶ Figure 5(a) shows this effect in an MR image obtained from the pelvic area. It is critical to perform bias correction on MR images prior to any registration.^{36,37} The accuracy of image registration not only depends on spatial and geometric similarity but also on the similarity of the intensity values of the same tissue across different images. There are segmentation-based methods that estimate the bias field as an image parameter using expectation–maximization (EM) algorithm.^{38–40} Other approaches are involved with image features instead, where N-4 ITK is the most commonly used method among them.^{9,41}

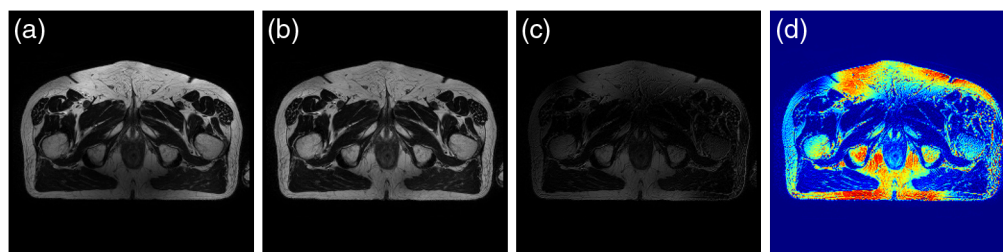


Fig. 5 (a) and (b) represent the same MR slice with Bias field and after bias removal. (c) and (d) are the absolute difference and heat map of the respective bias field.

This approach was first proposed by Tustison et al.,⁴¹ and is available in SimpleITK library. We have shown the results of bias field removal in Figs. 5(c), and 5(d).

2.2.3 Registration of MR data

Similar to CT, MR scanning may leads to mis-aligned images that require image registration. In fact, since MR scanning is much slower than CT, MR slices within the same image sequence are more likely to be offset from one another. While slice-level registration is required when processing 3D images within one channel, image-level registration becomes also essential while processing various modalities all together.^{42,43} A critical example could be applications where image fusion or domain adaptation is aimed between MR and CT images obtained using different fields of view during separate retrospective studies.¹⁶

2.2.4 MR data standardization

To have comparable intensity values in MR images, we must enforce image standardization so that for the same MR protocol within the same body region, a particular intensity value represents a certain tissue across all different slices and patients. Many studies in the literature use the standardization method proposed by Nyul and Udupa⁴⁴ to alleviate this problem. However, prior to Nyul pre-processing, we may: (i) shift the intensity values into a positive range of numbers, and (ii) use the probability distribution function (PDF) of the image intensity values to cut off the upper or lower narrow tails of the histogram. The latter can potentially help to remove very rare incidents of a certain intensity values caused by noise. Afterward, images can be standardized through a two-step Nyul pre-processing method which illustration is shown in Figs. 6 and 7.

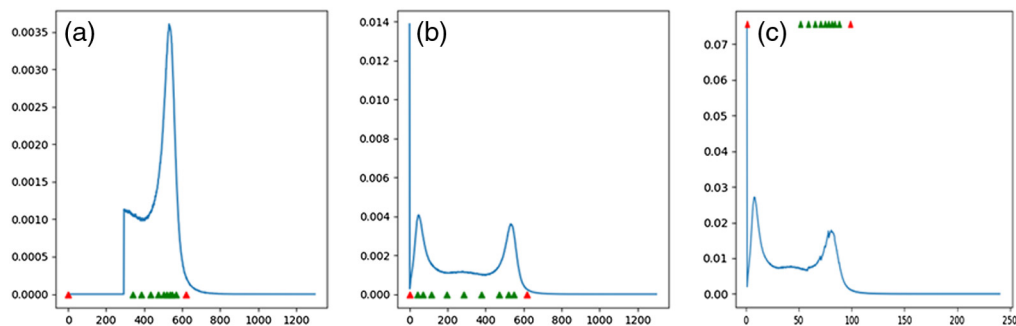


Fig. 6 (a), (b), and (c) represent the foreground PDF with standard landmarks, original PDF, and the resulted mapped PDF associated with Fig. 7.

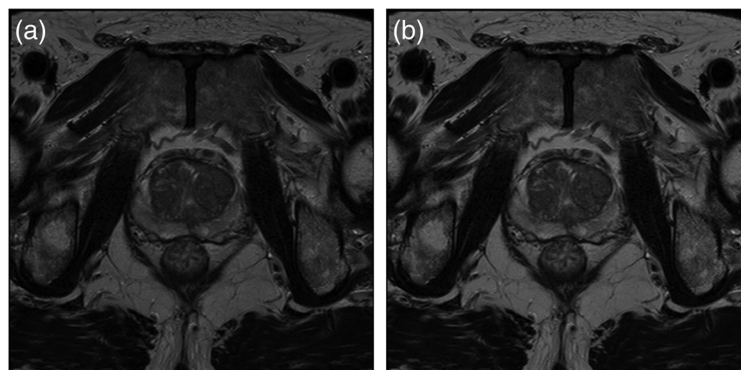


Fig. 7 (a) and (b) represent the same MR slice before and after Nyul preprocessing using 10 green landmarks in Fig. 6.

Table 3 Evaluating the effect of pre-processing on classification and Segmentation tasks in terms of validation accuracy. Bolded values indicate the improved results.

| Task | Image type | Criteria | With pre-processing (%) | No pre-processing (%) |
|----------------|------------|----------------|-------------------------|-----------------------|
| Classification | MR | Accuracy | 73.30 | 68.74 |
| | CT | Accuracy | 82.28 | 77.72 |
| Segmentation | MR | Mean Abs. Err. | 2.73 | 47.64 |
| | | Dice | 98.64 | 81.74 |
| | CT | Mean Abs. Err. | 3.68 | 19.99 |
| | | Dice | 98.25 | 95.25 |

The first step implies the learning phase where parameters of the standardizing transformation are learned from images in the training set. Next is the transforming phase where the intensity values of all images (training, validation, and validation) are mapped into new values using the standardizing transformation computed during the learning phase. Doing so, the PDF of each image would match a standard PDF. It is worthy to note that such standard PDF for MR data is approximated by a bi-modal distribution where the first bump usually represents the background. In practice, the Nyul method is focused on mapping the foreground or the second bump in histogram using a certain number of the landmarks [10 landmarks in Fig. 6(c)].

2.3 Demonstration

Finally, to demonstrate the efficacy of the aforementioned pre-processing methods in improving the performance of deep neural network algorithms, we performed four experiments with the aim of classification and segmentation in CT and MR images, i.e., two experiments for each image modality. These tasks can be described as (1) binary bone lesion classification (benign versus malignant) of 2-D lesion patches extracted from the CT data of proven prostate cancer patients; (2) binary 2-D slice-level classification (class 0 implies containing lesion of at most PIRADS 1,2 and class 1 describes slices with higher risk lesions PIRADS 3,4,5) of T2-weighted MR scans of prostate cancer patients from ProstateX dataset;⁴⁵ (3) visceral fat segmentation of 2-D slices from abdominal CT-scans; and (4) whole prostate segmentation in 2-D slices of T2-weighted MR images from ProstateX dataset.⁴⁵ For classification and segmentation purposes, we used two popular algorithms, VGG-16, and U-net, respectively. Our classifiers are trained taking advantage of transfer learning on the ImageNet, while U-net was trained from scratch in each case. More details are provided in Appendix B.

The results with and without the pre-processing are shown in Table 3 that confirm the necessity of appropriate pre-processing steps in the suggested order. While classification accuracy in both patch-level (CT) and slice-level (MRI) has increased significantly, the profound effect of pre-processing on segmentation performance is acutely presented both in terms of mean absolute error and dice score. The results in Table 3 also signify the demand for pre-processing in MRI images compared to CT images. Figure 8 shows an example of T2-MR slice where pre-processing could prevent under-segmentation of the whole prostate by U-net.

3 Discussion and Conclusions

Certainly, one major area in which AI is ascendant is the field of radiology. Conventionally, radiologists have exploited their scientific knowledge and visual ability to interpret medical images for diagnostic decisions. While such tasks are laborious and time-intensive, the results

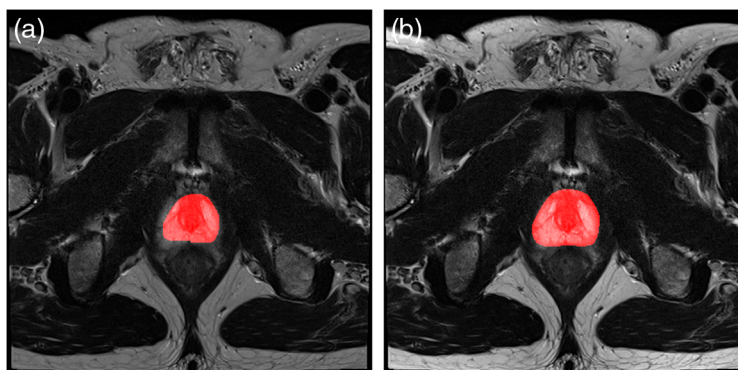


Fig. 8 An example of whole prostate Segmentation (a) without pre-processing (mean absolute error of 28.4%) and (b) with pre-processing (mean absolute error of 4.3%).

are excessively error-prone.² Current shortcomings necessitate an alternative solution. With the advent of deep neural networks, certain radiologist's tasks can be assigned to pre-trained machines instead. However, such a revolutionary substitution toward automation highly depends on how radiologists can understand the potentials of these algorithms. Currently, visual computation tasks in radiology addressed with deep neural networks mainly include but are not limited to detection, classification, and segmentation. In due time, it is critical for every radiologist to learn to: (1) define a problem with regards to the capabilities of deep learning algorithms; (2) collect and label the relevant data consistent with the algorithm to address the problem; (3) perform the necessary pre-processing steps to prepare the input data; and (4) choose an appropriate algorithm which will be trained on the provided data to solve the problem. There is a lack of thorough references available to educate radiologists on pre-processing as the first step. It is necessary for both radiologists and data scientists to understand these preprocessing steps so they can work effectively together to create ideal solutions. We hope this compilation of codes from the public domain will be useful.

4 Appendix A: Registration Details

4.1 Registration of CT data

As it can be seen in Fig. 3, both patients in Fig. 3 are adjusted to have the same axial view in 512×512 axial plane which results in different x and y spacing. On the other hand, with similar z -spacing (slice thickness) of 1 mm, images (a) and (b) are turn out to have a different number of slices, 663 and 778, respectively, while presenting a similar view in the z direction. With all these considerations, due to slight differences in measurement specification and organ-specific spatial organization of the patients, (x, y, z) physical spacing can not be used as a reference for registration.

5 Appendix B: Implementation Details

5.1 Bone lesion classification in CT scans

A dataset of 2879 annotated bone lesions from CT scans of 114 patients diagnosed with prostate cancer from the National Cancer Institute, National Institutes of Health was used for this experiment. We utilized 85 cases with 2224 lesions to train a pre-trained VGG-16 classifier and validated our results on 29 cases with 655 lesions. We used lesion annotations in terms of their labels of either being benign or malignant to define the classification task. Through 80 epochs of training, we used ADAM optimizer to minimize our binary cross-entropy loss function, initiated from the base learning rate of $3e - 3$ with 0.2 iterative drop, in learning rate in every 5 epochs.

5.2 *PI-RADS classification of T2-weighted MR images from ProstateX dataset*

We used whole prostate segmentation of 347 patients from publicly available Prostate X dataset including the T2-weighted MR scans along with internal annotation for lesions using 5-scale PI-RADS categorizations. We extracted 2D slices ($N = 2866$) from T2-weighted MR images where each 2D slice is assigned a class 0 ($N = 1394$) or 1 ($N = 1472$). In our dataset, class 0 implies that lesions of at most PI-RADS 2 are contained within the 2D slice while 2D slices with class 1 carry higher risk lesions with at least PI-RADS 3. The data were split into a training cohort of 300 patients and 47 patients for evaluation. We defined our loss function based on categorical cross-entropy, minimized it during 60 epochs of training, using ADAM optimizer with a base learning rate of $1e - 3$ along with 0.2 iterative drop in every 7 epochs.

5.3 *Visceral fat segmentation in CT-scans of abdominal area*

We used CT scans (from the abdomen region) of 131 patients obtained from multiple centers (Mayo clinic) with their respective visceral fat labels for segmentation. We randomly split these patients into two groups of train ($N = 115$) and validation ($N = 16$) to respectively train and validate our U Net-based segmentation network. To facilitate the training for the complicated task of visceral fat segmentation, we extended the limited number of the annotated CT slices, using augmentation methods such as image translation (width and height), reflection (horizontal flip), and zooming to increase (by 4 folds) our set of training images. Defining a pixel-wise binary cross-entropy as our loss function, we used SGD to increase the generalizability of our model with the base learning rate of $5e - 4$ and iterative decrease in learning rate. Through 100 epochs for training, we maintained 8 multiples of the learning rate in every 12 epochs.

5.4 *Whole prostate segmentation on T2-weighted MR images from ProstateX dataset*

We used 99 cases of publicly available T2-weighted MR scans from Radboud University Medical Center in Nijmegen, Netherlands, with internal annotation for whole prostate segmentation. We extracted 2D slices ($N = 2866$) from T2-weighted MR images. The data were split into a training cohort of 85 patients for training and 14 patients for evaluation. We forced training by 100 epochs for minimization of pixel-wise binary cross-entropy using SGD optimizer with base learning rate of $5e - 4$ and 0.2 iterative drop every 15 epochs.

Disclosure

No conflicts of interest, financial or otherwise, are declared by the authors.

Acknowledgments

This work supported in part by the Intramural Research Program of the National Cancer Institute, National Institutes of Health (NIH). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Code, Data, and Materials Availability

Supporting the research results reported in the manuscript, we provided the organized preprocessing codes for MR and CT images at https://github.com/NIH-MIP/Radiology_Image_Preprocessing_for_Deep_Learning.

References

1. A. Hosny et al., “Artificial intelligence in radiology,” *Nat. Rev. Cancer* **18**(8), 500–510 (2018).
2. R. Fitzgerald, “Error in radiology,” *Clin. Radiol.* **56**(12), 938–946 (2001).
3. P. Lakhani et al., “Hello world deep learning in medical imaging,” *J. Digital Imaging* **31**(3), 283–289 (2018).
4. U. C. Birbiri et al., “Investigating the performance of generative adversarial networks for prostate tissue detection and segmentation,” *J. Imaging* **6**(9), 83 (2020).
5. R. Maithra et al., “Transfusion: understanding transfer learning for medical imaging,” in *Adv. Neural Inf. Process. Syst.*, pp. 3342–3352 (2019).
6. M. Larobina and L. Murino, “Medical image file formats,” *J. Digital Imaging* **27**(2), 200–206 (2014).
7. M. F. McNitt-Gray, E. Pietka, and H. K. Huang, “Image preprocessing for a picture archiving and communication system,” *Investig. Radiol.* **27**(7), 529–535 (1992).
8. K. Grm et al., “Strengths and weaknesses of deep learning models for face recognition against image degradations,” *LET Biometrics* **7**(1), 81–89 (2017).
9. Z. Yaniv et al., “SimpleITK image-analysis notebooks: a collaborative environment for education and reproducible research,” *J. Digital Imaging* **31**(3), 290–303 (2018).
10. X. Chen et al., “Medical image segmentation by combining graph cuts and oriented active appearance models,” *IEEE Trans. Image Process.* **21**(4), 2035–2046 (2012).
11. Y. Zhang et al., “Reducing metal artifacts in cone-beam CT images by preprocessing projection data,” *Int. J. Radiat. Oncol. Biol. Phys.* **67**(3), 924–932 (2007).
12. S. Patil and V. R. Udupi, “Preprocessing to be considered for MR and CT images containing tumors,” *IOSR J. Electr. Electron. Eng.* **1**(4), 54–57 (2012).
13. M. Diwakar and M. Kumar, “A review on CT image noise and its denoising,” *Biomed. Signal Process. Control* **42**, 73–88 (2018).
14. A. Parker, R. V. Kenyon, and D. E. Troxel, “Comparison of interpolating methods for image resampling,” *IEEE Trans. Med. Imaging* **2**(1), 31–39 (1983).
15. J. X. Jim, H. Pan, and Z. P. Liang, “Further analysis of interpolation effects in mutual information-based image registration,” *IEEE Trans. Med. Imaging* **22**(9), 1131–1140 (2003).
16. T. Poggio, T. A. Poggio, and F. Anselmi, *Visual Cortex and Deep Networks. Learning Invariant Representations*, MIT Press (2016).
17. M. C. Florkow et al., “The impact of MRI-CT registration errors on deep learning-based synthetic CT generation,” *Proc. SPIE* **10949**, 1094938 (2019).
18. J. A. Maintz and M. A. Viergever, “A survey of medical image registration,” *Med. Image Anal.* **2**(1), 1–36 (1998).
19. X. Cao et al., “Image registration using machine and deep learning,” in *Handbook of Medical Image Computing and Computer Assisted Intervention*, pp. 319–342, Academic Press, Cambridge, Massachusetts (2020).
20. J. Kube et al., “Windowing CT,” 2019, <https://radiopaedia.org/articles/windowing-ct>.
21. E. Provenzi, Ed., “Color image processing,” in *MDPI* (2018).
22. J. Howard, “Don’t see like a radiologist,” 2019, <https://www.kaggle.com/jhoward/don-t-see-like-a-radiologist-fastai>.
23. J. V. Manjón, “MRI preprocessing,” in *Imaging Biomarkers*, L. Martí-Bonmatí and A. Alberich-Bayarri, Eds., pp. 53–63, Springer, Cham (2017).
24. M. D. Pattany et al., “Motion artifact suppression technique (MAST) for MR imaging,” *J. Comput. Assist. Tomogr.* **11**(3), 369–377 (1987).
25. D. Tamada et al., “Motion artifact reduction using a convolutional neural network for dynamic contrast enhanced MR imaging of the liver,” *Magn. Reson. Med. Sci.* **19**(1), 64 (2020).
26. G. Gerig et al., “Nonlinear anisotropic filtering of MRI data,” *IEEE Trans Med Imaging* **11**, 221–232 (1992).
27. R. D. Nowak, “Wavelet-based Rician noise removal for magnetic resonance imaging,” *IEEE Trans. Image Process.* **8**(10), 1408–1419 (1999).
28. A. Pizurica et al., “A versatile wavelet domain noise filtration technique for medical imaging,” *IEEE Trans. Med. Imaging* **22**(3), 323–331 (2003).

29. L. Bao et al., "Sparse representation-based MRI denoising with total variation," in *9th Int. Conf. Signal Process.*, IEEE, pp. 2154–2157 (2008).
30. L. Bao et al., "Structure-adaptive sparse denoising for diffusion-tensor MRI," *Med. Image Anal.* **17**(4), 442–457 (2013).
31. P. Coupé et al., "An optimized block-wise nonlocal mean denoising filter for 3-D magnetic resonance images," *IEEE Trans. Med. Imaging* **27**(4), 425–441 (2008).
32. J. V. Manjón et al., "Robust MRI brain tissue parameter estimation by multistage outlier rejection," *J. Magn. Reson. Med.* **59**(4), 866–873 (2008).
33. J. V. Manjón et al., "Multicomponent MR image denoising," *Int. J. Biomed. Imaging* **2009**, 756897 (2009).
34. A. Tristán-Vega and S. Aja-Fernández, "DWI filtering using joint information for DTI and HARDI," *Med. Image Anal.* **14**(2), 205–218 (2010).
35. J. Rajan et al., "Nonlocal maximum likelihood estimation method for denoising multiple-coil magnetic resonance images," *Magn. Reson. Imaging* **30**(10), 1512–1518 (2012).
36. S. Pereira et al., "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Trans. Med. Imaging* **35**(5), 1240–1251 (2016).
37. U. Bagci, "Lecture 4: Pre-processing Medical Images (ii)," Lecture notes, Medical Image Computing, CAP 5937, Center for Research in Computer Vision, University of Central Florida, Orlando, Florida (2017).
38. K. Van Leemput et al., "Automated model-based bias field correction of MR images of the brain," *IEEE Trans. Med. Imaging* **18**(10), 885–896 (1999).
39. W. M. Wells et al., "Adaptive segmentation of MRI data," *IEEE Trans. Med. Imaging* **15**(4), 429–442 (1996).
40. S. Guillemaud and M. Brady, "Estimating the bias field of MR images," *IEEE Trans. Med. Imaging* **16**(3), 238–251 (1997).
41. N. J. Tustison et al., "N4ITK: improved N3 bias correction," *IEEE Trans. Med. Imaging* **29**(6), 1310–1320 (2010).
42. J. Chappelow et al., "Multimodal image registration of ex vivo 4 Tesla MRI with whole mount histology for prostate cancer detection," *Proc. SPIE* **6512**, 65121S (2007).
43. F. Cornud et al., "TRUS–MRI image registration: a paradigm shift in the diagnosis of significant prostate cancer," *Abdom. Imaging* **38**(6), 1447–1463 (2013).
44. L. G. Nyul and J. K. Udupa, "Standardizing the MR image intensity scales: making MR intensities have tissue-specific meaning," *Proc. SPIE* **3976**, 496–504 (2000).
45. G. Litjens et al., "ProstateX challenge data," *Cancer Imaging Arch.* (2017).

Samira Masoudi is a post-doctoral fellow at the National Institutes of Health. She received her BS degree in electrical engineering from Ferdowsi University, MS degree in biomedical engineering from Sharif University of Technology, and PhD in electrical engineering from the University of Wyoming, in collaboration with University of Central Florida, in 2019. She has several publications in the field of medical signal and image processing. Her current research interest is focused on the application of deep learning for prostate cancer research.

Stephanie A. Harmon received her BS degree in physics from Illinois Institute of Technology and PhD in medical physics from the University of Wisconsin in 2016. She joined Leidos Biomedical Research as a postdoctoral scientist within the molecular imaging branch (MIB) at the National Cancer Institute. Currently, she is a staff scientist at the NIH Artificial Intelligence Resource (AIR). Her research interests are in computational characterization of cancer and cancer-related outcomes through combination of multi-scale data incorporating information from medical and pathologic imaging.

Sherif Mehralivand is a urologist specialized in urologic oncology at the National Cancer Institute, Bethesda, MD. His main research interests are prostate cancer imaging, multiparametric MRI, PET-CT imaging, and the focal therapy of prostate cancer. His current research is focused on the development and deployment of artificial intelligence-based diagnosis systems in medicine. He is also exploring opportunities for the deployment of augmented reality systems for the improvement of targeted biopsy and focal therapy procedures.

Stephanie M. Walker is currently a fourth-year medical student at the University of Kansas School of Medicine and recent graduate of the NIH Medical Research Scholars Program, a year-long residential research training program for medical students where she worked in the Molecular Imaging Program within the National Cancer Institute studying prostate cancer imaging techniques.

Harish RaviPrakash is a research fellow at the Department of Radiology and Imaging Sciences at the National Institutes of Health Clinical Center. He received his PhD in computer science in 2020 from the University of Central Florida, advised by Dr. Ulas Bagci. His research interests include medical image analysis, statistical machine learning, and deep learning for translational AI.

Ulas Bagci is a faculty member at University of Central Florida (UCF). His research interests are artificial intelligence, machine learning and their applications in biomedical and clinical imaging. He has more than 200 peer-reviewed articles in related topics, serves as an external committee member of the AIR at the NIH, an area chair for MICCAI, an associate editor of top-tier journals in his fields, and an editorial board member of *Medical Image Analysis*.

Peter L. Choyke is a senior investigator and chief of the molecular imaging branch of the NCI in Bethesda, MD. He received his MD degree from Jefferson Medical College and postgraduate training at Yale and the University of Pennsylvania. His particular focus is prostate cancer imaging using both MRI and PET. He is interested in the applications of AI to medical imaging and digital pathology to improve the reliability of diagnosis.

Baris Turkbey directs the AIR at the National Cancer Institute, National Institutes of Health. He received his MD degree and completed his residency in diagnostic and interventional radiology at Hacettepe University in Ankara, Turkey. He joined the molecular imaging branch of the National Cancer Institute in 2007. His main research areas are imaging, biopsy techniques, and focal therapy for prostate cancer, along with image processing using artificial intelligence. He is a steering committee member of PIRADS and serves in the genitourinary section of ACR Appropriateness Criteria Panel.