Pseudo Computed Tomography Image Generation from Brain Magnetic Resonance Image for Radiation Therapy Treatment Planning Using DCNN-UNET

S. Sreeja*
Research Scholar, Department of Computer Science, University of Kerala, Karyavattom Campus, Trivandrum, Kerala, India. E-mail: sreebhav@gmail.com

D. Muhammad Noorul Mubarak
Head of Department, Department of Computer Science, University of Kerala, Karyavattom Campus, Trivandrum, Kerala, India.
Email: noorul.muby@gmail.com

Received May 18, 2021; Accepted August 20, 2021
ISSN: 1735-188X
DOI: 10.14704/WEB/V18SI05/WEB18256

Abstract

Relative to Computed Tomography (CT), the increased soft tissue contrasts of magnetic resonance imaging (MRI) makes it a suitable imaging method to decide radiation therapy (RT). When MRI scans are used for therapy planning, a CT scan is still required for dosage calculation and x-ray-based patient placement. This raises workload, leads to incertitude owing to the requisite of image registration inter-modality and requires needless irradiation. Even though it would be advantageous to only use MR images, a way of estimating a pseudo-CT (pCT) must be used to generate electron density mapping and patient reference imagery. So, this paper brings an effective deep learning model to generate synthesized CT from MRI images using the following steps; a) data acquisition where CT and MRI scan images are collected, b) preprocessing of images to avoid the anomalies and noises using techniques like outlier elimination, data smoothening and data normalizing, c) feature extraction and selection using Principle Component Analysis (PCA) & regression method, d) generating pCT from MRI using Deep Convolutional Neural Network and UNET (DCNN-UNET). Further, we assessed metrics such as DC, SSIM, MAE and MSE for this model. However, our suggested model outperforms with an accuracy of 95%.

Keywords

Computed Tomography, Convolutional Neural Network, Deep Convolutional Neural Network, Deep Learning, Hounsfield Unit, Magnetic Resonance Imaging, Principle Component Analysis, Pseudo Computed Tomography.

Introduction

In the radiation treatment (RT) chain, computed tomography (CT) with three-dimensional attenuation coefficient maps is critical for treatment plan making and dosage computation. These are used to calculate dosages based on organ and tissue characteristics (1). Modern methods, such as intensity modulated radiation treatment (IMRT) and volumetric modulated radiation therapy (VMAT), use anatomical images to correctly define the goals and the organ at risk (OAR) for the administration of doses (Chandarana et al. (2018); Johnstone et al. (2018)). The usage of magnet resonance imaging (MRIs) in clinical practices has been increased as a result of the strong contrast between the soft tissue and clearer organ boundaries as compared to CT imaging. Furthermore, numerous studies have shown that operational MRI details, such as diffusion-weighted imaging (DWI) and dynamic contrast enhanced imaging, might be useful in identifying active tumor sub volumes in cancers of both the head and neck (Adjeiwaah et al. (2019)). For the proper determination of the target volume and OAR to be integrated into the CT reference image, magnet resonance (MR) images now have a rigid or flexible registration in the RT chain. Dose estimations are based on electron density measurements from CT images (Adjeiwaah et al. (2019); Dean et al. (2012)). MR-related errors with CT image recordings nonetheless contribute to systematic uncertainty which leads to a substantial influence on dosimetry, particularly for tiny tumors close to OARs (Edmund et al. (2017); Arabi et al. (2016)). In the radiation therapy (RT) process, MRI-only RT planning is used to avoid these mistakes in RT planning as well as lower therapy costs. MRI-RT alone eliminates the need for MR and CT image registration, reducing the number of imaging sessions and their associated costs in the process (CT imaging). This reduces the dosage received, especially for the patients who require several scans during the therapy (Arabi et al. (2017)).

The major problems of replacing MRI with CT are that the values of MRI intensity do not equate to electron densities and typical MRI sequences do not collect bone signal. Consequently, it is important to have a technology that can generate CT comparable details from MR images for the estimation of the dose throughout treatment planning and digitally reconstructed radiograph (DRR) patient setup. Pseudo-CTs or Synthetic CTs (pCT/sCT) are terms used to describe magnetic resonance (MR)-based findings. In hybrid Positron Emission Tomography / Magnetic Resonance Imaging (PET/MRI) systems, accurate pCTs from MR images are required in addition to MR-only radiation to correct PET attenuation. In addition, the reduced geometric validity of MRI may still be a concern when using pCT for these applications, since the pCT produced may inherit any geometric distortion present in the original MR image. The typical methodology for MRI pCT generation without DL method is shown in figure 1.

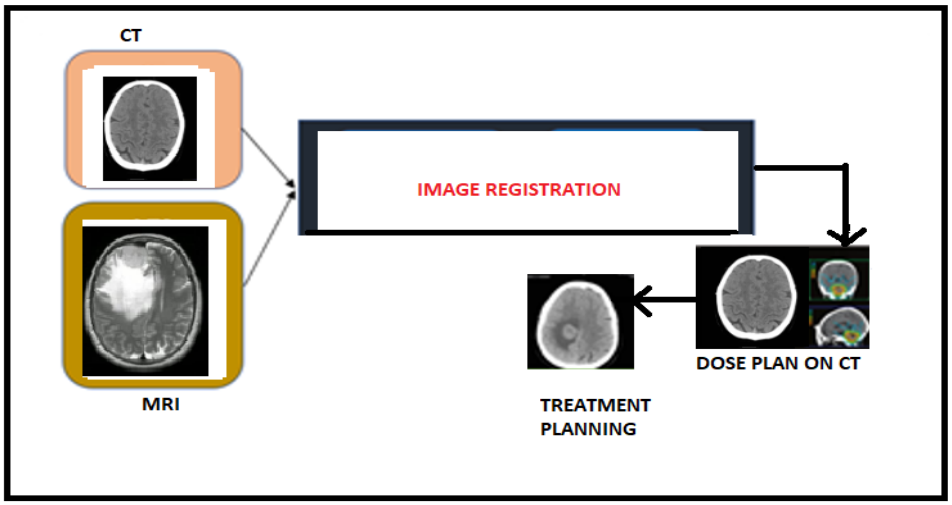

Figure 1 Traditional pCT from MRI process

Many alternative approaches for the automated creation of pseudo-CTs from MR images were suggested in the research. Tissue segmentation strategies, learning techniques, and atlas methods are the three main categories. The first segment of images, or the categorization of the MR image voxels into a distinct collection of tissue types, can be addressed by Tissue Segmentation (e.g. air, bones, fat, and soft tissues). Tissue segmentation is not a straightforward matter, as Hsu et al. (Arabi et al. (2016)) have pointed out, and an amount of MR images is usually insufficient to divide all major tissue types. Additionally, standard MR sequences cannot distinguish bone from air reliably. Therefore, several MR sequences, notably the specialist UTE sequences, are required to be used for most tissue segmenting techniques. This might lead nevertheless to additional time for acquiring images and more complex workflows. However, available learning methods use statistic learning or model fitting approaches to create a map function that integrates MR voxel intensities (or density patterns) with the related CT numbers (Edmund et al. (2014); Korhonen et al. (2014)). As with tissue segmentation, it is inadequate to generate valid predictions using standard MR images alone. Some techniques need manual segmentation of the bone volume is required to create distinct models in various areas (Kapanen et al. (2013)). Further, multiple sequences are utilized to differentiate bone voxels from air, in particular sequences from UTE (Johansson et al. (2011); Navalpakkam et al. (2013)) as an alternative. In addition, other coordinate-related features and patterns of local intensity to MR-intensity values have been created and used as extra inputs to increase the productivity of the pCT model (Rank et al. (2013)).

We offer a new pseudo-CT generation learning approach in this paper, which uses recent advances in deep learning (DL) as well as convolutional neural networks (CNNs). CNN (Ronneberger et al. (2015); Chen et al. (2014)) are multi-layered models that are completely

trained and incorporate complicated, high-dimensional interactions between input and output. They have been utilized for computer-assisted identification and for the segmentation of biomedical images to resolve issues in many domains of computer vision (Zeiler et al. (2014); Simonyan et al. (2014); Badrinarayanan et al. (2015)). The CNN's capacity to learn complicated end-to-end models rather than handmade components is appealing. As a result, we utilize CNN to develop a direct mapping of images between MR images to their CTs respectively. This differs from all previous pCT learning methods based on hand-design and voxel-by-voxel CT values.

Key Highlights

This paper focus on generating a pseudo-CT from MRI using an effective Deep Learning model in which following are the highlights;

a. This study is intended to enhance the positioning of patients and the dosage in brain radiation alone MRI, utilizing acquired MRIs.
b. Exploration of the ability to remove and replace the CT simulator with CT images from the brain of the MRI and calculate the density of the electrons for the planning of the patient therapy.
c. A new Deep Convolutional Neural Network- U-Net (DCNN-U-Net) approach is presented for pCT creation and its efficiency assessed in a series of images of patients with brain tumors.
d. Used to investigate the link exists between intensity of MRI and CT- Hounsfield unit (HU) in various areas.
e. Find a compact and effective model, which can offer details on the electron density of MRI images to produce synthetic CT images at a reasonable error rate that might be utilized during clinical RT therapy.

Organization of the paper: As we already come across the introductory part in Section 1, rest of this paper is as follows; Section 2 depict the related works, followed by Section 3 with methodology, section 4 depicts the implementation and results , Section 5 depicts the discussion and finally Section 6 depicts conclusion.

Related Works

In the calculation of MRI, T. Korsholm et al. (2014) claimed that a dosage error of 2% is medically reasonable (considering a 1 % dose computation error when using CT). Furthermore, patient location validation using this technology to create reference images is difficult due to a lack of bone segmentation. The electron density or Hounsfield Unit (HU)

value of each class may also be used to separate the tissues in the MR image into discrete classes. In most cases, there are two or three classes: bone and soft tissue, and in certain circumstances, air. These approaches have been described with enhanced dosimetry in comparison with the use of an override of homogenous density, especially brain, prostate, head and neck.

The semiautomatic approach of bone segmentation in the skull was tested by Stanescu et al. (2008). A point here was put near the structure, which needed to be segmented. The structure was then divided into thresholds. The authors highlighted the need of manual modifications in certain cases, especially to the bottom region of the skull.

Before overriding bulk density, Paula et al. (2019) utilized an atlas-based segmenting approach to divide the bone. If required, manual change has again been performed. However they concluded that manual contour adaptation is not ideal and these may lead to a greater usage of bulk density approaches in a future clinical workflow.

Atlas-based methods usually employ a conventional single MRI sequence for the production of a pCT. This provides a minimal scan duration, decreasing the risks of moving patients. This also implies that it is easy to use the scan technique in a clinical setting. The pCT manufacturing process may be completely automated and reference images to verify the placement and the automatic contouring of OAR can be created. Sjolund et al. (2015) pointed out that atlas-based technologies for imagery artifacts are reasonably resilient because they rely on past knowledge about training.

Further, prostate planning, for instance, the simplest atlas approaches employ one or the medium atlas, such as established by Dowling et al. (2019). The average atlas approach is used to co-register paired MRI and CT images from the patients' database. The average MRI atlas is then created, along with a corresponding organ shape. In order to calculate the distortion that must be forced on each MRI scan in the database to obtain the average CT atlas, identical deformations for the respective registered CTs may be created, and the average of them can be computed. Moreover, to create a pCT for an incoming MR image, the average MRI atlas is captured in the MRI scan and the average CT atlas then used to apply these deformations, yielding the pCT. It is also possible to communicate the organ's shape. The most recent application for CT synthesis is the Cone Beam Computed Tomography (CBCT)-to-CT conversion through DL. Liu and al. (2019) integrated CBCT-CT pair organ change into a cycle-GAN generator, revealing that their contribution substantially improved the predictions of the same network without the AG. CBCT-CT pairs in the context of pancreatic adaptive RT. Okty et al. (2020) presented a novel

mechanism termed "Attention Gate" (AG) to focus on target structures that may differ in shape and size in relation to recent developments in the field of deep education.

It would be rather fascinating if a CNN were to automatically create a pCTs measurement. Matt et al. (2020) stated that a multi-task network and a Bayesian based probabilistic model would be employed with uncertainty for automatic creation of a pCTs measurement. In recent years, two more projects have advocated that uncertainty should be used either through the integration of self-trained networks or by drop-off variation. To date the subject of deep learning insecurity, estimate for pCT generation has only been lightly addressed. Future study aimed at creating criteria in which failure instances may be automated with the use of uncertainty provision would be intriguing. The CT-rescan or pCT-adjustment is considered viable in patients with incorrect pseudo-CTs.

Methodology

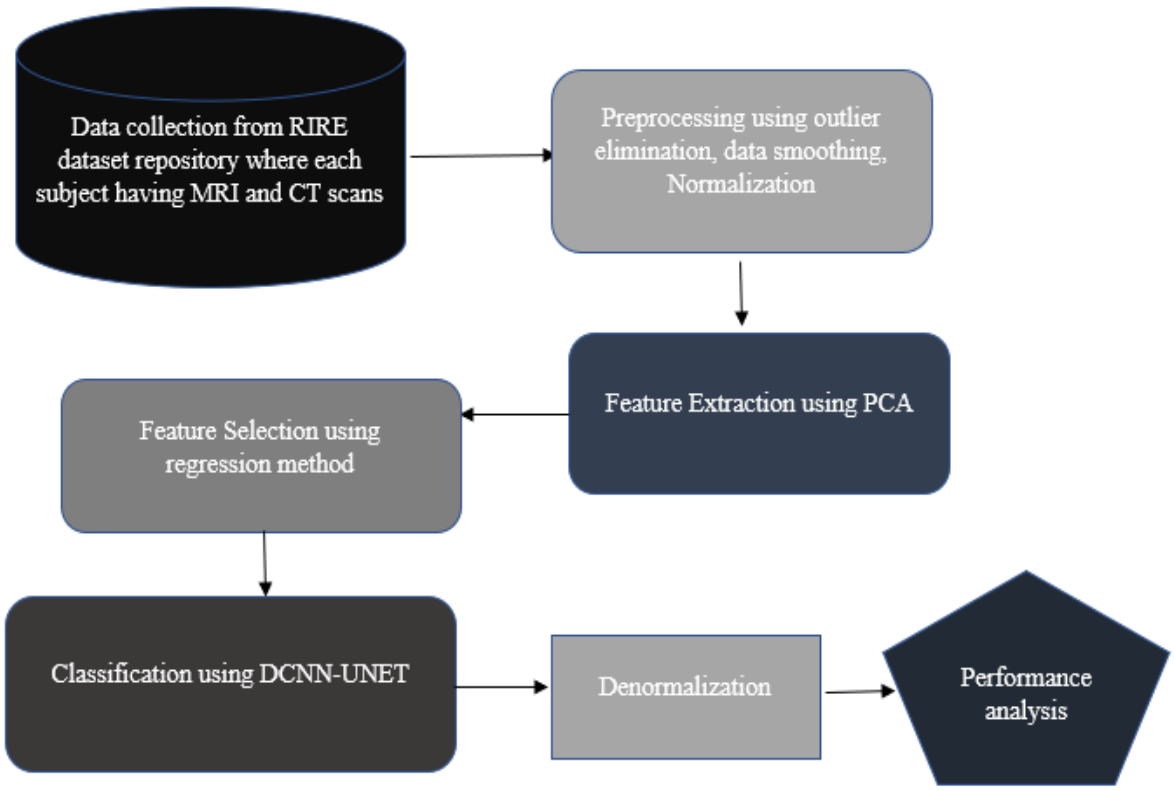

Figure 3 Proposed Block diagram for generating pCT

Figure 3 depict the block diagram of proposed system starts with data collection of both MRI and CT scan from available publicly RIRE dataset. A Magnetization Prepared Rapid Gradient Echo (MP-RAGE) sequence was used to produce T1-weighted MR data for this database. This is a fast echo-gradient procedure where a pre-preparation pulse is used to

increase contrast well before acquisition phase. Dimensions of the MR-data were 256x256x128 at 0.98mm × 0.99mm x 1.484mm on average. The matching CT scans were made of 3 mm diameter slice width with slice sizes 512 × 512, with a mean resolution of 0.419 mm x 0.419 mm. This raw data require pre-processing in order to eliminates noises and other anomalies with the help of techniques for outlier elimination, data smoothening and normalization. These pre-processed images are then utilized for extraction process using PCA and later, selection of features are done using regression method. Finally to generate pCT, DCNN-UNET is used.

Dataset Collection

We collected the MRI and CT scans from publicly available dataset called The Retrospective Image Registration Evaluation (RIRE) dataset (Dataset can be downloaded from: [28]). In the RIRE database, images from 19 individuals are included in T1 and T2. These subjects are trained and the rest are tested. For 800 epochs, 477 training images with 1.156 batch size are trained. Figure 4 (a, b) depict the CT and MRI respectively. No categories of brain tumor data set instances that is been used for this implementation.

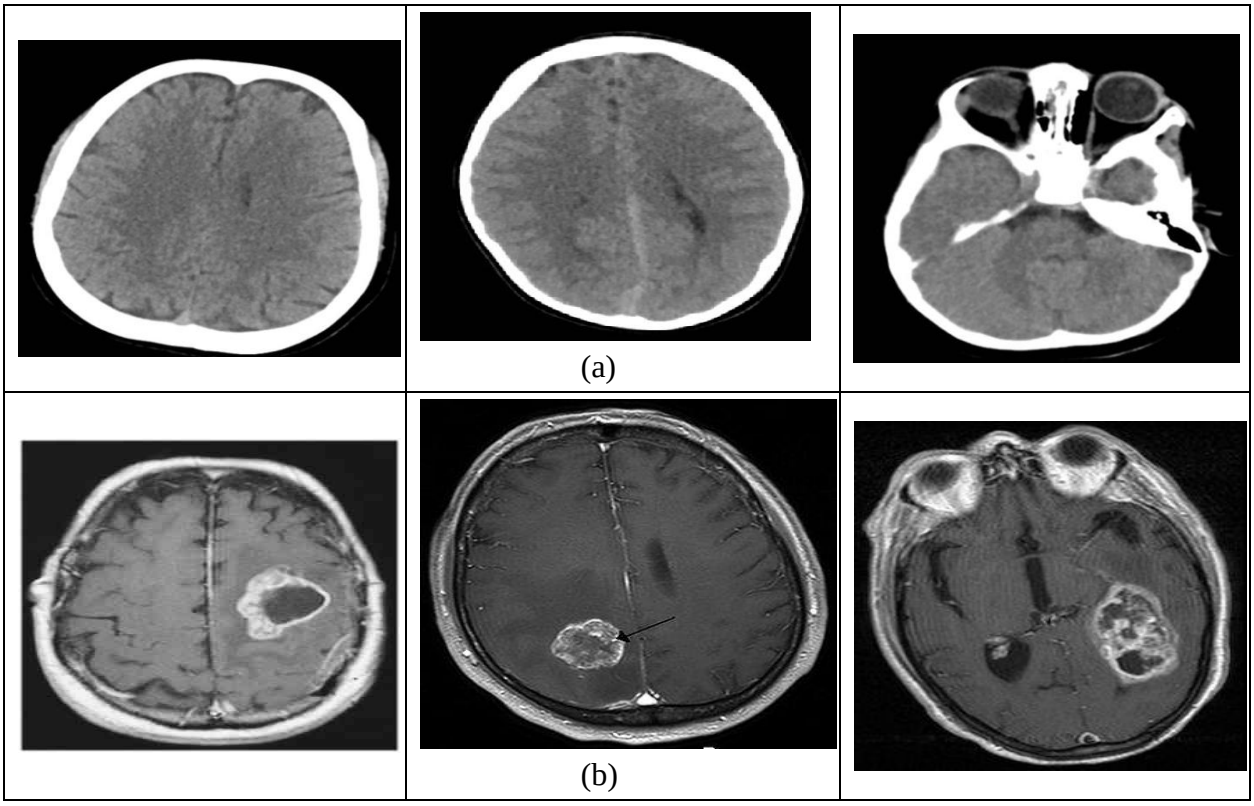

Figure 4 a) Instances of dataset showing CT, b) Instances of dataset showing MRI

Pre-processing

In preprocessing usually, image noise removal or image enhancement occurs. It helps to segment the part of the medical image affected by pathology. Zhang et al. (2016) has addressed these issues where the noise elimination was done by employing a median filter and the high-frequency factor shown in the picture. In addition, image enhancement was accomplished with the techniques of outlier elimination, data smoothing and normalization.

1. Outlier Elimination

Despite the availability of a wide range of state-of-the-art de-noising techniques, accurate noise reduction from MRI images remains a difficulty. A wavelet-based algorithm is used to denoise and preserve the true signal in the frequency domain. This builds the scaling coefficients freelance of the signal and therefore are often simply removed. We have used the Wavelet toolbox in MATLAB (2016) and lifting wavelet transform (LWT) (Lukas et al. (2004)) functions that are lwt2(), ilwt2(), and lwtcoef2(). These functions helped in removal of noise from MRI image. In addition, these functions facilitated to recover weak signals from noise and the processed image were cleaned up without blurring or losing the clarity.

2. Data Smoothing

A loss in spatial specificity is the biggest drawback of smoothing data. Smoothing distributes the signal in every subject, so we do not have to be as particular as to where it is located. This may or may not be a matter of worry depending on the spatial extent of your activation. Despite this decrease in spatial precision, smoothing MRI data is beneficial for a number of reasons. These may mostly be classified into statistical (smoothing helps you discover activation) reasons and unfair reasons (Gelvez et al. (2018)) (smoothing influences how you interpret your results). The Gaussian filter, where the dimensions of a gaussian smoothing kernel are usually represented as half its entire breadth (FWHM), was used for data smoothing. This blur is also significantly dependent on the standard variability employed in a Gaussian smoothing 3D convolution mask as well as on the size of that Gaussian mask or kernel. To this end, the synthetic base exposed to the procedure was considered as standard deviation.

$$G(i,j,k) = \frac{1}{(\sqrt{2\pi})^2\ \sigma_i\sigma_j\sigma_k} e^{-\left(\frac{i^2}{2\sigma_i^2}+\frac{j^2}{2\sigma_j^2}+\frac{k^2}{2\sigma_k^2}\right)}$$

3. Data Normalization

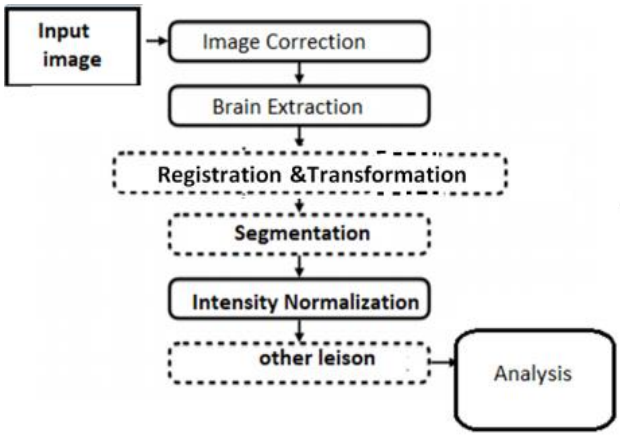

Figure 5 Pipeline of Intensity Normalization

T1-w, T2-w, PD, and FLAIR intensities are recorded in arbitrary units in conventional MRI (Lukas et al. (2004)). Though the same technique is followed, images are not consistent between scanners, subjects, or visits; and thereby algorithm performance, prediction, and inference are get affected. Moreover, the most basic tasks, such as thresholding a picture, may be complicated. However, the process of intensity normalization reduces the intensity differences between persons to a manageable level. The zscore_img is a function in neurobase, and uses mean WB and standard deviation WB computed from all voxels in the brain mask to Z-score each voxel.

$$T1WB = T1 - \mu WB\ \sigma WB$$

However, it requires an image as well as a binary mask. The brain mask's default setting is to use all voxels. The intensity acquired after normalization is seen in Figure 6.

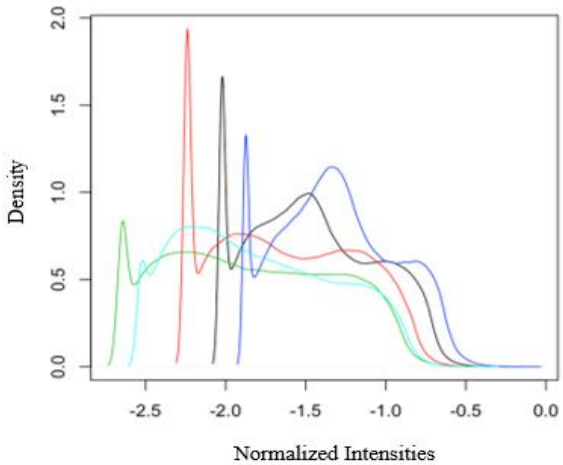

Figure 6 Normalization

Feature Extraction and Feature Selection

The Principal Component Analysis (PCA) is among the most effective image identification and encoding techniques. It is used to minimize the data's dimensionality and to extract features (Chandarana et al. (2018)). Feature extractions are carried out throughout both the training and testing phases. The basic goal of an MR image recognition system is to find as many similarities as possible between training and test MR images. Let Ω1 be the M x N pixel resolution training image of image 1 with a pixel resolution of M x N. (M rows, N columns). First, turn the image into a pixel vector Φ1 by concatenating each of the M rows into a single vector. The vector Φ1's length (or dimensions) will be M x N. The PCA algorithm transforms the vector Φ1 into a vector ω1 with a dimensionality of d, where d << M x N. These feature vectors ωi are computed and saved for each training image Ωi. Further, PCA is used to construct the feature vector ωj of the test image Ωj during the testing stage. The correlations between ωj and all of the feature vectors i'sωi's in the training set are calculated in order to identify the test image Ωj. The Euclidean distance is used to calculate the similarity between feature vectors. The image recognizer's output is the identification of the most comparable ωi. If i = j, the MR image j has been correctly recognized; otherwise, if i ≠ j, the MR image j as misclassified. For feature selection regression method is used in which features from subtype of the dataset is predicted based on the knowledge acquired from its counterpart. Moreover, Gradient descent algorithm is used for finding the parameter values for linear regression that outputs lowest values J (θ**)** value.

1. Modeling Relation of CT and MRI

The technique of voxel is based on the premise, at least for those voxels of the same tissue type, that there are a link between a CT number of a voxel and an MRI intensity. However, CT values may differ quite much from those of the same MRI value, even from the same cell type. Our goal is thus to get a second set of MRI images in the second set, which can enable us to differentiate these voxels and map them to matching CT numbers (Ronneberger et al. (2015)) with the same MRI value (in the first set), but with a high different of CT numbers.

All voxels within the same region are pooled together as training data to construct the regional prediction model for each section of the anatomical structure (bony, soft and mixed regions). Each voxel's data is comprised of its CT value, its standardized MR1 image intensity value S1, and its normalized MR2 image value S2. The prediction template is training to match the two MRI intensities of a voxel to its CT number for each area in the

HU scale. In this work, we employed a polynomial function in the nth two variable level which depends on the two MRIs of a voxel in order to forecast the predicted CT number of the voxel:

$$PCT(S_1, S_2) = \sum_{i_1, \lambda_2 - 0}^{i_1 + i_2 \leq n} c_{i_1 i} S_2^{h_1} S_2^{i_2} = c_{0,0} + c_{1,0} S_1 + c_{0,P} S_2 + c_{1,} S_1 S_2 + \cdots + c_{n,0} S_1^n + c_{0,x} S_2^n$$

MRI data from each area are divided into various value bins of MR1 and MR2 in order to decrease the computer overhead and noise. For MR1 and MR2, the same number of bins (Nbin) is used, and the bin width determined by the maximal normalized values for MR1 and MR2. Each voxel is thus assigned to one of two bin indexes, i for the MR1 bundle and j for the MR2 bundle, or (i j). In these two dimensions, there might be numerous voxels with different CT numbers within a single (ij) bin. The bin's average CT number may be described as

$$CT_{i,j}{}^{-} = \frac{1}{N_{ij}} \sum_{i=1}^{N_{i,j}} CT_{ij,k'}$$

Where i-th MR 1 bin and j-th MR2 bin have Nij and CTi j,k is the CT k-th voxel number inside that bin, and CTi j,k , the CT k-th voxel number in that bin. We found that pCT (s1.2) best reflects the average CT number CTij, as a function of S1 and S2 MRI intensities, after the regression analysis on the training data of each region for determination of Ci1-i2 in polynomial. Further, the 'NonlinearModelFit' function of Mathematica is used for regression analysis (Wolfram, Champaign, United States).

2. Generation of pCT

We have designed a DCNN model to train a mapping capability directly in this study. Each patient's MR/CT couple with suitable CT slices can be trained by collecting all MR slices. Once the model has been trained, the findings are slice-by-slice applied to a fresh MR image and the final pCT may be constructed. Due to restrictions on GPU commodity GPU memory and because of insufficient training data in this work, a comprehensive DCNN model seems ineffective. It may sometimes be unneeded as a slice provides extensive contextual data.

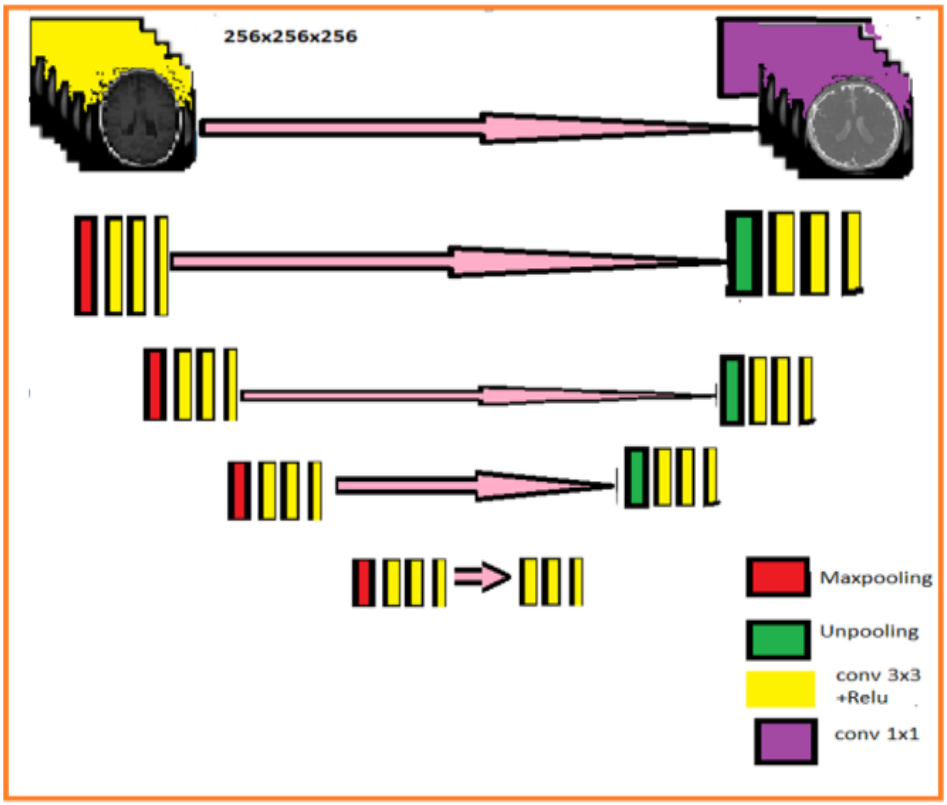

Figure 7 Overall architecture of the proposed pCT DCNN model

Figure 7 illustrates the model DCNN pCT with a convolutional layer (3x3) in each yellow box (with a rectified linear unit as the activation function). Every red box shows the max- poolling layer and every green box shows the un-pooling layer. The 2D image size and the depth of the feature map of each convolutional layer are at the top of each yellow box. The purple box shows the final 1x1 convolutional layer which produces the pCT output forecast (Chen et al. (2014)). The model consists of two primary components: the encoding part (left half) and decoding part, similar to Ronneberger et al. (2015) (right half). The encoding is a classical CNN that can extract from an input MR image a more complex hierarchy of features. The decoding element transforms features and recreates the pCT prediction in low to high resolution. The complete network output is a 2D image with the same size as the input image.

The establishment of a direct link is a significant invention from (Ronneberger et al. (2015)) that we implemented here (in Figure 7, as pink arrows). This will enable the convolutional layers of the second portion to exploit the previous part's high-resolution features. This method makes it easier to create high-resolution predictions for the decoding component. If high resolution (top of network) features are sufficient to produce trustworthy CT predictions, the model may learn to ignore ground level features intelligently (at the bottom

of the network). The functional extraction element of our model may be initialized by copying current VGG model weights learnt on a wide variety of image data by duplicating an existing architecture (i.e., 1.3 million natural images consisting of 1000 different object categories (Simonyan et al. (2014)). While the VGG model was trained on entirely different images, the low to moderate capabilities gained in a CNN model are pretty generic, and using a well-trained model to start a new task is an effective way to help us find good solutions with less training information. The components of the network in Figure 7 are explained in the following sections. Figure 8 depicts the U-net design.

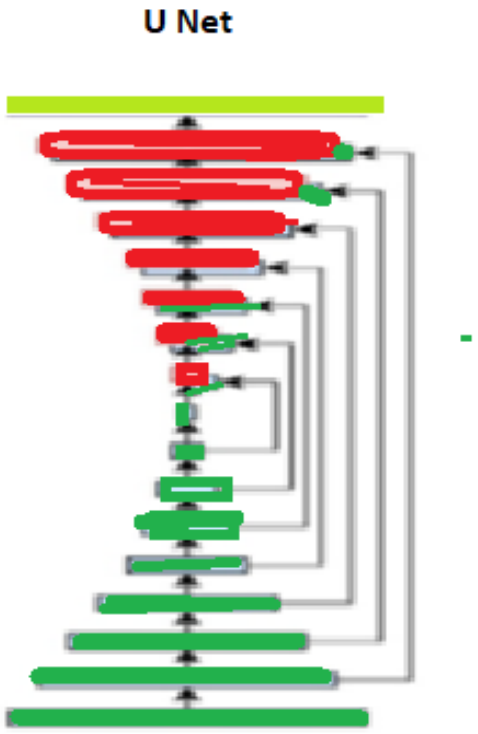

Figure 8 Diagrammatic representation of U-net

The encoding component is used to construct a hierarchy of features by using a range of convolution and pooling layers. The basic component of a CNN model is a convolutional layer. Each convolutional layer carries out 2D input convolutions using filters that run via a non-linear activation function. The operation is represented mathematically as,

$$H_k = \max (0, W_k * X + b_k), k \in [0, k-1]$$

Where W_k and b_k denote the i-th filter weights and biases, respectively, while '*' denotes the convolution process. The filter index in the total number of filters is indicated by the subscript k, k [0, K-1]. The entry is denoted by X, while the output of the k-th filter, also

known as the k-th channel of the convolutional layer output feature map, is denoted by H_k. The max (0,·), primary for the convolution output, corresponds to the non-linear function of the rectified linear unit (ReLU). Input X for the first convolutional layer is the MR image directly. The input is for the following layers from the output characteristic map. In every convolutional layer the number of filters is given, however, the weighting and bias (W_k and B_k's) are free.

One of the key aspects of CNNs' success is the ability to learn the weight and the preference of every function map instead than providing data-driven functional extractors relying on the set of manually produced functions. 3x3 filters are utilized for every convolutional layer according to the VGG model in the coding pathway. It should be noted that 3x3 refers just to the spatial size of each convergence filter. The same number of components as each channel is a vector for every weight. If R input channels, total number of free layer parameters equal to Kx3x3xR+K are accessible. We utilize zero-padding to keep the output (spatial) size of the input for this convolution. The output characteristic map shows the 2D size and number of channels on each layer.

Figure 7 shows that the function map needs a considerable amount of computer storage since every character map comprises dozens or hundreds of channels, directly related to the size of the image. The memory demand is a limiting issue which limits the creation of a DCNN model directly since there are large number of filters or channels to extract functions. In addition, more parameters in a DCNN model as well as more training data are needed. A max-pooling operational layer is used after 2 to 3 convolutional layers to lower the spatial resolving of the feature maps, thus enabling succeeding convolutional layers to learn features of a broader image context. The pooling process also allows features to be generated which are invariant for local disruptions. A 2x2 window and step 2 pool is used in the VGG model (non-overlapping window). This procedure includes the maximal functionality over 2x2 sub-windows of non-overlap size across each channel. The following can be formally formulated:

$$Z_{k,\,i,\,j} = \max_{p,\,q \,\epsilon\, [0,\,1]^2} h_{k,\,2i+p,\,2j+q}$$

Where (i,j) the output map's spatial index and (k) the channel index. In each spatial dimension the max-pooling process reduces the size of the map by half and figure 7 describes this. There are no free parameters to learn for the maximum pooling layer. However, the places where the greatest value is achieved in each pooling window must be saved for subsequent unpooling operations as a binary mask of the same size (Zeiler et al. 2014)), and illustrated in Figure 9.

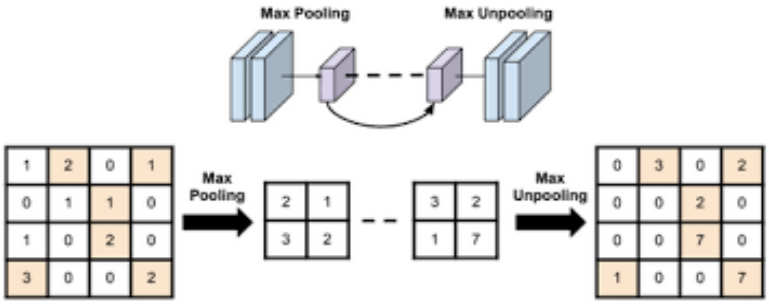

Figure 9 Pooling and Unpooling operations

Our DCNN model decoding section shall be selected to reflect the encoding portion identical to (Ronneberger et al. (2015) and (Simonyan et al. (2014)) with additional layer of 1x1, at the end of which each vector of 64 elements in the preceding layer mapped as a CT-number. In contrast to the extraction feature which uses max-based layers to lower the spatial resolution of function maps, unpolitical layers are used for decoding information from broad to narrower resolution. The accompanying max-pooling layer's un-pooling can be considered a reversal action. It uses the binary mask to double the spatial dimension of the input characteristic and instantaneously transfer the characteristic value from the input to the output at the maximum position for each pooling window, as illustrated in Figure 9. The remainder of the output is made up of nulls.

The un-pooling layer produces a little amount of output. Further, the convolutional layers use convergence mechanisms to transform the thin output into thicker representations. However, it was noticed a set of learning parameters separates all of these convolutional levels of the decoding component from the weight of the encoding part. Moreover, additional connections are established, as described in (Badrinarayanan et al. (2015)), to duplicate high resolution features from the encoding part and mix them with the output of un-pooling before they are transmitted to subsequent convolutions, as shown in figure 7. According to (Badrinarayanan et al. (2015)). When the encoding and decoding sections are combined, the network has a total of 27 layers and about 34.9 million parameters.

The whole network may be viewed as a sophisticated end-to-end mapping function that transforms an MR image to a corresponding CT image. The end-to-end mapping function is learned by minimizing a loss or prediction error between the predicted pictures F (θ; X) and the corresponding ground truth CT images Y. Further, estimation of network parameters θ= {W_1, b_1, W_2, b_2…} is necessary to learn the end-to-end mapping function. For a set of MR images Xi and their corresponding CT images Yi, use Mean Absolute Error (MAE) as the loss function:

$$L(\theta) = \frac{1}{N}\sum_{i=1}^{N} \| Y_i - F(X_i;\theta) \|$$

where N is the total number of training images. When employed as a losing function for MAE (L1-norm), the learning becomes more resistant to external data such as noise or other artefacts in the images, or due to poor matching of MR and CT image.

3. Implementation of Model

This research work utilizes the MATLAB programming to implement the work flow. The MR images acquired enable accurate and quick pseudo-CT creation utilizing perceptual loss through a generative adversarial network (GAN). Therefore, these images might be employed in clinical practice for dosage planning.

Validation of DCNN

To test the performance of the proposed DCNN approach a 6-fold cross-validation process was employed. In 6 groups of the same size, the 19 patients have changed. The other five groups will be used as training data for the training of a DCNN-model, with one group as a test set. The pCT is applied to each test participant's MR-image once the model is trained. The accuracy of the projected pCT is evaluated with an absolute mean error in the head region against the real CT for each participant.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |CT(i) - pCT(i)|$$

Where n is the overall MR head area voxel number. The head-region is defined by a binary mask produced automatically from the MR-image. Same as medium squared error (ME), the average voxel error (ME) are calculated:

$$ME = \frac{1}{N}\sum_{i=1}^{N} \left(CT(i) - pCT(i)\right)$$

and

$$MSE = \frac{1}{N}\sum_{l=1}^{N} |CT(i) - pCT(i)|^2$$

Performance Analysis

The model was built with the help of hardware specifications such Windows 10 OS, CPU, AMD Ryzen 7 5800H – 9 5900HS, Graphics: Nvidia GeForce RTX 3060 – 3080, RAM: up to 32 GB for the MATLAB model. With the data divided into 48 iterated batches, which takes 50 to complete all the training samples, known as a neural network training epoch,

one PCT from MRI takes 37 seconds. To evaluate the accuracy of the study described, a voxel-like comparison of the HU from CT and synthetic CT was made. Further, the mean absolute error (MAE) and the mean error (ME) for the whole body were calculated. The comparison between the pCT-CT image and the respective actual CT image for the structural similarity (SSIM) index with DICE coefficients (DC) shows the stereo-compatibility technique being superior to the registration method. In addition, utilizing Atlas and learning techniques, we analyzed our proposed DCNN-Unet. The DC and SSIM values of the different models are shown in Table 1. Figure 11 shows DC values of several models graphically represented (Gholamiankhah et al. (2021)).

Table 1 DC & SSIM values

Models	DC	SSIM
Atlas method	0.82±0.04	0.84±0.06
Learning-based methods	0.85±0.04	0.82±0.06
DCNN-UNET (Ours)	0.94±0.02	0.86±0.05

Figure 11 DC, SSIM vs. Models

Table 2 depict the MAE, MSE values of various models under final subject results. Figure 12 (a, b) displays the graphical representation of MAE values (HU) MAE values (HU) under final subjects. Figure 13 depict the graphical representation of 19 subjects flow of various models. Figure 14 depict the graphical representation of computation time taken by various models. Figure 15 depict the accuracy comparison of various models (Han et al. (2017)). Figure 16 (a, b) shows the normal MRI and pCT from MRI image.

Table 2 MAE, MSE values of various models

Models	MAE (HU)	MSE (HU)
Atlas method	76.7±17.8	165.3±33.0
Learning-based methods	78.0±0.8	114.1±27.5
DCNN-UNET (Ours)	67.5±17.3	188±33.7

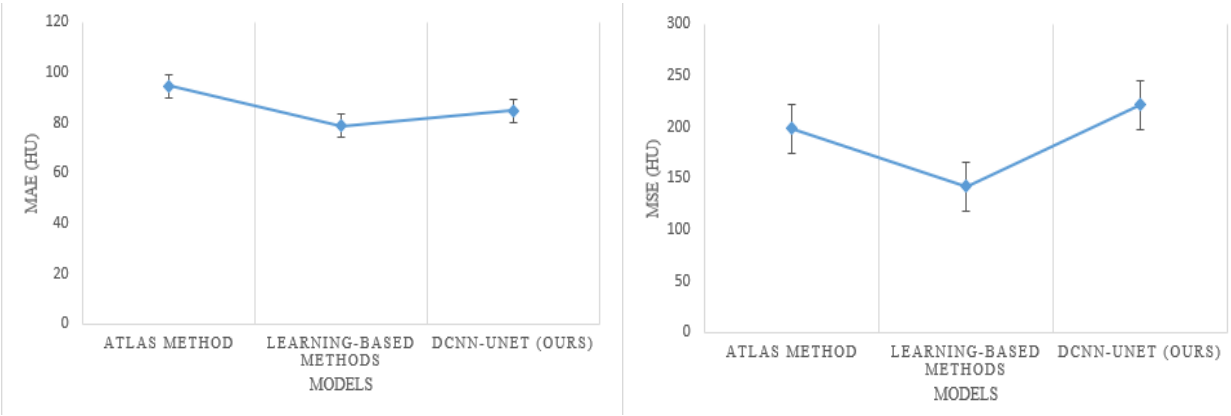

Figure 12 a) MAE (HU) vs Models, b) MSE (HU) vs Models

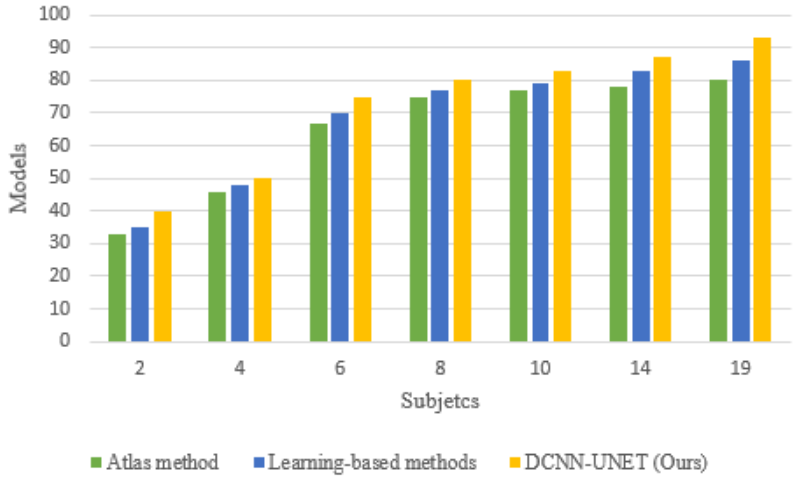

Figure 13 Models' vs Subjects based performance

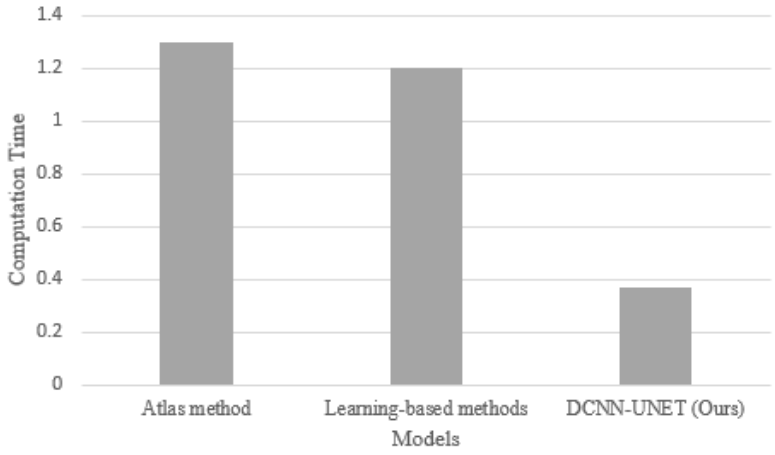

Figure 14 Models vs Computation Time per pCT generation

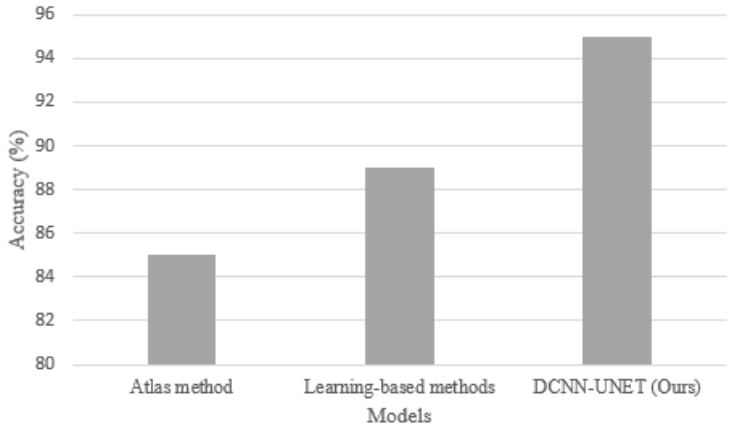

Figure 15 Models vs Accuracy

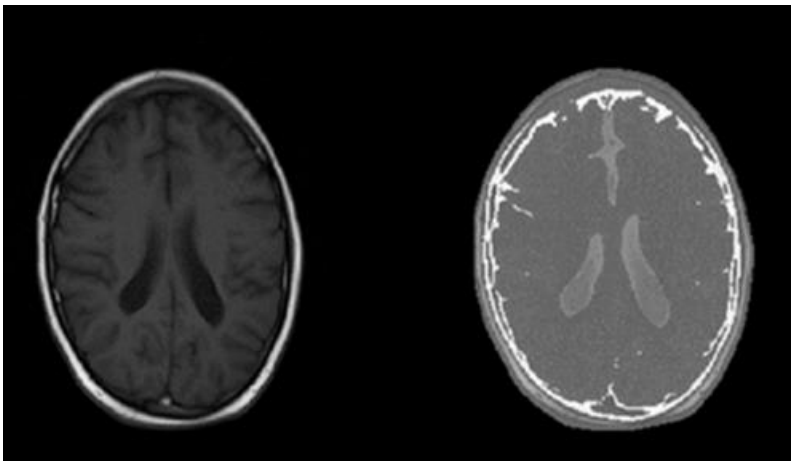

Figure 16 a) Normal MRI scan image, b) PCT CT image from MRI

Discussion

This study proposes a DCNN model for the use of standard T1-weighted MR images with the aim of synthesized CT production. Although a profound neural network pattern usually needs a significant quantity of training data, the usage of the existing pre-trained model using the transfer learning principle achieves very high performance with our little data. The method of DCNN does not require a linear or deformable image recording between individuals and quickly learns how to map the corresponding CT image from inside the MR image area. The results of the assessment showed that a DCNN approach with atlas refining and fusion is significantly higher than that of an Atlas technique. The patch-based technique is intended to be based on a comparison of the various atlases or patch basic approaches reported in literature in order to detect comparable applicants. It's usual as well. A small local patch comprises extremely limited image data and employs the crude intensity of an image to reduce discrimination since the characteristics of the data may be very repetitious. Rather, the DCNN model learns a hierarchy of image properties of diverse sizes and

complexity from a whole image segment. In comparison with previous models our model offers HU an MAE of 67.5±17.3, which is considerably good.

A major advantage of the DCNN approach is the fast computation time during model deployment. Although a model may take days to train, it has to be done once only and for all by multiple GPUs. It takes only a couple of seconds for a single GPU to apply the pCT model. Methods based upon model or atlas might be rather sluggish, otherwise. The DCNN approach also benefits from the fact that a huge quantity of the training data may be readily accommodated. In fact, because of their enormous model capacity, deep CNN models are generally recognized to substantially benefit from big data. Therefore, we anticipate that the performance of the proposed DCNN technology can considerably increase when more training data become accessible. The load is just longer because of an increasing quantity of information on the model, however, the size and speed of the final model remain the same.

Conclusion

In order to eliminate the unnecessary irradiations and workloads, a effective DL model is need for generating pCT from MRI. Hence we proposed DCNN-UNET for generating pCT from MRI which includes process of data collection from RIRE, then preprocessing to avoid noises and normalizing the data. Segmentation of images to gains the intensity of voxels and extracting these images based on the intensity values and finally generating effective pCT from MRI using DCNN model in which while evaluating our proposed model outperforms better with measures like MSE, MAE, DC, SSIM and gains an accuracy of 95% over models like atlas based and learning based. This paper will be also useful for other researchers who can dig deep to find interesting new concepts and thereby increasing the efficient of the DL models.

References

Chandarana, H., Wang, H., Tijssen, R.H.N., & Das, I.J. (2018). Emerging role of MRI in radiation therapy. *Journal of Magnetic Resonance Imaging, 48*(6), 1468-1478.

Johnstone, E., Wyatt, J.J., Henry, A.M., Short, S.C., Sebag-Montefiore, D., Murray, L., & Speight, R. (2018). Systematic review of synthetic computed tomography generation methodologies for use in magnetic resonance imaging–only radiation therapy. *International Journal of Radiation Oncology* Biology* Physics, 100*(1), 199-217.

Adjeiwaah, M., Bylund, M., Lundman, J.A., Söderström, K., Zackrisson, B., Jonsson, J.H., & Nyholm, T. (2019). Dosimetric impact of MRI distortions: a study on head and neck cancers. *International Journal of Radiation Oncology* Biology* Physics, 103*(4), 994-1003.

Dean, C.J., Sykes, J.R., Cooper, R.A., Hatfield, P., Carey, B., Swift, S., & Morgan, A.M. (2012). An evaluation of four CT–MRI co-registration techniques for radiotherapy treatment planning of prone rectal cancer patients. *The British journal of radiology, 85*(1009), 61-68.

Edmund, J.M., & Nyholm, T. (2017). A review of substitute CT generation for MRI-only radiation therapy. *Radiation Oncology, 12*(1), 1-15.

Arabi, H., & Zaidi, H. (2016). Whole-body bone segmentation from MRI for PET/MRI attenuation correction using shape-based averaging. *Medical physics, 43*(11), 5848-5861.

Arabi, H., & Zaidi, H. (2017). Comparison of atlas-based techniques for whole-body bone segmentation. *Medical image analysis, 36,* 98-112.

Arabi, H., & Zaidi, H. (2016). Magnetic resonance imaging-guided attenuation correction in whole-body PET/MRI using a sorted atlas approach. *Medical image analysis, 31,* 1-15.

Mehranian, A., Arabi, H., & Zaidi, H. (2016). Quantitative analysis of MRI-guided attenuation correction techniques in time-of-flight brain PET/MRI. *Neuroimage, 130,* 123-133.

Arabi, H., & Zaidi, H. (2021). Non-local mean denoising using multiple PET reconstructions. *Annals of nuclear medicine, 35*(2), 176-186.

Zheng, W., Kim, J.P., Kadbi, M., Movsas, B., Chetty, I.J., & Glide-Hurst, C.K. (2015). Magnetic resonance–based automatic air segmentation for generation of synthetic computed tomography scans in the head region. *International Journal of Radiation Oncology* Biology* Physics, 93*(3), 497-506.

Edmund, J.M., Kjer, H.M., Van Leemput, K., Hansen, R.H., Andersen, J.A., & Andreasen, D. (2014). A voxel-based investigation for MRI-only radiotherapy of the brain using ultra short echo times. *Physics in Medicine & Biology, 59*(23), 7501-7519.

Kapanen, M., & Tenhunen, M. (2013). T1/T2*-weighted MRI provides clinically relevant pseudo-CT density data for the pelvic bones in MRI-only based radiotherapy treatment planning. *Acta Oncologica, 52*(3), 612-618.

Korhonen, J., Kapanen, M., Keyriläinen, J., Seppälä, T., & Tenhunen, M. (2014). A dual model HU conversion from MRI intensity values within and outside of bone segment for MRI-based radiotherapy treatment planning of prostate cancer. *Medical physics, 41*(1), 011704.

Johansson, A., Garpebring, A., Karlsson, M., Asklund, T., & Nyholm, T. (2013). Improved quality of computed tomography substitute derived from magnetic resonance (MR) data by incorporation of spatial information–potential application for MR-only radiotherapy and attenuation correction in positron emission tomography. *Acta oncologica, 52*(7), 1369-1373.

Johansson, A., Karlsson, M., & Nyholm, T. (2011). CT substitute derived from MRI sequences with ultrashort echo time. *Medical physics, 38*(5), 2708-2714.

Navalpakkam, B.K., Braun, H., Kuwert, T., & Quick, H.H. (2013). Magnetic resonance–based attenuation correction for PET/MR hybrid imaging using continuous valued attenuation maps. *Investigative radiology, 48*(5), 323-332.

Rank, C.M., Hünemohr, N., Nagel, A.M., Röthke, M.C., Jäkel, O., & Greilich, S. (2013). MRI-based simulation of treatment plans for ion radiotherapy in the brain region. *Radiotherapy and Oncology, 109*(3), 414-418.

Rank, C.M., Tremmel, C., Hünemohr, N., Nagel, A.M., Jäkel, O., & Greilich, S. (2013). MRI-based treatment plan simulation and adaptation for ion radiotherapy using a classification-based approach. *Radiation Oncology*, *8*(1), 1-13.

Korsholm, M.E., Waring, L.W., & Edmund, J.M. (2014). A criterion for the reliable use of MRI-only radiotherapy. *Radiation Oncology*, *9*(1), 1-7.

Stanescu, T., Jans, H.S., Pervez, N., Stavrev, P., & Fallone, B.G. (2008). A study on the magnetic resonance imaging (MRI)-based radiation treatment planning of intracranial lesions. *Physics in Medicine & Biology*, *53*(13), 3579.

Kemppainen, R., Suilamo, S., Tuokkola, T., Lindholm, P., Deppe, M.H., & Keyriläinen, J. (2017). Magnetic resonance-only simulation and dose calculation in external beam radiation therapy: a feasibility study for pelvic cancers. *Acta Oncologica*, *56*(6), 792-798.

Sjölund, J., Forsberg, D., Andersson, M., & Knutsson, H. (2015). Generating patient specific pseudo-CT of the head from MR using atlas-based regression. *Physics in Medicine & Biology*, *60*(2), 825.

Greer, P., Martin, J., Sidhom, M., Hunter, P., Pichler, P., Choi, J.H., & Dowling, J. (2019). A multi-center prospective study for implementation of an MRI-only prostate treatment planning workflow. *Frontiers in oncology*, *9*, 826.

Oktay, O., Nanavati, J., Schwaighofer, A., Carter, D., Bristow, M., Tanno, R., & Nori, A. (2020). Evaluation of deep learning to augment image-guided radiotherapy for head and neck and prostate cancers. *JAMA network open*, *3*(11).

Shafai-Erfani, G., Wang, T., Lei, Y., Tian, S., Patel, P., Jani, A.B., & Yang, X. (2019). Dose evaluation of MRI-based synthetic CT generated using a machine learning method for prostate cancer radiotherapy. *Medical Dosimetry*, *44*(4), e64-e70.

Hemsley, M., Chugh, B., Ruschin, M., Lee, Y., Tseng, C.L., Stanisz, G., & Lau, A. (2020, October). Deep Generative Model for Synthetic-CT Generation with Uncertainty Predictions. *In International Conference on Medical Image Computing and Computer-Assisted Intervention*, *Springer, Cham*, 834-844.

Zhang, S., & Xu, G. (2016). A novel approach for brain tumor detection using MRI Images. *Journal of Biomedical Science and Engineering*, *9*(10), 44-52.

MATLAB tool (2016)-09806-07443-53955-64350-21751-41297

Lukas, L., Devos, A., Suykens, J.A., Vanhamme, L., Howe, F.A., Majós, C., & Van Huffel, S. (2004). Brain tumor classification based on long echo proton MRS signals. *Artificial intelligence in medicine*, *31*(1), 73-89.

Gelvez, E., Vera, M., Huérfano, Y., Valbuena, O., Salazar, W., Vera, M.I., & Contreras, Y. (2018). Smoothing filters in synthetic cerebral magnetic resonance images: *A comparative study*, *13*(4), 187-190.

Zhou, Y. (2017). *Generating virtual CT from MRI using fully convolutional neural networks with improved structure quality.*

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *In International Conference on Medical image computing and computer-assisted intervention*, 234-241.

Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A.L. (2014). Semantic image segmentation with deep convolutional nets and fully connected crfs. *arXiv preprint arXiv:1412.7062.*

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556.*

Zeiler, M.D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. *In European conference on computer vision*, 818-833.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556.*

Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, *39*(12), 2481-2495.

Gholamiankhah, F., Mostafapour, S., & Arabi, H. (2021). Deep learning-based synthetic CT generation from MR images: comparison of generative adversarial and residual neural networks. *arXiv preprint arXiv:2103.01609.*

Han, X. (2017). MR-based synthetic CT generation using a deep convolutional neural network method. *Medical physics*, *44*(4), 1408-1419.

Agrahari, A., Chaudhary, C.P., & Singh, S.N. (2018). Domain analysis of D-lib magazine: A bibliometric study. *Webology*, *15*(1), 61-76.