Feasibility of Genetic Programming for the Optimization of Tissue-Type-Segmented Maps in the Generation of Synthetic CT in Radiation Therapy Treatment Planning

Matthew Witten and Owen Clancey
Department of Radiation Oncology
NYU Winthrop Hospital
Mineola, New York, USA
Matthew.Witten@nyulangone.org, Owen.Clancey@nyulangone.org

***Abstract*—Modern radiation therapy treatment planning has traditionally relied upon computed tomography (CT) in modeling the interaction of megavoltage (MV) photons with the various tissues and organs of the body. CT image data provide both detailed information about individual patient anatomy, as well as a voxel-by-voxel three-dimensional grid of Hounsfield units (HU), which specifies, at all points within the patient, essentially the difference between the attenuation coefficient of the tissue within that voxel from the attenuation coefficient of water, normalized to that of water. From the HU value, the relative electron density can be inferred, and as the relative electron density is the major determinant of the interaction of the tissue with MV photons, the radiation dose distribution can then be calculated. Recently, there has been interest in using magnetic resonance imaging (MR) in lieu of CT, as MR provides superior soft-tissue contrast; however, MR does not provide any electron density information. Various approaches have been essayed to create a synthetic CT (synCT) from the MR data. In the present study, genetic programming was used to construct mappings of MR data to HU data for seven tissue types: bladder, cancellous bone, cortical bone, fat, muscle, prostate, and rectum. These maps were then applied to randomly chosen points in five patient data sets to calculate the synCT HU values, which were then compared with the actual HU values from CT images of those same patients. The method produced mean absolute errors (MAE) of 9.28 HU, 33.24 HU, 75.32 HU, 18.64 HU, 17.12 HU, 11.76 HU, and 18.40 HU for the respective tissue types, and these MAE values are less than those of previous approaches, indicating superior performance. Although the method of the present study does require more manual input, the superior performance is compensatory. Further study is necessary to confirm accuracy on entire MR data sets, and to ensure there is no sample variance effect on the current results.**

Keywords—genetic programming; radiation therapy; synthetic CT

I. INTRODUCTION

At the outset, it should be noted that the intent of the authors is to address a real-world problem using genetic programming methods. Though the optimization will be discussed in detail, the present study eschews more theoretical discussions of operators, recombination, termination conditions, bloat control, and the like. The goals here are to solve the problem of generation of synthetic computed tomography (synCT) data from magnetic resonance imaging (MR) data for the purposes of radiation therapy treatment planning, and to compare the ultimate performance of the genetic programming method adopted in the present study to that of the other prominent approaches.

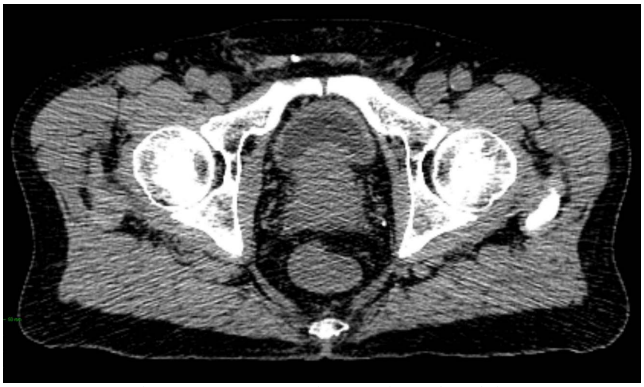

Fig. 1 Transverse CT image of a male pelvis.

Radiation therapy is most often delivered via high-energy megavoltage (MV) photons. In this energy range, the interaction of the radiation with tissue is governed by inelastic Compton scattering, wherein a photon imparts a portion of its energy to a liberated electron, and that electron in turn deposits energy locally. This local deposition of energy is commonly referred to as radiation dose. The deposition of radiation dose induces biological effects, including apoptosis (cell death), which is the direct cause of the killing of tumors and of injury to normal organs. The magnitude of the Compton scattering interaction is directly related to the electron density of the tissue, so that a knowledge of the differential distribution of electron densities within the body can be used to calculate the radiation dose distribution.

978-1-7281-6929-3/20/$31.00 ©2020 IEEE

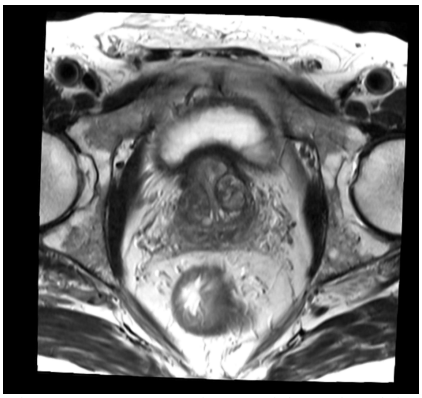

Fig. 2 Transverse MR image of the male pelvis.

The distribution of relative electron densities can be calculated directly from CT data. CT data sets contain, at each point in the patient, the Hounsfield unit (HU) value, defined in (1):

$$HU = 1000 * \frac{(\mu - \mu_{water})}{\mu_{water}}, \quad (1)$$

where HU is the Hounsfield unit value, μ is the linear attenuation coefficient of the tissue at the point in question, and μ_{water} is the linear attenuation coefficient of water. In addition, CT imaging provides detail regarding the internal anatomy of the patient, so that the radiation oncologist can segment (or contour), either manually or supplemented with some automation, the tumor and critical organs. For these reasons, CT imaging has become integral to the radiation therapy planning process. An example of a transverse CT image of the male pelvis is displayed in Fig. 1 on the previous page.

MR imaging provides superior soft-tissue contrast to that of CT, and as such, MR studies of the patient are registered with CT studies to provide aid to the radiation oncologist in segmenting the tumor and other organs. The MR is necessary for accurate delineation of the volume of the structures, and the CT is necessary in order to calculate the radiation dose distribution. An example of a transverse MR image is shown in Fig. 2. Because the registration of the MR and CT data sets introduces geometric error, there has been much interest in MR-only radiation therapy treatment planning [1-5]. These approaches use atlas-based methods and complicated algorithms, such as neural networks [2].

The present study is an attempt to address the question: is there a relatively simple alternative method, without the necessity of the cumbersome and lengthy process of training a neural network, and using the existing workflow in a typical radiation oncology department, to generate accurate synthetic CT data directly from an MR study of a patient? The approach adopted was to utilize genetic programming and optimize separate MR-to-HU maps for seven tissue types (bladder, cancellous bone, cortical bone, fat, muscle, prostate, and rectum) to generate synCT HU values. This novel method uses a previously untried optimization of a direct functional MR-to-HU mapping. The result is an approach that yielded more accurate HU values than the current methods.

II. Materials And Methods

A. CT and MR Image Data

CT data were acquired with a 16-slice Optima 580RT CT scanner (GE Healthcare Systems, Chicago, IL). MR data sets were acquired with a 1.5T Magnetom MRI scanner (Siemens Medical Solutions USA, Malvern, PA). Image data for patients undergoing radiation therapy for carcinoma of the prostate were used. A total of ten pairs of CT-MR image sets for ten distinct patients were utilized; five pairs were used to establish the MR-to-HU tissue-specific maps, and five pairs were used to test the accuracy of the synCT HU values calculated from the tissue-specific maps.

B. Registration, Segmentation, and Point Sampling

CT-MR image registration was accomplished via computer-aided fusion supplemented with manual adjustment by the radiation oncologist. The studies were imported into the Precision v1.1 treatment planning software (Accuray, Inc., Sunnyvale, CA), and a rigid-body automated registration was applied to align the image sets. Further fine-tuning of the registration was performed manually by the radiation oncologist, so that the registration of the soft tissues (e.g. prostate, bladder, and rectum) was optimized.

Subsequent to the image registration process, the radiation oncologist segmented (i.e. outlined by manipulating software drawing tools) the organs: bladder, prostate, rectum, femoral heads, and bowel. It is important to note that all these steps are part of the normal treatment planning workflow, so thus far no sacrifices in planning throughput have been made to accommodate the experiment carried out in this study.

In anticipation of the need to optimize the tissue-specific MR-to-HU maps, points were selected at random from five of the registered CT-MR pairs. For each tissue type, the MR pixel intensity value was recorded, as well as the corresponding CT HU value. A number of points were sampled for each tissue-type: bladder (34), cancellous bone (12), cortical bone (32), fat (36), muscle (38), prostate (49), rectum (35). The number of points sampled reflected the perceived variation in HU range, so that for each tissue type, the full range was captured in the set of sampled points.

C. Hardware and Software for Genetic Programming

All GP optimization was performed on a ThinkPad P50 laptop (Lenovo, Beijing, CN) with 32 GB RAM and an Intel Xeon 8-core processor running MATLAB R2016b (The Mathworks, Natick, MA).

D. Genetic Programming

For each tissue-type, a mapping, f_{tissue} was optimized such that:

$$f_{tissue}: f_{tissue}(MR) = HU. \quad (2)$$

In (2), $MR \in \mathbb{Z}$ is the intensity of the pixel in the MR image, and $HU \in \mathbb{Z}$ is the HU value at the pixel in the registered CT data set.

At this point, it was apparent that the nature of the problem was two-fold: classify the tissue type, and after correct classification, use the correct map to predict the HU value. Since the classification was achieved by the segmentation of the organs, what remains is to optimize the mapping. The problem now reduces to a symbolic regression, for which GP is well-suited [6].

For each tissue-type, a GP optimization to determine the mapping was performed. In each case, the fitness function was the sum of the squares of the deviations of the predicted synthetic HU values from the actual HU values, as in (3) below:

$$f_{fitness} = \sum\left(HU_{synthetic} - HU_{actual}\right)^2. \quad (3)$$

The function set selected was:

$$\{+, -, *, sin, cos, log\}, \quad (4)$$

while the terminals were just the MR pixel intensities, namely:

$$\{MR: MR \in \mathbb{Z}\}. \quad (5)$$

The GP implemented was that of GPLAB v4. The population size was 100, and the number of generations for each algorithm run was 100. The trees were created using ramped half-and-half initialization. Tree growth was initially controlled by setting the allowed maximum depth equal to 6.

The reproduction rate was set to 0.1, i.e. this was the rate at which individuals were directly copied into the succeeding population in the next generation in the absence of being acted upon by the genetic operators [7]. The sampling method for parent selection was a modified tournament selection operator called "lexictour" [7], which uses, in a lexicographic ordering, a primary (evaluated fitness) and secondary (tree size) objective. According to Silva [7], the technique is effective in problems where many individuals are expected to have equal fitness values. The expected number of children for each individual was calculated using the "rank85" method, so that each individual's expected number of children was determined by its rank in the population [7]. There was no elitism, and so the population was replaced in each generation.

The probability of crossover was set at a static value of 0.5, as was the probability of mutation. The crossover operator was the usual swap of branches from trees of parents subsequent to random node selection. The mutation operator acted to create point mutations.

Bloat was controlled with both a fixed maximum tree depth, set at 17, and a dynamic maximum tree depth, set at 6.

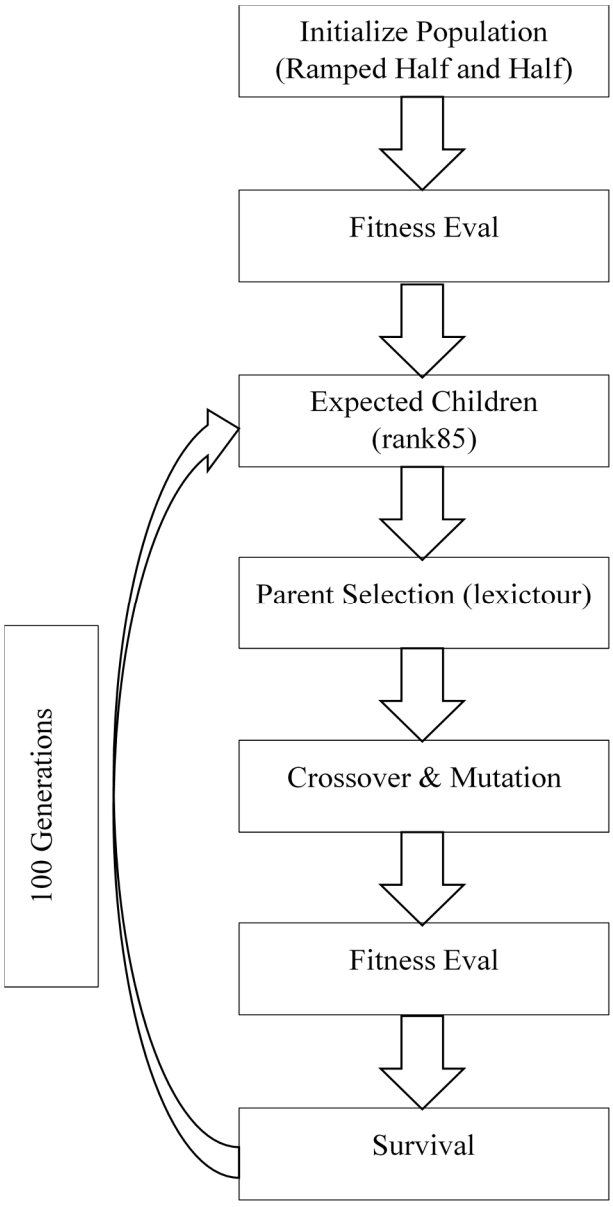

Fig. 3 Flow diagram for GP optimization

As explained by Silva [7], the dynamic depth technique has been demonstrated to show better results than that of lexicographic parsimony pressure. In addition, the combination of the dual dynamic and strict depth limits yields yet superior performance [7]. A flow diagram of the GP optimization is displayed in Fig. 3.

III. Results

The results of the optimization are presented in Fig. 4 on the following page. As there are seven optimization runs for each of the seven tissue types (bladder, cancellous bone, cortical bone, fat, muscle, prostate, and rectum), to present plots for all the runs such that the displays would be large enough to be intelligible would occupy far too much space, so the results for fat are shown as representative of the results for all tissue types. For the fat GP symbolic regression, there are

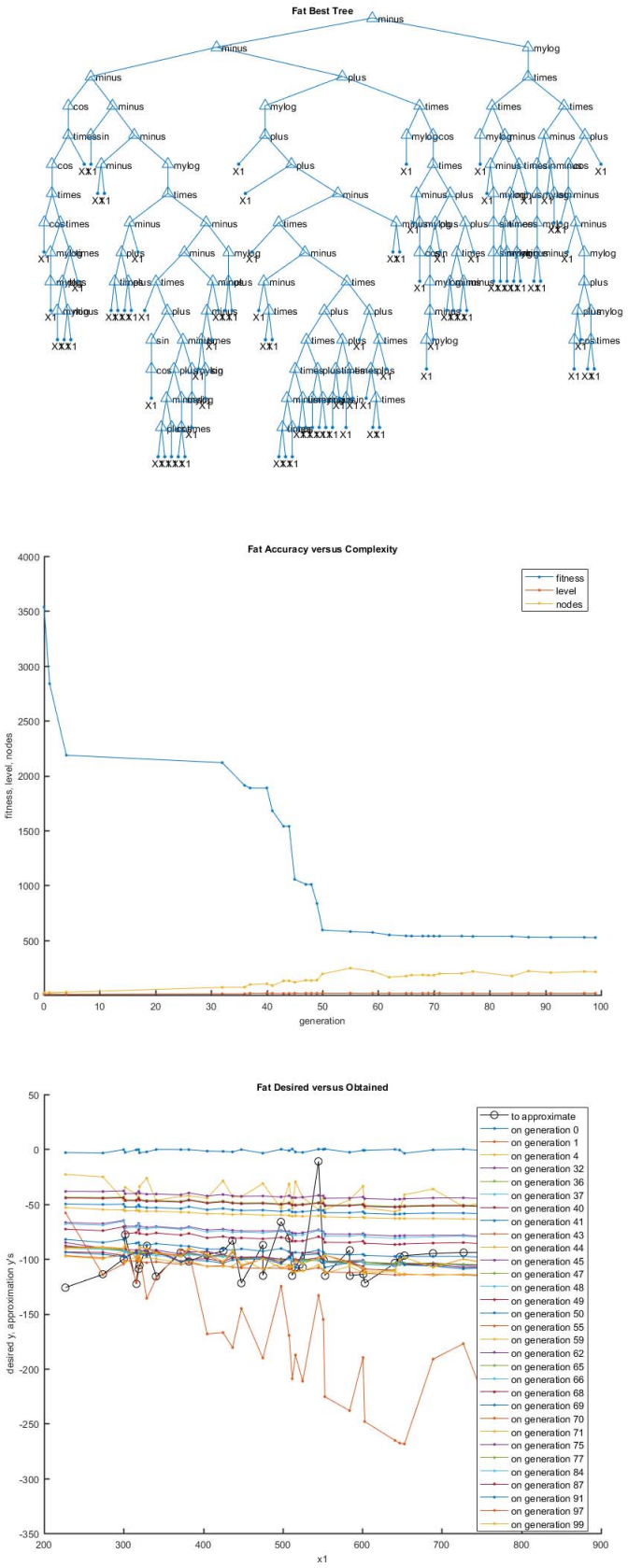

Fig. 4 Results of the optimization for fat tissue type. Top: Best Tree. Middle: Accuracy versus Complexity. Bottom: Desired Versus Obtained HU Values.

plots of the best tree, the accuracy versus complexity, and the desired versus obtained HU values. The accuracy versus complexity display plots the best fitness, number of nodes, and tree depth as a function of generation number. Note that in the best tree plot, the function $mylog$ is simply the natural logarithm of absolute value of the MR pixel intensity, and the terminal $X1$ represents the MR pixel intensity.

A. General Remarks Regarding GP Optimization Results

The best trees for all tissue types are somewhat bloated, but calculating the synCT data using them is trivial, and the quality of the results, as will be shown, is excellent. Considering the middle plot in Fig. 4, it is observed that for a modest increase in complexity, there is a payoff in more fit solutions, so the increase in tree depth and number of nodes is certainly tolerable. Finally, considering the bottom plot in Fig. 4, the synCT HU values calculated from the MR pixel intensities show good agreement with the actual CT data.

GP optimization yielded similar results for all tissue types. There is an observable marked improvement in fitness earlier in the run, with less improvement as the optimization proceeds to later generations. The number of nodes and depth of trees tended to increase throughout the run, but the bloat control of the dual fixed and dynamic depth limits succeeded in preventing the bloat from becoming overwhelming. The poorest performance observed was for the rectal tissue type. The fitness did not really improve that much throughout the run, and the trees tended to increase in complexity without an increase in the accuracy of the calculated synCT values. This may be attributable to less-than-ideal parameter settings, HU values that contained noise, or both. An accuracy versus complexity plot for the rectal tissue type is displayed in Fig. 5.

B. Calculated SynCT Data

For each tissue type, five points were randomly selected from each of five patient MR-CT image pairs, for a total of twenty-five points per tissue type. The GP-optimized map was then applied to calculate the synCT HU values. The synthetic HU data was then compared with the actual CT HU values at each point. Error histograms for all tissue types are shown in Fig. 6.

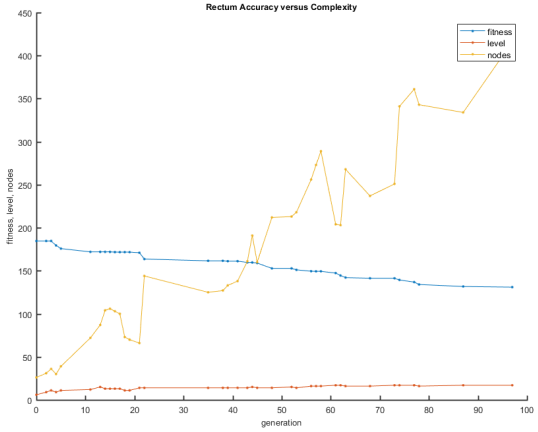

Fig. 5 Accuracy Versus Complexity for the rectum tissue type.

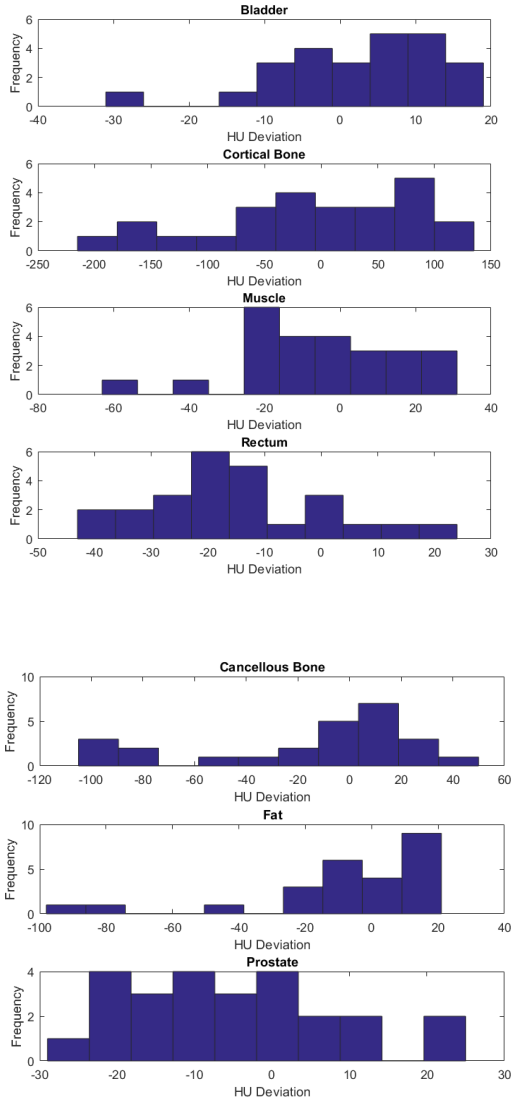

Fig. 6 HU Deviation Histograms for all tissue types.

The poorest performances of the GP-optimized maps were associated with the cancellous bone and cortical bone tissue types. This is unsurprising, as the MR-to-CT registration performed between the image data sets was fine-tuned by the radiation oncologist to ensure the most accurate registration of the soft tissues (prostate, bladder, rectum, muscle, and fat). Consequently, the match of the bony anatomy can be less than ideal, leading to erroneous MR-to-synCT HU values.

C. Comparison Of GP Results With Other Algorithms

Arabi et al. published a helpful survey of current synCT generation algorithms [8]. The approaches detailed in the paper include four atlas-based techniques: median value of atlas

images (ALMedian), atlas-based local weighted voting (ALWV), bone enhanced atlas-based local weighted voting (ALWV-Bone), iterative atlas-based local weighted voting (ALWV-Iter). The survey also included a machine learning technique using a deep convolution neural network (DCNN). The performance of the GP approach can be compared directly for some tissue types, and indirectly for others, by comparing the mean absolute errors (MAE) and standard deviations (STD) of HU values. Where no comparable results exist, 'N/A' is displayed for the tissue type. The comparison is shown in Table I below:

TABLE I. MAE AND STD OF SYNCT HU VALUES (MAE ± STD)

	GP	ALMedian	ALWV	ALWV-Bone	ALWV-Iter	DCNN
Bladder	9.3 ± 6.6	30.0 ± 17.6	24.1 ± 13.6	26.4 ± 12.7	25.2 ± 10.1	18.4 ± 6.6
Rectum	18.4 ± 11.5	93.5 ±71.2	88.1 ± 60.8	100.0 ± 62.0	114.8 ± 63.6	78.3 ± 69.2
Cortical Bone	75.3 ± 52.7	161.1 ± 30.0	134.2 ± 24.0	163.8 ± 25.0	130.2 ± 23.4	119.9 ± 22.6
Canc. Bone	33.2 ± 33.2	N/A	N/A	N/A	N/A	N/A
Muscle	17.1 ± 13.7	N/A	N/A	N/A	N/A	N/A
Fat	18.6 ± 23.7	N/A	N/A	N/A	N/A	N/A
Prostate	11.8 ± 8.6	N/A	N/A	N/A	N/A	N/A

It is clear from the comparison that the GP approach shows promise versus all the other methods. This is not unexpected, since all the other approaches are more automated; indeed, in all the other cases, the segmentation is performed automatically, i.e. there is no human radiation oncologist contouring the organs on the images [8]. Clearly, this expert knowledge enhances the GP approach of the authors, yielding more accurate synCT HU values. In the authors' experience, many clinics still employ this manual segmentation procedure, so the addition of the GP-optimized tissue-specific maps would be a natural fit if it were to be adopted. Further, the statistics quoted above for the algorithms other than GP are for larger ensembles of data points, so that some of the effect seen above could be caused by sample variance (though an attempt was made to minimize any such effect by randomly sampling the data points in each organ).

IV. CONCLUSION

In this study, genetic programming was used to optimize tissue-specific MR-to-HU maps for the generation of synthetic CT data; tissue types were: bladder, cancellous bone, cortical bone, fat, muscle, prostate, and rectum. MR-CT image studies were registered using automated fusion, and then fine-tuned. A GP algorithm was employed to optimize the MR-to-HU mappings. The optimized maps were then used to calculate synCT HU values at randomly selected points. The synCT data were then compared with the actual CT HU values to assess the performance of the algorithm.

The GP method outperformed other atlas-based methods, as well as a deep convolution neural network, when the mean absolute errors and standard deviations of the synCT values were compared. This was not unexpected, since the problem of generating a synCT involves both classification of the tissues as well as predicting the HU values from the MR pixel intensities. The GP method relies on expert human knowledge, so this obviates the need for an automated solution to the classification problem, which is nontrivial.

The GP method shows promise but needs further study. Larger sample sizes of points need to be tested to ensure that the results are not subject to a sample variance effect. An automated segmentation algorithm could be applied to classify the organs, so that the process does not involve human interaction. Finally, radiation dose distributions should be calculated on the full synCT studies to assess the accuracy of the dose to the tissues.

REFERENCES

[1] A. Wood et al., " Spectral characterization of tissues in high spectral and spatial resolution MR images: Implications for a classification - based synthetic CT algorithm," *Med. Phys*., 44: 1865-1875, 2017.

[2] X. Han, "MR - based synthetic CT generation using a deep convolutional neural network method," *Med. Phys*., 44: 1408-1419, 2017.

[3] H. Emami et al., "Generating synthetic CTs from magnetic resonance images using generative adversarial networks," *Med. Phys*., 45: 3627-3636, 2018.

[4] R. Farjam, "Multiatlas approach with local registration goodness weighting for MRI - based electron density mapping of head and neck anatomy," *Med. Phys*., 44: 3706-3717, 2017.

[5] C. Siversson, "Technical Note: MRI only prostate radiotherapy planning using the statistical decomposition algorithm," *Med. Phys*., 42, 2015.

[6] A. E. Eiben and J.E. Smith, Introduction to Evolutionary Computing, 2nd ed., Springer-Verlag: Berlin: 112-113, 2015.

[7] S. Silva, "GPLAB: A Genetic Programming Toolbox for MATLAB," Toolbox documentation, April 2007, retrieved from http://mech.fsv.cvut.cz/~leps/teaching/mmo/data/gplab.manual.3.pdf, 2018.

[8] H. Arabi, H., "Comparative study of algorithms for synthetic CT generation from MRI: Consequences for MRI - guided radiation planning in the pelvic region," *Med. Phys.*, 45: 5218-5233, 2018.