

Estimation of Contrast Agent Concentration from Pulsed-Mode Projections to Time Contrast-Enhanced CT Scans

Isabelle M. Heukensfeldt Jansen^{*a}, Eri Haneda^a, Bernhard Claus^a,
Jed Pack^a, Albert Hsiao^b, Elliot McVeigh^c, and Bruno De Man^a

^aGE Research, One Research Circle, Niskayuna, NY, USA 12309; ^bDept. of Radiology at University of California San Diego, 9500 Gilman Dr, La Jolla, CA, USA 92093; ^cDept. of Bioengineering, Medicine, Radiology at University of California San Diego, 9500 Gilman Dr, La Jolla, CA, USA 92093

ABSTRACT

Cardiac CT exams are some of the most complex CT exams due to the need to carefully time the scan to capture the heart during a quiescent cardiac phase and when the intravenous contrast bolus is at its peak concentration in the left and/or right heart. We are interested in developing a robust and autonomous cardiac CT exam, using deep learning approaches to extract contrast and cardiac phase timing directly from projections. In this paper, we present a new approach to estimate contrast bolus timing directly from a sparse set of CT projections.

We present a deep learning approach to estimate contrast agent concentration in left and right sides of the heart directly from a set of projections. We use a virtual imaging framework to generate training and test data, derived from real patient datasets. We finally combine this with a simple analytical approach to decide on the start of the cardiac CT exam.

Keywords: bolus tracking, deep learning, enhancement prediction, scan timing

1. INTRODUCTION

Cardiac CT exams such as Coronary CT Angiography (CCTA) are some of the most complex CT exams due to the need to carefully time the scan to capture the heart during the quiescent cardiac phase (when the heart is relatively still) and when the contrast bolus in the heart chambers is at its peak concentration to achieve good contrast enhancement. The overall exam duration and the complexity of performing these exams (combined with limited reimbursement levels) have limited patient access to cardiac CT to academic hospitals and specialized cardiac imaging centers. Timing the CT scan to coincide with the peak contrast concentration can be done using a separate ‘timing bolus’ acquisition or with ‘bolus tracking’. Both approaches have pros and cons and require highly trained operators to achieve consistent bolus enhancement.

Our overall project goal is to develop a smart cardiac CT scanner that autonomously determines the optimal scan time interval without ECG, traditional bolus tracking or timing bolus, using real-time deep learning analytics of sparsely pulsed projections^{1,2}. Recent advances in X-ray tube technology have provided the ability to acquire pulsed-mode projections (PMPs) or only a few projections per gantry rotation. While not sufficient to perform image reconstruction, these PMPs can be utilized along with deep learning to predict the contrast agent concentration in specific compartments of the heart. Here we present results of a deep learning approach to estimate contrast agent concentration in left and right sides of the heart and to determine the start of the CCTA scan.

2. BACKGROUND

2.1 Standard-of-care CCTA Protocol

In CCTA exams, an intravenous (IV) power injector and an ECG monitor are connected to the patient. A scout scan is performed for patient positioning, followed by a low-dose CT scan to determine a region-of-interest for tracking the contrast bolus. In a first standard-of-care protocol, a timing bolus (or test bolus) is administered as a ‘trial run’³, involving a separate injection of a small amount of contrast agent (10–20 ml) followed by a saline flush. After injection,

*isabelle.jansen@ge.com

a predetermined number of low-dose bolus timing scans are performed with narrow collimation and with a relatively long inter-scan delay (to minimize radiation dose). The enhancement is measured in a region of interest (ROI) to determine time-to-peak enhancement, which is used to compute the delay time between the administration of the main bolus and the start of diagnostic CCTA scan with wide collimation.

A second standard-of-care protocol uses real-time bolus tracking: the timing bolus is omitted, and the time-of-peak-enhancement is predicted in real-time. When a pre-defined threshold is reached, a diagnostic delay (e.g., 7-sec) is added to give a breath-hold command, open up the collimation, and re-position the patient, before starting the diagnostic CCTA scan³. The time-to-peak contrast enhancement depends on many factors, including target vessel, patient anatomy and cardiac output. The enhancement is measured in an ROI, and because of the necessary diagnostic delay, the threshold (usually 100–150 HU) is far below the peak enhancement (250–600 HU)¹⁻³. As a result, a longer bolus (and hence larger contrast volume) is used for robust scanning, i.e.: to avoid missing the peak.

2.2 Proposed CCTA protocol

Our proposed CCTA protocol uses IV contrast administration and a scout scan for patient positioning but eliminates the ECG as well as the additional CT scans to track the bolus concentration in ROIs. The patient and collimator are immediately positioned in the right location for the actual CCTA scan. After injection of the main bolus, a X-ray tube is pulsed ON and OFF, using a very low duty cycle for low radiation dose. For example, 1, 2, 4, 8, or 16 300- μ s pulses may be performed per 0.28 sec rotation, resulting as many CT projections as there were pulses. Deep learning networks analyze these projections in real-time and determine a running estimate of the current contrast agent concentration as well the current cardiac phase (% R-to-R-peak). A separate timing algorithm uses these estimates to decide on the time for the breath-hold command and the start of the actual CCTA scan. This approach promises to eliminate or greatly reduce the diagnostic delay since the patient and collimator are already in the right position, so the peak enhancement may robustly be identified even for a smaller bolus and a narrower plateau. The PMPs minimize the additional radiation dose since no full-rotation scans are performed. The approach also leads to robust automation relying on deep learning algorithms to time the CCTA exam.

3. MATERIALS AND METHODS

We briefly summarize the virtual imaging framework used to generate training data (section 3.1), we then describe the convolutional neural network (CNN) to predict the contrast level present in a PMP (section 3.2), and finally we derive a simple scan timing algorithm (section 3.3).

3.1 Training data generation

We created a virtual imaging framework for creating cardiac CT projections at any combinations of view angles, cardiac phase, and bolus contrast timing as presented in detail in [4]. In summary, we did this by developing five-dimensional cardiac CT models from multi-phase clinical cardiac (cine) CT scans by segmenting the heart compartments and identifying a blood flow propagation map in each compartment. To model contrast dynamics at multiple bolus time points from datasets that were acquired at (approximately) a single bolus time point, we segmented the cardiac compartments, we parametrized the voxels inside those compartments based on their location along the flow direction, and then incremented the voxel values to model different bolus distributions based on location and time point. We then defined multiple instantiations of CT exams based on specific timing of cardiac cycle, contrast bolus, and CT scan and generated virtual CT projection data. The model contained segmentations for the right atrium (RA), right ventricle (RV), pulmonary artery (PA), left atrium (LA), left ventricle (LV), ascending aorta (AA), and descending aorta (DA). Figure 1 shows a specific CT exam instance as a function of time. The top row shows the CT gantry rotation angle. The second row shows the patient ECG signal. The 7 colored curves show CT number averaged over each of the 7 cardiac compartments. For each curve, we can clearly observe a rising edge, a plateau, and a more gradual decay. The large delay between compartments in the right left sides of the heart is due to the pulmonary circulation. We did not simulate dispersion for the left side of heart in this experiment. The bottom row shows the time of the injection, the start of the pulsed-mode projections (PMPs), the breath-hold command, and the actual CCTA scan.

In this work, a number of sample instantiations were defined by selecting and combining different patients, heart rates and contrast dynamics. For each sample, a series of virtual CT projection data was generated along two view angles, anterior-posterior (AP) and posterior-anterior (PA). For each series, a baseline reference image was defined as the

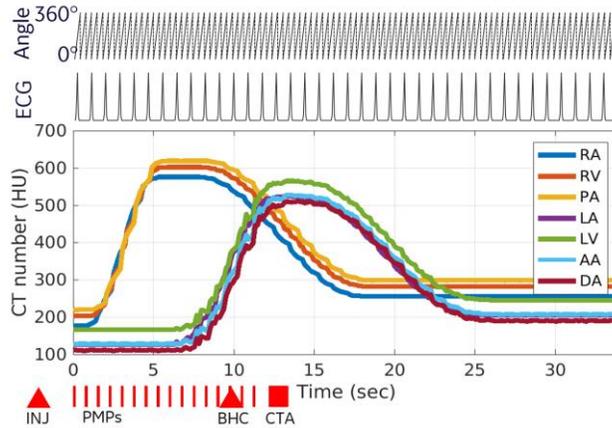


Figure 1. Example of a CT exam instance as a function of time. From top to bottom: gantry rotation angle, ECG signal, average CT number in the 7 cardiac compartments, and the timing of injection (INJ), pulsed-mode projections, (PMPs), breath-hold command (BHC), and CCTA scan. The gantry rotation angle, ECG signal, and PMPs are shown for illustration purpose. The spacing is not exact.

average of all the PMPs from the matching view angle that had been estimated as not yet having any enhanced contrast from the bolus. Network input images were then created from grouping 5 PMPs (representing the 5 “latest” PMPs in a real-time scenario) together in a sliding window and stacking the PMPs with the baseline image matching the view angle of the last PMP in the series for a 128x128x6 input image. A sample input comprised of a reference image and 2 PMPs is shown in Figure 2.

Training, validation, and testing datasets were created to be independent of each other by separating the simulated dataset such that all the images from an individual cardiac model were assigned to only one of the sets. The training, validation, and testing sets comprised of 32, 4, and 4 models respectively, for a total of 300 simulated series of 250 PMPs each, or 61,254 possible samples in the training set when each series was broken down into sets of 5 PMPs to generate input images. Each sample was labeled with two volume-weighted averages, one of the contrast in the RA-RV chambers (right half) and one of the contrast in the LA-LV chambers (left half).

3.2 Neural Network Architecture and Training

The network comprised of a CNN followed by fully-connected layers to perform regression on the input images. In the CNN, pairs of convolutional layers were followed by batch normalization, 2x2 max pooling, and a relu activation function. This was repeated three times, using first a pair of convolutional layers with 16 features each, then 32 and 64 features each for the second and third repetition. The output of the CNN was then flattened and fed through fully connected layers followed by a relu activation function. Fig. 3 shows the network architecture. The network output was a pair of values representing the contrast level in the right and left halves. The network was trained using a MSE loss function and ran for 500 epochs using an Adam optimizer. Each epoch consisted of 1,000 random samples from the training set images.

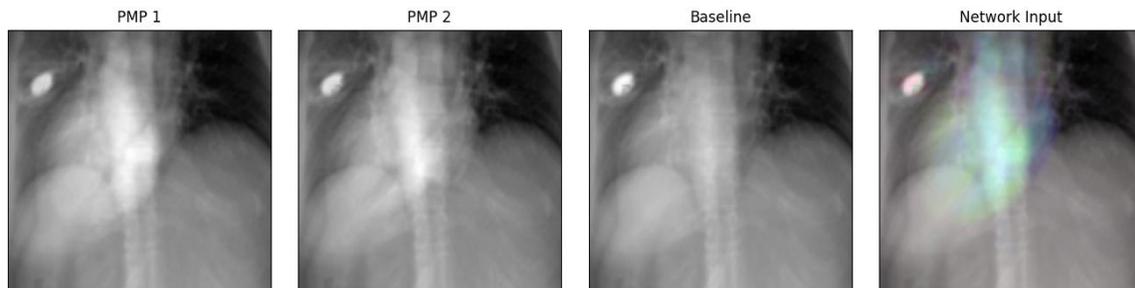


Figure 2. A sample input image using 2 PMPs. In this configuration, the presence of a reference baseline created with pre-bolus images means the stacked layers show the presence of the bolus as color on a greyscale image.

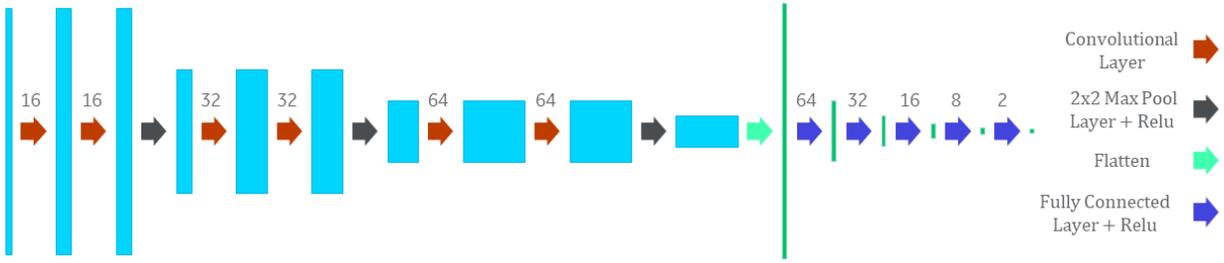


Figure 3. Structure of the neural network. Where relevant, the number of features or nodes are shown.

3.3 CCTA Scan Timing

Each series in the validation and testing set was broken down into a sequence of images representing a sliding window over the entire series. Each image was evaluated by the network to create a time series of the contrast levels for both the right and left chambers. Each time series was fed into an algorithm that looped over the series. For each PMP, all the predictions up to that PMP were evaluated to determine if peak contrast levels had been reached without any further information, to simulate a real-time scenario where peak contrast levels may not be known.

A simple timing algorithm was created to determine a trigger point based on thresholds defined on the predicted enhancement.

4. RESULTS

Network training converged to a R-squared value of 0.99 for the training set and 0.91 for the validation set. Fig. 4 shows the results of network training when the labels are taken as individual datapoints, for the training dataset (left) and validation dataset (right). The broader distribution of the validation dataset compared to that of the training dataset is a classic symptom of overfitting the data. Figure 5 shows representative results of several series in the validation set when the predictions are plotted in a time series for each set. In Fig 5a, the predictions match the ground-truth labels closely. Figures 5b-c show the two most common modes of error. In Fig. 5b, the bolus curve in the left heart is scaled to a lower value than the ground truth labels. In Fig 5c, the right heart bolus curve is confounded with the bolus curve in the left heart and has an additional bump in the distribution when the left-heart contrast is high. While the scaling error in Fig 5b was observed for both right heart and left heart estimates, the error shown in Fig 5c was exclusively seen as the right-heart predictions gaining a secondary bump when the left heart contrast is bright. In both cases, the overall shape of the bolus rise time is preserved and is hence suitable for timing optimization algorithms.

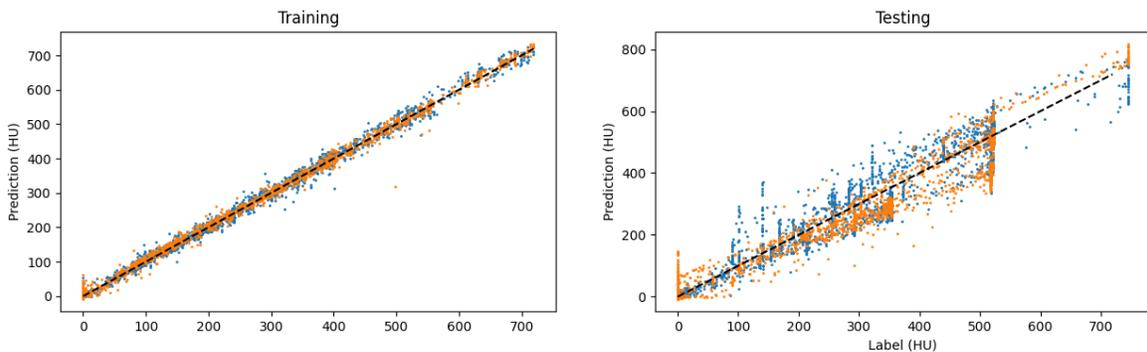


Figure 4. Prediction vs Labeled contrast levels in the training and validation sets. For each plot, orange and blue represent the right and left chambers respectively.

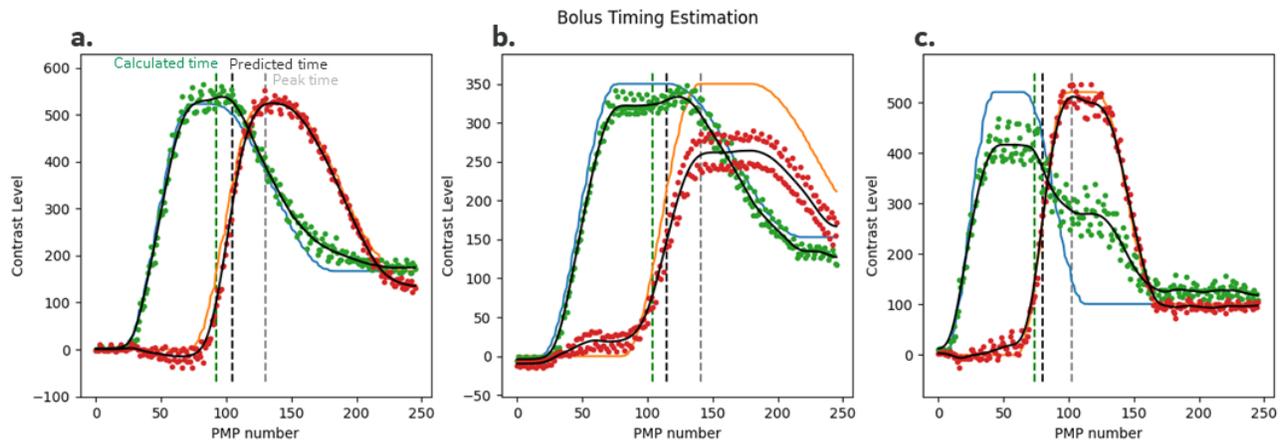


Figure 5. Sample validation outputs of the network. Blue represents the right chamber enhancement with corresponding green predictions, while orange represents the left chamber enhancement with corresponding red predictions. Predictions are smoothed using a gaussian filter (resulting in black curves). The green and black dashed lines represent the desired and the trigger points based on a fixed threshold, while the grey dashed line shows the start of the peak enhancement in the left side of the heart. a.) The prediction closely matches the ground-truth; b.) the prediction in one chamber is off by a scaling factor; c.) the prediction in the right chambers is confounded with the left chambers, resulting in an addition bump after the bolus in the right chambers has reached the converged bolus level.

The green and black dashed lines represent the desired and the trigger points based on a fixed threshold, while the grey dashed line shows the start of the peak enhancement in the left side of the heart. The RMS error was 11 PMPs in the testing set when compared to the same threshold applied to the ground-truth data, which corresponds with an RMS error of 1.6 seconds. More work is needed to optimize the trigger points and derive the time for the breath-hold command and for the start of the CCTA scan.

5. CONCLUSION

A neural network provides a powerful tool for analyzing the sparse projection data acquired with PMPs. For the purpose of scan timing, the rise and decay times are more critical than the precise levels of the contrast enhancement, which will guide us in designing improved networks.

This study used a simple threshold-based timing algorithm, in line with similar algorithms used in more traditional bolus tracking. A more sophisticated timing algorithm might derive trigger points for the breath hold command approximately 3-4 seconds in advance, then continuing to monitor the bolus curve to begin the scan once the actual peak is reached. By monitoring PMPs instead of an ROI, the equipment delays created by patient repositioning and opening the collimation is negated. This work shows promising results for the use of PMPs in optimizing timing for CCTA scans.

ACKNOWLEDGEMENT

Research reported in this publication was supported by the NIH/NHLBI grant R01HL153250. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

REFERENCES

- [1] R. L. Hallett en D. Fleischmann, “Tools of the trade for CTA: MDCT scanners and contrast medium injection protocols”, *Techniques in vascular and interventional radiology*, vol 9, no 4, bl 134–142, 2006.
- [2] S. Oda et al., “Low contrast and radiation dose coronary CT angiography using a 320-row system and a refined contrast injection and timing method”, *Journal of cardiovascular computed tomography*, vol 9, no 1, bl 19–27, 2015.
- [3] J.-E. Scholtz en B. Ghoshhajra, “Advances in cardiac CT contrast injection and acquisition protocols”, *Cardiovascular diagnosis and therapy*, vol 7, no 5, bl 439, 2017.
- [4] E. Haneda et al., “A five-dimensional cardiac CT model for generating virtual CT projections for user-defined bolus dynamics and ECG profiles”, in *The 7th International Conference on Image Formation in X-Ray Computed Tomography, Baltimore, MD, USA, June 12-16, 2022*.