Breast Cancer Patient Auto-Setup Using Residual Neural Network for CT-Guided Therapy

CHENNING LIU1, CHENAN LIU2, FENGQUAN LV3, KAIBANG ZHONG4, AND HONGJIN YU1
1Department of Gynaecology and Obstetrics, Affiliated Dongguan Maternal and Child Health Care Hospital, Southern Medical University, Dongguan 523000, China
2Shanxi Medical University, Taiyuan 030002, China
3Department of Radiation Oncology, Guangdong Provincial People’s Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, China
4Department of Otorhinolaryngology, Tungwah Hospital of Sun Yat-sen University, Dongguan 523000, China
Corresponding author: Hongjin Yu (yu000811@163.com)

ABSTRACT Patient setup will influence the treatment of the breast cancer in radiation therapy. Improving the accuracy of the tumor target localization is vital for the cancer treatment. In this study, we focus on the breast patient setup and develop an accurate tumor localization method based on the deep learning in radiation therapy. The proposed method used a double residual neural network model to achieve the high precision and efficiency patient tumor localization. In the network training, the model attempt to localize the breast and then detect the landmarks inside the localized region. After the model training, we used an iterative filter scheme for calculating a transformation to the daily CT. Therefore, the gray value distribution can match well with the training image. The final landmark positions were obtained after the iteration. The translation errors in the daily CT were determined using the detected landmarks. We used the digital CT phantom images and the real patient CT images to evaluate the proposed method. Then result of the breast patient setup was shown to be clinically acceptable. The mean and standard deviation setup errors were 0.64 ± 1.40 mm, 0.15 ± 1.28 mm, and −0.46 ± 1.17 mm in the anterior-posterior, left-right, and superior-inferior, respectively. In conclusion, we proposed an accurate patient setup method, which shown a very promising alternative for marker-free breast auto-setup.

INDEX TERMS Breast setup, daily computed tomography, residual network, image-guided radiation therapy.

I. INTRODUCTION

Many researchers studied the accurate tumor localization in radiation therapy. Among all tumor localization methods, X-ray imaging is the most widely used image for localizing the tumor. The X-ray guided prostate localization methods can be divided into two major folds: two-dimensional (2D) X-ray projection and 3D CBCT image. The limitation of the 2D X-ray is that the correlation of the bony pelvis anatomy and the position of the breast tumor is weak because independent inter-fractional motion exists between the bony anatomy and the breast tumor [9]. The limitation of the 3 CBCT image is that (a) the inherently low image contrast of the breast tumor and (b) substantial daily breast tumor position changes and some breast tumor deformations [10]. Therefore, the required further manual intervention for correcting the registration errors is thus becoming a barrier to breast tumor localization.

In this paper, we try to find an accurate translation parameter for the patient setup to achieve the breast cancer...
targeting. Different from the registration approaches based on the image intensity, in this paper, we establish a patient-specific deep learning model to achieve the breast auto-setup. The setup errors between the planning CT (pCT) and the daily CT is obtained by aligning the corresponding breast landmarks which are outputted from the deep learning model [8], [11]–[13]. It can address the drawbacks in the traditional methods through the following contributions: 1) Instead of requiring a large patient dataset, the training data in the article is generated using a patient-specific data augmentation strategy, which can account for the changes of the image content between the daily CT and the pCT. We develop a new architecture of network named double residual network aiming at the landmark detection on the breast image. Specifically, the double residual network consists of two deep residual convolutional neural networks (Resnet), with each focusing on one specific task. In the first step, we propose a Resnet based regression model, aiming to localize breast. To make precise landmarks detection, in the second step, we enlarge the breast image in the located image, which is outputted from the first step. The enlarged breast image is then inputted into another Resnet model. Importantly, with the advantage of deep feature extraction using the Resnet, the proposed network can detect the landmarks accurately, even if there are no fiducial markers (FMs) implantation. Besides, the double residual network can simultaneously detect large-scale landmarks in real time because of the parallel computation. 2) We proposed an iterative prediction strategy to the daily CT image in predicting (testing) the landmarks’ position, which can make the most of the constraint of the relative distance between each landmark. The iterative prediction allows us to incorporate prior knowledge into the landmark prediction more effectively [14]–[16]. In this sense, the breast setup is no longer a conventional image registration problem. Rather, it becomes a breast detection/localization problem. This approach gives different perspectives to the patient auto-setup in the daily CT guided radiotherapy and sheds some new insight into the organ localization [17]–[22].

In Summary, the major contributions of our work were in three folds:

1) We proposed a patient-specific model for tumor localization without requiring a large patient dataset. It can mitigate the human labor in the data annotation.

2) A new architecture of network using residual network was proposed. With the advantage of deep feature extraction using the Resnet, the proposed network can detect the landmarks accurately, even if there are no FMs implantation.

3) An iterative prediction strategy was proposed to the daily CT image in predicting (testing) the landmarks’ position, which can make the most of the constraint of the relative distance between each landmark.

We organize this paper as follows. In section 2, we provide a patient-specific pCT augmentation model to generate the training data and describe a double residual network suitable for the breast auto-setup. In section 3, we demonstrate the technical feasibility of the proposed framework by testing 270 pCT cases and 80 clinical breast daily CT cases. Section 4 summarizes and discusses the results. Section 5 concludes the proposed method.

II. METHODS

A. TRAINING DATA GENERATION

In this section, we attempt to deal with the challenging problem of the limited training data in detecting landmarks with breast CT image. As a typical procedure of the breast external beam radiation therapy, the breast contours were first manually outlined on the pCT image. Therefore, the landmarks can be selected from the on the breast contours as the label data. There is only one pair of training data if we do not add other patients’ pCT image. Therefore, we enlarge the dataset through the use of image augmentation technique, which is usually implemented in deep neural network [23]. In this study, a patient-specific image augmentation model for pCT was proposed to generate a large amount of training data without relying on the other patients’ image data. For taking into account the different positions of the patients in the treatment setup, the proposed augmentation model first applies the random geometric transformation to the pCT image. Different from the conventional augmentation technique, to improve the testing accuracy of landmarks detection on the daily CT image, we develop a random image transformation processing accounting for the image modality and content changes. The random image transformation does not require any parameter learning, which can be easily implemented to the other convolutional neural network (CNN) regression tasks. The details of the random geometric and image transformation are explained as follows.

1) RANDOM GEOMETRIC TRANSFORMATIONS

The most popular current practice for data augmentation is to perform the image geometric transformations. In this paper, the methods of the geometric transformations include rotation, translation, and rescaling. The rotation angles are randomly chosen from $-9^\circ$ to $9^\circ$ in the yaw, pitch, and the roll direction. The translation distances are randomly generated from $-60$ mm to $60$ mm in the in the Anterior-Posterior (AP), the Left–Right (LR), and the Superior–Inferior (SI) direction. The scaling factor are randomly from 0.8 to 1.2. The selected parameters here are reflected the possible geometric changes in the patient breast setup.

2) RANDOM IMAGE TRANSFORMATIONS

The image transformations include the random Gaussian filtering and random erasing. The image transformations are conducted with a certain probability in training. For an image $I$ in a mini-batch, the probability of it undergoing random image transformation are $p$, and the probability of it being kept unchanged is $1-p$. Gaussian filtering selects a standard deviation $\sigma$ of the Gaussian distribution randomly in the range specified by the minimum $\sigma_l$ and the maximum $\sigma_u$. 

\begin{align*}
\sigma &= \frac{\sigma_u - \sigma_l}{2} \times \text{rand}() + \sigma_l
\end{align*}
After that, the random erasing is applied to the image[24]. A cube region \( I_r \) is randomly selected to be erased. The voxel values in the erased areas are set to zeros. Suppose that the voxel size of the training image \( V = l \times w \times h \). We randomly initialize the edge length of erasing cube region to \( r \), which is in range by minimum \( r_l \) and maximum \( r_u \). Then, we randomly initialize the point \( P = (x_r, y_r, z_r) \) in \( I \). We set the region \( I_r = I (x_r, y_r, z_r, x_r + r, y_r + r, z_r + r) \) as the selected rectangle region if \( I_r \) do not cover the breast \( (I_r \cap C = \emptyset) \). Otherwise repeat the above process until an appropriate \( I_r \) is selected.

The parameter setting and the procedure of the random image transformation is summarized in Table 1.

| Algorithm 1: random image transformation procedure |
|---------------------------------------------------|
| **Input:** Input 3D image: \( I \); The breast contour: \( C \); 3D Image size: l=512, w=512, and h=220; Filtering and Erasing probability: \( p_1 = 0.5 \), \( p_2 = 0.5 \); Filtering standard deviation range: \( \sigma_l = 0 \), \( \sigma_u = 5 \); Erasing Cube edge length range: \( r_l = 5 \), \( r_u = 40 \); |
| **Output:** Transformation image \( I_\delta \) |
| 1. \( p_\tau = \text{rand}(0,1) \); \( p_\sigma = \text{rand}(0,1) \); |
| 2. if \( p_\tau \leq p_\tau \) then |
| 3. \( \sigma = \text{rand}(\sigma_l, \sigma_u) \); \( I_1 = G(\sigma) \ast I \); |
| else |
| 4. \( I_1 = I \); |
| 5. end |
| 6. end |
| 7. if \( p_\sigma \leq p_\sigma \) then |
| 8. while true do |
| 9. \( r = \text{rand}(r_l, r_u) \); \( x_r = \text{rand}(0, l - r) \); |
| 10. \( y_r = \text{rand}(0, w - r) \); \( z_r = \text{rand}(0, h - r) \); |
| 11. if \( I_1 \cap C = \emptyset \) then |
| 12. \( I_r = I_r \ast (x_r, y_r, z_r, x_r + r, y_r + r, z_r + r) \); |
| 13. \( I_1 \cap C = \emptyset \); \( I_\delta = I_\delta ; \text{return} \ I_\delta \); |
| 14. end |
| 15. end |
| 16. else |
| 17. \( \text{return} \ I_\delta = I_\delta \); |
| 18. end |

### B. THE ARCHITECTURE OF THE DOUBLE RESIDUAL NETWORK

In this section, we focus on detecting the landmarks accurately on the breast CT image using the training data generated from the above section. As demonstrated in Fig. 1, to detect the landmarks on the breast image, we develop a double residual network framework. The task of the first-step residual network regression model is to find the 3D coordinate of the breast center and localize the breast. The second step is to measure the landmarks’ 3D positions on the breast shown in Fig. 1(b). The sub-sections illustrate the architecture of the double residual network framework in detail.

1) **FIRST STEP: LOCALIZING THE BREAST IN CT IMAGE**

We use the local pCT image, as the input data. The entire pCT is cropped into a region of interest (ROI) in the first step as shown in Fig. 1(b). The breast will be inside the ROI because the initial setup correction is carried out to the patient using the sign on the patient skin or the fixed device before scanning the daily CT. Therefore, the breast will not move faraway in the setup treatment relative to planning. In the training stage, the label of the breast center is calculated by averaging all the 3D coordinates of the selected landmarks on the breast.

As demonstrated in Fig. 1(a), the 1st step of the double residual network is to localize the breast. (b) the second task of the double residual network is to detect the landmarks in the extracted image which is outputted from the first step.
dimensions as seen in the solid arrows, the identity shortcut is used directly. If the dimensions are being increased, the projection shortcut shown in the dotted arrows is used to match dimensions by using the convolutional layer. The detail explanation of shortcut can be found in the Ref[25].

2) SECOND STEP: DETECTING THE LANDMARKS IN THE LOCAL CT IMAGE
Given the 3D coordinates which are outputted from the 1st step, we can extract a 3D image cover the breast shown in Fig. 1(a). The edge length of the extracted 3D image is the same as the planning target volume. The task of the 2nd step is to detect the landmarks in the extracted 3D image. In other words, given the extracted 3D image, the network in the 2nd step aims at formulating a non-linear mapping to output the landmarks’ 3D coordinates on the breast [26]. To improve the accuracy of the landmark detection, we enlarge the voxel matrix of the 3D breast image with a smaller physical voxel size before feeding the image into the non-linear mapping as seen in Fig 1(b). The architecture of the non-linear mapping in the 2nd step can be similar with the 1st step that treats the 3D images as input and the 3D coordinates as output.

3) PARAMETER SELECTION AND NETWORK TRAINING
The double residual network is a regression model. The cropped pCT images are taken as the input data. The input data is first connected with the $7 \times 7 \times 7$ convolutional layer, and then followed by 8 residual blocks. We double the number of filters in every two residual blocks shown in Fig. 1. The filter size is $n \times n \times n$ in the residual blocks. Because the performance may be influenced by the number and the size of the filter, we will carry out the double residual network with different parameters in the di. Training the double residual network is to find a non-linear mapping for detecting the landmarks on the CT image. All the weighting and bias in the network can be estimated by optimizing a loss function. The training in the 1st step can be written as

$$
\Theta_1^* = \arg \min_{\Theta_1} \| \bar{Y}_{p1} - f_1 (\Theta_1, \bar{X}_{p1}) \|_2^2,
$$

where $f_1$ is the architecture of the Resnet with the initial weighting and bias $\Theta_1$. The input training set is $\bar{X}_{p1}$ and the output label set is $\bar{Y}_{p1}$. $\| \bar{Y}_{p1} - f_1 (\Theta_1, \bar{X}_{p1}) \|_2^2$ is the mean square error (MSE). After the training in the 1st step, the 3D coordinates of the breast center $\bar{Y}_{pred1} = f (\Theta_1, \bar{X}_{p1})$ is outputted. The extracted 3D images $\bar{X}_{p2}$ and their labels $\bar{Y}_{p2}$ are obtained using the output $\bar{Y}_{pred1}$ in the 1st step. The training in the 2nd step can be written as

$$
\Theta_2^* = \arg \min_{\Theta_2} \| \bar{Y}_{p2} - f_2 (\Theta_2, \bar{X}_{p2}) \|_2^2,
$$

where $f_2$ is the architecture of the Resnet with the initial weighting and bias $\Theta_2$. The input training set is $\bar{X}_{p2}$ and the output label set is $\bar{Y}_{p2}$. $\| \bar{Y}_{p2} - f_2 (\Theta_2, \bar{X}_{p2}) \|_2^2$ is the MSE. We set the initial learning rate at $10^{-2}$.

C. PREDICTING (TESTING) ON THE DAILY CT IMAGE
Due to the different image modality between the pCT and the daily CT, the results are not accurate if the prediction is directly carried out on the daily CT image. To improve the prediction accuracy, an image transformation is applied to the daily CT image so that its gray value distribution can be as close as the training image. However, the prediction results are sensitive to the parameter of the image transformation. To address this problem, we proposed an iterative prediction strategy to find a preferable parameter for daily CT image transformation. Instead of directly predicting the landmarks on the daily CT image, we incorporate the relationship between each landmark into a constraint, which aims at maintaining the shape of the predicted landmarks on the daily CT image. Because the relative difference of each voxel inside the breast will not distort exceed 3 mm as reported by various publication [27]–[29], the relative distance between each predicted landmark should be not far away from the corresponding distance on the pCT image. With the prior knowledge from the pCT image, the landmarks in the prediction image is no longer independent.

For a given daily CT image, we carry out the method of the image transformation mentioned in Algorithm 1. Instead of setting the parameter randomly, in this section, the image transformation parameter $\bar{t} = \{\sigma, r, x_r, y_r, z_r\}$ is determined by a proposed objective function. To find a preferable $\bar{t}^*$ from the feasible solution that maintaining the shape of the predicted landmarks on the daily CT image, the goal of the objective function is given by

$$
\bar{t}^* = \arg \min_{\bar{t}} \left\| \bar{D}_{pred2} (\bar{t}) - \bar{D}_p \right\|_2^2,
$$

where $\bar{D}_{pred2}$ and $\bar{D}_p$ are the average value of the relative difference between each landmark in the daily CT and the reference pCT image, respectively. $\bar{D}_{pred}$ and $\bar{D}_p$ can be written as

$$
\bar{D}_{pred2} (\bar{t}) = \frac{1}{N} \sum_{j=1}^{N} \left( \sum_{i=1}^{N} \| \bar{y}_{pred2,i} - \tilde{y}_{pred2,i} (\bar{t}) \|_2 \right) / N^2,
$$

$$
\bar{D}_p = \frac{1}{N} \sum_{j=1}^{N} \left( \sum_{i=1}^{N} \| \bar{y}_p - \tilde{y}_p (\bar{t}) \|_2 \right) / N^2,
$$

where $\bar{y}_{pred2,i} (\bar{t})$ is the $i$th predicted daily CT landmark outputted from the double residual network through the use of the image transformation with parameter $\bar{t}$. $\bar{y}_p (\bar{t})$ is the $i$th reference pCT landmark. $N$ is the number of the landmarks. Equation 3 is a non-linear and non-differentiable function because of introducing the double residual network. Existing optimization algorithm, such as Newton’s method and gradient descent, cannot solve this problem. In this work, the objective function is solved using the mesh adaptive direct search algorithm which is a class of derivative-free direct search method. The details of the algorithm have been given in Refs. [30] and [31] and will not be repeated here. With the
initial parameter setting \( \hat{\theta}_0 \), the iterative prediction process is self-updating during each iteration. If the relative difference between each landmark in the daily CT and the reference pCT image \( \| \hat{\mathbf{D}}_{\text{pred}}(\hat{\theta}) - \hat{\mathbf{D}}_p \|_2 \geq 3 \), the iteration process continues using Eq. 3. Otherwise, the preferable \( \hat{\theta}^* \) is returned. If the maximum iteration number is reached, the process will also stop.

D. EVALUATION OF THE ALGORITHM PERFORMANCE

1) DATASET

In this paper, the daily CT images are acquired at different times during the setup treatment of the patient. To make a quantitative evaluation on the proposed method without relying on the FMs, the reference translational parameters are determined by the experienced physician. To train and verify the proposed marker-free breast auto-step, all the CT images are divided into three sets: the training set, the validation set, and the testing set. In the training stage, the number of the landmarks is 10. The training and the validation sets are a group of 900 pCT images and 100 pCT images, respectively. These data are augmented from the patient-specific pCT image using the proposed method mentioned in section 2. The testing set has two parts: the testing digital CT set and the testing daily CT set, which is explained as follows.

2) EVALUATION OF THE TESTING DIGITAL CT SET

The testing digital CT set is a group of 270 cases of the dCT which are acquired using the augmentation model. Because the images and the existing landmarks are transformed with the known transformation matrices, the ground truth of the landmarks on the testing digital CT images are exist. Therefore, the ground truth of the translational parameters can be calculated. We can assess the accuracy of the translational parameters on the testing digital CT set quantitatively by comparing the proposed method with the ground truth.

3) EVALUATION OF THE TESTING DAILY CT SET

The testing daily CT set is a group of 80 cases of the daily CT images, which were augmented from the patient with 20 cases of daily CT acquired at a different time. Different parameters of the translation are applied in the testing daily CT augmentation. However, the evaluation of the translational parameters is often a difficult task on the testing daily CT set because of the lack of the ground truth. To address this problem, we employed an experienced physician to position the patient and acquired the ground-truth of the translational parameters.

III. RESULTS

A. THE TESTING DIGITAL CT STUDIES

The proposed method is first evaluated on the testing digital CT images. The quantitative analysis of the translational parameter errors is shown in figure 2. For the analysis of the translational error, figure 2(a) plots the translation errors between the ground truth and the proposed method in 3D. The frequency histograms, figures 2(b) indicate the ensemble mean and standard deviation (SD) for the distribution of translation errors. Figure 2(c) shows the 2D correlations of the proposed method and the reference. The analysis of translation error from row 1 to row 3 are in the direction of AP, LR, and SI, respectively.

We can make the following observations. First, the ensemble mean and SD for the distribution of the error is 0.06 mm (SD 0.87 mm), -0.27 mm (SD 0.70 mm), and 0.16 mm (SD 0.56 mm) in the direction of AP, LR, and SI, respectively. Second, all the Pearson’s correlation coefficients are very close to one, which implies the relationship between the ground truth and the proposed method is positive linear correlation. Third, we find that most of the translational errors are less than 2 mm. The largest translational errors are less than 4 mm. The results are found to be clinically acceptable.

B. THE TESTING DAILY CT STUDIES

We now illustrate the results of the proposed method on the testing daily CT images. As mentioned in section 2, the testing daily CT set is a group of 80 cases of the daily CT images, which were augmented from the patient with 20 cases of daily CT acquired at a different time.

The proposed method is also evaluated on the testing daily CT images using the similar evaluation method with test digital CT images. The quantitative analysis of the translational parameter errors is shown in figure 3. In figure 3(a), most of the translation errors are less than 2.5 mm. The largest translational errors is less than 4 mm. Although the results in testing daily CT images are larger than that in the testing digital CT image, the results are still found to be clinically acceptable. Figures 3(b) indicates the ensemble mean and standard deviation (SD) for the distribution of translation errors. The ensemble mean and SD for the distribution of the translational error is 0.44 mm (SD 1.64 mm), 0.14 mm (SD 1.34 mm), and -0.33 mm (SD 1.50 mm) in the direction of AP, LR, and SI. Figure 3(c) shows the 2D correlations of the proposed method and the reference and all the Pearson’s correlation coefficients are also very close to one. We also compared our work with the most recent publish works [32], which was < 1 mm in 89.1% of the daily CT. The proposed results were improved to 89.4% of the daily CT.

IV. DISCUSSION

The accuracy of the landmark detection will directly influence the auto-setup. Therefore, we attempt to focus on improve the performance of the double residual network. In this section, we first flesh out the difference between the double residual network and the conventional network, and then analyze the influences of the parameters.

A. IMPACT OF THE NETWORK ARCHITECTURE

There are at least two differences between the proposed double residual network method and the previous regression-based method. First, instead of using the conventional stacked layers, we adopt a regression model using the Res-net in
FIGURE 2. The plot of translational errors between the proposed method and the ground truth on the 270 testing pCT cases. The translation errors of the proposed method are shown in (a). The histograms of the translation errors are shown in (b). The 2D correlations of the translation errors are shown in (c). Row 1 to 3 in (b) and (c) represents the direction in AP, LR, and SI, respectively. ch is outputted from the first step.

FIGURE 3. The plot of translational errors between the proposed method and the ground truth on the 80 testing daily CT cases. The translation errors of the proposed method are shown in (a). The histograms of the translation errors are shown in (b). The 2D correlations of the translation errors are shown in (c). Row 1-3 in (b) and (c) represents the direction in A-P, L-R, and S-I, respectively.

the double residual network, which can alleviate degradation problem. Figure 4(a) shows the changes of the root mean square error (RMSE) concerning different epoch achieved by the proposed double residual network and the conventional stacked layers without residual blocks in the validation set. It is evident that the RMSE achieve an error of 2 mm in
the proposed method, which is better than that without residual block. Second, instead of directly detect the landmarks, the proposed two step network can improve the accuracy. The experimental results are shown in figure 4(b). The RMSE using the proposed method is smaller than the method with only one-step at the end of the epoch.

![Figure 4](image-url)

**FIGURE 4.** Change of the RMSE concerning different epoch on the validation set. (a) Plain network versus residual network; (b) only one-step network versus double residual network.

### B. IMPACT OF THE FILTER NUMBER AND THE FILTER SIZE

The filter number and the size of the double residual network need to be optimized. Although the proposed method may be more difficult to reach a convergence before the 10th epoch, the final RMSE is lower. As compared the performance with different filter number, we find that the RMSE is the lowest when \( k = 4 \). The possible reason is that as the depth of the network increases, the network enables to extract a deeper feature of the image. However, as the filter number increases, the computer memory needs to be enlarged. Therefore, based on existing computer configuration, we set \( k = 4 \). Figure 5(b) shows the influence of the filter size.

![Figure 5](image-url)

**FIGURE 5.** Change of the RMSE concerning different epoch on the validation set. (a) Plain network versus residual network; (b) only one-step network versus double residual network.

### C. IMPACT OF THE ITERATIVE PREDICTION STRATEGY

As mentioned in section 2, we proposed an iterative prediction strategy to find a preferable parameter for daily CT image transformation. To evaluate the impact of this strategy on the testing daily CT image, we add the experiment that predicting the landmarks directly on the daily CT image. The proposed method achieves lower setup errors on the testing daily CT images than the method without iterative prediction.

### D. LIMITATIONS AND FUTURE WORK

First, 3D neural networks are included residual network. Therefore, the training processing is consuming and it takes 21 hours for one case. In the future, we will focus on the efficiency improvement in the training step using down-sampling or other preprocessing technique on the training image. Second, only 20 cases of the testing daily CT image were evaluated in this study. As one of our future work, we will test more cases to further assess the robustness of the algorithm. Of course, the proposed method not only can apply in the breast setup, but also any other organ such as spine, brain, liver and so on. Third, image artifacts such as the scatter, ring, metal artifact will hamper the accuracy of the setup [33]–[44].

### V. CONCLUSION

The marker-free breast setup was formulated as a landmark alignment problem and the concept of the patient-specific image augmentation model was introduced to account for the changes of the image content between the daily CT and the pCT. Instead of requiring a large patient dataset, the approach allows us to train the network with limited data. Besides, this new architecture of network can simultaneously detect the landmarks fast. Furthermore, an iterative prediction and the determination of the parameters in the setup treatment were proposed. In this sense, the breast setup is no longer a conventional image registration problem. Rather, it becomes a breast detection/localization problem. This approach gives different perspectives to the patient auto-setup in the daily CT guided radiotherapy and sheds some new insight into the image-guided radiation therapy. The results suggest that it has strong potential to replace the current image registration process and makes it possible to full-automatic and the high-precise image-guided radiation therapy without the manual intervention.

### REFERENCES

[1] C. Gaisberger, P. Steininger, B. Mitterlechner, S. Huber, H. Weichenberger, F. Sedlmayer, and H. Deutschmann, “Three-dimensional surface scanning for accurate patient positioning and monitoring during breast cancer radiotherapy,” Strahlentherapie und Onkologie, vol. 189, no. 10, pp. 887–893, Oct. 2013.

[2] A. M. Kirby, P. M. Evans, E. M. Donovan, H. M. Convery, J. S. Haviland, and J. R. Yarnold, “Prone versus supine positioning for whole and partial-breast radiotherapy: A comparison of non-target tissue dosimetry,” Radiotherapy Oncol., vol. 96, no. 2, pp. 178–184, Aug. 2010.

[3] M. Laaksomaa, M. Kapanen, M. Haltamo, T. Skytta, S. Peltona, S. Hyödynmaa, and P.-L. Kellokumpu-Lehtinen, “Determination of the optimal matching position for setup images and minimal setup margins in adjuvant radiotherapy of breast and lymph nodes treated in voluntary deep inhalation breath-hold,” Radiat. Oncol., vol. 10, no. 1, p. 76, Dec. 2015.

[4] C. A. Nalder, A. M. Bidmead, C. D. Mubata, D. Tait, and C. Beardmore, “Influence of a vac-fix immobilization device on the accuracy of patient positioning during routine breast radiotherapy,” Brit. J. Radiol., vol. 74, no. 879, p. 249, 2001.

[5] R. Prabhakar, G. K. Rath, P. K. Julka, T. Ganesh, K. P. Haresh, R. C. Joshi, S. Senthamizhchelvan, S. Thulkar, and G. S. Pant, “Simulation of dose to surrounding normal structures in tangential breast radiotherapy due to setup error,” Med. Dosimetry, vol. 33, no. 1, pp. 81–85, Mar. 2008.
[6] M. Riboldi, G. Baroni, R. Orecchia, and A. Pedotti, “Enhanced surface registration techniques for patient positioning control in breast cancer radiotherapy,” *Technol. Cancer Res. Treatment*, vol. 3, no. 1, p. 51, 2004.

[7] Z. F. Mohammed and A. A. Abdulla, “Thresholding-based white blood cells segmentation from microscopic blood images,” *UHD J. Sci. Technol.*, vol. 4, no. 1, pp. 9–17, 2020.

[8] Y. Liang, W. Zhao, Y. Zhang, H. Hirost, M. K. Buyyounouski, S. L. Hancock, H. Bagshaw, Q. Zhang, Y. Xie, and L. Xing, “A deep learning framework for prostate localization in cone beam CT guided radiotherapy,” *Med. Phys.*, vol. 47, no. 9, pp. 4233–4240, Sep. 2020, doi: 10.1007/s00331-019-06

[9] C. Hargrave, T. Deegan, M. Poulsen, T. Bednarz, F. Harden, and K. Mengersen, “A feature alignment score for online cone beam CT based image guided radiotherapy for prostate cancer,” *Med. Phys.*, vol. 45, no. 7, pp. 2899–2911, Jul. 2018.

[10] M. Chao, T. Li, E. Schreibmann, A. Koong, and L. Xing, “Automated contour mapping with a regional deformable model,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 70, no. 2, pp. 599–608, Feb. 2008, doi: 10.1016/j.ijrobp.2007.09.057.

[11] C. Chen, W. Xie, J. Franke, P. A. Gruetzner, L. P. Nolte, and G. Zheng, “Automatic X-ray landmark detection and shape segmentation via data-driven joint estimation of image displacements,” *Med Image Anal.*, vol. 18, no. 3, pp. 487–494, Apr. 2014, doi: 10.1016/j.media.2014.01.002.

[12] B. Ibragimov, R. Korez, B. Likar, F. Pernus, L. Xing, and T. Vrtovec, “Segmentation of pathological structures by landmark-assisted deformable models,” *IEEE Trans. Med. Imag.*, vol. 36, no. 7, pp. 1457–1469, Jul. 2017.

[13] J. Zhang, M. Liu, and D. Shen, “Detecting anatomical landmarks from limited medical imaging data using two-stage task-oriented deep neural networks,” *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4753–4764, Oct. 2017, doi: 10.1109/TIP.2017.2721106.

[14] J.-H. Kim, D. T. Nguyen, J. T. Booth, C.-Y. Huang, T. Fuangrod, P. Poulsen, R. O’Brien, V. Caillet, T. Eade, A. Kneebone, and P. Keall, “The accuracy and precision of kilovoltage intrafraction monitoring (KIM) six degree-of-freedom prostate motion measurements during patient treatments,” *Radiotherapy Oncol.*, vol. 126, no. 2, pp. 236–243, Feb. 2018.

[15] C. Garibaldi, G. Piperno, A. Ferrari, A. Surgo, M. Muto, S. Ronchi, A. Bazani, F. Pansini, M. Cremonesi, B. A. Jereczek-Fossa, and R. Orecchia, “Translation and rotational localization errors in cone-beam CT based image-guided lung stereotactic radiotherapy,” *Phys. Medica*, vol. 32, no. 7, pp. 859–865, Jul. 2016.

[16] X. Liang, N. Li, Z. Zhang, S. Yu, W. Qin, Y. Li, S. Chen, H. Zhang, and Y. Xie, “Shading correction for volumetric CT using deep convolutional neural network and adaptive filter,” *Quant. Imag. Med. Surg.*, vol. 9, no. 7, pp. 1242–1254, Jul. 2019, doi: 10.1002/qims.2019.05.19.

[17] M. Serban, C. Lambert, R. Ruo, C. Maietta, W. Parker, M. Evans, and B. Ibragimov, “A feature alignment score for online cone beam CT based image guided radiotherapy for prostate cancer,” *Med. Phys.*, vol. 45, no. 7, pp. 2912–2922, Jul. 2018.

[18] J. H. Siewerdsen, M. J. Daly, B. Bakhtiar, D. J. Moseley, S. Richard, H. Keller, and D. A. Jaffray, “A simple, direct method for X-ray scatter estimation and correction in digital radiography and cone-beam CT,” *Med. Phys.*, vol. 33, no. 1, pp. 187–197, Dec. 2005.

[19] X. Liu, C. C. Shaw, T. Wang, L. Chen, M. C. Altunbas, and S. C. Kappadath, “An accurate scatter measurement and correction technique for cone beam breast CT imaging using scanning sampled measurement (SSM) technique,” in *Med. Imag. Phys. Med. Imag.*, vol. 16, no. 1, pp. 770–778, Jun. 2016, Art. no. 614234.

[20] J. M. Schallenkamp, M. G. Herman, J. J. Kruse, and T. M. Pisansky, “A simple, direct method for X-ray scatter estimation and correction in digital radiography and cone-beam CT,” *Med. Phys.*, vol. 33, no. 1, pp. 187–197, Dec. 2005.

[21] X. Liu, Y. Jiang, W. Zhao, Z. Zhang, C. Luo, J. Xiong, S. Yu, X. Yang, J. Sun, Q. Zhou, T. Niu, and Y. Xie, “Scatter correction for a clinical cone beam CT system using an optimized stationary beam blocker in a single scan,” *Med. Phys.*, vol. 46, no. 7, pp. 3165–3179, Jul. 2019, doi: 10.1002/mp.13568.

[22] C. Bert, K. G. Metheany, K. Doppke, and G. T. Y. Chen, “A phantom evaluation of a stereo-vision surface imaging system for radiotherapy patient setup,” *Med. Phys.*, vol. 32, no. 9, pp. 2753–2762, Aug. 2005.
CHENNING LIU received the master’s degree from Guangdong Medical University, China. She is currently a Gynaecologist with the Dongguan Maternal and Child Health Care Hospital, China. Her research interests include diseases of uterus and breast.

CHENAN LIU received the bachelor’s degree in medicine from Shanxi Medical University, China, where he is currently pursuing the master’s degree. His major interests include cancer imaging and immune therapy.

FENGQUAN LV received the bachelor’s degree from Southern Medical University, China. He is currently with the Department of Radiation Oncology, Guangdong Provincial People’s Hospital. His research interest includes image-guided radiation therapy.

KAIBANG ZHONG received the master’s degree from Southern Medical University, China. He is currently a Physician with the Tungwah Hospital of Sun Yat-sen University, Dongguan, China. His research interests include medical image diagnostic and processing.

HONGJIN YU received the M.D. degree. She has over 30 years of experience in treatment of gynaecological and breast diseases. She is currently a Chief Physician with the Dongguan Maternal and Child Health Care Hospital.

* * *