J Digit Imaging (2017) 30:751–760
DOI 10.1007/s10278-017-9981-6

A Method to Recognize Anatomical Site and Image Acquisition View in X-ray Images

Xiao Chang[1] · **Thomas Mazur**[1] · **H. Harold Li**[1] · **Deshan Yang**[1]

Published online: 16 June 2017

Abstract A method was developed to recognize anatomical site and image acquisition view automatically in 2D X-ray images that are used in image-guided radiation therapy. The purpose is to enable site and view dependent automation and optimization in the image processing tasks including 2D-2D image registration, 2D image contrast enhancement, and independent treatment site confirmation. The X-ray images for 180 patients of six disease sites (the brain, head-neck, breast, lung, abdomen, and pelvis) were included in this study with 30 patients each site and two images of orthogonal views each patient. A hierarchical multiclass recognition model was developed to recognize general site first and then specific site. Each node of the hierarchical model recognized the images using a feature extraction step based on principal component analysis followed by a binary classification step based on support vector machine. Given two images in known orthogonal views, the site recognition model achieved a 99% average *F*1 score across the six sites. If the views were unknown in the images, the average *F*1 score was 97%. If only one image was taken either with or without view information, the average *F*1 score was 94%. The accuracy of the site-specific view recognition models was 100%.

Keywords Image processing · Classification · Machine learning · Principal component analysis · Image-guided radiation therapy

✉ Deshan Yang
yangdeshan@wustl.edu

[1] Department of Radiation Oncology, Washington University School of Medicine, St. Louis, MO, USA

Introduction

The patient positioning using kV imaging onboard a medical linear accelerator represents the state of the art for image-guided radiation therapy (IGRT). After the patient is positioned on the treatment couch on the treatment day based on skin markers, the radiation therapist operates the onboard X-ray image system to acquire two X-ray images in orthogonal cardinal directions (e.g., in the anterior-posterior direction and in the right-lateral direction). The two X-ray images will then be registered to the digital reconstructed radiography (DRR) images, which were previously computed using treatment planning CT images in the same cardinal directions, to determine the amount of the treatment couch table translational shifts. After the couch table shifts are applied so that the patient on the couch table is aligned to the planned treatment position, the therapist can then turn on the radiation treatment beams. This procedure is referred to as 2D-2D image guidance, by which the patient positional alignment of the treatment day is determined under the guidance of two 2D X-ray images acquired in the orthogonal directions.

To ensure the image guidance accuracy, the radiation therapists commonly perform 2D image registrations manually at treatment machines. The radiation oncology physician will then spend 1 h or so daily to review the image guidance results manually for his/her patients performed by the therapist. Examples of incorrect patient setup detected by physicians include (1) the head of a head-neck cancer patient being not centered or straight and (2) a breast cancer patient being registered based on bones instead of breast surface. Incorrect or inaccurate alignments detected by physicians in the post-treatment reviews are estimated to be ~5% among all the 2D IGRT cases at authors' institution. Considering over tens of thousand treatments annually using 2D-2D image guidance at authors' intuition, there is a clear need to automate this manual

and time-consuming process to ensure accurate and safe treatment.

There have been multiple efforts to automate 2D X-ray image registration [1, 2]. However, studies have shown that automatic image registration methods may work relatively well for certain treatment sites, e.g., head, chest, and breast cases, but perform poorly for some other treatment sites, e.g., pelvis [3]. Therefore, 2D registration needs to be optimized for each treatment site to ensure accuracy and robustness [4, 5]. To our knowledge, there is no published method on automated anatomical site detection in 2D X-ray images. We are therefore motivated to develop a method to determine anatomical site automatically. Our aim is, by providing an anatomical site detection method, to enable site-dependent optimization, and therefore to improve both robustness and accuracy in 2D registration methods. Our ultimate future goal is to replace the manual registration and manual secondary confirmation processes with automatic, robust, and accurate 2D image registration algorithms therefore to improve the positional accuracies in radiation treatments for cancer patients.

To improve 2D image registration's robustness, it is also important to recognize image acquisition orientations, i.e., image views, automatically because the X-ray images acquired from the anterior or posterior direction of the patient are significantly different from those acquired from the left or right lateral direction due to varying body thickness. There are limited studies in the literature that are relevant to image view detection. These studies focused on chest radiographs only, using either heuristic algorithm [6] or artificial neural network [7]. Template-based method [8], nearest-neighbor classifier [9], and artificial neural network [10] have also been proposed to identify posterior-anterior and lateral views of the chest radiographs.

In this study, to detect anatomical site and image view automatically in 2D IGRT images, we developed a hierarchical multiclass recognition model. Principal component analysis (PCA) was employed to extract robust and low dimensional feature representation. Support vector machine (SVM) was then employed to model binary classifiers. A dataset including clinical images from six anatomical sites was used to train and test the model.

Methods and Materials

Materials

With an IRB approval, 360 kV X-ray images acquired from 180 patients with two each patient (one in AP view and one in RT view) were included in this study. The images were retrospectively obtained as DICOM image files from MOSAIQ (Elekta, Stockholm, Sweden), the treatment management system (TMS) used at author's institution. The treatment site and view information for each image as also obtained from the patient treatment prescriptions and treatment delivery records in MOSAIQ. Six treatment sites were included in this study (the brain, head-neck, breast, lung, abdomen, and pelvis), and 30 patients each site were randomly selected. Note that the patients received radiation treatment on multiple days, and the two X-ray images for each patient were randomly selected from the same treatment day.

The images were originally acquired in seven treatment rooms. Each room is equipped with different treatment machines from a single vendor (Varian Medical System, Palo Alto, CA, USA) but of different models (three 10-year-old Trilogy models, one 5-year-old IX model, and three 1 to 3-year-old TrueBeam models). The kV X-ray imaging system on each treatment machine consists of a kV X-ray generator, which operates at 40 to 160 kVp and configurable ms and mA settings, and a flat-panel X-ray detector (1024 × 768 pixels, 40 × 30 cm, 0.39 mm pixel resolution) of different ages. The imaging parameters (kVp, mAs) were manually configured by the radiation therapist at the time of image acquisition, depending on body sizes, anatomical sites, and imaging views. These imaging parameters were not stored in MOSAIQ and therefore not available to be considered in this study. Note that common imaging parameter settings were often used: 80–120 kVp, 40–60 ms, and 80–120 mA for the thorax and breast; 100–140 mVp, 150–200 ms, and 90–100 mA for the abdomen and pelvis; and 60–70 kVp, 50 ms, and 100 mA for the head, neck, and brain. The distance from the X-ray source to the patient's midline is 100 cm and the source-to-detector distance is 150 cm.

As shown in Fig. 1, the 2D X-ray images obtained in the 2D image guidance process have significantly inferior image quality than diagnostic X-ray images, e.g., chest radiographs. Common problems are image intensity inconsistency, anatomical overlap between different treatment sites, and interference resulted from treatment accessories, such as thermoplastic head masks, patient immobilization devices, and motion management devices. The images for three patients are presented for each site to demonstrate image variation for the same site.

Overview of the Site and View Recognition Method

We approached the site and view recognition problem as an image classification problem. Treatment sites or views can be defined as classes. As shown in the flowchart in Fig. 2, the method has two phases: model training and testing. There are two steps in the model training phase: (1) to transform the high dimensional X-ray image representative vectors into low dimensional vectors and (2) to train the class recognition model with the low dimensional vectors. In the testing phase, i.e., the application of the recognition model to recognize an X-ray image, a new X-ray image is projected onto the lower-

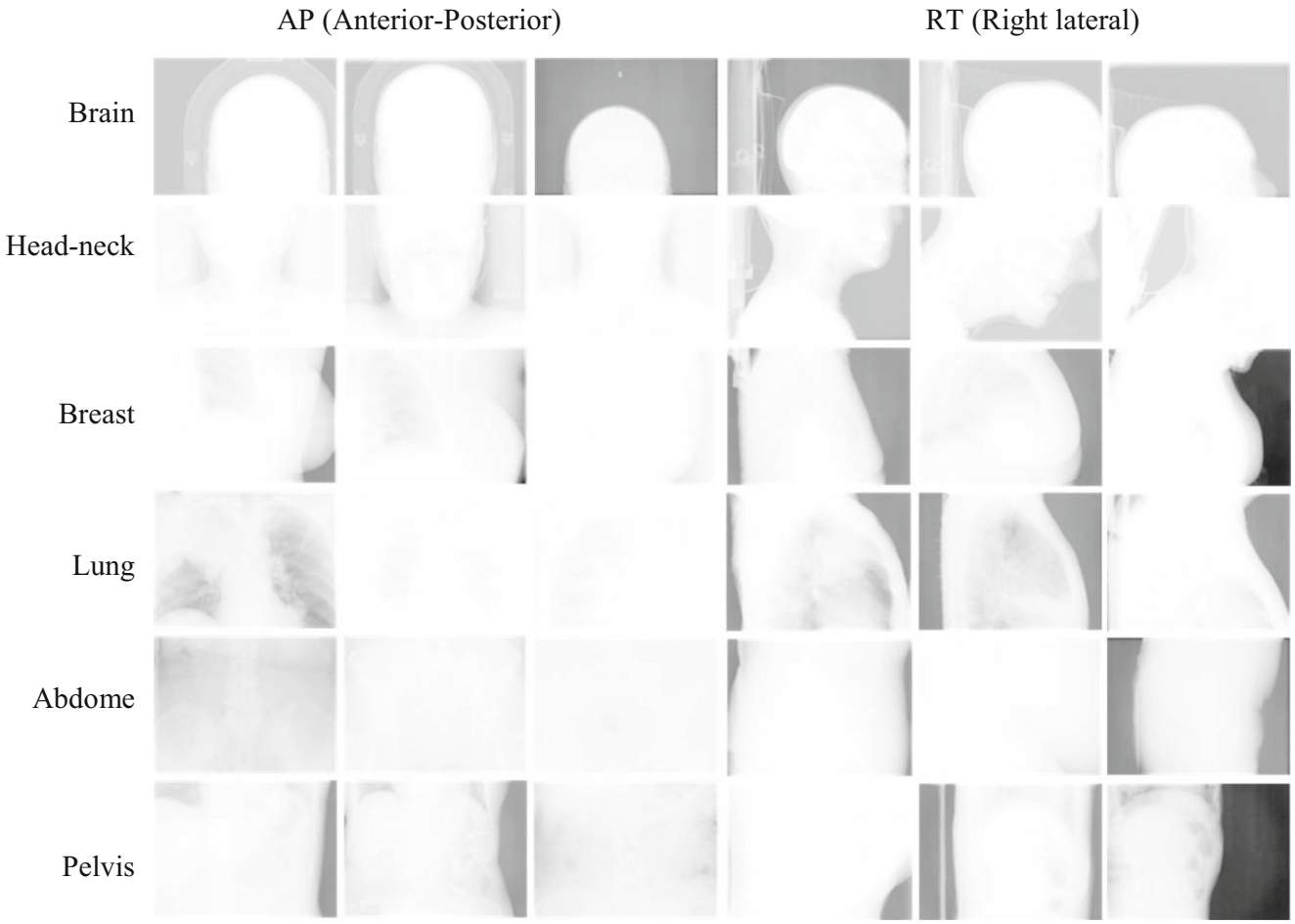

Fig. 1 Sample X-ray images from six sites in two views. AP and RT X-ray images of the same site for three patients are shown in each row. The first three images from left in every row are AP images. The last three images from left in every row are RT images. From top to bottom, the images in each row are from the brain, head-neck, breast, lung, abdomen, and pelvis, respectively

dimensional space with the transformation matrix learned in the training phase to yield a low 290,400 dimensional vector, which is then used by the learned classifier to recognize the class of the image.

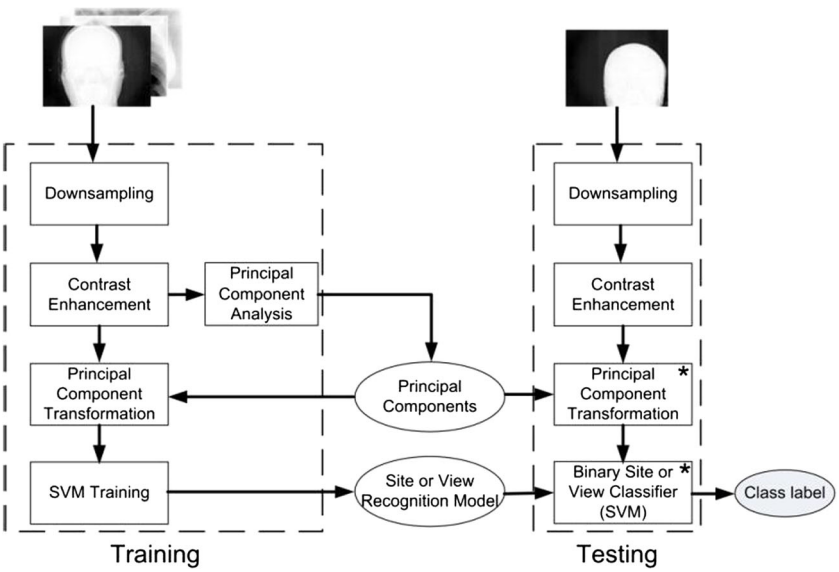

Fig. 2 The proposed site or view recognition workflow. The modules labeled with a *star* are applied to every binary classifier in the proposed hierarchical site and view recognition model

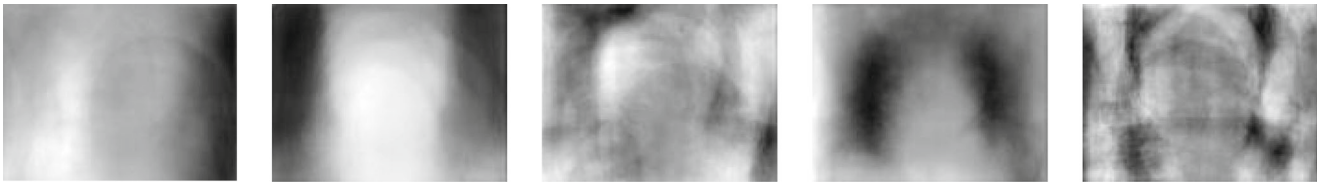

Fig. 3 Demonstration of the eigenvectors obtained by PCA on the sample X-ray images for all treatment sites. One can see the brain, head-neck, breast, lung, and abdomen/pelvis in the images (a–e), respectively

Prior to training and testing, the X-ray images were down-sampled from 1024 × 768 to 128 × 96 pixels by using a bi-cubic interpolation algorithm [11]. The purpose of down-sampling was to reduce space and time complexity of both training and test process; however, down-sampling may worsen accuracy. We tested the flat site recognition algorithm using the original images and the images with lower resolutions (512 × 384, 256 × 192, 128 × 96, and 64 × 48) to determine an optimal resolution for site and view recognition task. We found that the accuracy was identical for images with 1024 × 768, 512 × 384, 256 × 192, and 128 × 96 resolutions and worse for a 64 × 48 resolution. Therefore, the optimal resolution was deemed to be 128 × 96.

A window/level contrast enhancement [12] was then employed to enhance the contrast of the down-sampled image. The level was set at the middle of the intensity range. The upper boundary of the window was set at the lowest intensity value of those 1% pixels with highest intensity values. The lower boundary of the window was set at the highest intensity value of those 1% pixels with lowest intensity values.

Feature Representation and Dimension Reduction

The pixel intensities of X-ray images after down-sampling and contrast enhancement are used as features to represent images. The rows of a processed X-ray image are concatenated into a vector of 12,288 elements. To significantly reduce the dimension of representation vectors so that the site or view can be detected faster and more accurately, PCA is employed to project the X-ray image feature representation vector onto a low dimensional feature space [13]. The new feature space is spanned by the principal components, which are obtained by maximizing the sum of the variance of sample images. The uncorrelated eigenvectors of the data covariance matrix are used to map the data into a new representation. Larger eigenvalues indicate the larger total variance in the sample data set. We use m eigenvectors corresponding to the m largest eigenvalues to project the vectorized image data x onto the PCA space $\mathscr{M}$ such that:

$$t = U^T x \tag{1}$$

where, U is the transformation matrix in which i-th column is a eigenvector corresponding to i-th largest eigenvalue, and t i s a projection of an vector x representing image in $\mathscr{M}$, having m elements. Site or view recognition is then performed in the PCA space $\mathscr{M}$, where the data points are optimally distributed.

The eigenvectors obtained by performing PCA on the collected X-ray images, rearranged back to 2D images of 128 × 96 pixels, are shown in Fig. 3. We computed the eigenvectors using the X-ray images for all six treatment sites. One can recognize the general treatment sites in these images of eigenvectors to be the brain, head-neck, breast, lung, and abdomen/pelvis in Fig. 3a–e, respectively. Fig. 3e, which

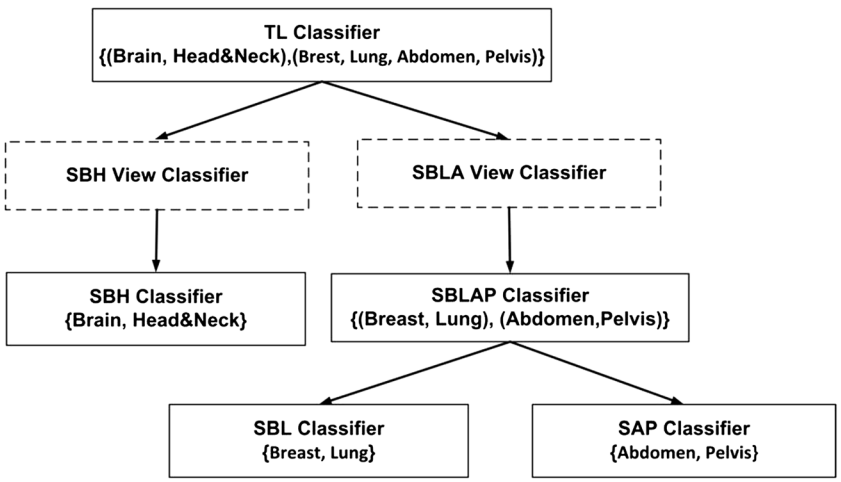

Fig. 4 Hierarchical site and view recognition model. *Dashed boundary lines* refer to optional modules

represents both pelvis and abdomen, demonstrates that simple single-step, multi-site recognition might be ineffective in distinguishing pelvis images from abdomen images due to the high similarity between these two sites.

Site and View Recognition with a Hierarchical Recognition Model

Introduction—a Basic Binary Classifier

We employed a support vector machine (SVM) [14] to build a recognition model in the PCA space $\mathcal{M}$ for recognizing sites and views. By finding an optimal discriminative function with largest separation between the two classes, an SVM efficiently performs a non-linear classification by implicitly mapping inputs into high-dimensional feature spaces. The general linear discriminative function is written as follows:

$$y = w^T \Phi(t) \tag{2}$$

where, w is a projection direction in the feature space, $\Phi(t)$ is a vector of the basis functions that could be nonlinear functions, and $\boldsymbol{t}$ is a vector representation of an image. An image is assigned to one of the two categories according to the sign of y value.

A Hierarchical Model for Multi-Site Recognition

Flat multiclass SVM classifiers, e.g., one-against-one and one-against-rest SVM classifiers, are often used to recognize multiple classes (sites or views in this study), and voting processes [15] are used to determine a final classification from the results of multiple binary classifiers. Such flat multiclass SVM classifiers have severe deficiencies, notably including that a one-against-one SVM classifier might be applied to unrelated classes thus introducing a potential source for misclassification. One-against-rest SVM classifiers suffer from the problem that the scale of the confidence values may differ between the binary classifiers.

To avoid the potential problem of flat multiclass SVM classifiers and to improve the overall recognition accuracy, we have therefore designed a hierarchical multiclass recognition model. As illustrated in Fig. 4, the hierarchical model is a binary tree with binary SVM classifiers as nodes. The model is designed to recognize general sites first and then to recognize the more specific sites within a recognized general site. A general site is defined as a collection of multiple sites with similar appearances. We group the treatment sites in two levels, from general to more specific, as (1) the brain and head-neck (SBH), (2) the breast, lung, abdomen, and pelvis (SBLAP), and under SBLAP, (2.1) the breast and lung (SBL), and (2.2) the abdomen and pelvis (SAP).

The differences between the proposed hierarchical model and the previous general hierarchical SVM model [16] are (1) considering the appearance similarity between the site images to design the topology of hierarchical model and (2) designing a hierarchical subspace corresponding to the hierarchical classification model by using PCA. That is, the PCA transformation is performed for each binary classification model.

Three site recognition scenarios are supported by the hierarchical models: (1) two X-ray images in orthogonal views are available with known view information, (2) two X-ray images in orthogonal views are available but the view information is unknown, and (3) a single X-ray image is available with or without the view information. The nodes in Fig. 4 with dashed boundary lines are optional nodes for the second scenario.

For the first scenario, the view information is known, and the view classifiers (indicated by the dashed boundary line) will not be applied. For the second scenario, an image will be randomly picked from two available unknown views, and a general site will be recognized by using the TL classifier before the views are recognized using the view classifier of a general site.

Similarly, for the third scenario, if a single image is given with view information, the view recognition will not be applied. If the view information is not available for the single image, the site will be recognized before the site-specific view classifier is applied to recognize the view of the image. The site-specific view recognition accuracy is 100%, as shown in our results. The overall accuracy of site recognition with two images in orthogonal views is greater than with a single image in any view, because the information provided in two images is more comprehensive.

The site classifiers below the view classifiers use both images in AP and RT views if two images are available. The feature vectors of two views are concatenated together (AP before RT) before PCA is employed to reduce the feature dimensions. If one or both images are in views other than the required AP or RT view (e.g., in LT or left), the images will be flipped accordingly and be converted from LT to RT, or PA to AP.

Implementation

We implemented the hierarchical recognition algorithms in Python, with the binary classifiers specifically implemented by using PCA and SVM algorithms within the Scikit-learn library.[1]

For each binary recognition model, we used the eigenvectors corresponding to the top N largest eigenvalues to create a transformation matrix. The number N for each binary recognition model was determined by the eigenvectors achieving

[1] http://scikit-learn.org/stable/

Table 1 Performances of hierarchical site recognition models using three evaluation metrics. Each entry in the table is a mean value of a metric of 20 cross-validation experiments

Methods	Metrics	Brain	Head-neck	Breast	Lung	Abdomen	Pelvis	Average
AP-RT	Precision	1.0	1.0	0.975	1.0	1.0	1.0	0.9958
	Recall	1.0	1.0	1.0	0.95	1.0	1.0	0.9916
	*F*1	1.0	1.0	0.9833	0.9744	1.0	1.0	0.9930
TU	Precision	1.0	1.0	0.9473	0.9298	1.0	0.9824	0.9766
	Recall	0.9736	1.0	0.9736	0.9473	0.9736	1.0	0.9780
	*F*1	0.9824	1.0	0.9561	0.9368	0.9824	0.9894	0.9745
OU	Precision	0.9857	0.9833	0.9188	0.9033	0.9214	0.9428	0.9417
	Recall	0.9666	0.95	0.9333	0.8833	0.9	0.9833	0.9361
	*F*1	0.9741	0.9651	0.926	0.9107	0.9087	0.9601	0.9408

TU two unknown views, OU one unknown view

best results in the measurement of *F*1 score in the cross-validation experiments with the training data of the model. We found that less than 30 eigenvectors corresponding to largest eigenvalues are sufficient to create a PCA model with good performance. Therefore, we choose N in the range from 1 to 30.

For SVMs, we employed a series of leave-one-out experiments to determine the optimal choices of the kernel functions and associated parameters. In the experiments, we choose between linear and Gaussian kernel functions, the tradeoff coefficient C of the SVM from $\{2^k | k \in [-3, 3]\}$, and the variance of the Gaussian kernel function δ in $\{2^k | k \in [-4, 3]\}$. The X-ray image set introduced in Section Materials was used to test the performance of the proposed site and view recognition methods. A cross-validation experiment was conducted in which the image set was evenly divided into 30 sets such that we use 29 images (per site) to train a model in each set, and then use the remaining one image to test the model.

Evaluation

Three metrics including precision, recall, and *F*1 score were used to evaluate the proposed methods. Precision is the number of true predicted positive instances divided by the number of all predicted positive instances. Recall is the number of true predicted positive instances divided by the number of all positive instances. The *F*1 score is defined by the following:

$$F1 = 2\left(\frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}\right) \tag{3}$$

The *F*1 score considers both precision and recall. Its best value is 1 and its worst is 0. All of the recognition performances given in the tables are the average results of 30 cross-validation experiment.

The paired sample t test [17] was employed to determine whether performances of two models are statistically significantly different, e.g., whether a hierarchical site recognition model and a flat site recognition model are statistically significantly different. For each model, the *F*1 scores were obtained from a 30-fold cross-validation experiment. The paired sample t test was then performed on the two samples obtained from the cross-validation experiment of two models. The significance level was set as 0.05.

The more images used to train the site recognition model, the more robust the models could become. We plotted the learning curve of the site recognition models to test the model's performance change with increasing number of the training images. The cross-validation experiments were performed ten times. The mean and standard deviation of the

Table 2 Performances of the binary classifiers in the hierarchical site recognition model AP-RT where two known views are available for each site. In a same given row, two neighboring cells with same shading style show the *F*1 scores of two "sites" (general sites or more specified sites) handled by a binary classifier. Each cell crosses one or more columns of sites. Each cell shows the *F*1 score of recognizing a general or specific site with a binary classifier

	Brain	Head-neck	Breast	Lung	Abdomen	Pelvis
Top Layer	1.0				1.0	
Middle Layer	1.0	1.0	1.0		1.0	
Bottom Layer			0.9833	0.9744	1.0	1.0

Table 3 Performances of the flat site recognition methods using three evaluation metrics. Each entry of the table is a mean value of a metric of 20 cross validation experiments

Methods	Metrics	Brain	Head-neck	Breast	Lung	Abdomen	Pelvis	Average
AP-RT-Flat	Precision	0.95	1	0.975	0.95	0.883333	0.975	0.9513
	Recall	1	0.975	0.975	0.95	0.9	1	0.9625
	*F*1	0.9740	0.9873	0.9666	0.95	0.89	0.9833	0.9585
OU-Flat	Precision	0.9714	0.969	0.8907	0.893	0.8678	0.9071	0.9165
	Recall	0.9666	0.9333	0.8666	0.8333	0.8666	0.9666	0.9055
	*F*1	0.9664	0.9483	0.8734	0.8505	0.8604	0.9315	0.9051

computed *F*1 score were then shown on the learning curve. In the first experiment, the training set included the images for 12 patients from six sites with two patients each site. In subsequent experiments, the images for 12 additional patients from six sites were added into the training set prior to re-training the model. Only the classifiers with images in two views were included in this learning curve experiment.

Results

The proposed hierarchical site and view recognition models were experimentally tested for the three scenarios outlined in Section of the hierarchical recognition model. The models applied for the three scenarios are named AP-RT, TU (two unknown views), and OU (one unknown view), respectively. The experimental results of the three scenarios are given in Table 1. AP-RT achieved the best average accuracy and 100% recognition accuracy on four of six sites. For the cases of two unknown views, the view prediction errors affect the accuracy of recognizing more specific sites in the lower layer.

The performance of the binary classifiers in the hierarchical model for the AP-RT scenario, which would be the most used scenario clinically, is shown in Table 2. All the higher-layer binary classifiers have 100% recognition accuracy, and the breast-lung classifier in the lower layer also shows very good performance. The performance of the higher layer classifiers assures a good overall performance of the hierarchical model.

We compared the performances of the proposed hierarchical classification models and the flat multiclass classification models for the first and third scenarios. The flat models are named as AP-RT-Flat and OU-Flat, respectively. The experimental results of these two methods, are given in Table 3, together with the results listed in Table 1, demonstrate that the proposed hierarchical classification models have better performance. The second scenario was skipped from the comparison due to additional and unrelated uncertainties associated with view recognition. The p value from a paired sample t test on experimental results of AP-RT and AP-RT-Flat was 0.045. The p value from a paired sample t test on experimental results of OU and OU-Flat was 0.048. The p values suggest that the hierarchical site recognition models statistically significantly outperform the flat site recognition models.

Two confusion matrices are shown in Fig. 5. Each entry of the confusion matrix is a number of instances that the site specified in a given row were actually identified as the site labeled by a corresponding column.

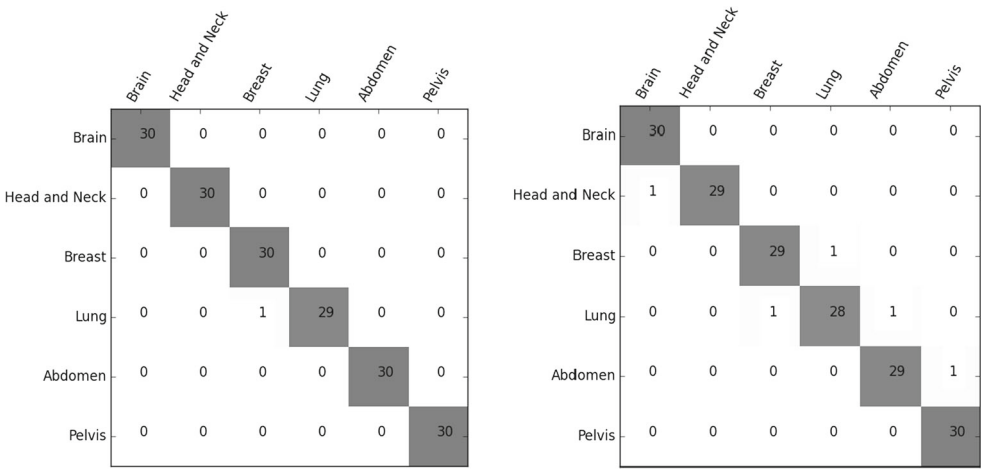

Fig. 5 The confusion matrices of the site recognition results of the hierarchical AP-RT mode and the AP-RT-Flat model. Each row corresponds to a true class. Each column corresponds to a predicted treatment site. Each entry in the matrices is the total number of test instances of 20 cross-validation experiments

a Confusion Matrix for AP-RT

b Confusion Matrix for AP-RT-Flat

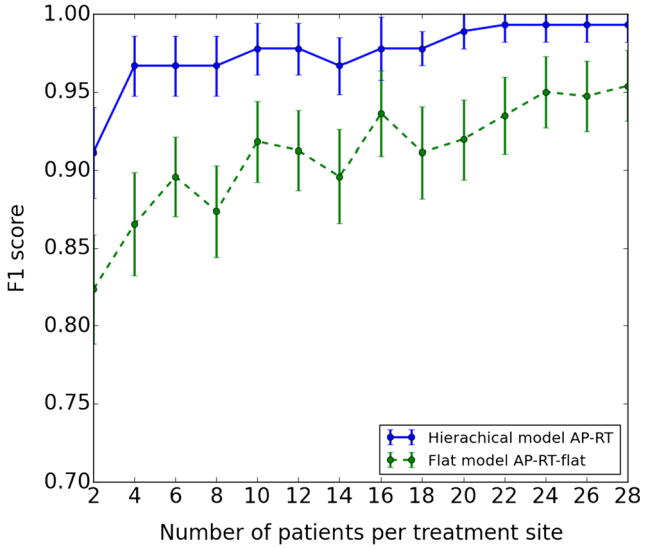

Fig. 6 Learning curves of the AP-RT and AP-RT-Flat models. For each number of patients (N_{Pat}), the number of images included in training the models was $N_{Pat} \times 2$ (views) × 6 (sites)

The confusion matrix for AP-RT (Fig. 5a) shows that only one lung image was misclassified as a breast image. The cause of the error was confirmed to be that the lung image in AP view had similar appearance with some breast images in AP view.

More recognition errors can be found in the confusion matrix of AP-RT-Flat in Fig. 5b. The classifier became confused between the brain and head-neck, breast and lung, lung and abdomen, and abdomen and pelvis. Compared to Fig. 5a, one can see that most recognition errors of AP-RT-Flat were corrected in AP-RT. This comparison confirms that the hierarchical design of the multiclass recognition model is able to reduce the recognition errors between sites with distinct appearances, e.g., the lung and abdomen, and to improve the recognition results between sites with similar appearances, e.g., the brain and head-neck, and abdomen and pelvis.

We conducted experiments to generate the learning curve of the hierarchical site recognition model AP-RT and the flat site recognition model AP-RT-Flat. From the two learning curves are shown in Fig. 6, one can see that the recognition accuracies improve as the number of training images increase. The hierarchical model only needs images of four patients from each site to achieve an averaged accuracy of 96%. The flat model AP-RT-Flat needs images from 28 patients per site to achieve an accuracy of 95%, averaged for the six sites.

Most X-ray images have low contrast, making it difficult for analysis by eye and computer algorithm. Therefore, we applied window/level contrast enhancement algorithm to improve the contrast of X-ray images. The results of performing the three hierarchical site recognition models with the original X-ray images are given in Table 4. Comparing the results in Tables 1 and 4, we can see that the performance of site recognition is improved by processing the X-ray images with the contrast enhancement algorithm. The p value from a paired sample t test on experimental results of the AP-RT site recognition model with and without image contrast enhancement without image contrast enhancement was 0.477. This suggested that the contrast enhancement step did not statistically improve the site recognition accuracy of the AP-RT model. Similarly, a p value = 0.12 was obtained for the TU model and suggested that contrast enhancement did not statistically improve the site recognition accuracy of the TU model. However, a p value = 0.01 was obtained for the OU models and suggested that image contrast enhancement could significantly improve the accuracy of the site recognition on a single X-ray image.

We also tried other more advanced contrast enhancement methods, including adaptive histogram equalization (AHE) [11] and contrast limited adaptive histogram equalization (CLAHE) [18]. Our results however suggested that the site and view recognition performances have not been further improved by using AHE and CLAHE.

We also tested the accuracy in recognizing the right lateral (RT) and left lateral (LT) views in the view recognition step. To generate data in order to train the view classifier, we randomly selected half of the RT view images from six sites in the

Table 4 Performances of the hierarchical site recognition models on the original X-ray images without contrast enhancement

Methods	Metrics	Brain	Head-neck	Breast	Lung	Abdomen	Pelvis	Average
AP-RT	Precision	1.0	1.0	0.9666	1.0	1.0	0.95	0.9861
	Recall	0.925	1.0	1.0	0.975	0.975	1.0	0.9791
	*F*1	0.95	1.0	0.98	0.9833	0.9833	0.97	0.9777
TU	Precision	0.9473	1.0	0.9473	0.8508	0.9649	0.9210	0.9385
	Recall	0.8421	0.9736	1.0	0.8684	0.9736	1.0	0.9429
	*F*1	0.8771	0.9824	0.9684	0.8491	0.9614	0.9508	0.9315
OU	Precision	0.9333	1.0	0.825	0.95	0.7916	0.8583	0.8930
	Recall	0.9	1.0	1.0	0.875	0.85	0.775	0.9
	*F*1	0.9066	1.0	0.885	0.9	0.805	0.7983	0.8825

TU two unknown views, *OU* one unknown view

data set and flipped them into LT views. A leave-one-out method was employed to test the view recognition accuracy. The experimental results indicate that the accuracy of recognizing view in RT and LT is 100% for each site.

Discussion

The proposed hierarchical model has demonstrated overall good performance in site and view recognition using online kV images acquired from a variety of treatment sites. The model is useful as a step to automatically determine site and view information in the onboard kV images, thereby allowing site and view dependent optimization in the subsequent image processing tasks. The primary clinical application is an automated 2D-2D image registration for more precise patient positioning in IGRT. Collaborating with a RT software vendor, we envision the following steps in terms of clinical implementation. The proposed method will first be integrated into clinical systems to provide anatomical site and view information automatically, and the 2D-2D image registration algorithms in the clinical systems will then be improved to select optimal parameters automatically based on the site and view information. In this way, the 2D-2D registration methods in the clinical image guidance systems would be more robust and accurate. The improved and more robust 2D-2D registration methods are expected to be able to outperform manual registration in many cases, and, in other cases, be used confidently as a pre-registration step preceding manual registration. In either scenario, the accuracy of the image-guided treatment setup can be improved. The proposed method could also be useful for supporting a secondary independent verification of patient setup that has been performed using a similar 2D-2D image registration method. The third clinical application is enabling automatic and site-dependent image contrast enhancement to assist manual RT image reviewing.

In the future work, we will extend the proposed model to recognize more sites, e.g., the shoulder and leg, and to recognize localized areas of a disease site, e.g., right and left sides of the lung or breast. We will also extend the model to distinguish images in more views including left lateral and posterior-anterior. The robustness of the proposed approach could be further improved by developing more image preprocessing functions to identify the artifacts in images such as the bowtie filter and couch table, which could affect site and view recognition accuracy.

Conclusions

In this study, we developed a hierarchical site and view recognition model to support three clinical scenarios of site and view recognition. The results demonstrate that the proposed method achieved good performance in both automated site and view recognition in radiation therapy localization X-ray images. The proposed method could be useful for automating site and view dependent X-ray image processing for various clinical applications and for automating the DICOM tag correction.

Acknowledgements Research reported in this study was supported by the Agency for Healthcare Research and Quality (AHRQ) under award 1R01HS0222888. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency of Healthcare Research and Quality.

References

1. Asvestas PA, et al: ESTERR-PRO: a setup verification software system using electronic portal imaging. Int J Biomed Imaging 2007, 2006
2. Wu J, Samant SS: Novel image registration quality evaluator (RQE) with an implementation for automated patient positioning in cranial radiation therapy. Med Phys 34(6):2099–2112, 2007
3. Sutherland K, et al: Detection of patient setup errors with a portal image–DRR registration software application. J Appl Clin Med Phys 12(3), 2011
4. Li G, et al: Clinical assessment of 2D/3D registration accuracy in 4 major anatomic sites using on-board 2D Kilovoltage images for 6D patient setup. Technol Cancer Res Treat 14(3):305–314, 2015
5. Nie K, et al: Site-specific deformable imaging registration algorithm selection using patient-based simulated deformations. Med Phys 40(4):041911, 2013
6. Pietka E, Huang H: Orientation correction for chest images. J Digit Imaging 5(3):185–189, 1992
7. Boone JM, Seshagiri S, Steiner RM: Recognition of chest radiograph orientation for picture archiving and communications systems display using neural networks. J Digit Imaging 5(3):190–193, 1992
8. Arimura H, et al: Development of a computerized method for identifying the posteroanterior and lateral views of chest radiographs by use of a template matching technique. Med Phys 29(7):1556–1561, 2002
9. Lehmann TM, et al: Determining the view of chest radiographs. J Digit Imaging. 16(3):280–291, 2003
10. Boone JM, et al: Automated recognition of lateral from PA chest radiographs: saving seconds in a PACS environment. J Digit Imaging 16(4):345–349, 2003
11. Pizer SM, et al: Adaptive histogram equalization and its variations. Comput Vis Graph Image Process 39(3):355–368, 1987
12. Acharya T, Ray AK: Image processing: principles and applications. Hoboken: Wiley, 2005
13. Turk M, Pentland A: Eigenfaces for recognition. J Cogn Neurosci 3(1):71–86, 1991
14. Hastie T, Tibshirani R, Friedman J: The elements of statistical learning. 2 ed. Springer series in statistics. New York: Springer New York Inc, 2006
15. Hsu C-W, Lin C-J: A comparison of methods for multiclass support vector machines. IEEE Trans Neural Netw 13(2):415–425, 2002

16. Vural V, Dy JG: A hierarchical method for multi-class support vector machines. in Proceedings of the twenty-first international conference on Machine learning. ACM, 2004
17. McDonald JH: Handbook of biological statistics, Vol. Vol. 2. Baltimore: Sparky House Publishing, 2009
18. Zuiderveld K: Contrast limited adaptive histogram equalization. In graphics gems IV. Cambridge: Academic Press Professional, Inc, 1994