

# Organ Localization & Identification in Thoracic CT Volumes using 3D CNNs Leveraging Spatial Anatomic Relations

Rajath Elias Soans and James A Shackleford

Electrical & Computer Engineering, Drexel University, 3141 Chestnut St, Philadelphia, PA 19104, USA

## ABSTRACT

In this paper, we present a model to obtain prior knowledge for organ localization in CT thorax images using three dimensional convolutional neural networks (3D CNNs). Specifically, we use the knowledge obtained from CNNs in a Bayesian detector to establish the presence and location of a given target organ defined within a spherical coordinate system. We train a CNN to perform a soft detection of the target organ potentially present at any point,  $\mathbf{x} = [r, \theta, \phi]^T$ . This probability outcome is used as a prior in a Bayesian model whose posterior probability serves to provide a more accurate solution to the target organ detection problem. The likelihoods for the Bayesian model are obtained by performing a spatial analysis of the organs in annotated training volumes. Thoracic CT images from the NSCLC–Radiomics dataset are used in our case study, which demonstrates the enhancement in robustness and accuracy of organ identification. The average value of the detector accuracies for the right lung, left lung, and heart were found to be 94.87%, 95.37%, and 90.76% after the CNN stage, respectively. Introduction of spatial relationship using a Bayes classifier improved the detector accuracies to 95.14%, 96.20%, and 95.15%, respectively, showing a marked improvement in heart detection. This workflow improves the detection rate since the decision is made employing both lower level features (edges, contour etc) and complex higher level features (spatial relationship between organs). This strategy also presents a new application to CNNs and a novel methodology to introduce higher level context features like spatial relationship between objects present at a different location in images to real world object detection problems.

**Keywords:** CNNs, deep learning, bayes classifier, spatial context, prior knowledge, organ identification, tensorflow

## 1. INTRODUCTION

The physical mechanisms driving many medical imaging techniques pose non-negligible risks to the patient. In the case of diagnostic X-ray computed tomography, the acquisition of increasingly detailed anatomical information—such as contrast ratio, field of view, and signal-to-noise ratio—comes at the cost of increased exposure to DNA damaging ionizing radiation. This cost of acquisition has been the subject of study for some time as the balance between actionable information and the potential detriment to the patient required for its retrieval remains a subject of active investigation.<sup>1-4</sup> Consequently, imaging protocols are often situationally adapted to reduce acquisition time and radiation dose, which results in reduced field of view, inconsistent setup geometry, poor contrast, and motion/acquisition artifacts.<sup>5</sup> Such variability across acquisition protocols has driven the need for organ identification/segmentation algorithms that are robust to such widely varying sources of noise.

The field of supervised machine learning has focused on solving the general problem of model approximation by means of parameter estimation from large sets of labeled example data—commonly referred to as *training data*. Furthermore, special care taken to ensure that such models avoid learning noise patterns in the training data has led to the development of very effective noise rejection techniques. Consequently, machine learning has become a valuable tool in the development of computer vision techniques involving the classification of noisy images.<sup>6</sup> These attributes make machine learning derived models highly attractive for organ identification, localization, and segmentation within noisy diagnostic CT image volumes.

The primary challenge in applying machine learning based techniques to the biomedical imaging domain, however, is the difficulty of acquiring adequately large and diverse labeled training datasets such that the resulting learned models

---

Further author information: (Send correspondence to James A Shackleford)  
E-mail: shack@drexel.edu, Telephone: 215-571-4269

are appropriately general—i.e. not overfit to the data.<sup>7,8</sup> Obtaining such a large and diverse training dataset requires the involvement of expert clinicians as well as computational experts with appropriate HIPPA certification and/or clinical collaborators willing to perform patient deidentification.<sup>9</sup> Although, the scale of data required can be dramatically reduced through the application of augmentation techniques, which simulate relevant realistic, common deformations; thereby effectively amplifying the amount of training data available.<sup>10</sup>

In recent years, convolutional neural networks (CNNs) have exhibited markedly improved performance over state of the art algorithms in the task of image classification.<sup>10</sup> Typically once trained, these networks are used to assign a categorical class label to a candidate image. Such architectures learn extractors for image features that best provide separability between classes, with each subsequent network layer encoding a progressively higher level pattern. For instance the first layer may detect edges, a second layer may detect patterns formed by connections between edges, and a third layer may recognize objects formed using those patterns.<sup>11</sup> Applications of CNNs to medical image segmentation and classification problems have successfully demonstrated the segmentation of organs at risk (OAR) in the head and neck region for radiology treatment planning,<sup>12</sup> identification of lung structure in chest radiographs,<sup>13</sup> and the delineation of the urinary bladder.<sup>14</sup>

In this paper, we introduce a method for increasing the robustness of a CNN classifier designed to detect thoracic organs by imposing a model that enforces local spatial context. Due to the limited availability, quantity, diversity, and cleanliness of labeled CT training data volumes, the accuracy of the CNN classifier may suffer; resulting in incorrect class labels at certain anatomical sites. We therefore use the CNN output only as a means for constructing a Bayesian prior that indicates a location in space where the target organ may be present with a given probability. The final categorical probability for a site is obtained as a posterior after imposing the known spatial distribution between neighboring organs as likelihood information.

The remaining sections are organized as follows. Section 2, discusses the methods used in this study. Here, insight is provided regarding the network architecture and adopted training procedure used for the presented experiments. Section 3 presents the results obtained in our training and testing phases. Finally, a brief discussion and future research suggestions are presented in Section 4.

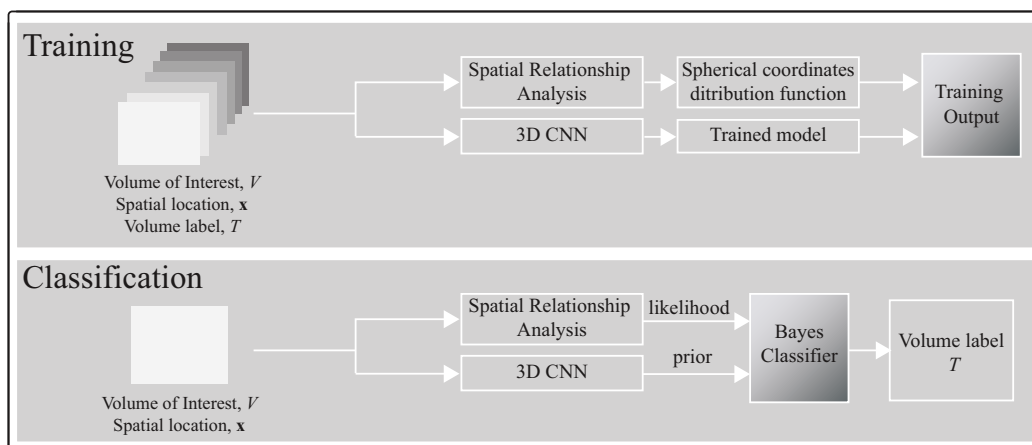


Figure 1: **Model Workflow Overview.** (top) *Model Training.* In addition to training the CNN in the traditional fashion, spatial likelihood distributions are constructed for each labeled organ with respect the other labeled organs. (bottom) *Organ Classification.* The CNN is provided a subvolume,  $V$ , taken from the full thoracic CT volume and produces an organ class probability. This class probability is fed to the Bayes classifier layer as a prior, which is then reevaluated given the evidence provided by the likelihood distribution constructed during training to produce the final volume label  $T$ .

## 2. METHODS

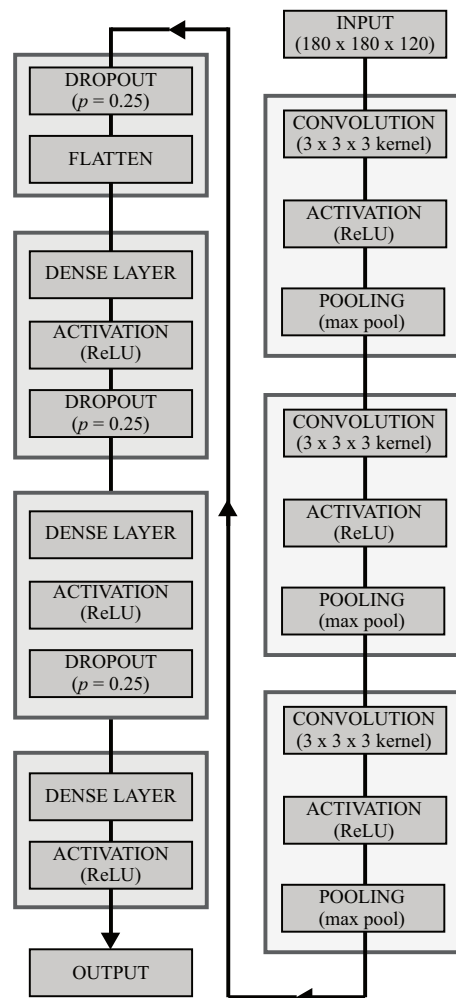
### 2.1 Convolutional Neural Network Architecture

The convolutional neural network described here is used to obtain the probability that a subvolume extracted from a full thoracic CT volume contains a specific organ, namely the left lung, right lung, or heart. From here on, we shall refer to

these subvolumes as potential volumes of interest (VOIs). As shown in Figure 2, the employed CNN architecture consists of an input layer, which accepts a VOI, immediately followed by a 3D convolution (CONV) layer, which uses a  $3 \times 3 \times 3$  kernel. Immediately following this convolution layer is a rectified linear unit (ReLU) layer implementing the activation function  $f(x) = \max(0, x)$ , which serves to introduce element wise non-linearity into the network.<sup>15</sup> Following the ReLU layer is a max pooling (POOL) layer, which serves to employ a simple means of spatial downsampling. This three layer sequence consisting of CONV  $\rightarrow$  ReLU  $\rightarrow$  POOL is repeated two more times, as shown in Figure 2, for a total of nine layers immediately following the input layer.

Immediately following the initial nine layers implementing convolution and max pooling are three fully connected hidden layers using the ReLU activation function. As shown, dropout is performed throughout the hidden layers in an attempt to mitigate overfitting to the training data.<sup>8</sup> The output derived from these final dense layers is a vector containing the probability of detecting the target organ within the VOI fed to the input layer. This vector is used as a categorical prior to a Bayesian classifier, which forms the final layer of the network.

The Bayesian classifier serves to compute the posterior probabilities of the VOI belonging to the various organ classes using a Gaussian likelihood distribution describing the spatial relationship between the target organ and a reference organ. The details of this Bayesian classifier are covered in depth in Section 2.5. Figure 1 illustrates the overall workflow.



**Figure 2: 3D CNN Network Graph.** Our CNN consists of three 3D convolution and max pooling units. After subsequent dropout and flattening, the output from these higher layers are fed to a fully connected layer where dropout ( $p = 0.25$ ) is employed throughout to avoid possible overfitting. Class probabilities are produced at the output of the CNN, which are subsequently used as prior probabilities to the Bayesian classification layer.

## 2.2 CNN Training

Training was performed using data from The Cancer Imaging Archive (TCIA). Specifically, the TCIA NSCLC–Radiomics thoracic CT dataset was used, which consists of 422 thoracic CT volumes each from a different non-small cell lung cancer (NSCLC) patient. Figure 3 shows a 3D isosurface rendering of a random CT volumes from this dataset. VOIs containing the heart (HT), right lung lobe (RL), and the left lung lobe (LL) were manually extracted from these volumes and labeled by hand. CNN detectors were then trained to identify the presence of these organs within an arbitrarily selected volume of interest from any given thoracic CT volume. Given the limited training data of 422 thoracic CT volumes, data augmentation was performed to mitigate overfitting. For this purpose, we developed a uniform cubic B-spline based data augmentation methodology that serves to emulate anatomical variation representative of population variance and differing respiration state. Given this augmented training data, CNN training was performed using the stochastic gradient descent optimizer provided by TensorFlow.<sup>16</sup>

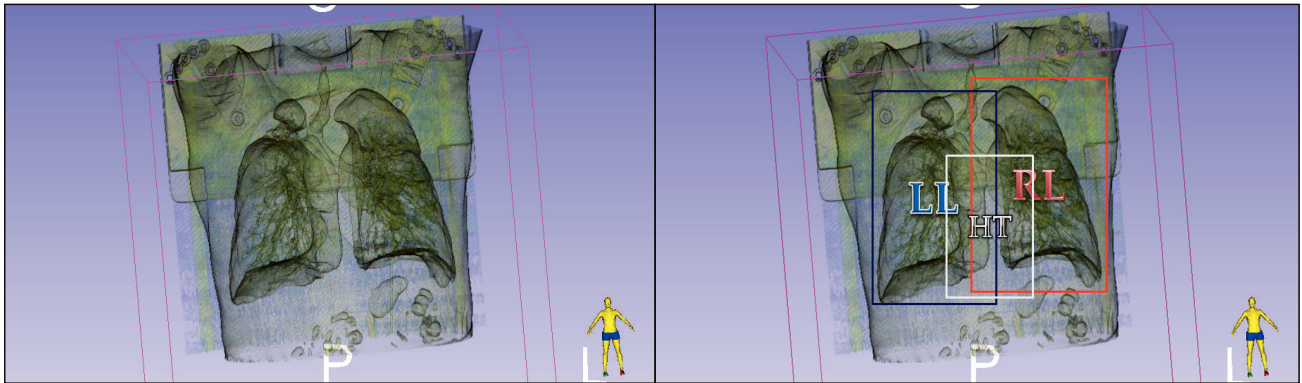


Figure 3: **Example Thoracic CT Volume.** (left) A randomly selected thoracic scan from the TCIA NSCLC–Radiomics dataset used in this study. Isosurface rendering with alpha channeling for visualization. (right) VOIs containing the target organs LL = Left lobe, RL = Right Lobe, HT = Heart.

## 2.3 Volume of Interest (VOI) Extraction

The main objective of this operation is to extract a sub-volume that encapsulates the target organ. Cuboid VOIs,  $V = [v_x, v_y, v_z, v_l, v_d, v_h]^T$ , are identified in the thoracic CT volume,  $I$ , such that  $V \subset I$  where  $[v_x, v_y, v_z]$  represents the point at the front-top-left corner of  $V$  in Cartesian coordinates. Similarly,  $[v_l, v_d, v_h]$  represents the length, depth, and height of  $V$ . An identified VOI,  $V$ , is assigned a label,  $t \in T = \{RL, LL, HT, NONE\}$ , depending on the target organ contained: right lung, left lung, heart, or nothing, respectively. Additionally, each VOI,  $V$ , is associated with a point,  $\mathbf{x} = [r, \theta, \phi]^T$  defined within a spherical coordinate system, which is related to the Cartesian coordinate system by the transform:

$$\begin{bmatrix} r \\ \theta \\ \phi \end{bmatrix} = \begin{bmatrix} \sqrt{x^2 + y^2 + z^2} \\ \arccos(\frac{z}{r}) \\ \arctan(\frac{y}{x}) \end{bmatrix}, 0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi \quad (1)$$

where  $x$ ,  $y$  and  $z$  are the coordinates of the VOI center in the Cartesian system with respect to image origin  $\mathbf{x}_0 = [0, 0, 0]^T$  i.e., top left corner of the first frame in the image volume:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} v_x + \frac{v_l}{2} \\ v_y + \frac{v_d}{2} \\ v_z + \frac{v_h}{2} \end{bmatrix} \quad (2)$$

Figure 4 illustrates the positioning of  $V$  at point,  $\mathbf{x}$ , in the coordinate axes.

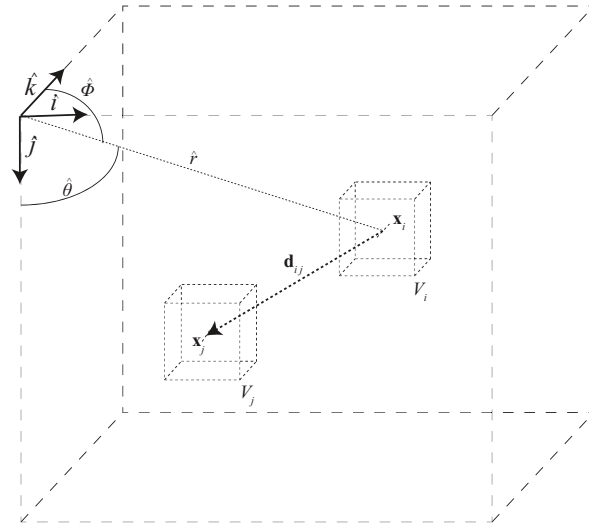
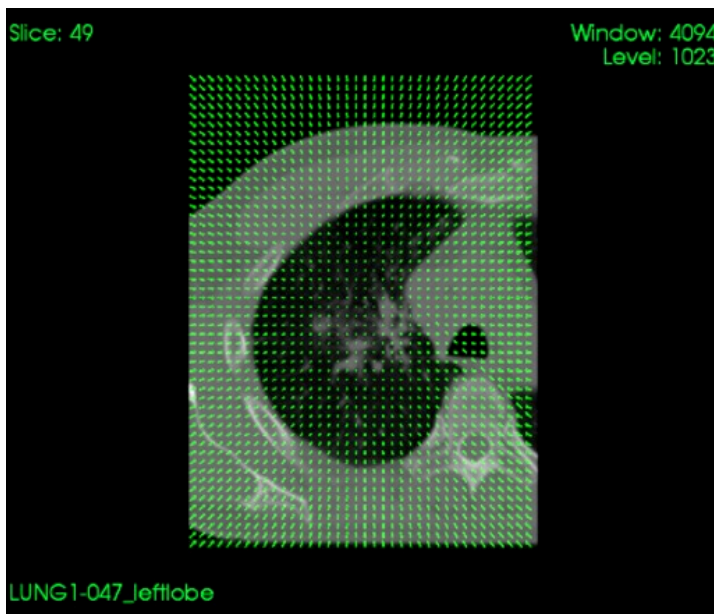
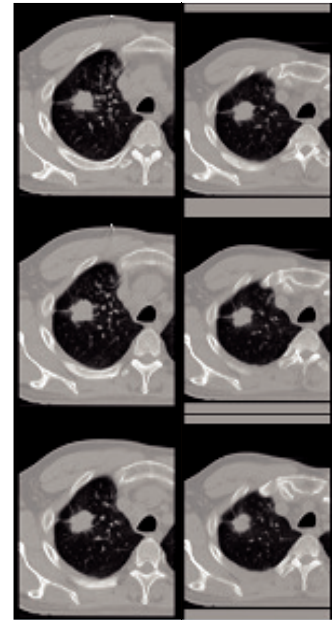


Figure 4: **Inter-VOI relationship in spherical space.** The spatial relationship between VOIs is defined by the vector  $\mathbf{d}_{ij}$ . The probability of finding a VOI with a specific label at a location relative to another VOI is described by a Bayesian likelihood computed using the labeled training data.



(a) Visualization of VOI with motion vector field.



(b) Augmentation visualization.

Figure 5: **Data Augmentation.** (a) The motion vector fields are generated to describe realistic variations in breathing patterns and size/shape. (b) Warped slices (right) show the variations introduced by our augmentation protocol applied to the original slices (left).

## 2.4 Improving Training Using Data Augmentation

The TCIA NSCLC-Radiomics dataset consists of 422 thoracic CT volumes. To compensate for this small number of true samples as well as the lack of physical diversification naturally found in larger human populations, a new method of augmentation was developed. In this method, elastic deformations are applied to training CT volumes in order to generate an augmented dataset more capable of training a CNN that maintains generality.<sup>10</sup>

Using this new method, augmentation of the original NSCLC thoracic CT volumes was achieved by employing elastic

deformations parameterized by the uniform cubic B-spline basis. This same deformation model is routinely solved to recover intra-patient respiratory motion as well as inter-patient anatomical mappings as they apply to a variety of clinical applications.<sup>17</sup> Consequently, employing this model allows for the realistic simulation of common anatomical variations in the anatomy found across more representative patient populations—such as heart size and the respiratory state of the lungs at the time of image acquisition. Figure 5 illustrates this data augmentation process. The resulting augmented volumes shown in Figure 5(b) were obtained using the Plastimatch<sup>18</sup> open-source medical image computation software.

## 2.5 Improving Detection Using Spatial Relationship Analysis

Any two VOIs,  $V_i$  and  $V_j$  encapsulating target organs determined by their labels,  $t_i$  and  $t_j$  are associated to points  $\mathbf{x}_i = [r_i, \theta_i, \phi_i]^T$  and  $\mathbf{x}_j = [r_j, \theta_j, \phi_j]^T$ , respectively, during VOI extraction. Considering that  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are two points in a spherical coordinate system, a vector  $\mathbf{d}_{ij}$  can be established between them as:

$$\mathbf{d}_{ij} = d_r \hat{r} + d_\theta \hat{\theta} + d_\phi \hat{\phi} \quad (3)$$

where  $d_r$  is the magnitude of the vector in the direction of unit vector,  $\hat{r}$ . We define vector  $\mathbf{d}_{ij}$  as the spatial relationship between  $V_i$  and  $V_j$ . Image volumes comprising the training set are used to establish the probability of finding a target organ at  $\mathbf{x}_i$  with respect to another at  $\mathbf{x}_j$ . The Gaussian distribution describing this relationship is used as the likelihood in the final Bayesian classification layer of our network.

## 2.6 Bayes Classifier

The naïve Bayes classifier is used as a final amended layer to our CNN. The implementation is straightforward. As described in Section 2.1, the class prior  $P(T = t)$  required by the Bayes classifier is obtained as the output from the previous fully connected layers in the network shown in Figure 1. This prior probability is then used in conjunction with the

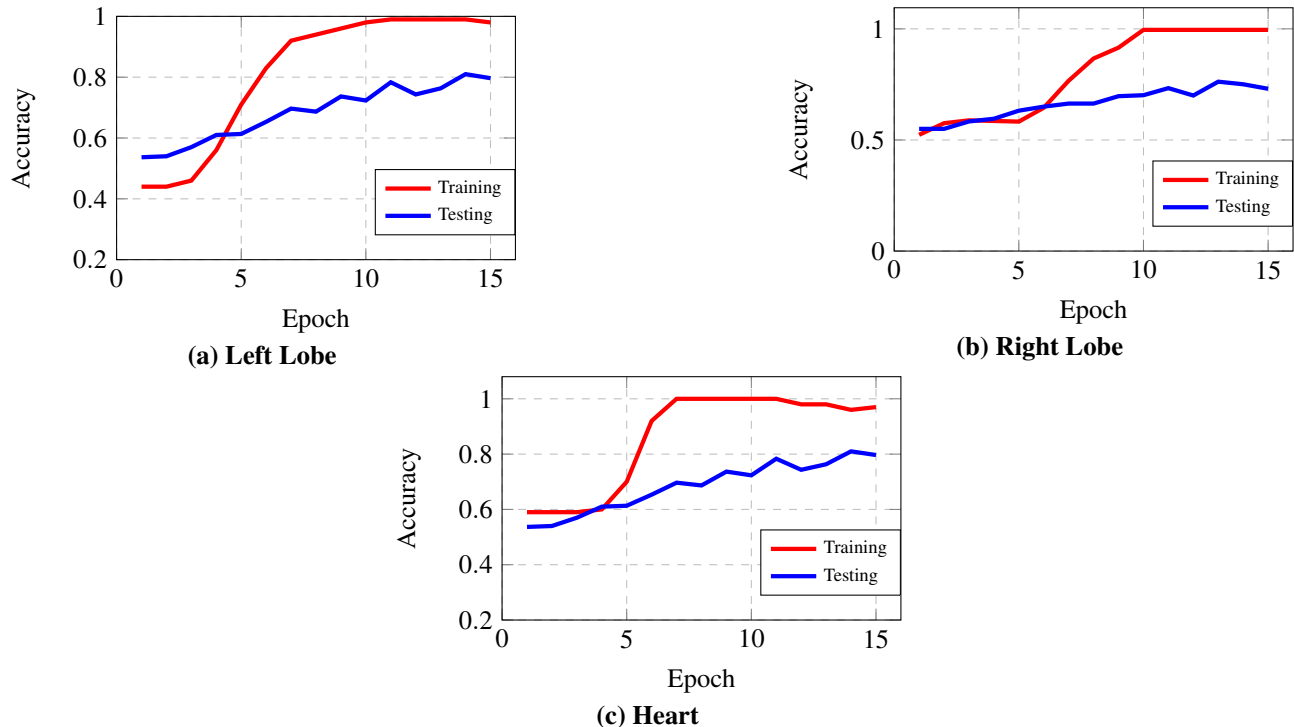


Figure 6: **Training Accuracy vs epoch with non-augmented data.** The variation of accuracy over 15 epochs of training for (a) the left lung, (b) the right lung, and (c) the heart. Notice how the accuracy changes are significant in the first few epochs.

marginal likelihood  $P(\mathbf{d}_{ij}|t)$ , which is easily obtained from the known spatial separation between organs in the training data. Applying the classic Bayes Rule for an arbitrary VOI,  $V_i$ , located at point,  $\mathbf{x}_i$  in space:

$$P(t|\mathbf{d}_{ij}) = \frac{P(\mathbf{d}_{ij}|t)P(t)}{P(\mathbf{d}_{ij})} \quad (4)$$

where the class prior probability,  $P(t)$ , is simply the probability that  $V$  belongs to the class  $T = t$ . This is obtained by supplying  $V$  to the input layer of the CNN illustrated in Figure 1, which produces  $P(T = t)$  at its output. The marginal likelihood  $P(\mathbf{d}_{ij}|t)$ , or evidence as some call it, is the probability of  $V_i$  existing spatially at point  $\mathbf{x}_i$  given that it belongs to the class  $T = t$ . As previously mentioned, this can be computed from the training examples by using VOIs labeled as belonging to class  $T = t$ . The term  $P(\mathbf{d}_{ij})$  is a measure of how probable it is for  $V_i$  to exist at any relative point  $\mathbf{D}_{ij}$  given all possible labels  $T$ . Finally, the posterior probability  $P(t|\mathbf{d}_{ij})$  provides the probability that  $V_i$  belongs to the class  $T = t$  given new evidence that it exists at point  $\mathbf{D}_{ij} = \mathbf{d}_{ij}$  within the CT volume relative to  $V_j$ . It should be noted that since the point  $\mathbf{d}_{ij}$  is taken in relation to another organ whose location is believed to be known with a high probability, there exists unique a Bayesian model for each such  $V_j$  “anchor organ.”

### 3. RESULTS

The proposed CNN was characterized using the thoracic testing volumes obtained from the TCIA NSCLC dataset. TensorFlow<sup>16</sup> was used to perform the characterization experiments on a machine equipped with dual HyperThreaded 3.2 GHz Intel Xeon Octo-Core E5-2630v3 processors, 512 GB DDR4 RAM, and dual NVIDIA GeForce GTX 980 GPUs.

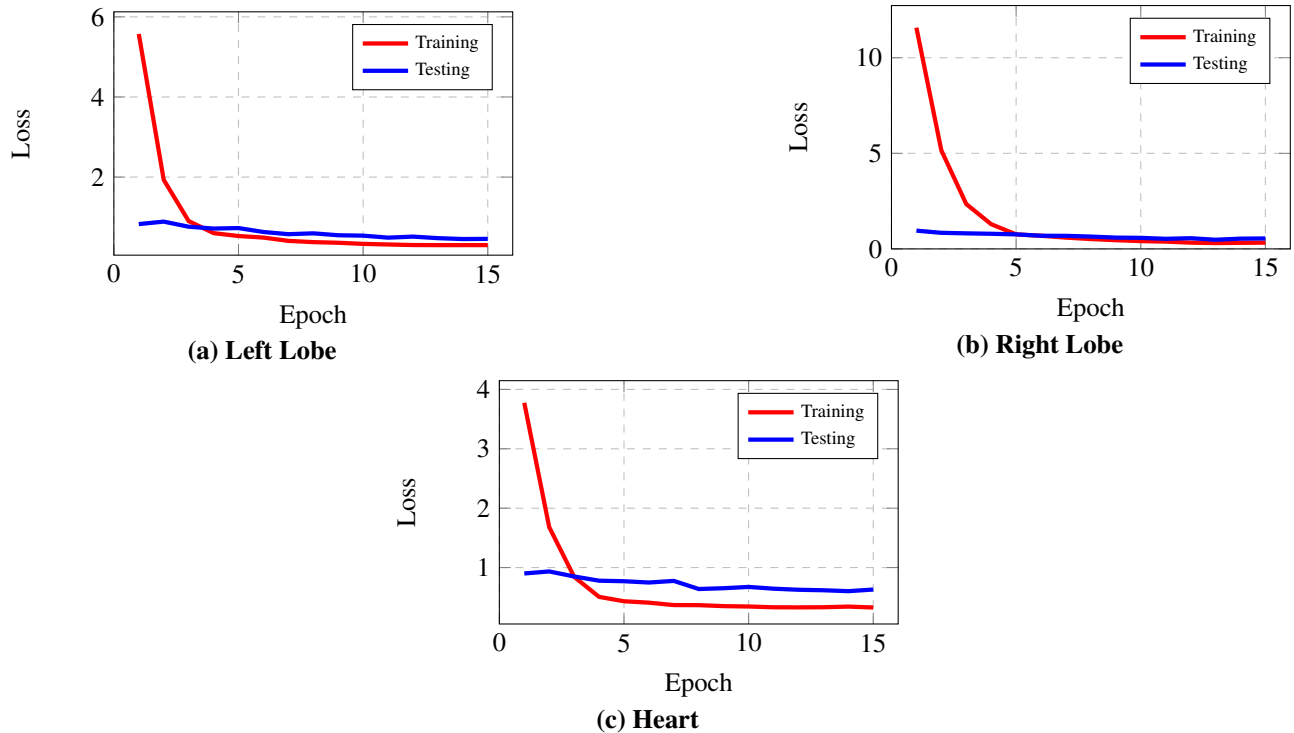


Figure 7: **Loss vs epoch on non-augmented data.** Categorical cross-entropy loss values are calculated after each epoch of training and profiled over 15 epochs for our target organs specifically, (a) the left lung, (b) the right lung, and (c) the heart. The loss value is expected to decrease since after each epoch of training the model performs better classification.

Training and testing sets were formed by partitioning the 422 NSCLC thoracic CT volumes into 300 training volumes and 122 testing volumes. The accuracy obtained over 15 epochs of training for our CNN detectors is profiled in Figure 6. The variation in loss function during training is also profiled and shown in Figure 7.

After partitioning, augmentation was performed to increase the training data from 300 volumes to 10,000 volumes and testing data from 122 to 2500 volumes. Care was taken to assure that all 10,000 training volumes were derived from the original 300 training-exclusive data volumes. Similarly, 2500 testing volumes were derived from 122 original testing-exclusive data volumes. The accuracy and loss profiled over 15 epochs of training with augmented data are shown in Figures 8 and 9 respectively.

Throughout characterization, the number of true positives (target VOIs with correctly classified organs), true negatives (non-target VOIs correctly classified as NONE), false positives (non-target VOIs incorrectly classified as containing an organ), and false negatives (target VOIs incorrectly classified as NONE) were collected. The performance of the organ detector was measured in terms of sensitivity:

$$SN = \frac{TP}{TP + FN} \quad (5)$$

specificity:

$$SP = \frac{TN}{FP + TN} \quad (6)$$

and accuracy:

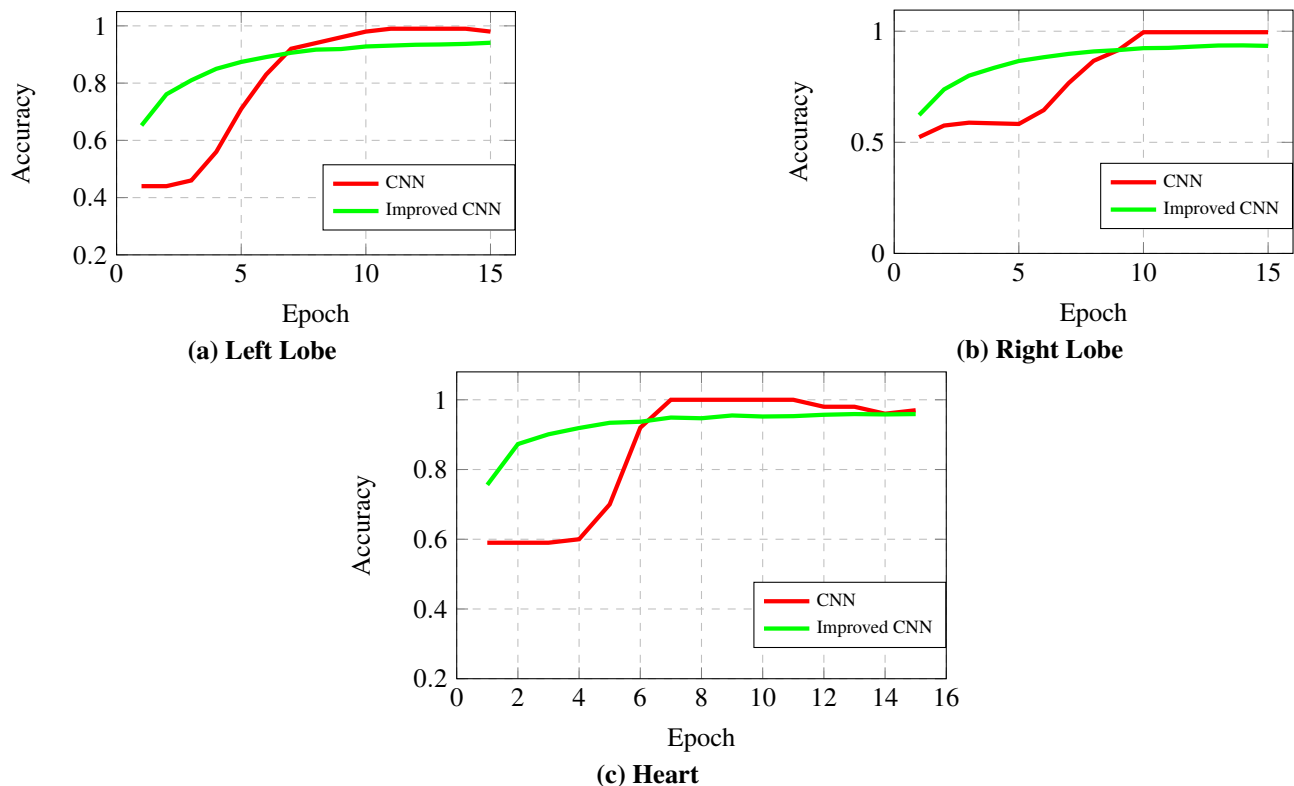


Figure 8: **Training Accuracy vs epoch on augmented data.** The variation of accuracy over 15 epochs of training for (a) the left lung, (b) the right lung, and (c) the heart. The accuracy changes are significant in the first few epochs for both networks. The improved CNN does not exhibit signs of over-fitting unlike the typical CNN which exhibits 100% training accuracy and much lower testing accuracy.

$$AC = \frac{TP + TN}{TP + FN + FP + TN} \quad (7)$$

These figures of merit are represented in the receiver operating characteristic (ROC) curves shown in Figure 10, which illustrate the variation of a detector's sensitivity as a function of its specificity. The area under these ROC curves (i.e. the AUC) is an excellent measure of detector performance and is closer to one for systems performing accurate detection.

Figure 10 shows ROC curves for the various organ detectors both with and without the Bayesian classification layer. These curves demonstrate that the lungs classify with high accuracy and specificity. Additionally, in the low false positive rate region of the heart ( $\sim 1\%$  to  $3\%$ ), the Bayesian model provides a substantial increase in true positive rate, sometimes as high as a 47 percentage point improvement, without sacrificing specificity. This improvement in detector performance is further demonstrated by Table 1, which compares the AUC for each organ detector with and without the Bayesian layer. As shown, a statistically significant improvement is seen for each organ as a result of the Bayesian layer; of these the heart exhibits a substantial AUC improvement, increasing from 0.9076 to 0.9515.

#### 4. CONCLUSIONS AND DISCUSSION

This research concerns the integration of contextual information into a CNN model designed to assign categorically defined organ labels to subvolumes within a thoracic CT image volume. This has been achieved through the addition of a final Bayesian classification layer to the end of the CNN, which leverages contextual evidence to improve the accuracy of organ class assignment to candidate subregions. Specifically, likelihood information modeling the probability of specific

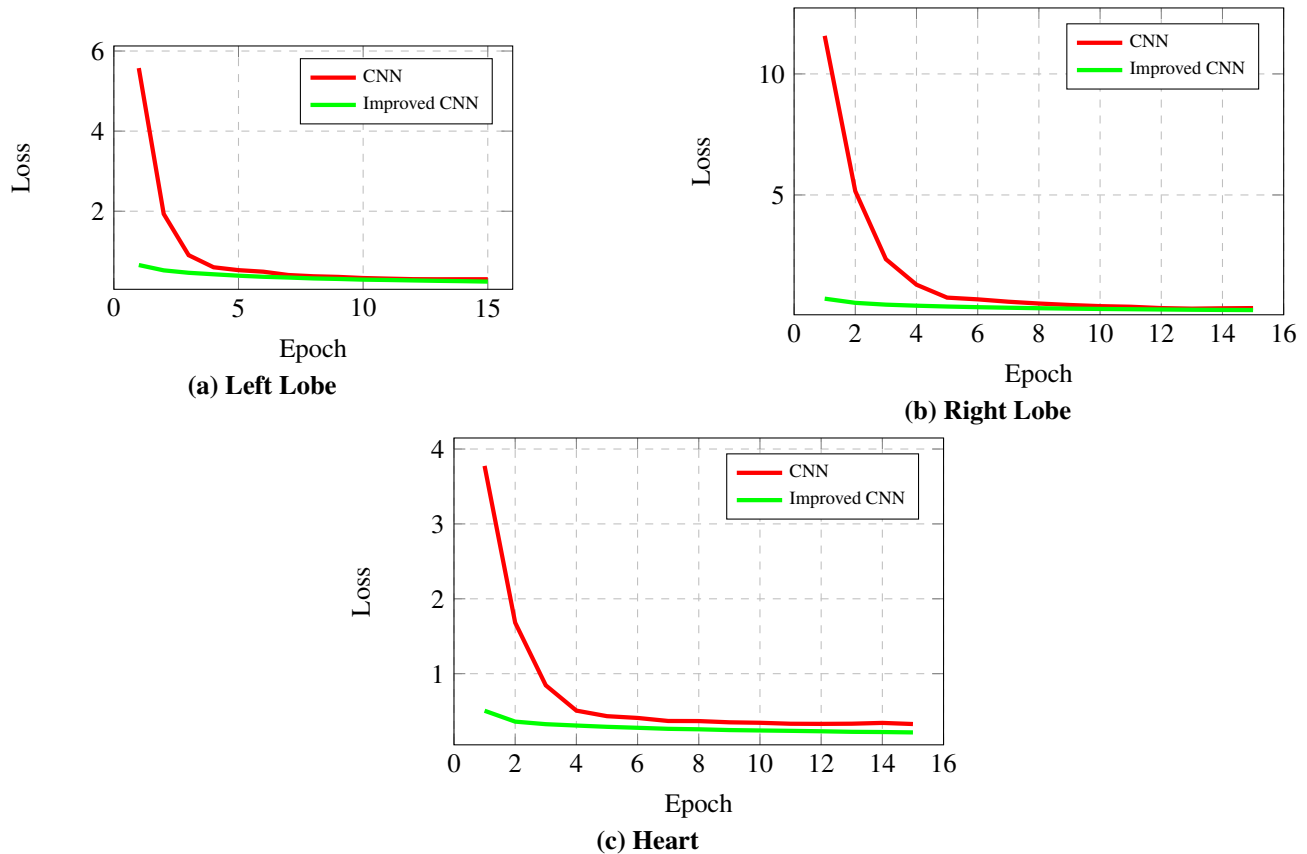
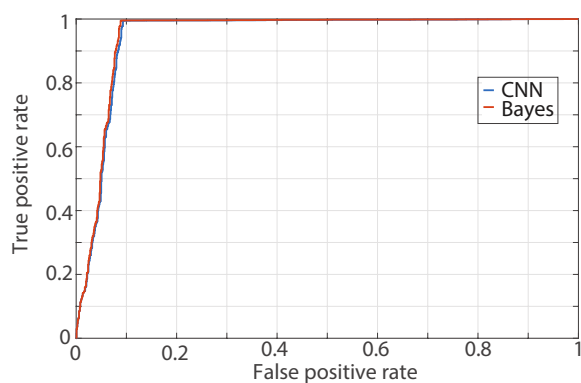
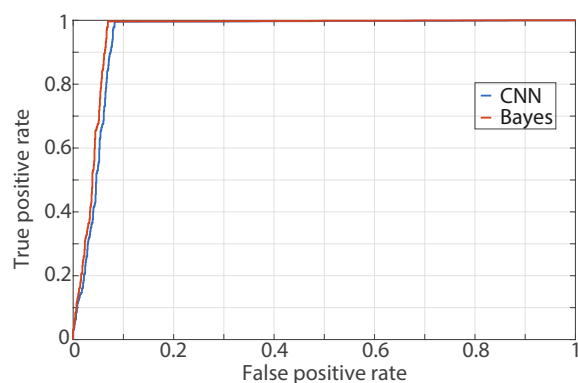


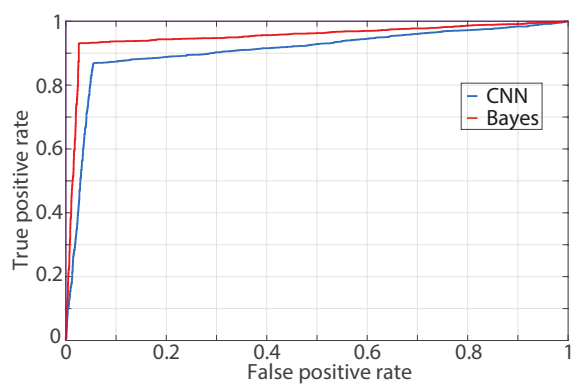
Figure 9: **Loss vs epoch on augmented data.** Categorical cross-entropy loss values are calculated after each epoch of training and profiled over 15 epochs for our target organs specifically, (a) the left lung, (b) the right lung, and (c) the heart. As, expected the loss value decreases after each epoch of training.



(a) **Left Lobe**



(b) **Right Lobe**



(c) **Heart**

Figure 10: **ROC Curves** for the detection of (a) the left lung, (b) the right lung, and (c) the heart. The bayes classifier performs better than the CNN classifier indicated by increase in the AUC. The changes in the curve in towards the lower FPR area indicate more true positives have been detected by bayes than CNN in the low sensitivity region which is relatively a more relevant region to medical applications.

Table 1: Area under ROC for organ detectors before and after employing spatial relationship evidence

| ORGAN           | POST CNN | POST CNN+BAYES |
|-----------------|----------|----------------|
| Left Lung Lobe  | 0.9487   | 0.9514         |
| Right Lung Lobe | 0.9537   | 0.9620         |
| Heart           | 0.9076   | 0.9515         |

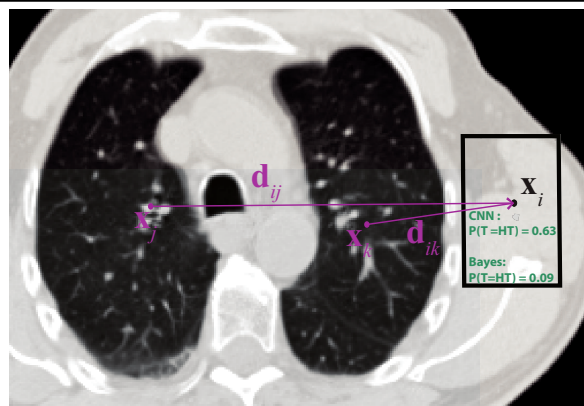


Figure 11: **An axial CT volume slice exhibiting the correction of a poor initial heart detection.** The boxed region is the cross section of a VOI identified by the raw CNN to contain the heart with 64% certainty. Using this as a class prior for a Bayesian output layer with marginal likelihoods describing the relation of the heart with respect to the lungs appropriately decreases this certainty from 64% to 9%.

organ classes existing at spatial locations in relation to one another is used to provide these improved posterior class membership probabilities.

Experiments demonstrating the efficacy of the final Bayesian classification layer in the detection of the left lung, right lung, and heart have shown the addition of the layer to be successful in reducing the false positive rate while increasing the true positive rate. This effect is particularly pronounced in the detection of the heart, which exhibits an increase in the true positive rate as large as 0.47 while maintaining a specificity as high as 0.97. Similar improvements are not seen in the lungs as the standard CNN architecture already exhibits good detection using learned feature extractors alone, operating with an AUC of  $\sim 0.95$  and  $\sim 0.96$  before and after the Bayesian layer respectively.

As demonstrated by Figure 11, it is reasonable to believe that the enriching effect of the Bayesian layer is more pronounced when detecting the heart due to the relative similarity in Hounsfield units possessed by muscle tissue. In other words, such tissues have roughly similar radiodensity, causing them to exhibit poorly pronounced contrast in X-ray CT images ( $\sim 35$ – $55$  H.U.), which may cause more distinctive features to go unlearned by the CNN without a more robust training dataset. However, this case of poor CNN heart classification serves to well exemplify the motivation behind the employment of spatial relationship information in the Bayesian layer, which increases the area under the heart ROC curve shown in Figure 10(c) from 0.9076 to 0.9515. As a result, false hypothesis such as erroneous heart classifications are corrected and weak true heart classifications are reinforced by the employment of spatial proximity evidence.

This success in leveraging spatial anatomic relations among organs within the thorax provides a basis for improving automatic detection of organs in diagnostic medical imaging setups that are non-standard, noisy, or otherwise challenging.

## 5. ACKNOWLEDGEMENTS

The authors thank Pingge Jiang and Micheal Marino for their assistance in labeling thoracic CT volumes. This material is based upon work supported by the National Science Foundation under Grant No. 1553436 and 1642380.

## REFERENCES

- [1] Naidich, D. P., Marshall, C. H., Gribbin, C., Arams, R. S., and McCauley, D. I., "Low-dose ct of the lungs: preliminary observations.," *Radiology* **175**(3), 729–731 (1990).
- [2] Brenner, D. J., "Radiation risks potentially associated with low-dose ct screening of adult smokers for lung cancer 1," *Radiology* **231**(2), 440–445 (2004).
- [3] Poletti, P.-A., Platon, A., Rutschmann, O. T., Schmidlin, F. R., Iselin, C. E., and Becker, C. D., "Low-dose versus standard-dose ct protocol in patients with clinically suspected renal colic," *American Journal of Roentgenology* **188**(4), 927–933 (2007).
- [4] Gervaise, A., Osemont, B., Louis, M., Lecocq, S., Teixeira, P., and Blum, A., "Standard dose versus low-dose abdominal and pelvic ct: comparison between filtered back projection versus adaptive iterative dose reduction 3d," *Diagnostic and interventional imaging* **95**(1), 47–53 (2014).
- [5] Matthews, S., "Imaging pulmonary embolism in pregnancy: what is the most appropriate imaging protocol?," *The British journal of radiology* (2014).
- [6] Prince, S. J., [*Computer vision: models, learning, and inference*], Cambridge University Press (2012).
- [7] Dietterich, T. G., "Ensemble methods in machine learning," in [*International workshop on multiple classifier systems*], 1–15, Springer (2000).
- [8] Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R., "Dropout: a simple way to prevent neural networks from overfitting.," *Journal of Machine Learning Research* **15**(1), 1929–1958 (2014).
- [9] Moore, S. M., Maffitt, D. R., Smith, K. E., Kirby, J. S., Clark, K. W., Freymann, J. B., Vendt, B. A., Tarbox, L. R., and Prior, F. W., "De-identification of medical images with retention of scientific research value," *RadioGraphics* **35**(3), 727–735 (2015).
- [10] Krizhevsky, A., Sutskever, I., and Hinton, G. E., "Imagenet classification with deep convolutional neural networks," in [*Advances in neural information processing systems*], 1097–1105 (2012).
- [11] Masters, T., [*Deep Belief Nets in C++ and CUDA C: Volume 1: Restricted Boltzmann Machines and Supervised Feedforward Networks*], CreateSpace Independent Publishing Platform (2015).
- [12] Ibragimov, B. and Xing, L., "Segmentation of organs-at-risks in head and neck ct images using convolutional neural networks," *Medical Physics* (2016).
- [13] Hasegawa, A., Lo, S.-C. B., Freedman, M. T., and Mun, S. K., "Convolution neural-network-based detection of lung structures," in [*Medical Imaging 1994*], 654–662, International Society for Optics and Photonics (1994).
- [14] Cha, K. H., Hadjiiski, L., Samala, R. K., Chan, H.-P., Caoili, E. M., and Cohan, R. H., "Urinary bladder segmentation in ct urography using deep-learning convolutional neural network and level sets," *Medical physics* **43**(4), 1882–1896 (2016).
- [15] Dahl, G. E., Sainath, T. N., and Hinton, G. E., "Improving deep neural networks for lvcsr using rectified linear units and dropout," in [*Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*], 8609–8613, IEEE (2013).
- [16] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., et al., "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467* (2016).
- [17] Shackleford, J., Kandasamy, N., and Sharp, G., "On developing B-spline registration algorithms for multi-core processors," *Physics in Medicine and Biology* **55**(21), 6329 (2010).
- [18] Shackleford, J., Shusharina, N., Verberg, J., Warmerdam, G., Winey, B., Neuner, M., Steininger, P., Arbisser, A., Golland, P., Lou, Y., et al., "Plastimatch 1.6 – current capabilities and future directions," in [*Proc. MICCAI 2012 Image-Guidance and Multimodal Dose Planning in Radiation Therapy workshop*], (2012).