



Cross-domain Medical Image Translation by Shared Latent Gaussian Mixture Model

Yingying Zhu¹(✉), Youbao Tang¹, Yuxing Tang¹, Daniel C. Elton¹,
Sungwon Lee¹, Perry J. Pickhardt², and Ronald M. Summers¹

¹ Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology
and Imaging Sciences, National Institutes of Health, Clinical Center, Bethesda,
MD 20892, USA

{yingying.zhu,rms}@nih.gov

² School of Medicine and Public Health, University of Wisconsin,
Madison, WI 53706, USA

Abstract. Current deep learning based segmentation models generalize poorly to different domains due to the lack of sufficient labelled image data. An important example in radiology is generalizing from contrast enhanced CT to non-contrast CT. In real-world clinical applications, cross-domain image analysis tools are in high demand since medical images from different domains are generally used to achieve precise diagnoses. For example, contrast enhanced CT at different phases are used to enhance certain pathologies or internal organs. Many existing cross-domain image-to-image translation models show impressive results on large organ segmentation by successfully preserving large structures across domains. However, such models lack the ability to preserve fine structures during the translation process, which is significant for many clinical applications, such as segmenting small calcified plaques in the aorta and pelvic arteries. In order to preserve fine structures during medical image translation, we propose a patch-based model using shared latent variables from a Gaussian mixture. We compare our image translation framework to several state-of-the-art methods on cross-domain image translation and show our model does a better job preserving fine structures. The superior performance of our model is verified by performing two tasks with the translated images - detection and segmentation of aortic plaques and pancreas segmentation. We expect the utility of our framework will extend to other problems beyond segmentation due to the improved quality of the generated images and enhanced ability to preserve small structures.

1 Introduction

Developing deep learning based segmentation models which can generalize to different domains has been in high demand since different type of medical images are usually collected in real clinical practice to achieve a precisely diagnosis. For

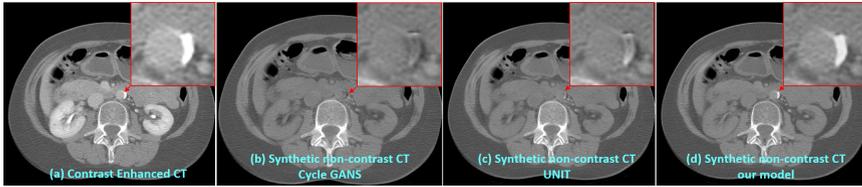


Fig. 1. (a) Real post-contrast CT scan. (b) Synthetic non-contrast CT using CycleGAN [21]. (c) using shared latent variables from a Gaussian distribution [4, 8]. (d) Gaussian mixture model.

example, a patient might have a non-contrast and a contrast-enhanced CT scan generated by injecting an intravenous contrast agent to highlight different internal structures at different time points. For instance, the arteries are enhanced in early the early phase and the kidneys are enhanced in the late phase as the contrast agent is metabolized in the kidneys.

Although many existing works [5, 13–17] have used image-to-image translation techniques to assist in medical image analysis tasks, less work been done to address cross-domain image segmentation due to the lack of sufficient labelled data from different domains. Only a few studies have used synthetic images generated by cross-modality image-to-image translation (e. g., CT and MRI) for cross-domain image (e. g., organ) segmentation [20, 22]. However, image segmentation across CT and MRI scans is of less clinical importance due to its poor performance caused by the large difference between the two modalities. Moreover, CT scans are typically used to scan a large range (fast but low resolution) while MRI scans are often targeted at small regions (slow but high resolution). Paired CT and MRI scans for the same region of the body are rarely collected in clinical practice. Such works usually focus on large organ (e. g., heart) segmentation. However, hardly any work has been done for the small structure segmentation across domains, for example calcified plaque segmentation in the aorta and pelvic arteries under different contrast levels. This is a clinically important problem, since calcified plaque in the arteries is a strong predictor of heart attack [9].

Applying existing image translation models to improve calcified plaque segmentation on different domains is impractical since these model show inconsistent performance for preserving fine/tiny structures after image translation (as shown in Fig. 1, the calcified plaques are blurry and covered by neighboring structures by CycleGAN [21] and UNIT [4, 8]). We hypothesize this is because UNIT assumes a shared latent Gaussian variable across domains and real medical images actually lie in a shared Gaussian mixture model since different internal structures (image patches) lies in different local clusters as shown on the right side of Fig. 1.

In order to address these problems, we proposed a patch-based domain invariant method using shared latent variables from a Gaussian mixture model for image translation [7, 23]. In order to quantitatively evaluate the image

translation performance of our model, we applied it to calcified plaque detection/segmentation and pancreas segmentation on both non-contrast and contrast enhanced CT images. We compared our model to several image translation networks. Experimental results showed that our model performs much better than the baseline (without image translation) and competing methods (e. g., CycleGAN [21] and UNIT [8]) by improving the performance of both tasks. It is worth noting that our model is trained using unpaired images across different domains, which means it can be easily adapted to real-world clinical practice where paired images are relatively rare.

2 Method

2.1 Unsupervised Image-to-Image Translation Networks (UNIT)

Let $x_1 \in \mathcal{X}_1$ and $x_2 \in \mathcal{X}_2$ be two images from non-contrast and contrast enhanced CT, respectively. The image size of x_1, x_2 is 512×512 in our dataset. Liu et al. [8] proposed two sets of variational auto-encoders for the two different domains and translate the images across the two domains via a shared latent space \mathcal{Z} . The shared latent space \mathcal{Z} is conditionally independent from the two domains and is enforced to follow a Gaussian distribution with unit variance. Intuitively, this latent space \mathcal{Z} encodes the underlying morphological structure of objects and is domain invariant. For example, the latent space \mathcal{Z} may depend on the shape of internal organs which is invariant across image domains. They implemented this by sharing the latent space \mathcal{Z} layers of two variational autoencoders. The UNIT model [8] trained using the whole image, which results in loss of detailed structures as shown in Fig. 1(b). Similarly, the CycleGAN method also translates imgs between domains using a shared latent space and cycle consistent loss but can not preserve the detailed structures (see Fig. 1(c)).

2.2 Patch-Based Mixtures Gaussian Image-to-Image Translation

We proposed a patch based method and extracted many small randomly sampled patches from each image. The optimal patch size is determined using testing with the validation set. We used a patch size of 32×32 in all our experiments. We extracted random patches at the same location from x_1, x_2 respectively. We extracted the image features using a pre-trained CNN on these small patches and found that they lie in different local clusters as shown in Fig. 2. Intuitively, image patches from different organs or internal structures will be clustered into different local clusters. Based on these observations, we proposed to model the domain independent shared latent space \mathcal{Z} using a mixture Gaussian model, $\mu_z = [\mu_{z,1}, \dots, \mu_{z,k}]$, $\Sigma_z = [\Sigma_{1,z}, \dots, \Sigma_{z,K}]$ are the mean and variance for different Gaussian components.

$$z \sim \sum_{k=1}^K \pi_k \mathcal{N}(z | \mu_{k,z}, \Sigma_{k,z}), \text{ s.t. } \sum_{k=1}^K \pi_k = 1, \quad (1)$$

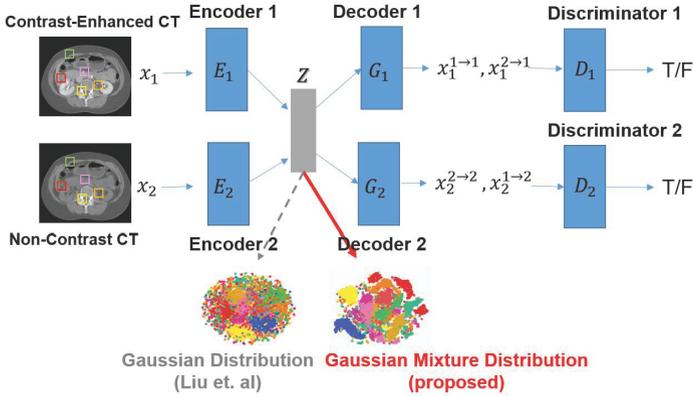


Fig. 2. The framework of our proposed model compared to UNIT [8] model. Our model assumes that the shared latent variable lies in a Gaussian mixed distribution (different patches lies in different local clusters and Liu et. al. [8] assume a single Gaussian distribution.

K is the number of Gaussian components following [2]. It is worth noting that K is determined by the validation dataset on the downstream segmentation/detection task. We follow [8] and use 6 sub-networks: two domain image encoders: E_1, E_2 , two domain image generators G_1, G_2 , and two domain adversarial discriminators D_1, D_2 as shown in Fig. 2. We use VGG-16 as the encoder and a reversed VGG-16 structure as the decoder. We use the variational inference model from *Edward* [18] to solve the parameters in our model. The encoder outputs a set of mean vectors for each Gaussian component: $E_1(x_1, \theta_1), E_2(x_2, \theta_2)$, $\Sigma_1 = [\Sigma_{1,1}, \dots, \Sigma_{1,K}], \Sigma_2 = [\Sigma_{2,1}, \dots, \Sigma_{2,K}]$ are the variance matrix for each Gaussian component. $\Theta_1 = [\theta_{1,1}, \dots, \theta_{1,K}], \Theta_2 = [\theta_{2,1}, \dots, \theta_{2,K}]$ are the weights for using a shared encoder to output the mean for different Gaussian components. The distribution of latent code z given x_1, x_2 are listed as,

$$q_1(z|x_1) \sim \sum_{i=1}^K \pi_k \mathcal{N}(z|E_1(x_1, \theta_{1,k}), \Sigma_{k,1}), \sum_{k=1}^K \pi_k = 1,$$

$$q_2(z|x_2) \sim \sum_{i=1}^K \pi_k \mathcal{N}(z|E_2(x_2, \theta_{2,k}), \Sigma_{k,2}), \sum_{k=1}^K \pi_k = 1.$$

The reconstructed image $\hat{x}_1^{1->1} = G_1(z \sim q_1(z|x_1))$ for two variational autoencoder (VAE) (E_1, G_1) and (E_2, G_2) are $\hat{x}_2^{2->2} = G_2(z \sim q_2(z|x_2))$, we have the VAE loss:

$$\mathcal{L}_{VAE_1}(E_1, G_1, \Theta_1, \Sigma_1, \Sigma_z, \mu_z) = \lambda_1 \text{KL}(q_1(z|x_1)||p(z)) \quad (2)$$

$$- \lambda_2 \mathbb{E}_{z \sim q_1(z|x_1)} [\log p_{G_1}(x_1|z)]$$

$$\mathcal{L}_{VAE_2}(E_2, G_2, \Theta_2, \Sigma_2, \Sigma_z, \mu_z) = \lambda_1 \text{KL}(q_2(z|x_2)||p(z)) \quad (3)$$

$$- \lambda_2 \mathbb{E}_{z \sim q_2(z|x_2)} [\log p_{G_1}(x_2|z)],$$

where λ_1, λ_2 are the parameters controls the weights for the objective terms and the KL divergence terms penalized derivation of the distribution if the latent variable from the prior distribution.

Our model also has two generative adversarial networks: $GAN_1 = \{G_1, D_1\}$, $GAN_2 = \{G_2, D_2\}$. D_1, D_2 are constrained to output true if images are sampled from the first or second domain respectively and output false if the images are generated from G_1, G_2 respectively. We have the following conditional GAN objective functions, which constrain the translated images to resemble images in their respective target domains:

$$\mathcal{L}_{GAN_1}(E_2, G_1, D_1, \Theta_1, \Sigma_1, \Sigma_z, \mu_z) = \lambda_0 \mathbb{E}_{x_1 \sim P_{\mathcal{X}_1}} [\log D_1(x_1)] \quad (4)$$

$$+ \lambda_0 \mathbb{E}_{z \sim q_2(z|x_2)} [\log D_1(G_1(z))]$$

$$\mathcal{L}_{GAN_2}(E_1, G_2, D_2, \Theta_2, \Sigma_2, \Sigma_z, \mu_z) = \lambda_0 \mathbb{E}_{x_2 \sim P_{\mathcal{X}_2}} [\log D_2(x_2)] \quad (5)$$

$$+ \lambda_0 \mathbb{E}_{z \sim q_1(z|x_1)} [\log D_2(G_2(z))],$$

Similarly to the previous equation, the hyperparameter λ_0 balances the impact of the GAN objective function. We also incorporate a cycle consistency constraint to ensure that twice translated images resemble the original image and a KL divergence term which penalizes the latent code from deviating too far from the prior distribution.

$$\mathcal{L}_{CC_2}(E_1, G_1, E_2, G_2, \Theta_1, \Theta_2, \Sigma_1, \Sigma_2, \Sigma_z, \mu_z) \quad (6)$$

$$= \lambda_3 \text{KL}(q_1(z|x_1)||p(z))$$

$$+ \lambda_4 \text{KL}(q_2(z|x_1^{1-\gamma_2})||p(z)) - \lambda_4 \mathbb{E}_{z \sim q_2(z|x_1^{1-\gamma_2})} [\log p_{G_1}(x_1|z)]$$

$$\mathcal{L}_{CC_2}(E_2, G_2, E_1, G_1, \Theta_1, \Theta_2, \Sigma_1, \Sigma_2, \Sigma_z, \mu_z)$$

$$= \lambda_3 \text{KL}(q_2(z|x_2)||p(z)) \quad (7)$$

$$+ \lambda_4 \text{KL}(q_1(z|x_2^{2-\gamma_1})||p(z)) - \lambda_4 \mathbb{E}_{z \sim q_1(z|x_2^{2-\gamma_1})} [\log p_{G_2}(x_2|z)]$$

Combining all the above objective functions, our final objective functions is:

$$\arg \min(E_1, E_2, G_1, G_2, \Theta_1, \Theta_2, \Sigma_1, \Sigma_2, \Sigma_z, \mu_z) \max(D_1, D_2) \quad (8)$$

$$\mathcal{L}_{VAE_1}(E_1, G_1, \Theta_1, \Sigma_1, \Sigma_z, \mu_z) + \mathcal{L}_{VAE_2}(E_2, G_2, \Theta_1, \Sigma_1, \Sigma_z, \mu_z)$$

$$+ \mathcal{L}_{CC_1}(E_1, G_1, E_2, G_2, \Theta_1, \Theta_2, \Sigma_2, \Sigma_z, \mu_z)$$

$$+ \mathcal{L}_{CC_2}(E_1, G_1, E_2, G_2, \Theta_1, \Theta_2, \Sigma_2, \Sigma_z, \mu_z)$$

$$+ \mathcal{L}_{GAN_1}(E_1, G_1, D_1, \Theta_1, \Sigma_1) + \mathcal{L}_{GAN_2}(E_2, G_2, D_2, \Theta_2, \Sigma_2)$$

3 Experiments

Evaluation Tasks. In order to evaluate the image translation model quantitatively, we evaluated the performance of our image translation models on two challenging image segmentation tasks: calcified plaque detection/segmentation and pancreas segmentation on both the non-contrast and contrast enhanced CT scans given labelled CT scans only from one domain. We also compared our model to two recent state-of-the-art image translation models: CycleGAN [21] and UNIT [4, 8], which have been broadly applied to image translation on natural images.

Image Translation Training. For training we used 140 unpaired CT scans (70 non-contrast, 70 contrast enhanced) taken from renal donor patients at the University of Wisconsin Medical Center. We applied our image translation model to generate synthetic contrast enhanced CT scans from the labelled non-contrast CT scans and used them as augmented data to improve our plaque segmentation/detection performance and pancreas segmentation on both contrast enhanced and non-contrast CT scans.

The image translation training dataset was separated into 10 folds and one fold was used as validation data for selecting hyperparameters including the number of Gaussian mixture components K and $\lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4$. The validation dataset was evaluated for downstream plaque detection tasks using the model trained from Sect. 3.1. The optimal value of K is very important for our final image translation results. A small K value can lead to blurring of local image structures and a large K value can add to computation cost and more uncertainty on the output images. We used $K = 25$ based on the validation dataset performance. It is worth noting that the setting of K can vary between datasets. In order to select the best generated image for cross-domain image segmentation task, we also train a quality control network using an independent CT dataset selected from DeepLesion data [19] to remove the synthetic CT images with artifacts.

3.1 Calcified Plaque Detection and Segmentation

Labelling calcified plaques in the CT scans is very time consuming since the plaques are very small, frequently on the order of just a few pixels in CT images. In our experiments, we only have labelled training CT images from low dose non-contrast CT scans. We trained a 2D detection and segmentation model [6] on 75 low dose CT scans which contained a total of 25,200 images (transverse cross sections/slices), including 2,119 with plaques. The training dataset was divided into 10 folds and we used 9 folds as the training dataset and 1 fold as the validation dataset for parameter selection. We shuffled the training and validation dataset and trained 10 2D Mask R-CNN [3] models and applied these models to our independent testing dataset. We report the mean and standard derivation across all 10 models in Table 1. For this work we labelled an additional testing dataset with 30 contrast enhanced CT scans and 30 non-contrast scans

from a different dataset collected at University of Wisconsin Medical Center. It had plaque labeled manually (7/30 of these scans contained aortic plaques, with a total of 53 plaques overall).

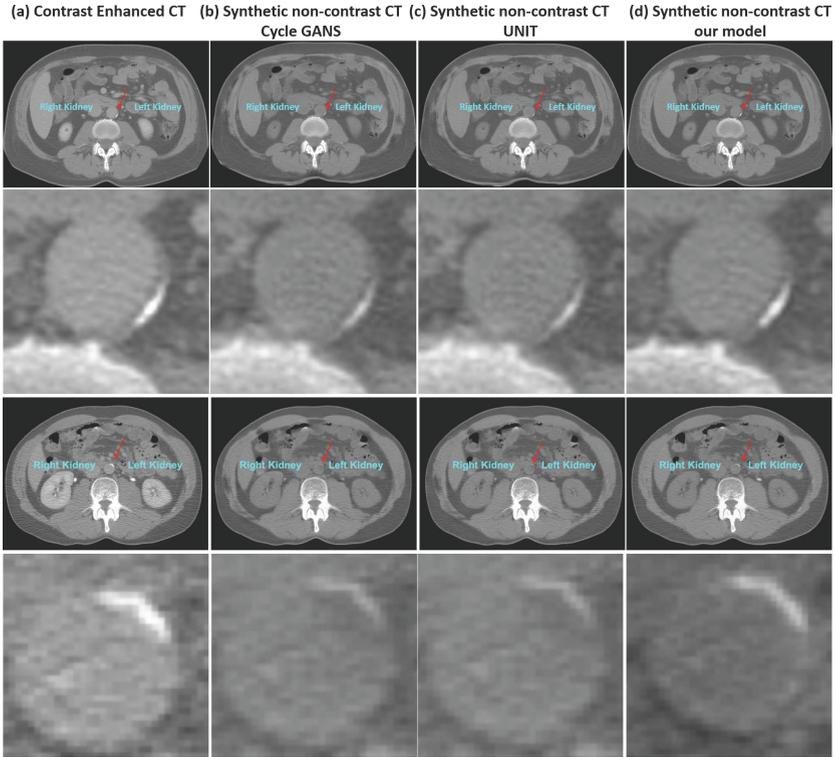


Fig. 3. Visualization of synthetic non-contrast CT scans from contrast enhanced CT (a) scans by (b) CycleGAN (c) UNIT (d) our model. We show two cases of median quality generated images by all competing methods.

We selected an synthetic image which shows the median performance on image translation (as shown in Fig. 3). Figure 3 (a) shows contrast enhanced CT scans in the late phase (the kidneys are enhanced with bright pixel values). There is a very small plaque in this image which was translated into non-contrast CT scans by cycleGAN [21], UNIT [8] and our model. The whole translated images looks very similar by all competing methods, however, the calcified plaque pixel brightness is better preserved by our approach.

Quantitative Results. The calcified plaque detection and segmentation results are shown in Table 1. Our model achieved similar plaque detection and segmentation performance to the real pre-contrast CT scans (precision decreased about $> 1.5\%$, recall decreased about $> 4\%$) and dice coefficients drops about > 0.1 .

Table 1. Plaque detection and segmentation results. The first column gives detection results for the original data without image translation. The 2nd, 3rd, and 4th columns give results for non-contrast (NC) and contrast enhanced (CE) plaque detection after non-contrast to contrast enhanced image translation with different image translation models. SYN= synthetic images.

Method	Baseline	CycleGAN [1]	UNIT [4,8]	Ours
Training	NC	NC & SYN CE	NC & SYN CE	NC & SYN CE
Testing	NC	NC	NC	NC
Precision	78.4 ± 1.82%	79.5 ± 1.75%	80.2 ± 1.68%	80.7 ± 1.65%
Recall	82.4 ± 2.45%	83.1 ± 2.26%	84.7 ± 2.13%	85.2 ± 2.03%
Dice	0.724 ± 0.245	0.733 ± 0.223	0.751 ± 0.218	0.756 ± 0.193
Testing	CE	CE	CE	CE
Precision	48.6 ± 3.52%	61.5 ± 2.87%	64.7 ± 2.64%	78.2 ± 2.58%
Recall	54.3 ± 3.64%	64.3 ± 3.21%	69.8 ± 3.05%	81.2 ± 2.87%
Dice	0.452 ± 0.251	0.534 ± 0.236	0.566 ± 0.198	0.676 ± 0.176

The detection and segmentation model trained on synthetic images generated from UNIT [8] and CycleGAN [21], by contrast, shows a > 15% drop in precision, a > 13% drop in recall and drop > 0.18 in Dice coefficients caused by loss of fine structures.

3.2 Pancreas Segmentation

Pancreas segmentation is very important for the diagnosis of pancreas cancer and surgical planning. Pancreas segmentation is challenging since the pancreas is very small compared to other internal organs and has large variance in its shape and orientation. Most existing pancreas segmentation approaches focus on segmenting pancreas only in contrast enhanced CT where the pancreas structures are more enhanced and have clearer boundaries. Current public pancreas segmentation data are only labelled on contrast enhanced CT. We combined two public contrast enhanced CT datasets for pancreas segmentation. The first one includes 82 labelled contrast enhanced CT scans from the Cancer Imaging Archive database and second one has 281 contrast enhanced CT scans from the Medical Segmentation Decathlon database [10,11].

We use 10-fold cross validation and report the mean and standard derivation across the 10 folds. In order to improve pancreas segmentation on non-contrast CT images, we generated non-contrast CT from these contrast enhanced CT and used them to train a cross-domain 3D segmentation model. We use the multiple scale 3D Unet model proposed in [12] and compared with a 3D U-Net trained using synthetic non-contrast CT scans generated the different image translation models (CycleGAN [21], UNIT [4,8] and our model). We use 24 non-contrast CT scans annotated by an expert radiologist as the non-contrast pancreas segmentation testing dataset.

Table 2. Pancreas segmentation results with false positives pixel numbers and Dice scores on contrast enhanced CT and non-contrast CT. The first column show the baseline model which is only trained using contrast enhanced CT scans. Th second, third, and last column show the results trained by contrast enhanced CT and synthetic non-contrast CT generated by CycleGAN [21], UNIT [4, 8] and our method respectively. NC = non-contrast CT. CE = contrast enhanced CT. SYN = synthetic.

Method	Baseline	CycleGAN [21]	UNIT [4, 8]	Ours
Training	CE	CE & SYN NC	CE & SYN NC	CE & SYN NC
Testing	CE	CE	CE	CE
Precision	85.7 ± 1.82%	87.2 ± 1.62%	88.1 ± 1.53%	89.5 ± 1.44%
Recall	86.8 ± 2.1%	88.7 ± 1.97%	89.5 ± 1.83%	90.7 ± 1.68%
Dice	0.728 ± 0.173	0.728 ± 0.154	0.731 ± 0.142	0.734 ± 0.136
Testing	NC	NC	NC	NC
Precision	78.1 ± 2.83%	82.8 ± 2.67%	83.2 ± 2.45%	84.7 ± 2.25%
Recall	81.5 ± 3.0%	84.3 ± 2.81%	86.2 ± 2.75%	87.2 ± 2.56%
Dice	0.642 ± 0.183	0.684 ± 0.172	0.697 ± 0.163	0.725 ± 0.153

Quantitative Results. Table 2 shows the quantitative results of cross domain pancreas segmentation results on both contrast enhanced CT and non-contrast CT using different image translation methods (CycleGAN [21], UNIT [8] and our model). As shown in the top part of the table, the pancreas segmentation results on contrast enhanced CT scans are similar for all competing methods. Adding synthetic non-contrast CT scans in the training can also improve the pancreas segmentation on contrast enhanced CT, and our method shows slight improvement compared to all other methods. For pancreas segmentation on non-contrast CT, adding synthetic non-contrast CT images in the training can significantly improve the segmentation Dice score and reduce false positive pixels. For example, CycleGAN and UNIT show an average improvement of Dice score $> 0.04/ > 0.05$ and reduction in false positive pixels $> 4000/ > 4500$ compared to the baseline model. Our model shows the best performance and achieves a > 0.08 improvement in Dice score and a reduction of > 5000 false positive pixels on average compared to the baseline model.

4 Conclusion

In this work, we proposed an image translation model using shared latent variables from a Gaussian mixture distribution to preserve fine structures on medical images. We applied our method to two challenging medical imaging segmentation tasks: cross domain (non-contrast and contrast enhanced CT) calcified plaque detection/segmentation and pancreas segmentation. We demonstrated that our model can translate the medical images across different domains with better preservation of fine structures compared to two state-of-the-art image translation models for natural images. In the future work, we will explore the application

of this model to translating medical images across multiple domains, for example contrast enhanced CT scans at different phases from non-contrast CT scans. Possible applications of this method are generating synthetic images to reduce radiation dose or creating 100% contrast enhanced CT scans from 10% dose contrast enhanced CT scans to reduce the dose of intravenous contrast agent to be used on patients.

Acknowledgments. This research was supported in part by the Intramural Research Program of the National Institutes of Health Clinical Center. We thank NVIDIA for GPU card donations.

References

1. Clark, K., et al.: The cancer imaging archive (TCIA): maintaining and operating a public information repository. *J. Digit. Imaging* **26**(6), 1045–1057 (2013)
2. Dilokthanakul, N., et al.: Deep unsupervised clustering with gaussian mixture variational autoencoders. arXiv e-prints [arXiv:1611.02648](https://arxiv.org/abs/1611.02648) (2016)
3. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask R-CNN. In: International Conference on Computer Vision (ICCV), pp. 2980–2988 (2017)
4. Huang, X., Liu, M.Y., Belongie, S., Kautz, J.: Multimodal unsupervised image-to-image translation. In: European Conference on Computer Vision (ECCV), pp. 172–189 (2018)
5. Jin, D., Xu, Z., Tang, Y., Harrison, A.P., Mollura, D.J.: CT-realistic lung nodule simulation from 3D conditional generative adversarial networks for robust lung segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 732–740 (2018)
6. Liu, J., Yao, J., Bagheri, M., Sandfort, V., Summers, R.M.: A semi-supervised CNN learning method with pseudo-class labels for atherosclerotic vascular calcification detection. In: International Symposium on Biomedical Imaging (ISBI), pp. 780–783 (2019)
7. Liu, L., Nie, F., Wiliem, A., Li, Z., Zhang, T., Lovell, B.C.: Multi-modal joint clustering with application for unsupervised attribute discovery. *IEEE Trans. Image Process.* **27**(9), 4345–4356 (2018)
8. Liu, M.Y., Breuel, T., Kautz, J.: Unsupervised image-to-image translation networks. In: Advances in Neural Information Processing Systems (NeurIPS), pp. 700–708 (2017)
9. Parab, S.Y., Patil, V.P., Shetmahajan, M., Kanaparthi, A.: Coronary artery calcification on chest computed tomography scan-anaesthetic implications. *Indian J. Anaesth.* **63**(8), 663 (2019)
10. Roth, H., et al.: DeepOrgan: multi-level deep convolutional networks for automated pancreas segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9349, pp. 556–564. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24553-9_68
11. Simpson, A.L., et al.: A large annotated medical image dataset for the development and evaluation of segmentation algorithms. arXiv e-prints [arXiv:1902.09063](https://arxiv.org/abs/1902.09063) (2019)
12. Sriram, S.A., Paul, A., Zhu, Y., Sandfort, V., Pickhardt, P.J., Summers, R.: Multilevel U-Net for pancreas segmentation from non-contrast CT scans through domain adaptation. In: Hahn, H.K., Mazurowski, M.A. (eds.) Medical Imaging 2020: Computer-Aided Diagnosis. SPIE, March 2020

13. Tang, Y.B., Oh, S., Tang, Y.X., Xiao, J., Summers, R.M.: CT-realistic data augmentation using generative adversarial network for robust lymph node segmentation. In: *Medical Imaging: Computer-Aided Diagnosis*, vol. 10950, p. 109503V (2019)
14. Tang, Y., et al.: CT image enhancement using stacked generative adversarial networks and transfer learning for lesion segmentation improvement. In: *International Workshop on Machine Learning in Medical Imaging*, pp. 46–54 (2018)
15. Tang, Y., Tang, Y., Xiao, J., Summers, R.M.: XLSor: a robust and accurate lung segmentor on chest x-rays using criss-cross attention and customized radiorealistic abnormalities generation. In: *International Conference on Medical Imaging with Deep Learning*, pp. 457–467 (2019)
16. Tang, Y.X., Tang, Y.B., Han, M., Xiao, J., Summers, R.M.: Abnormal chest X-ray identification with generative adversarial one-class classifier. In: *International Symposium on Biomedical Imaging*, pp. 1358–1361 (2019)
17. Tang, Y., Tang, Y., Sandfort, V., Xiao, J., Summers, R.M.: TUNA-Net: task-oriented unsupervised adversarial network for disease recognition in cross-domain chest X-rays. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 431–440 (2019)
18. Tran, D., Kucukelbir, A., Dieng, A.B., Rudolph, M., Liang, D., Blei, D.M.: Edward: a library for probabilistic modeling, inference, and criticism. arXiv e-prints [arXiv:1610.09787](https://arxiv.org/abs/1610.09787) (2016)
19. Yan, K., Wang, X., Lu, L., Summers, R.M.: DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning. *J. Med. Imaging* **5**(3), 1–11 (2018)
20. Zhou, T., Fu, H., Chen, G., Shen, J., Shen, J., Shao, L.: Hi-net: Hybrid-fusion network for multi-modal MR image synthesis. *IEEE Trans. Med. Imaging* **PP**(99), 1 (2020)
21. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *International Conference on Computer Vision (ICCV)*, pp. 2223–2232 (2017)
22. Zhu, Y., Elton, D.C., Lee, S., Pickhardt, P., Summers, R.: Image translation by latent union of subspaces for cross-domain plaque detection. arXiv e-prints [arXiv:2005.11384](https://arxiv.org/abs/2005.11384) (2020)
23. Zhu, Y., Huang, D., la Torre Frade, F.D., Lucey, S.: Complex non-rigid motion 3D reconstruction by union of subspaces. In: *Proceedings of CVPR*, June 2014