

Construction of a Multi-modal Model of Pancreatic Tumors by Integration of MRI and Pathological Images using Conditional Cycle α -GAN

Tomoshige Shimomura¹, Mauricio Kugler¹, Tatsuya Yokota¹, Chika Iwamoto², Kenoki Ohuchida²,
Makoto Hashizume², and Hidekata Hontani¹

¹Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, Aichi, 466-8555, Japan

t.shimomura.525@nitech.jp, {kugler.mauricio, t.yokota, hontani}@nitech.ac.jp

²Kyushu University, 744 Motoooka Nishi-ku, Fukuoka, 819-0395, Japan

Abstract

In this study, we constructed the multi-modal model of a pancreas tumor by learning the correspondence between MRI voxels and pathology image patches with a conditional cycle α -GAN. In this framework, we constructs encoder-decoder networks that translate MRI voxels and pathology image patches each other, and two discriminators for both modalities. When a voxel in a pancreas tumor region in an MRI image is selected, this model can generate various corresponded pathology image patches non-invasively, and vice versa. We made a training dataset by registering between an MRI image and a 3D pathology image and trained our multi-modal model. Using trained model, anyone can observe the change of pathology image with respect to the MRI values, and we found the behaviors are closely related to the growing process of pancreas tumor.

1. Introduction

Modeling the correlations between pathology images and MRI images has been investigated. The former images can be used for definitive diagnosis and the latter images can be obtained non-invasively. Models that represent the correlations of these images would improve the confidence of diagnosis and can be used for predicting the histopathological status from the corresponding MRI images.

In this study, we construct a multi-modal model that predicts pathology image patches in the condition of the voxel in an MRI image of a pancreas tumor of a KPC mouse. For constructing such a model, we made a training dataset by registering between the tumor region in the MRI image and the 3D pathology image reconstructed from a spatial of the 2D pathology images [1] as shown in Figure 1. We created a model with a generator trained on a conditional cycle

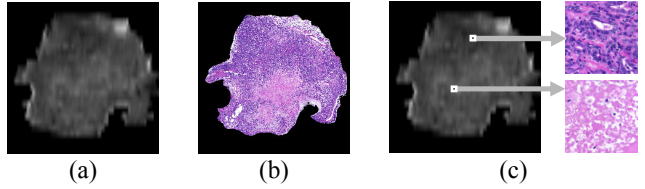


Figure 1. Registration between the tumor region in the MRI image and the 3D pathology image. (a): An MRI image of the tumor region. (b): The corresponding pathology image. (c): The correspondence between the MRI voxels and the pathology image patches.

α -GAN (cc- α -GAN) extended from an α -GAN [2].

2. Method

We construct a multi-modal model using a cc- α -GAN combining α -GANs [2] and cycle GANs [3]. This GANs consist of encoder, generator, discriminator and code discriminator with parameters η , θ , ϕ and ω , respectively. We define the loss function for the encoder and the generator:

$$\begin{aligned} \mathcal{L}(\theta, \eta) = & \mathbb{E}_{\mathbf{x} \sim p(\mathbf{x})} \mathbb{E}_{\mathbf{y}, \mathbf{z} \sim q_{\eta}(\mathbf{y}, \mathbf{z} | \mathbf{x})} \left[\right. \\ & \lambda_1 \|\mathbf{x} - \mathcal{G}_{\theta}(\mathbf{y}, \mathbf{z})\|_1 + \lambda_2 \|\mathbf{y}_{\text{real}} - \mathbf{y}\|_2^2 \\ & \left. - \log \frac{\mathcal{D}_{\phi}(\mathcal{G}_{\theta}(\mathbf{y}, \mathbf{z}))}{1 - \mathcal{D}_{\phi}(\mathcal{G}_{\theta}(\mathbf{y}, \mathbf{z}))} - \log \frac{\mathcal{C}_{\omega}(\mathbf{y}, \mathbf{z})}{1 - \mathcal{C}_{\omega}(\mathbf{y}, \mathbf{z})} \right] \\ & + \mathbb{E}_{\mathbf{y}, \mathbf{z} \sim p(\mathbf{y}, \mathbf{z})} \mathbb{E}_{\hat{\mathbf{y}}, \hat{\mathbf{z}} \sim q_{\eta}(\mathbf{y}, \mathbf{z} | \mathcal{G}_{\theta}(\mathbf{y}, \mathbf{z}))} \left[\right. \\ & \lambda_2 \|\mathbf{y} - \hat{\mathbf{y}}\|_2^2 + \lambda_3 \|\mathbf{z} - \hat{\mathbf{z}}\|_2^2 \\ & \left. - \log \frac{\mathcal{D}_{\phi}(\mathcal{G}_{\theta}(\hat{\mathbf{y}}, \hat{\mathbf{z}}))}{1 - \mathcal{D}_{\phi}(\mathcal{G}_{\theta}(\hat{\mathbf{y}}, \hat{\mathbf{z}}))} - \log \frac{\mathcal{C}_{\omega}(\hat{\mathbf{y}}, \hat{\mathbf{z}})}{1 - \mathcal{C}_{\omega}(\hat{\mathbf{y}}, \hat{\mathbf{z}})} \right] \end{aligned} \quad (1)$$

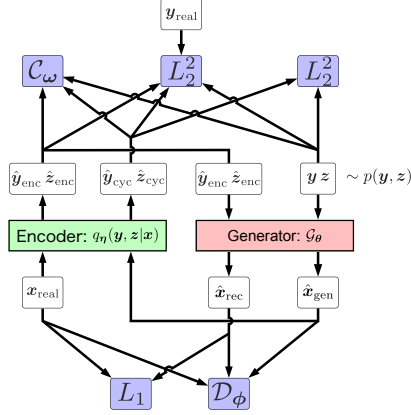


Figure 2. cc- α -GAN framework

$\mathbf{x} \in \mathbb{R}^{96 \times 96}$	Pathology image patch
$\mathbf{y} \in \mathbb{R}^2$	MRI voxel intensity and 2D Euclidean distance from the contour of the pancreas tumor region
$\mathbf{z} \in \mathbb{R}^{1024}$	Latent variable sampled from normal distribution
$q_\eta : \mathbf{x} \rightarrow (\mathbf{y}, \mathbf{z})$	Encoder
$\mathcal{G}_\theta : (\mathbf{y}, \mathbf{z}) \rightarrow \mathbf{x}$	Generator
\mathcal{D}_ϕ	Discriminator for \mathbf{x}
\mathcal{C}_ω	Code discriminator for (\mathbf{y}, \mathbf{z})

Table 1. The description of each symbol

where the coefficients λ_1 , λ_2 and λ_3 are scale parameters, and \mathbf{y}_{real} is the conditional variable corresponding to $\mathbf{x} \sim p(\mathbf{x})$. The discriminator and the code discriminator are trained with cross entropy loss as standard GANs.

We use a training dataset made from a MRI image and corresponding 3D pathology image of a pancreas tumor. Each sample consists of pairs of a pathology image patch as \mathbf{x} and the voxel information of corresponding MRI image as \mathbf{y} , which has the voxel intensity y_1 and the 2D Euclidean distance y_2 from the contour of the pancreas tumor region in the MRI image. The latent variable \mathbf{z} follows the standard normal distribution with 1024 dimensions. Figure 2 shows the framework of a cc- α -GAN and Table 1 summarizes the description of each symbol.

3. Experiments and Results

We trained our networks with the dataset composed of about one hundred thousand samples with augmentations for 20 epochs using Adam optimizer and a batch size of 100. The learning rate was 5×10^{-4} and the coefficients λ_1 , λ_2 and λ_3 were set to 50, 10 and 10, respectively.

Using trained generator, we generate pathology image patches from conditional variables y_1 and y_2 sampled from

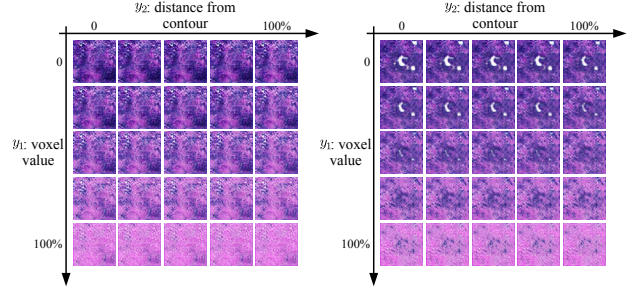


Figure 3. Image matrices of generated pathology images from conditional variables. Each image matrix has own latent variable. The patches from same latent variable have similar texture and same conditional variable have similar style.

every 25% of the cumulative frequency of the each histogram and same latent variable which is randomly sampled. Figure 3 shows the result of using two different latent variables. We found the patches from same latent variable have similar texture and same conditional variable have similar style. In addition, if y_1 and y_2 increase, generated patches become pink like necrosis portions. The tendency is similar to the characteristic of the actual images that a voxel in MRI images with high intensity corresponds to necrosis portion because of moisture arising in necrosis. As a result, we succeeded in pseudo reproducing the process of growing to necrosis of tumors.

4. Conclusion

We constructed a multi-modal model of pancreas tumors that can predict pathology image patches from a voxel in the corresponding MRI image non-invasively using a cc- α -GAN. Our model succeeded in pseudo reproducing the process of growing to necrosis of tumors. The future works are improving accuracy and developing the evaluation method.

References

- [1] M. Kugler, Y. Goto, Y. Tamura, N. Kawamura, H. Kobayashi, T. Yokota, C. Iwamoto, K. Ohuchida, M. Hashizume, A. Shimizu, and H. Hontani. Robust 3d image reconstruction of pancreatic cancer tumors from histopathological images with different stains and its quantitative performance evaluation. *International Journal of Computer Assisted Radiology and Surgery*, Jul 2019.
- [2] M. Rosca, B. Lakshminarayanan, D. Warde-Farley, and S. Mohamed. Variational Approaches for Auto-Encoding Generative Adversarial Networks. *arXiv e-prints*, page arXiv:1706.04987, June 2017.
- [3] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2242–2251, 2017.