



Construction of a Generative Model of H&E Stained Pathology Images of Pancreas Tumors Conditioned by a Voxel Value of MRI Image

Tomoshige Shimomura^{1(✉)}, Kugler Mauricio¹, Tatsuya Yokota¹,
Chika Iwamoto², Kenoki Ohuchida², Makoto Hashizume²,
and Hidekata Hontani¹

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

simomura@iu.nitech.ac.jp, {kugler.mauricio,t.yokota,hontani}@nitech.ac.jp

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

Abstract. In this paper, we propose a method for constructing a multi-scale model of pancreas tumor of a KrasLSL.G12D/+; p53R172H/+; PdxCre^{tg}/+ (KPC) mouse that is a genetically engineered mouse model of pancreas tumor. The model represents the correlation between the value at each voxel in the MRI image of the tumor and the pathology image patches that are observed at each portion corresponds to the location of the voxel in the MRI image. The model is represented by a cascade of image generators trained by a Laplacian Pyramid of Generative Adversarial Network (LAPGAN). When some voxel in a pancreas tumor region in an MRI image is selected, the cascade of generators outputs patches of the pathology images that can be observed at the location corresponds to the selected voxel. We trained the generators by using an MRI image and a 3D pathology image, the latter was first reconstructed from a spatial series of the 2D pathology images and was then registered to the MRI image.

1 Introduction

Modelling the correlations between pathology images and MRI images has been investigated (e.g. [9]). 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 histopathological status from the corresponding MRI images. In this study, we construct a non-parametric model that represents the correlation between the voxel value of an MRI image and corresponding histopathological images of a pancreas tumor of a KPC mouse. This model plays an important role in a system that acquires various kinds of information currently obtained from histopathological images with probability information from MR images.

For constructing such a model, we employ a conditional Laplacian Pyramid of Generative Adversarial Network (LAPGAN) [1].

A Generative Adversarial Network (GAN) [4] can construct a sufficiently representative latent model of target images, while simultaneously learning a generator and a discriminator: The generator can create sample images that are intended to come from the same distribution with the training data and the discriminator examines samples to determine whether they are real or fake. The latent model of target images can be represented by a manifold, from which the generator can create the fake images by sampling [10]. In many cases, a noise signal, \mathbf{z} , is input to the generator and the noise signal, \mathbf{z} , can be used as the local coordinate system on the manifold [2]. Generators learned by conditional GANs take not only noise signals but also other condition signals as inputs, where the condition signals restrict the output sample images. The condition signals also can be used as the local coordinate system on the manifold. When the input values of the condition signals are fixed, the generator creates fake images that correspond to a sub-manifold restricted by the conditions. The manifolds with the local coordinate systems can represent the correlations between the target images and the condition signals [10]. In this study, we construct a generator that takes noise signals and a voxel value of an MRI image as input and outputs corresponding pathology image patches sampled from the sub-manifold determined by the MRI voxel value.

The spatial resolution of an MRI image and that of a microscope pathology image are largely different. Each single voxel of an MRI image corresponds to a large image patch of the pathology image. We hence assume that each voxel value correlates with low-resolution features of the pathology images and employ a LAPGAN, which can generate a cascade of image generators, each of which creates a sample fake image that represents the difference between a high-resolution image and the corresponding low-resolution input image [1].

2 Method

2.1 Outline of the Proposed Method

Figure 1 shows the outline of the construction of a multi-scale pancreas tumor model from an MRI image and the corresponding pathology images of a KPC mouse. The MRI image of a whole body of a KPC mouse was captured just before the pancreas tumor was extracted. A 3D pathology image of the tumor was reconstructed from a spatial series of 2D microscope images of the tumor and the 3D pathology image was non-rigidly registered to the tumor region in the MRI image in order to obtain a set of training data for the LAPGAN, in which each datum is a pair of the voxel value of the MRI image and the corresponding image patch in the microscope image. Applying conditional LAPGAN to the training data, we construct a cascade of generators, which can generate a pathology patch image that corresponds to the input voxel value of the MRI image.

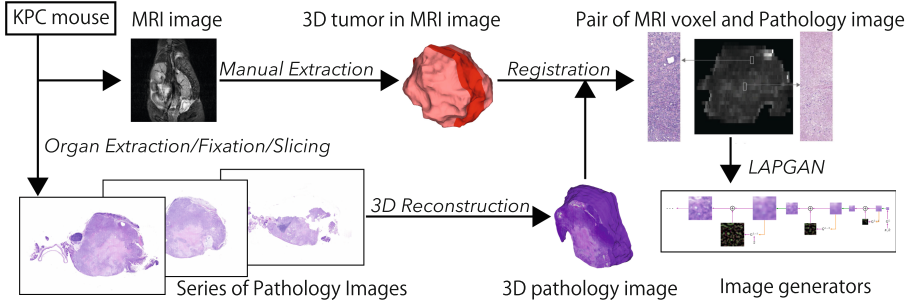


Fig. 1. Construction of a multiscale model of pancreas tumor using an MRI image and the corresponding pathology images

2.2 Images Used in the Experiments

The training images for the conditional LAPGAN was obtained as follows. An MRI image of the whole body of the KPC mouse was captured just before the organs including the whole part of the tumor was extracted. The spatial resolution of the MRI image was $0.1536 \text{ mm} \times 0.1536 \text{ mm} \times 0.5 \text{ mm}$. The tumor was spherical and its diameter was about 2 cm. Two MRI images of the extracted organs were obtained before and after the organs were formaline-fixed. Registering the two MRI images, we found that the organs around the tumor shrunk but the tumor itself did not deform in the formaline fixation.

The extracted organs were paraffin-embedded after the formaline-fixation. We first cut the paraffin block into five small blocks, of which the thickness was about 5 mm, and one of the small blocks, which contained the center portion of the tumor, was sliced into a spatial series of about 800 thin sections. The thickness of the section was set to $4 \mu\text{m}$. The number of the sections obtained from the $5 \text{ mm} = 5000 \mu\text{m}$ thick block was less than $1250 = 5000/4$ because of the loss generated by the slicing. We dyed the thin sections by the Hematoxylin and Eosin (H&E) stain. The microscopy images of these stained sections were then captured with the spatial resolution, $0.22 \mu\text{m} \times 0.22 \mu\text{m}$.

2.3 Reconstruction of a 3D Microscope Image

For the 3D reconstruction of the microscope image, we employed the non-rigid registration method proposed in [6]. Let the given 2D microscope images be denoted by $I_1(\mathbf{x}), I_2(\mathbf{x}), \dots, I_M(\mathbf{x})$, where M is the total number of the 2D microscope images and $\mathbf{x} = [x, y]^T$ denotes the 2D image coordinates. It is assumed that the given images are roughly aligned, for example, by a rigid registration method. Let the deformation mapping computed for $I_i(\mathbf{x})$ be denoted by ϕ_i ($i = 1, 2, \dots, M$). Then, the 3D image, $J(x, y, z)$, is reconstructed as $J(x, y, z = i) = I_i(\phi_i^{-1} \circ \mathbf{x})$. The mapping, ϕ_i , is computed from a set of landmarks located in $I_i(\cdot)$. Let \mathbf{p}_i^j denote the 2D image coordinates of the j -th landmark ($j = 1, 2, \dots, N$) in the i -th image, $I_i(\cdot)$ and let the coordinates of

the *destination* of the j -th landmark in $I_i(\cdot)$, to which the landmark should be mapped by ϕ_i , be denoted by \mathbf{q}_i^j . Then, with the regularization with respect to the deformation rigidity, the mapping is obtained by solving a minimization problem: $\phi_i = \arg \min_{\phi} \sum_j \|\mathbf{q}_i^j - \phi \circ \mathbf{p}_i^j\|^2$. The method we employed [6] determines the destination, \mathbf{q}_i^j , of each landmark by smoothing the *trajectory* of the j -th landmark in the 3D image space.

It should be noted that each of the mappings, ϕ_i , is determined not by referring to only consecutive two images but by referring to all the given images.

2.4 Registration Between MRI Image and Pathology Image

The tumor region in the MRI image and the tumor region in the reconstructed 3D microscope image were registered. A mutual information based non-rigid registration method [8] was employed. Assuming that the deformation of the tumor in the 3D microscope image mainly occurred when each thin section of the tumor specimen was placed on the slide glass, we restricted the movement of the control points to a plane where the z -coordinate is constant: Let the mapping to be computed be denoted by ψ and let $\mathbf{X}' = \psi \circ \mathbf{X}$, where the three-vectors, $\mathbf{X}' = [x', y', z']^\top$ and $\mathbf{X} = [x, y, z]^\top$, denote the 3D coordinates in the reconstructed microscope image, $J(\cdot)$. We computed ψ that maximizes the mutual information under a condition that $z' = \psi \circ z = z$ is satisfied. The mapping, ψ , then keeps the plane, $z = i$, corresponding to each microscope image, $I_i(\cdot)$, flat.

As the result of the registration described above, each voxel in the MRI image is corresponded to a specific region in the 3D microscope image as shown in Fig. 2.

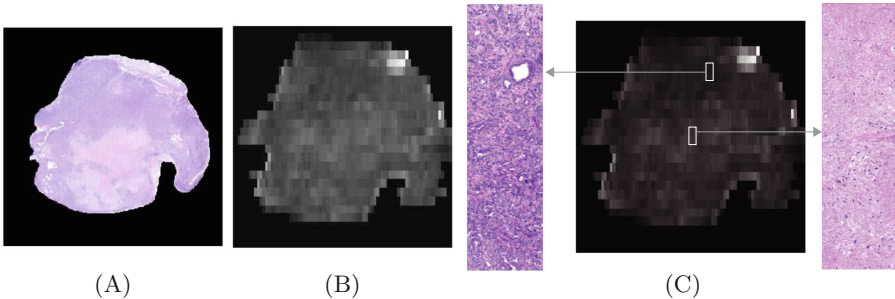


Fig. 2. Registration between the tumor region in the MRI image and the reconstructed 3D pathology image. (A): A pathology image of the tumor region. The pale pink portion includes necrosis region. (B): The corresponding MRI slice image. As shown in (C), the registration makes the correspondence between each voxel in the MRI image and a portion in the pathology image. (Color figure online)

2.5 Construction of Training Image Data for LAPGAN

Let the index of the voxel in the MRI image be denoted by $m \in \mathbb{N}$ and let the 3D region in the reconstructed 3D microscope image, $J(\psi^{-1} \circ \mathbf{X})$, that corresponds to the m -th voxel of the MRI image be denoted by Γ_m . Let the portion in the plane, $z = i$, included in Γ_m be denoted by Γ_{mi} . The region, Γ_{mi} , in the deformed i -th microscope image, $J(x', y', z' = i)$, corresponds to the m -th voxel in the MRI image. Let the value of the m -th voxel of the MRI image be denoted by v_m and let a set of 256×256 image patches included in Γ_{mi} be denoted by $\{\mathcal{I}_{mis}(\mathbf{x}) | s = 1, 2, \dots, S_{mi}\}$, where S_{mi} denotes the number of patches sampled from Γ_{mi} . We sample the patches only from the H&E stained images and augmented the patches by applying random rotation to \mathcal{I}_{mis} for increasing the number of training data. Then, we obtain a set of training data,

$$\mathcal{D} = \{(v_m, \mathcal{I}_{mis}) | m, i, s \in \mathbb{N}\}, \quad (1)$$

in which each datum is a pair of the voxel value of the MRI image and the corresponding patch in H&E stained microscope images.

For avoiding the mode collapse, we quantize the MRI voxel value, v_m , by applying a K -means clustering method so that the variety of the data that have same condition is increased. Let the number of the clusters be denoted by K and let the clusters of voxel values be denoted by \mathcal{C}_k ($k = 1, 2, \dots, K$). Then, we can obtain K sets, \mathcal{D}_k , of training image patches from \mathcal{D} such that

$$\mathcal{D}_k = \{\mathcal{I}_{mis} | v_m \in \mathcal{C}_k; m, i, s \in \mathbb{N}\}, \quad (2)$$

where the image patches included in \mathcal{D}_k correspond to the MRI voxel values that are included in \mathcal{C}_k . We use the index, k , as the condition for the LAPGAN.

The LAPGAN constructs a series of image generative models within a Laplacian pyramid framework. In the Laplacian pyramid framework, an image is represented in a coarse-to-fine fashion, that is, by a series of band-passed images plus a low-frequency residual. Let D_{\downarrow} denote a downsampling operation with a factor, two (2): When the size of an input image, I , is $W \times W$, then $D_{\downarrow} \circ I$ is a new image of size $W/2 \times W/2$. Following the paper [1], we first built a Gaussian pyramid, $g(\mathcal{I}_{mis})$ from each image patch, \mathcal{I}_{mis} , such that $g(\mathcal{I}_{mis}) = [\mathcal{I}_{mis}^0, \mathcal{I}_{mis}^1, \dots, \mathcal{I}_{mis}^L]$, where

$$\mathcal{I}_{mis}^{l+1} = D_{\downarrow} \circ G_{\sigma} \circ \mathcal{I}_{mis}^l, \quad (3)$$

G_{σ} denotes a Gaussian smoothing with the variance, σ^2 , $\mathcal{I}_{mis}^0 = \mathcal{I}_{mis}$, L denotes the number of the levels in the pyramid, and $l = 0, 1, \dots, L - 1$ denotes the level. From the Gaussian pyramid, $g(\mathcal{I}_{mis})$, we then constructed the series of the band-passed images, \mathcal{B}_{mis}^l , which can be computed from \mathcal{I}_{mis}^l .

The LAPGAN we implemented constructs a series of $L + 1$ image generators: The l -th generator ($l = 0, 1, \dots, L - 1$) generates the band-passed image, \mathcal{B}_{mis}^l , from \mathcal{I}_{mis}^{l+1} and the last (the L -th) generator can generate the lowest-frequency image residual, \mathcal{I}_{mis}^L , from a Gaussian noise image. The cascade of the $L + 1$ generators, in the descending order of l , can generate the original image, \mathcal{I}_{mis} .

The l -th generator ($l = 0, 1, \dots, L - 1$) is constructed from sets of the training data, $\{\mathcal{D}_k^l | k = 1, 2, \dots, K\}$, where $\mathcal{D}_k^l = \{(\mathcal{I}_{mis}^{l+1}, \mathcal{B}_{mis}^l) | v_m \in \mathcal{C}_k, m, i, s \in \mathbb{N}\}$. The dataset, \mathcal{D}_k^l , consists of the pairs of the downsampled image of the low-frequency residual, \mathcal{I}_{mis}^{l+1} , and the band-passed image, \mathcal{B}_{mis}^l , both are obtained from the microscope images that correspond to the voxel value, $v_m \in \mathcal{C}_k$, of the MRI image. The last (the L -th) image generator is constructed from the sets of the images, $\{\mathcal{D}_{mis}^L | k = 1, 2, \dots, K\}$, where $\mathcal{D}_k^L = \{\mathcal{I}_{mis}^L | v_m \in \mathcal{C}_k, m, i, s \in \mathbb{N}\}$. The last generator does not need the band-passed images for the training.

2.6 Conditional LAPGAN

We constructed $L + 1$ image generators, G^0, G^1, \dots, G^L , by using the LAPGAN. The inputs of the last generator, G^L , are the Gaussian noise image, \mathbf{z} , and the index, k , of the class of the corresponding MRI voxel value. The output is a residual image of the lowest-frequency, \mathcal{I}^L , which is indistinguishable from the training images, $\mathcal{I}_{mis}^L \in \bar{\mathcal{D}}_k^L$. One can generate variety of such the indistinguishable images by changing the input Gaussian noise images. The inputs of the other generators, G^l ($l = 0, 1, \dots, L - 1$) are the lower-frequency image, \mathcal{I}^{l+1} , the Gaussian noise image, \mathbf{z} , and the index, k . The output of G^l is a band-passed image, \mathcal{B}^l , that can generate the higher-frequency image, \mathcal{I}^l , from the input lower-frequency image, \mathcal{I}^{l+1} as $\mathcal{I}^l = \mathcal{B}^l + O_{sm} \circ U_{\uparrow} \circ \mathcal{I}^{l+1}$, where the resultant image, \mathcal{I}^l , is indistinguishable from the training images, $\mathcal{I}_{mis}^l \in \mathcal{D}_k^l$. Changing the input Gaussian noise image, one can generate variety of images that are indistinguishable from the training data for the discriminator.

Let the discriminators corresponding to G^l be denoted by D^l . Given the dataset, $\{\bar{\mathcal{D}}_k^l | k = 1, 2, \dots, K\}$, the LAPGAN constructs the L -th generator, G^L , and discriminator, D^L , by solving the problem, $\min_G \max_D F^L(G, D)$, where (some superscripts, L , are abbreviated)

$$\begin{aligned} F^L(G, D) &= \mathbb{E}_{\mathcal{I}, k \sim P_{\text{data}}(\mathcal{I}, k)} [\log D(\mathcal{I}, k)] \\ &+ \mathbb{E}_{\mathbf{z} \sim P_{\text{noise}}(\mathbf{z}), k \sim P_{\text{data}}(k)} [\log(1 - D(G(\mathbf{z}, k), k))]. \end{aligned} \quad (4)$$

Given the dataset, $\{\mathcal{D}_k^l | k = 1, 2, \dots, K\}$, the LAPGAN constructs the l -th generator, G^l for $l = 0, 1, \dots, L - 1$, by solving the problem, $\min_G \max_D F^l(G, D)$, where (again, some superscripts are abbreviated):

$$\begin{aligned} F^l(G, D) &= \mathbb{E}_{\mathcal{B}, k \sim P_{\text{data}}(\mathcal{B}, k)} [\log D(\mathcal{B}, k)] \\ &+ \mathbb{E}_{\mathcal{I}, k \sim P_{\text{data}}(\mathcal{I}, k), \mathbf{z} \sim P_{\text{noise}}(\mathbf{z})} [\log(1 - D(G(\mathcal{I}, \mathbf{z}, k), k))]. \end{aligned} \quad (5)$$

We employed CNNs for the generators and discriminators. We initialized the networks by using the method proposed in [3] and employed Adam [7] and the batch normalization [5] for the stochastic optimization.

3 Results

Setting $K = 4$, we divided voxel values of the MRI image into four (4) clusters. Examples of the pathology images, \mathcal{I}_{mis} , included in \mathcal{D}_1 (the brightest portion)

and \mathcal{D}_4 (the darkest portion) are shown in Fig. 3. The dataset included about one million image patches. The distribution of the patch patterns are different among the clusters, \mathcal{D}_k . Setting the number of the cascade $L = 3$, we constructed $L + 1$ generators. The cascade of the generators, $\mathbf{G}^L, \mathbf{G}^{L-1}, \dots, \mathbf{G}^0$, can output a fake pathology image from a given Gaussian noise, \mathbf{z} , and the condition, $k \in [1, K]$. Figure 4 shows examples of the randomly chosen generated images. Corresponding to the bright voxel values of the MRI image, the cascade of the generators sampled patches similar to those in the necrosis portion with a higher probability.

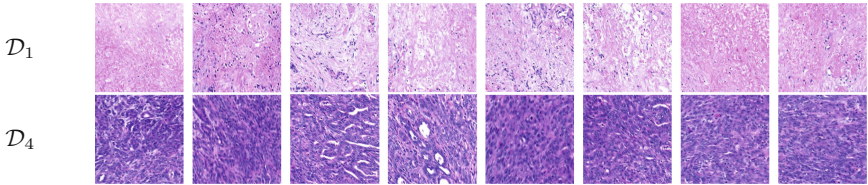


Fig. 3. Examples of the training pathology images, \mathcal{I}_{mis} , included in \mathcal{D}_1 and \mathcal{D}_4

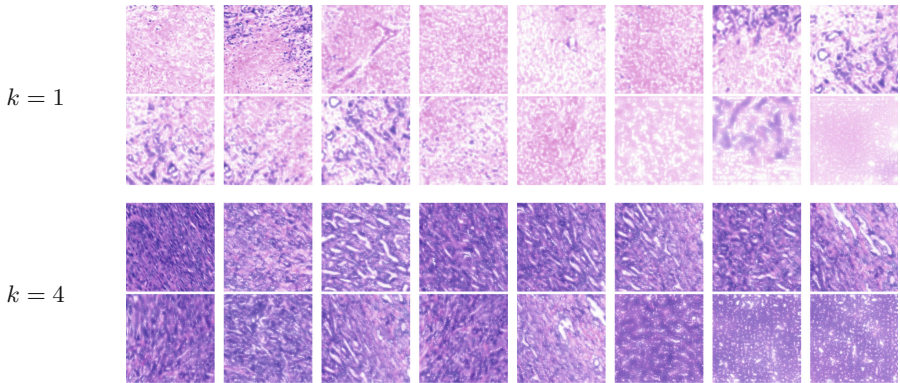


Fig. 4. Examples of the fake pathology images generated by the cascade of the generators for the condition $k = 1$ and $k = 4$.

4 Discussion and Conclusion

We constructed a multiscale model of pancreas tumor that can generate a H&E stained pathology image from a voxel value of the corresponding MRI image. For obtaining a set of training images, we first reconstructed a 3D pathology

image of the pancreas tumor and then registered it to the tumor region in the MRI image. For the construction of the multiscale model, we employed the conditional LAPGAN. The resultant generators output pathology image patches that look like those in the necrosis region with a high probability when the brighter voxel value of the MRI image is input. We constructed the model from a partial region of only one pancreas tumor observed from one KPC mouse. The future works include to annotate histopathological/genetic information to each portion of the histopathology image and to construct a model that can derive the histopathological/genetic information with confidence from each voxel of a given MR image.

References

1. Denton, E.L., Chintala, S., Fergus, R., et al.: Deep generative image models using a Laplacian pyramid of adversarial networks. In: *Advances in Neural Information Processing Systems*, pp. 1486–1494 (2015)
2. Gauthier, J.: Conditional generative adversarial nets for convolutional face generation. *Class Project for Stanford CS231N: Convolutional Neural Networks for Visual Recognition, Winter semester 2014(5)*, 2 (2014)
3. Glorot, X., Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pp. 249–256 (2010)
4. Goodfellow, I., et al.: Generative adversarial nets. In: *Advances in Neural Information Processing Systems*, pp. 2672–2680 (2014)
5. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. In: *International Conference on Machine Learning*, pp. 448–456 (2015)
6. Kawamura, N., et al.: Landmark-based reconstruction of 3D smooth structures from serial histological sections. In: *Proceedings of SPIE*, vol. 10581, pp. 1–7. *International Society for Optics and Photonics* (2018)
7. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. *arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980)* (2014)
8. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imaging* **18**(8), 712–721 (1999)
9. Shinmoto, H., et al.: Small renal cell carcinoma: MRI with pathologic correlation. *J. Magn. Reson. Imaging* **8**(3), 690–694 (1998)
10. Zhang, Z., Song, Y., Qi, H.: Age progression/regression by conditional adversarial autoencoder. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2 (2017)