2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

A generative adversarial network-based CT image standardization model for predicting progression-free survival of lung cancer

Qingxia Wu, Wenhui Huang, Shuo Wang, He Yu, Liushu Wang, Zhangjie Wu, Yongbei Zhu, Zhenyu Liu, He Ma, Jie Tian, *Fellow, IEEE*

***Abstract*— Progression-free survival (PFS) prediction using computed tomography (CT) images is important for treatment planning in lung cancer. However, the generalization ability of current analysis methods is usually affected by the scanning parameters of CT images, such as slice thickness and reconstruction kernel. In this paper, we proposed a generative adversarial network (GAN)-based model to convert heterogenous CT images into standardized CT images with uniform slice thickness and reconstruction kernel to increase the generalization of the predictive model. This model was trained in 173 patients with multiple CT sequences including both thin/thick voxel-spacing and sharp/soft reconstruction kernel. Afterward, we built a 3D-CNN model to predict the individualized 1-year PFS of lung cancer using the standardized CT images in 281 patients. Finally, we evaluated the predictive model by 5-fold cross-validation and the mean area under the receiver operating characteristic curve (AUC). After transforming to the heterogenous CT images into the uniform thin-spacing and sharp kernel CT images, the AUC value of the 3D-CNN model improved from 0.614 to 0.686. Furthermore, this model can stratify the patients into high-risk and low-risk groups, where patients in these two groups showed significant difference in PFS ($P < 0.001$).**

I. INTRODUCTION

Lung cancer remains the deadliest cancer worldwide with only 10% $\sim$ 20% 5-year survival rate. According to the American Cancer Society's report, the lung cancer death rate has declined by almost one-half from 2014 to 2018 [1]. One of the contributors is the great progress in treatment regimens, such as molecularly targeted therapy. The molecular-targeted agents have shown potential in prolonging patients' progression-free survival (PFS). However, some patients emerged drug resistance after a period of targeted cancer therapy [2]. Therefore, early assessment of PFS with regard to targeted therapy is crucial for treatment adaptations.

This work was supported by the Ministry of Science and Technology of China (2017YFA0205200), the National Natural Science Foundation of China (62027901, 81930053, 82001913, and 91859203), and China Postdoctoral Science Foundation (2019TQ0019 and 2020M670101).

+Q. Wu (89757elizabeth@gmail.com) is with the United Imaging Research Institute of Intelligent Imaging, Beijing, China. * is the corresponding author. + are co-first authors.

+W. Huang (18027110926@163.com), H. Ma, and *J. Tian (tian@ieee.org) are with the College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, Liaoning, China.

S. Wang, and L. Wang are with the Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine and Engineering, Beihang University, Beijing, China.

H. Yu is with the Department of Respiratory and Critical Care Medicine, West China Hospital, Sichuan University, Chengdu, China.

Z. Liu. and is with the CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China.

In clinical practice, computed tomography (CT) is widely used for lung cancer diagnosis and treatment monitor. Previous studies have shown that CT texture features are associated with PFS of lung cancer [3]. However, various CT protocols have a great impact on the CT images and limit the further generalization of related research [4]. In terms of CT protocols, slice thickness and reconstruction kernel are very important, since they can affect the spatial resolution and noise pattern of the CT images.

Recently, Generative Adversarial Networks (GANs) have shown great potential in transforming medial images to solve some clinical problems, such as using synthetic images to improve liver lesion classification [5] and to enhance liver tumor detection [6]. Inspired by the related works, we aimed to use CycleGAN to generate CT images with uniform slice thickness and reconstruction kernel. Afterward, we proposed a 3D-CNN model to predict 1-year PFS of lung cancer after receiving targeted therapy. Finally, we compared the performance improvement of the model using uniform synthetic CT images over the model using non-uniform original CT images. Furthermore, we assessed the prognostic ability of our model regarding PFS by the Kaplan-Meier method.

In this paper, Section II presents our method to generate uniform CT images and the comparison between our method with interpolation in predicting 1-year PFS of lung cancer patients. Section III focuses on experimental details and results, Section IV shows the conclusion.

II. METHOD

Our method included: 1) CycleGAN-based image transformation, and 2) 3D-CNN-based 1-year PFS prediction. The image transformation phase aimed at standardizing the heterogeneous CT images into CT images of uniform slice spacing and reconstruction kernel. We trained two CycleGAN, one was used to transform 5-mm soft kernel images to 1-mm sharp kernel images, and the other was used to transform 5-mm sharp kernel images to 1-mm sharp kernel images. Afterward, we used uniformized 1-mm sharp kernel images (synthetic images + original images) to predict the 1-year PFS of lung cancer by 3D-CNN model. The workflow is shown in Fig.1.

A. Uniform CT images synthesis

In order to fit the following 3D-CNN model, we used bicubic interpolation to standardize all CT images into 240 × 360 × 48 before image generation. After that, we built a CycleGAN to convert CT images from the 5-mm sharp/soft

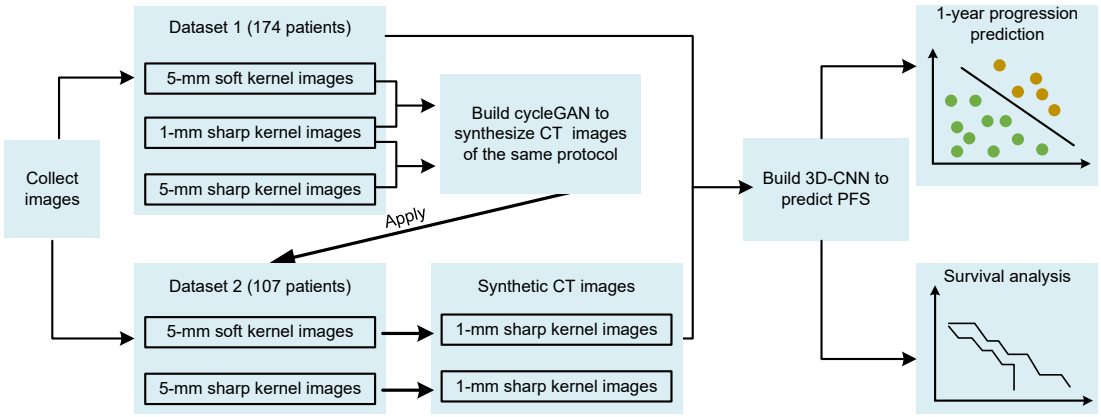

Fig. 1. The workflow of the study.

kernel images to 1-mm sharp kernel images. The architecture is shown in Fig.2. The generator was a ResNet-like architecture. The first three convolutional layers included 64, 128, and 256 filters with stride 7, 3, and 3, respectively. Between each convolutional layer, we used instance normalization layers with LeakyReLU activation function. Then the network was connected with nine residual blocks. Each residual block included two convolutional layers and instance normalization layers with LeakyReLU activation function. In the last convolutional layer, we used tanh activation function to restrict the output image values between [-1,1].

In the discriminator, the first four convolutional layers including 64, 128, 256, and 512 filters were used to learn the image features, and the last convolutional layer with 1 filter was used to calculate the output. Between each convolutional layer, we used instance normalization layers with LeakyReLU activation function. We combined every three adjacent image slices to form a 3-channel input size (240 × 360 × 3) to train the model.

B. Progression-free survival analysis

After training the CycleGAN model in Dataset 1, we applied the model to Dataset 2 to transform 5-mm sharp/soft kernel images to 1-mm and sharp kernel CT images. Then we combined the original 1-mm sharp kernel images from Dataset 1 with the synthetic 1-mm sharp kernel images from Dataset 2 to predict 1-year PFS by building a 3D-CNN model. The architecture of the model was similar to 3D-ResNet18. We added a convolutional layer in the branch of the second to the fourth residual blocks. The input size was 240 × 360 × 48, and other detail parameters are shown in Table 1. In Table I, N means the number of filters, k means kernel size, and s means stride size.

To avoid over-fitting, we augmented the training samples on the fly during the training process by rotating a random angle between 0 and 10, translating a random shift between -15 and 15 voxels in each axis, flipping randomly.

TABLE I
PARAMETERS OF THE 3D-CNN MODEL

Layers	Output size	Parameters
Convolution	64 @ 120 × 180 × 24	64N, k7s2
Max pooling	64 @ 60 × 90 × 12	window2, s2
Residual block 1	64 @ 60 × 90 × 24	Conv1: 64N, k3s1 Conv2: 64N, k3s1, concatenation Conv3: 64N, k3s1 Conv4: 64N, k3s1, concatenation
Residual block 2	128 @ 30 × 45 × 6	Conv5: 128N, k3s2 Conv6: 128N, k3s1, concatenation Conv7: 128N, k3s2, concatenation Conv8: 128N, k3s1 Conv9: 128N, k3s1, concatenation
Residual block 3	256 @ 15 × 23 × 3	Conv10: 256N, k3s2 Conv11: 256N, k3s1, concatenation Conv12: 256N, k3s2, concatenation Conv13: 256N, k3s1 Conv14: 256N, k3s1, concatenation
Residual block 4	512 @ 8 × 12 × 2	Conv15: 512N, k3s2 Conv16: 512N, k3s1, concatenation Conv17: 512N, k3s2, concatenation Conv18: 512N, k3s1 Conv19: 512N, k3s1, concatenation
Global average pooling	512	
Fully connected	1	Sigmoid, L2 regularization

III. EXPERIMENTS AND RESULTS

A. Dataset

A total of 281 advanced lung cancer patients from West China Hospital of Sichuan University from 2009 to 2012 were enrolled in this study under the approval of the institutional review boards. All patients were treated with molecularly targeted agents, such as Gefitinib, Erlotinib, and Icotinib. After targeted cancer therapy, all patients were followed up for at least 1 year. The ground truth in this study is 1-year PFS. The PFS was measured from the time of therapy to the time of progression or recurrence. At the

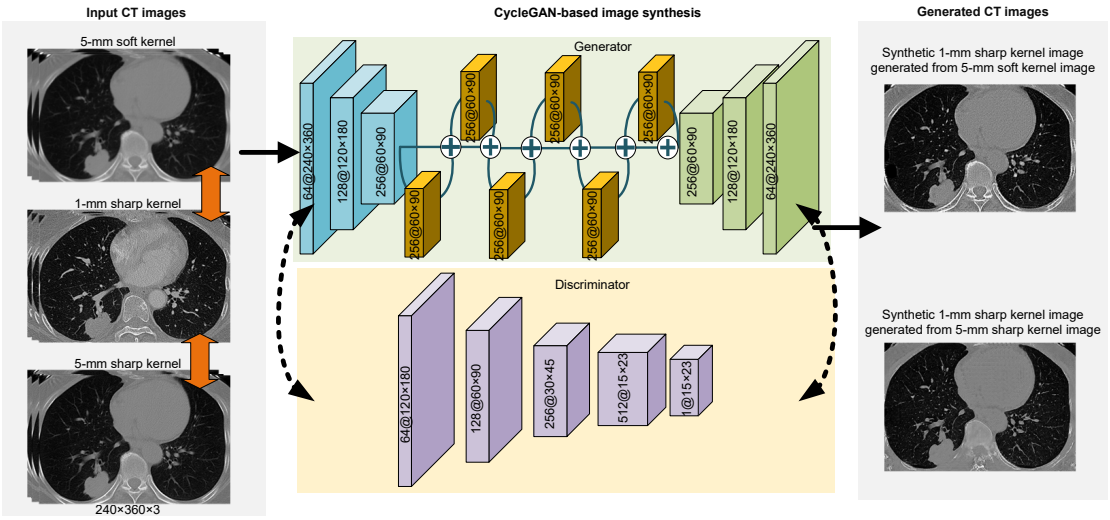

Fig. 2. The architecture of proposed CycleGAN model.

last follow-up, 132 patients recurred within one year, 149 patients did not recur within one year.

Every patient underwent chest CT scans by using Siemens SOMATOM Definition AS scanner. The acquisition parameters were as followings: tube voltage = 120 kV, tube current = 200 mAs, rotation time = 0.7 s, matrix = 512 × 512, and intervals = 0.75 mm. The reconstruction kernels had the following two types :1) sharp kernel (B60f and B80f), and 2) soft kernel (B30f and B31f). The slice thickness was 1 mm or 5 mm. Note that all thin-spacing images (1-mm slice thickness) only had the sharp reconstruction kernel.

In Dataset 1, 174 patients had both 1-mm sharp kernel images, 5-mm sharp kernel images and 5-mm soft kernel images. In Dataset 2, 108 patients only had 5-mm sharp or soft kernel images.

B. Data preprocessing

Firstly, we clipped the voxel intensity to [-1000, 400], and then normalized into [-1, 1). Afterward, we built an FPN model with DenseNet121 backbone to get a lung volume mask. The training details can be found in our previous studies [7]. After training for 40 epochs, we applied the DenseNet121-FPN model for lung segmentation in CT images. The cubic bounding box of the segmented lung area was cropped and resized into 240 × 360 × 48 voxel size for further analysis.

All networks were implemented by Python 3.7 and Keras 2.2 with TensorFlow 2.3 backend and trained on the machine with NVIDIA TITAN X GPU.

C. Synthesize uniform CT images

We used Dataset 1, which contained 174 patients with 8352 training images to train the CycleGAN to convert 5-mm soft/sharp kernel images to 1-mm sharp kernel images. We used mean-squared error (MSE) as the adversarial loss, mean absolute error (MAE) as the cycle consistency loss and identity loss, Adam optimizer with a fix learning rate of 0.0002, and batch size of 5 for training. To measure the image quality, peak signal-to-noise ratio (PSNR) and mean-squared error (MSE) were used. The evaluation results between the synthetic and original 1-mm sharp kernel images are shown in Table II, and the sample results of synthetic images are shown in Fig.3. The results exhibited that using 5-mm soft kernel images had a slightly better image quality than using 5-mm sharp kernel images to convert 1-mm sharp kernel images.

TABLE II
EVALUATION SIMILARITY BETWEEN GENERATED AND REAL 1-MM SHARP KERNEL IMAGES

Images	MSE		PSNR	
	mean	sd	mean	sd
Synthetic 1-mm sharp kernel images generated from 5-mm soft kernel images	0.03	0.01	19.87	1.44
Synthetic 1-mm sharp kernel images generated from 5-mm sharp kernel images	0.04	0.01	18.89	1.71

D. Progression-free survival analysis

We combined Dataset 1 and 2 (281 patients) to train the 3D-CNN model to predict 1-year PFS. In Dataset 1, we used original 1-mm sharp kernel images, and in Dataset 2, we compared the models using the original 5-mm soft/sharp kernel images with the synthetic 1-mm sharp kernel images. We used 5-fold cross-validation to evaluate the model performance and calculated the mean area under the receiver operating characteristic curve (AUC) and mean accuracy (ACC) to assess the PFS prediction of the model. We used kaiming initialization with normal distribution ('he normal') method to initialize the parameters in the model. Adam optimizer

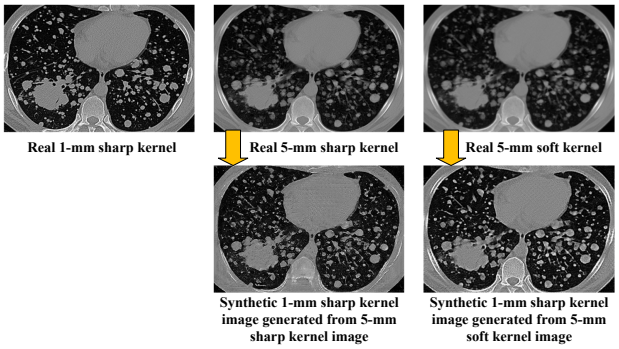

Fig. 3. The sample results of the image transformation.

with an initial learning rate of 0.0001 and exponential decay was used to train the model for up to 50 epochs. Due to the GPU memory limitation, the batch size was set to 2. The results are shown in Table III. The model using the original heterogenous CT images (1-mm sharp kernel images mixed with original 5-mm soft/sharp kernel images) yielded AUC around 0.61 and ACC around 0.60. When transforming the heterogenous CT images into uniform 1-mm sharp kernel images, the model performance was improved, AUC range from 0.671 to 0.686. These results illustrated that using CT images of uniform slice thickness and reconstruction kernel achieved better performance than directly using heterogenous CT images with various slice thickness and reconstruction kernels.

To further explore the prognostic ability of our model, we assessed the 3D-CNN model (original 1-mm sharp images and synthetic 1-mm sharp kernel images transformed from 5-mm soft kernel images) with regard to PFS by the Kaplan-Meier method. This model successfully stratified the patients into high-risk and low-risk groups using the median predicted value of all patients (Fig.4). The PFS between the two groups is significantly different (P <0.001). This result indicated that our 3D-CNN model had a good prognostic value of identifying lung cancer patients who have high-risk of progression after receiving molecularly targeted therapy.

IV. CONCLUSION

In this paper, we proposed a CycleGAN-based model to convert heterogenous CT images into uniform CT images with the same slice thickness (1 mm) and reconstruction kernel (sharp kernel). Afterward, we compared the prognostic model using the synthetic unified CT images with the model using heterogenous CT images to predict 1-year PFS of lung cancer. We built a 3D-CNN model to predict the individualized progression. Our results suggested that transforming the 5-mm CT images with various reconstruction kernels into identical CT images with 1-mm slice thickness and sharp reconstruction kernel can improve the performance of predicting PFS. Furthermore, we investigated the prognostic ability of our model by Kaplan-Meier method. The results illustrated that our 3D-CNN model can significantly stratify

TABLE III
RESULTS OF 1-YEAR PFS PREDICTION

Input images	AUC		ACC	
	mean	sd	mean	sd
original 1-mm sharp kernel images + original 5-mm soft kernel images[a]	0.614	0.05	0.595	0.03
original 1-mm sharp kernel images+ original 5-mm sharp kernel images[a]	0.628	0.04	0.607	0.02
original 1-mm sharp kernel images + synthetic 1-mm sharp kernel images[b]	0.671	0.06	0.644	0.05
original 1-mm sharp kernel images + synthetic 1-mm sharp kernel images[c]	0.686	0.05	0.660	0.03

[a]standardized by interpolation;[b]generated from 5-mm sharp kernel images;[c]generated from 5-mm soft kernel images.

advanced lung cancer patients treated with molecularly targeted treatment into high/low-risk groups. For the high-risk patients, they are more resistant to the molecularly targeted therapy need more intervention in an early phase.

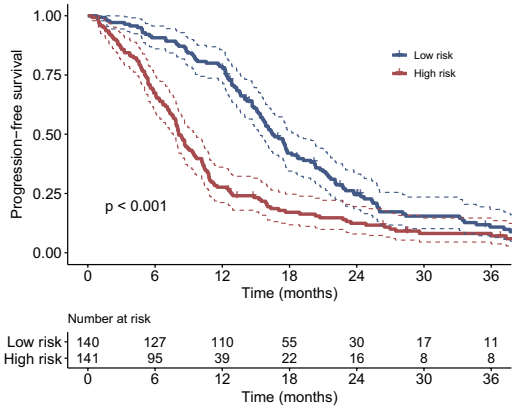

Fig. 4. The Kaplan-Meier curves of the 3D-CNN model.

REFERENCES

[1] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, Cancer Statistics, 2021, CA. Cancer J. Clin., vol. 71, no. 1, pp. 7–33, 2021.
[2] Z. F. Lim and P. C. Ma, Emerging insights of tumor heterogeneity and drug resistance mechanisms in lung cancer targeted therapy, J. Hematol. Oncol., vol. 12, no. 1, pp. 1–18, 2019.
[3] M. Ravanelli et al., CT texture analysis as predictive factor in metastatic lung adenocarcinoma treated with tyrosine kinase inhibitors (TKIs), Eur. J. Radiol., vol. 109, no. October, pp. 130–135, 2018.
[4] J. Choe et al., Deep Learning–based Image Conversion of CT Reconstruction Kernels Improves Radiomics Reproducibility for Pulmonary Nodules or Masses, Radiology, vol. 292, no. 2, pp. 365–373, 2019.
[5] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification, Neurocomputing, vol. 321, pp. 321–331, 2018.
[6] J. Zhao et al., Tripartite-GAN: Synthesizing liver contrast-enhanced MRI to improve tumor detection, Med. Image Anal., vol. 63, p. 101667, 2020.
[7] S. Wang et al., A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis., Eur. Respir. J., vol. 56, no. 2, p. 2000775, 2020.