A framework for 3D radiotherapy dose prediction using the deep learning approach

Lam Thanh Hien[1], Ha Manh Toan[2], Do Nang Toan[2]
[1]Faculty of Information Technology, Lac Hong University, Bien Hoa, Dong Nai, Vietnam
[2]Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam

Article Info

Article history:

Received Jan 4, 2024
Revised Jun 20, 2024
Accepted Jul 2, 2024

Keywords:

3D dose prediction
Cancer treatment
Radiotherapy
ResNet
U-Net

ABSTRACT

Cancer is known as a dangerous disease to humans with a very high death rate. There are a lot of cancer treatment methods that have been studied and applied in the world. One of the main methods is using radiation beams to kill cancer cells. This method, also known as radiotherapy, requires experts having a high level of skill and experience. Our work focuses on the 3D dose prediction problem in radiotherapy by proposing a framework aiming to create a medical intelligent system for this problem. To do that, we created a convolutional neural network based on ResNet and U-Net to generate the predicted radiation dose. To improve the quality of the training phase, we also applied some data processing techniques based on the characteristics of the 3D computed tomography (CT) data. The experiment used the dataset from patients who were cancer-treated with radiotherapy in the OpenKBP competition. The results achieved good evaluating metrics, the first is by the Dose-score and the second is by the dose-volume histogram (DVH) score. From the training result, we built the medical system supporting 3D dose prediction and visualizing the result as slices in heatmap form.

Corresponding Author:

Do Nang Toan
Institute of Information Technology, Vietnam Academy of Science and Technology
18 Hoang Quoc Viet, CauGiay District, Hanoi 10072, Vietnam
Email: dntoan@ioit.ac.vn

1. INTRODUCTION

Cancer is currently one of the most dangerous diseases for people all over the world. In 2020, the World Health Organization (WHO) [1] had a statistics report presenting nearly ten million deaths related to cancer. The figure meant nearly one-sixth of deaths and the report also showed that there are approximately 400,000 children developing cancer each year. According to [2] in 2019 and 2020, Vietnam recorded nearly 183 thousand new cancer patients and approximately 123 thousand deaths. Many reasons may cause cancer, such as using alcohol, using cigarettes, and unhealthy life routines. From the medical view, cancer starts when there is an appearance of cancer cells, and then they would grow to create malignant tumors. The cancer cells would invade and eliminate healthy organs in the human body. So, the main idea in many cancer treatments is the way to destroy cancer cells and remove cancer tumors. And one of the most common methods is radiotherapy.

Radiotherapy tries to destroy cancer cells by using the power of radiation therapy beams in treatment machines. By the way, we can prevent the growth of the disease. An advantage of radiation therapy is that this method can be applied to almost all parts of the human body, but it also is a complex process requiring high-level experts [3]. This method can take hours to calculate the radiotherapy dose distribution and adjust to get the final optimal planning. To do that, practitioners need to be skillful and experienced to reach the accuracy of each detail of the whole process.

This paper concerns dose prediction, a method to help make a high-quality plan quickly. There were some studies that attempted to apply traditional machine-learning models to this problem. In 2016, McIntosh and Purdie [4] applied the regression forests to predict dose. Results, the overall dose prediction accuracies reached 78.68%, 64.76%, and 86.83% respectively. Next in 2021, Zhou *et al.* [5] presented a study using a support vector machine to predict doses in the case of cervical cancer with the mean squared error and Gamma metrics. The experimental data came from 50 patients in the affiliated Hospital of Southwest Medical University over the 2-years period between 2016 and 2018.

Today, the deep learning approach was applied successfully in many studies of the medical field with various different data types. In 2021, Furtado [6] conducted a study to segment abdominal organs in magnetic resonance imaging data using two deep-learning architectures DeepLabV3 and fully convolutional network (FCN). Also in 2021, Chen *et al.* [7] designed Squeeze-and-Excitation convolutional neural network (CNN) to classify lung nodules from computed tomography images. In 2023, Toan *et al.* [8] had research using several convolution neural networks on chest X-ray images for tuberculosis diagnosis. Furthermore, recent deep learning research in healthcare is also oriented towards building applicable systems. For example, Prabakaran and Selvaraj [9] presented an intelligent healthcare system using an ensemble approach to diagnose lung disease. The same year, Alquran *et al.* [10] showed a hybrid automated system to classify liver tumors on magnetic resonance images.

In recent times, the deep-learning approach has been applied widely in the radiotherapy dose prediction problem. In 2021, a 3D Dense-U-Net model was presented by Liu *et al.* [11] in the report with respectively 2.93% and 2.42% of the average deviations from the maximum and average dose values of organs at risks (OARs) and planning target volumes (PTVs). The authors experimented on the data of 124 patients treated with TomoTherapy. Another study is a convolutional neural network of Ahn *et al.* [12]. Their architecture was based on the U-Net model and adjusted to fit the dose prediction problem in radiotherapy. 50 samples of therapy plans for breast cancer were used to test that method and the authors also conducted the comparison with RapidPlan™. The work of Babier *et al.* [13] with the generative adversarial network (GAN) model was conducted on annotated computed tomography (CT) images from 127 samples in the accepted treatment plans for oropharyngeal cancer cases. In another attempt, many versions of convolutional neural networks were tested in the work of Gronberg *et al.* [14]. They trained and tested their model on the data of the OpenKBP challenge [15] with some different options of hyperparameters and augmentation methods. In the same dataset, Soomro *et al.* [16] published an architecture combining dilated DenseNet and ResNet to predict dose called DeepDoseNet. In experiments, some loss functions were used and the quality of that model was scored in terms of dose-score and dose-volume histogram (DVH) score. Zimmermann *et al.* [17] applied one cycle learning to their customized U-Net version and also experimented on the dataset of the OpenKBP challenge. In that study, some changes were tried with loss functions and Mish activation function, and at that time, their results ranked 4th and 2nd respectively on dose score and DVH score in the OpenKBP challenge.

This paper is interested in the radiation therapy dose prediction problem. In detail, we use CT images of patients treated with radiotherapy in our framework and apply the deep learning approach. The results of deep-learning methods are highly related to the training dataset, in both quantity and quality. Particularly for medical images, there are many cases where it is difficult to distinguish between regions of internal organs due to poor contrast and different configurations of imaging machines. Moreover, medical images are generally more scarce than regular images due to privacy and security constraints. So, on the one hand, we need normalization techniques to reduce noise to get normalized data. On the other hand, we also have to deal with the scarcity of data by applying data augmentation. In this work, our study aims to build a dose prediction system in the context of radiotherapy treatment by proposing a framework connecting all tasks. Firstly, we designed a customized convolution neural network for dose prediction and applied some preprocessing and augmentation techniques to enhance this model. Secondly, we analyzed the predicted results in many faces, using standard metrics such as dose-score and DVH-score, visualizing the predicted results and comparing them to the ground truth data, providing running time information. Lastly, we built a complete system to support 3D radiotherapy dose prediction and visualize the result in different dimensions.

2. METHOD

This section describes how we conducted our research and the experimental procedure. Inside, we would present the proposed framework and then detail the components in this. They include data preprocessing and augmentation techniques, the custom convolutional neural network, and the dose visualization technique.

2.1. Proposed framework

The framework includes the training and the testing phases as in Figure 1. The training phase aims to improve the quality of the model based on the training data. In this work, the training data includes CT images, annotated information about PTVs and OARs. OARs and PTVs are masks to point out the annotated regions. So, the input data has 11 channels with CT images, OARs, and PTVs. The input data would be preprocessed and augmented before pushing to the model in the training task. During the training phase, the state of the model's weights would be evaluated at each epoch with some metrics. Here, the dose score and the DVH score are two indicators used for scoring the model's checkpoints to choose the best state.

The testing phase aims to the prediction function of the study's problem which can be applied in the medical software of radiotherapy. Here, the input data also includes CT images, OARs, and PTVs. The input data would be preprocessed and then pushed into the trained model to get the predicted dose. This dose would be applied to the constraint about the value range to get the final result. This dose also would be displayed as a heatmap to help people easily understand.

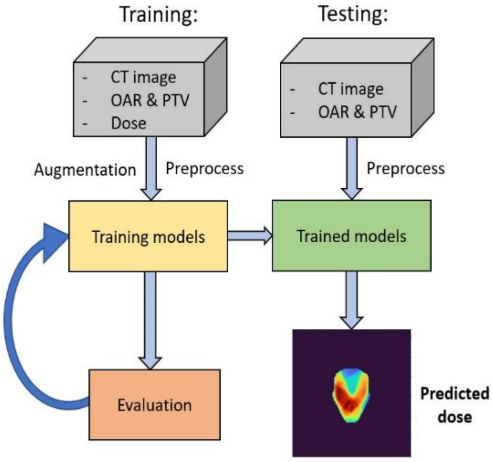

Figure 1. The proposed framework

2.2. Data processing techniques

The data used in this work were CT images of patients who had head-and-neck cancer and were treated by radiotherapy treatment. Before the deep learning task, we tried to apply some data processing techniques to CT images. This step would help normalize image data to improve the final results. In comparison with normal images, because of the difference between capturing machines, CT image data has some particular characteristics. Pixel values are related to the material of captured objects, such as bone, and high-fat tissues. This information is referenced in Hounsfield units [18].

For augmentation, in the case of this study, the data were 3D CT scans. In CT images, the empty space usually gets values close to zero. Because we are only interested in the body in the CT image, we would apply transformations with translation and flipping and no change in the scale of the CT image. In this way, the 3D content of the human body is still preserved. Another reason is that the image-capturing pose is fixed due to medical indications. This means that objects in CT images often have stable size.

The first technique is flipping: As described in Figure 2, this technique will rotate images randomly by an angle equaling 180 degrees. So, the image result will have a different pose versus origin. In this case, we can quickly apply this technique and do not need to calculate image content at gaps in the corner regions of 3D volume. The second is translation: As described in Figure 3, this technique will translate the CT image with a random displacement and not lose any body parts. This technique is also simple and fast to perform.

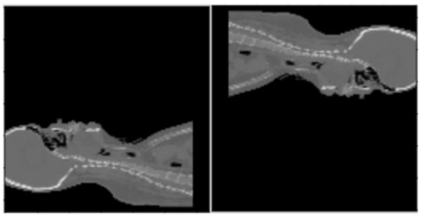

Figure 2. Image flipping example: the original image (left) and the flipped image (right)

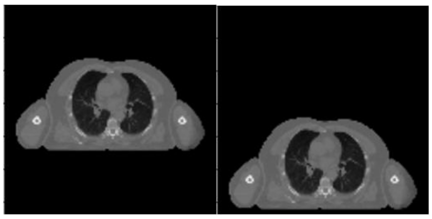

Figure 3. Image translation example: the original image (left) and the translated image (right)

2.3. The custom convolutional neural network

In this work, we designed an architecture that is inspired by the study of Ronneberger *et al.* [19] in 2015. We also used a residual block presented by He *et al.* [20] in their architecture for the image classification problem. Figure 4 shows the design of a residual block.

In this work, the structure has two parts as described in Figure 5. The first is the down-sampling part and the second is the up-sampling part. In this architecture, residual blocks would be used in both parts. In Figure 5, the input data is at the beginning of the down-sampling part as described on the left-top side, and the output data is at the end of the up-sampling part as described on the right-top side.

The size of the input data is 128×128×128×11. The first three numbers are the size of the CT input image and the last number is the amount of channels. Among the 11 channels, there is one channel for the CT input image, seven channels for OARs, and three channels for PTVs. Residual blocks are designed of two convolution layers. Each layer has the same 3×3×3 sliding window size, 1 stride, and also is padded to keep size. LeakyReLU functions are set following convolution layers. Besides, in the up-sampling part, a deconvolution layer is used after each residual block. The end of the architecture is added with a ReLU function. This function helps keep each pixel from negative values because of the meaning of radiation doses. At last, the output data size is 128×128×128.

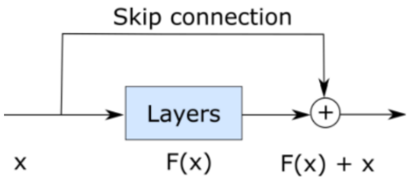

Figure 4. Residual block

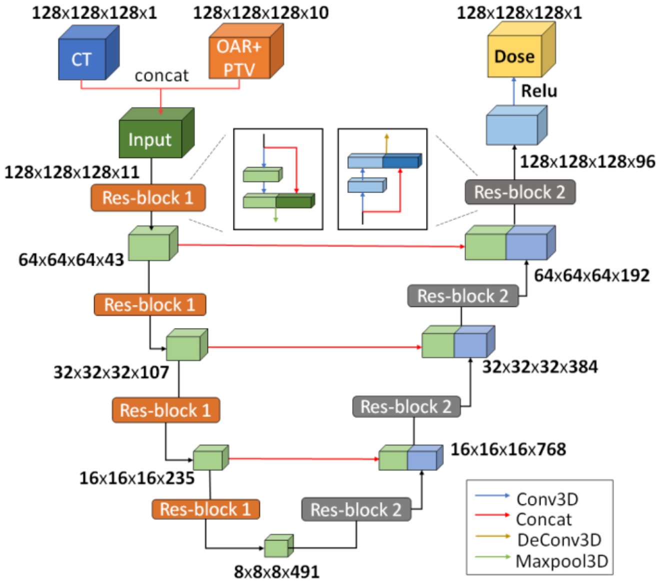

Figure 5. Our custom CNN architecture

2.4. Dose visualization

In this work, the output dose is in 3D raster format. Thus, all results would be displayed as images of each slice. Because the CT image has one channel, the displayed image is grayscale. But human vision is not sensitive to the changes in gray levels, we often distinguish well in colors. So, we use the colormap technique to enhance the view of the results.

In implementation, the 256-color look-up table is applied to displayed images. Each dose data slice needs to be transformed into the grayscale to be able to apply the colormap technique. In the experiment, we coded this module with the support of Matplotlib [21]. Matplotlib is the Python library that supports some options of colormaps. An example of dose visualization was described in Figure 6.

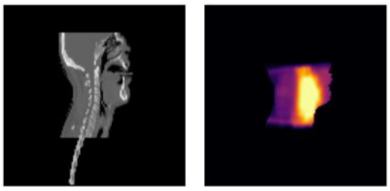

Figure 6. Dose visualization of an example in the dataset: the CT data (left), the visualization result (right)

3. RESULTS AND DISCUSSION

3.1. Experiment setup

3.1.1. Dataset

The dataset used in our work was collected from 340 patients who had head-and-neck cancer and were treated by radiotherapy treatment. This data has a source from the open-access database TCIA. TCIA contains medical images supporting cancer research and belongs to the University of Arkansas. The authors pushed all data into the data-cleaning step. Each patient's dose influence matrix was obtained using 6 MV step-and-shoot intensity modulated radiation therapy (IMRT). This task was conducted with equispaced coplanar fields at 0°, 40°, 80°, 120°, 160°, 200°, 240°, 280°, and 320°. All fields were divided into 5×5 mm beamlets. The OpenKBP challenge organized this data for the competition. So, the dataset was organized with the training set, the testing set and the validation set with sizes of 200, 40, and 100 respectively. Each patient's record includes a CT image with a size of 128×128×128, annotations for organs at risk, annotations for planning target volumes, and information about the treated radiotherapy dose distribution. In detail, organs at risk consist spinal cord, brain stem, left parotid gland, right parotid gland, larynx, mandible, and esophagus. Planning target volumes consist of 56 Gy, 63 Gy, and 70 Gy regions. Figure 7 shows a data sample in this dataset.

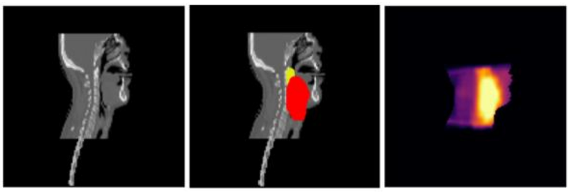

Figure 7. Data example: CT image (left), target regions visualization (middle), and dose visualization (right)

3.1.2. Evaluation metrics

To evaluate results, we need metrics. This experiment used the DVH-score and the dose-score to analyze the dose-predicted results. All metrics are calculated between the true dose and the predicted dose.

The dose-score is calculated using the volume of interest receiving radiation therapy with the mean of the absolute difference between the true value and the predicted value. The dose-score needs to be calculated in each voxel. The formula of the dose-score is described as (1):

$$\alpha_p = \frac{\|s_P - \hat{s}_P\|_1}{|v_P|} \tag{1}$$

The score α_p is calculated for patient p. In detail, patient p has the volume treated radiotherapy v_P with the size of $|v_P|$. About doses, s_P is the predicted dose, and lastly, $\hat{s}_P$ is the true dose.

The DVH score is calculated as the absolute distance between the DVH values of the true dose and the predicted dose. For OAR, the task is to estimate the dose average and the dose maximization corresponding to organs. Similarly, for PTV, the task is to estimate the dose corresponding to the target volumes at 1 percent, 95 percent, and 99 percent rates. The DVH-score is calculated as (2):

$$\beta_P = \|D(s_P) - D(\hat{s}_p)\|_1 \tag{2}$$

The score β_P is calculated for patient p. About the predicted result, s_P is the predicted dose with $D(s_P)$ is the DVH value of s_P. About the ground truth, $\hat{s}_p$ is the true dose and $D(\hat{s}_p)$ is the DVH value of $\hat{s}_p$.

3.1.3. Training configuration

Input data consists of 11 channels with each channel measuring 128×128×128, and the output data size is 128×128×128×1. For input, the first channel is a CT image, the next 7 channels contain OAR information, and the last 3 channels contain PTV information. Training task was conducted using Adam's algorithm [22]. Our study was implemented in Python language and using the deep-learning platform TensorFlow [23]. The experiment software run on both the Google Colab system [24] and the Kaggle system [25], which are supported by the computation power of NVIDIA graphics processing units (GPUs). The loss function is the mean of absolute error. Inside, we set up the learning rate by $1e-3$, the decay by $1e-4$, the momentum β1 by 0.9, and the momentum β2 by 0.99. The formula of the mean absolute error loss function is calculated as (3):

$$\mathrm{L} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}|\mathrm{y_i} - \hat{\mathrm{y}}_\mathrm{i}| \tag{3}$$

In which, n is the number of samples, y_i is the predicted dose, and $\hat{y}_i$ is the true dose of sample i.

In the experiment, we design one more case of model training to clarify the effect of preprocessing and augmentation techniques. This case used original data, which had no augmentation and preprocessing. The detail was described in Table 1. In our hypothesis, the results from case 2 should be better than case 1. The result figures would be the evidence proving the effect of preprocessing and augmentation in the experiment. For training and scoring the model, we used the training set with 200 volume samples and the testing set with 100 volume samples.

Table 1. Two cases of model training

Training case	Description
1	Using an original dataset
2	Using a dataset which was applied augmentation and preprocessing

3.2. Result and evaluation

To begin with, we investigated the training results of two cases in this study. The line graphs of the loss functions are presented in Figure 8. In general, the training losses were reduced during the training period. In case 2, the loss value was 0.607552 at the first epoch, decreased to 0.316148 at the third epoch, and continued to decline at a lower rate. The training loss value in case 1 was also reduced but it had some small fluctuations in the line graph. Specifically, the lowest value always belonged to the loss value of case 2. It reflects our theoretical hypothesis because in case 2, the model was trained with better training data, which was performed preprocessing and augmenting.

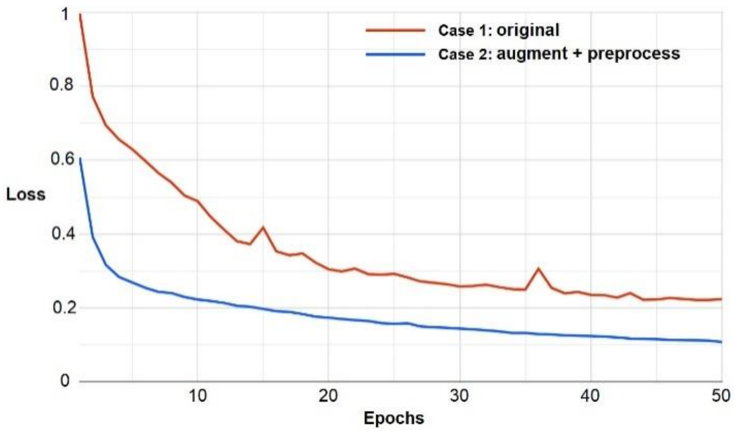

Figure 8. The training results of two cases

The next task is to evaluate the trained models in training cases. The detailed results are shown in Table 2. To do this, we used a 100-sample test set and 2 measures of DVH-score and dose-score. Easy to see, case 1 is worse than case 2 in both measures. About DVH-score, this score in case 1 is 18.931728 while in case 2 is 2.011907. About dose-score, this score in case 1 is 11.307606 while in case 2 is 3.111332. All figures pointed out that the result in case 2 is better than the results in case 1. This conclusion also conforms to our hypothesis when setting up training cases.

Table 2. Evaluation of two cases in the testing phase

Training case	Dose-score	DVH-score
1	11.307606	18.931728
2	3.111332	2.011907

To get clearer, we conducted an analysis of the predicted result in one sample. In detail, the analysis was performed on the result of our custom model that was trained with preprocessed and augmented data, which means case 2. The randomly chosen sample is sample 258 in the dataset. The corresponding dose-score is 2.592035 and the corresponding DVH-score is 1.805577. We focused on the visualization of the predicted result. And because the input data is CT slices and the output is the 3D volume of doses, we would analyze when viewing slices of data. By using visualization as mentioned in subsection 2.4, we would observe each slice of predicted dose as a heatmap, as displayed in Figure 9.

In Figure 9, the higher radiation dose intensity would correspond to the brighter color in the heatmap, and also vice versa. In this way, we can observe the distribution of radiation dose directly with our eyes. The next task is to represent the comparison between the true dose and the predicted dose, as described in Figure 10.

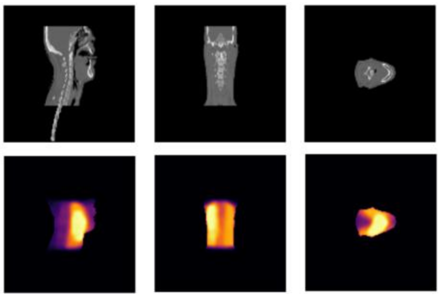

Figure 9. Three views of the prediction result of sample 258: the CT images (first row) and the dose visualizations (second row)

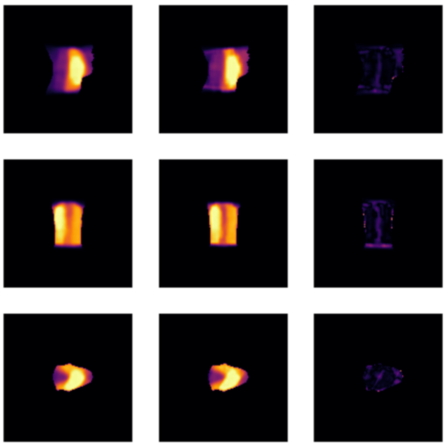

Figure 10: The comparison of the prediction results with the ground truth data of sample 258: the ground truth (left), the predicted dose (middle), and the difference image (right)

In Figure 10, the first two columns are the predicted doses and the true doses which are displayed as heatmaps. The last column is the difference images between the predicted doses and the true doses. Each difference image is computed by subtracting the corresponding predicted dose from the corresponding true dose and getting the absolute value of the subtraction result. The difference images are also shown as heatmaps and easily to see, values of pixels in the difference images are low over the majority of the area. About the comparison of case 1 and case 2, we also present in Figure 11.

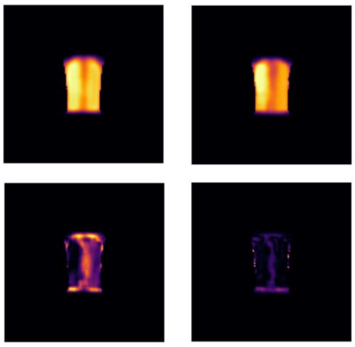

Figure 11. The comparison between cases of sample 258: case 1 (left) and case 2 (right), each column includes the predicted dose visualization and the difference image

In Figure 11, we observe the predicted doses in the first row and the difference images, which are the result of a comparison between the predicted doses to the ground truth in the second row. In detail, information about case 1 is presented in the first column and information about case 2 is presented in the second column. When observing the second row, we can see that case 2 is better than case 1 because the difference image in case 2 has lower intensity than case 1 in general. The next task is the information on the running time as described in Table 3. In detail, we presented the running time of one epoch in the training phase and the running time of one sample prediction. About prediction, we supplied time information when running in both the GPU machine and the central processing unit (CPU) machine.

To estimate the running time in cases, we used the Google Colab platform supported by Tesla K80 GPU and our personal laptop with the simple configuration of Intel I5 with 1.60 GHz. As presented in Table 3, the running time in the CPU machine is not long and it is acceptable in our real test. This information is helpful for us because GPU machines are rather expensive and not easy to be ready for deployment in general. Another estimation is the comparison with the result of DeepDoseNet in the study [16]. In that research, the authors experimented with some different loss functions. In this study, we performed a comparison between our result in case 2 with the result of DeepDoseNet in the same context experiment of the loss function - the mean absolute error loss. All details are shown in Table 4.

Table 3. Running time information

Case	Machine	Approximate time
One epoch training of case 1	GPU	10.37 minutes
One epoch training of case 2	GPU	30.45 minutes
Predicting one sample of case 2	GPU	1.24 second
Predicting one sample of case 2	CPU	11.15 seconds

Table 4. Comparison with the DeepDoseNet in case of the same loss function

Model	Dose-score	DVH-score
Our model (case 2)	3.111332	2.011907
The DeepDoseNet	3.5	2.3

From the theory of the DeepDoseNet, Soomro *et al.* [16] also applied the residual blocks. However, our model was designed with more layers. In addition, Soomro *et al.* [16] did not perform augmenting data and they also did not talk about preprocessing data. Finally, we built a medical system for predicting dose as described in Figure 12. This software would output the dose prediction and display it as a heatmap on the application canvas. Besides, we also support viewing dose data in different dimensions.

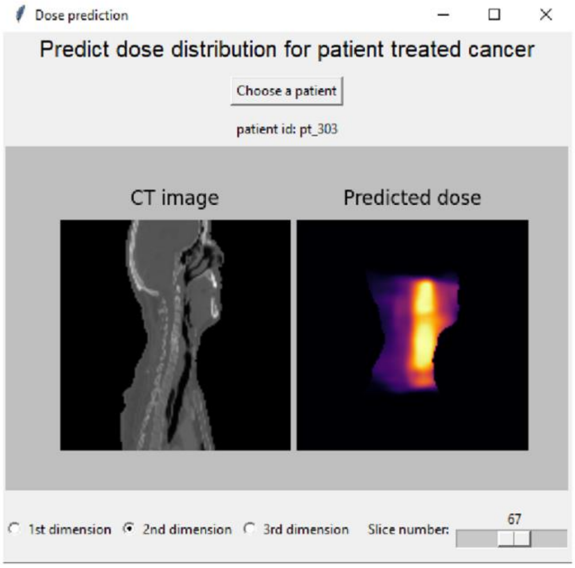

Figure 12. The application interfaces

4. CONCLUSION

Cancer is a dangerous disease so requires a lot of attempts to prevent it. One of the leading cancer treatments is radiotherapy which applies radiation therapy beams to eliminate cancer cells. This study investigated the applying of the deep learning approach to radiotherapy dose prediction. We found that it was necessary for suitable deep architecture and data-processing techniques. The proposed method in this study uses a network architecture for 3D dose prediction based on Res-U-Net 3D and preprocessing and augmentation techniques based on analysis of data characteristics. Our study demonstrates the performance when compared to the DeepDoseNet study which does not apply data-processing techniques. Lastly, we also developed the 3D dose prediction application supporting dose visualization. This paper obtained some success but to be able to build a medical intelligent system that can work effectively in reality, we also have to do a lot of work. The success of this work is one of the first steps in our way of preventing cancer in the intelligent medical industry.

ACKNOWLEDGEMENTS

The research content in this paper is sponsored by the Institute of Information Technology (Viet Nam Academy of Science and Technology) under the project code “CS24.01”.

REFERENCES

[1] WHO, “Cancer,” *World Health Organization*, 2022. https://www.who.int/news-room/fact-sheets/detail/cancer (accessed Dec. 12, 2023).

[2] WHO, “Cancer today,” *World Health Organization*, https://gco.iarc.fr/today/data/factsheets/populations/704-viet-nam-fact-sheets.pdf (accessed Dec. 12, 2023).

[3] S. Gianfaldoni, R. Gianfaldoni, U. Wollina, J. Lotti, G. Tchernev, and T. Lotti, “An overview on radiotherapy: from its history to its current applications in dermatology,” *Open Access Macedonian Journal of Medical Sciences*, vol. 5, no. 4, pp. 521–525, Jul. 2017, doi: 10.3889/oamjms.2017.122.

[4] C. McIntosh and T. G. Purdie, “Contextual atlas regression forests: multiple-atlas-based automated dose prediction in radiation therapy,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 4, pp. 1000–1012, Apr. 2016, doi: 10.1109/TMI.2015.2505188.

[5] P. Zhou *et al.*, “Support vector machine model predicts dose for organs at risk in high-dose rate brachytherapy of cervical cancer,” *Frontiers in Oncology*, vol. 11, Jul. 2021, doi: 10.3389/fonc.2021.619384.

[6] P. Furtado, “Magnetic resonance sequences: experimental assessment of achievements and limitations,” *Journal of Advances in Information Technology*, vol. 12, no. 1, pp. 66–70, 2021, doi: 10.12720/jait.12.1.66-70.

[7] Y. Chen, W. Du, X. Duan, Y. Ma, and H. Zhang, “Squeeze-and-excitation convolutional neural network for classification of malignant and benign lung nodules,” *Journal of Advances in Information Technology*, vol. 12, no. 2, pp. 153–158, 2021, doi: 10.12720/jait.12.2.153-158.

[8] H. M. Toan, L. T. Hien, N. D. Vinh, and D. N. Toan, “Detecting tuberculosis from Vietnamese X-Ray imaging using transfer learning approach,” *Computers, Materials & Continua*, vol. 74, no. 3, pp. 5001–5016, 2023, doi: 10.32604/cmc.2023.033429.

[9] Prabakaran and P. Selvaraj, “Advance IoT intelligent healthcare system for lung disease classification using ensemble techniques,” *Computer Systems Science and Engineering*, vol. 46, no. 2, pp. 2141–2157, 2023, doi: 10.32604/csse.2023.034210.

[10] H. Alquran *et al.*, “Liver tumor decision support system on human magnetic resonance images: a comparative study,” *Computer*

Systems Science and Engineering, vol. 46, no. 2, pp. 1653–1671, 2023, doi: 10.32604/csse.2023.033861.

[11] Y. Liu *et al.*, "Dose prediction using a three-dimensional convolutional neural network for nasopharyngeal carcinoma with tomotherapy," *Frontiers in Oncology*, vol. 11, Nov. 2021, doi: 10.3389/fonc.2021.752007.

[12] S. H. Ahn *et al.*, "Deep learning method for prediction of patient-specific dose distribution in breast cancer," *Radiation Oncology*, vol. 16, no. 1, Dec. 2021, doi: 10.1186/s13014-021-01864-9.

[13] A. Babier, R. Mahmood, A. L. McNiven, A. Diamant, and T. C. Y. Chan, "Knowledge-based automated planning with three-dimensional generative adversarial networks," *Medical Physics*, vol. 47, no. 2, pp. 297–306, Feb. 2020, doi: 10.1002/mp.13896.

[14] M. P. Gronberg, S. S. Gay, T. J. Netherton, D. J. Rhee, L. E. Court, and C. E. Cardenas, "Dose prediction for head and neck radiotherapy using a three-dimensional dense dilated U-Net architecture," *Medical Physics*, vol. 48, no. 9, pp. 5567–5573, Sep. 2021, doi: 10.1002/mp.14827.

[15] A. Babier *et al.*, "OpenKBP: the open-access knowledge-based planning grand challenge and dataset," *Medical Physics*, vol. 48, no. 9, pp. 5549–5561, Sep. 2021, doi: 10.1002/mp.14845.

[16] M. H. Soomro, V. G. L. Alves, H. Nourzadeh, and J. V. Siebers, "DeepDoseNet: a deep learning model for 3D dose prediction in radiation therapy," *arXiv preprint arXiv:2111.00077*, 2021.

[17] L. Zimmermann, E. Faustmann, C. Ramsl, D. Georg, and G. Heilemann, "Dose prediction for radiation therapy using feature-based losses and one cycle learning," *Medical Physics*, vol. 48, no. 9, pp. 5562–5566, Sep. 2021, doi: 10.1002/mp.14774.

[18] Radiopaedia, "Hounsfield unit," *Radiopaedia*, 2024. https://radiopaedia.org/articles/hounsfield-unit (accessed Feb. 19, 2024).

[19] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 2015, pp. 234–241.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[21] Matplotlib, "Choosing Colormaps in Matplotlib," Matplotlib, https://matplotlib.org/stable/tutorials/colors/colormaps.html (accessed Feb. 19, 2024).

[22] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[23] M. Abadi *et al.*, "TensorFlow: large-scale machine learning on heterogeneous systems," *Preprint arXiv:1603.04467*, 2016.

[24] "Welcome to Colab," *Colab*, https://colab.research.google.com (accessed Feb. 19, 2024).

[25] "Level up with the largest AI & ML community," *Kaggle*, https://www.kaggle.com (accessed Feb. 19, 2024).

BIOGRAPHIES OF AUTHORS

Lam Thanh Hien joined an M.Sc. in applied informatics at the INNOTECH Institute, France, and received a degree in 2004. In 2017, he earned a Ph.D. degree at the Vietnam Academy of Science and Technology. Now, he works at Lac Hong University in the headmaster's role. His studies interests relate to machine learning, computer vision, and deep learning. He can be contacted at email: lthien@lhu.edu.vn.

Ha Manh Toan learned applied mathematics and informatics at the College of Science, Vietnam National University, Hanoi, and received a degree in 2009. In 2015, he earned an M.Sc. degree at the University of Engineering and Technology, Vietnam National University, Hanoi. Now, he is a researcher at the Vietnamese Academy of Science and Technology. His studies interests relate to machine learning, computer vision, and deep learning. He can be contacted at email: hmtoan@ioit.ac.vn.

Do Nang Toan studied applied mathematics and informatics at Hanoi University and received a degree in 1990. In 2001, he earned a Ph.D. degree at the Vietnam Academy of Science and Technology. Now, he is an associate professor at the Vietnamese Academy of Science and Technology. His studies interests relate to machine learning, computer vision, and virtual reality. He can be contacted at email: dntoan@ioit.ac.vn.