Bladder Cancer Treatment Response Assessment in CT Urography by Using Deep-Learning and Radiomics

Di Sun*[a], Lubomir Hadjiiski[a], Heang-Ping Chan[a], Richard H. Cohan[a], Elaine M. Caoili[a], Ajjai Alva[b], Kenny H. Cha[a], Ravi K. Samala[a], Chuan Zhou[a]

[a]Department of Radiology, University of Michigan, 1500 E Medical Center Dr, Ann Arbor, MI, USA 48109; [b]Department of Internal Medicine-Hematology/Oncology, University of Michigan, 1500 E. Medical Center Drive, Ann Arbor, MI, USA 48109

ABSTRACT

We are developing deep-learning convolutional neural network (DL-CNN) and radiomics models to assist physicians in treatment response assessment in CT urography of bladder cancer to neoadjuvant chemotherapy (NAC). We collected a total of 264 pre- and post- treatment lesion pairs of 227 patients from University of Michigan hospital with IRB approval. The data were split into 3 sets by case: a training set including 35 complete responders (CRs) (T0 stage after treatment) and 113 non-complete responders (NCRs) (> T0 stage after treatment), a validation set including 5 CRs and 5 NCRs, and an independent test set including 19 CRs and 87 NCRs. The training set was used to train the models to classify CRs and NCRs and the selection of optimal models was guided by the validation set. The selected models were deployed on the test set to generate the likelihood score of CR of each pair. The classifying performance was evaluated by the area under ROC curve (AUC). Hybrid ROIs extracted from the lesions in the pre- and post- treatment scan pairs were used as input to the DL-CNN model. The optimal DL-CNN model achieved an AUC of 0.75 ± 0.06 on the test set. For the radiomics model, the random forest classifier was applied to the features extracted from the pre- and post-treatment lesions. The optimal radiomics model achieved an AUC of 0.76 ± 0.05. A combined DL-CNN and radiomics model increased the AUC to 0.77 ± 0.06. The results indicated the feasibility of using the DL-CNN model and radiomics model for assessing treatment response of bladder cancer.

Keywords: computer-aided diagnosis; bladder cancer; treatment response assessment; deep learning; radiomics; neoadjuvant chemotherapy

Abbreviations: deep-learning convolutional neural network (DL-CNN); complete responders (CRs); non-complete responders (NCRs); neoadjuvant chemotherapy (NAC)

1. INTRODUCTION

Bladder cancer is the 4th most common cancer in men. In 2022, about 81,180 new cases of bladder cancer were diagnosed and about 17,100 cases died from it in the US[1]. The 5-year relative survival rate was 96% for in situ stage, 70% for localized stage, 38% for regional stage, and 6% for distant stage[1]. Studies have shown that early diagnosis and treatment can improve the survival rate. Neoadjuvant chemotherapy (NAC) may decrease the tumor volume, making the tumor more operable for radical cystectomy, and reduce the probability of metastatic disease[2-4]. However, NAC has associated toxicities including neutropenic fever, sepsis, mucositis, nausea, vomiting, malaise, and alopecia[5]. It is therefore of great importance to assess the response of the bladder lesions at early treatment cycles to identify lesions that do not respond so that the patient can seek alternative treatment early without suffering the toxicities of the entire course of chemotherapy.

The purpose of this study is to develop DL-CNN and radiomics models to assist physicians in assessing the response of bladder cancer to NAC.

*disun@umich.edu

Medical Imaging 2023: Computer-Aided Diagnosis, edited by Khan M. Iftekharuddin, Weijie Chen, Proc. of SPIE Vol. 12465, 1246533 · © 2023 SPIE
1605-7422 · doi: 10.1117/12.2654659

2. MATERIALS AND METHODS

2.1 Dataset

With IRB approval, we collected retrospectively pre- and post- chemotherapy CTU scans of 227 patients from University of Michigan hospital, from which 264 pre- and post- treatment lesion pairs were identified. The pathological cancer stage after treatment was used as the gold standard for determining if a lesion would have complete response after the whole course of NAC treatment. Fifty-nine of the 264 cancer lesions were clinically determined to be at T0 stage (complete response) after NAC. Figure 1 is an example of 3D CTU scan and axial slice of a CTU scan.

(a) (b)

Figure 1: Examples of CTU scan. (a) A 3D volume rendering of CTU scan in which the bladder and ureters are visible. (b) An axial slice of a CTU scan in which the bladder is partially filled with intravenous (IV) contrast material.

2.2 Preprocessing

For each scan pair, our auto-initialized cascaded level sets system (AI-CALS) was used to segment the lesions, from which multiple regions of interest (ROIs) of 32×16 pixels were extracted[6]. One ROI from the pre-treatment scan and one from the post-treatment scan formed a hybrid ROI (hROI) of 32×32 pixels, resulting in a number of hROIs from each lesion pair (Figure 2). The pathological cancer stage after treatment was collected as reference standard for determining if a patient had complete response to NAC treatment (stage T0) or not (stage >T0), which also determined the label of all hROIs from the lesion.

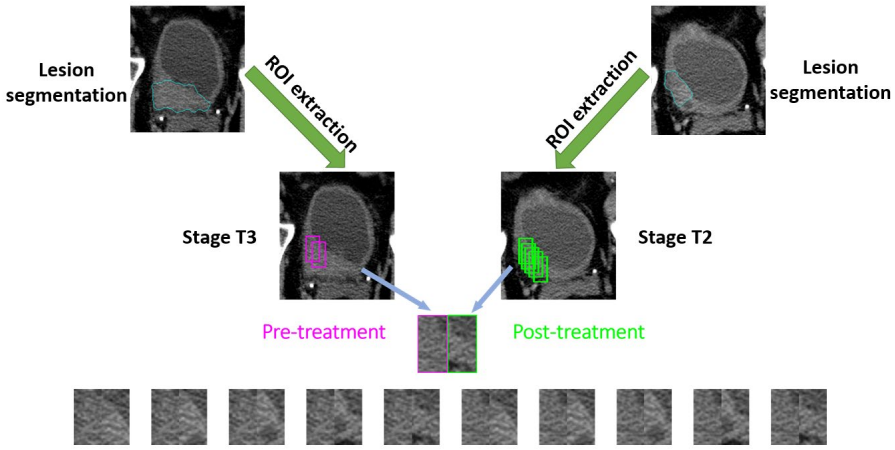
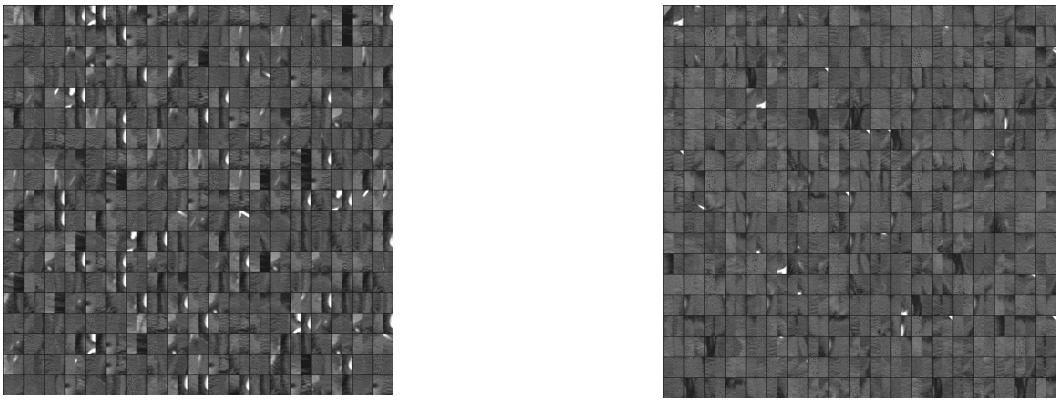


Figure 2: Hybrid ROIs generated from a pre- and post-treatment lesion pair. In this example the cancer was stage T3 in pre-treatment, and stage T2 in post-treatment scan, resulting in a label of non-complete response (>T0 stage after treatment).

The data set was split by case into training, validation, and test sets. The training set consisted of 23,488 hROIs from 148 pairs (35 at T0, 113 at >T0). The validation set consisted of 2,048 hROIs from 10 pairs (5 at T0, 5 at >T0). The test set consisted of 108,608 hROIs from 106 pairs (19 at T0, 87 at >T0). Figure 3 shows the examples of the hROIs subset.



(a) (b)

Figure 3: Subsets of hybrid ROIs. (a) A subset of complete responders (T0). (b) A subset of non-complete responders (>T0).

2.3 Classification

DL-CNN Model: We implemented and validated a DL-CNN structure in TensorFlow[7] to assess the cancer treatment response (T0 versus >T0)[8]. The DL-CNN used in this study consisted of two convolutional layers (C1 and C2), two locally connected layers (L3 and L4), and a fully connected layer (FC10). Layers C1 and C2 had 5 × 5 kernels with a stride of 1, followed by max pooling with a 3 × 3 filter of stride 2 and local response normalization layer. Transfer learning was employed such that the model was pre-trained with the CIFAR10 image set before fine-tuning with the bladder cancer training set. After the DL-CNN was trained to classify CRs and NCRs by using the training set, the trained models were deployed on the validation set to select the optimal model. Then the optimal model was deployed on the hybrid ROIs of the test set, the likelihood score of complete response of each lesion pair was generated by taking the minimum of the scores of all hybrid ROIs from the same lesion pair. The classification process is shown in Figure 4.

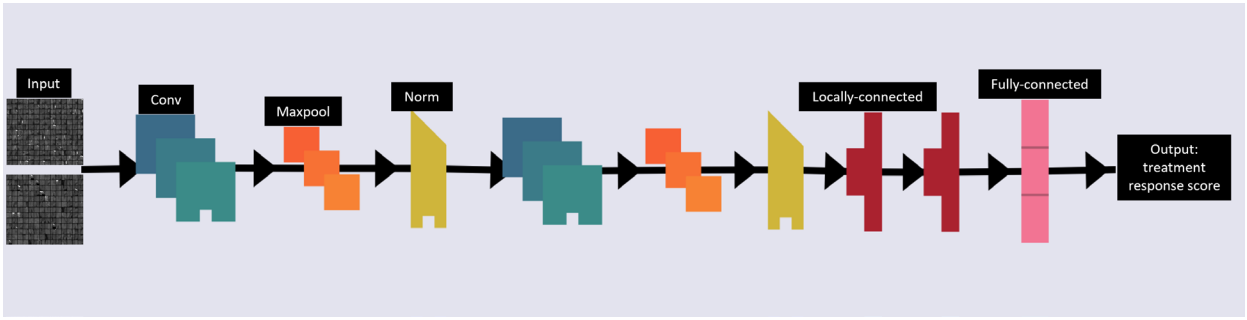


Figure 4: Classification process. The hROIs were the input to DL-CNN. DL-CNN assessed the hROIs and outputted the cancer treatment response score.

Radiomics model: For the radiomics model, we extracted 91 features from the segmented pre- and post-treatment lesions, including shape, size, and texture features[8]. The random forest classifier was employed to classify CRs and NCRs, guided by the validation set to select the parameter settings. The model with the selected optimal parameters setting was deployed on the test set to perform the classification, generating the likelihood score of complete response of each pair.

Combined model: We combined the response scores from the DL-CNN and radiomics models of each lesion pair by taking the smaller value out of the two scores. The area under the ROC curve (AUC) was calculated as the performance measure.

3. RESULTS

The AUC was 0.75 ± 0.06 for the DL-CNN model and 0.76 ± 0.05 for the radiomics model. The combined DL-CNN-radiomics model achieved an AUC of 0.77 ± 0.06. Figure 5 shows some examples of pre- and post-treatment lesion pairs and the response predictions by the DL-CNN and radiomics models.

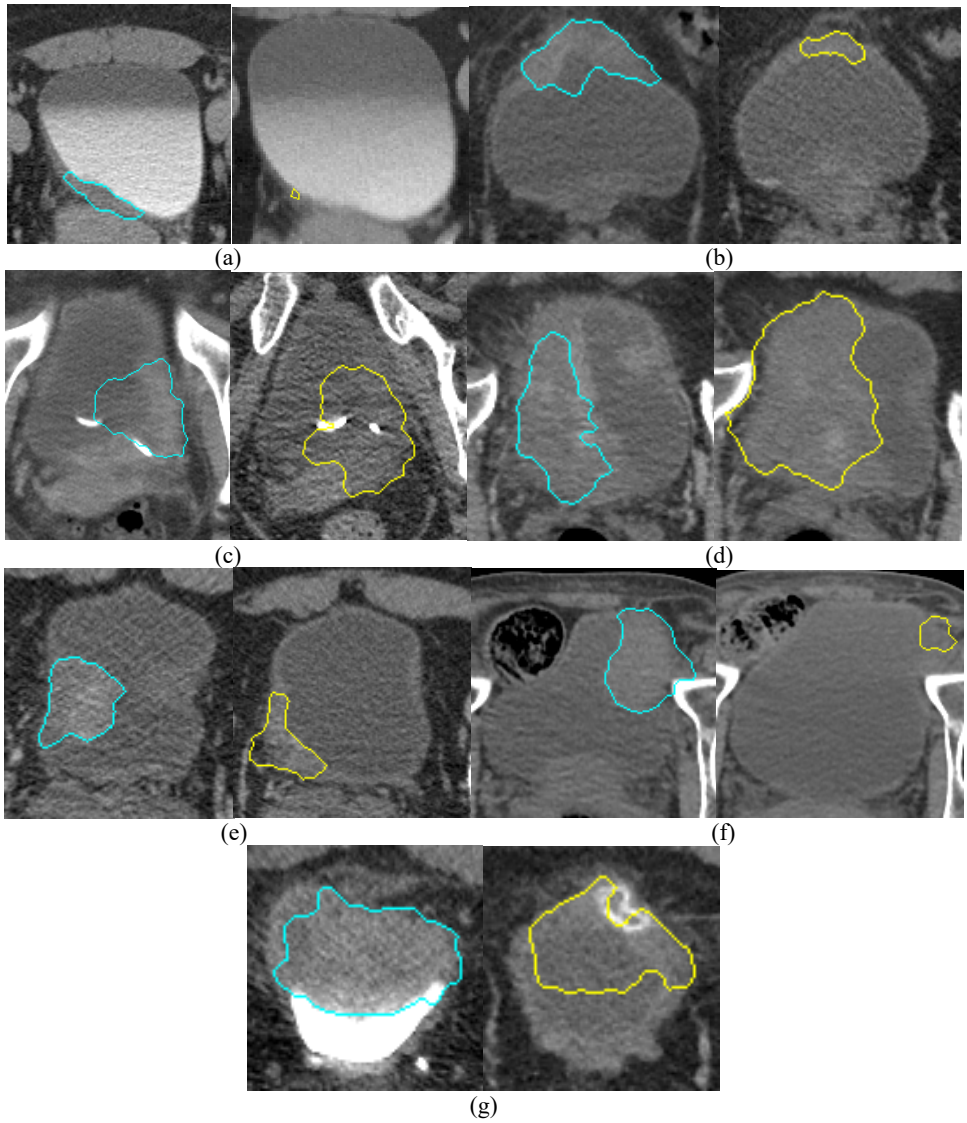


Figure 5: Examples of pre- and post- treatment lesion pairs and the response prediction by our trained models. The lesion of each scan was segmented by AI-CALS and was marked as blue (pre-treatment scan) and yellow (post-treatment scan) outline. Cases (a) and (b) were complete responders. Both the DL-CNN and the radiomics models provided correct estimates. Cases (c) and (d) were non-complete responders. Both the DL-CNN and radiomics models provided the correct estimates. Case (e) was a complete responder. Both the DL-CNN and the radiomics model estimates were wrong. Case (f) is a non-complete responder. The DL-CNN model made the correct estimate but the radiomics model provided the wrong estimate. Case (g) is a non-complete responder. The estimate of the radiomics model was correct but the estimate of the DL-CNN model was wrong.

4. CONCLUSION

This study demonstrated the feasibility of developing DL-CNN model and radiomics model in estimation of bladder cancer treatment response in CTU. The combined model showed improvement, although small, in performance compared to the DL-CNN model or the radiomics model alone, indicating that the deep features and radiomics features have complementary information and combining them has advantage. Further study is underway to improve the models.

ACKNOWLEDGMENT

This work is supported by National Institutes of Health Grant number U01-CA232931.

REFERENCES

[1] American cancer society. Cancer facts & figures 2022. American Cancer Society: Atlanta, GA, USA, 2022.
[2] Fagg, S.L.; Dawson-Edwards, P.; Hughes, M.; Latief, T.; Rolfe, E.; Fielding, J., "Cis-diamminedichloroplatinum (ddp) as initial treatment of lnvasive bladder cancer," British journal of urology 56: 296-300 (1984).
[3] Raghavan, D., B. Pearson, G. Coorey, D. Arnold, J. Smith, J. Donovan, P. Langdon and W. Woods., "Intravenous cis-platinum for invasive bladder cancer: Safety and feasibility of a new approach," Medical Journal of Australia 140: 276-78 (1984).
[4] Meeks, J. J., J. Bellmunt, B. H. Bochner, N. W. Clarke, S. Daneshmand, M. D. Galsky, N. M. Hahn, S. P. Lerner, M. Mason and T. Powles., "A systematic review of neoadjuvant and adjuvant chemotherapy for muscle-invasive bladder cancer," European urology 62: 523-33 (2012).
[5] Witjes, J., M. Wullink, G. Oosterhof and P. De Mulder., "Toxicity and results of mvac (methotrexate, vinblastine, adriamycin and cisplatin) chemotherapy in advanced urothelial carcinoma," European urology 31: 414-19 (1997).
[6] Hadjiiski, L., Chan, H.P., Caoili, E.M., Cohan, R.H., Wei, J. and Zhou, C., "Auto-initialized cascaded level set (AI-CALS) segmentation of bladder lesions on multidetector row CT urography," Academic Radiology 20.2: 148-155 (2013).
[7] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M. and Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., and Zheng, X., " Tensorflow: a system for large-scale machine learning," In Osdi 16: 265-283 (2016).
[8] Sun, D., Hadjiiski, L., Alva, A., Zakharia, Y., Joshi, M., Chan, H.P., Garje, R., Pomerantz, L., Elhag, D., Cohan, R.H., Caoili, E.M., Kerr, W.T., Cha, K.H., Kirova-Nedyalkova, G., Davenport, M.S., Shankar, P.R., Francis, I.R., Shampain, K., Meyer, N., Barkmeier, D., Woolen, S., Palmbos, P.L., Weizer, A.Z., Samala, R.K., Zhou, C., and Matuszak, M., "Computerized Decision Support for Bladder Cancer Treatment Response Assessment in CT Urography: Effect on Diagnostic Accuracy in Multi-Institution Multi-Specialty Study," Tomography 8.2: 644-656 (2022).