IPC

Indonesian Physics Communication

Department of Physics, Universitas Riau, Pekanbaru, Indonesia
p-ISSN 1412-2960 | e-ISSN 2579-521X

Vol. 21 | No. 2
July 2024

Web: https://kfi.ejournal.unri.ac.id
E-mail: kfi@ejournal.unri.ac.id

Random forest algorithm for precision dose prediction in brain cancer radiotherapy

Luqyana Adha Azwat*, Prawito Prajitno, Dwi Seno Kuncoro Sihono, Dewa Ngurah Yudhi Persada

[1]Department of Physics, Universitas Indonesia, Depok 16424, Indonesia
[2]Department of Physics, MRCCC Siloam Hospitals Semanggi, South Jakarta 12930, Indonesia

*Corresponding author: luqyana.adha@ui.ac.id

ABSTRACT

Improving dose optimization during clinical planning using the treatment planning system for radiotherapy patients is crucial, yet executing this process can be time-consuming and reliant on the expertise of medical physicists. This research focuses on dose prediction employing machine learning for the planning target volume (PTV) and organ at risk (OAR) in cases of brain cancer treated with the volumetric modulated arc therapy planning technique. Utilizing DICOM planning data from brain cancer cases, this study utilizes extracted radiomic and dosiomic values as inputs and outputs for the research, employing a random forest algorithm model. Evaluation of the model reveals its effectiveness in predicting doses for PTV in brain cancer and OAR, with predicted homogeneity index and conformity index values of 0.14 ± 0.04 and 0.95 ± 0.01, respectively, compared to clinical values of 0.14 ± 0.13 and 0.94 ± 0.13. Thus, the random forest model demonstrates proficiency in predicting doses for brain cancer PTV and OAR, with an mean square error value of 0.017.

Keywords: Mean square error; OAR; PTV; p-value; random forest

Received 08-05-2024 | Revised 10-06-2024 | Accepted 24-06-2024 | Published 26-07-2024

INTRODUCTION

Brain cancer significantly impacts public health in Indonesia, with 5,964 cases reported in 2020 according to Global Burden of Cancer data [1]. Although not the most common cancer, it requires attention due to the brain's role as the central nervous system, controlling the body's movements and necessitating the protection of many healthy organs during treatment [2]. Radiotherapy, including the advanced technique volumetric modulated arc therapy (VMAT), is an effective treatment for brain cancer. VMAT delivers precise radiation doses to the tumor while minimizing damage to surrounding healthy tissue, offering benefits in speed, precision, and efficiency compared to conventional methods [3]. Proper radiation therapy planning is essential for effectiveness and minimizing damage to healthy tissue [4].

The treatment planning system designs optimal radiation plans, ensuring accurate targeting and minimal impact on healthy tissue. To enhance planning quality and efficiency, artificial intelligence technologies like knowledge based planning with machine learning, including random forest algorithms, are used for accurate dose prediction in VMAT techniques. This study will predict doses for the planned target volume (PTV) and organ at risk (OAR) using a random forest model with k-fold validation and DICOM data [5].

RESEARCH METHODS

Data and Research Tools

This study utilized 178 radiation therapy planning data in DICOM file format from brain cancer cases. The hardware used was an Asus laptop with a Ryzen 7 processor, 8 GB RAM, 500 GB SSD, and Windows 11 OS. Software utilized included3D Slicer, Python, and Microsoft Excel.

Pre-processing

Data Radiomic and dosiomic feature extraction was conducted using 3D Slicer. For dosiomic, features such as its shape were extracted, involving dose analysis foreach organ obtained from DICOM dose data. Dosiomic data included V_{95}, V_{PTV}, $D_{2\%}$, D50%, $D_{98\%}$, D_{mean}, and D_{max}. The organs involved in this study were the PTV and its OAR, the brainstem. Dosiomic data were normalized with the prescriptiondose to ensure scale consistency, then merged into a CSV format.

Build the Random Forest Model

Radiomic and dosiomic data were imported into Python for training and testing.70% of the data were used for training, and 30% for testing. The prediction model utilized radiomic features to compute dosiomic data during testing, with prediction results derived from the average value of all estimators [6].

Evaluating the Random Forest Model

Evaluation involved several steps. First, assessing error values in the validation stage to gauge the fit of the random forest model with actual clinical data, yielding the mean squared error (MSE). Subsequently, testing the planning quality using homogeneity index (HI) and conformity index (CI) parameters with the following formulas:

$$HI = \frac{D_{2\%} - D_{98\%}}{D_{50\%}} \quad (1)$$

$$CI = \frac{V_{95}}{V_{PTV}} \quad (2)$$

RESULTS AND DISCUSSION

During the pre-processing stage, radiomic and dosiomic data were obtained. Dosiomic data extracted underwent normalization against the prescription dose of each patient. This normalization process ensures data within the same value range.

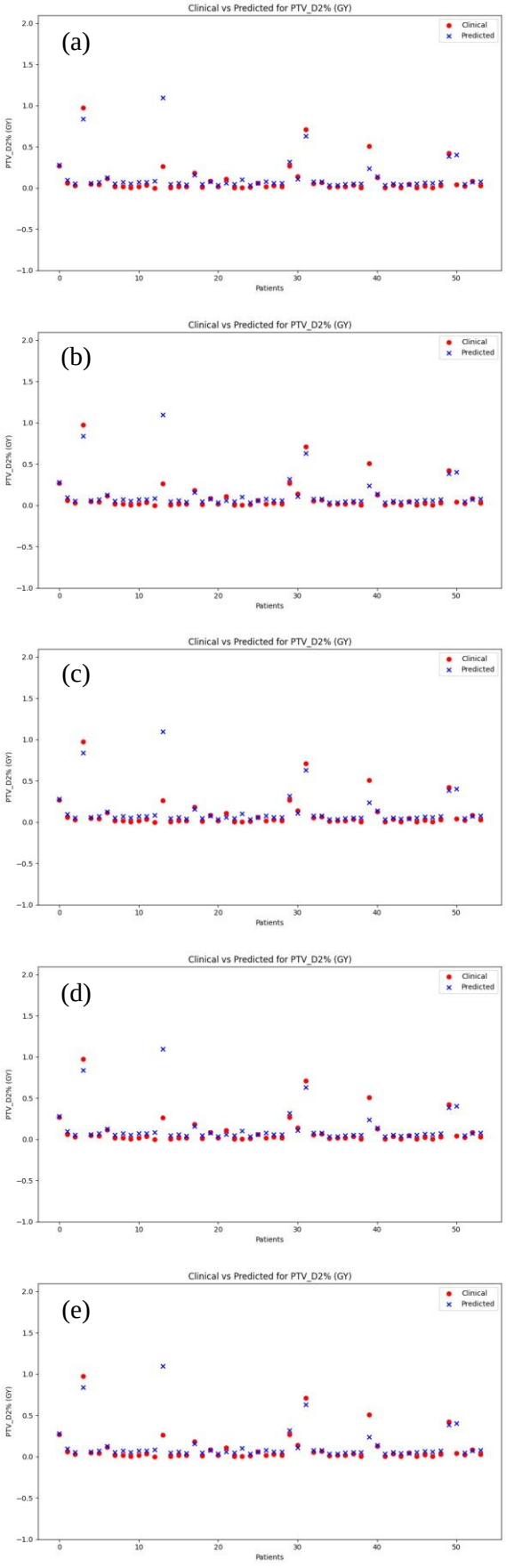

Figure 1. Clinical versus predicted for: (a) PTV $D_{2\%}$; (b)PTV $D_{50\%}$; (c) PTV $D_{98\%}$; (d) brainstem mean; and (e) brainstem max.

Normalization exclusively applied to dosiomic data, as radiomic data did not require normalization due to their association with organ shape information, necessitating consistency with the original data. Out of 124 data points, 70% were allocated for training and 30% for testing. The feature values from the machine learning training data influenced the regression outcomes in this random forest method. The results indicate clinical and predicted outcomes for PTV and OAR brainstem as follows in Figure 1.

From Figure 1, it can be observed that the overall prediction results are quite close to the clinical values, with an average MSE value of 0.0177. MSE is used to evaluate the model's performance, where a smaller MSE value indicates that the model has predictions closer to the actual values. It can be concluded that the machine learning model's performance is capable of predicting PTV and OAR organs well, resulting in an MSE error value very close to 0 [7].

In the radiation therapy planning stage, important parameters to consider are the conformity of the dose distribution with the target shape, determined by the CI value, and the homogeneity of the dose distribution within the target volume, determined by the HI value. With clinical HI and CI values of 0.14±0.13 and 0.94 ± 0.13, respectively, while the predicted values are 0.14 ± 0.04 and 0.95 ± 0.01, according to ICRU 83 recommendations [8], the HI value should be 0, and for CI, according to ICRU 62 recommendations [9], the CI value should be 1.

CONCLUSION

The Random Forest model accurately predicts dose distribution with an error value of 0.017 and shows no significant difference between clinical outcomes and predicted results. As a result, this Random Forest model can be used as a reference to streamline the radiation therapy planning process for medical physicists.

REFERENCES

1. Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A. & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, **71**(3), 209–249.
2. Abdullah, F., Ganiem, A. R., Rasyid, A., Munir, B. & Wiratman, W. (2023). Pedoman praktik klinik neurologi 2023, Department of Neurology.
3. Verbakel, W. F., Cuijpers, J. P., Hoffmans, D., Bieker, M., Slotman, B. J. & Senan, S. (2009). Volumetric intensity-modulated arc therapy vs. conventional IMRT in head-and-neck cancer: a comparative planning and dosimetric study. *International Journal of Radiation Oncology*Biology*Physics*, **74**(1), 252–259.
4. Van Herk, M., Remeijer, P., Rasch, C. & Lebesque, J. V. (2000). The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. *International Journal of Radiation Oncology*Biology*Physics*, **47**(4), 1121–1135.
5. Savargiv, M., Masoumi, B., & Keyvanpour, M. R. (2021). *A new random forest algorithm based on learning automata*. *Computational Intelligence and Neuroscience*, **2021**(1), 5572781.
6. Purwanto, A. D., Wikantika, K., Deliar, A., & Darmawan, S. (2022). Decision tree and random forest classification algorithms for mangrove forest mapping in Sembilang National Park, Indonesia. *Remote Sensing*, **15**(1), 16.
7. Song, Y., Hu, J., Liu, Y., Hu, H., Huang, Y., Bai, S., & Yi, Z. (2020). Dose prediction using a deep neural network for accelerated planning of rectal cancer radiotherapy. *Radiotherapy and Oncology*, **149**, 111–116.
8. Menzel, H. G. (2012). The international commission on radiation units and

measurements. *Journal of the ICRU*, **12**(2), 1–2.
9. International Commission on Radiological Units. (1961). *Report of the International Commission on Radiological Units and Measurements (ICRU)*, 1959 (Vol. 78). US Department of Commerce, National Bureau of Standards.