Prediction of Radiotherapy Dose Distribution for Glioblastoma Cancer Using Convolutional Neural Network Model

.

# ABSTRACT

|  |
| --- |
| **Aims:** This study aims to predict radiotherapy dose distribution for glioblastoma patients using Machine Learning with a Convolutional Neural Network (CNN) model.  **Study design:** This research used an experimental design with a quantitative approach to predict radiotherapy dose distribution for glioblastoma patients using a CNN model. The study involved training and testing the CNN model on medical imaging data from The Cancer Imaging Archive (TCIA), evaluating its performance based on Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Structural Similarity Index Measure (SSIM). The results were analyzed to determine the model’s accuracy in replicating actual dose distributions, providing a data-driven assessment of its predictive capability.  **Place and Duration of Study:** This research was conducted in the Department of Physics at Udayana University from October 2024 to January 2025.  **Methodology:** The research involved 180 patient datasets divided into 126 training data and 54 testing data. The CNN architecture is implemented using the Google Collaboratory platform. Model evaluation is performed using MSE, RMSE, and SSIM to measure the accuracy of dose distribution prediction.  **Results:** The MSE, RMSE, and SSIM values obtained from the CNN model are 0.00015795, 0.01256, and 0.979718, respectively. The low MSE and RMSE values, which are close to 0, indicate that the model can predict radiotherapy dose maps with minimal error. Additionally, the SSIM value approaching 1 show that the CNN model's predictions are nearly identical to the actual dose distribution data. Visually, three 2D slices from the 3D data—axial, coronal, and sagittal slices—generated by the CNN model prediction exhibit patterns that closely resemble the ground truth dose maps.  **Conclusion:** The CNN model demonstrates effectiveness in predicting the dose distribution for glioblastoma radiotherapy, achieving highly accurate evaluation metrics. Low MSE and RMSE values indicate minimal prediction errors in radiotherapy dose maps, while an SSIM close to 1 suggests the CNN model's predictions closely match the actual dose distribution. Visually, the model exhibit patterns highly similar to the actual dose map. |

*Keywords: CNN, Dose Distribution, Glioblastoma, Radiotherapy, Machine Learning*

# 1. INTRODUCTION

In the human body, the brain functions as one of the most vital organs, aiding in decision-making and controlling all actions. It serves as the central regulator of the nervous system and must be protected from all types of diseases (Saha et al., 2024). Due to the increasing global population, brain cancer is among the most fatal medical conditions worldwide (Verma et al., 2023). Brain cancer is a mass that arises from the uncontrolled proliferation of brain cells and the loss of the brain's regulatory system. Approximately 700,000 people worldwide suffer from brain cancer, with 86,000 new cases identified in 2019. Nearly 16,380 people died from brain cancer in 2019 (Rehman et al., 2021). Glioblastoma is the most common malignant brain cancer, accounting for 80% of all malignant brain tumors (Montaha et al., 2023).

Glioblastoma is a type of primary brain cancer that attacks glial cells, also known as glioma. It is classified as a grade IV glioma, characterized by aggressive cell growth, necrosis (cell death), and rapid progression without prior lesions. The survival rate for this cancer is estimated to be approximately 12 months after diagnosis (Mutamimah et al., 2022). Its high tendency to spread makes it more challenging to treat (Tan et al., 2020). Common treatment methods for brain cancer patients include oncological surgery, chemotherapy, and radiotherapy. In glioblastoma cases, tumor removal surgery is sometimes ineffective due to the possibility of residual tumor tissue remaining in the brain despite the procedure (Nurwati & Prasetya, 2014). Radiotherapy is considered the most effective treatment for brain cancer cases (Mutamimah et al., 2022).

Radiotherapy, commonly used as a curative, adjuvant, or palliative treatment, is one of the most essential modalities for cancer patients, with more than 50% of them undergoing this treatment (Liu et al., 2019). It can be delivered through two methods, external radiotherapy where radiation is applied externally using a treatment machine and brachytherapy where radioactive sources are temporarily or permanently placed inside the body (Harun et al., 2022). Before radiation therapy is administered, a Treatment Planning System (TPS) is used to optimize treatment parameters such as the number of fields, beam angles, and dose distribution, ensuring that the maximum dose is delivered to the tumor while minimizing exposure to surrounding healthy organs (Winarno et al., 2021). In external radiotherapy, techniques like intensity-modulated radiotherapy (IMRT), volumetric modulated arc therapy (VMAT), and helical tomotherapy (HT) have become standard treatment approaches for various cancers (Liu et al., 2019).

With advancements in external radiotherapy techniques, treatment planning has become more complex, leading to increased planning time and variations in quality. One major challenge is inverse planning, where clinical requirements are converted into a one-dimensional dose-volume histogram (DVH). The rationality behind setting DVH objectives and their relative weights varies among medical physicists, leading to inconsistencies in planning quality. While clinical protocols and planning templates help establish default DVH goals and weights, adjusting these parameters to achieve optimal clinical outcomes remains challenging, particularly when trade-offs are required (Liu et al., 2019).

Several efforts have been made to predict radiotherapy dose distribution using machine learning (ML) by leveraging previous patient plans, structural data, or dose information to automatically predict DVH curves and dose constraints. While these methods have demonstrated predictive accuracy and improved planning consistency, they still rely on manual feature extraction for input into DVH models (Liu et al., 2019). Recently, a new paradigm has emerged in three-dimensional dose distribution prediction using deep learning, which integrates imaging results with structured and unstructured planning data. This approach utilizes connected layers that are versatile, fast-converging, and architecturally simple. However, it tends to generalize poorly when dealing with high-dimensional data. One promising alternative is the Convolutional Neural Network (CNN), which is specifically designed for image segmentation problems and can be implemented using architectures such as U-Net, V-Net, or dilated convolution networks (Kearney et al., 2018).

Patient-specific dose distribution prediction is crucial for accelerating radiotherapy planning by medical physicists, ensuring efficient treatment for glioblastoma patients. This study aims to evaluate the performance of machine learning in predicting patient-specific radiotherapy dose distribution for glioblastoma treatment, contributing to the advancement of automated dose planning in radiation oncology. Thus, this study not only builds upon previous research in dose prediction but also focuses on improving the accuracy and efficiency of radiotherapy planning, which is expected to enhance treatment precision and patient outcomes.

# 2. MATERIALS AND METHODS

## 2.1 Materials

The dataset used in this study was obtained from The Cancer Imaging Archive (TCIA) under the University of Arkansas for Medical Science. The dataset titled Burdenko-GBM-Progression Version 1 (Zolotova et al., 2023), consists of a total of 180 patients, with 126 cases used for training and 54 cases for testing. It includes MRI and CT images, organ contour structures, radiotherapy planning data, and patient dose distribution maps. In this study, MRI/CT images with organ contours (RTSTRUCT) serve as the input, while radiotherapy dose distribution (RTDOSE) is used as the label, forming the foundation for the CNN model to be developed.

## 2.2 Methods

### 2.2.1 Data Preprocessing

Efficient data loading and preprocessing are essential for training a CNN-based model for radiotherapy dose estimation. The dataset comprises DICOM files, including RTSTRUCT (radiotherapy contours) and RTDOSE (dose distribution), stored in a structured directory. Initially, data is split (70-30%) for training and testing, ensuring balanced evaluation. Given the large size of 3D medical images, a data generator loads small batches to prevent memory overflow. Preprocessing begins with contour extraction from RTSTRUCT, identifying critical regions of interest (ROIs) like tumors and organs at risk. Dose maps from RTDOSE undergo normalization (0-1 scaling) to maintain numerical consistency and improve model stability. The data is then resized using trilinear interpolation and reformatted to (100, 100, 100, 1) for CNN compatibility. The preprocessed data is fed into the model via the data generator, ensuring efficient training while preserving anatomical structures.

### 2.2.2 CNN Model

The model consists of multiple 3D convolutional layers (Conv3D) designed to extract spatial features from radiotherapy dose and structure images. It begins with an input layer (100, 100, 100, 1), followed by a Conv3D layer with 32 filters (3×3×3 kernel), ReLU activation, and ‘same’ padding to preserve spatial dimensions. A MaxPooling3D layer (2×2×2) then reduces feature dimensions to enhance computational efficiency. The second Conv3D layer increases to 64 filters, followed by another MaxPooling3D layer. The third Conv3D layer further increases to 128 filters, capturing more complex patterns. To reconstruct spatial information, UpSampling3D layers restore feature resolution, followed by Conv3D layers reducing filters to 64, and a final Conv3D layer with one filter and linear activation for dose map prediction. The model is compiled using the Adam optimizer with mean squared error (MSE) loss, as it predicts numerical dose values. Training is managed with a data generator, allowing efficient memory use while handling large datasets. The dataset is split 70:30 for training and testing, and early stopping prevents overfitting by monitoring val\_loss, stopping training if no improvement occurs for eight consecutive epochs. Post-training, the model is evaluated on a test batch, comparing predictions with ground truth using MSE, RMSE, and SSIM to assess structural similarity. Finally, predicted dose maps are visualized in 2D slices to provide insights into the model’s accuracy and performance. The Structural Similarity Index Measure is designed to assess visual quality and structural similarity between two images. The SSIM formula captures three key aspects: luminance, contrast, and structure. The overall SSIM value is computed by averaging SSIM scores across multiple local windows within the image, producing a score between -1 and 1, where a value closer to 1 indicates a higher similarity (Liang et al., 2021).

### 2.2.3 Quantitative Evaluations

The performance of the machine learning model is evaluated using metrics such as MSE, RMSE, and SSIM. Mean Squared Error is the average of the squared differences between predicted and actual values. MSE is crucial for identifying outliers and data imbalances within the dataset. Since MSE is always non-negative, a lower value indicates better model fit. A smaller MSE implies a smaller trade-off between bias and variance, improving prediction accuracy (Hodson, 2022). The MSE value can be calculated using the following equation.

Where are actual value are predicted value, and *n* represent amount of data. Root Mean Squared Error is the square root of MSE, restoring the error unit to the original scale of the target variable. While RMSE is still influenced by outliers, it is less sensitive compared to MSE. Due to this characteristic, RMSE is useful for detecting outliers in machine learning models. It is also intuitive since it is expressed in the same units as the predicted variable, making it a widely used metric for model comparison. In other words, when comparing multiple models and algorithms, the one with the lowest RMSE is considered the most accurate (Hodson, 2022). The RMSE value can be calculated using the following equation.

The Structural Similarity Index Measure is designed to assess visual quality and structural similarity between two images. The SSIM formula captures three key aspects, luminance, contrast, and structure. The overall SSIM value is computed by averaging SSIM scores across multiple local windows within the image, producing a score between -1 and 1, where a value closer to 1 indicates a higher similarity (Liang et al., 2021). The SSIM value can be calculated using the following equation.

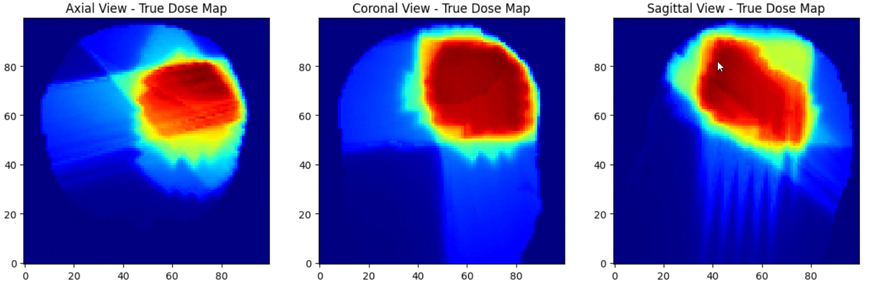
Where ​ and ​ represent the mean intensity of images x and y which represent luminance, ​ and ​ denote the variance of images x and y which represent contrast, ​ is the covariance between x and y indicating structural similarity. Additionally, ​ and ​ are small constants added to stabilize the formula, particularly when the mean or variance is close to zero. are typically defined as and are typically defined as , where is the dynamic range of pixel values (L=255 for 8-bit images), also ​ and ​ are small constants (commonly set to dan ).

# 3. RESULTS AND DISCUSSION

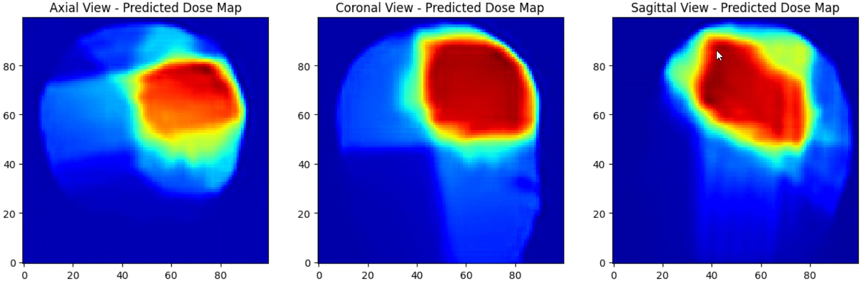
In this research, the CNN model training process was repeated to achieve an optimal level of accuracy. The training time ranged between 5 to 8 hours, with an average duration of 1,247 seconds per epoch. The CNN model also implemented the early stopping technique with a patience value of 8, causing the training process to automatically stop, with the best weights appearing between epochs 15 and 21. This approach prevents the model from overfitting to the training data, ensuring good generalization on new data while reducing unnecessary epochs and accelerating the training process. From these iterations, the best result was obtained with an accuracy of 97.53%. Since the CNN model in this study is solely used to predict the radiotherapy dose distribution for glioblastoma cancer without classification, accuracy is defined as the percentage of pixels where the difference between the predicted and actual values is smaller than the error threshold. Accuracy is calculated by dividing the number of pixels meeting this condition by the total number of pixels in the image, using an error threshold of 0.01.

The CNN model evaluation results showed that the MSE, RMSE, and SSIM values were 0.00015795, 0.01256, and 0.979718 respectively. The MSE and RMSE values being very close to zero indicate that the model can predict radiotherapy dose maps with minimal error. MSE measures the average squared difference between the predicted and actual values, while RMSE provides a more intuitive representation as it is in the same scale as the original data. In this context, an RMSE value of 0.01256 signifies that the average prediction error is extremely small, demonstrating the model’s strong performance in reconstructing dose distribution. Furthermore, an SSIM value of 0.979718 indicates a very high structural similarity between the model’s predictions and the actual dose distribution. SSIM evaluates structural similarity between two images by considering luminance, contrast, and structure. An SSIM value close to 1 suggests that the model’s predictions are nearly identical to the actual data in terms of dose distribution patterns. This is crucial in radiotherapy, as small differences in dose distribution can significantly impact treatment effectiveness. Overall, these evaluation results confirm that the CNN model performs exceptionally well in predicting radiotherapy dose maps, achieving low error rates and high structural similarity.

In addition to evaluation metrics such as MSE, RMSE, and SSIM, the training results also include visualizations of the predicted radiotherapy dose distribution for glioblastoma cancer. These visualizations consist of three 2D slices from the 3D data, namely axial, coronal, and sagittal slices. Visually, the ground truth dose distribution map, shown in Figure 1, can be compared with the CNN model’s predicted dose distribution map, presented in Figure 2.

****

**Fig. 1. Ground truth dose distribution in axial slice, coronal slice, and sagittal slice**

****

**Fig. 2. Predicted dose distribution in axial slice, coronal slice, and sagittal slice**

From Figure 1 and Figure 2, it can be observed that the dose distribution predicted by the CNN model closely resembles the actual dose map. The contours of high and low dose regions appear well-aligned, indicating that the CNN model has effectively learned the dose distribution patterns. Additionally, the predicted results show a high level of similarity to the ground truth, as seen in the axial, coronal, and sagittal slices.

# 4. CONCLUSION

The CNN model demonstrated excellent performance based on several evaluation metrics, including accuracy, MSE, RMSE, and SSIM. The high accuracy indicates that the model predicts with minimal error. The low MSE and RMSE values suggest that the average difference between the predicted values and the ground truth is relatively small, while the high SSIM value confirms a strong spatial similarity between the predicted and actual dose distributions. Additionally, visualization results show that the predicted dose distribution closely resembles the ground truth, confirming that the model successfully captures relevant patterns in the data. In conclusion, the CNN model proves to be effective in predicting dose distribution for glioblastoma radiotherapy, achieving highly accurate evaluation results.

# CONSENT

All authors declare that ‘written informed consent was obtained from the patient (or other approved parties) for publication of this case report and accompanying images. A copy of the written consent is available for review by the Editorial office/Chief Editor/Editorial Board members of this journal.

# ETHICAL APPROVAL

All authors hereby declare that all experiments have been examined and approved by the appropriate ethics committee and have therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

# REFERENCES

Harun, H. M., Jannah, N., Idawati, & Ahmad, Z. F. (2022). Evaluasi pengobatan radioterapi pada pasien kanker. *Journal Syifa Sciences and Clinical Research (JSSCR)*, *4*(3), 662–670. https://ejurnal.ung.ac.id/index.php/jsscr/article/view/15794

Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development*, *15*(14), 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022

Kearney, V., Chan, J. W., Haaf, S., Descovich, M., & Solberg, T. D. (2018). DoseNet: A volumetric dose prediction algorithm using 3D fully-convolutional neural networks. *Physics in Medicine and Biology*, *63*(23). https://doi.org/10.1088/1361-6560/aaef74

Liang, X., Nguyen, D., & Jiang, S. B. (2021). Generalizability issues with deep learning models in medicine and their potential solutions: illustrated with cone-beam computed tomography (CBCT) to computed tomography (CT) image conversion. *Machine Learning: Science and Technology*, *2*(1), 015007. https://doi.org/10.1088/2632-2153/abb214

Liu, Z., Fan, J., Li, M., Yan, H., Hu, Z., Huang, P., Tian, Y., Miao, J., & Dai, J. (2019). A deep learning method for prediction of three-dimensional dose distribution of helical tomotherapy. *Medical Physics*, *46*(5), 1972–1983. https://doi.org/10.1002/mp.13490

Montaha, S., Azam, S., Rakibul Haque Rafid, A. K. M., Hasan, M. Z., & Karim, A. (2023). Brain Tumor Segmentation from 3D MRI Scans Using U-Net. *SN Computer Science*, *4*(4), 1–10. https://doi.org/10.1007/s42979-023-01854-6

Mutamimah, R., Susilo, & Sardjono, Y. (2022). Aplikasi Program PHITS Versi 3.21 untuk Analisis Dosis Radiasi pada Terapi Kanker Otak dengan Metode Proton Therapy. *Unnes Physics Education Journal*, *11*(1), 26–35. http://journal.unnes.ac.id/sju/index.php/upej

Nurwati, S., & Prasetya, R. I. (2014). Prosiding Pertemuan dan Presentasi Ilmiah-Penelitian Dasar Ilmu Pengetahuan dan Teknologi Nuklir. *Pusat Sains Dan Teknologi Akselerator-BATAN Yogyakarta*, *6*, 10–11.

Rehman, A., Khan, M. A., Saba, T., Mehmood, Z., Tariq, U., & Ayesha, N. (2021). Microscopic brain tumor detection and classification using 3D CNN and feature selection architecture. *Microscopy Research and Technique*, *84*(1), 133–149. https://doi.org/10.1002/jemt.23597

Saha, P., Das, S. K., & Das, R. (2024). A Review on Machine Learning and Deep Learning Based Systems for the Diagnosis of Brain Cancer. *SN Computer Science*, *5*(1). https://doi.org/10.1007/s42979-023-02360-5

Tan, A. C., Ashley, D. M., López, G. Y., Malinzak, M., Friedman, H. S., & Khasraw, M. (2020). Management of glioblastoma: State of the art and future directions. *CA: A Cancer Journal for Clinicians*, *70*(4), 299–312. https://doi.org/10.3322/caac.21613

Verma, A., Gupta, N., Bhatele, P., & Khanna, P. (2023). JMCD Dataset for Brain Tumor Detection and Analysis Using Explainable Deep Learning. *SN Computer Science*, *4*(6). https://doi.org/10.1007/s42979-023-02308-9

Winarno, Nurmansya, V. A., & Miskiyah, Z. (2021). Radioterapi Kanker Cervix Dengan Linear Accelerator (LINAC). *Jurnal Biosains Pascasarjana*, *23*(2), 75. https://doi.org/10.20473/jbp.v23i2.2021.75-86

Zolotova, S. V., Golanov, A. V., Pronin, I. N., Dalechina, A. V., Nikolaeva, A. A., Belyashova, A. S., Usachev, D. Y., Kondrateva, E. A., Druzhinina, P. V., Shirokikh, B. N., Saparov, T. N., Belyaev, M. G., & Kurmukov, A. I. (2023). Burdenko’s Glioblastoma Progression Dataset (Burdenko-GBM-Progression) (Version 1) [Data set]. The Cancer Imaging Archive. <https://doi.org/10.7937/E1QP-D183>