Prediction of Radiotherapy Dose Distribution for Glioblastoma 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), Structural Similarity Index Measure (SSIM), Dice Similarity Coefficient (DSC), Peak Signal to Noise Ratio (PSNR), Normalized Cross-Correlation (NCC). 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, SSIM, DSC, PSNR, and NCC values obtained from the CNN model are 0.00015795, 0.01256, 0.979718, 0.9711, 32dB, and 0.96289 respectively. The low MSE and RMSE values indicate minimal prediction error, while the high SSIM confirms strong structural similarity between the predicted and actual dose maps. The DSC demonstrates excellent spatial overlap, and the high PSNR reflects high-quality dose reconstruction. Additionally, the NCC highlights strong correlation with the ground truth. Visually, the axial, coronal, and sagittal slices closely resemble the actual dose distributions, further validating the model’s accuracy.  **Conclusion:** The CNN model demonstrates effectiveness in predicting the dose distribution for glioblastoma radiotherapy, achieving highly accurate evaluation metrics. 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 increase in the number and aging of the world's population, the incidence of brain cancer has increased. Glioblastoma is one of the deadliest diseases 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 (Rehman et al., 2021). Approximately 1 million individuals in the United States are currently living with a primary brain tumor (National Brain Tumor Society, 2024). In 2023, approximately 24,810 new cases of brain cancer were diagnosed in the United States, with an estimated 18,990 related fatalities (Siegel et al., 2023). Glioblastoma is the most common malignant brain cancer, accounting for 50% of all malignant brain tumors (National Brain Tumor Society, 2024).

Glioblastoma is a type of primary brain cancer, also known as malignant diffuse glioma of high grade. It is classified as a grade 4 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). The high tendency to diffuse growth in brain tissue makes treatment more difficult (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). Radiation therapy in combination with temozolomide chemotherapy is considered the most effective treatment for brain cancer (Barker et al., 2012).

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. A restricted license agreement has been implemented to ensure compliance with regulations and ethical standards set by the dataset provider for the use of this data. 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. In general, a summary of radiotherapy dose distributions among 180 patients is presented in Table 1.

**Table 1. Radiotherapy dose summary statistics**

|  |  |  |  |
| --- | --- | --- | --- |
| **Metric** | **Lowest** | **Highest** | **Average** |
| Mean Dose (Gy) | 0.128024 | 9.399805 | 4.748748 |
| Max Dose (Gy) | 2.168881 | 82.261153 | 61.005860 |
| Dose Coverage (%) | 22.957565 | 54.255686 | 40.219731 |

## 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. Considering the variability in image resolution of RTDOSE, all dose maps are resampled to a uniform shape of (128, 128, 128). To ensure consistency across samples, the dose values are normalized between 0 and 1 using min-max scaling. This normalization step contributes to improved model stability by maintaining consistent numerical distributions across the dataset. To enhance generalization and mitigate overfitting, data augmentation techniques are applied during training. These augmentations include random flipping along the three spatial dimensions and small-angle rotations (±10°) around the sagittal, coronal, or axial planes. 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 CNN model is designed to process 3D volumetric medical images and predict radiotherapy dose distributions with high precision. The architecture comprises several key components. Convolutional layers are employed to extract spatial features from input images, while an attention mechanism, implemented through a custom AttentionLayer with a Conv3D attention filter, enhances the model’s ability to focus on critical dose regions. Max-pooling layers are used to progressively downsample feature maps, whereas upsampling layers restore spatial resolution in the final dose prediction. The model concludes with a Conv3D output layer, utilizing a linear activation function to generate the final predicted dose distribution. To optimize performance, the model is compiled using the Adam optimizer and trained with a custom weighted loss function. This loss function assigns higher penalties to dose regions exceeding a predefined threshold (0.1 Gy), ensuring more accurate dose estimation in clinically significant areas.

To ensure robust model evaluation, a 5-fold cross-validation strategy is implemented. The dataset is partitioned into training and testing subsets in each fold, allowing for comprehensive performance assessment. Multiple evaluation metrics are used to quantify the accuracy of dose predictions. MSE and RMSE are employed to measure absolute prediction errors, while the SSIM evaluates the perceptual similarity between predicted and ground-truth dose maps. The Dice Score assesses the spatial overlap between predicted and actual dose regions, and the PSNR provides insight into the quality of dose predictions. Additionally, NCC quantifies the correlation between predicted and true dose distributions, while a custom accuracy metric calculates the percentage of pixels falling within a predefined error threshold. To facilitate interpretation of the model’s predictions, dose distributions are visualized across the axial, coronal, and sagittal planes.

### 2.2.3 Quantitative Evaluations

The performance of the machine learning model is evaluated using metrics such as MSE, RMSE, SSIM, DSC, PSNR, and NCC. 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 ).

Dice Score or Dice Similarity Coefficient (DSC) is the main validation metric of spatial overlap index. The DSC measures the spatial overlap between two segmentations (Zou et al., 2004), where A and B target regions and ∩ is the intersection, it can be calculated using the following equation.

Peak Signal to Noise Ratio (PSNR) is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation (Ambika et al., 2022). The unit of PSNR is decibel and it can be calculated using the following equation.

Normalized Cross-Correlation (NCC) is a measure of similarity of two series or comparison of the processed image and reference image (Ambika et al., 2022) and it can be calculated using the following equation.

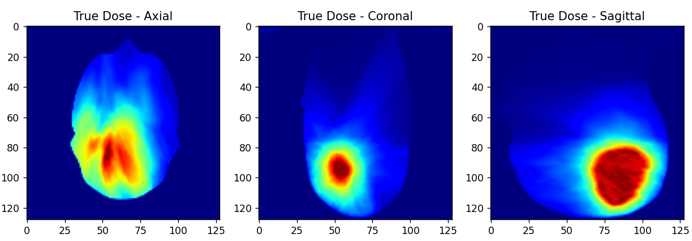
# 3. RESULTS AND DISCUSSION

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. 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.

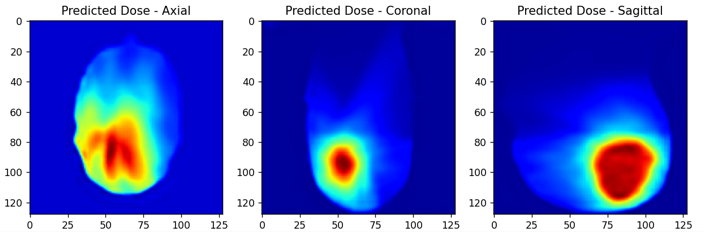
The evaluation metrics for the CNN model demonstrate strong predictive accuracy in radiotherapy dose distribution. The model achieved MSE of 0.00015795 and RMSE of 0.01256, both of which indicate minimal deviations between predicted and actual dose values. The low RMSE suggests that the model effectively captures dose variations with high numerical accuracy, ensuring precise dose estimation. Additionally, the SSIM was 0.979718, highlighting that the predicted dose distributions closely resemble the actual dose maps in terms of luminance, contrast, and structure. SSIM values near 1.0 are desirable in radiotherapy applications as they ensure that anatomical structures and dose gradients are accurately preserved.

The DSC of 0.9711 further supports the model’s accuracy by measuring the spatial overlap between predicted and actual dose regions. A DSC value close to 1.0 indicates a high degree of correspondence, which is essential in ensuring that critical structures receive the planned dose while avoiding unnecessary exposure. Compared to the study by (Irannejad et al., 2024), which reported DSC values in the average of 0.842 for deep learning-based dose prediction models, the proposed CNN model demonstrates a competitive or slightly improved performance in spatial dose matching. Furthermore, the Peak Signal-to-Noise Ratio (PSNR) of 32 dB suggests high fidelity in dose prediction, where higher PSNR values indicate that the predicted dose maps contain minimal noise and maintain high-quality reconstruction. Typical PSNR values reported in prior research vary between 30–35 dB, depending on dataset quality and preprocessing techniques. The Normalized Cross-Correlation (NCC) of 0.96289 reinforces this finding, showing a strong correlation between predicted and actual dose distributions. NCC values close to 1.0 suggest that the overall dose patterns remain intact, aligning with previously reported NCC values of 0.95–0.97 in dose prediction studies.

Overall, these results confirm the effectiveness of the proposed CNN model in accurately predicting radiotherapy dose distributions. The high SSIM, DSC, PSNR, and NCC values indicate strong spatial agreement and structural preservation, making this approach highly applicable in clinical settings for automated dose planning. Future studies could explore further optimization strategies, such as transfer learning or hybrid architectures, to enhance prediction robustness across diverse patient datasets. In addition to evaluation metrics, 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.

The inherently imbalanced nature of radiotherapy dose distributions can bias models toward over-predicting common dose levels while failing to accurately estimate rarer but clinically significant dose regions. On the computational side, training CNN models on high-dimensional 3D medical images requires substantial memory and processing power, necessitating specialized hardware such as high-performance GPUs or TPUs. The need for large-scale 3D convolutions, deep network architectures, and extensive hyperparameter tuning further exacerbates computational demands, making model training both time-consuming and resource-intensive.

# 4. CONCLUSION

The model demonstrates high accuracy with minimal error, as evidenced by its strong performance in metrics such as SSIM, DSC, PSNR, and NCC. Although this study focuses on predicting dose distributions directly from RTDOSE data, the proposed method is generalizable and can be adapted to other dose prediction tasks, including different calculation algorithms and treatment modalities. The ability to generate high-quality dose predictions in real time has significant clinical implications, particularly in adaptive radiotherapy and automated treatment planning. Future research can explore the integration of this approach with Monte Carlo-based simulations or hybrid deep learning models to further enhance accuracy and generalizability. The use of deep learning in dose prediction represents a promising avenue for advancing precision radiotherapy and improving patient outcomes.

# CONSENT

All authors declare that ‘written informed consent was obtained from the patient (or other approved parties).

# 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.

**Disclaimer (Artificial intelligence)**

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

# REFERENCES

Ambika, Biradar, R. L., & Burkpalli, V. (2022). Encryption-based steganography of images by multiobjective whale optimal pixel selection. *International Journal of Computers and Applications*, *44*(12), 1140–1149. https://doi.org/10.1080/1206212X.2019.1692442

Barker, C. A., Chang, M., Chou, J. F., Zhang, Z., Beal, K., Gutin, P. H., & Iwamoto, F. M. (2012). Radiotherapy and concomitant temozolomide may improve survival of elderly patients with glioblastoma. *Journal of Neuro-Oncology*, *109*(2), 391–397. https://doi.org/10.1007/s11060-012-0906-4

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

Irannejad, M., Abedi, I., Lonbani, V. D., & Hassanvand, M. (2024). Deep-neural network approaches for predicting 3D dose distribution in intensity-modulated radiotherapy of the brain tumors. *Journal of Applied Clinical Medical Physics*, *25*(3), 1–10. https://doi.org/10.1002/acm2.14197

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

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

National Brain Tumor Society. (2024). *Brain Tumor Facts*. https://braintumor.org/brain-tumors/about-brain-tumors/brain-tumor-facts/

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

Siegel, R. L., Miller, K. D., Wagle, N. S., & Jemal, A. (2023). Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, *73*(1), 17–48. https://doi.org/10.3322/caac.21763

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

Zou, K. H., Warfield, S. K., Bharatha, A., Tempany, C. M. C., Kaus, M. R., Haker, S. J., Wells, W. M., Jolesz, F. A., & Kikinis, R. (2004). Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index. *Academic Radiology*, *11*(2), 178–189. https://doi.org/10.1016/S1076-6332(03)00671-8

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>