

Machine Learning Approach Detects Brain Tumor Boundaries

Data from thousands of patients with glioblastoma were used to develop an accurate model for detecting tumor boundaries.

December 14, 2022 By Vicki Contie and National Institutes of Health

Glioblastoma is an aggressive and hard-to-treat type of brain cancer. It's the most common type of brain cancer in adults. But because it affects fewer than 10 in 100,000 people each year, it's considered to be a rare disease.

Defining the boundaries of glioblastoma tumors is important for treatment. One key region represents the breakdown of the blood-brain barrier inside the tumor. Another, called the tumor core, could be relevant for surgical removal. It is also typically measured to assess treatment response. A third region, the whole tumor, represents infiltrated tissue that might be treated with radiation. Identifying these regions with precision can be difficult, especially in facilities without many cases of the disease.

Despite years of progress in understanding glioblastoma, survival rates have only slightly improved over the past two decades. One roadblock has been the difficulty of collecting large and diverse data sets for this rare cancer. Big data sets could potentially give new insights. But sharing such data across institutions poses challenges for patient privacy and other legal reasons.

To overcome these obstacles, a research team led by Dr. Spyridon Bakas of the University of Pennsylvania developed a method for learning from glioblastoma data across institutions worldwide. The approach is called federated machine learning. It allows institutions to collaborate on artificial intelligence and machine learning projects without sharing sensitive patient data. The findings were described in [Nature Communications](#) on December 5, 2022.

Machine learning depends on computer algorithms that are continuously improved and refined as they analyze vast numbers of data points, looking for patterns that can reliably diagnose or predict outcomes. In federated machine learning, these algorithms are trained across multiple sites or servers. With this approach, there is no need for institutions to upload and share sensitive information in a centralized database.

The research team set out to create a federated machine learning model to define the boundaries of glioblastoma tumors. They used a multi-step process. Initial algorithms were developed and

refined based on expert judgment of brain-imaging data in a publicly available data set. More patient data were then added from multiple federated sites to validate the model and improve its accuracy. In its final stage, the model included data from more than 6,300 patients with glioblastoma at 71 sites worldwide.

Compared to the preliminary model, the final model led to a 33% improvement in pinpointing the tumor core and a 16% improvement in identifying the whole tumor. Detection of blood-brain barrier breakdown improved by 27%.

These improvements show that incorporating rare data from multiple sites can enhance machine learning outcomes. The researchers note that this approach could be useful in fields where data can be hard to come by, such as with rare disorders or underrepresented populations.

“This is the single largest and most diverse data set of glioblastoma patients ever considered in the literature, and was made possible through federated learning,” Bakas says. “The more data we can feed into machine learning models, the more accurate they become, which in turn can improve our ability to understand, treat, and remove glioblastoma in patients with more precision.”

This [research summary](#) was originally published by the National Institutes of Health on December 13, 2022.

© 2026 Smart + Strong All Rights Reserved.

<http://beta.docker.cancerhealth.com/article/machine-learning-approach-detects-brain-tumor-boundaries>