

Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm

Mateusz Buda, Ashirbani Saha, Maciej A. Mazurowski

Corresponding author: Mateusz Buda (mateusz.buda@duke.edu)

The details of preprocessing

Images varied significantly in size between patients. Therefore, we scaled them first to obtain images with smaller spatial dimension equal 256 pixels and then applied center crop to 256x256 square. This operation never cut the brain.

Next, we used in-house deep learning solution to perform skull-stripping. It was based on the same network architecture and used the same data pipeline as for tumor segmentation. Ground truth was obtained from automatic skull stripping MATLAB Statistical Parametric Mapping toolbox [27]. However, for 20 cases it gave unsatisfactory results. These cases were excluded from training our skull stripping network.

After resizing, we applied contrast stretching that required basic histogram analysis. For skull stripping, we set 50% of low value pixels to 0 and then stretch other values to have the peak of histogram at pixel with value 100 that corresponds to grey matter. For tumor segmentation, we only apply stretching to move maximum value of the histogram to pixel 100. It is done after skull stripping that already sets all pixel values outside the brain to 0. Contrast stretching was done per patient for all slices at once and separately for each available image channel (FLAIR, pre-contrast, and post-contrast).

Final preprocessing step was data normalization. We performed mean subtraction and divided entire training data by its standard deviation. For testing, we use values of mean and standard deviation computed on the training set and we repeat it for each test fold in cross-validation split.