Fast Glioblastoma Detection in Fluid-attenuated inversion recovery (FLAIR) images by Topological Explainable Automatic Machine Learning

Matteo Rucco*[1]

[1]United Technology Research Center, Via Praga 5, 38121 Trento (Tn), Italy.

Abstract

Glioblastoma multiforme (GBM) is a fast-growing and highly invasive brain tumor, it tends to occur in adults between the ages of 45 and 70 and it accounts for 52 percent of all primary brain tumors. Usually, GBMs are detected by magnetic resonance images (MRI). Among MRI images, Fluid-attenuated inversion recovery (FLAIR) sequence produces high quality digital tumor representation. This sequence is very sensitive to pathology and makes the differentiation between cerebrospinal fluid (CSF) and an abnormality much easier. Fast detection and segmentation techniques are needed for overcoming subjective medical doctors (MDs) judgment. In this work, a new methodology for fast detection and segmentation of GBM on FLAIR images is presented. The methodology leverages topological data analysis, textural features and interpretable machine learning algorithm, it was evaluated on a public available dataset. The machine learning classifier uses only eight input numerical features and it reaches up to the 97% of accuracy on the detection task and up to 95% of accuracy on the segmentation task. Tools from information theory were used for interpreting, in a human readable format, what are the main numerical characteristics of an image to be classified ill or healthy.

Keywords Glioblastoma; Fluid-attenuated inversion recovery; Brain; Tumor; Topological Data Analysis; Persistent Homology; Persistent Entropy; Intepretable Machine Learning; Explainable Machine Learning; Automatic Machine Learning; Co-occurrence matrix; Textural features; The Cancer Imaging Archive

*Corresponding Author: matteo.rucco@utrc.utc.com

1 Introduction

Gliomas are the most common primary brain tumors, originating from glial cells. They are characterized by a diffuse and infiltrative growth pattern, as invasive glioma cells often migrate along myelinated white matter (WM) fiber tracts. This is a major cause of their appalling prognosis: tumor cells invade, displace, and possibly destroy WM. The non-invasive detection of microscopic infiltrations, as well as the identification of aggressive tumor components within spatially heterogeneous lesions, are of outstanding importance for surgical and radiation therapy planning or to assess response to chemotherapy. Magnetic resonance (MR) imaging plays an important role in the detection and evaluation of brain tumors. For more than 30 years [13] conventional MR imaging, typically T1w images before and after paramagnetic contrast administration, T2w images and FLAIR images have been largely used to evaluate brain neoplasms. MR allows to localize the lesions, helps to distinguish tumors from other pathologic processes, and depicts basic signs of response to therapy, such as changes in size and degree of contrast enhancement. Manual tumor detection and segmentation on MR images is time-consuming and can have a great interobserver variability; automatic segmentation is more reproducible and efficient when robustness is particularly desirable, such as in monitoring disease progression or in the longitudinal evaluation of emerging therapies [7]. Texture analysis [8, 32], an image-analysis technique that quantifies gray-level patterns by describing the statistical relationships between the intensity of pixels, has demonstrated considerable potential for cerebral lesion characterization and segmentation [36]. For automatic glioma detection and segmentation in MRI, several algorithms have been already proposed. The most recent works can be grouped into superpixel based segmentation and deep learning based segmentation. In [43], the authors presented an interesting bottom-up approach that aims to combine graphical probabilistic model (i. e, Conditional Random Fields - CRF -) for capturing the spatial interactions among image superpixel regions and their measurements. A number of features including statistics features, the combined features from the local binary pattern as well as gray level run length, curve features, and fractal features were extracted from each superpixel. The features were used for teaching machine learning models to discriminate between healthy and pathological brain tissues. In [41], the authors used the concept of superpixels based segmentation for instrumenting a machine learning algorithm (i.e., support vector machine - SVM) that uses both textural and statistical features. The authors reported excellent average Dice coefficient, Hausdorff distance, sensitivity, and specificity scores when applying their method on T2w sequences from the Multimodal Brain Tumor Image Segmentation Benchmark 2017 (BraTS2017) dataset. In [37], the authors exploited a new method for brain tumor detection by combining features computed from Fluid-Attenuated Inversion Recovery (flair) MRI in association with the graph-based manifold ranking algorithm. The algorithm counts three mains steps: in the first phase, superpixel method is used to convert the homogeneous pixels in the form of superpixels. Rank of each superpixel or node is computed based on the affinity against certain selected nodes as the background prior in the second phase. The relevance of each node with the background prior is then computed and represented in the form of tumor map. In [38] the authors have reported on an approach that computes for each superpixel a number of novel image features including intensity-based, Gabor textons, fractal analysis and curvatures within the entire brain area in FLAIR MRI to ensure a robust classification. The authors compared Extremely randomized trees (ERT) classifier with support vector machine (SVM) to classify each superpixel into tumour and non-tumour.

Deep learning based solutions are becoming the new tools for brain segmentation. In [23],the authors used autoencoders for instrumenting an automatic segmentation of increased signal regions in fluid-attenuated inversion recovery magnetic resonance imaging images. In [25], the authors have provided a solution for dealing to the limited amount of available data from ill brains. They have trained a one-class classifier algorithm based on deep learning for segmenting brain tumors from fluid attenuation inversion recovery MRI. The technique exploits fully convolutional neural networks, and it is equipped with a battery of augmentation techniques that make the algorithm robust against low data quality, and heterogeneity of small training sets. Beside segmentation, deep-learning based solutions have been exploited for skull-stripping, and tissues identification (white matter, gray matter, etc. .) [22, 16],

A topological space is a powerful mathematical concept for describing the connectivity of a space. Informally, *a topological space* is a set of points, each of which equipped with the notion of *neighboring.* Algebraic topology uses typical concepts of abstract algebra to study topological spaces [19, 27]. In the last decade a new suite of tools, based on algebraic topology, for data exploration and modeling haven been invented [6, 44, 14]. The data science community refers to these tools as *Topological Data Analysis* (TDA). TDA has been used in different domains: biology, manufacturing, medicine and others [35]. Topological entropy, namely Persistent Entropy, is equipped with suitable mathematical properties, that permits to describe complex systems [3] and it has been applied in different experiments, e.g. the analysis of biological images [21] and the analysis of medical signals [31].

The aim of this study was to evaluate the possibility of discriminating healthy and pathological tissue on FLAIR MR images, by the use of Statistical

Texture Analysis and Topological Data Analysis. In fact, FLAIR reveals a wide range of lesions, including cortical, periventricular, and meningeal diseases that were difficult to see on conventional images. Inspired by the paradigm ”let the data speak for themselves" introduced in [17] by Gould, We have focused only on textural and topological features, since we believe this are the exact combinatorial representation of what the MDs see when looking at the gray scale images. The discriminating power of statistical texture features and of topological features was then exploited for the development of a supervised tumor detection and segmentation methodology by means of machine learning algorithms. In order to achieve the highest accuracy possible, we have challenged automatic machine learning framework, i.e., TPOT. TPOT takes as input the raw data and it compares different data analysis pipelines and it down-selects the one with the highest accuracy. Moreover, our solution provides human readable information that can be used by MDs to understand what are the main textural and topological characteristics of a tissue to be classified ill or healthy. The paper is organized as follows. in Section 2 we introduce the relevant background, namely the fundamental concepts of *superpixels, textural features* and *topological features.* Section 3 introduces the innovative methodologies for GBM detection and segmentation. In Section 4 we report on the implementation and application of the methodologies on slices extracted from public domain FLAIR images. Final thoughts about the results and next steps are discussed in the last Section 5.

2 Background

2.1 Superpixel

Superpixel is a powerful tool widely adopted in image segmentation. Each superpixel is a collection of pixels with common characteristics (e.g., gray scale). For an example of super segmentation with superpixels see Figure 2.1. Superpixel enables to extract more information than working at single pixel. Also, each superpixels are a compact representation of image region of interest (ROI) and that can be very useful for computationally demanding problems (e.g., dealing with high resolution images). For a complete overview of superpixel we refer to [20, 42, 1]. In this work, we used the *Simple Linear Iterative Clustering* (SLIC) algorithm for computing superpixels[1]. The algorithm starts with K regularly spaced cluster centers and it moves them to seed locations corresponding to the lowest gradient position in a grid of neighborhood pixels. Each pixel in the image is associated with the nearest cluster center whose search area overlaps

[1] `https://www.peterkovesi.com/projects/segmentation/` - last visit on 12/4/2019

this pixel. After all the pixels are associated with the nearest cluster center, a new center is computed as the average color vector of all the pixels belonging to the cluster. The process is repeated until all the pixels are associated to the clusters.

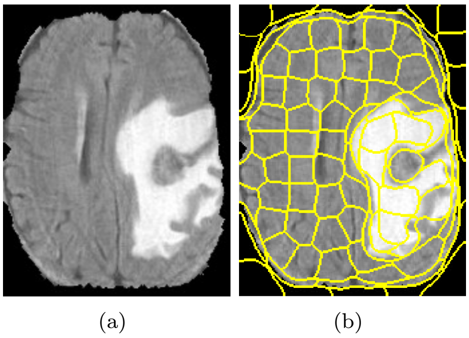


Figure 1: Example of superpixel over-segmentation on skull stripped flair image with 50 superpixels : left - original image; right - original image over-segmented with superpixels.

2.2 Gray-Level Co-occurrence matrix

Gray-Level Co-occurrence matrix, often referred as GLCM, is a statistical method of examining image texture by taking into account the spatial relationship of pixels. The GLCM counts how often pairs of pixel with specific values and in a specified spatial relationship occur in an image. From GLCM it is possible to extract textural information, e.g. correlation, energy, homogeneity, etc… [18].

2.3 Topological Data Analysis

Persistent homology

Homology is an algebraic machinery used for describing a topological space $\mathfrak{C}$. Informally, for a fixed natural number k, the $k-$*Betti number* β_k counts the number of $k-$dimensional holes characterizing $\mathfrak{C}$: β_0 is the number of connected components, β_1 counts the number of holes in 2D or tunnels in 3D[2], β_2 can be thought as the number of voids in geometric solids, and so on.

Persistent homology is a method for computing the $k-$dimensional holes at different spatial resolutions. Persistent holes are more likely to represent true features of the underlying space, rather than artifacts of sampling (noise), or due to particular choices of parameters. For a more formal description we refer the reader to [15]. In order to compute persistent homology, we need a distance

[2] Here nD refers to the $n-$dimensional space $\mathbb{R}^n$.

function on the underlying space. This can be obtained constructing *a filtration* on a simplicial complex, which is a nested sequence of increasing subcomplexes. More formally, a filtered simplicial complex K is a collection of subcomplexes $\{K(t) : t \in \mathbb{R}\}$ of K such that $K(t) \subset K(s)$ for $t < s$ and there exists $t_{max} \in \mathbb{R}$ such that $K_{t_{max}} = K$. The filtration time (or filter value) of a simplex $\sigma \in K$ is the smallest t such that $\sigma \in K(t)$.

Persistent homology describes how the homology of K changes along a filtration. A $k-$dimensional Betti interval, with endpoints $[t_{start}, t_{end})$, corresponds to a $k-$dimensional hole that appears at filtration time t_{start} and remains until time t_{end}. We refer to the holes that are still present at $t = t_{max}$ as *persistent topological features*, otherwise they are considered *topological noise* [2]. Figure 2.3 depicts the computation of persistent homology from a filtered simplicial complex made by four 2-simplices (i.e., filled triangles). The set of intervals representing birth and death times of homology classes is called the *persistence barcode* associated to the corresponding filtration. Instead of bars, we sometimes draw points in the plane such that a point $(x, y) \in \mathbb{R}^2$ (with $x < y$) corresponds to a bar $[x, y)$ in the barcode. This set of points is called *persistence diagram.* There are several algorithms for computing persistent barcodes, the principal of which are Gudhi [26] and jHoles [4]. For a complete overview of the available tools we refer to [30].

Euler Characteristics

Two or more topological objects, e.g.simplicial complexes, are homologically equivalent if they have the same sequence of Betti numbers. In other words, they are characterized by the same number of homological holes at each dimension. Two topological objects can be easily compared by comparing their Euler Characteristics:

Definition 2.1 (Euler Characteristic)

$$\chi = \sum_{i=0}^{i=n} (-1)^i \beta_i$$

Where β_i is the Betti numbers at the i-th homological group (e.g., β_0 is the Betti number at $H0$ for counting the number of connected componets, β_1 for counting the number of 2D holes, $beta_2$ for counting the number of 3D empty volumes etc. . .).

Persistent Entropy

In order to measure how much the construction of a filtered simplicial complex is ordered a new entropy measure, called *Persistent Entropy* (PE), was defined

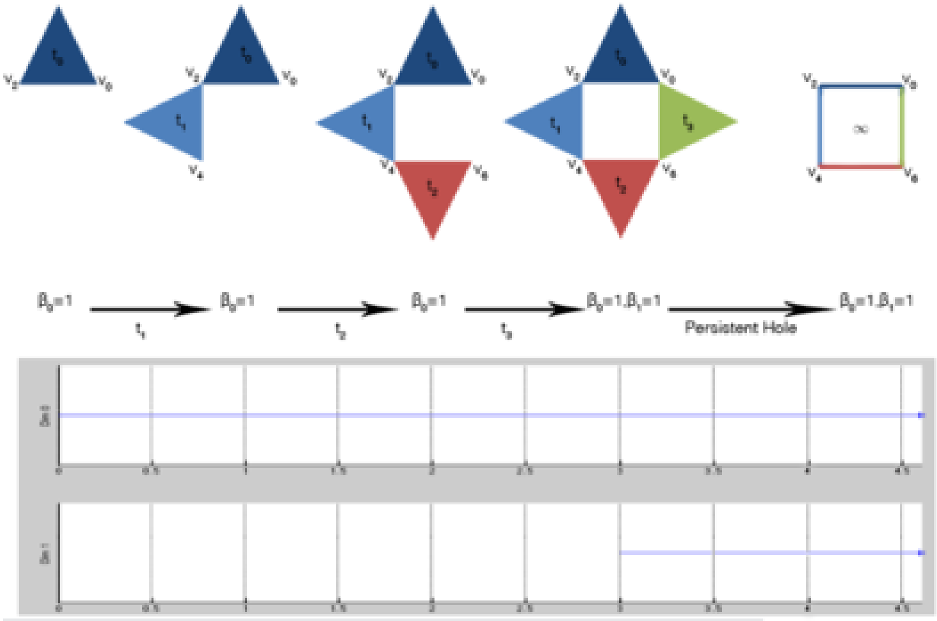


Figure 2: Pictorial representation of persistent homology computation of a filtered simplicial complex (top) with four 2D simplices - filled triangles. The persistent barcodes (bottom) catpure the evolution of the connected components (H0 group) and of 1D homological cycle (H1 group). The persistent Betti number sequence is $\beta_0 = 1; \beta_1 = 1$. This sequence means that there is only one connected components - all the triangles are connected, and only one topological loop that is the inner squared empty space of the simplicial complex. In this toy example the persistent 1D hole is made by 4 1D simplices (line segments) and 4 0D simplices (vertices), this are the so-called generators. The picture is from [4].

in [33]. A precursor of this definition was given in [10] to measure how different bars of a barcode are in length. Here we recall the definition.

Definition 2.2 (Persistent Entropy)
Given a filtered simplicial complex $\{K(t) : t \in F\}$, and the corresponding persistence barcode $B = \{a_i = [x_i, y_i) : i \in I\}$, the Persistent Entropy (PE) H of the filtered simplicial complex is defined as follows:

$$H = -\sum_{i \in I} p_i log_{10}(p_i)$$

where $p_i = \frac{\ell_i}{L}$, $\ell_i = y_i - x_i$, and $L = \sum_{i \in I} \ell_i$.

Note that, when topological noise is present, for each dimension of the persistence barcode there can be more than one interval, denoted by $[x_i \ , \ y_i)$ with $i \in I$. This is equivalent to say that, in the persistence diagram, the point $[x_i, y_i)$ could have multiplicity greater than 1 (see [15, page 152]).

In the case of an interval with no death time, $[x_i, +\infty)$, several approaches can be considered, such as extending real numbers including $+\infty$, removing or truncating infinite intervals or using extended persistence [12, 9, 5]. In this paper, we truncate infinite intervals and replace $[x_i, +\infty)$ by $[x_i\ ,\ m)$ in the persistence barcode, where $m = t_{\max} + 1$.

Note that the maximum PE corresponds to the situation in which all the intervals in the barcode are of equal length. In that case, $H = \log n$ if n is the number of elements of I. Conversely, the value of the PE decreases as more intervals of different length are present.

The stability theorem defined in [34] defines how to compare persistent entropies from different simplicial complexes. The main requirement is that the simplicial complexes have the same number of 0-simplices (nodes).

Generator Entropy

Each topological feature, regardless its persistency, is generated by the so-called generators [28]. In this work, we propose a new seminal statistics that summarizes the number of 0D generators in each topological features.

Definition 2.3 (Generator Entropy)

$$GH = -\Sigma_{i=1}^{i=N}(p_i log_{10}(p_i))$$

where N is the number of homological holes, n_i = *number of unique generators in the i-th homological hole,* $L = \Sigma_{i=1}^{i=N} n_i$ *and* $p_i = \frac{n_i}{L}$.

We envision that this new topological statistics could be used for detecting correlation between holes length and specific objects in the input space represented by the simplicial complex (e.g., tumor).

From image to filtered simplicial complex

We propose a new algorithm for transforming a gray scale image into a filtered simplicial complex. Eventually, the simplicial complex is analyzed by means of persistent homology. The input of the algorithm is a gray scale image or a part of it.

- The image is super-segmented with k superpixels, this returns a group of k superpixel regions: $S = \{s_1 \ldots, s_k\}$ where each superpixel is a group of pixels $s_1 = \{p_1, \ldots, p_n\}$. We remark that each pixel is described in term of its gray level, namely $gray(p_i)$.

- Superpixels became the vertices of a graph $G = (V \times E)$, $V = \{s_1, \ldots, s_n\}$, $E = \{e_{1,2}, \ldots e_{m,n}\}$. Two nodes are connected if the corresponding superpixels are adjacent.

- Maximal clique finding algorithm is executed on the graph. We used the Bron-Kerbosch algorithm[3]. The algorithm returns a list of maximal cliques $C = \{c_1, c_2, \ldots, c_n\}$, where each clique is a set of edges. We remark that $k-$maximal clique corresponds to a $k-1-$simplices. E.g., a 3 clique, that is a triangle, corresponds to a 2D simplices, that is a filled triangle.

- In order to transform the graph into a filtered simplicial complex, the following filter functions are defined:

  Definition 2.4 (Filter Functions)
  Filter function for the 0-simplices (vertices):

  $$f(s_i) = mean\{gray(p_1), \ldots, gray(p_n)\}$$

  Filter function for the 1-simplices (edges):

  $$f(e_{i,j}) = max\{f(s_i), f(s_j)\}$$

  Filter function for the $n_{\{}n > 1\}$-simplices (high dimensional simplices):

  $$f(c_i) = max\{f(e_i), f(e_j), \ldots, f(e_k)\}$$

- The weighted graph is analyzed as a filtered simplicial complex. In this setting, dark regions (low gray levels) are added the the beginning of the construction. While, brighter regions (usually pathological tissues) are added at the end of the construction.

- Compute persistent homology.

3 A new methodology for supervised brain tissue classification and segmentation

Methodology for glioblastoma detection

In this work we have implemented and tested the methodology depicted in Figure 3 and that is based on the algorithm described in Sec. 2. The methodology

[3] `https://uk.mathworks.com/matlabcentral/fileexchange/30413-bron-kerbosch-maximal-clique-finding-algorithm` - last visit on 4/12/2019.

takes as input a collection of FLAIR file. Each FLAIR file is completed with a 3D tumor mask, or in other words with the region of interest (ROI), that indicates where the tumor is located. The methodology counts the following steps:

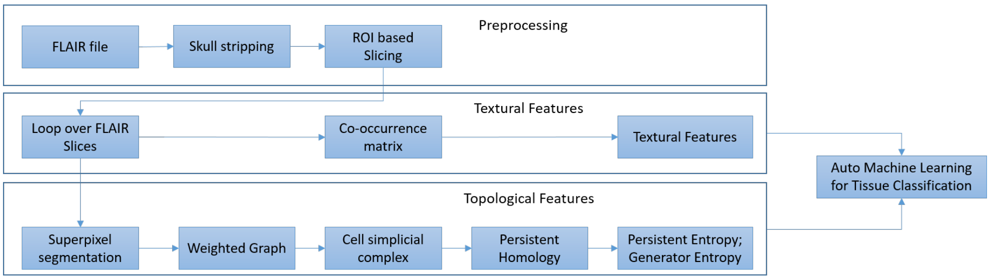


Figure 3: Data analysis workflow presented in this research paper. The pipeline is divided in four main blocks: preprocessing, textural and topological features extraction and finally automatic machine learning for brain tissue classification. Details are reported in the text of Sec. 3.

1. Preprocessing

    - The FLAIR file is optimized by removing fat tissues and by performing skull stripping.
    - Accordingly to the 3D ROI, the slices containing GBM are extracted from the FLAIR file. For each slice the information on the position of the tumor with respect to the center of the image is also stored.

2. Each slice is divided it in two parts the so-called lateral and contralateral. In this setting, lateral images contain the tumor, while contralateral images contain healthy tissue. The parts are resized such that they have the same sizes.

3. For each part the following statistics are computed:

4. Textural Features - note that, features *b* to *f* are computed from the GLCM:

    (a) Gray level - average grays intensity in each superpixel.

    (b) Contrast - it measures the local variability in the gray level.

    (c) Correlation - it measures the joint probability that a given pair of pixels is found.

    (d) Homogeneity - it measures the distance between the GCLM diagonal and element distribution.

(e) Energy

5. Topological features from the simplicial complex associated to the image under analysis:

   (a) Euler characteristics.

   (b) Persistent entropy.

   (c) Generator entropy.

6. Each part is represented by a feature vector of length 8 plus a label : $l = \{0, 1\}$ if the vector was computed from lateral or controlateral, respectively.

7. The dataset, aka the collection of feature vectors, is divided randomly in three different subsets: *training, testing* and *validation* that contain the 70%, 15% and 15% of samples respectively. We remark that the training set is used for training a machine learning (ML) classifier. The test set is used for improving the tuning of the algorithm and to increase the prediction reliability of the algorithm. The validation set is used for measuring the performances of the trained ML classifier that are reported by receiver operating characteristic curve (ROC) and area under curve (AUC). Selection of ML algorithm can be obtained by using Automatic Machine Learning approach.

8. Classifier is debugged by tools from information theory that compute both global and local features relevance. This allows to understand the relevance of each feature and to understand what are the numerical input characteristics related to the output.

Methodology for glioblastoma segmentation

A slightly modified version of the methodology can be used for the segmentation of a slice and without the need of dividing it into lateral and contralateral. After the preprocessing step, the image is initially simplified by superpixels. For each superpixel the textural features and topological features are computed, as described above. The intersection between the pixels in the superpixel region and the corresponsing pixels in the ROI image is computed. If the intersection is $\emptyset$ the class for superpixel is set to *healthy*, otherwhise the class is set to be *ill*.

4 Discussion

We have tested the methodology described in Sec. 3 on a public freely available dataset accessible via *The Cancer Imaging Archive (TCIA)* [11]. Fig. 4 shows two of the slices used in this work. In details, the dataset includes DICOM files of 20 subjects with primary newly diagnosed glioblastoma who were treated with surgery and standard concomitant chemo-radiation therapy (CRT) followed by adjuvant chemotherapy. The sequences are T1-weighted (pre and post-contrast agent), FLAIR, T2-weighted, ADC, normalized cerebral blood flow, normalized relative cerebral blood volume, standardized relative cerebral blood volume, and binary tumor masks (i.e., ROIs) [28]. Each patient is described by two MRI exams: within 90 days following CRT completion and at progression. We have processed the FLAIR sequences collected within the 90 days of CRT completion. We decided to focus on this set of sequences since the tumor is always present. The DICOM were transformed into NII files for enabling further analysis[4] . The preprocessing steps, namely the removal of fat tissues and skull stripping, were performed by using the deep-learning algorithm described in [24][5]. From FLAIR files a total of 2408 gray scale images were extracted. Note that each image contains the tumor.

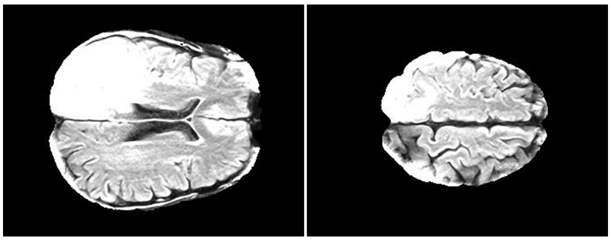


Figure 4: Example of slices contained in the dataset.

Glioblastoma detection

The images were partitioned in 2408 lateral parts plus 2408 contralateral parts for a total of 4816 samples. Textural and topological features were calculated with open source software Octave[6] and JavaPlex [39]. The dataset was randomly divided in the training, testing and validation subsets. The dataset was used for two different machine learning experiments. In the first experiment we have trained a multilayer supervised perceptron (MLP) [7] with a fixed architecture of a single hidden layer with 8 neurons each of them equipped with *lbfgs*. The algorithm was trained with the back-propagation approach. In this approach the data were preprocessed by applying standard scaler[8]. In the second experiment

[4] https://pypi.org/project/dicom2nifti/

[5] https://github.com/JanaLipkova/s3

[6] https://www. gnu. org/software/octave/

[7] https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html

[8] https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html

we used the automatic machine learning framework TPOT [29]. TPOT is a framework that uses genetic algorithm for comparing different machine learning solutions and it finds automatically the best pipeline that maximizes the accuracy. TPOT produces a pipeline that contains both the preprocessing steps - if needed - and the selected Machine Learning architecture. In both experiments and for increasing prediction reliability, 5-fold cross validation[9] approach was used during the training. Skater and Lime algorithms were used for for investigating the structure of the trained machine learning algorithms and for computing features relevance [40]. For a complete introduction to machine learning interpretation we refer to[10]

Glioblastoma segmentation

For evaluating the methodology for glioma segmentation each slice has been segmented in 250 superpixels. The dataset thus contains $250 \times 2408 = 602.000$ samples. Each sample is completed with its class label, as described in the methodology. Similarly to the methodology for the detection, the dataset was divided in three subsets and used for training and evaluating both a multi layer perceptron and automatic selected machine learning classifier by TPOT.

Results

Glioblastoma detection

The methodology described in Sec. 3 was evaluated over different number of superpixels, i.e., from 50 to 500 with an increment of 50. For each number of superpixels we have evaluated the accuracy by means of AUC of both the Artificial Neural Networks and TPOT models. The comparison of the performances is displayed in Figure 4. In general, TPOT model outperforms artificial neural network. TPOT reaches the maximum accuracy at superpixels = 500: TPOT AUC = 97. 1% and ANN AUC = 93. 7%. Table 1 reports the accuracy and the corresponding 95% confidence interval for each experiment and for both TPOT and ANN. The list of algorithms that were selected by TPOT are reported in Sec.3.

In both cases the accuracy are quite good. We argue that TPOT shows better performance because it adds preprocessing steps (e.g., StandardScaler[11],

[9] https://scikit-learn.org/stable/modules/cross_validation.html

[10] https://www.oreilly.com/ideas/interpreting-predictive-models-with-skater-unboxing-model-opacity

[11] https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html

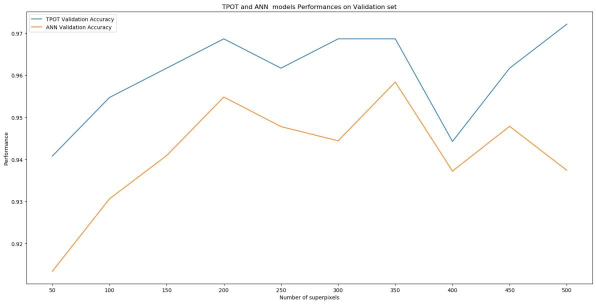


Figure 5: Accuracy comparison on the validation set of both TPOT and ANN accuracy over different number of superpixels.

	ANN		TPOT	
Superpixels	AUC	95% C. I.	AUC	95% C. I.
50	0. 913	[0. 858 - 0. 969]	0. 941	[0. 885 - 0. 978]
100	0. 931	[0. 892 - 0. 969]	0. 955	[0. 916 - 0. 991]
150	0. 941	[0. 924 - 0. 958]	0. 962	[0. 932 - 0. 993]
200	0. 955	[0. 933 - 0. 977]	0. 969	[0. 945 - 0. 996]
250	0. 948	[0. 924 - 0. 971]	0. 962	[0. 932 - 0. 993]
300	0. 944	[0. 910 - 0. 978]	0. 969	[0. 945 - 0. 996]
350	0. 958	[0. 930 - 0. 987]	0. 969	[0. 945 - 0. 996]
400	0. 937	[0. 920 - 0. 955]	0. 969	[0. 945 - 0. 996]
450	0. 948	[0. 920 - 0. 975]	0. 962	[0. 932 - 0. 993]
500	0. 937	[0. 898 - 0. 977]	0. 972	[0. 947 - 1. 0]

Table 1: Machine learning models accuracy

RobustScaler[12] and Normalizer[13]) that increase the separation of the features with respect the two classes. For the sake of completeness, raw data do not show any strong separation for what concern the two classes as depicted by the density plots in Figure 4.

The analysis of the average feature relevance with respect to the ML algorithm is depicted in Figure 4. For ANN the textural features *gray levels energy*, and *homogenity* are the most informative features. The two topological entropies, namely *Persistent Entropy (PE)* and *Generator Entropy (Gen)* compensate each other. PE is quite important for superpixels = 350 and it corresponds to the point in which the accuracy of the ANN reaches its maximum. For TPOT models the topological features are less important, while textural features *correlation*, *energy* and *contrast* are fundamental.

[12] `https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.RobustScaler.html`

[13] `https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.Normalizer.html`

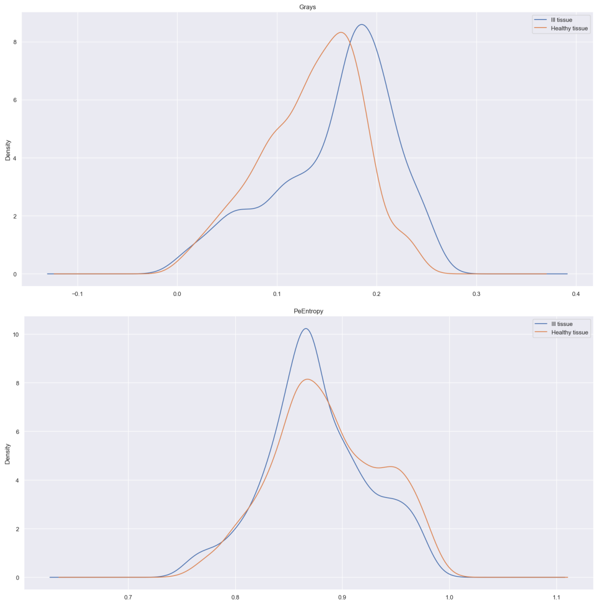


Figure 6: Comparison of density plots for ill and healthy tissues for gray levels (top) and persistent entropy (bottom). For typographic reasons we report only 2 out of 8 density plots.

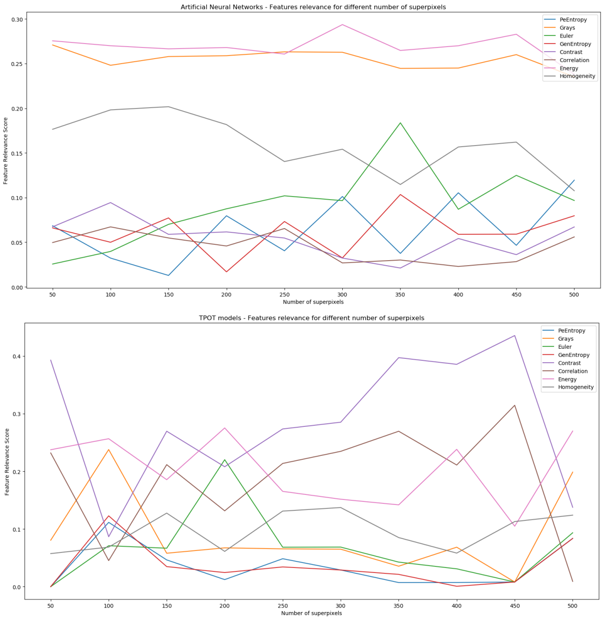


Figure 7: Global analysis of features relevance for ANN (top) and TPOT models (bottom).

Lime algorithm allows us to understand what are the numerical characteristic for a patient to be classified healthy or ill, see Figure 4. Positive (blue) indicate ill tissues, while negative (orange) indicate healthy tissue. The way to interpret the weights by applying them to the prediction probabilities. The bars indicate the weight for each feature and the corresponding value for the slice under test. For example, in case of ANN model, a slice to be classified as healthy shall have grays $>$ 0. 74, homogeneity $<=$ -0. 37, Persistent Entropy $<=$ -. 65, etc. . . . In case of TPOT model, if we set to zero the values for the features Energy, Contrast and GenEntropy we expect the classifier to predict the slice under test as *ill tissue* with probability 1.00 - 0.39 - 0.06 - 0.04 = 0.51 that would mean approximately random detection. This highlights that even if some features might have globally less relevance, they are locally fundamental for avoiding ”flipping coin" predictions.

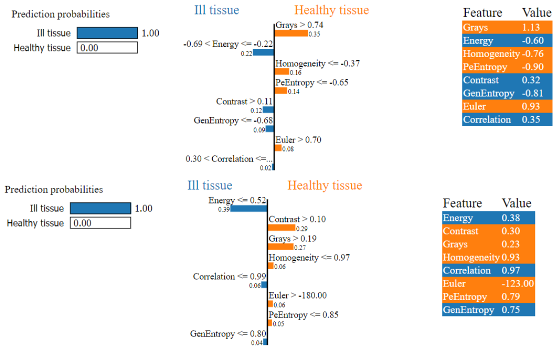


Figure 8: Local analysis of features relevance for ANN (top) and TPOT models (bottom).

Glioblastoma segmentation

The methodology for glioblastoma segmentation and reported in Sec. 3 was evaluated for a fixed number of superpixels = 250. For the sake of clarity, 250 is a good trade-off between the accuracy, the number of pixels in each superpixel and the timing of the experiments[14]. The accuracy on the validation set of ANN and of TPOT are reported in Table 2. Also for segmentation TPOT outperforms ANN. An example of the execution of the segmentation is depicted in Fig.4. We argue that even in this case the preprocessing steps introduced by TPOT improves the quality of the results. The pipeline selected by TPOT is reported below. Skater analysis is depicted in Fig. 4. The model selected by TPOT for the segmentation task is strongly conditioned by the topological features, while

[14] Note that, all the numerical experiments were executed on Apple MaBook Air equipped with CPU 1.6GHz, 8GB RAM@1600MHZ, macOS Mojave ver. 10.14.6, Python 3.7.3

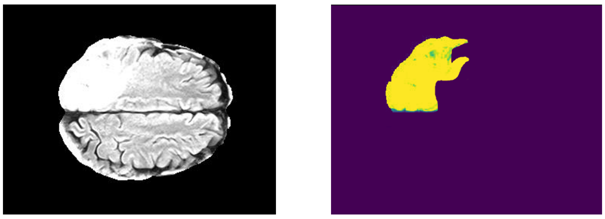


Figure 9: Output of the segmentation process. Input image (left) - Detected tumor (right). Yellow indicates the GBM. Green corresponds to internal tissues or other pathologies that are beyond the scope of this paper.

	ANN		TPOT	
Superpixels	AUC	95% C. I.	AUC	95% C. I.
250	0.94	[0.93 - 0.95],	0.95	[0.94 - 0.96]

Table 2: Machine learning models accuracy for the segmentation experiment

textural features are still the most relevant for the artificial neural network. Lime analysis is reported in Fig. 4. For TPOT the removal of persistent entropy would cause a miss-classification of the superpixel under test.

```
exported_pipeline = make_pipeline(
    MinMaxScaler(),
    StackingEstimator(estimator=KNeighborsClassifier ...
        (n_neighbors=17, p=1, weights="distance")),
    KNeighborsClassifier(n_neighbors=1, p=2, ...
        weights="uniform")
)
```

Listing 1: TPOT classifier for segmentation

5 Conclusions

We have described a method for semi-automatic detection and segmentation of cerebral gliomas in brain Flair sequences by means of a classifier trained on statistical texture features and topological features. The use of textural features in combination with topological data analysis and interpretable automatic machine learning in FLAIR sequences, with the aim of detecting brain tumor, is, at the best of our knowledge, a new approach. Fast and reliable assessment of glioma presence is critical for an accurate surgical and radiotherapy planning as well as to evaluate and quantify treatment-related effects and progression. Our method started with a procedure for image simplification by means of superpixels segmentation and thorough the characterization of healthy and pathologic tissue

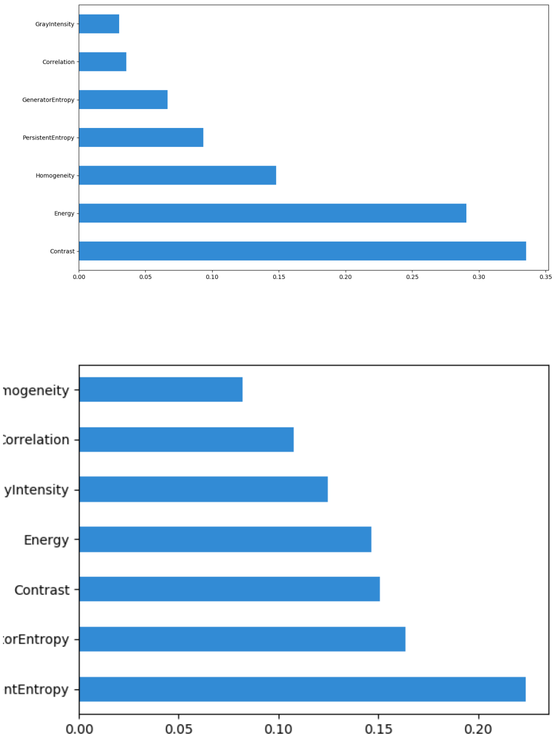
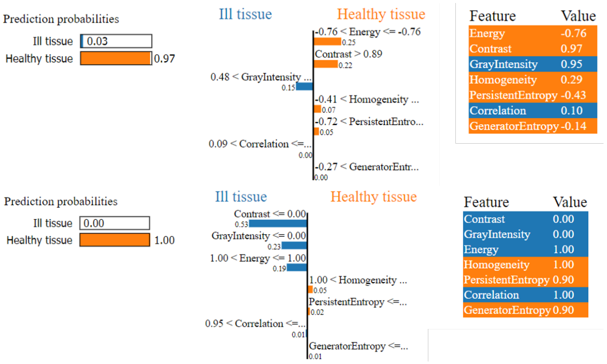


Figure 10: Global analysis of features relevance for ANN (top) and TPOT models (bottom) for the segmentation task.



Feature	Value
Energy	-0.76
Contrast	0.97
GrayIntensity	0.95
Homogeneity	0.29
PersistentEntropy	-0.43
Correlation	0.10
GeneratorEntropy	-0.14

Feature	Value
Contrast	0.00
GrayIntensity	0.00
Energy	1.00
Homogeneity	1.00
PersistentEntropy	0.90
Correlation	1.00
GeneratorEntropy	0.90

Figure 11: Local analysis of features relevance for ANN (top) and TPOT models (bottom) for the segmentation task.

by the use of textural and topological features. The features were computed over different number of superpixels. Then, both artificial neural network and automatic machine learning were used to train an automatic supervised classifier. The quality of classification was assessed by ROC curves, and AUC values were indeed quite large (detection: AUC for ANN = 0.96 - 95% C. I. : [0.93 - 0.98] , AUC for TPOT = 0.97 95% C. I. : [0.95 - 1.0]), (segmentation: AUC for ANN = 0.94, 95% CI [0.93, 0.95], AUC for TPOT = 0.95, 95% CI [0.93, 0.96]). The performances were evaluated on a public and freely available dataset. Finally, the trained systems were investigated with tools from information theory. Skater and Lime algorithms allowed to compute features relevance and for understanding what are the numerical characteristics of an input sample to be classified ill or healthy. Intepretation can be seen as an initial step to be compliant with the EU-GDPR 22nd article ("Automated individual decision-making, including profiling") for the "right to be informed"[15]. Among the topological features, we have introduced a new topological entropy, the so-called *generator entropy* that would summarize the role of the length of homological cycles for tissue classification. In the future, we envision several efforts: we would like to evaluate if the approach described in this paper allows to quantify and to monitor structural changes of the tumor during follow-up period. Eventually, we intend to investigate mathematical properties (minimum set of generators, stability theorem, etc...) of the new topological entropy [28].

6 Acknowledgement

The author would like to thank Prof. Matthias Zeppelzauer for the initial advise on using superpixels for image simplification. Dr. Giorgio De Nunzio, Dr. Marina Donativi and MD Antonella Castellano for long-standing conversations on related topics.

References

[1] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels. Technical report, 2010.

[2] Henry Adams and Andrew Tausz. Javaplex tutorial, 2011.

[3] Nieves Atienza, Rocio Gonzalez-Diaz, and Matteo Rucco. Persistent entropy for separating topological features from noise in vietoris-rips complexes. *Journal of Intelligent Information Systems*, 52(3):637–655, 2019.

[4] Jacopo Binchi, Emanuela Merelli, Matteo Rucco, Giovanni Petri, and Francesco Vaccarino. jHoles: A Tool for Understanding Biological Complex Networks via Clique Weight Rank Persistent Homology. *Electronic Notes in Theoretical Computer Science*, 306:5–18, 2014.

[15] `http://www.privacy-regulation.eu/it/22.htm`

[5] Peter Bubenik and Paweł Dłotko. A persistence landscapes toolbox for topological statistics. *Journal of Symbolic Computation*, pages –, 2016.

[6] Gunnar Carlsson. Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.

[7] Antonella Castellano, Marina Donativi, Roberta Rudà, Giorgio De Nunzio, Marco Riva, Antonella Iadanza, Luca Bertero, Matteo Rucco, Lorenzo Bello, Riccardo Soffietti, et al. Evaluation of low-grade glioma structural changes after chemotherapy using dti-based histogram analysis and functional diffusion maps. *European radiology*, 26(5):1263–1273, 2016.

[8] G Castellano, L Bonilha, LM Li, and F Cendes. Texture analysis of medical images. *Clinical radiology*, 59(12):1061–1069, 2004.

[9] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas J. Guibas, and Steve Oudot. Proximity of persistence modules and their diagrams. In *Proceedings of the 25th ACM Symposium on Computational Geometry, Aarhus, Denmark, June 8-10*, pages 237–246, 2009.

[10] Harish Chintakunta, Thanos Gentimis, Rocio Gonzalez-Diaz, Maria-Jose Jimenez, and Hamid Krim. An entropy-based persistence barcode. *Pattern Recognition*, 48(2):391–401, 2015.

[11] Kenneth Clark, Bruce Vendt, Kirk Smith, John Freymann, Justin Kirby, Paul Koppel, Stephen Moore, Stanley Phillips, David Maffitt, Michael Pringle, et al. The cancer imaging archive (tcia): maintaining and operating a public information repository. *Journal of digital imaging*, 26(6):1045–1057, 2013.

[12] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Extending persistence using poincaré and lefschetz duality. *Foundations of Computational Mathematics*, 9(1):79–103, 2009.

[13] Raymond Damadian. Tumor detection by nuclear magnetic resonance. *Science*, 171(3976):1151–1153, 1971.

[14] Herbert Edelsbrunner and John Harer. Persistent Homology – a Survey. *Contemporary mathematics*, 453:257–282, 2008.

[15] Herbert Edelsbrunner and John Harer. *Computational topology: an introduction.* American Mathematical Soc., 2010.

[16] Francisco J Galdames, Fabrice Jaillet, and Claudio A Perez. An accurate skull stripping method based on simplex meshes and histogram analysis for

magnetic resonance images. *Journal of neuroscience methods*, 206(2):103–119, 2012.

[17] Peter Gould. Letting the data speak for themselves. *Annals of the Association of American Geographers*, 71(2):166–176, 1981.

[18] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621, 1973.

[19] Allen Hatcher. *Algebraic Topology.* Cambridge University Press, 2002.

[20] Peter W Hawkes. *Advances in imaging and electron physics*, volume 133. Elsevier, 2004.

[21] María José Jimenez, Matteo Rucco, P Vicente-Munuera, P Gómez-Gálvez, and Luis M Escudero. Topological data analysis for self-organization of biological tissues. In *International Workshop on Combinatorial Image Analysis*, pages 229–242. Springer, 2017.

[22] Jens Kleesiek, Gregor Urban, Alexander Hubert, Daniel Schwarz, Klaus Maier-Hein, Martin Bendszus, and Armin Biller. Deep mri brain extraction: a 3d convolutional neural network for skull stripping. *NeuroImage*, 129:460–469, 2016.

[23] Panagiotis Korfiatis, Timothy L Kline, and Bradley J Erickson. Automated segmentation of hyperintense regions in flair mri using deep learning. *Tomography*, 2(4):334, 2016.

[24] Jana Lipková, Panagiotis Angelikopoulos, Stephen Wu, Esther Alberts, Benedikt Wiestler, Christian Diehl, Christine Preibisch, Thomas Pyka, Stephanie E Combs, Panagiotis Hadjidoukas, et al. Personalized radiotherapy design for glioblastoma: Integrating mathematical tumor models, multimodal scans, and bayesian inference. *IEEE transactions on medical imaging*, 38(8):1875–1884, 2019.

[25] Pablo Ribalta Lorenzo, Jakub Nalepa, Barbara Bobek-Billewicz, Pawel Wawrzyniak, Grzegorz Mrukwa, Michal Kawulok, Pawel Ulrych, and Michael P Hayball. Segmenting brain tumors from flair mri using fully convolutional neural networks. *Computer methods and programs in biomedicine*, 176:135–148, 2019.

[26] Clément Maria, Jean-Daniel Boissonnat, Marc Glisse, and Mariette Yvinec. The gudhi library: Simplicial complexes and persistent homology. In *International Congress on Mathematical Software*, pages 167–174. Springer, 2014.

[27] James R Munkres. *Elements of algebraic topology*, volume 2. Addison-Wesley Reading, 1984.

[28] Ippei Obayashi. Volume-optimal cycle: Tightest representative cycle of a generator in persistent homology. *SIAM Journal on Applied Algebra and Geometry*, 2(4):508–534, 2018.

[29] Randal S Olson and Jason H Moore. Tpot: A tree-based pipeline optimization tool for automating machine learning. In *Automated Machine Learning*, pages 151–160. Springer, 2019.

[30] Nina Otter, Mason A Porter, Ulrike Tillmann, Peter Grindrod, and Heather A Harrington. A roadmap for the computation of persistent homology. *EPJ Data Science*, 6(1):17, 2017.

[31] Marco Piangerelli, Matteo Rucco, Luca Tesei, and Emanuela Merelli. Topological classifier for detecting the emergence of epileptic seizures. *BMC research notes*, 11(1):392, 2018.

[32] Fabio Remondino and Sabry El-Hakim. Image-based 3d modelling: a review. *The photogrammetric record*, 21(115):269–291, 2006.

[33] Matteo Rucco, Filippo Castiglione, Emanuela Merelli, and Marco Pettini. Characterisation of the idiotypic immune network through persistent entropy. In *Proceedings of ECCS 2014*, pages 117–128. Springer, 2016.

[34] Matteo Rucco, Rocio Gonzalez-Diaz, Maria-Jose Jimenez, Nieves Atienza, Cristina Cristalli, Enrico Concettoni, Andrea Ferrante, and Emanuela Merelli. A new topological entropy-based approach for measuring similarities among piecewise linear functions. *Signal Processing*, 134:130–138, 2017.

[35] Matteo Rucco, Adane Letta Mamuye, Marco Piangerelli, Michela Quadrini, Luca Tesei, and Emanuela Merelli. Survey of topdrim applications of topological data analysis. In *CEUR Workshop Proceedings*, volume 1748, page 1814, 2016.

[36] Lothar R Schad, Stefan Blüml, and Ivan Zuna. Ix. mr tissue characterization of intracranial tumors by means of texture analysis. *Magnetic resonance imaging*, 11(6):889–896, 1993.

[37] Shiv Naresh Shivhare and Nitin Kumar. Brain tumor detection using manifold ranking in flair mri. In *Proceedings of ICETIT 2019*, pages 292–305. Springer, 2020.

[38] Mohammadreza Soltaninejad, Guang Yang, Tryphon Lambrou, Nigel Allinson, Timothy L Jones, Thomas R Barrick, Franklyn A Howe, and Xujiong Ye. Automated brain tumour detection and segmentation using superpixel-based extremely randomized trees in flair mri. *International journal of computer assisted radiology and surgery*, 12(2):183–203, 2017.

[39] Andrew Tausz, Mikael Vejdemo-Johansson, and Henry Adams. javaplex: a research platform for persistent homology. *Book of Abstracts Minisymposium on Publicly Available Geometric/Topological Software*, 7, 2012.

[40] Pengfei Wei, Zhenzhou Lu, and Jingwen Song. Variable importance analysis: a comprehensive review. *Reliability Engineering & System Safety*, 142:399–432, 2015.

[41] Yaping Wu, Zhe Zhao, Weiguo Wu, Yusong Lin, and Meiyun Wang. Automatic glioma segmentation based on adaptive superpixel. *BMC medical imaging*, 19(1):1–14, 2019.

[42] F Xing and L Yang. Machine learning and its application in microscopic image analysis. In *Machine Learning and Medical Imaging*, pages 97–127. Elsevier, 2016.

[43] Zhe Zhao, Guan Yang, Yusong Lin, Haibo Pang, and Meiyun Wang. Automated glioma detection and segmentation using graphical models. *PloS one*, 13(8):e0200745, 2018.

[44] Afra Zomorodian. Topological data analysis. In *Advances in Applied and Computational Topology. Proceedings of Symposia in Applied Mathematics*, volume 70, pages 1–39, 2007.

Appendix

Number of Superpixels	TPOT Model
50	$exported\_pipeline = make\_pipeline(DecisionTreeClassifier(criterion ='' ent$
100	$exported\_pipeline = make\_pipeline(PCA(iterated\_power = 8, svd\_solver =''$
150	$exported\_pipeline = make\_pipeline(RobustScaler(), StackingEstimator(estim$
200	$exported\_pipeline = make\_pipeline(StackingEstimator(estimator = Logistic$
250	$exported\_pipeline = make\_pipeline(RobustScaler(), KNeighborsClassifier(n$
300	$exported\_pipeline = make\_pipeline(StandardScaler(), RobustScaler(), Norma$
350	$exported\_pipeline = make\_pipeline(RobustScaler(), KNeighborsClassifier(n$
400	$exported\_pipeline = DecisionTreeClassifier(criterion ='' entropy'', max\_dep$
450	$exported\_pipeline = make\_pipeline(RobustScaler(), KNeighborsClassifier(n$
500	$exported\_pipeline = make\_pipeline(StandardScaler(), KNeighborsClassifier$

Table 3: Machine learning models selected by TPOT. TPOT produces pipelines by using Scikit-Learn API.