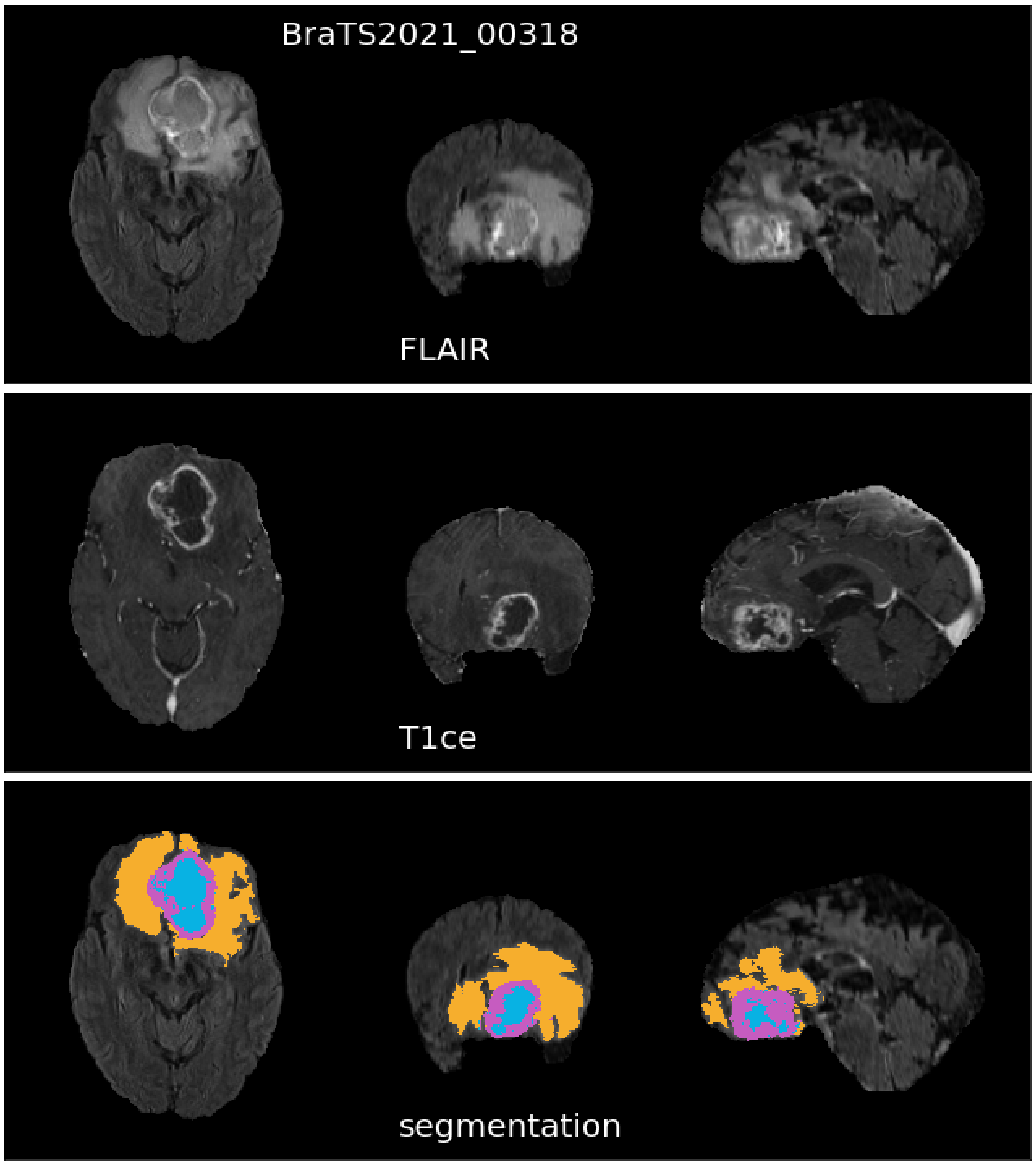
Infinite-dimensional Geometry: Theory and Applications
Week 5: Shape Analysis and Medical Applications
Erwin Schrödinger International Institute – 14/02/2025

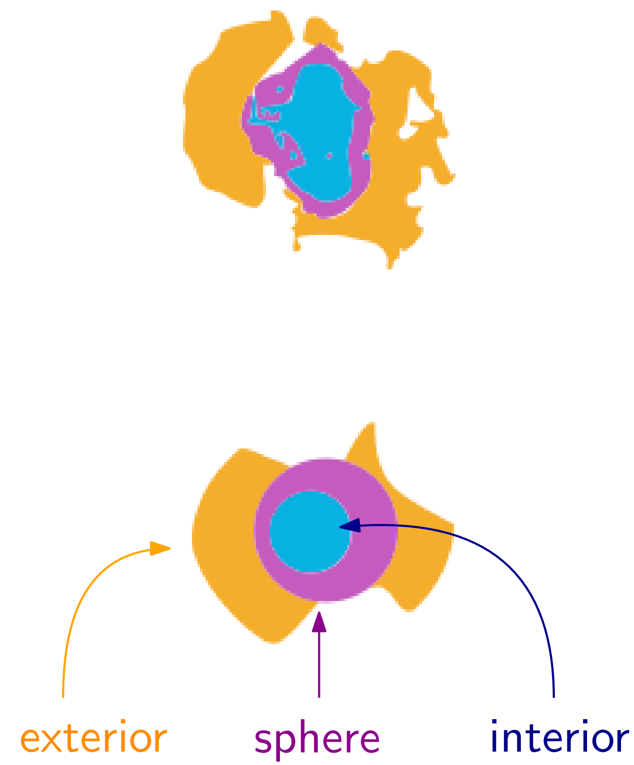
Train-Free Segmentation in MRI with Cubical Persistent Homology

Anton François – ENS Paris-Saclay
Raphaël Tinarrage – IST Austria

Objective: segment glioblastoma in MRIs (modalities Flair and T1ce).
Dataset: BraTS2021.



Three classes: Peritumoral Edema (ED), Tumorous Core (TC), Enhancing Tumor (ET).

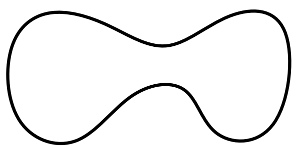
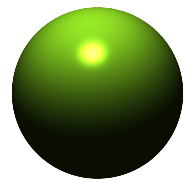
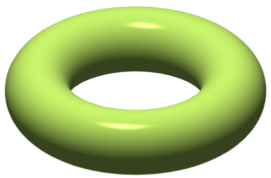
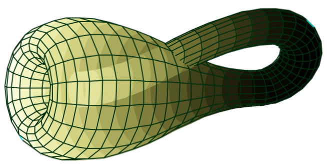


Homology groups

Let k be a field. The n^{th} singular **homology** with coefficients in k is a functor

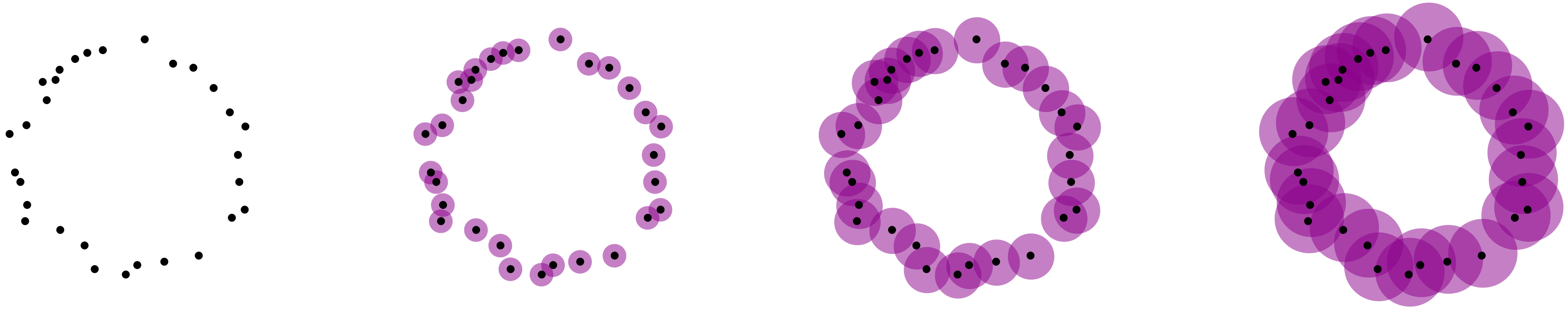
$$H_n \colon \mathbf{Top} \to k\text{-}\mathbf{Vect}$$

i.e.,
- to each topological space is associated a k-vector space $H_n(X;k)$,
- to each continuous map $f\colon X \to Y$ is associated a linear map $f_*\colon H_n(X;k) \to H_n(Y;k)$.

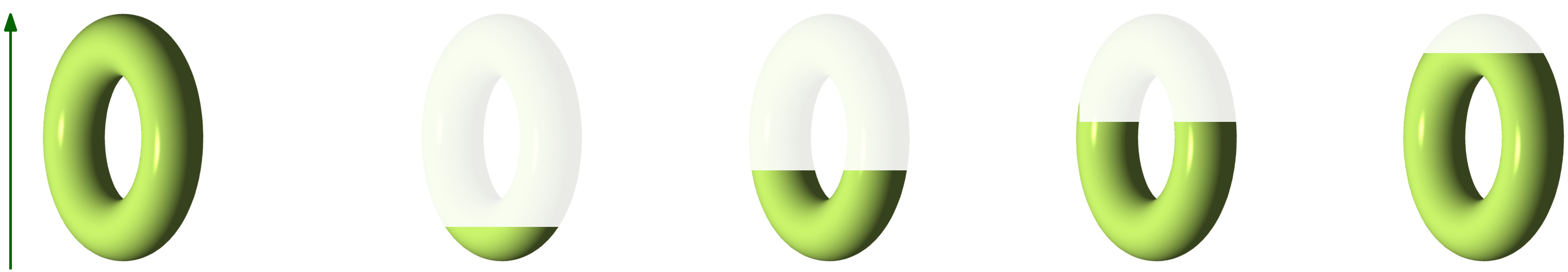
X	circle 	2-sphere 	torus 	Klein bottle 
$H_0(X;\mathbb{Z}/2\mathbb{Z})$	$\mathbb{Z}/2\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z}$
$H_1(X;\mathbb{Z}/2\mathbb{Z})$	$\mathbb{Z}/2\mathbb{Z}$	0	$(\mathbb{Z}/2\mathbb{Z})^2$	$(\mathbb{Z}/2\mathbb{Z})^2$
$H_2(X;\mathbb{Z}/2\mathbb{Z})$	0	$\mathbb{Z}/2\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z}$

Interpretation: H_0 counts connected components, H_1 counts holes, H_2 counts cavities.

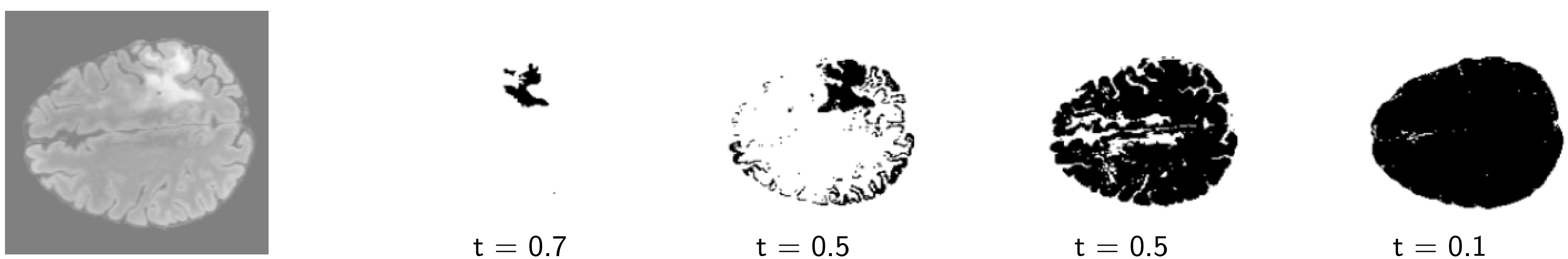
Let $X \subset \mathbb{R}^n$ finite. For $t \geq 0$, define the t-**thickening** $X^t = \{y \in \mathbb{R}^n \mid \exists x \in X, \|x - y\| \leq t\}$.



Let $f\colon \mathcal{M} \to \mathbb{R}$ continuous. For $t \in \mathbb{R}$, consider the t-**sublevel set** $f^t = f^{-1}((-\infty, t])$.



Let $I\colon [0,1]^3 \to [0,1]$ be an image. For $t \in [0,1]$, consider the t-**superlevel set** $I^t = I^{-1}([t,1))$.



Given a filtration

$$\cdots\cdots \to I^{t_1} \xhookrightarrow{i_{t_1}^{t_2}} I^{t_2} \xhookrightarrow{i_{t_2}^{t_3}} I^{t_3} \xhookrightarrow{i_{t_3}^{t_4}} I^{t_4} \hookrightarrow \cdots\cdots$$

one applies the homology functor

$$\cdots\cdots \to H_i(I^{t_1}) \xrightarrow{(i_{t_1}^{t_2})_*} H_i(I^{t_2}) \xrightarrow{(i_{t_2}^{t_3})_*} H_i(I^{t_3}) \xrightarrow{(i_{t_3}^{t_4})_*} H_i(I^{t_4}) \cdots\cdots$$

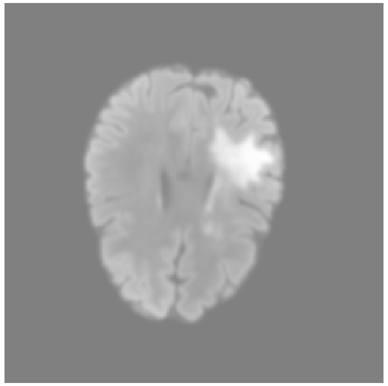
Tracking the cycles: Consider $c \in H_i(I^{t_0})$.

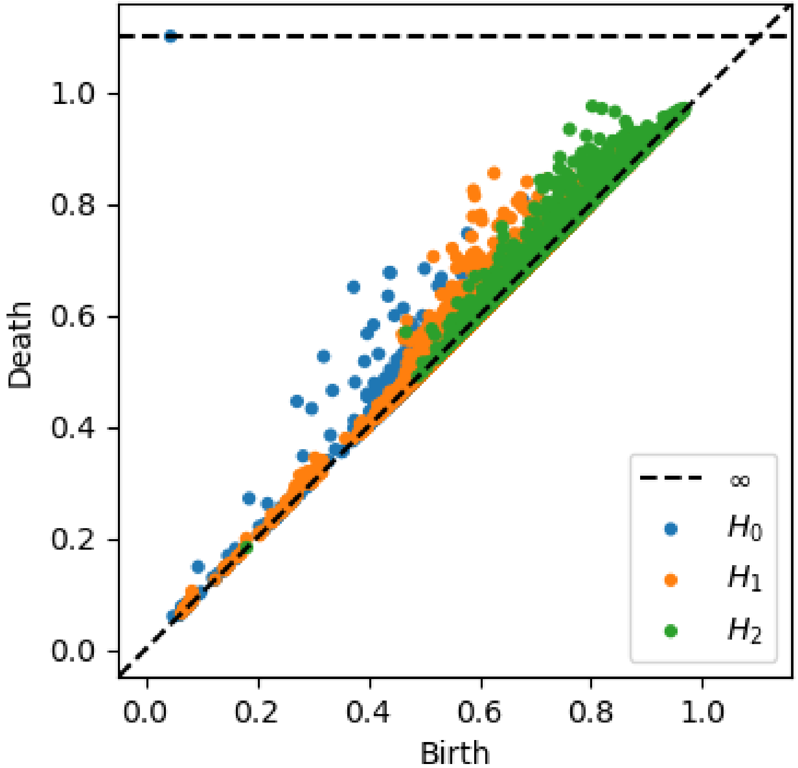
Its **death time** is: $\sup \left\{ t \geq t_0 \mid \left(i_{t_0}^{t}\right)_* (c) \neq 0 \right\}$,

Its **birth time** is: $\inf \left\{ t \leq t_0 \mid \left(i_{t}^{t_0}\right)_*^{-1} (\{c\}) \neq \emptyset \right\}$,

Its **persistence** is the difference.

One can define a **persistence diagram**.
It is a multiset of points (b, d), with $b \leq d$.





Given a filtration

$$\dashrightarrow I^{t_1} \xhookrightarrow{i_{t_1}^{t_2}} I^{t_2} \xhookrightarrow{i_{t_2}^{t_3}} I^{t_3} \xhookrightarrow{i_{t_3}^{t_4}} I^{t_4} \hookrightarrow \dashrightarrow$$

one applies the homology functor

$$\dashrightarrow H_i(I^{t_1}) \xrightarrow{(i_{t_1}^{t_2})_*} H_i(I^{t_2}) \xrightarrow{(i_{t_2}^{t_3})_*} H_i(I^{t_3}) \xrightarrow{(i_{t_3}^{t_4})_*} H_i(I^{t_4}) \dashrightarrow$$

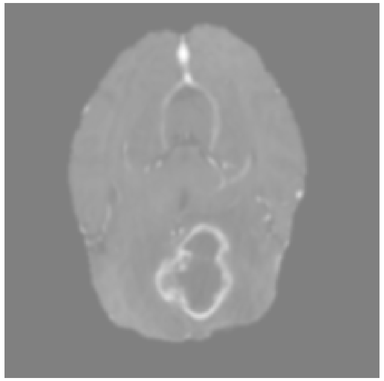
Tracking the cycles: Consider $c \in H_i(I^{t_0})$.

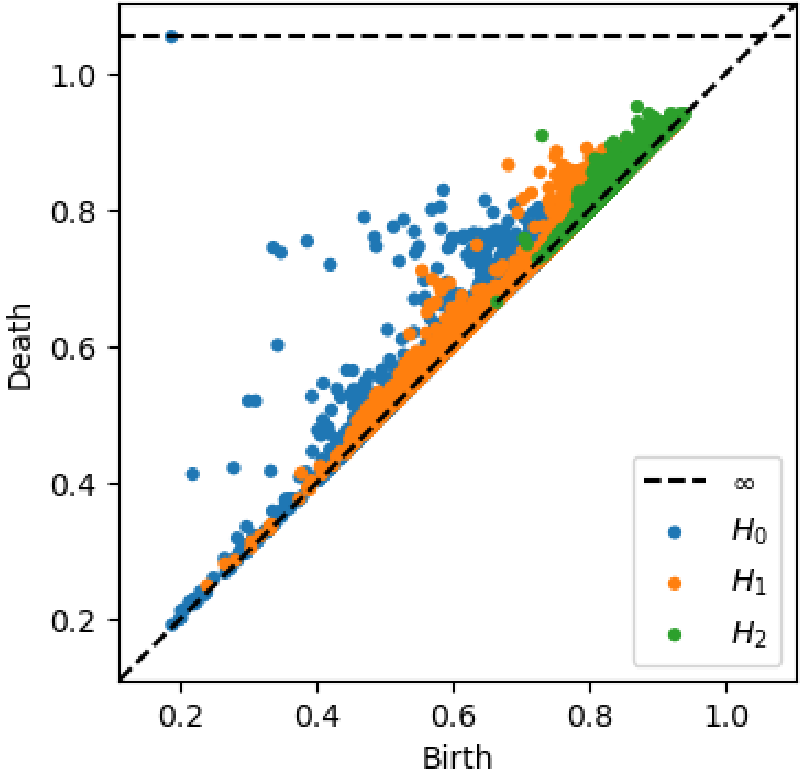
Its **death time** is: $\sup \left\{ t \geq t_0 \mid \left(i_{t_0}^{t}\right)_* (c) \neq 0 \right\}$,

Its **birth time** is: $\inf \left\{ t \leq t_0 \mid \left(i_{t}^{t_0}\right)_*^{-1} (\{c\}) \neq \emptyset \right\}$,

Its **persistence** is the difference.

One can define a **persistence diagram**.
It is a multiset of points (b, d), with $b \leq d$.





Definition: Let k be a field. A **persistence module** is a functor $(\mathbb{R}, \leq) \to k\text{-}\mathbf{Vect}$.
In other words, it is a pair

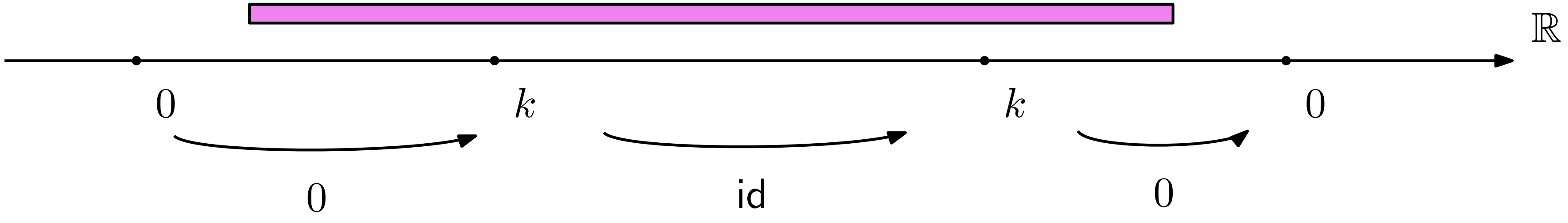
$$\mathbb{V} = \left((V^t)_{t\in\mathbb{R}},\ (v_s^t \colon V^s \to V^t)_{s\leq t\in\mathbb{R}}\right)$$

where V^t are vector spaces over k, and v_s^t are linear maps such that

- $\forall t \in \mathbb{R}$, $v_t^t = \mathrm{id}$,
- $\forall r, s, t \in \mathbb{R}$ such that $r \leq s \leq t$, one has $v_s^t \circ v_r^s = v_r^t$.

Definition: Let $S \subset \mathbb{R}$ be an interval. The **interval-module** associated to S is the persistence module $\mathbb{V}[S]$ with vector spaces and linear maps defined as

$$V^t = \begin{cases} k & \text{if } t \in S, \\ 0 & \text{else,} \end{cases} \quad \text{and} \quad v_s^t = \begin{cases} \mathrm{id} & \text{if } s, t \in S, \\ 0 & \text{else.} \end{cases}$$



Definition: Let k be a field. A **persistence module** is a functor $(\mathbb{R}, \leq) \to k$-$\mathbf{Vect}$. In other words, it is a pair

$$\mathbb{V} = \big((V^t)_{t\in\mathbb{R}},\ (v_s^t\colon V^s \to V^t)_{s\leq t\in\mathbb{R}}\big)$$

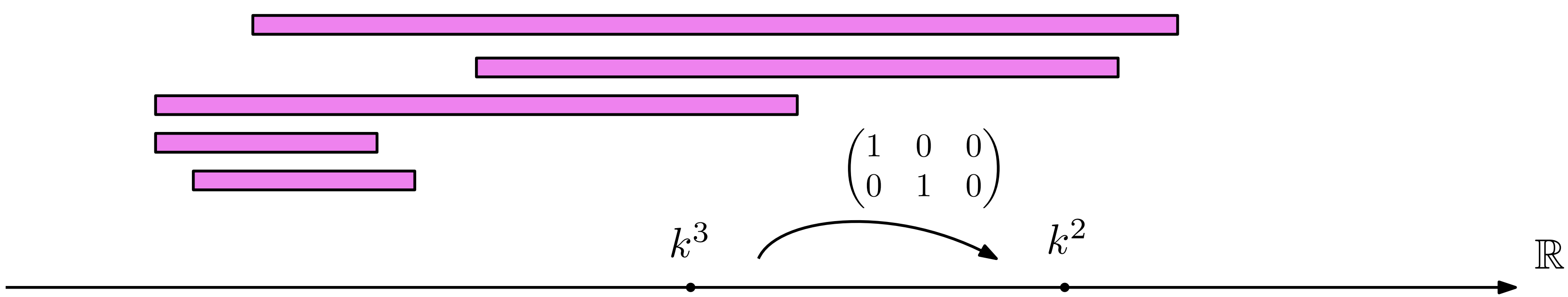
where V^t are vector spaces over k, and v_s^t are linear maps such that

- $\forall t \in \mathbb{R}$, $v_t^t = \mathrm{id}$,
- $\forall r, s, t \in \mathbb{R}$ such that $r \leq s \leq t$, one has $v_s^t \circ v_r^s = v_r^t$.

Definition: Let $S \subset \mathbb{R}$ be an interval. The **interval-module** associated to S is the persistence module $\mathbb{V}[S]$ with vector spaces and linear maps defined as

$$V^t = \begin{cases} k & \text{if } t \in S, \\ 0 & \text{else,} \end{cases} \qquad \text{and} \qquad v_s^t = \begin{cases} \mathrm{id} & \text{if } s, t \in S, \\ 0 & \text{else.} \end{cases}$$

One can sum interval-modules:



A persistence module $\mathbb{V}$ **decomposes into interval-modules** if there exists a multiset $\mathcal{B}$ of intervals such that

$$\mathbb{V} \simeq \bigoplus_{S \in \mathcal{B}} \mathbb{V}[S].$$

Theorem (Crawley-Boevey, 2015): A pointwise finite-dimensional persistence module decomposes into interval-modules.

[Zomorodian, Carlsson, Computing Persistent Homology, 2004]

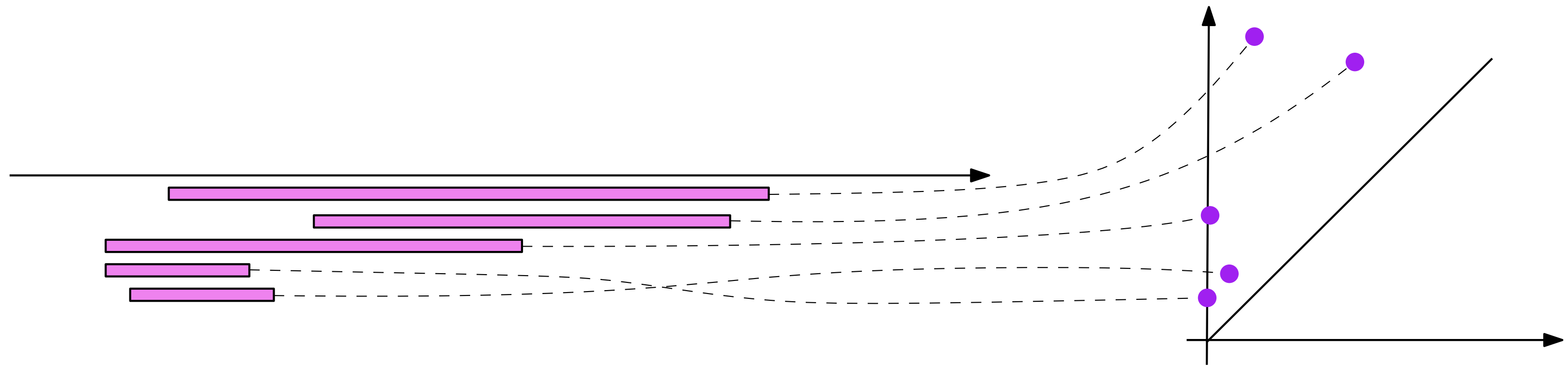
[Chazal, de Silva, Glisse, Oudot, The Structure and Stability of Persistence Modules, 2012]

[Crawley-Boevey, Decomposition of pointwise finite-dimensional persistence modules, 2015]

[Botnan, Crawley-Boevey, Decomposition of persistence modules, 2020]

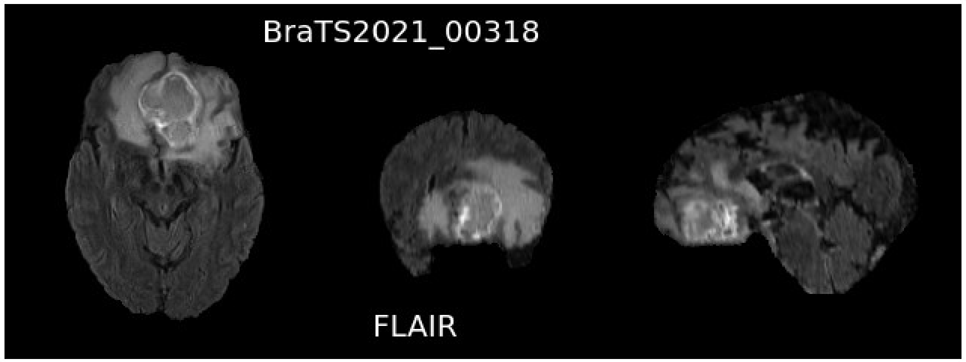
Theorem (consequence of Krull-Remak-Schmidt-Azumaya): If such a $\mathcal{B}$ exists, then it is unique.

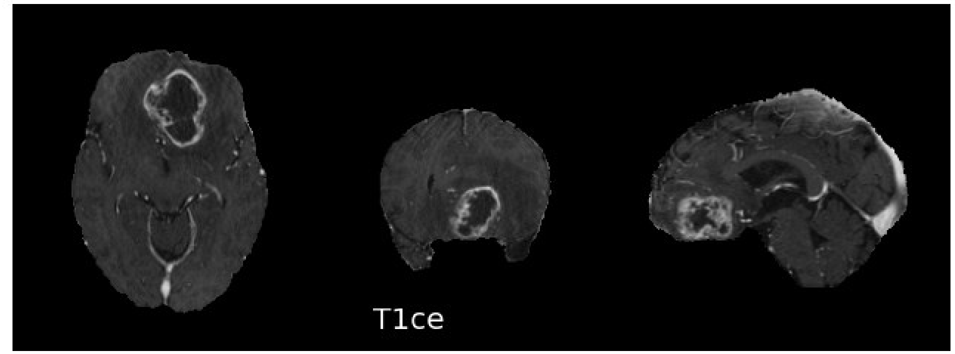
In this case, the multiset $\mathcal{B}$ is called the **persistence barcode** of $\mathbb{V}$.
Seen as a subset of $\mathbb{R}^2$, it is called the **persistence diagram**.

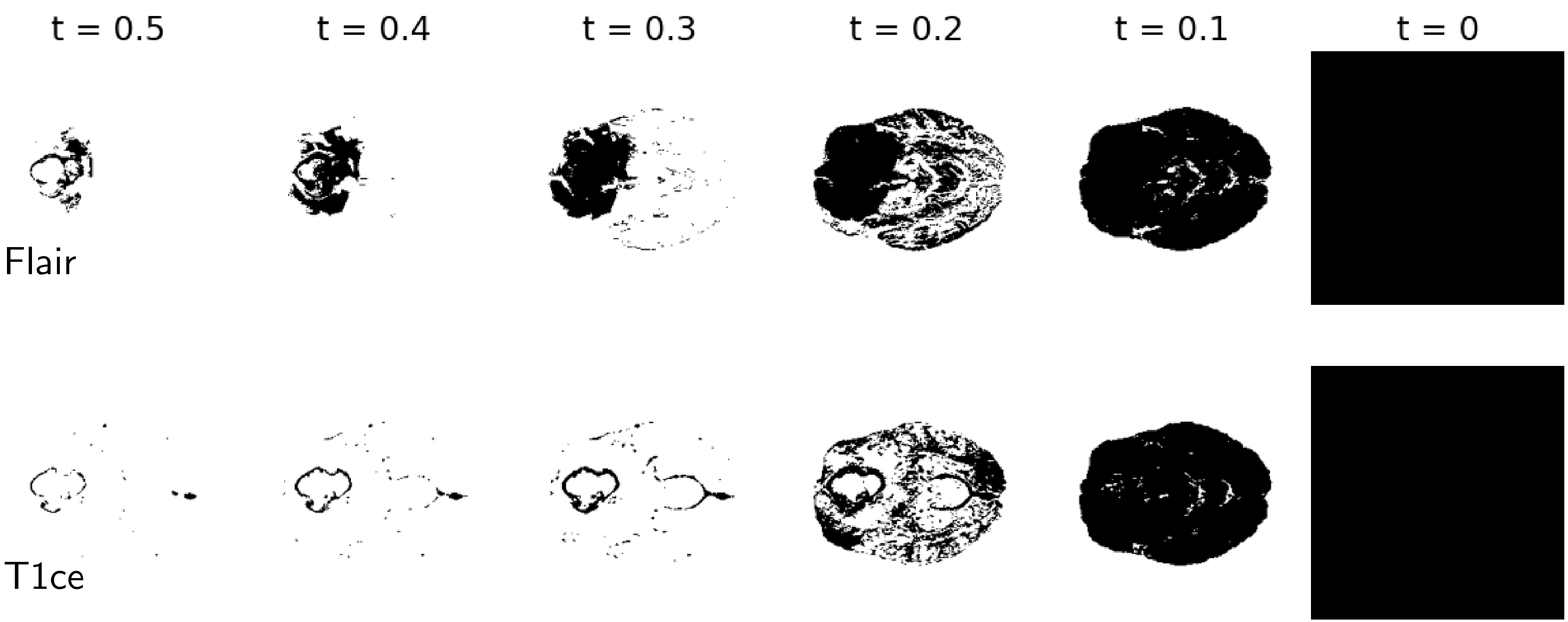


Superlevel set persistence of brain MRIs

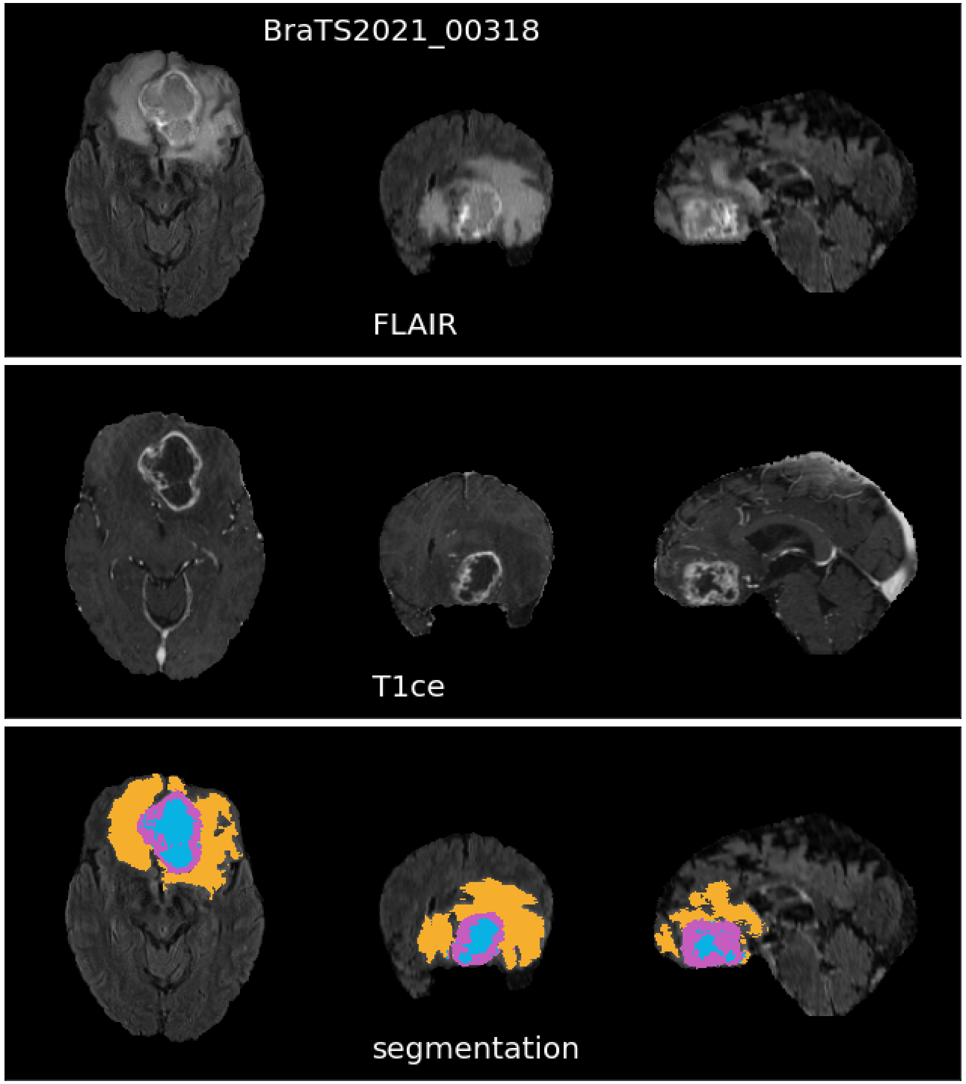
Consider the superlevel sets of Flair and T1ce modalities: $I^t = I^{-1}([t, 1))$, where $I\colon [0,1]^3 \to [0,1]$.







Superlevel set persistence of brain MRIs



Persistence of Flair: the whole tumor is represented by a persistent connected component.

Persistence of T1ce: the Enhancing Tumor induces a persistent cycle in H_2.

Our strategy:

1. Identification of whole tumor (in Flair),
2. Detection of Enhancing Tumor (in T1ce),
3. Deduction of other components (Peritumoral Edema, Tumorous Core)

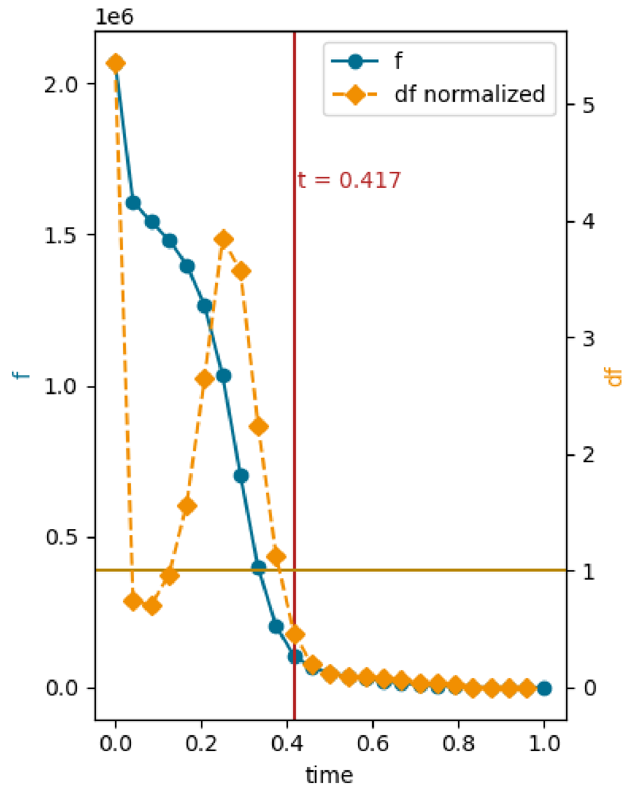
Notations: Images I_{FLAIR} and $I_{\mathrm{T1ce}} \colon \Omega \to [0,1]$.
The three components are denoted X_{ET}, X_{TC} and X_{ED}. Their union, X_{WT}, is the whole tumour.

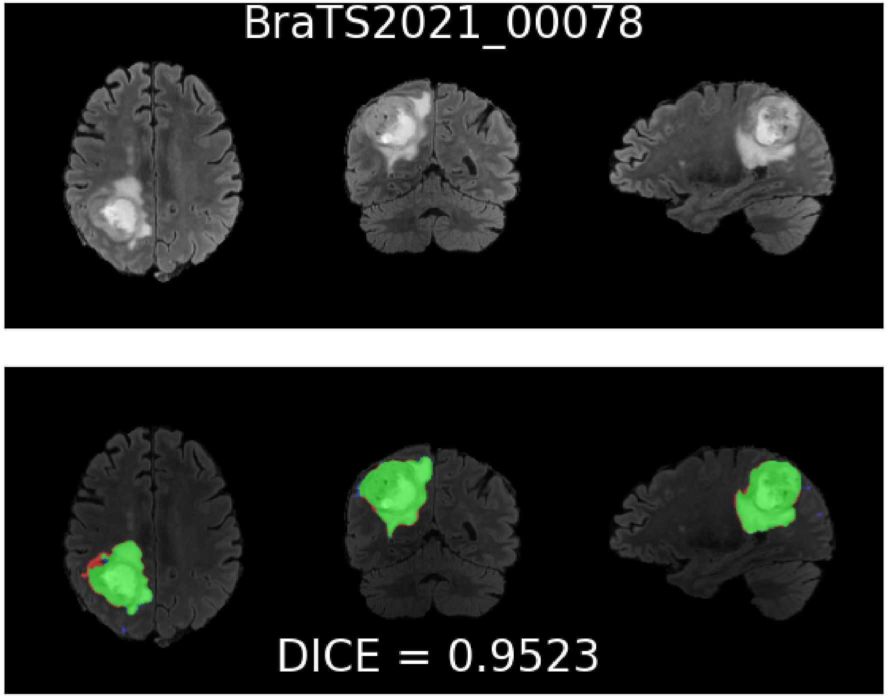
Idea: Select the largest hyper-intense region present in Flair, supposedly corresponding to X_{WT}.

Let $t \mapsto \# I^t_{\mathrm{FLAIR}}$ number of voxels of intensity $\geq t$, and $t \mapsto \mathrm{d}\# I^t_{\mathrm{FLAIR}}$ its derivative (normalized).

Identify the first value t (starting from 1) for which $\mathrm{d}\# I^t_{\mathrm{FLAIR}} \geq$ `dt_threshold` (fixed parameter).

Last define X_{WT} as the largest connected component of I^t_{FLAIR}.





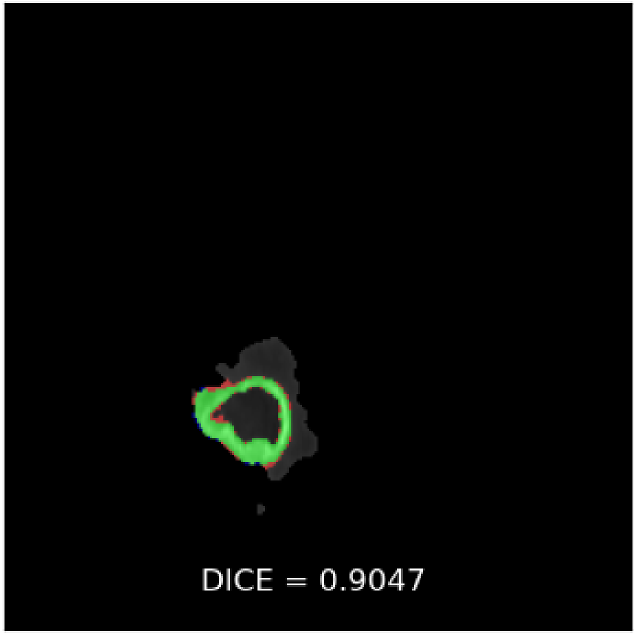
This is a sort of Otsu's binarization method.

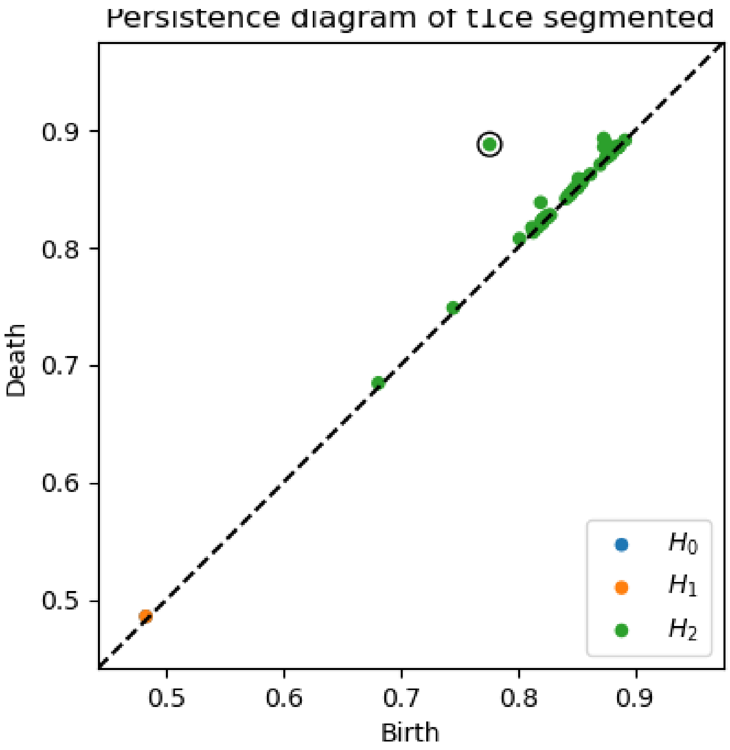
Idea: Select the spherical boundary of the tumour, supposedly corresponding to X_{ET}.

Compute the persistent homology of the superlevel sets of image I_{T1ce} restricted to X_{WT}.

Select the H_2-feature of highest persistence (i.e., point (t_b, t_d) that maximizes $|t_d - t_b|$).

Let $x_b \in \Omega$ be the voxel that gave birth to it, and define X_{ET} as its connected component in $I^{t_b}_{\mathrm{T1ce}}$.





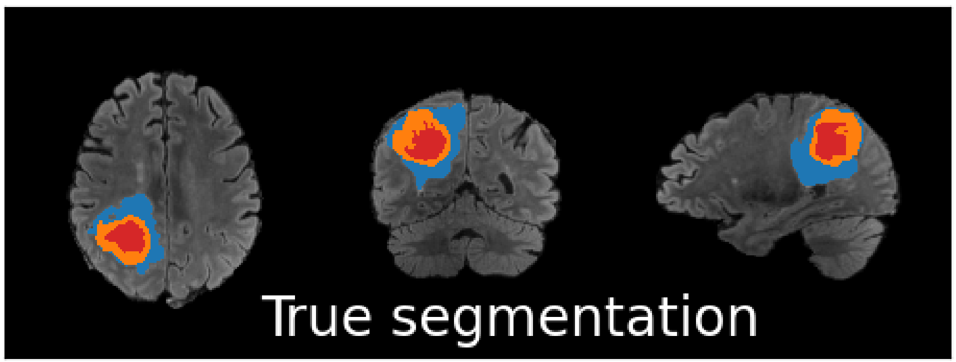
Remark: This connected component may not be a representative cycle of the homology class.

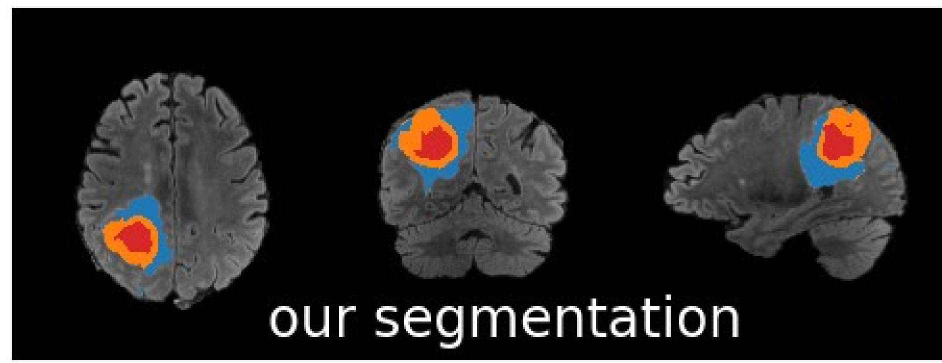
Idea: Select the interior and exterior of X_{ET}, supposedly corresponding to X_{TC} and X_{ED}.

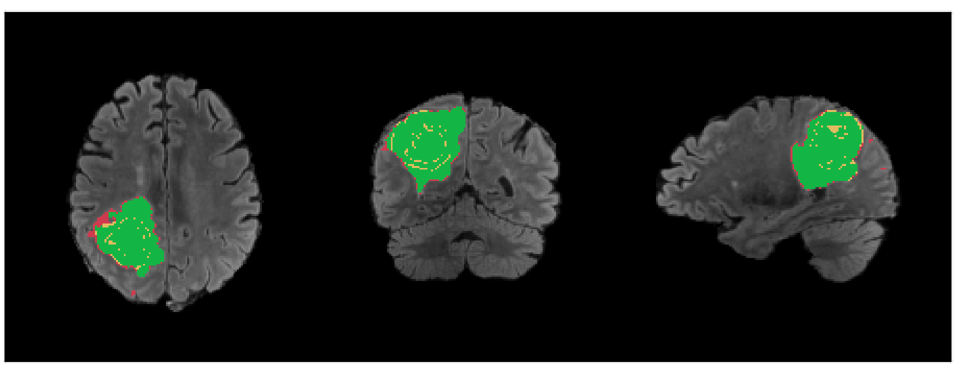
Consider the subset $X_{\mathrm{WT}} \setminus X_{\mathrm{ET}} \subset \Omega$, and compute its connected components.

The outer component (that in contact with the background) is saved in X_{ED}.

The others are considered inner and are added to X_{TC}.



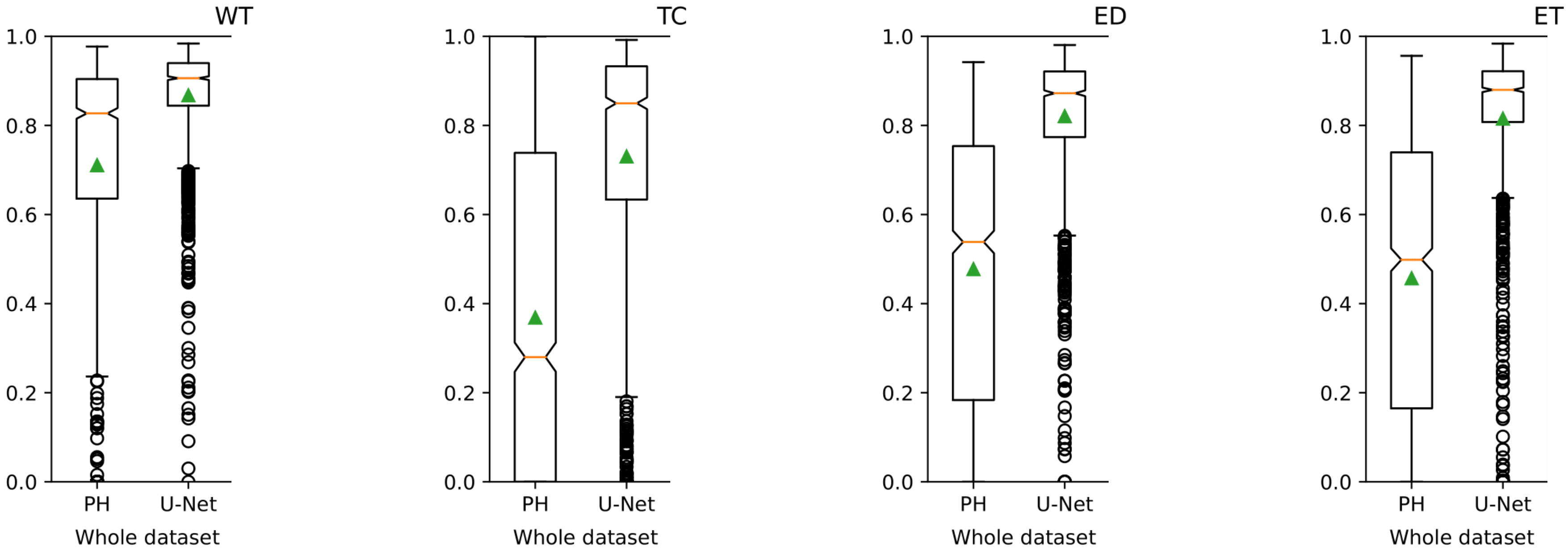




DICE: WT = 0.94, TC = 0.94, ET = 0.90, ED = 0.89

Dice coefficient between two binary images $X, Y : \Omega \to \{0,1\}$ is

$$\mathrm{Dice}(X,Y) = \frac{2\#(X \cap Y)}{\#X + \#Y}$$

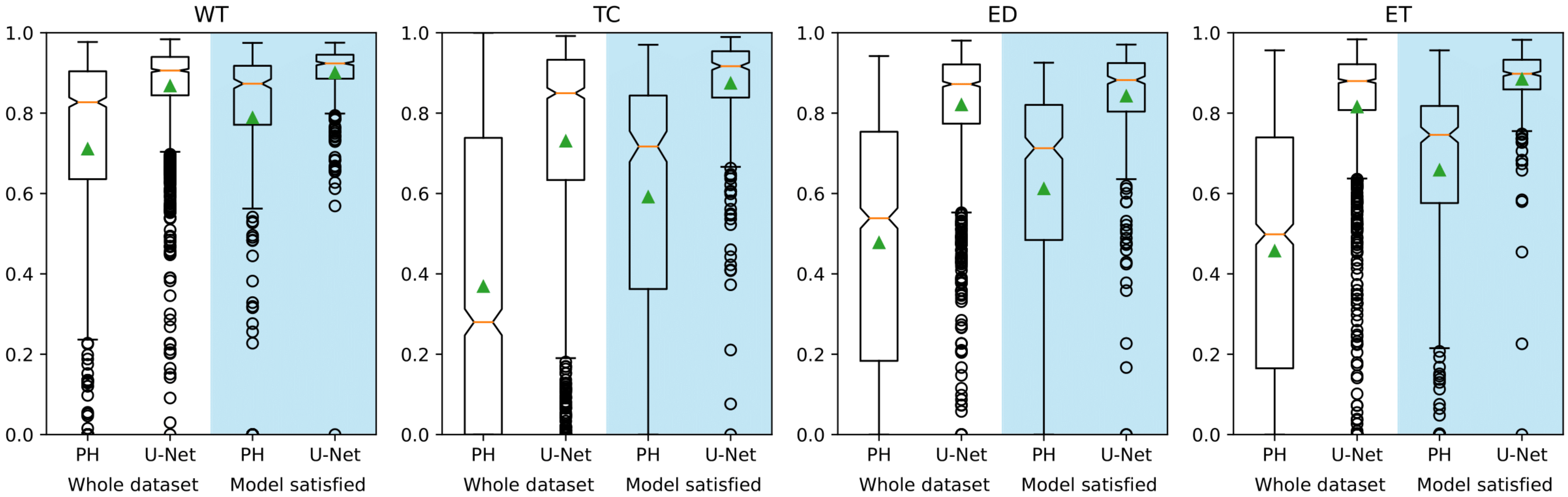


We compare our results with **U-net**, on the whole BraTS 2021 dataset (1521 MRIs).

Results

Dice coefficient between two binary images $X, Y : \Omega \to \{0, 1\}$ is

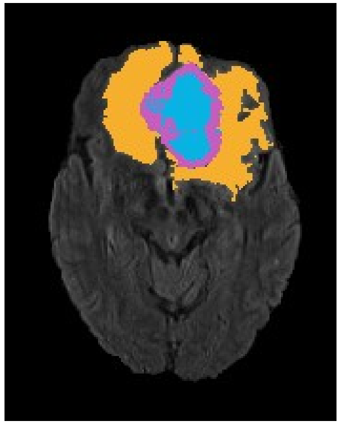
$$\mathrm{Dice}(X, Y) = \frac{2\#(X \cap Y)}{\#X + \#Y}$$



We compare our results with **U-net**, on the whole BraTS 2021 dataset (1521 MRIs).

In addition, we restrict the scores to the images satisfying our geometric model (31% of dataset).

Geometric model: Let X_{TW}, X_{ED}, X_{TC}, and X_{ET} be the classes of grountruth segmentation.



Peritumoral Edema (ED),
Tumorous Core (TC),
Enhancing Tumor (ET).

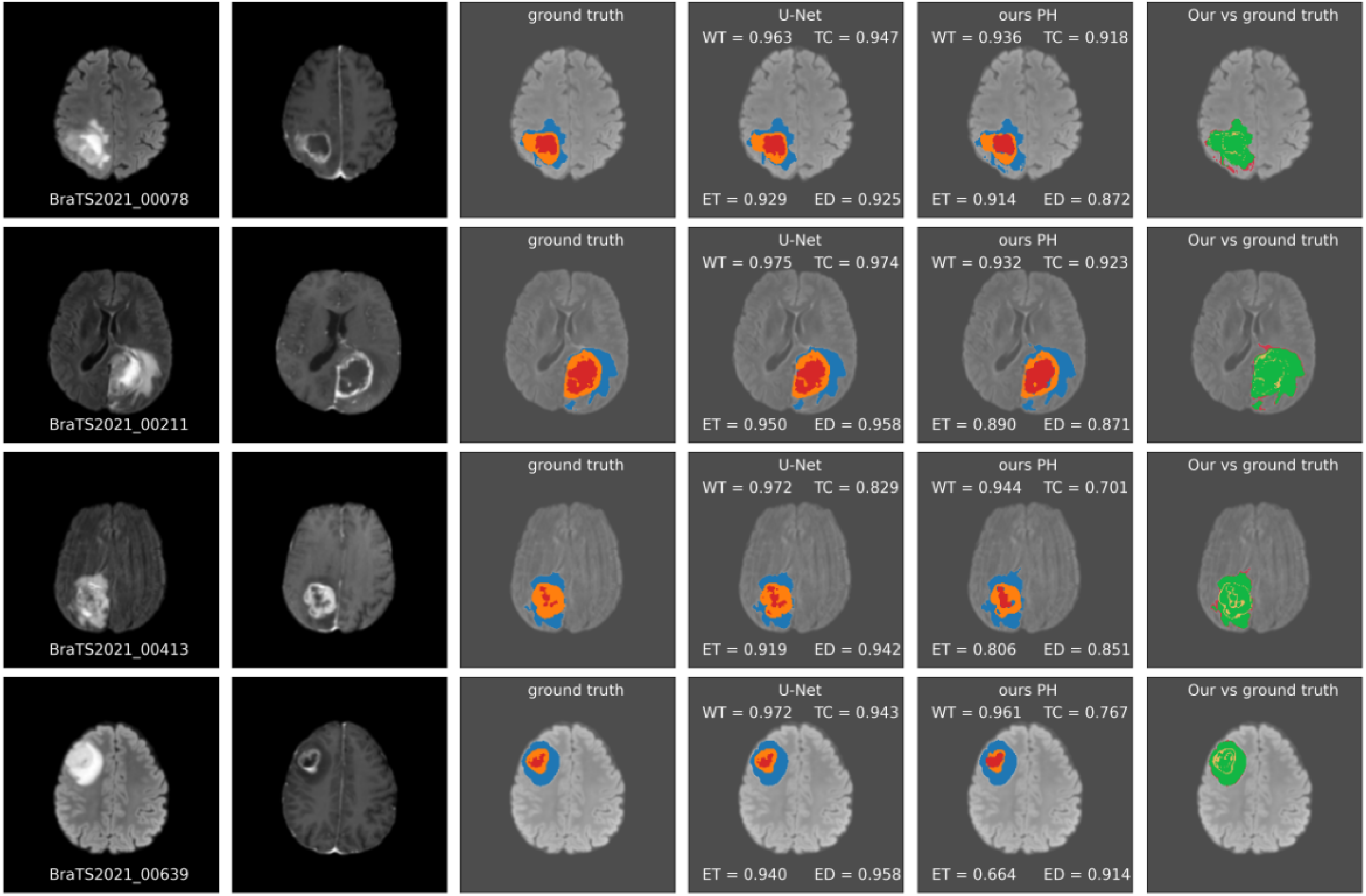
WT is a hyperintense cluster: X_{TW} consists of one connected component, or potentially more, the other ones being 10 times smaller. The most intense voxel WT in FLAIR belongs to X_{TC} or X_{ET}.

ET is sphere-like: After 3 binary dilations, X_{ET} divides the space into two connected components. Moreover, the most intense voxel of WT in T1ce belongs to X_{ET}.

TC (resp. ED) is inside (resp. outside): Applying a binary dilatation to X_{TC} (resp. X_{ED}) yields new pixels of which at least (resp. at most) half belongs to X_{ET}.

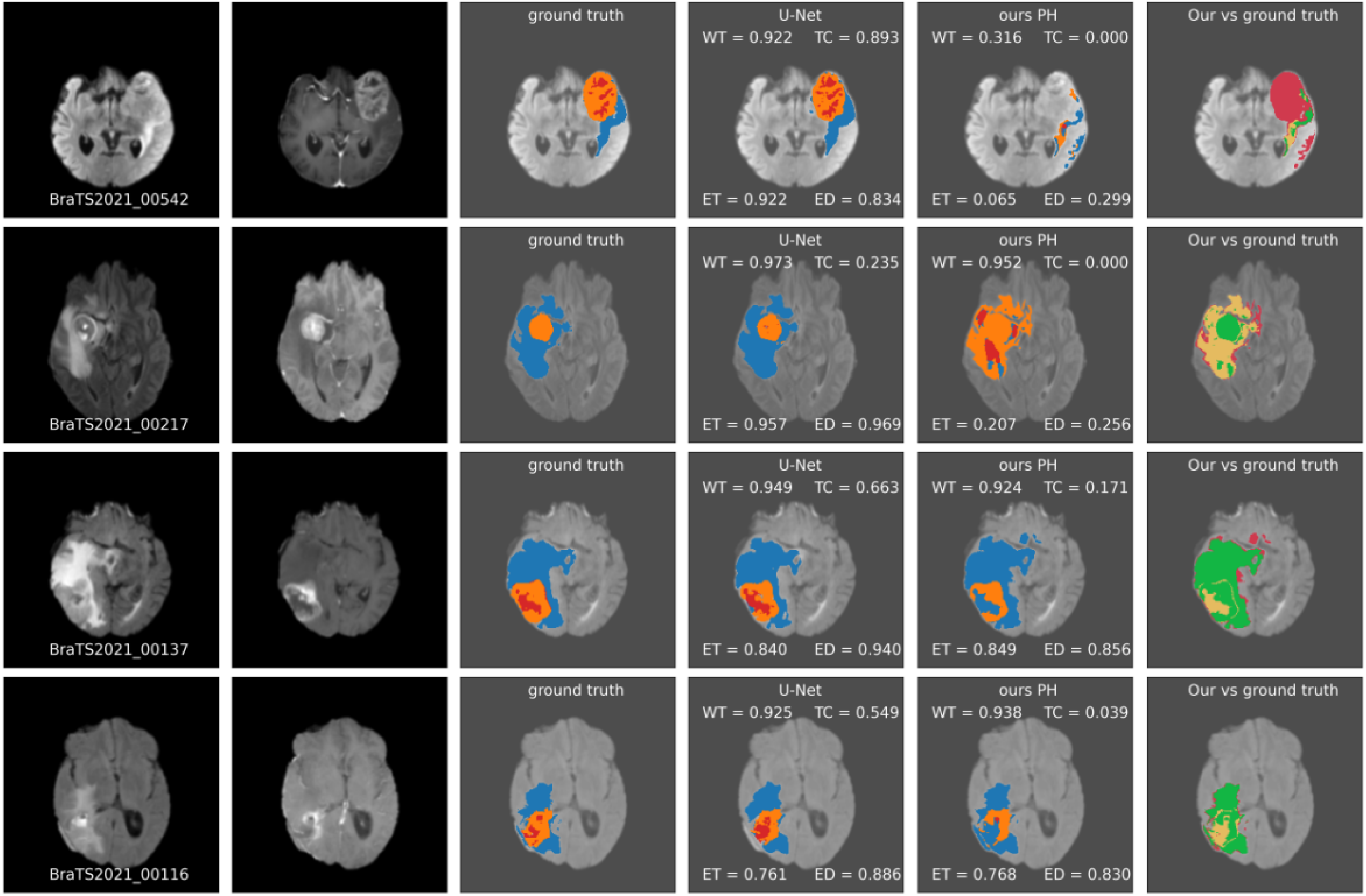
31% of the dataset satisfy this model.

Results



Cases where the model is valid

Results

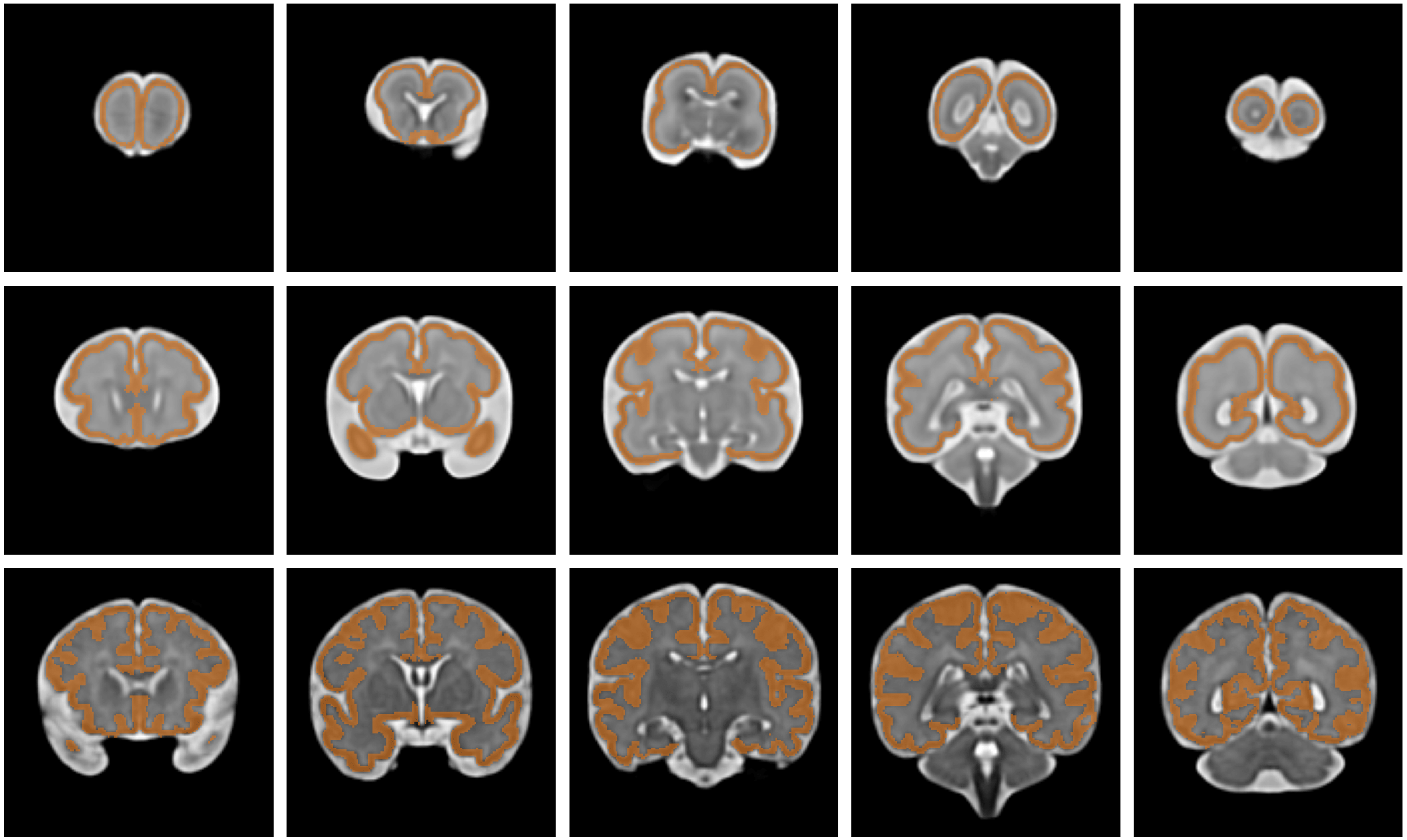


Cases where the model is not valid

Fetal plate segmentation

Objective: cortical plate segmentation in MRI (modality T2).

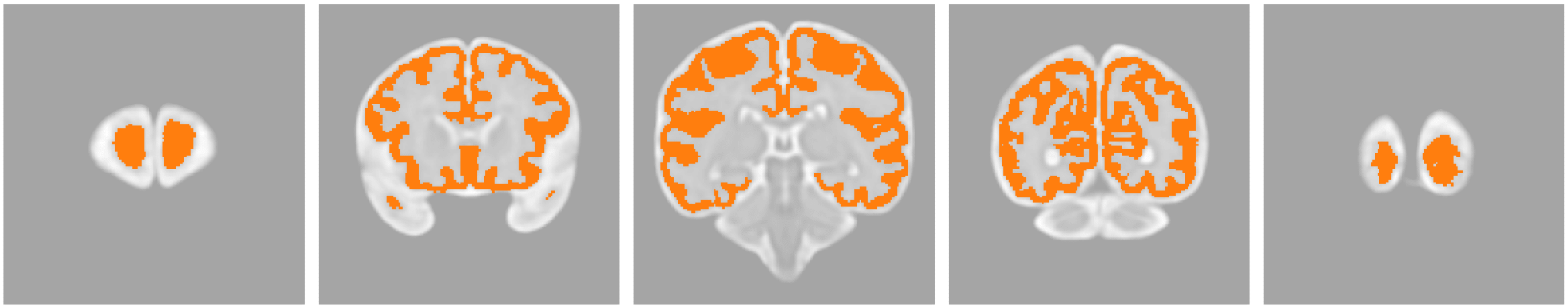
Dataset: Spatiotemporal Atlas (STA), one-week intervals between 21 and 38 weeks gestational age.



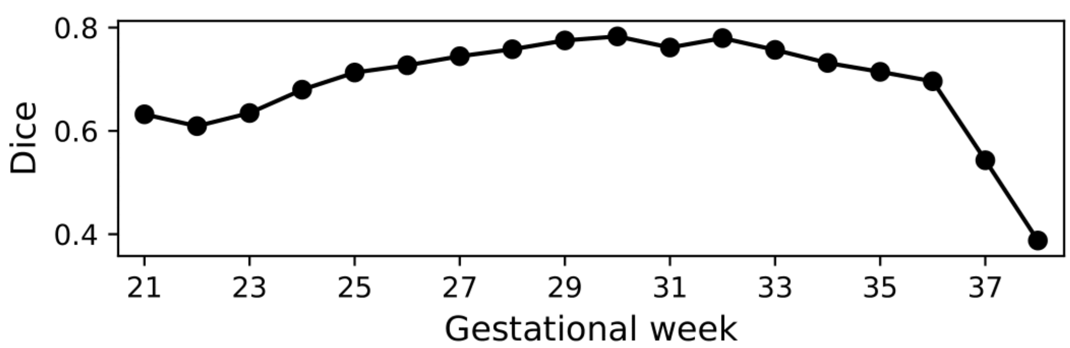
Cortical plate segmentations, for gestational week 21, 30, and 38.

Fetal plate segmentation

In cortical slices, the cortical plate may form a circle, two circles, or a simply connected object, or two connected components.



Strategy: Identify the topology via H_1-persistence.

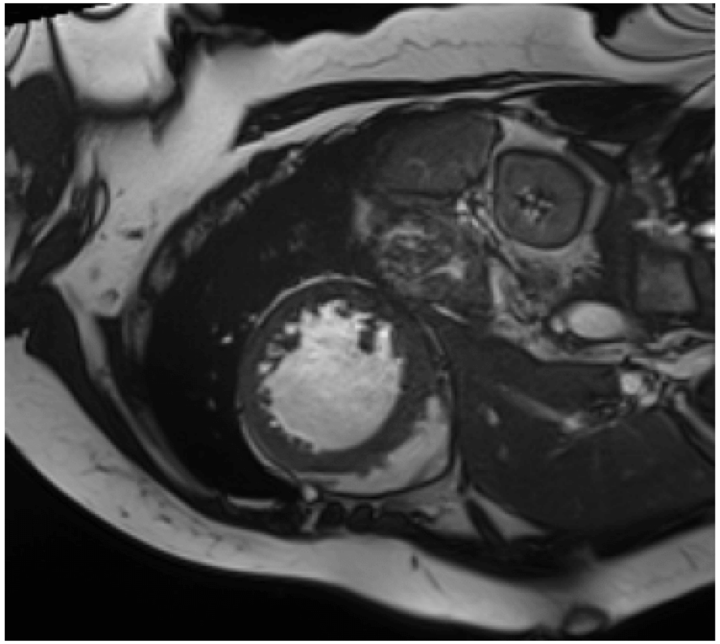


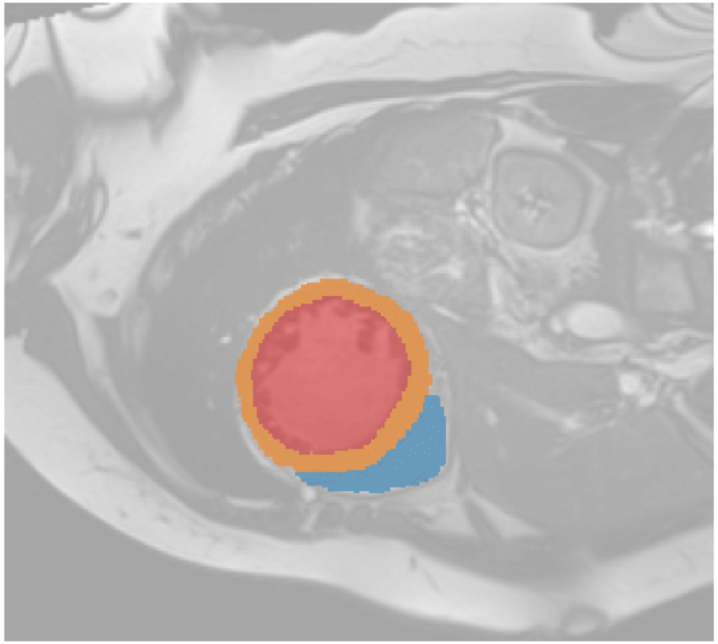
Cardiac segmentation

Objective: coronal segmentation in Magnetic Resonance Images (CMR).

Dataset: Automated Cardiac Diagnosis Challenge (ACDC).
150 patients, two scans (at end diastolic and end systolic phase).

Classes: Myocardium, Right Ventricule, Left Ventricule.

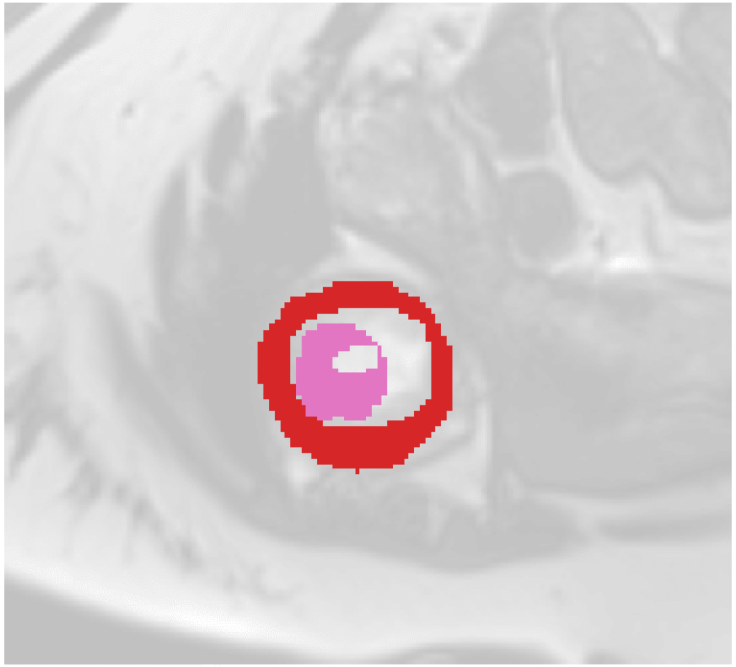




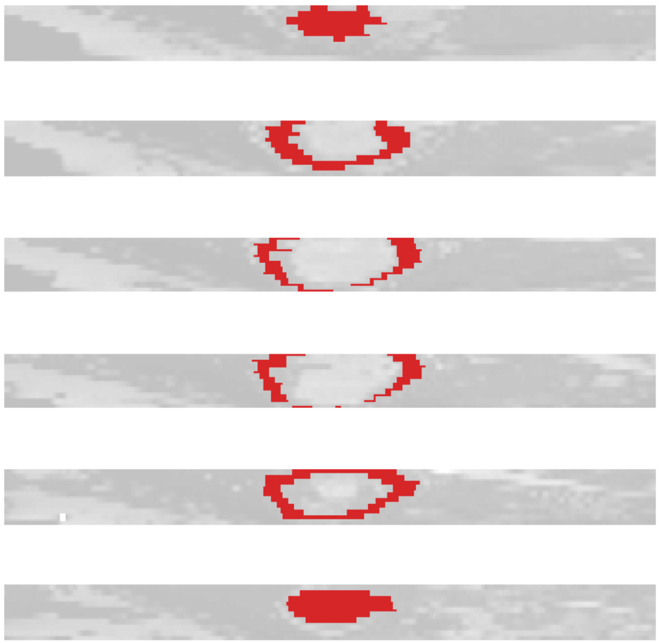
RV and LV: hyperintense.

Myocardium: hypointense, and form a cylinder.

One should study the CMR slice by slice.



Superposition of the segmentation of the myocardium in two consecutive axial slices.

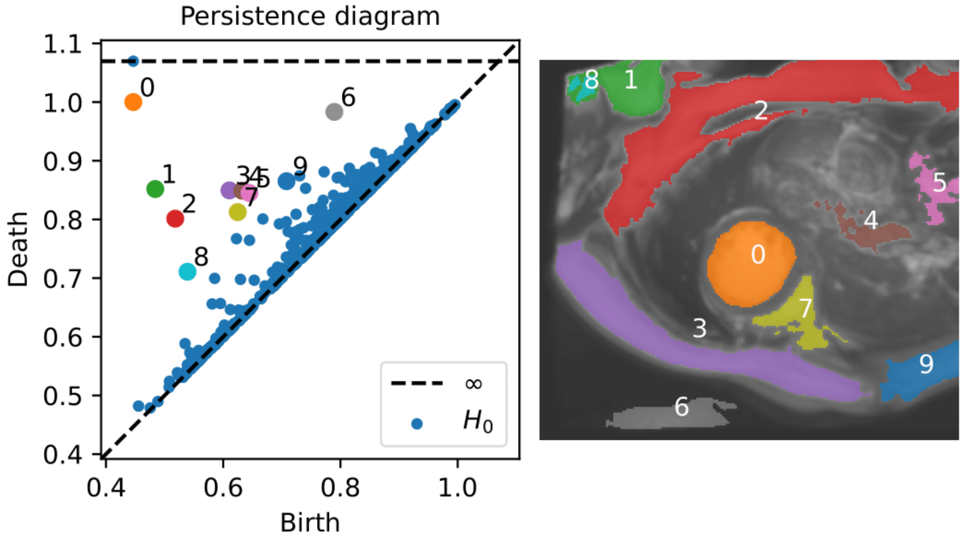


Several coronal slices, with myocardium in red.

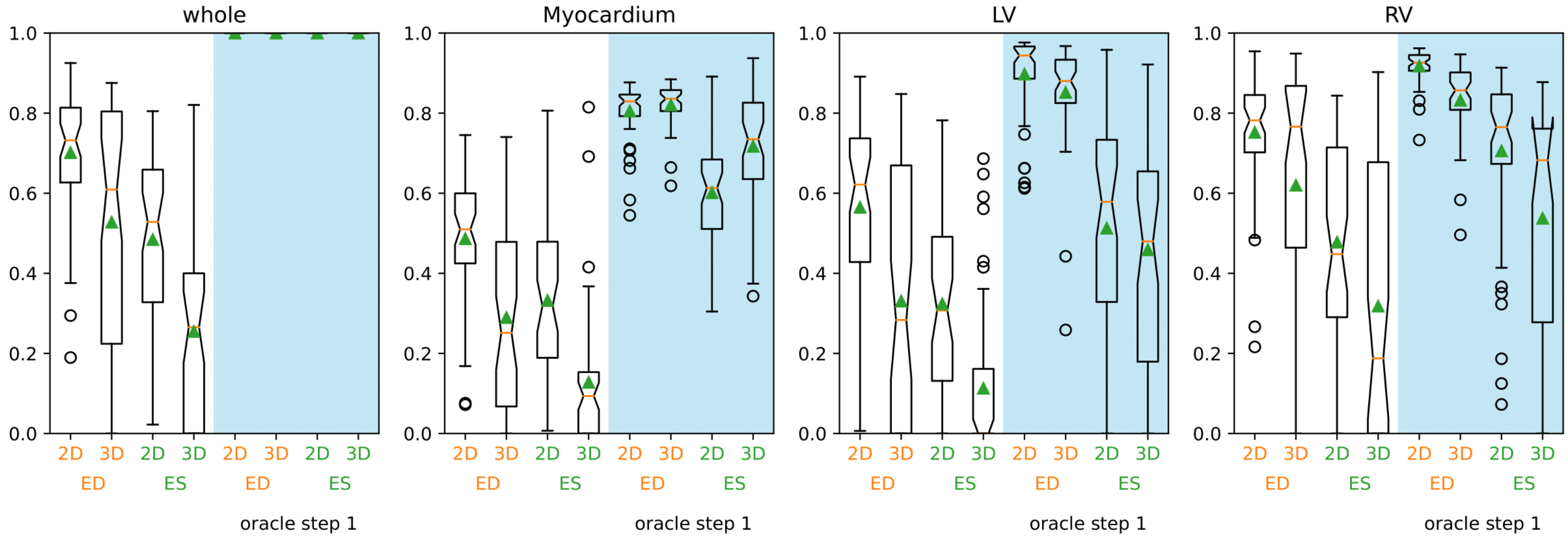
Strategy: Slice by slice,

1. Identification of LV as the most spherical connected component,
2. Detection of RV as the closest connected component to LV,
3. Dilate RV until it reaches LV, and identify the Mocardium as the most persisting H_1-cycle.

We obtain a first segmentation of the image via H_0-persistence.

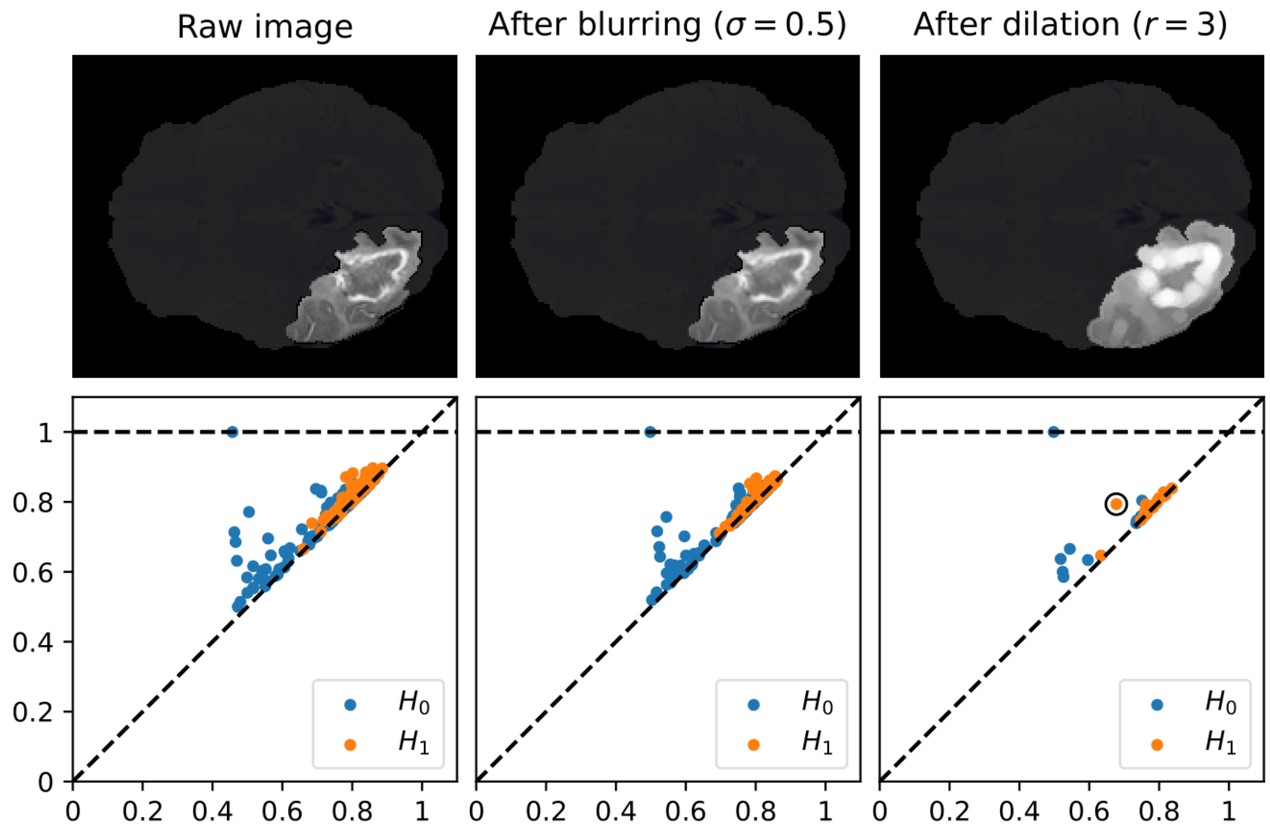


Results (DICE score):



Potential improvements

Preprocessing can enhance the cycles.



Representative cycle identification: We are not extracting representatives of homology classes, but only their connected components.

[Dey, Hirani, Krishnamoorthy, Optimal homologous cycles, total unimodularity, and linear programming, 2010]

[Escolar, Hiraoka, Optimal cycles for persistent homology via linear programming, 2016]

[Obayashi, Volume-optimal cycle: Tightest representative cycle of a generator in persistent homology, 2018]

[Li, Thompson, Henselman-Petrusek, Giusti, Ziegelmeier, Minimal cycle representatives in PH using linear programming, 2021]

[Cohen-Steiner, Lieutier, Vuillamy, Lexicographic optimal homologous chains and applications to point cloud triangulations, 2022]

Conclusion

3D MR images | **Topological Data Analysis** | **Automatic Segmentation**