Automatic Generation of Training Data for Brain Tissue Classification from MRI

Chris A. COCOSCO, Alex P. ZIJDENBOS, and Alan C. EVANS

chris.cocosco@mail.mcgill.ca

http://www.bic.mni.mcgill.ca/users/crisco/

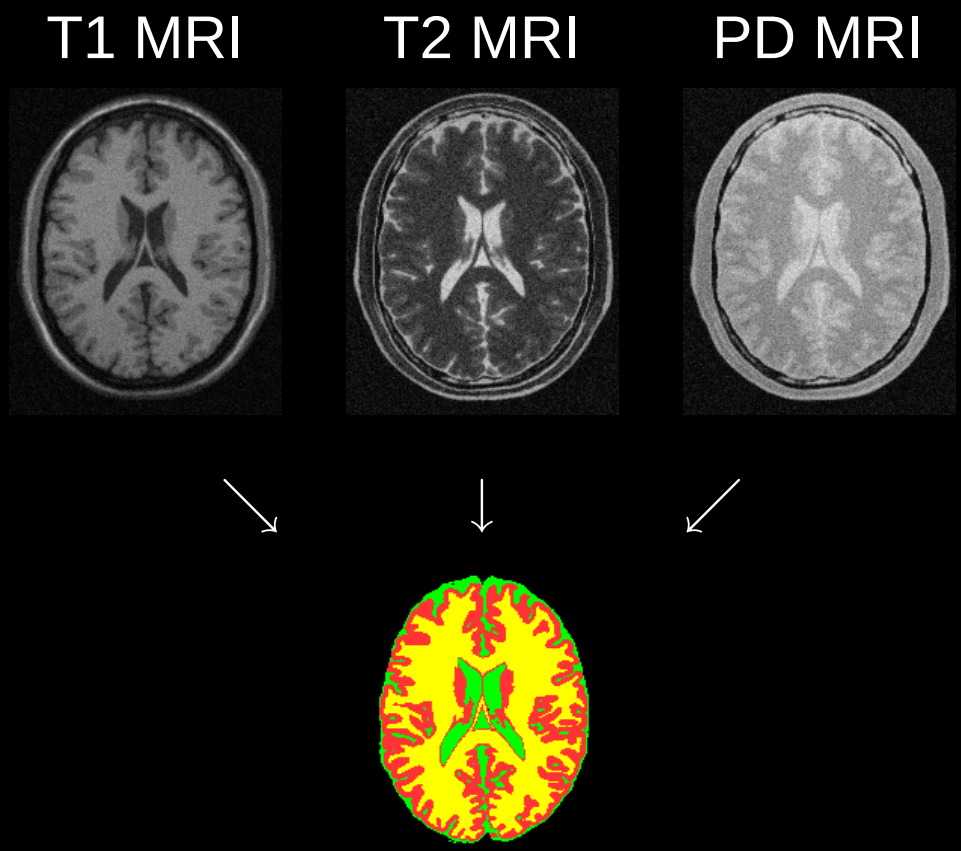
McConnell Brain Imaging Centre

Montreal Neurological Institute

McGill University

Brain tissue classification:

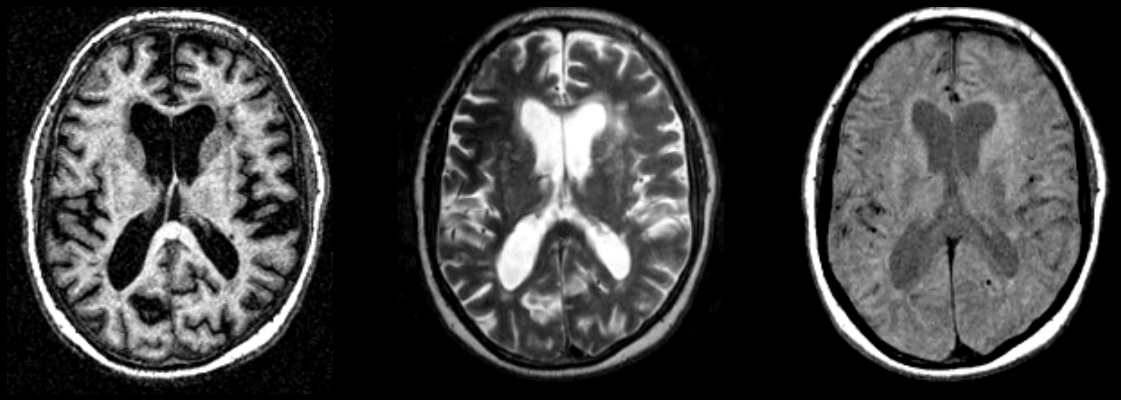
$\Rightarrow$ the procedure of labeling each image voxel as a tissue class.



4 classes: CSF, Grey-matter, White-matter, “background”

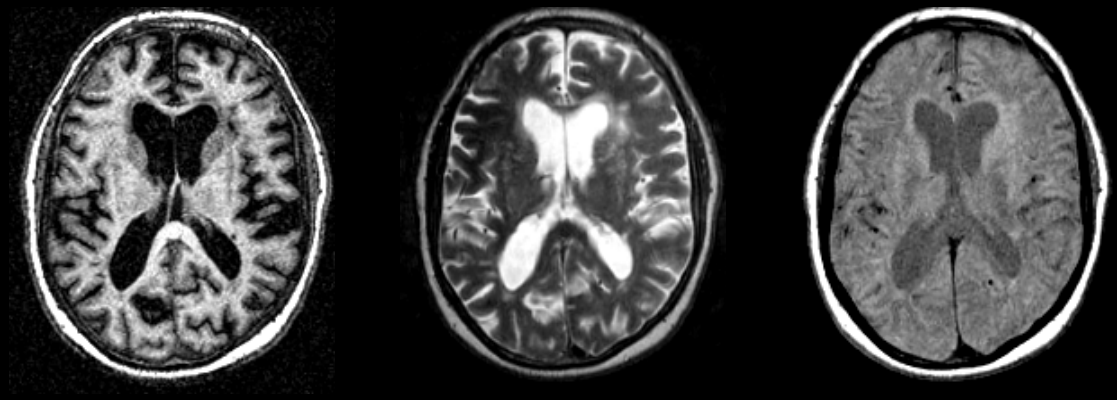
Preview:

Main problem: subject different than anatomical model.



Preview:

Main problem: subject different than anatomical model.



Outline:

- Requirements
- Existing methods
- Our method
- Validation and results

Target application:

- Quantitative measurements, such as:
  - normalized tissue/structure volume
  - atrophy measures
  - voxel-based morphometry
  - cortical surface
  - cortical thickness
  - …
- Studies on a large number of subjects (150 – 1000); data acquired at many different sites.

Requirements:

Tissue classification method should be:

- ACCURATE
  - application: quantitative measurements

Requirements:

Tissue classification method should be:

- ACCURATE
  - application: quantitative measurements
- FULLY AUTOMATIC
  - reproducibility; many datasets to process

Requirements:

Tissue classification method should be:

- ACCURATE
  - application: quantitative measurements
- FULLY AUTOMATIC
  - reproducibility; many datasets to process
- ROBUST against variability in:
  - subject brain's morphology
  - MRI data: image contrast, artifacts, . . .

Existing methods:

- Kamber et al. (IEEE TMI ’95)
- Van Leemput (IEEE TMI ’99), Ashburner (“SPM-99”)

$\hookrightarrow$ all use a probabilistic brain anatomy atlas
$\rightarrow$ problems with brain anatomies significantly different than atlas.

Existing methods: EM

EM-style schemes by Van Leemput (IEEE TMI ’99), Ashburner (“SPM-99”).

Drawback:
assume multi-variate Normal (“Gaussian”) intensity distributions.

$\longrightarrow$ poor assumption for multi-spectral brain aMRI ?

biology, acquisition artifacts . . .

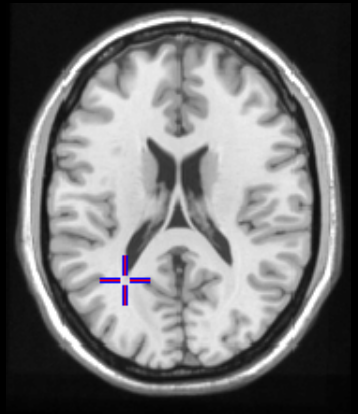
Existing methods: Kamber

Kamber (IEEE TMI ’95) used Tissue Probability Maps (TPM), defined in a stereotaxic space.

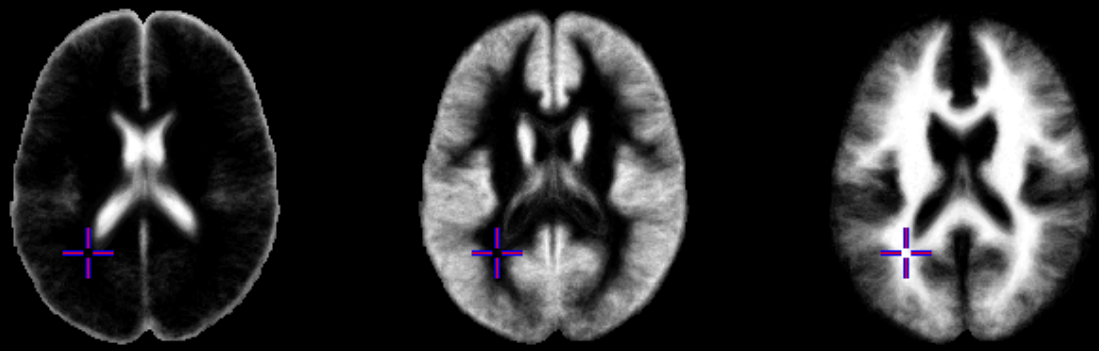
1. subject MRI spatially registered to stereotaxic space (linear registration)
2. select MRI intensity samples from spatial locations very likely to contain a given tissue type
3. use these samples to train a supervised classifier (such as: Bayes, neural network, kNN, . . .).

Stereotaxic space TPM:

Subject MRI:



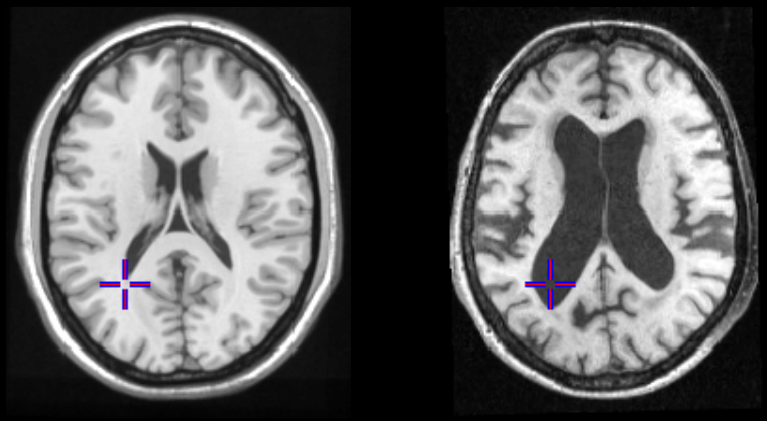
TPM:



CSF = 0% GM = 4% WM = 96%

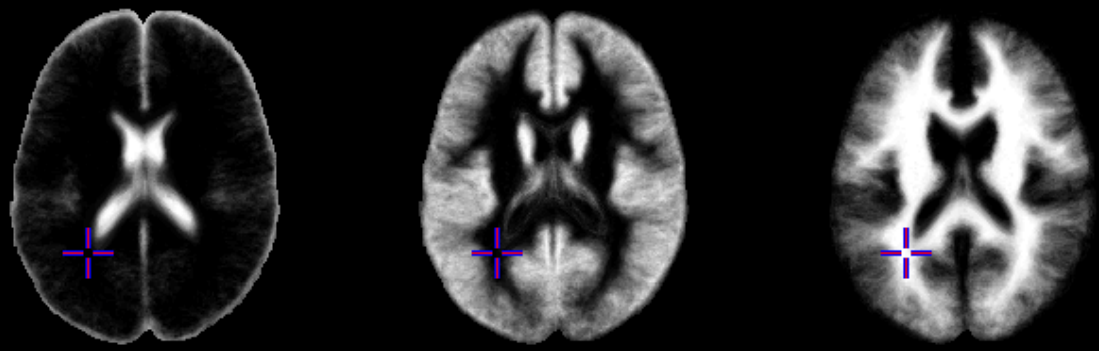
Stereotaxic space TPM:

young
normal
$\longrightarrow$



elderly,
Alzheimer's Disease
$\longleftarrow$

young
normal
population
$\Longrightarrow$

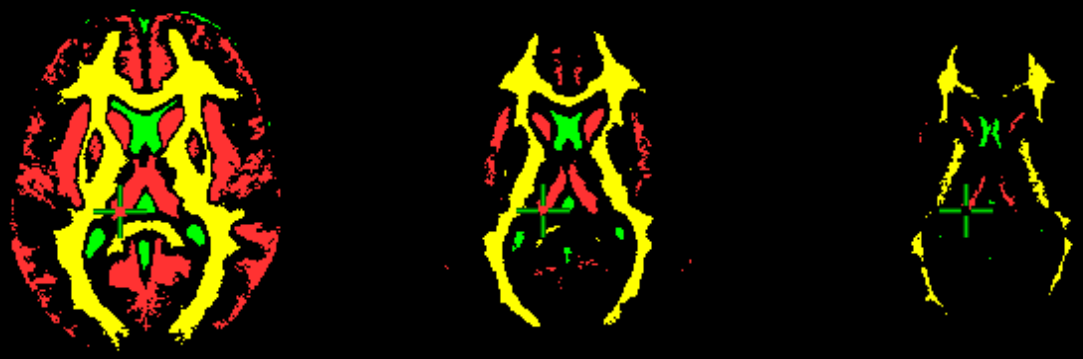


CSF = 0% GM = 4% WM = 96%

Training samples selection:

$\hookrightarrow$ choose spatial locations with TPM value $\geq \mathcal{T}$

$\mathcal{T} =$ 0.7 0.9 0.99

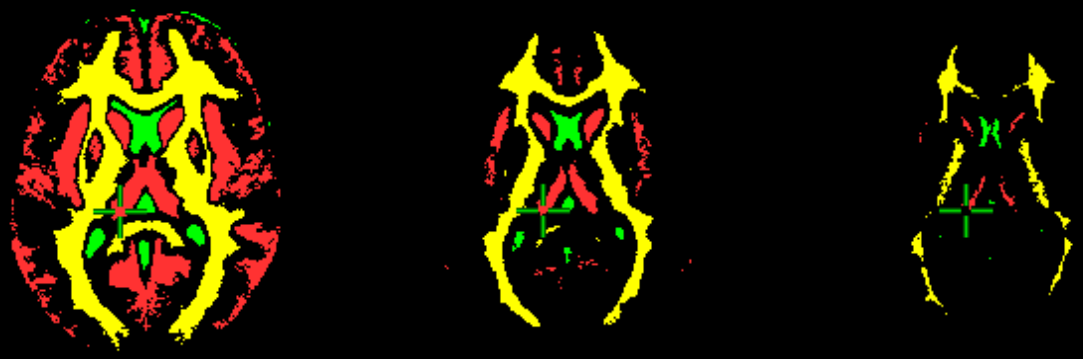


CSF
Grey−matter
White−matter

Training samples selection:

$\hookrightarrow$ choose spatial locations with TPM value $\geq \mathcal{T}$

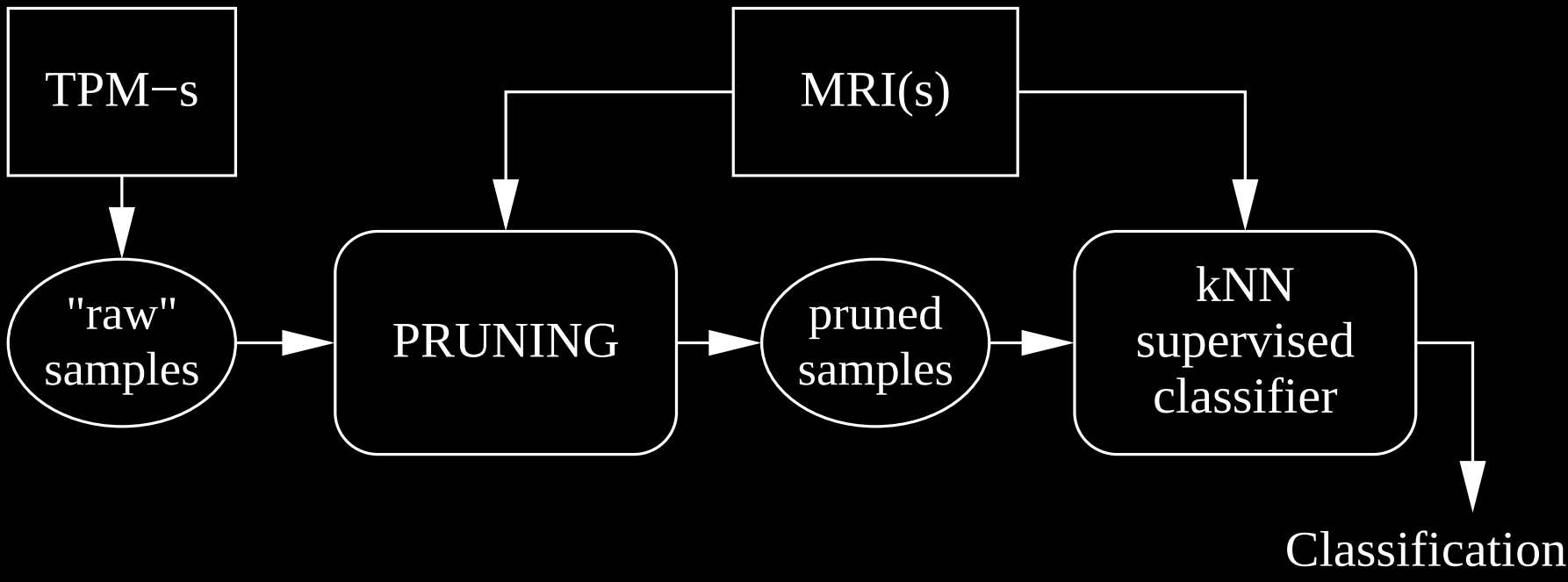
$\mathcal{T} =$ 0.7 0.9 0.99



CSF
Grey−matter
White−matter

- lower $\mathcal{T}$ desired for more spatial coverage $\Rightarrow$ robust estimation of tissue intensity distributions.
- higher $\mathcal{T}$ desired for reducing wrong class guesses.

Our novel method:



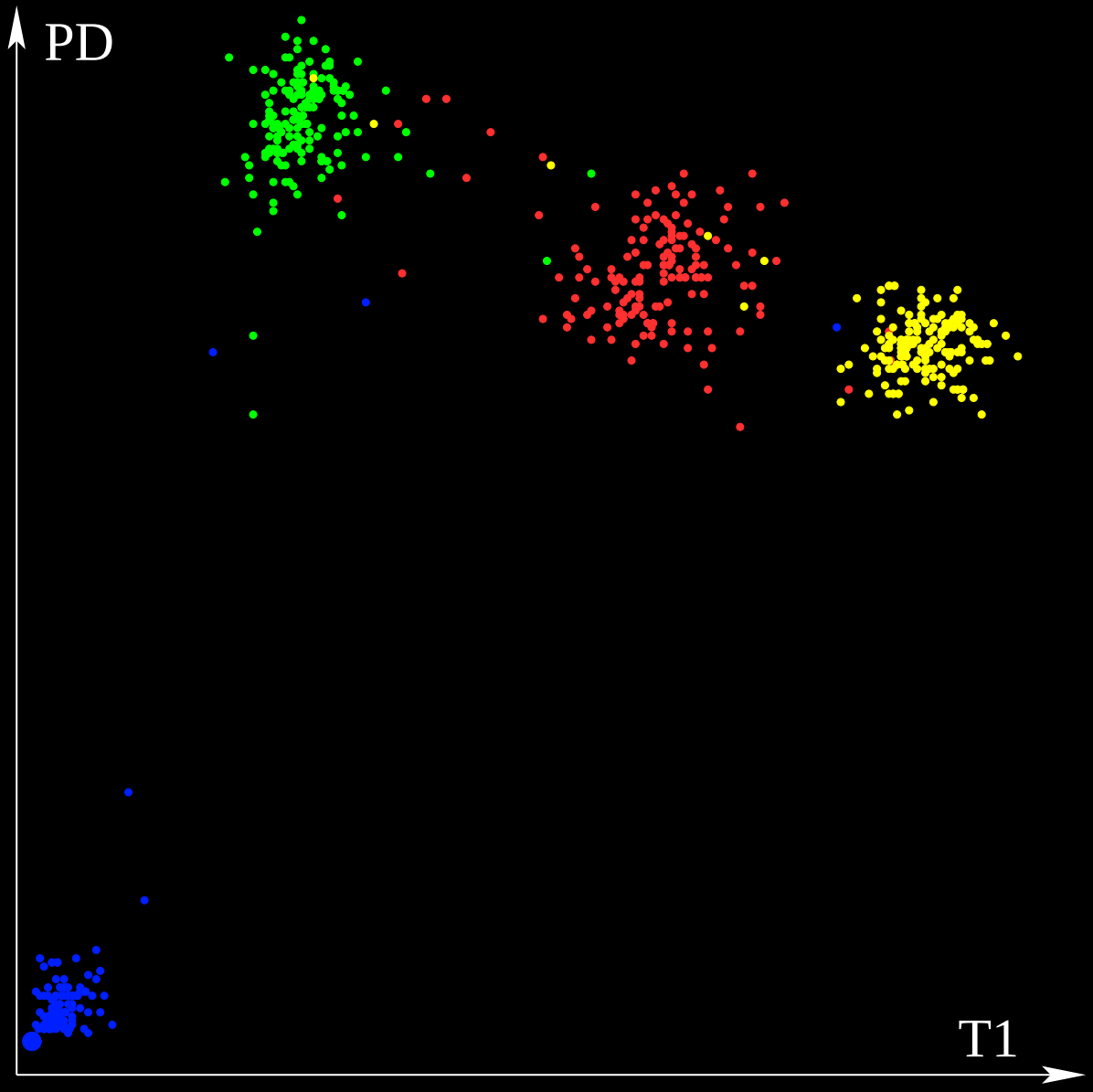
Pruning : removal of samples with incorrect class labels.

Our novel method:

- accommodates subject anatomies significantly different than model
- non-parametric: no assumptions about feature space (intensity) distributions
- allows for a lower TPM $\mathcal{T}$
  $\Rightarrow$ better estimation of intensity distributions
  $\Rightarrow$ accuracy, robustness

Pruning method:

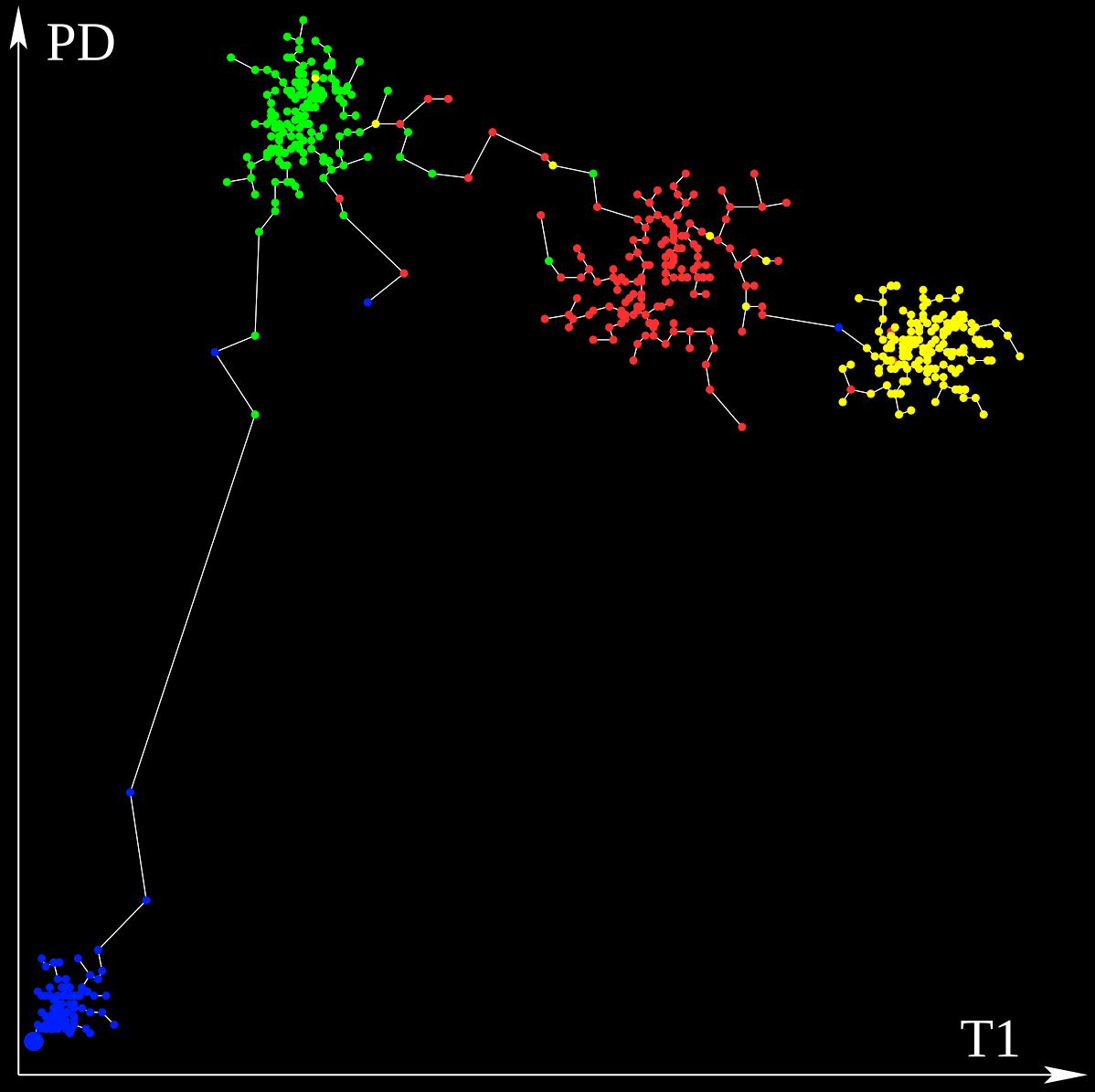
"raw" samples in feature space:



Pruning method:

[Step 1]

Minimum Spanning Tree

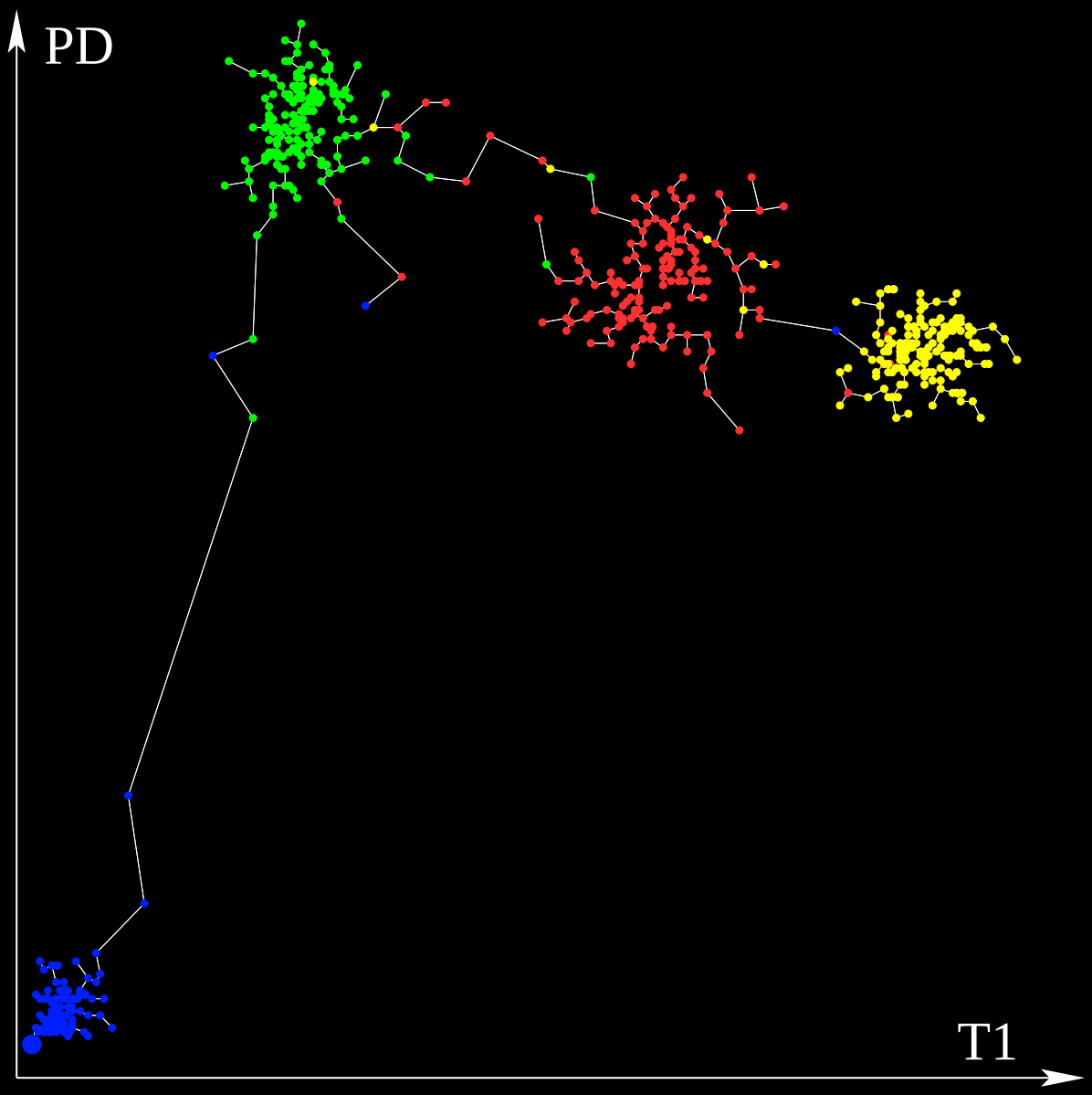


Pruning method:

[Step 2]

edge (i, j) is removed if $length(i, j) > T \times A(i)$ or if $length(i, j) > T \times A(j)$,

where $A(i) =$ average length of all other edges incident on node i

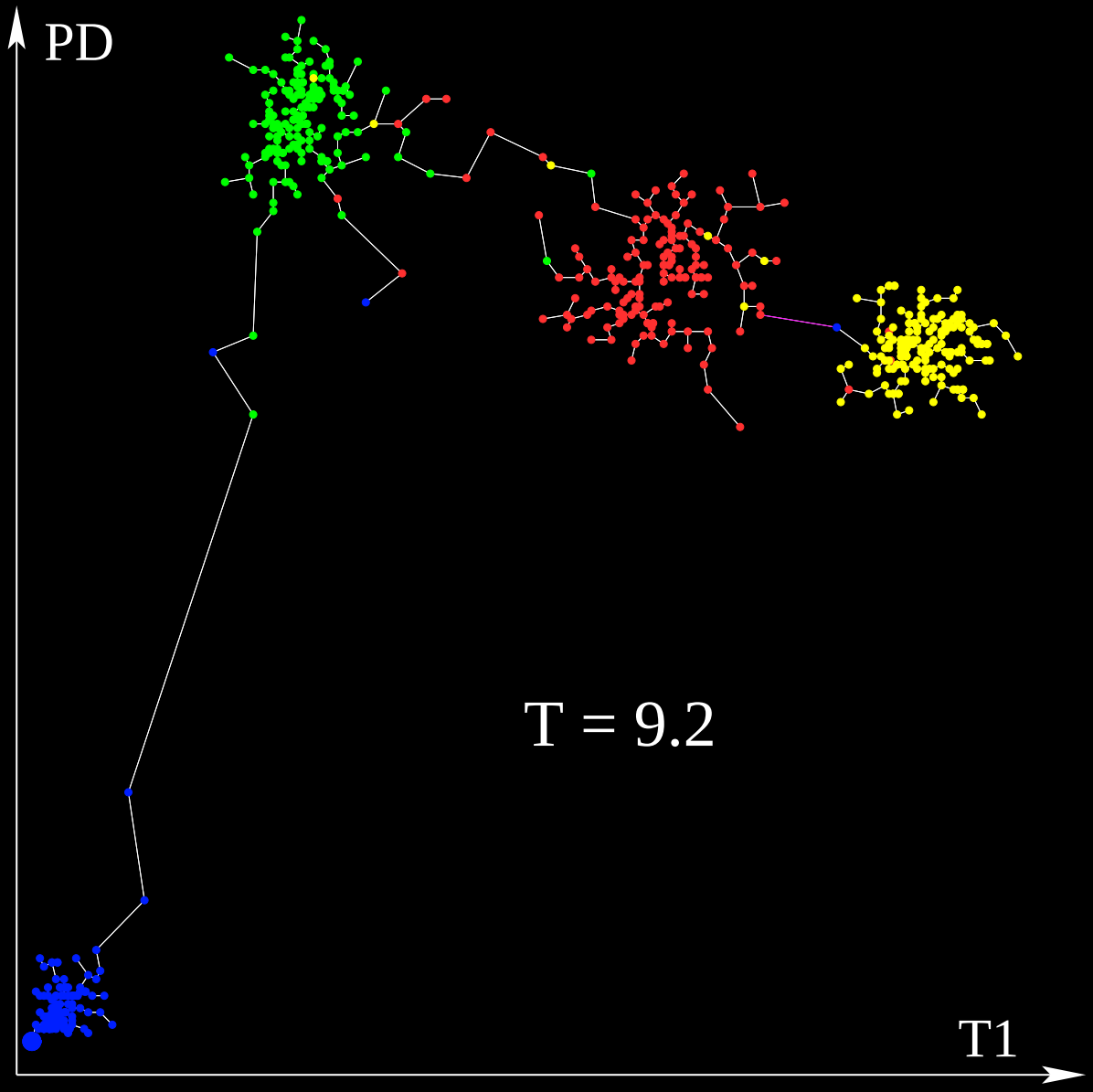


Pruning method:

[Step 2]

edge (i, j) is removed if $length(i, j) > T \times A(i)$ or if $length(i, j) > T \times A(j)$,

where $A(i) =$ average length of all other edges incident on node i

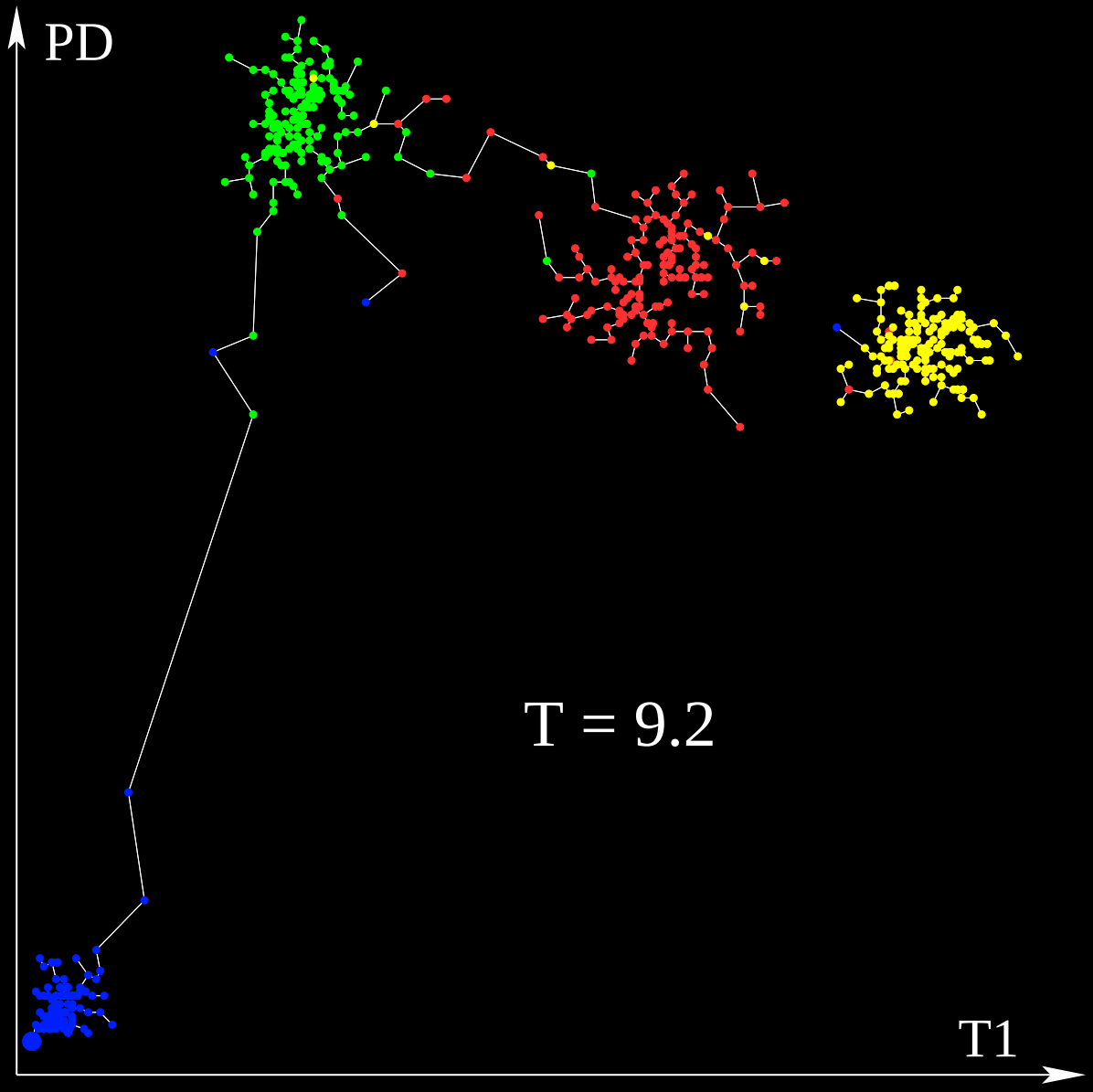


Pruning method:

[Step 2]

edge (i, j) is removed if $length(i, j) > T \times A(i)$ or if $length(i, j) > T \times A(j)$,

where $A(i) =$ average length of all other edges incident on node i

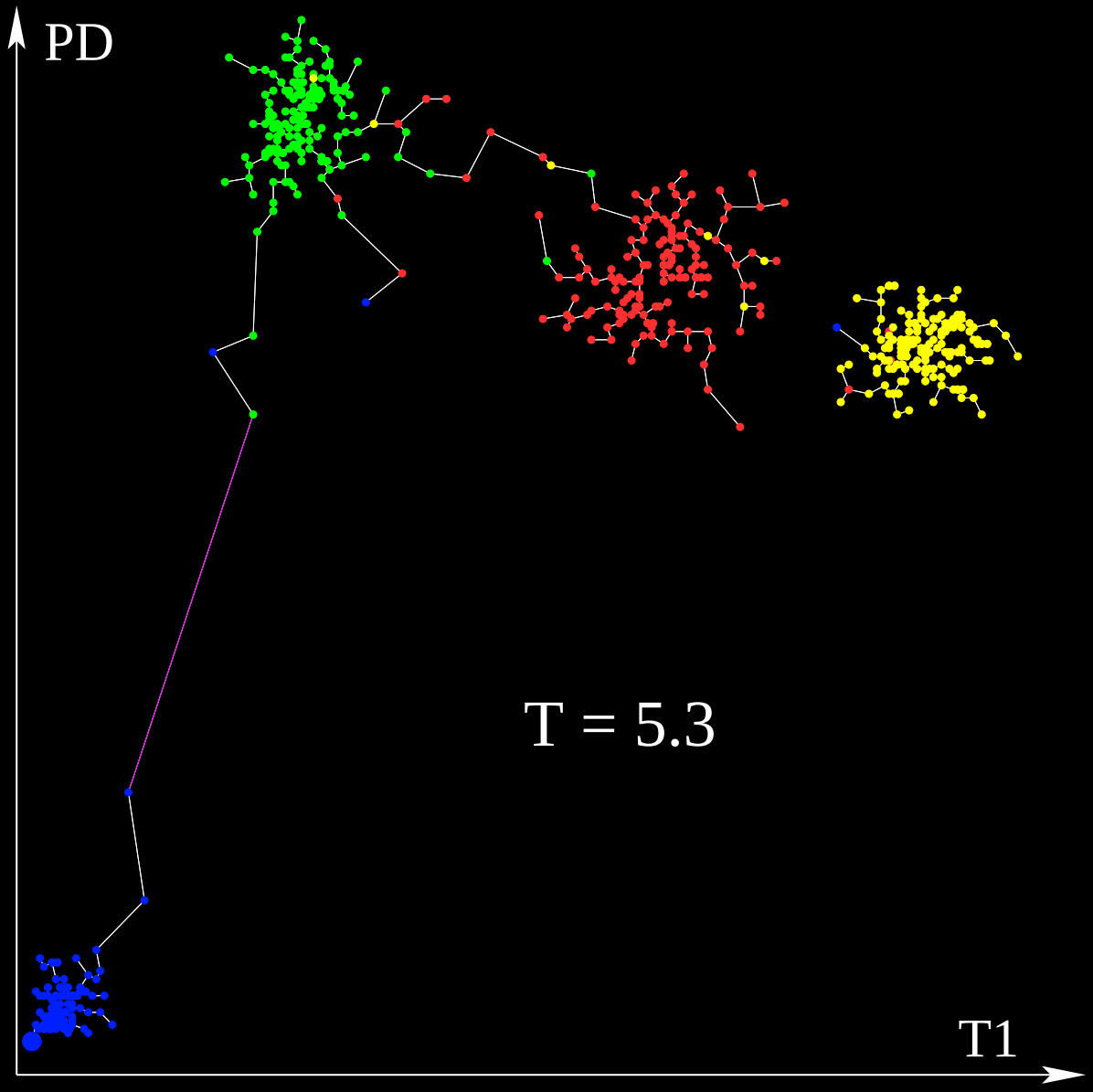


Pruning method:

[Step 2]

edge (i, j) is removed if $length(i, j) > T \times A(i)$ or if $length(i, j) > T \times A(j)$,

where $A(i) =$ average length of all other edges incident on node i

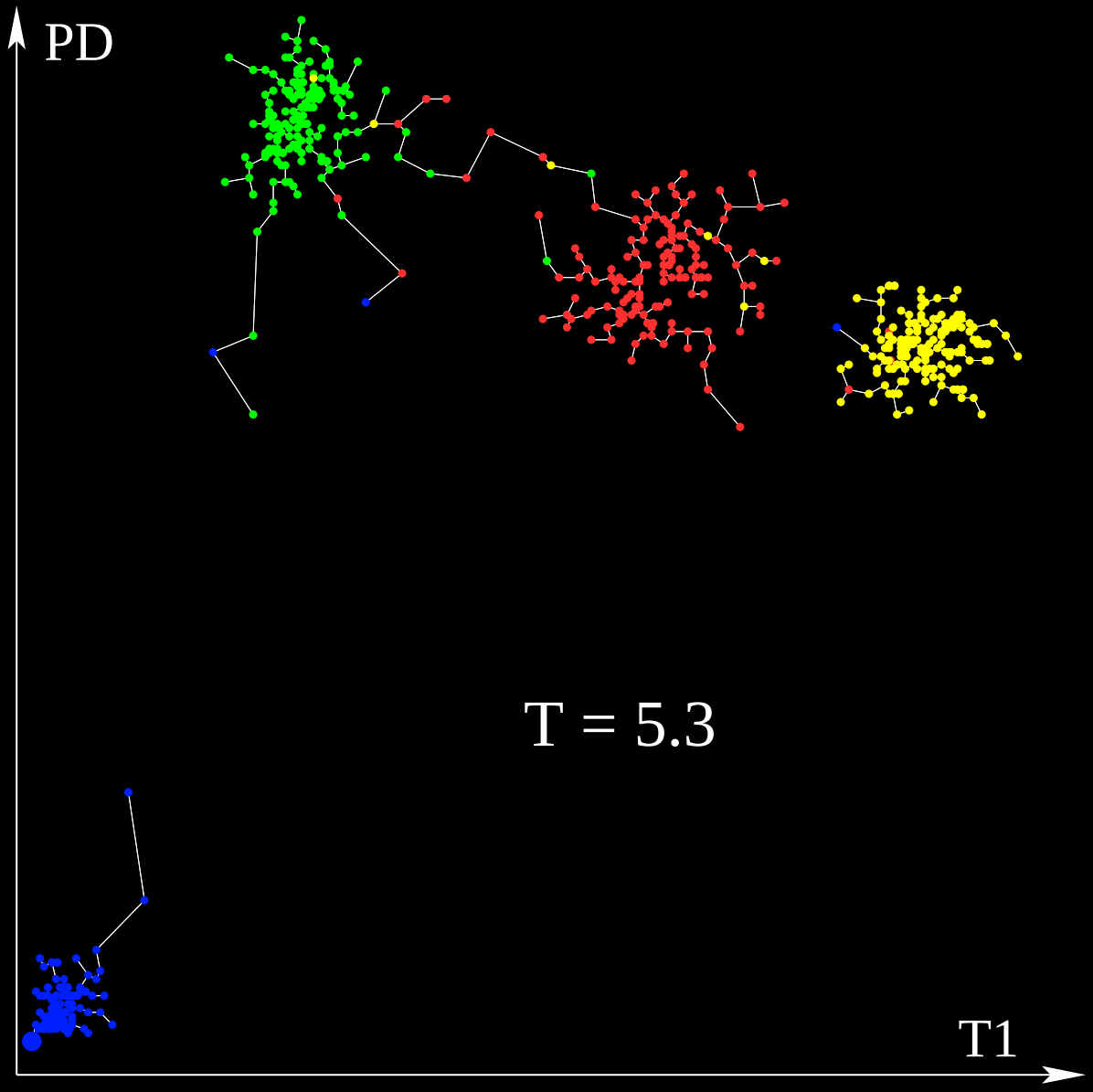


Pruning method:

[Step 2]

edge (i, j) is removed if $length(i, j) > T \times A(i)$ or if $length(i, j) > T \times A(j)$,

where $A(i) =$ average length of all other edges incident on node i



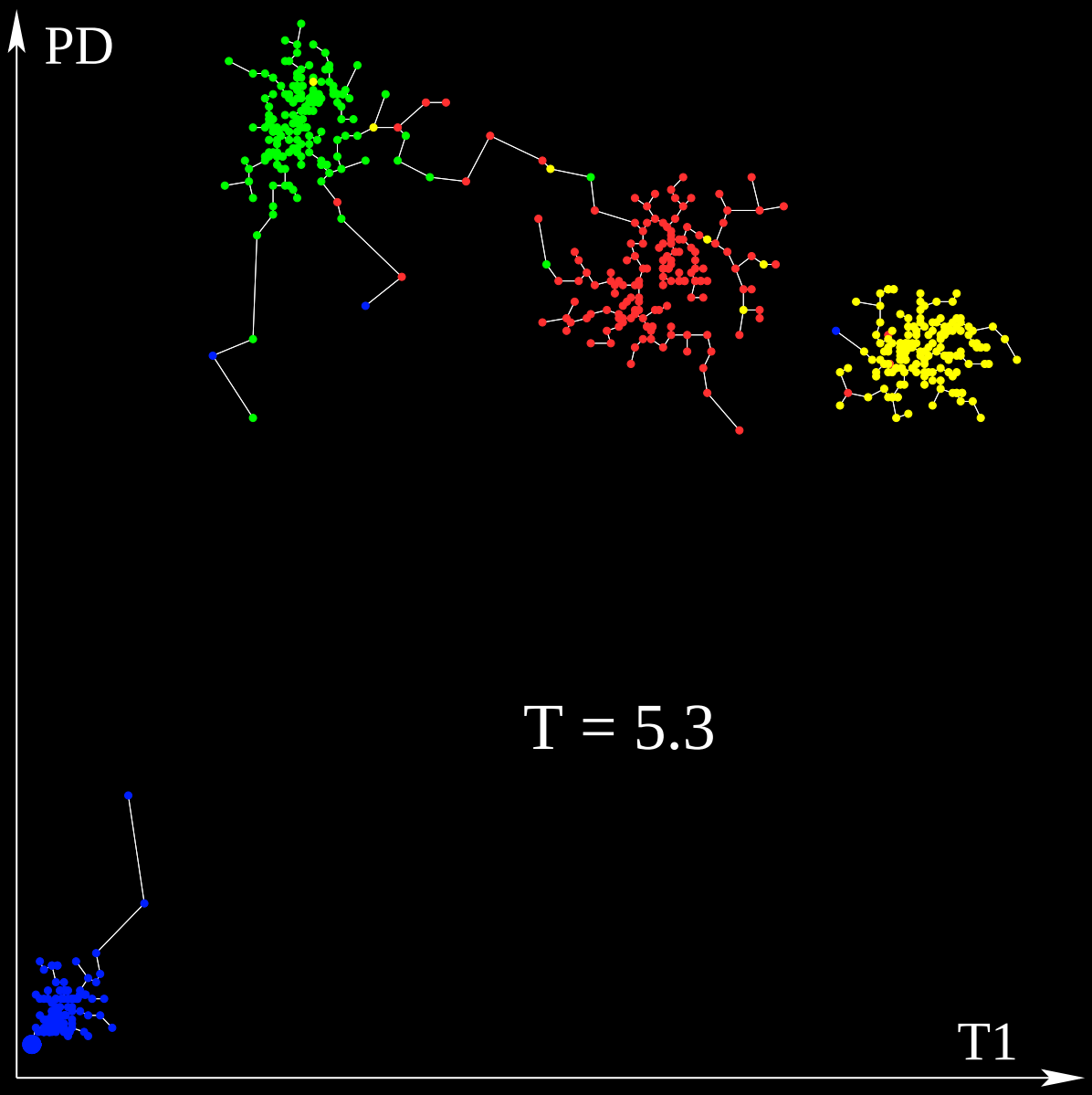
Pruning method:

[Step 2]

cluster = a connected component of graph.

CSF cluster = cluster with most CSF samples. . . .

Stop when BG, CSF, GM, WM clusters are distinct.



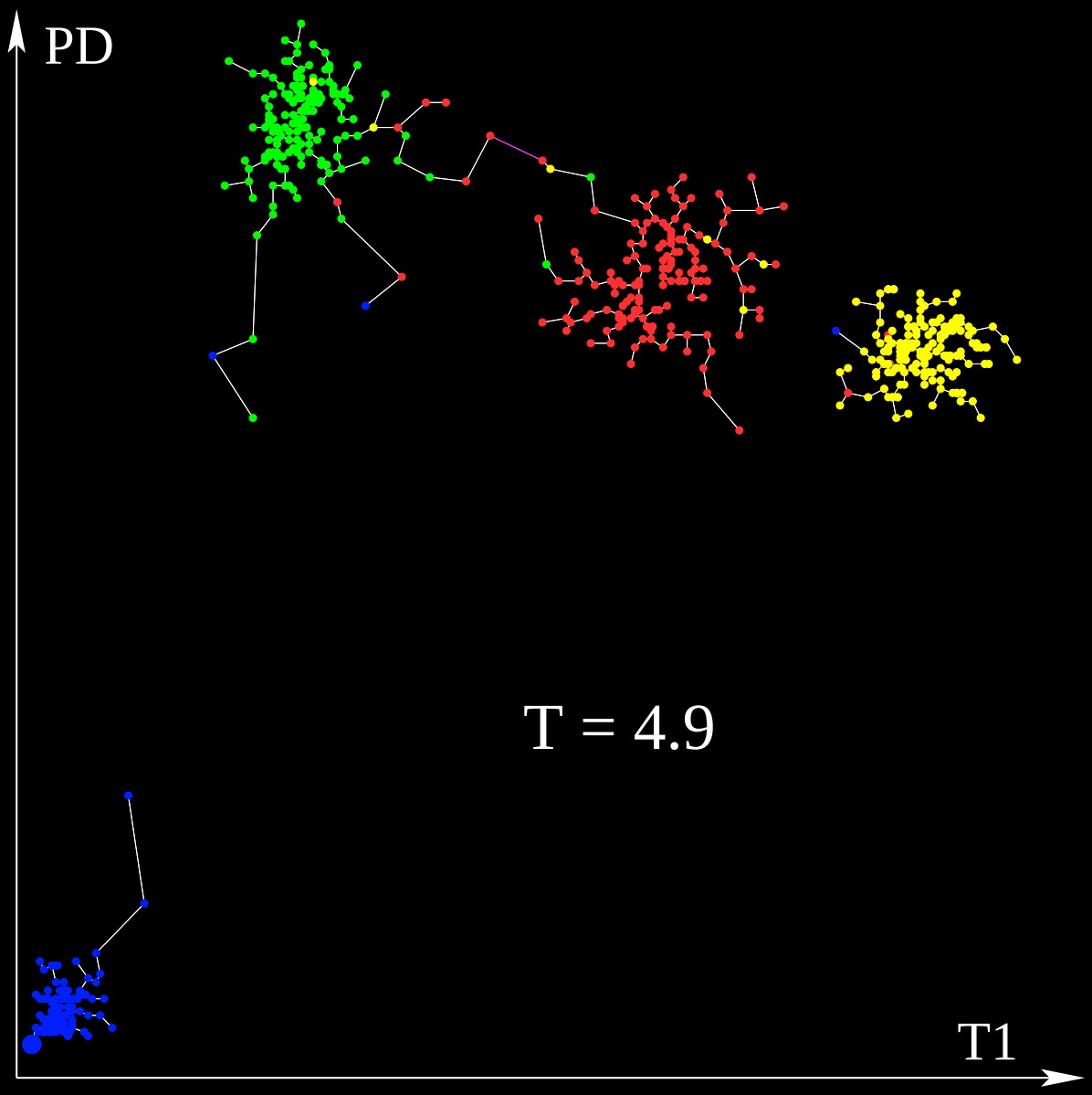
Pruning method:

[Step 2]

cluster = a connected component of graph.

CSF cluster = cluster with most CSF samples. . . .

Stop when BG, CSF, GM, WM clusters are distinct.



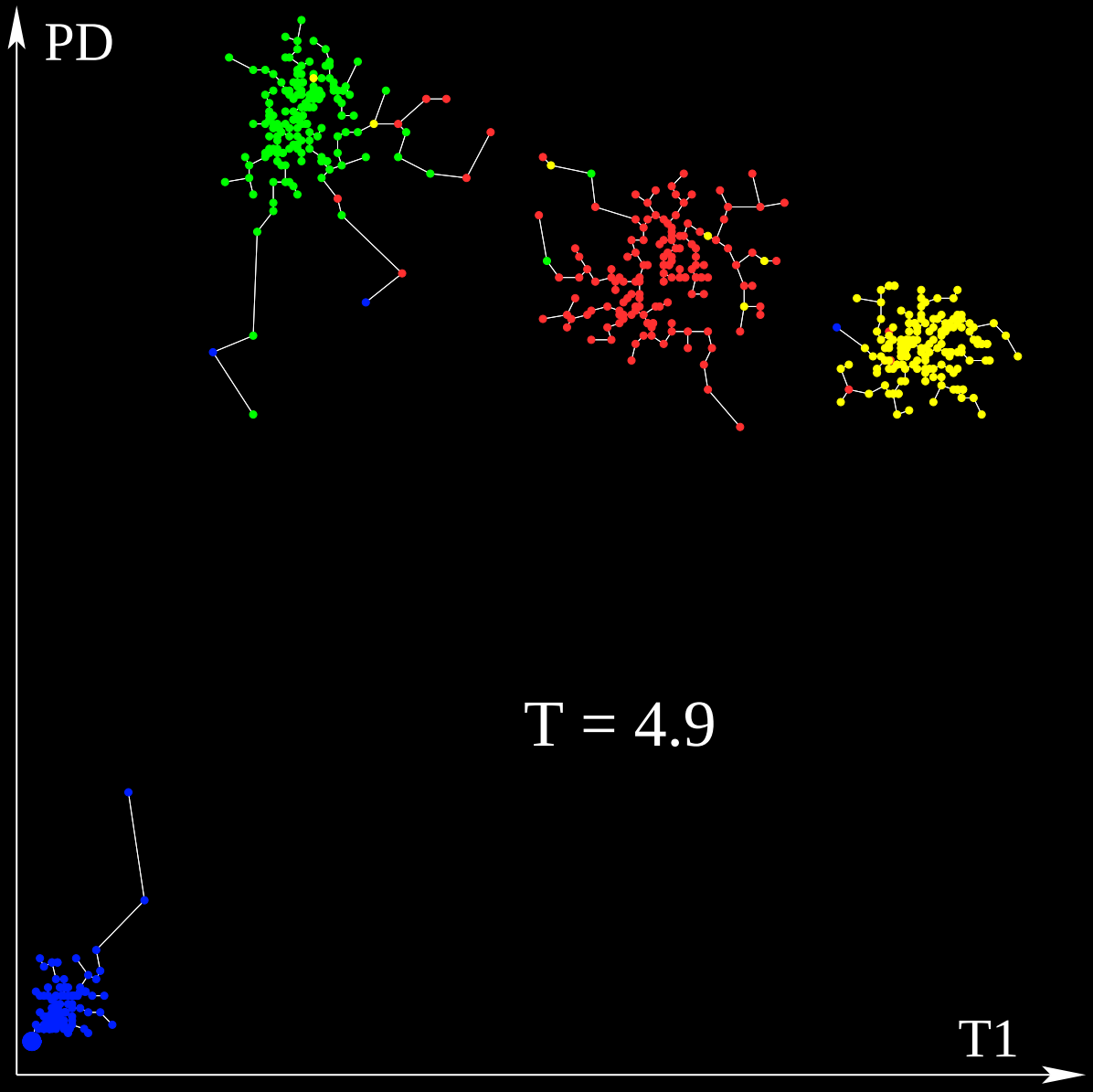
Pruning method:

[Step 2]

cluster = a connected component of graph.

CSF cluster = cluster with most CSF samples. . . .

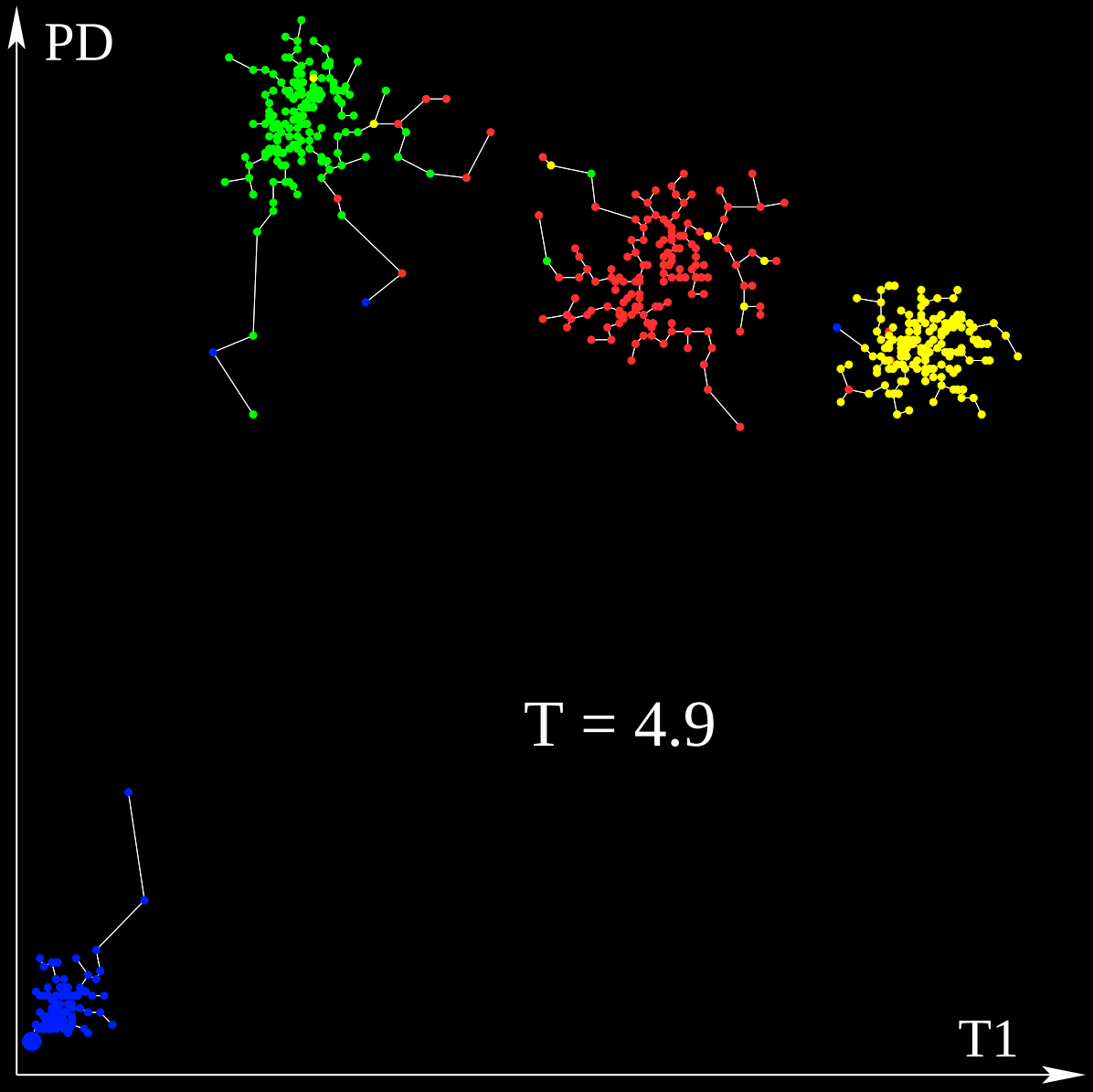
Stop when BG, CSF, GM, WM clusters are distinct.



Pruning method:

[Step 3]

$\Rightarrow$ discard samples that are not found in correct cluster.

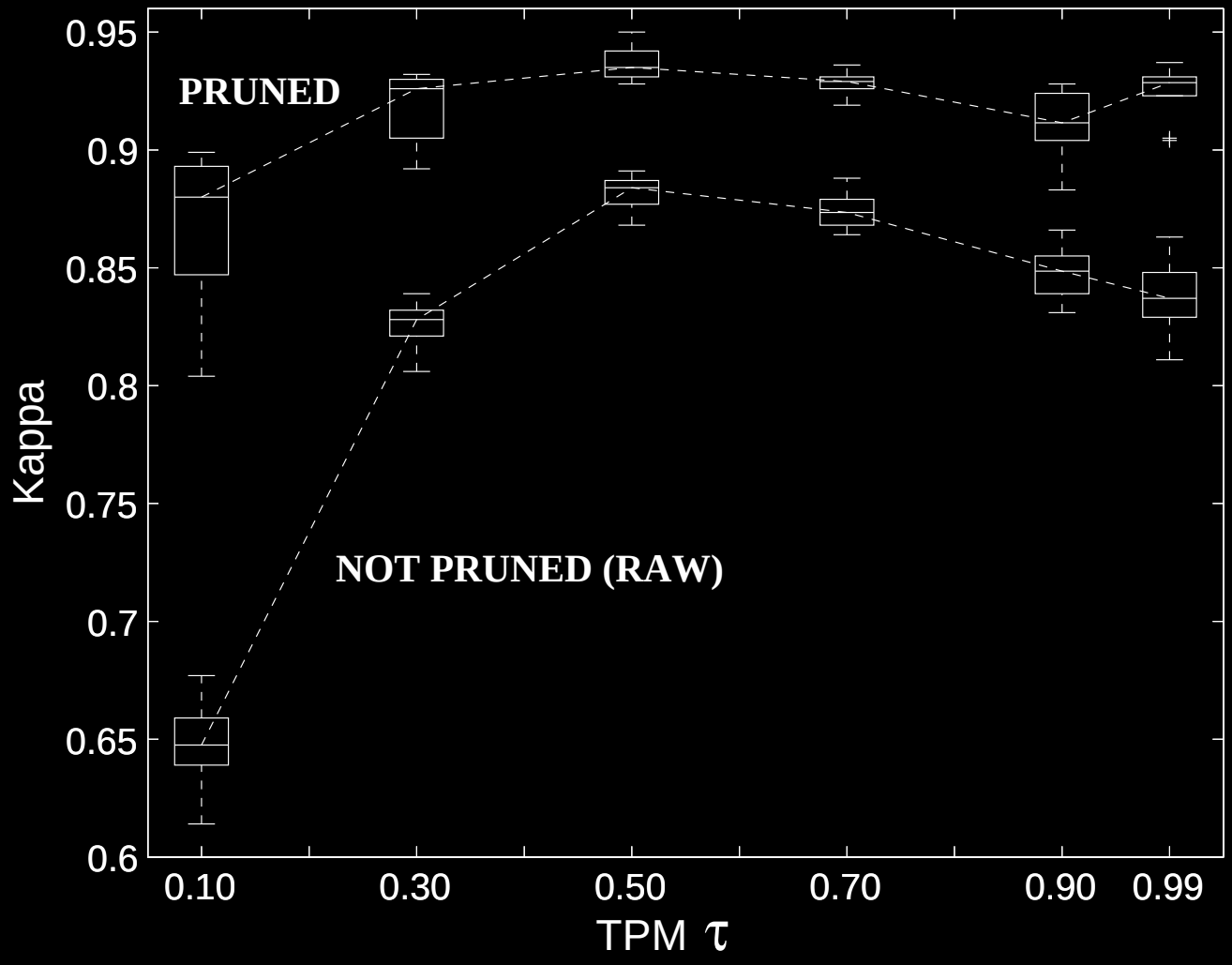


Validation: simulated MRI

- T1-T2-PD multi-spectral simulated MRI,
  10 different “elderly brain” phantoms
- young-normal model (TPM), N=53

- quantitative evaluation:
  Kappa = chance-corrected similarity measure between two image labelings (classifications).
  classified image $\Leftrightarrow$ “gold standard” (phantom)

Validation: simulated MRI



(elderly brain simulated MRI, young-normal model)

Validation: real MRI

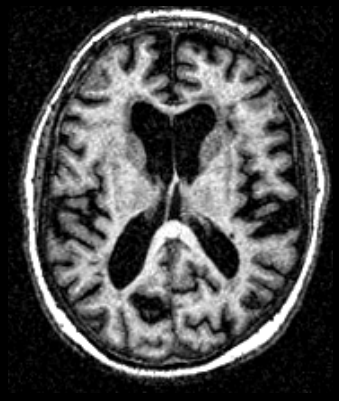
1. young & normal individual (T1+T2+PD, and also T1 only), against full-brain manual segmentation.

2. 31 Ischemia patients (T1+T2+PD).

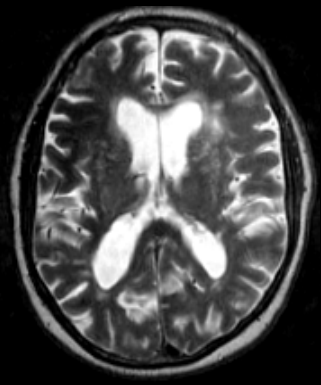
3. 11 Alzheimer's Disease (A.D.) elderly patients (T1+T2+PD).

Results: Ischemia

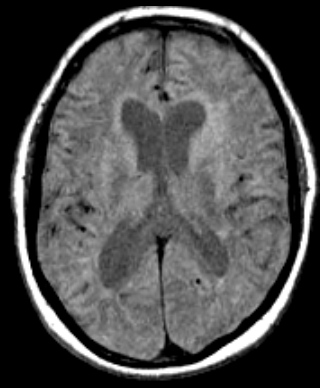
T1 MRI



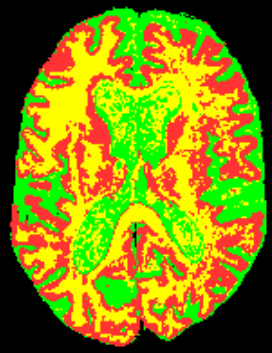
T2 MRI



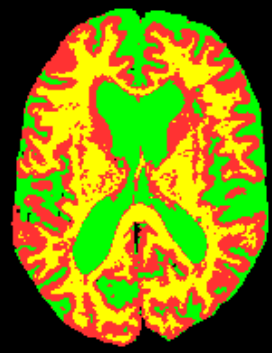
PD MRI



raw (not pruned):

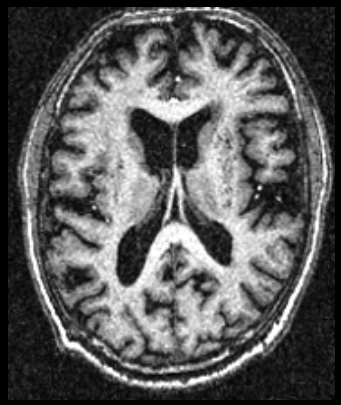


pruned:

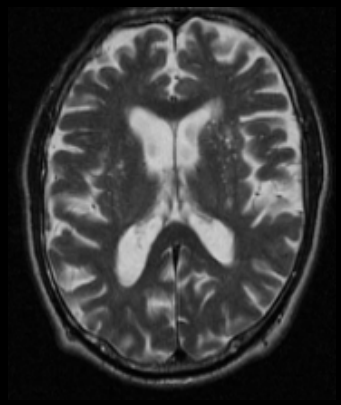


Results: Ischemia

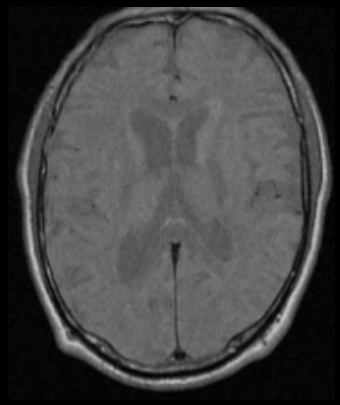
T1 MRI



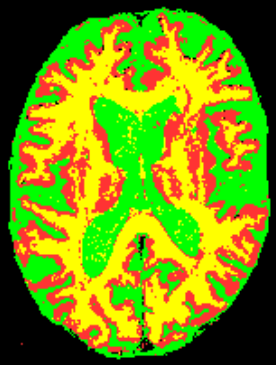
T2 MRI



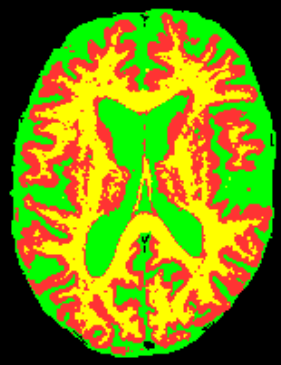
PD MRI



raw (not pruned):

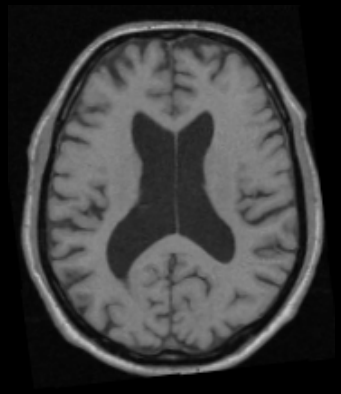


pruned:

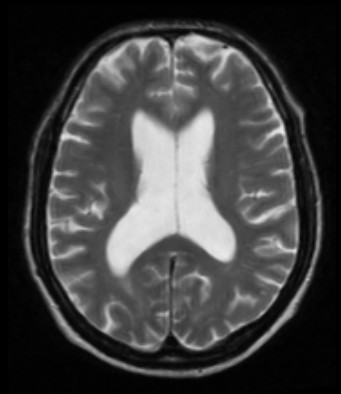


Results: Ischemia

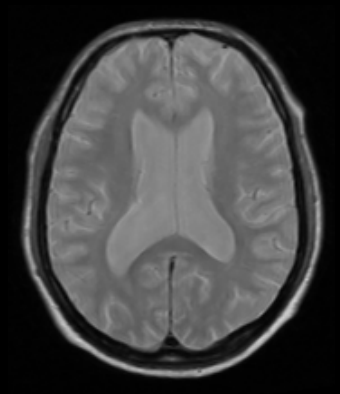
T1 MRI



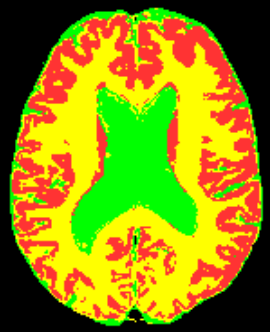
T2 MRI



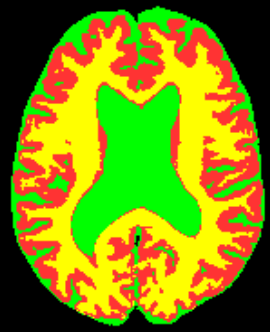
PD MRI



raw (not pruned):

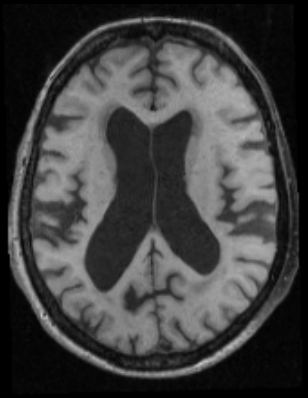


pruned:

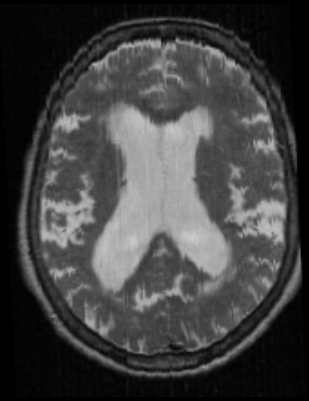


Results: A.D. elderly

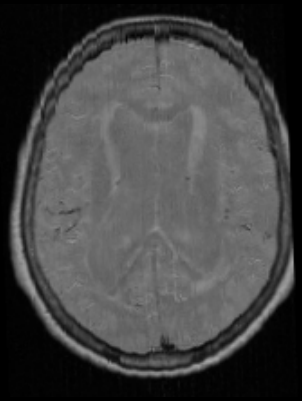
T1 MRI



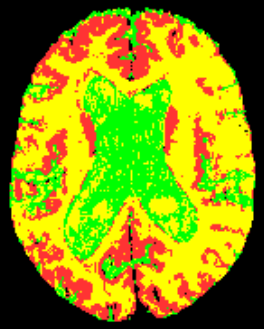
T2 MRI



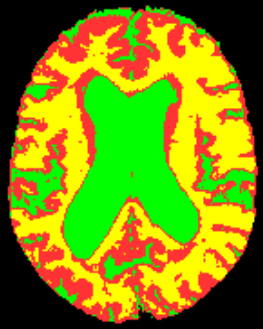
PD MRI



raw (not pruned):

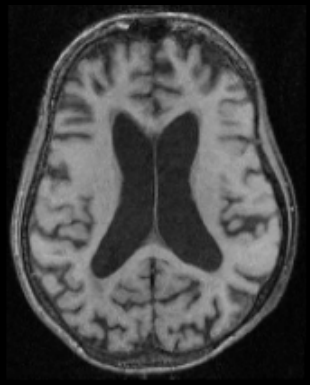


pruned:

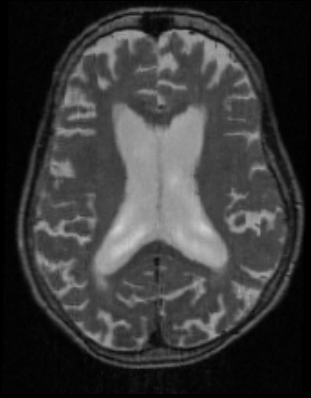


Results: A.D. elderly

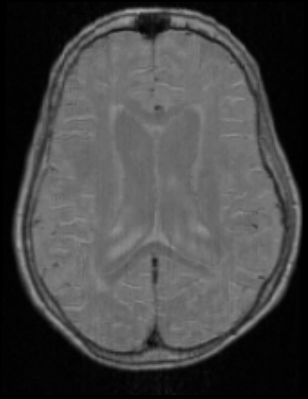
T1 MRI



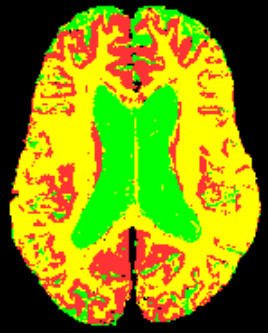
T2 MRI



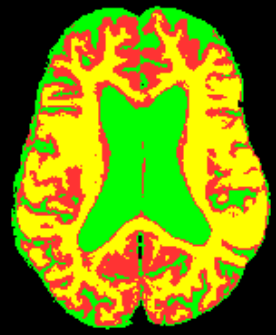
PD MRI



raw (not pruned):



pruned:



Future work:

Current limitations:

- inherent to intensity-only, discrete classification.
- due to overlap of tissue intensity distributions (brain biology, MRI partial volume).

Future work:

Current limitations:

- inherent to intensity-only, discrete classification.
- due to overlap of tissue intensity distributions (brain biology, MRI partial volume).

$\Rightarrow$ also use voxel neighbourhood information (e.g. image gradient), . . .

Summary:

- fully automatic brain tissue classification procedure.
- robust against anatomical variability.
- non-parametric: no assumptions about tissue intensity distributions ($\Rightarrow$ robust against imaging artifacts).
- validated qualitatively and quantitatively on simulated and on real MRI data.

Acknowledgements:

- John Sled, Steve Robbins, Peter Neelin, Godfried Toussaint, Noor Kabani, Louis Collins, Jean-François Mangin, Jason Lerch, Jennifer Campbell, Ives Levesque, Najma Khalili.
- the anonymous MICCAI-2002 reviewers.
- Alma Mater Student Travel Fund, Faculty of Graduate Studies and Research, McGill University, Montreal.

Automatic Generation of Training Data for Brain Tissue Classification from MRI

Chris A. COCOSCO, Alex P. ZIJDENBOS, and Alan C. EVANS

chris.cocosco@mail.mcgill.ca

http://www.bic.mni.mcgill.ca/users/crisco/

McConnell Brain Imaging Centre

Montreal Neurological Institute

McGill University