KU LEUVEN

imec

erc European Research Council Established by the European Commission

Tensor-based blind source separation in Magnetic Resonance Spectroscopic Imaging

Bharath Halandur Nagaraja

bhalandu@esat.kuleuven.be

Prof. Sabine Van Huffel

sabine.vanhuffel@kuleuven.be

10-9-2018

Contents Overview

| Cancer Diagnostics

UZ Leuven departments

Imaging & Pathology

Radiology

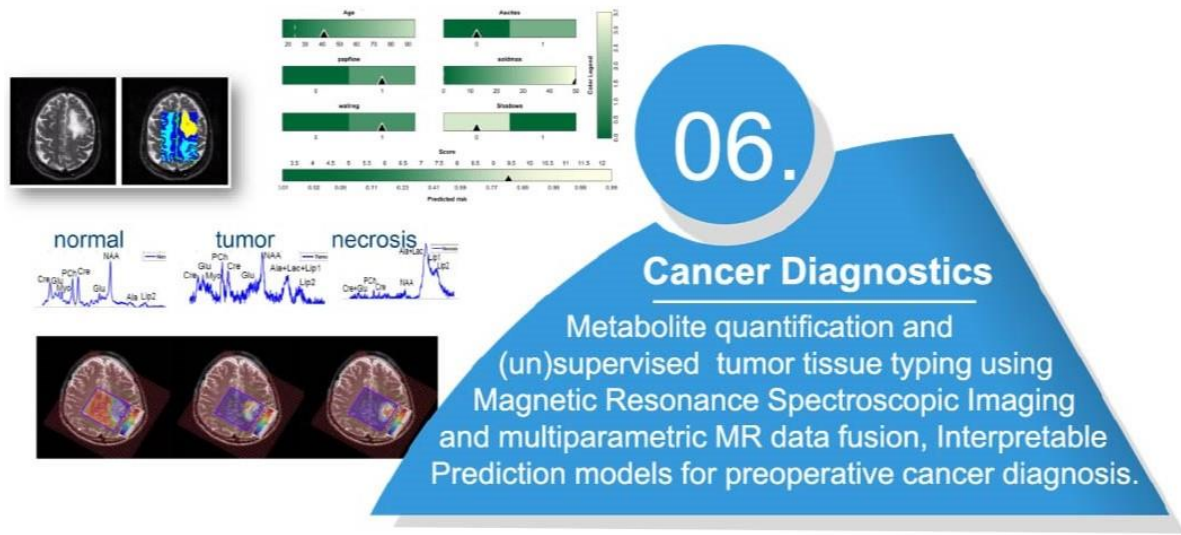
Neurosurgery

Development and Regeneration

UZ Gent department

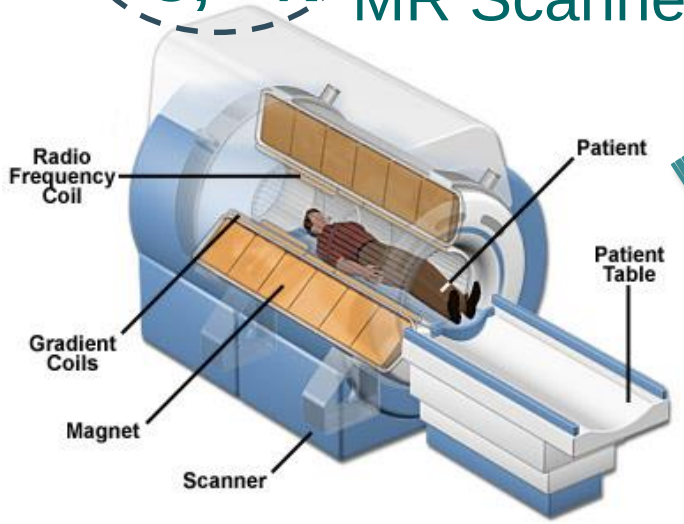
Radiology

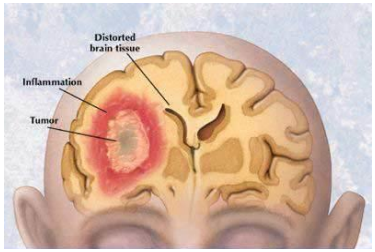
EXAMPLE



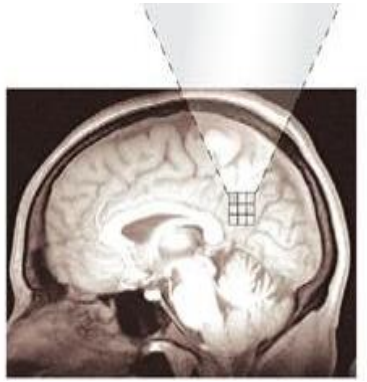
INTRODUCTION: MRS Acquisition and Purpose

^{1}H, ^{31}P, ^{13}C, ^{14}N MR Scanner



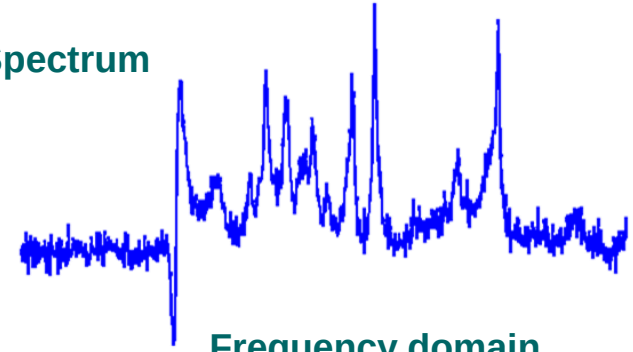


MRS signals



Diagnosis: Cancer, metabolic diseases, Alzheimer's disease

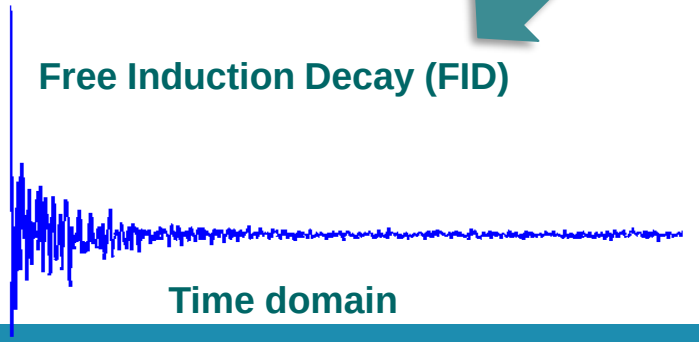
Spectrum



Frequency domain

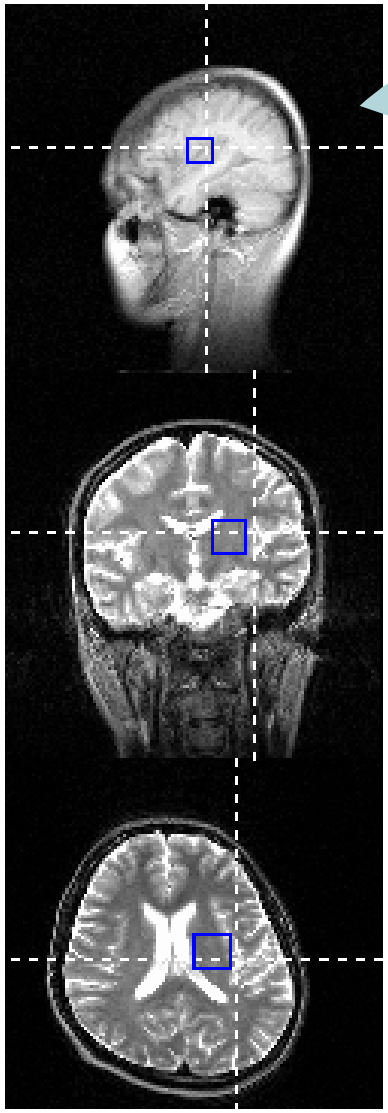
FFT

Free Induction Decay (FID)

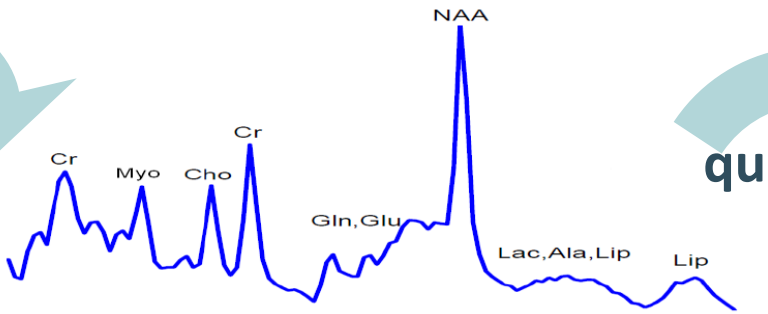


Time domain

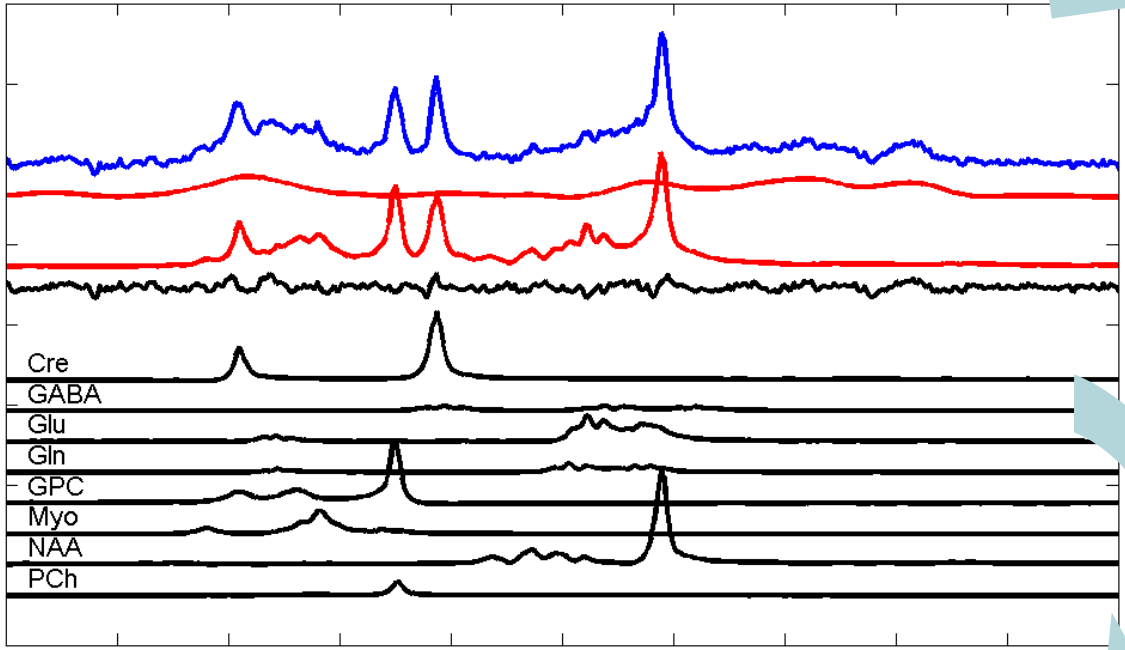
Metabolite quantification for MR Spectroscopy (MRS)



Single-voxel MRS

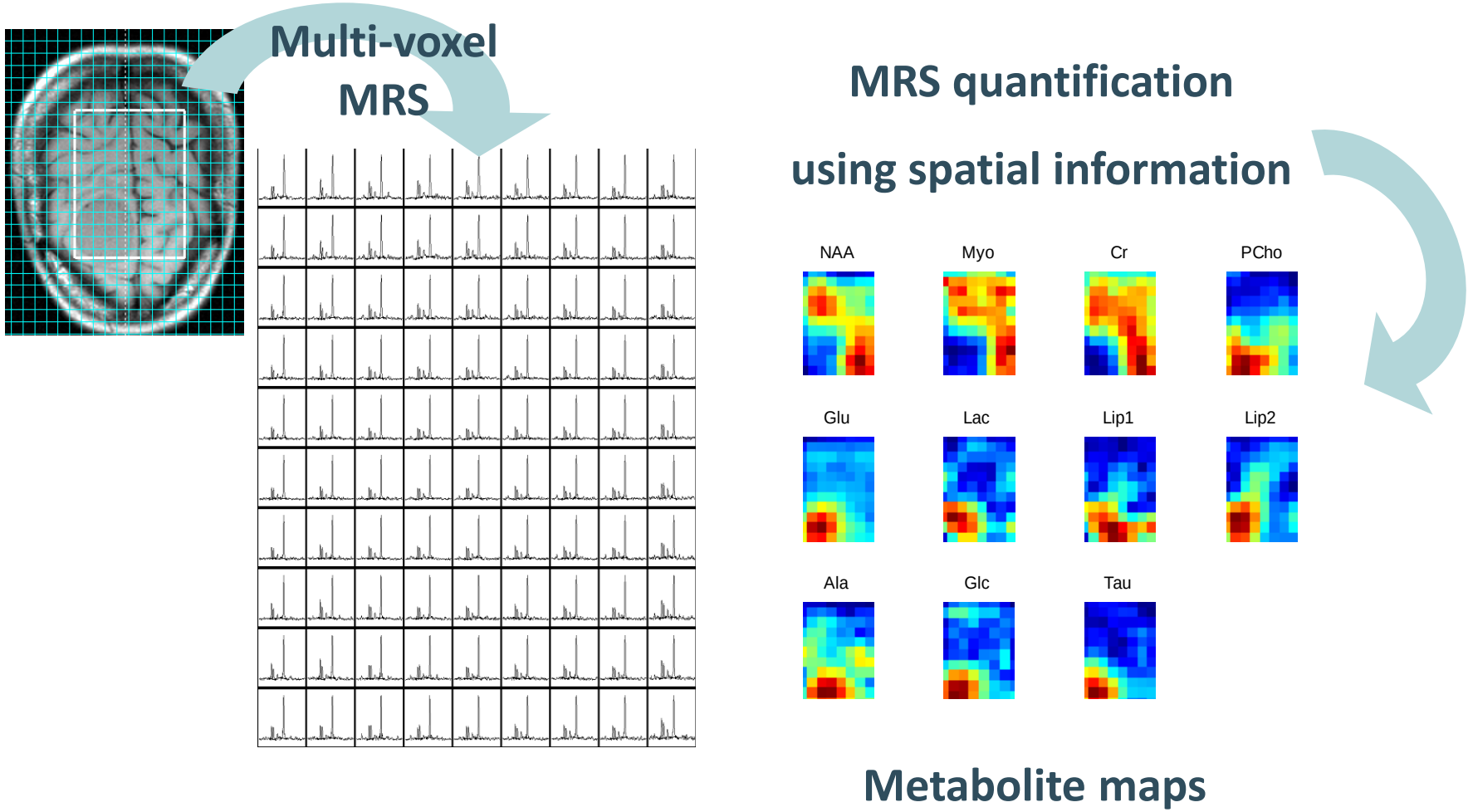


MRS quantification



Metabolite concentrations

Metabolite quantification for MRS Imaging (MRSI)

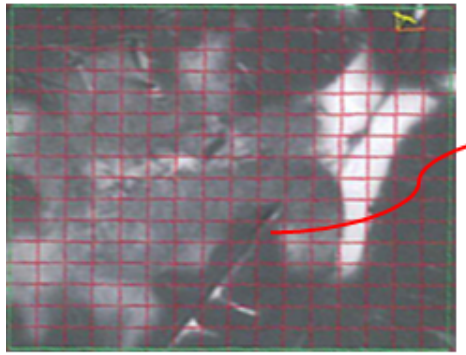


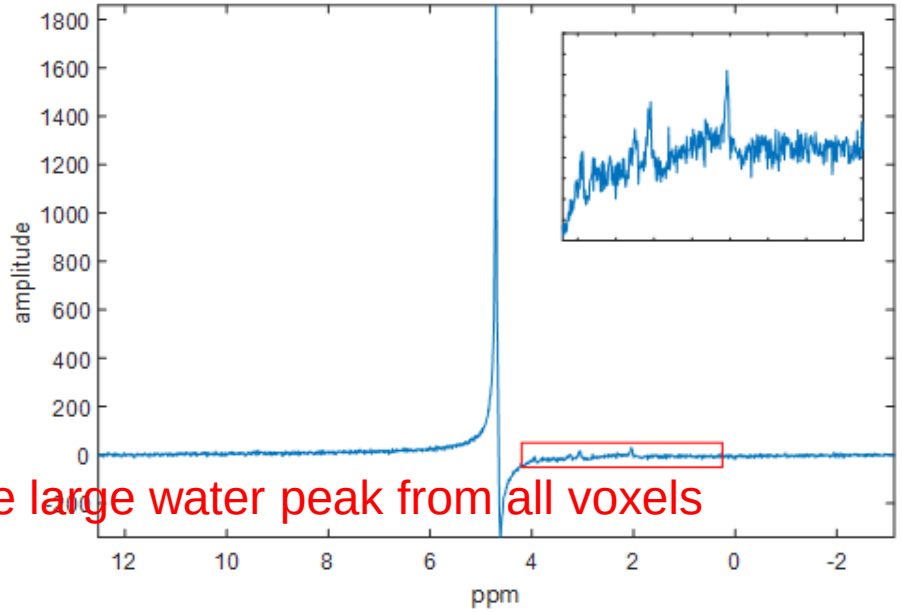
Metabolite concentrations = biomarkers of disease

Contents Overview

1. Introduction
   - Magnetic Resonance Spectroscopic imaging
2. **Residual Water Suppression**
   - Hankel-tensor based water suppression
   - Löwner-tensor based water suppression
3. Tumor tissue type differentiation
   - From MRSI
   - From MP-MRI
4. MRSI tumor voxel classification
5. Tumor segmentation from MP-MRI

Introduction- MRSI and water suppression





Aim: Suppress the large water peak from all voxels

Frequency domain Model

$$S(f) = \sum_{r=1}^{R} \frac{a_r e^{j\phi_r}/2\pi}{d_r + j2\pi(f - f_r)} + \eta(f)$$

$$S = WH^T$$

Spectra from Voxels

Source abundancies in the grid

Spectra of sources

Time domain Model

$$S(t) = \sum_{r=1}^{R} a_r e^{j\phi_r} e^{(-d_r + j2\pi f_r)t} + \eta(t)$$

HSVD based water suppression
Hankel Tensor + MLSVD

HSVD based water suppression

- Construct the Hankel matrix *H* from the time domain signal S.
- The Hankel matrix *H* has Vandermonde decomposition.

$$H = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ z_1^{L-1} & \cdots & z_R^{L-1} \end{bmatrix} \begin{bmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_R \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ z_1^{M-1} & \cdots & z_R^{M-1} \end{bmatrix}^T$$

$a_r e^{j\phi_r}$ $\qquad$ $e^{(-d_r + j2\pi f_r)nt}$

- Estimate the parameters {*Z and C*} using SVD.

$$H = USV^H \quad \text{truncate to R, } H_R = U_R S_R V_R^H$$

- Use the shift–invariance property to estimate the poles.

Solve $\quad U_R^{\uparrow} = U_R^{\downarrow} T \quad$ with $T = Q^{-1} Z Q$ and Z = diag{z_1, …, z_R}

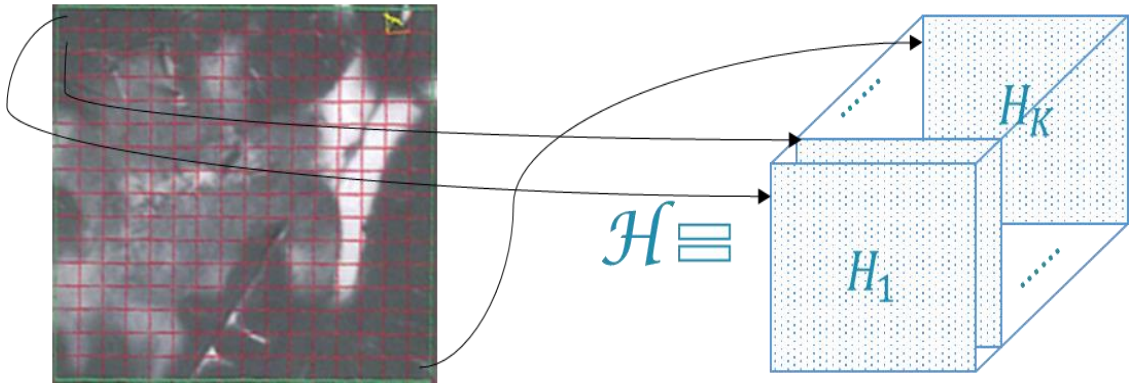
- Poles are estimated as $z = eig(Q^{-1} Z Q)$.

HSVD based water suppression

- Calculate the amplitudes a_k from estimated poles and measured signal using least squares.
- Estimate the frequency and damping factor from each pole.
- Estimate the water component using only the poles and amplitudes whose frequencies are outside the region of interest (0.25-4.2 ppm).
- Finally, suppress the water component by subtracting its estimate from the measured signal.
- The disadvantage of this method is that we have to repeat the procedure for all voxels in the MRSI grid.
- Neighboring voxels have similar water sources with different complex amplitudes→ Exploit using tensor approaches.

Hankel-tensor based water suppression- MLSVD

- Each individual component can be well approximated by complex-damped exponentials.
- For each voxel in the MRSI, construct a Hankel matrix from the free induction decay (FID) signal but now stack these to form a tensor.



- Parameters of the complex-damped exponentials can be obtained using MLSVD

$$\mathcal{H} = \mathcal{A} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}$$

Bharath, H. N., Debals, O., Sima, D. M., Himmelreich, U., De Lathauwer, L., and Van Huffel, S. Tensor based method for residual water suppression in 1H magnetic resonance spectroscopic imaging. in IEEE Transactions on Biomedical Engineering, 2018.

Hankel based water suppression- MLSVD

- The frequency and damping of individual sources can be obtained using shift-invariance property and eigen-value decomposition of mode-1 factor matrix $U^{(1)}$.
- Water sources are estimated by reconstructing those exponentials outside the region of interest (0.25-4.2 ppm).
- The total water component (including amplitudes) is estimated by fitting those water exponentials to the measured signal using least-squares.
- Finally, the water component is suppressed by subtracting this estimated water signal from the measured MRSI signal.

Löwner based water suppression

$$S(f) = \sum_{r=1}^{R} \frac{a_r e^{j\phi_r}/2\pi}{d_r + j2\pi(f - f_r)} + \eta(f)$$

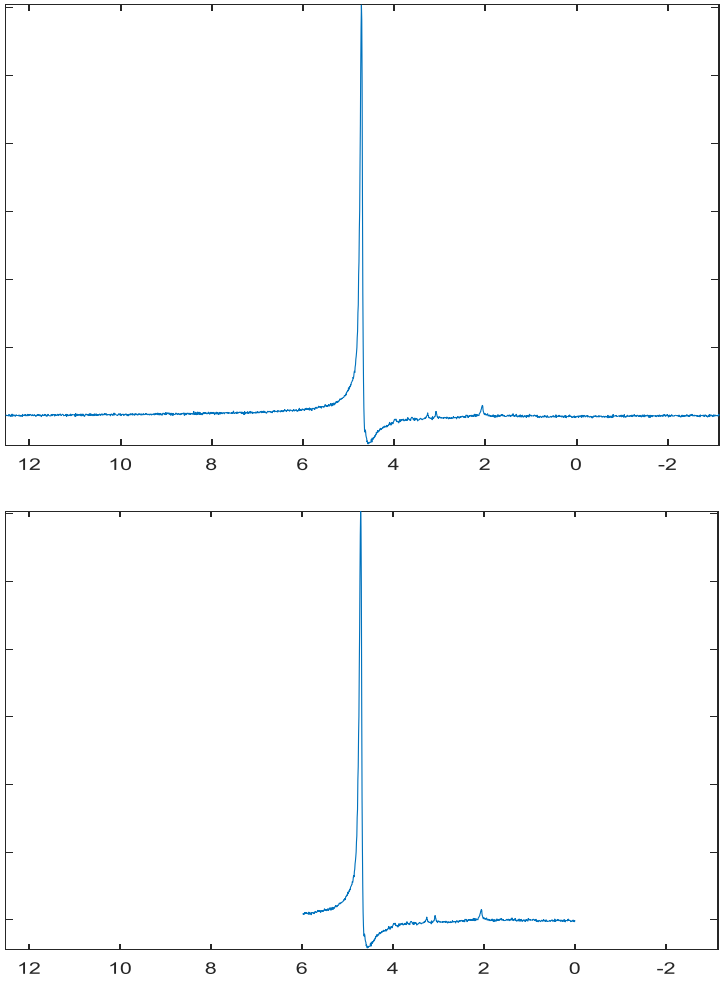
- For each voxel in the MRSI signal construct a Löwner matrix from truncated spectrum S_{wr}.

$$L = \begin{bmatrix} \frac{S_{wr}(x_1) - S_{wr}(y_1)}{x_1 - y_1} & \cdots & \frac{S_{wr}(x_1) - S_{wr}(y_J)}{x_1 - y_J} \\ \vdots & \ddots & \vdots \\ \frac{S_{wr}(x_I) - S_{wr}(y_1)}{x_I - y_1} & \cdots & \frac{S_{wr}(x_I) - S_{wr}(y_J)}{x_I - y_J} \end{bmatrix}$$

Using interleaved partitioning in two point sets X and Y:

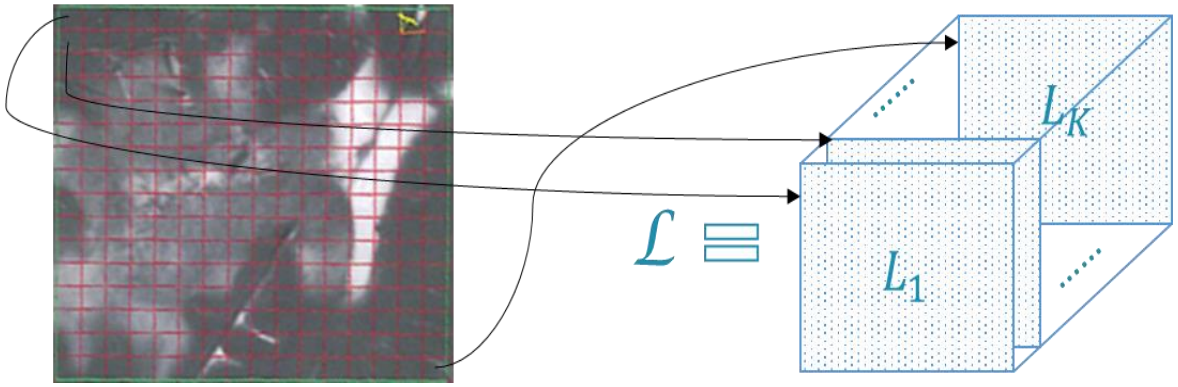
X = $\{x_1, x_2 \ldots\}$→{1,3,…..} and Y = $\{y_1, y_2 \ldots\}$ → {2,4,…..}

- If a Löwner matrix L is constructed from a rational function of degree *R*, the matrix L has rank *R*.



Löwner based water suppression- CPD

- For each voxel in the MRSI signal construct a Löwner matrix from the spectra and stack these to form a tensor.



- Each individual component can be well approximated by a degree-1 rational function, hence BSS reduces to CPD.

$$\mathcal{L}_s \approx \sum_{r=1}^{R} a_r \otimes b_r \otimes h_r \ , \qquad \mathrm{L}_R = a_r b_r^T{}_r$$

Bharath, H. N., Debals, O., Sima, D. M., Himmelreich, U., De Lathauwer, L., and Van Huffel, S. Tensor based method for residual water suppression in 1H magnetic resonance spectroscopic imaging. in IEEE Transactions on Biomedical Engineering, 2018.

Löwner based water suppression

$$L_r = A(:,r)B(:,r)^T$$

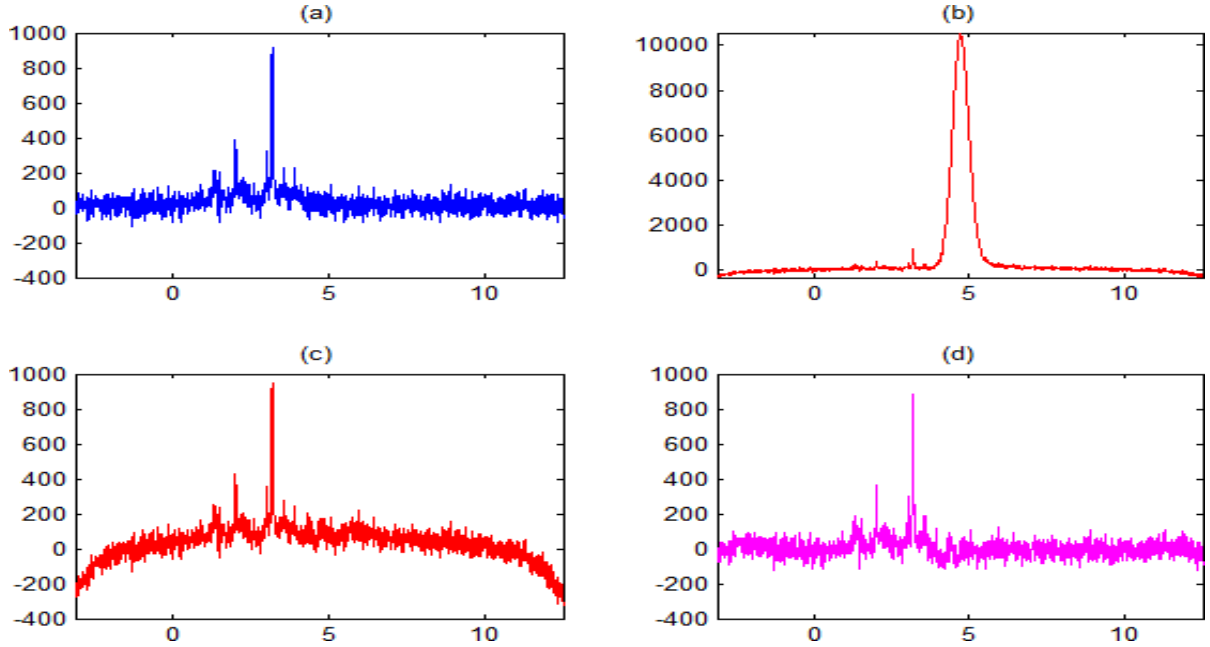
$$A(:,r) = \begin{bmatrix} \frac{a_{r1}}{(ix_1 + b_r)} \\ \cdots \\ \frac{a_{k1}}{(ix_I + b_r)} \end{bmatrix}, \qquad B(:,k)^T = \begin{bmatrix} \frac{a_{r2}}{(iy_1 + b_r)} & \cdots & \frac{a_{r2}}{(iy_J + b_r)} \end{bmatrix}$$

$$a_r = a_{r1} * a_{r2}$$

- Parameters a_{k1}, a_{k2} and b_k are estimated using least squares from A(:,k) and B(:,k).
- Water component is estimated using only the sources and amplitudes that are outside the region of interest (0.25-4.2 ppm).
- Finally the water component is suppressed by subtracting the estimated water signal from the measured signal.

Löwner based water suppression- Baseline

- Problem: In some voxels, water suppression will result in a baseline at the edges of the spectra.



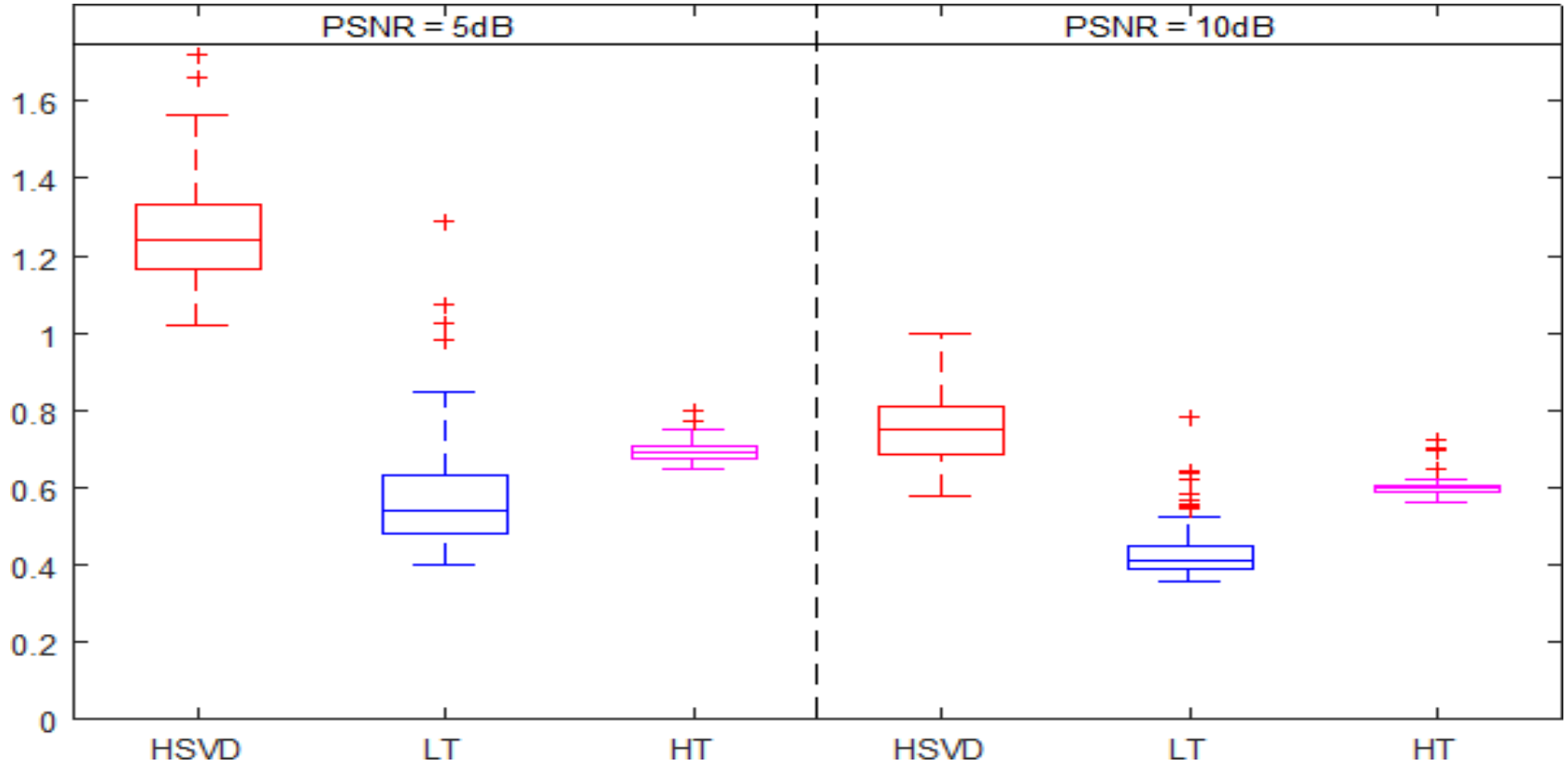
- Solution: Model the baseline using polynomial function and add it to the source matrix W.

$$W_{poly} = \begin{bmatrix} w_{11} & \cdots & w_{1R} & 1 & \cdots & f_1^d \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ w_{N1} & \cdots & w_{NR} & 1 & \cdots & f_N^d \end{bmatrix}$$

Hankel based water suppression- Results

Simulation:

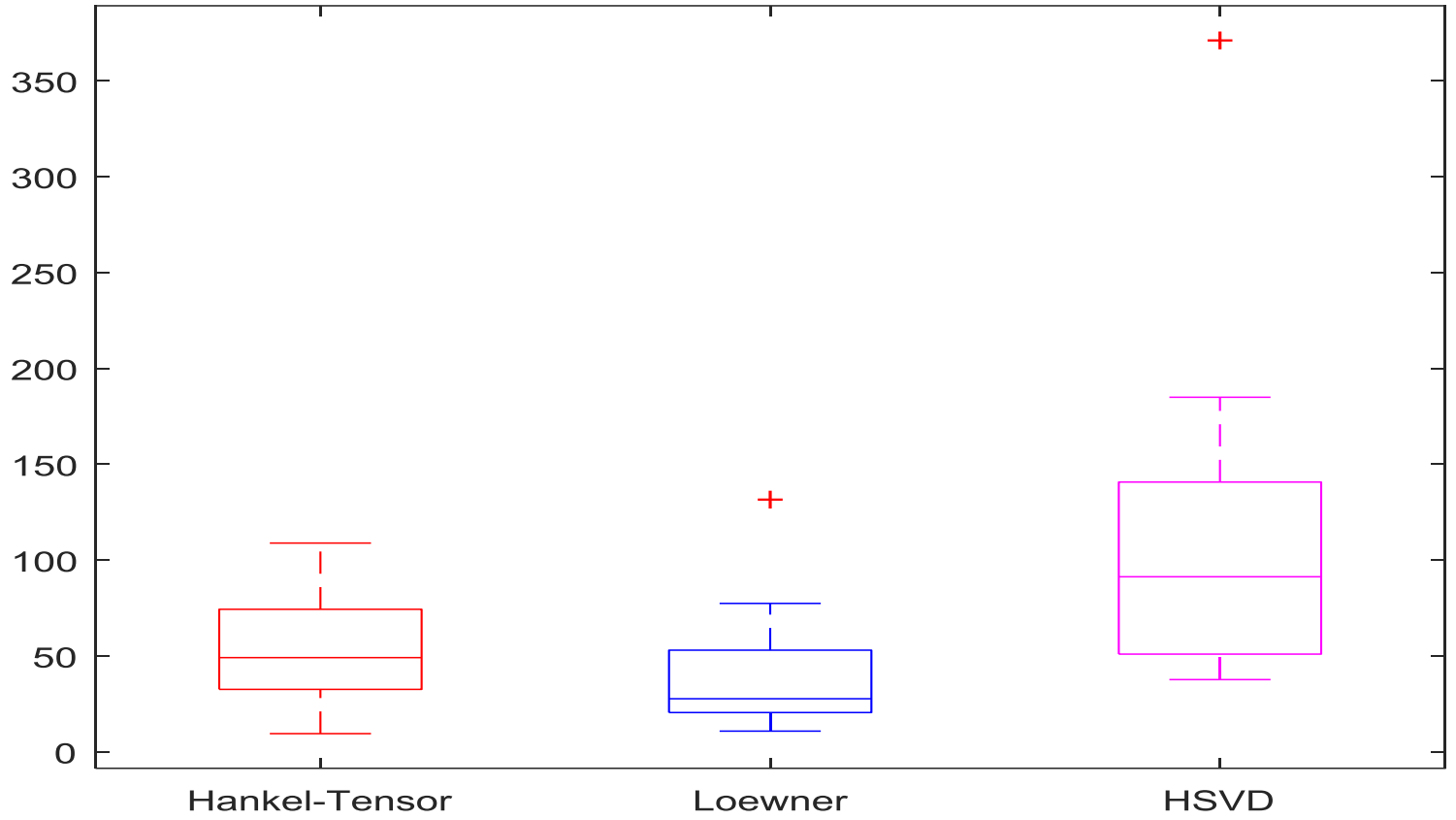
- Metabolite spectra generated using in-vitro signals.
- Residual water generated using scaled water reference signal measured in-vivo.



Box-plot of errors on simulated MRSI data

Hankel based water suppression- Results

- Water suppression methods are tested on 28 in-vivo measured MRSI signals.

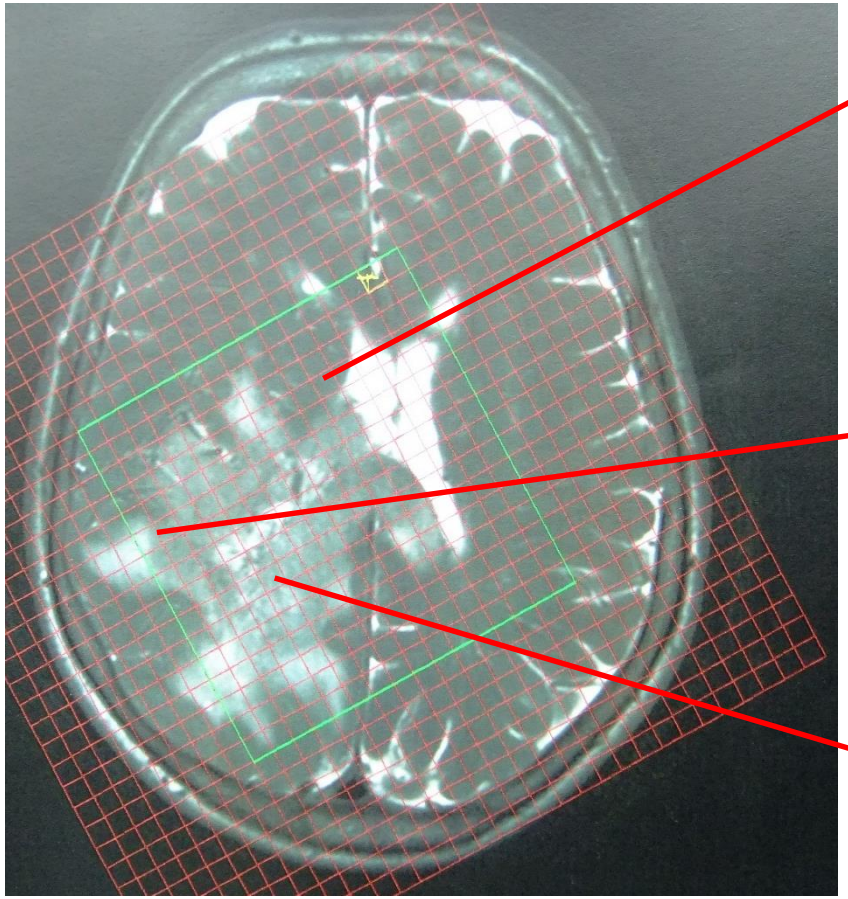


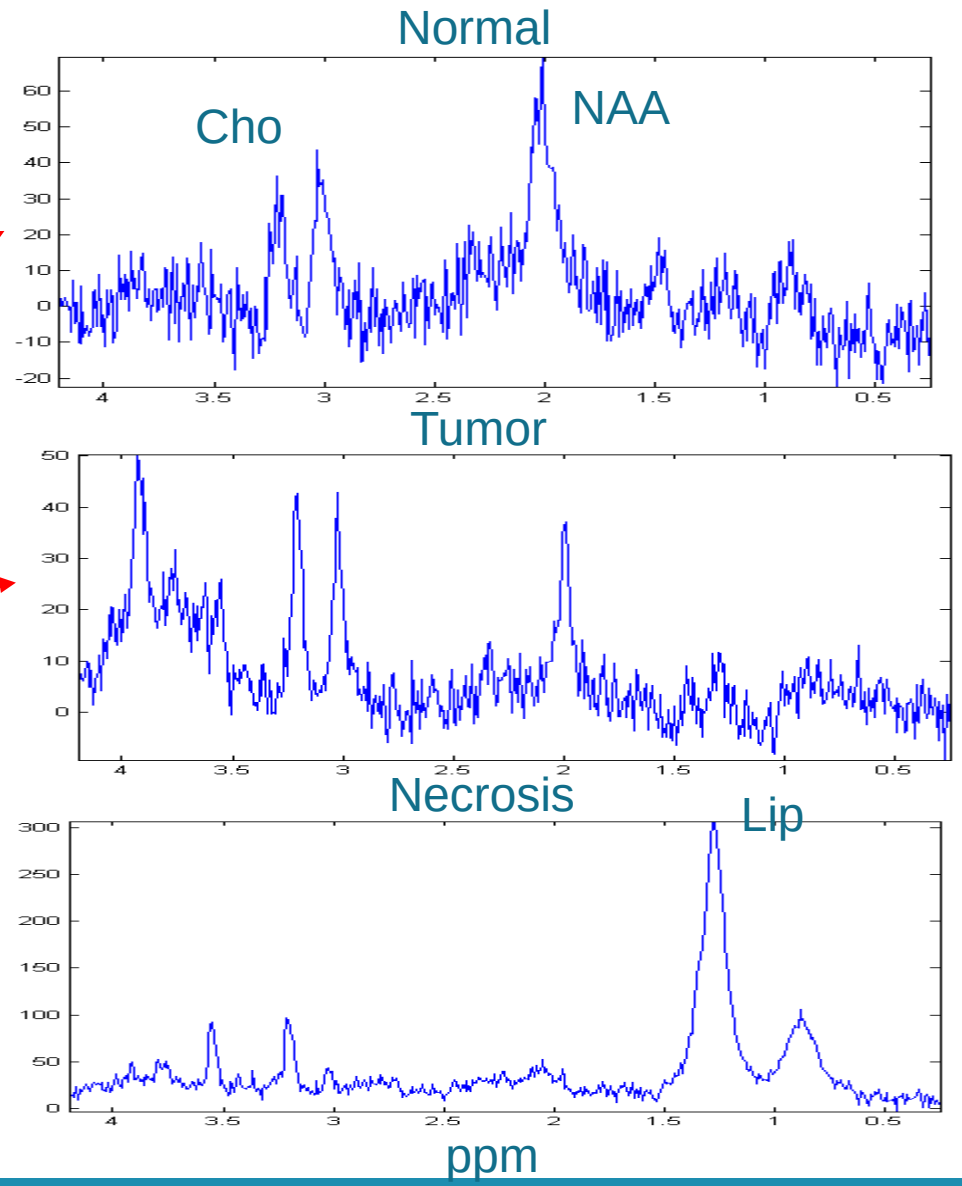
Box-plot displays differences in variance in suppressed region of the in-vivo MRSI data

Contents Overview

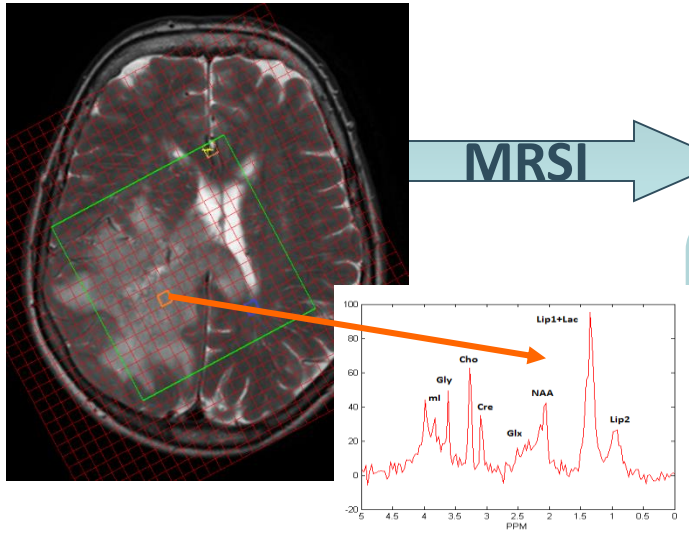
1. Introduction
   - Magnetic Resonance Spectroscopic imaging
2. Residual Water Suppression
   - Hankel-tensor based water suppression
   - Löwner-tensor based water suppression
3. **Tumor tissue type differentiation**
   - From MRSI
   - From MP-MRI
4. MRSI tumor voxel classification
5. Tumor segmentation from MP-MRI

Gliomas- Pathologic tissue spectrum





Unsupervised Brain Tumor Diagnosis using NMF



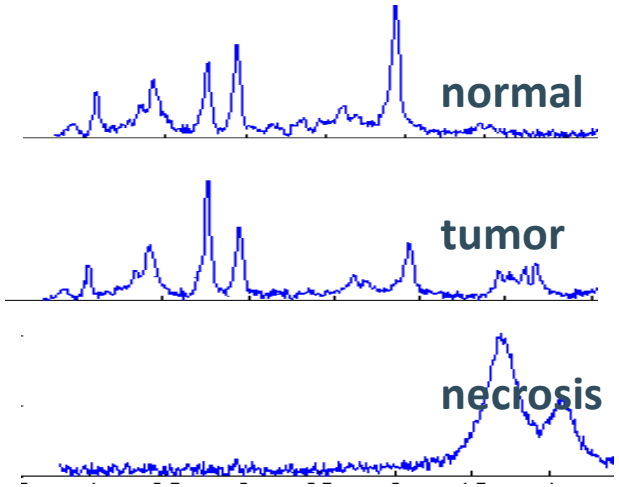
MRSI

Y = matrix of spectra, $Y \approx W H$

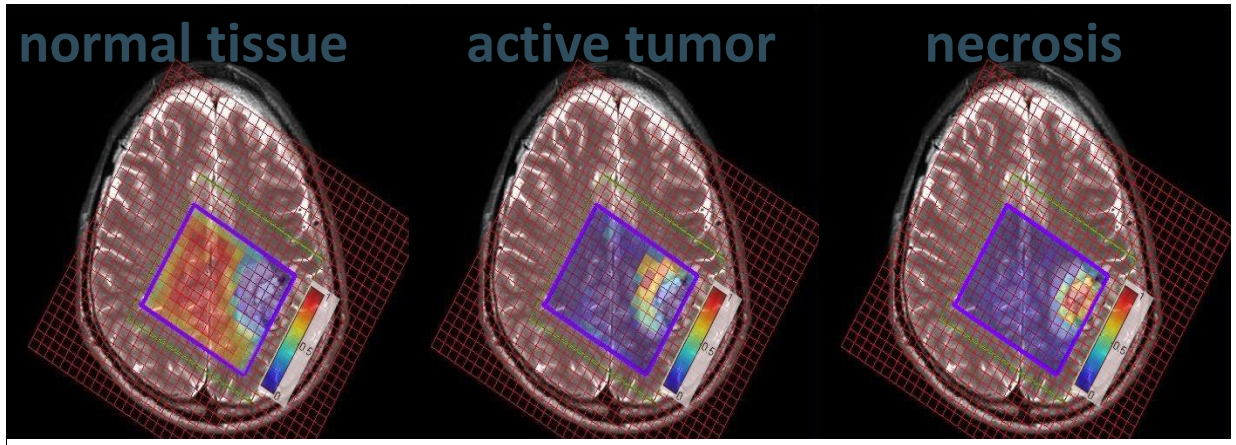
$$\min \| Y - W H \| \quad \text{such that } W \geq 0, H \geq 0$$

non-negative matrix factorization

W = tissue-specific spectral patterns:

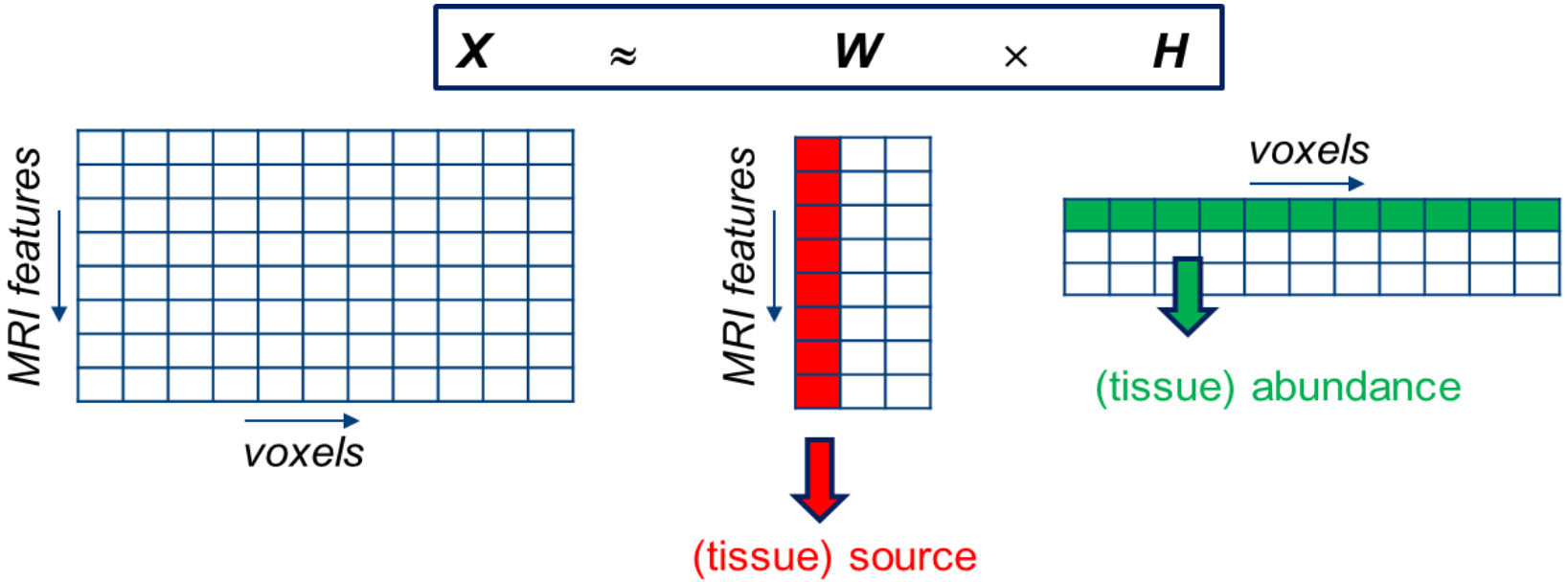


spatial distribution of tissue types:



glioblastoma multiforme patient

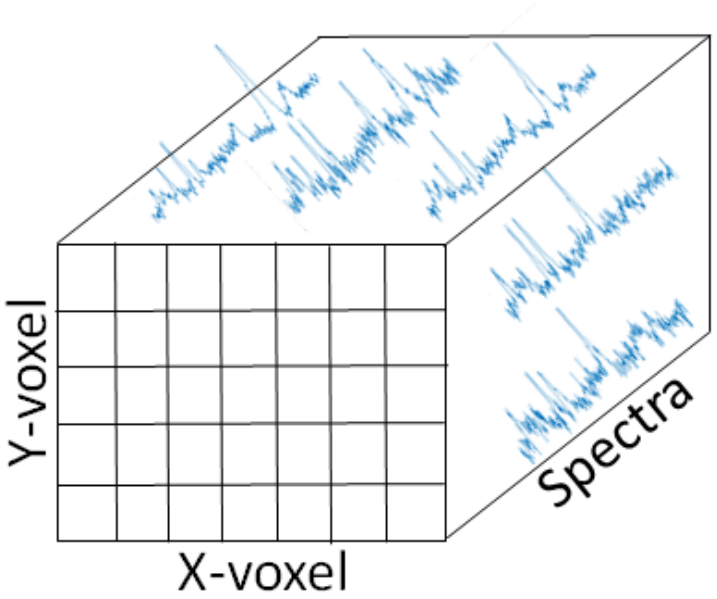
Non-negative matrix factorization (NMF)



- Non-negativity constraint: $X_{i,j}$, $W_{i,j}$, $H_{i,j} \geq 0$, $\forall i,j$
  - Columns of W ~ sources (tissue signatures)
  - Rows of H ~ source abundances (tissue weights)
- Unsupervised: directly applicable patient-by-patient, but tissue classes not a priori known

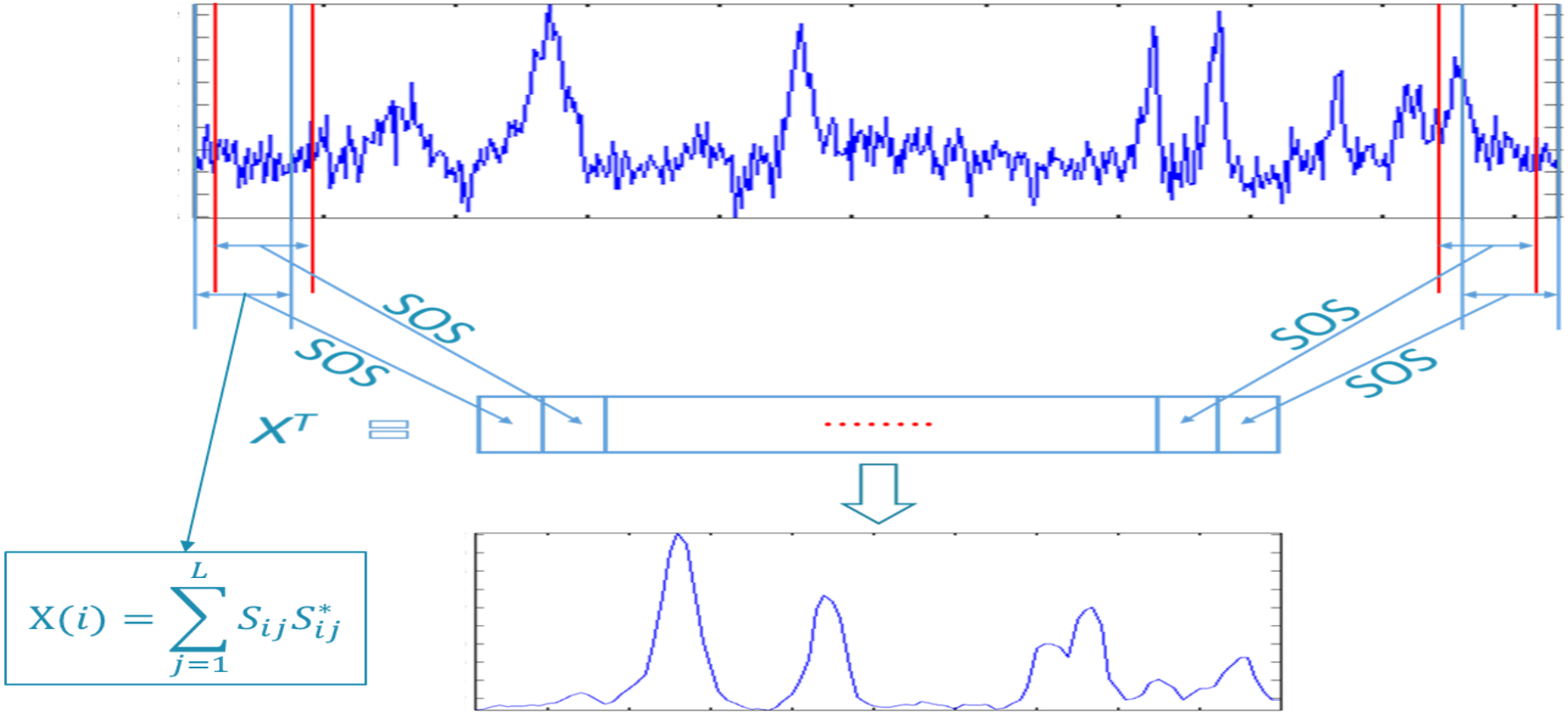
Spatial MRSI tensor

$\mathcal{P}$ =



- Use rank-$(L_r, L_r, 1)$ block term decomposition to estimate tissue specific spectra and their corresponding distribution.
  - Tissue distribution may not be of low rank.
  - Difficult to estimate the rank L_r for each term.

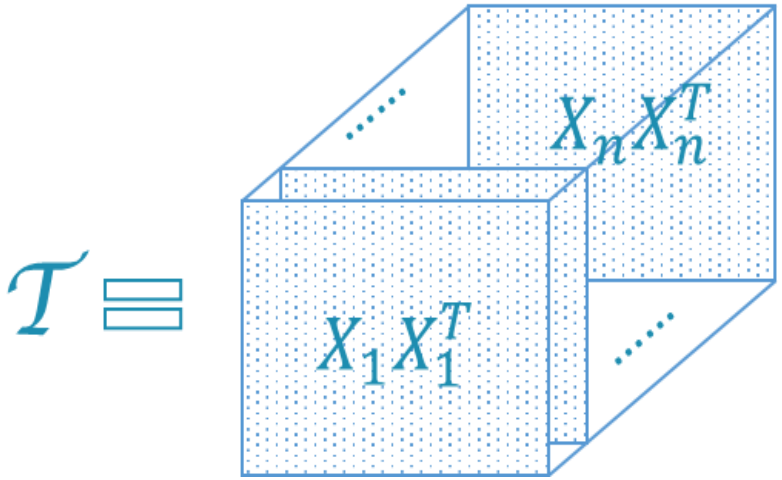
Tissue type differentiation from MRSI

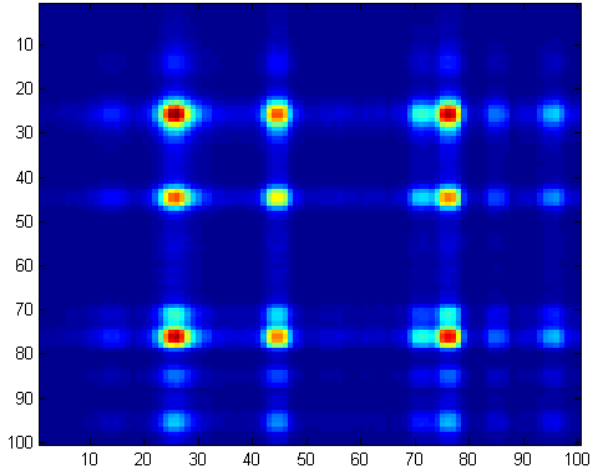


- Reduces the length of spectra without losing vital information required for tumor tissue type differentiation.
- This preprocessing gives more weight to the peaks and makes the signal smoother.

Bharath, H. N., Sima, D. M., Sauwen, N., Himmelreich, U., Lathauwer, L. D., and Van Huffel, S Nonnegative canonical polyadic decomposition for tissue-type differentiation in gliomas. IEEE Journal of Biomedical and Health Informatics 21, 4 (July 2017), 1124–1132

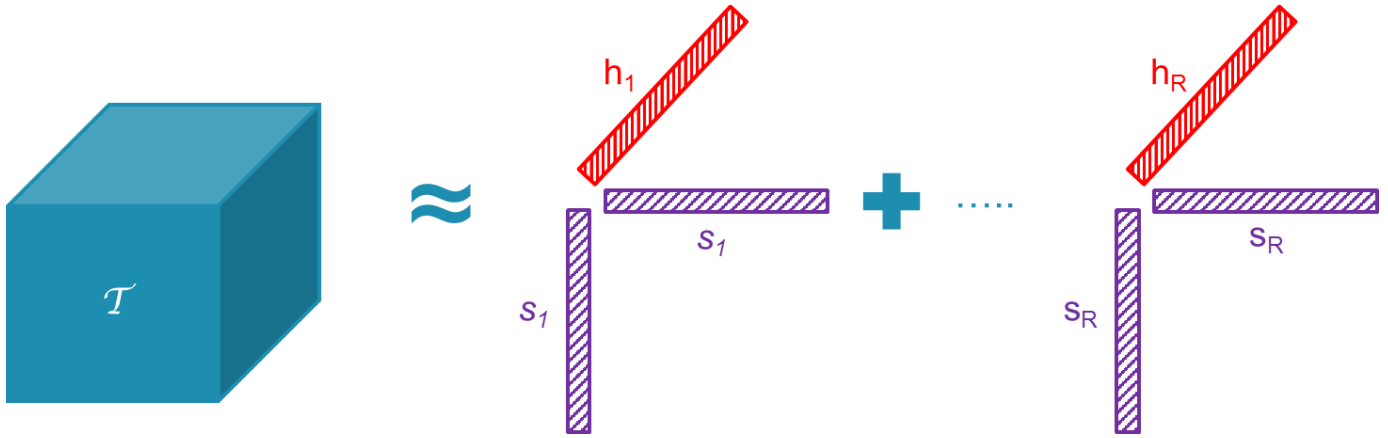
Tissue type differentiation from MRSI- XX^T tensor





- Construct a 3-D tensor by stacking XX^T from each voxel.
- MRSI tensor couples the peaks in the spectra because of the XX^T in the frontal slices.

Tissue type differentiation from MRSI: NCPD



- Non-negative constraint is applied on all 3-modes.
- To maintain symmetry in frontal slices common factor (S) is used in both mode 1 and mode 2.

$$T \approx [S, S, H] = \sum_{r=1}^{K} S(:, r) \circ S(:, r) \circ H(:, r)$$

- Non-negative CPD is performed in Tensorlab toolbox using structured data fusion.

l_1 Regularized NCPD

- Here, we assume that spectra corresponding to each voxel belongs to a particular tissue type, therefore the factor matrix H will be sparse.
- Non-negative CPD with l_1 regularization on the abundances H.

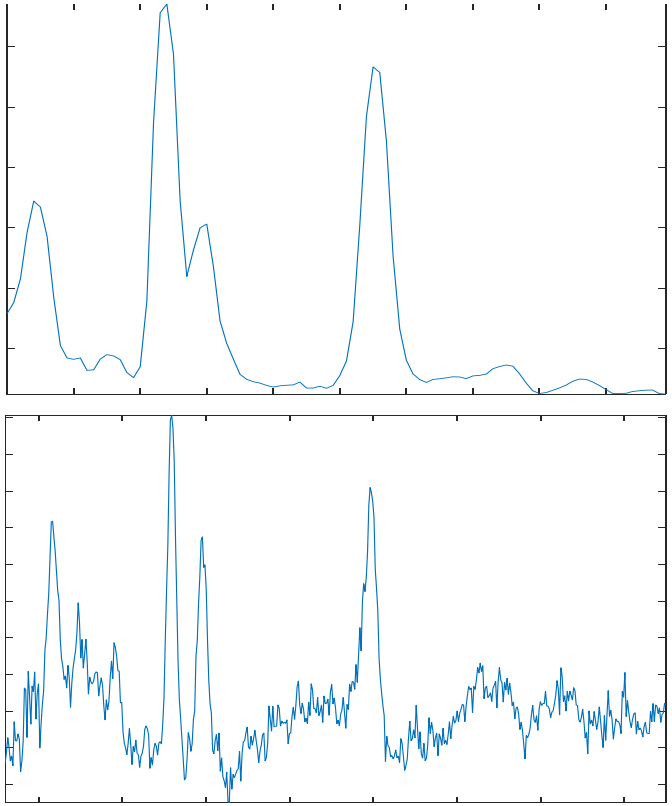
$$[S^*, H^*] = \min_{S,H} \left\| T - \sum_{r=1}^{K} S(:,r) \circ S(:,r) \circ H(:,r), \right\|_2^2 + \lambda |Vec(H)|_1$$

where λ controls the sparsity in H.

- Use more sources (higher rank) to accommodate for artifacts and variations within tissue types.
- Original source spectra are recovered using least squares:

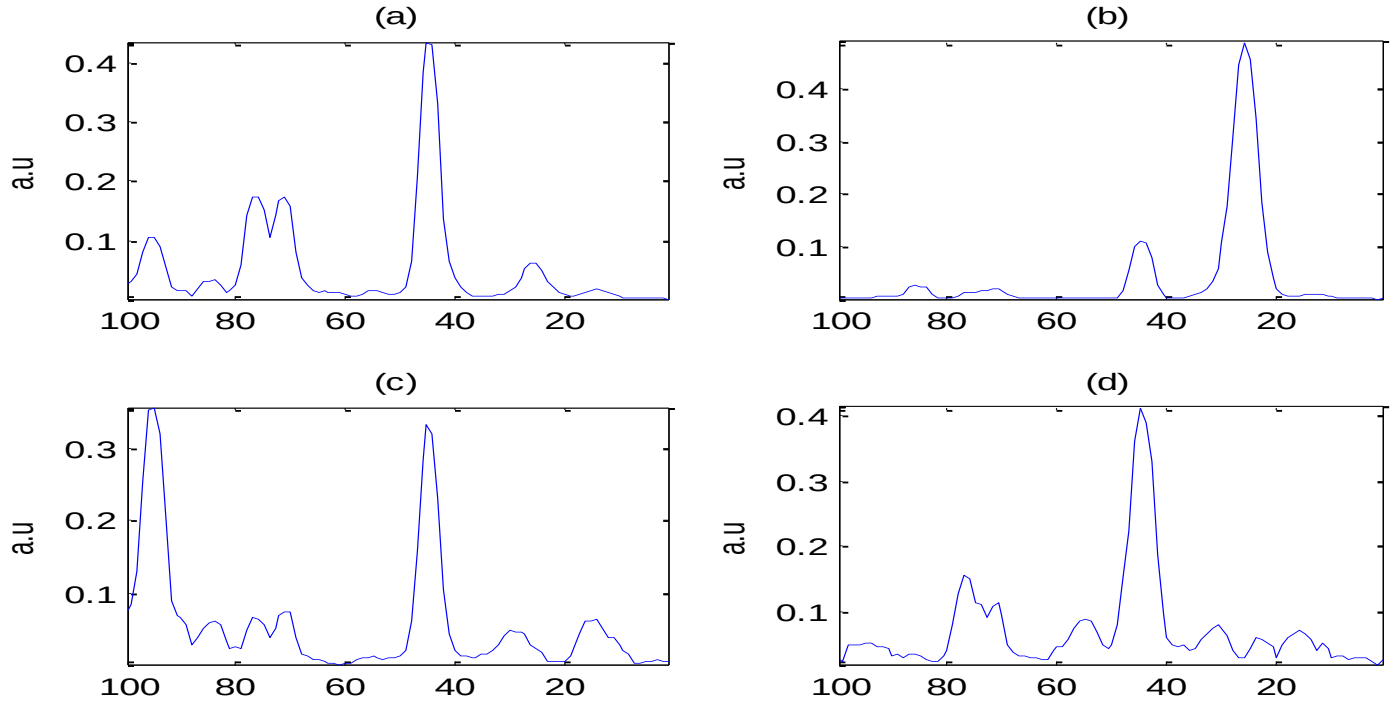
$$S = (H^{\dagger} Y^T)^T$$

$H^{\dagger}$ is the pseudo-inverse of *H* obtained from Non-negative CPD.



Initialization

- Take SVD of the complex spectra, $Y = U\Sigma V^H$.
- Construct the reduced spectrum from *R* dominant singular vectors. Use this to initialize *S* in the NCPD.
- *H* in NCPD is initialized using least squares, $H_{init} = (S_{init}^{\dagger} A)^T$.



Rank estimation

- Estimate the covariance matrix from the reduced spectra data matrix, *A*.

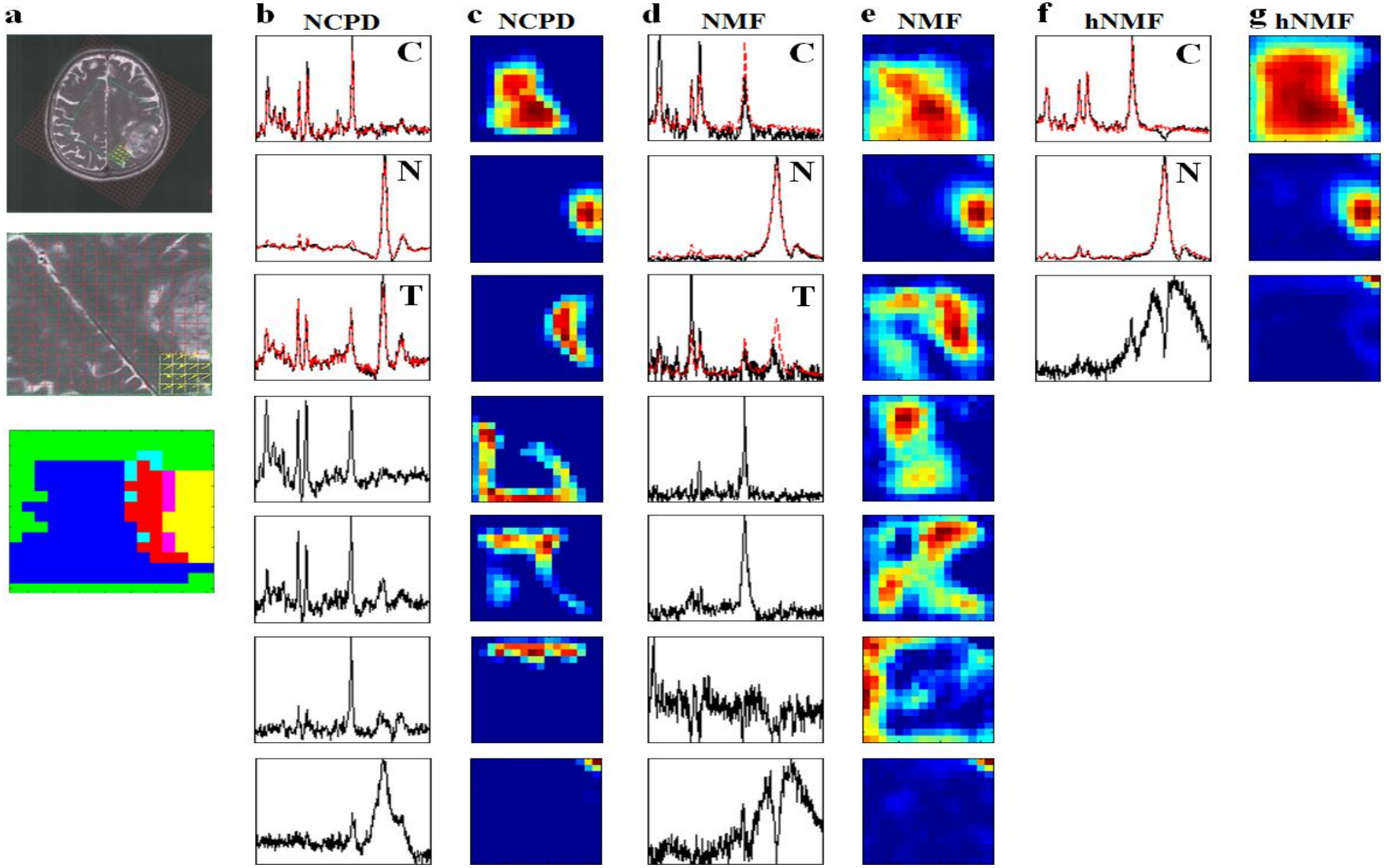
$$C = E[(A - E[A])(A - E[A])^T]$$

- The number of sources is estimated as the minimum number *R* such that the cumulative sum of eigenvalues represents at least 99% of the total energy.

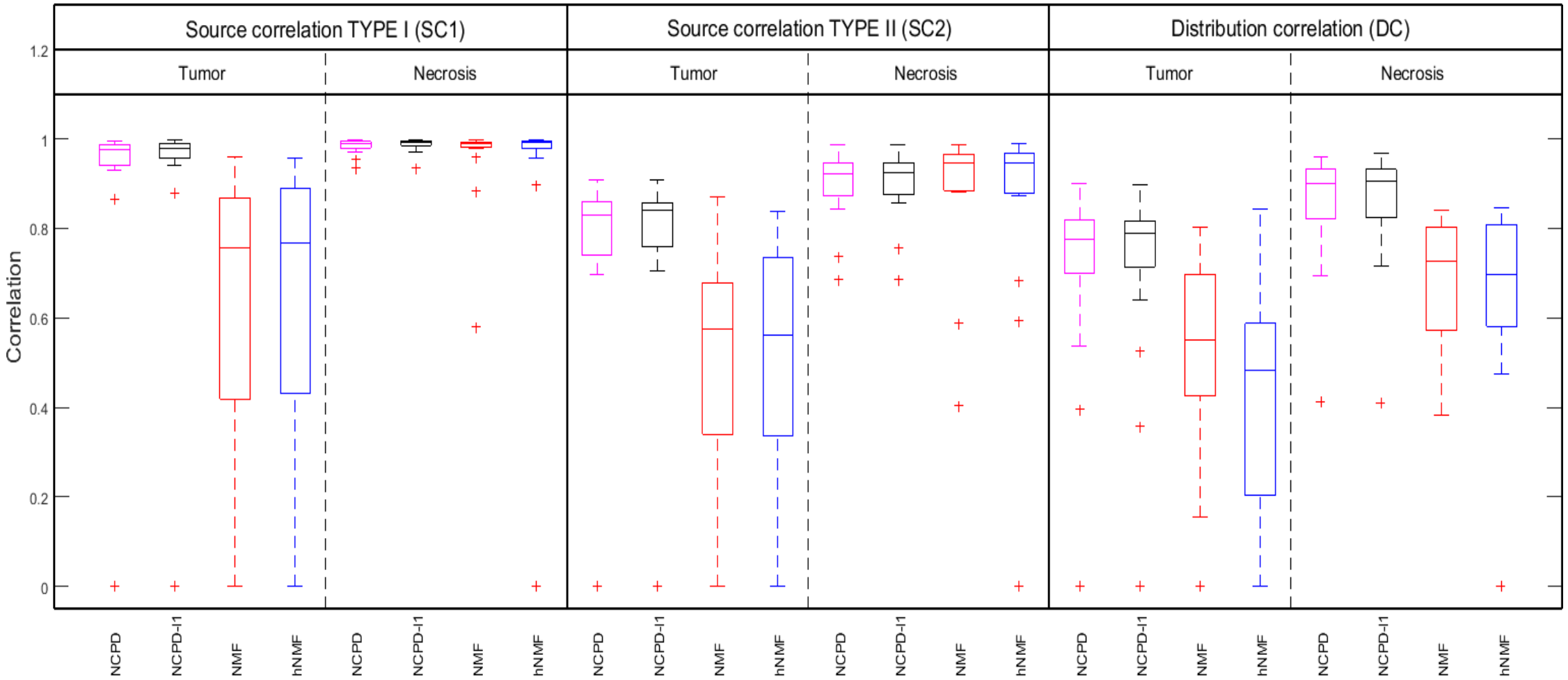
$$R^* = \arg\min_R \left[\frac{\sum_{i=1}^{i=R} \lambda_i}{\sum_{i=1}^{i=K} \lambda_i} \geq 0.99 \right]$$

- Incorporate prior knowledge about the maximum number of sources (R<=8).

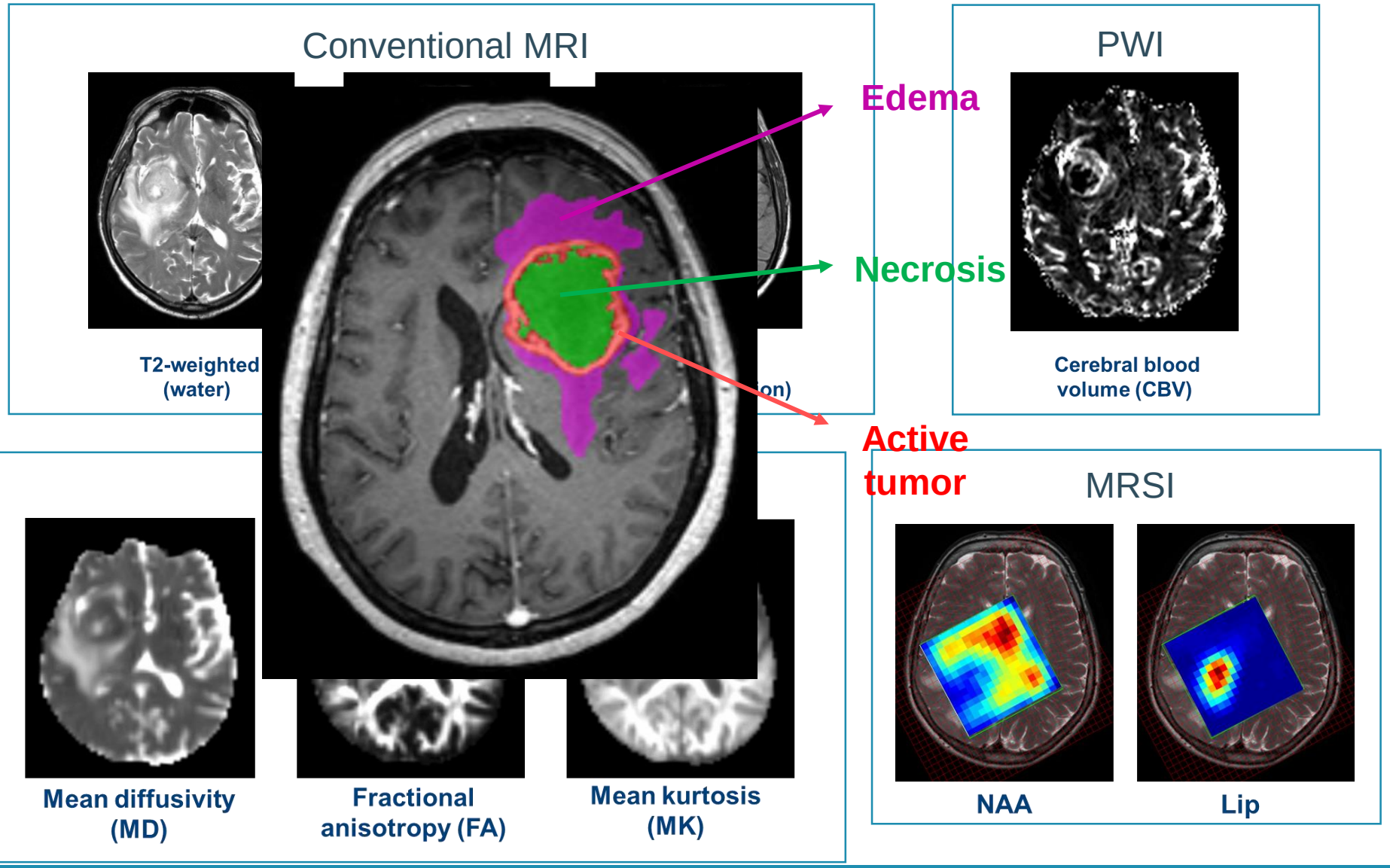
Tissue type differentiation from MRSI- Results



Tissue type differentiation from MRSI- Results

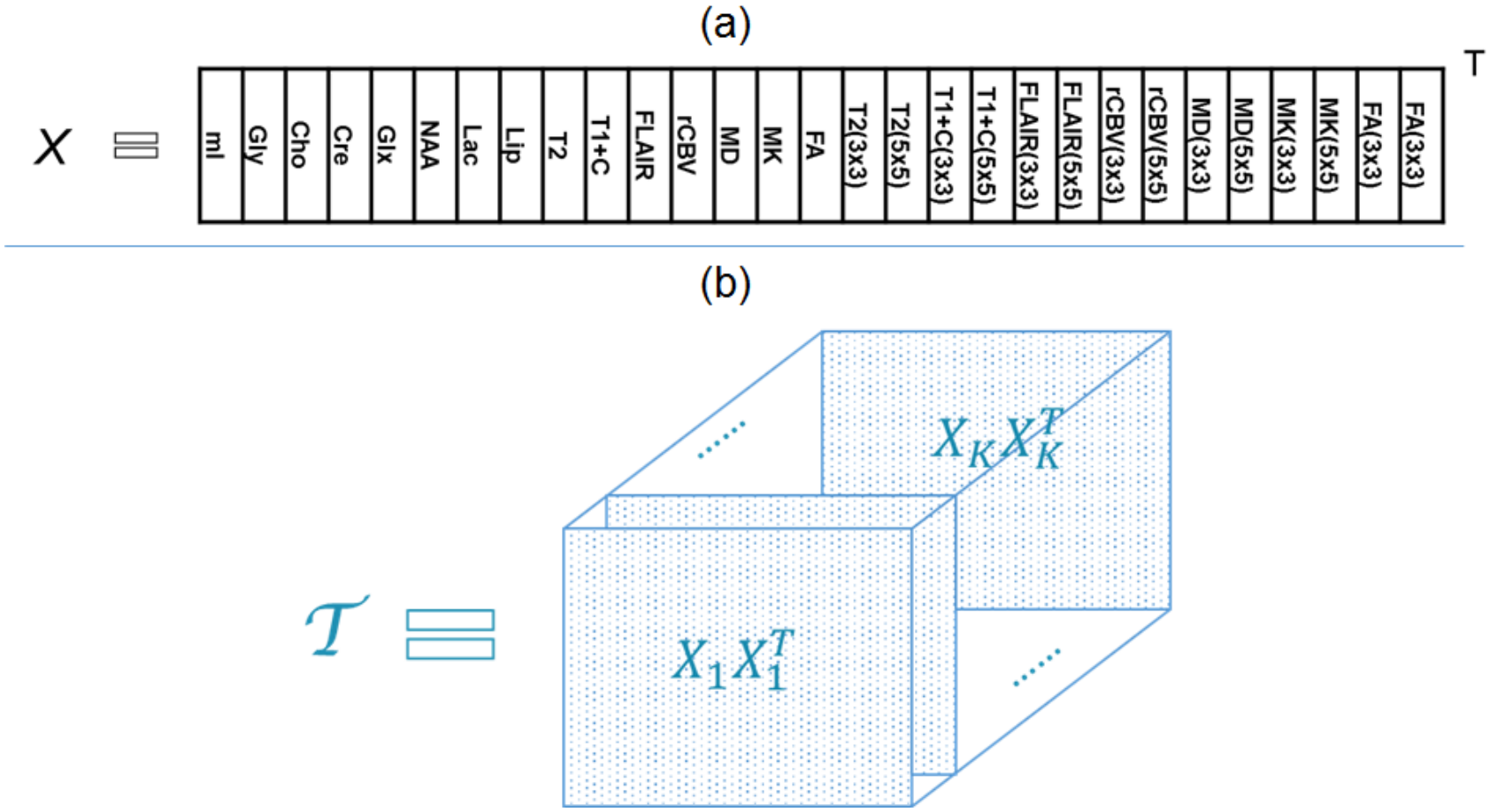


Tissue type differentiation from MultiParametric (MP)-MRI*

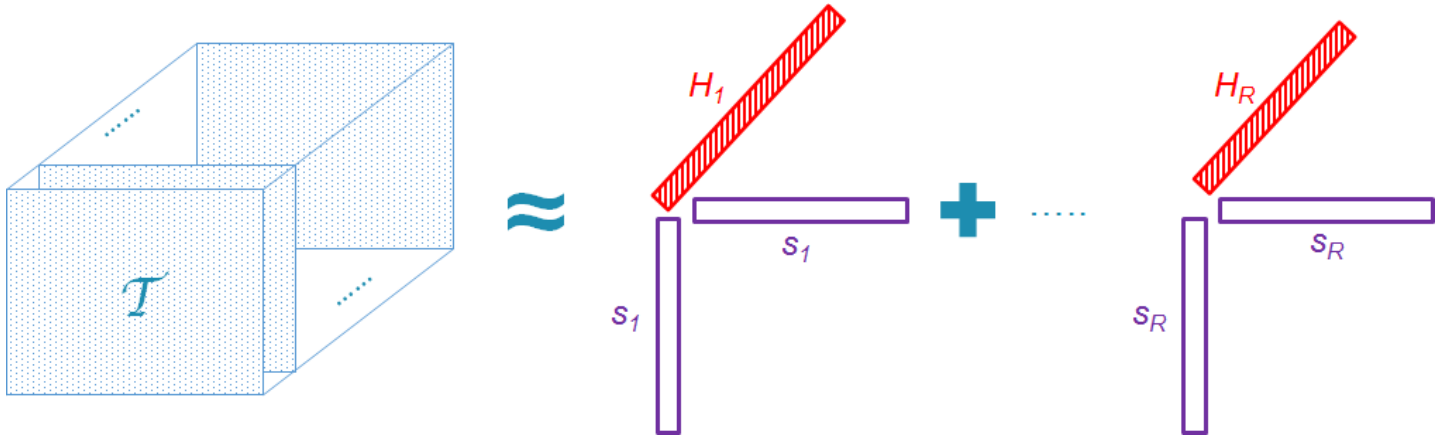


Bharath, H. N., Sauwen, N., Sima, D.M., Himmelreich, U., De Lathauwer, L., and Van Huffel, S. Canonical polyadic decomposition for tissue type differentiation using multi-parametric MRI in high-grade gliomas. European Signal Processing Conference (EUSIPCO) 2016, Budapest, Aug. 2016 (pp. 547-551).

Tissue type differentiation from MP-MRI:-XX^T tensor



Tissue type differentiation from MP-MRI:- CPD

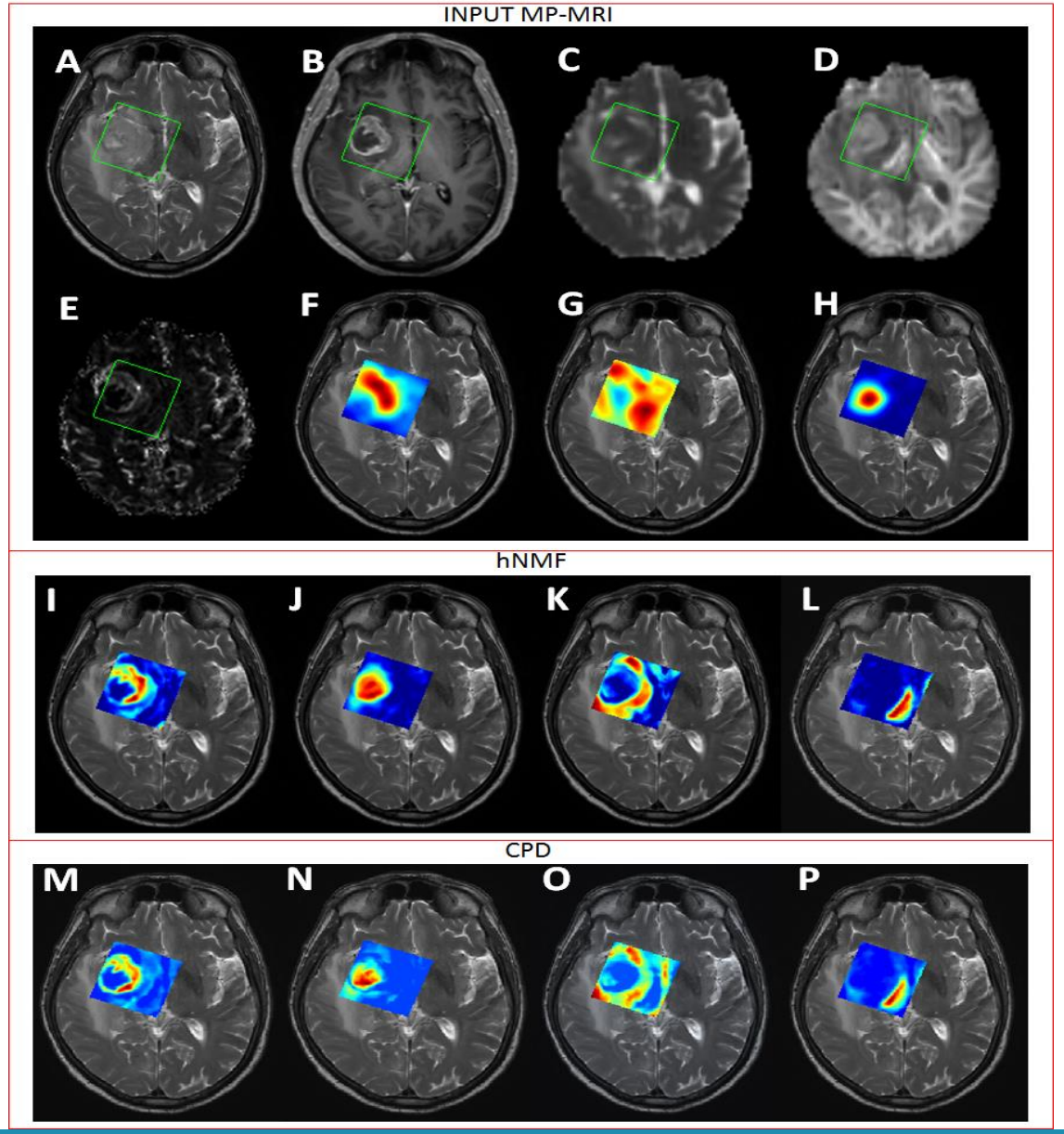


- To maintain symmetry in frontal slices common factor (S) is used in both mode-1 and mode-2.
- Non-negative constraint is applied on mode-3, H.
- Also, l_1 regularization in applied on the abundances H.

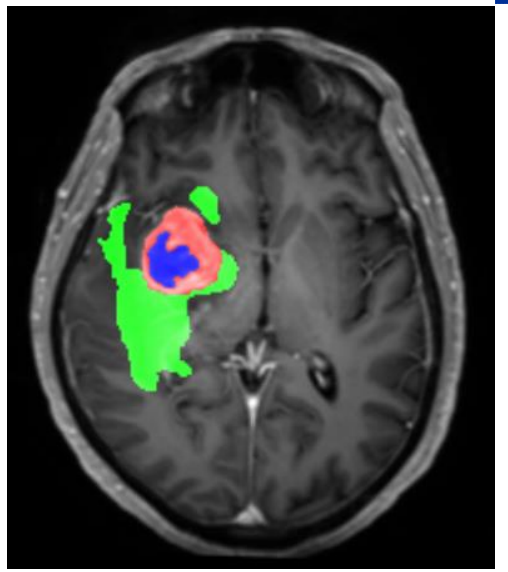
$$[S^*, H^*] = \underset{S,H \geq 0}{\arg\min} \left\| \mathcal{T} - \sum_{i=1}^{R} S(:,i) \; o \; S(:,i) \; o \; H(:,i) \right\|_2^2 + \lambda \|Vec(H)\|_1$$

- Solve using structured data fusion method in Tensorlab.

Results:



Segmentation accuracy



Based on manual segmentation by radiologist (for pathological tissue types)

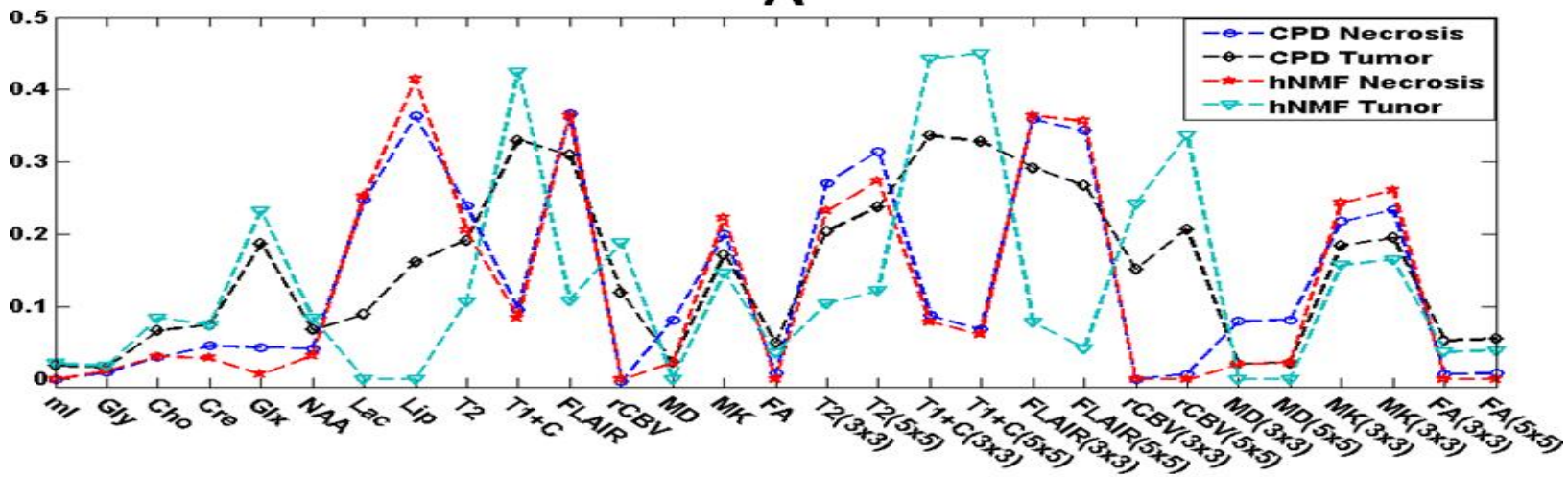
1) Dice-scores* (based on H)

$$Dice = \frac{2 * area(A \cap B)}{area(A) + area(B)}$$

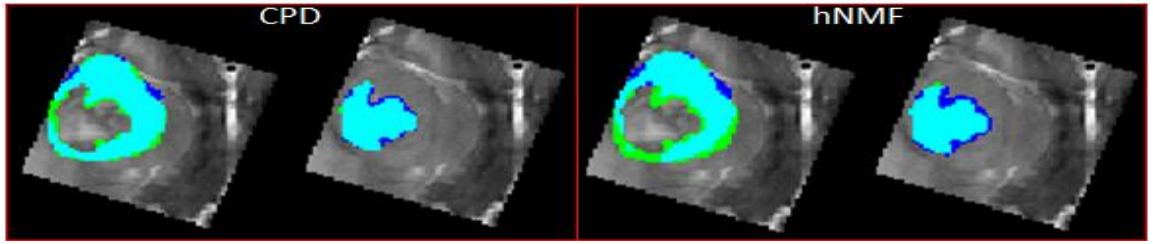
* Dice, Lee R. "Measures of the Amount of Ecologic Association Between Species". Ecology, 1945; 26 (3): 297–302.

Tissue type differentiation from MP-MRI:- results

A



B

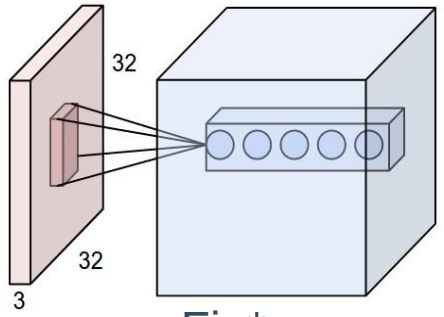


	Constrained CPD-l_1			hNMF		
	Dice Tumor	Dice Core	Tumor source Correlation	Dice Tumor	Dice Core	Tumor source Correlation
Mean	0.83	0.87	0.95	0.78	0.85	0.81
Standard deviation	0.07	0.1	0.05	0.09	0.13	0.19

Contents Overview

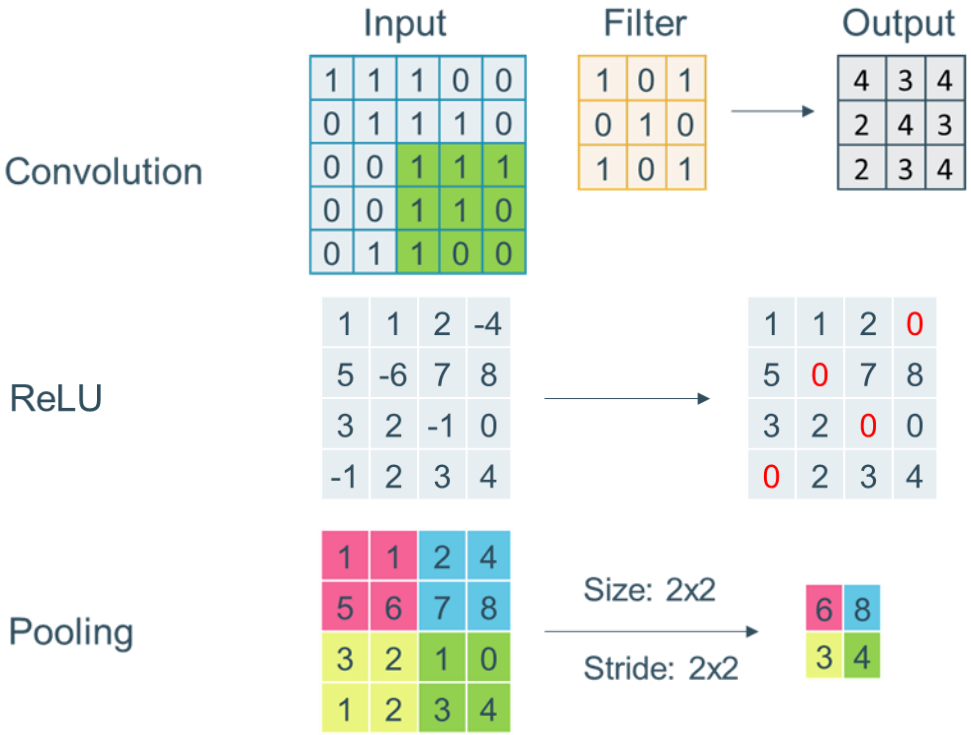
Tumor classification: CNN

- Convolutional neural network is a class of deep neural networks. Exploit spatial locality with shared weights.



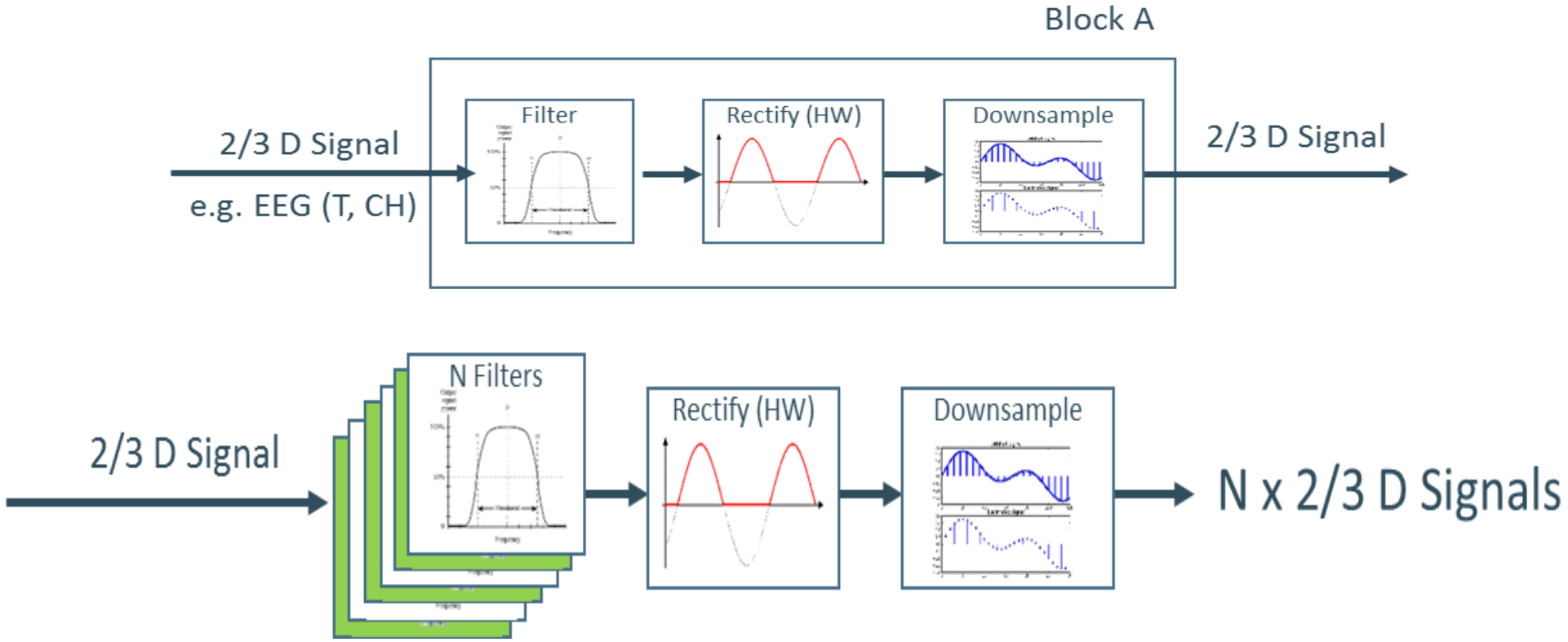
Fig*

- Three main layers
  - Convolutional layer:
    - Performs 2-D convolution
  - ReLU:
    - Applies non-linear activation function.
  - Pool:
    - Values over a window are reduced to a single value.



* https://en.wikipedia.org/wiki/Convolutional_neural_network

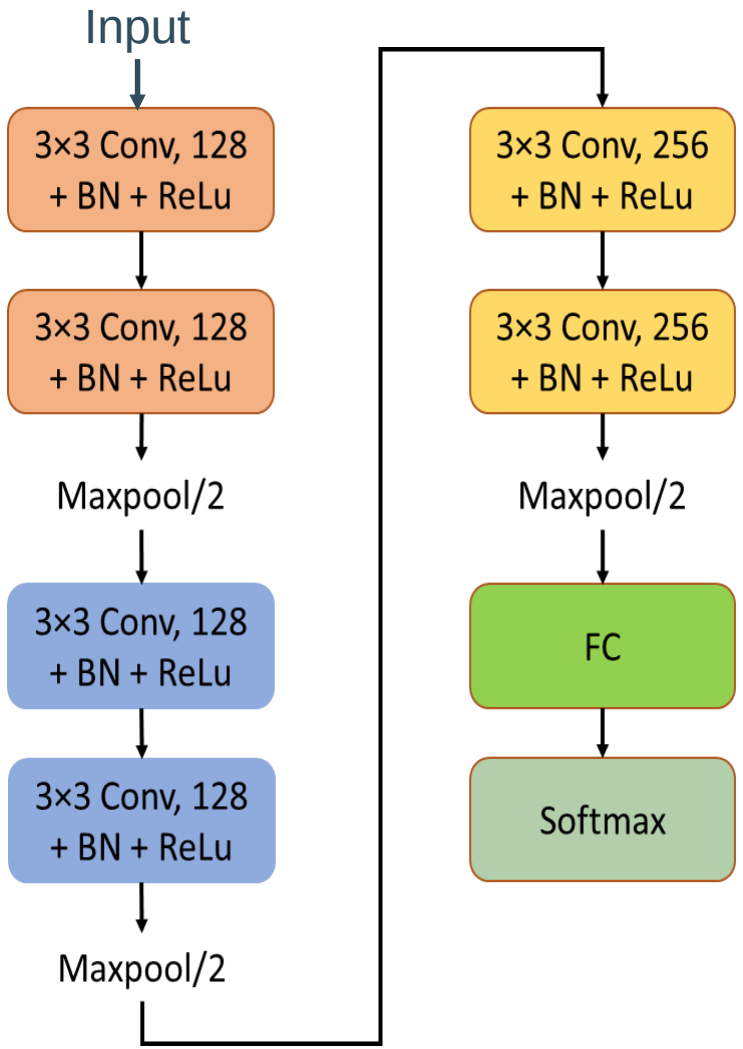
Convolutional Neural Network



Weights - 4D tensor ($I \times J \times M \times N$)
I, J filter size.
M- input dimension, N- output dimension.

Tumor classification: CNN

- CNN Architecture



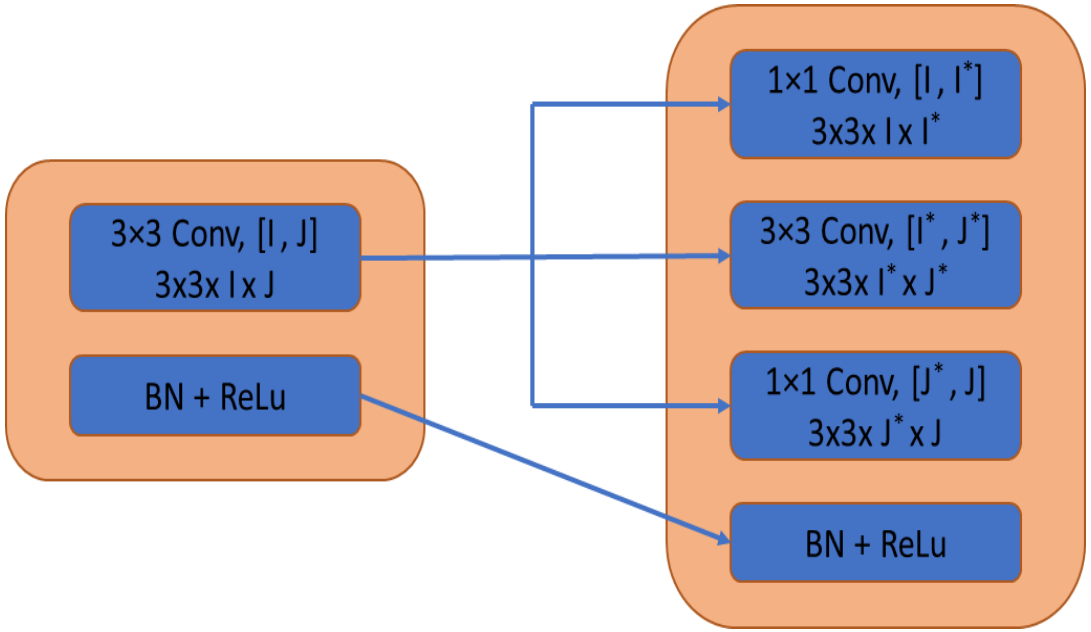
Tumor classification: CNN

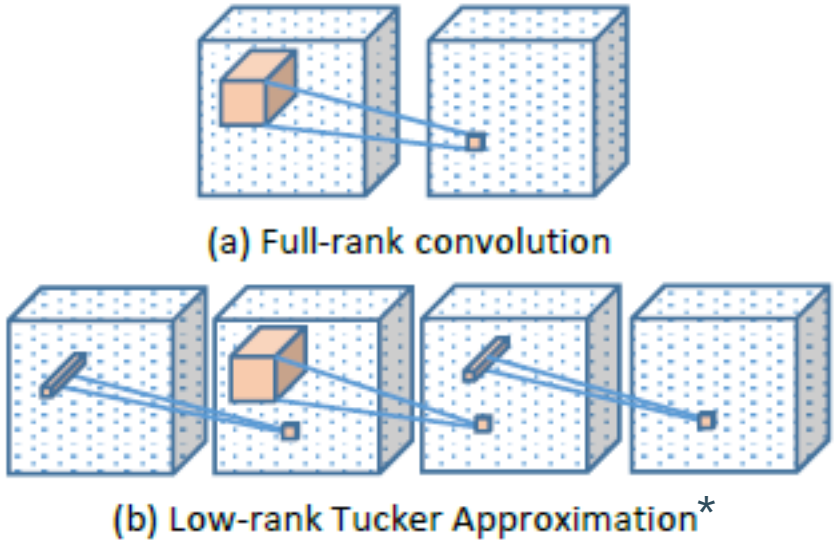
- Apply MLSVD to convolution tensor $\mathcal{T} \in 3 \times 3 \times I \times J$.

$$\mathcal{T} = S \times_1 A_1 \times_2 A_2 \times_3 A_3 \times_4 A_4$$

$$\mathcal{T} \approx S1 \times_3 \hat{A}_3 \times_4 \hat{A}_4 \quad , \quad S1 \approx \hat{S} \times_1 A_1 \times_2 A_2$$

$$\hat{S} = S(1{:}I, 1{:}J, 1{:}K^*, 1{:}L^*), \quad \hat{A}_3 = A_3(:, 1{:}K^*), \hat{A}_4 = A_4(:, 1{:}L^*), \qquad K^* < K \, , L^* < L$$





* Yunpeng, C., Xiaojie, J., Bingyi, K., Jiashi, F., and Shuicheng, Y. Sharing Residual Units Through Collective Tensor Factorization in Deep Neural Networks. arXiv preprint arXiv:1703.02180 (2017).

Tumor classification: CNN

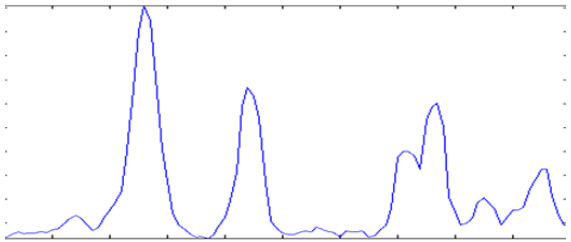
- First train the CNN for 10 epochs.
- Apply low-rank regularization to all the convolution layers except the first one.
- Reduce the dimension in the input and output mode such that the truncated tensor retains 80% of the variance in the respective mode.
- Replace convolution layers with low rank model and initialize the weights with truncated tensor and factor matrix.
- Train the low rank-rank CNN model for 40 more epochs.
- CNN design and training is performed in MatConvNet

Tumor classification: CNN

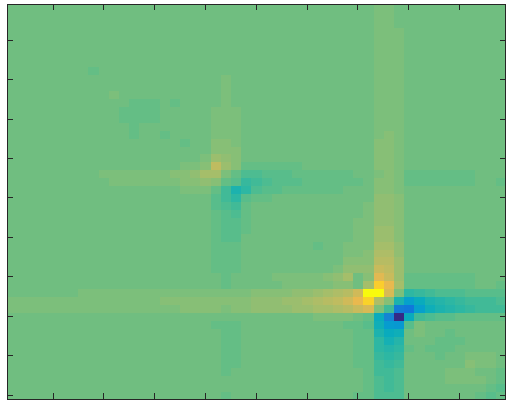
- Given a signal S evaluated at points N = {n_1, n_2,…..n_N}, Loewner matrix is defined as

$$L = \begin{bmatrix} \frac{S(x_1)-S(y_1)}{x_1-y_1} & \cdots & \frac{S(x_1)-S(y_J)}{x_1-y_J} \\ \vdots & \ddots & \vdots \\ \frac{S(x_I)-S(y_1)}{x_I-y_1} & \cdots & \frac{S(x_I)-S(y_J)}{x_I-y_J} \end{bmatrix}$$

With X = {n_1, n_3,…..} and Y = {n_2, n_4,…..}

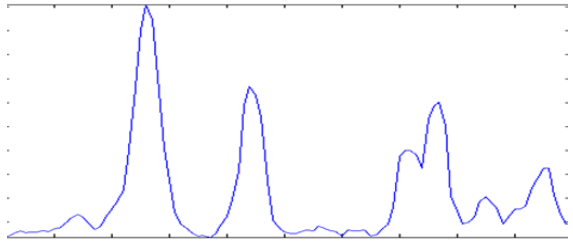


loewnerize



Tumor classification: Random forest

- The reduced spectra are used as features.
- Construct 200 tree random forest classifier with three classes: *tumor, normal and bad*.



- Class imbalance:
  - Tumor class is given more weight compared to others.
- Leave-one-out cross-validation
  - 17 patients (28 MRSI dataset).
  - 5904 voxels in total.

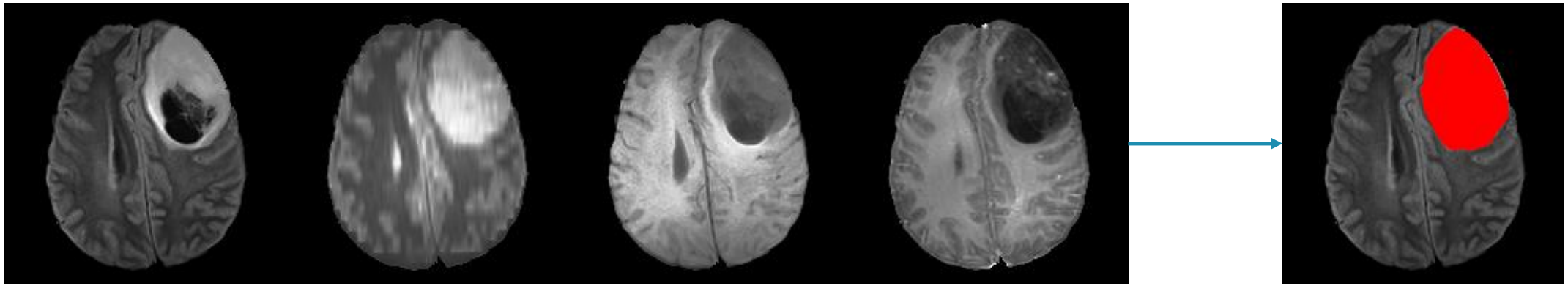
Results

Tumor	sensitivity	specificity	precision	F1 score
NCPD	81.01%	87.96%	70.36%	0.7437
Random forest	76.89%	**94.46%**	80.76%	0.7739
CNN	81.58%	94.22%	**82.70%**	**0.8089**
CNN + tensor regularization	**82.90%**	93.18%	81.06%	**0.8087**

Normal	sensitivity	Specificity	precision	F1 score
NCPD	71.69%	86.06%	63.02%	0.5851
Random forest	84.15%	**94.27%**	**84.08%**	0.7702
CNN	**87.39%**	92.96%	77.76%	**0.7862**
CNN + tensor regularization	85.85%	93.91%	78.86%	**0.7835**

Contents Overview

Multi-Parametric (MP) MRI: Tumor segmentation

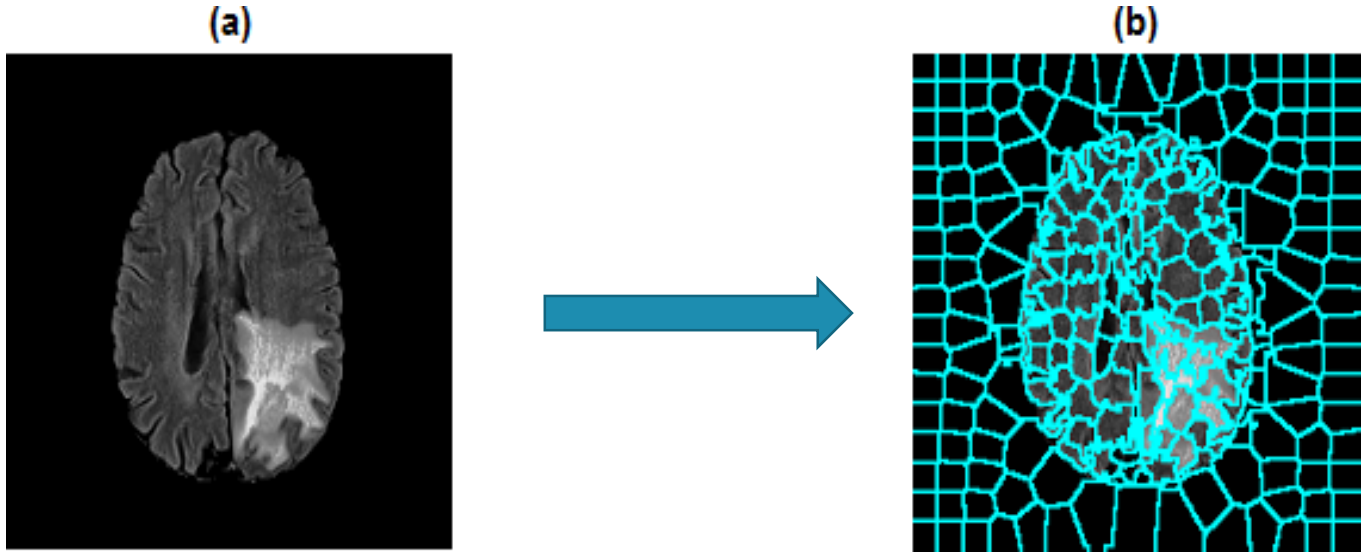


T2 FLAIR T1 T1c Expert Segmentation

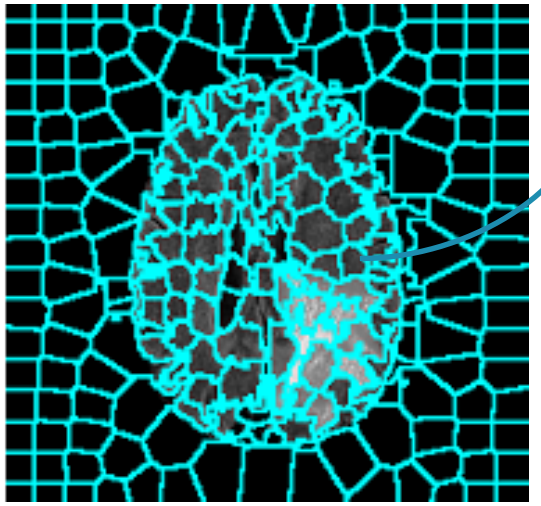
Bharath, H. N., Colleman, S., Sima, D.M., Van Huffel, S. (2018) Tumor Segmentation from Multimodal MRI Using Random Forest with Superpixel and Tensor Based Feature Extraction. In: Crimi A., Bakas S., Kuijf H., Menze B., Reyes M. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2017. Lecture Notes in Computer Science, vol 10670. Springer, Cham.

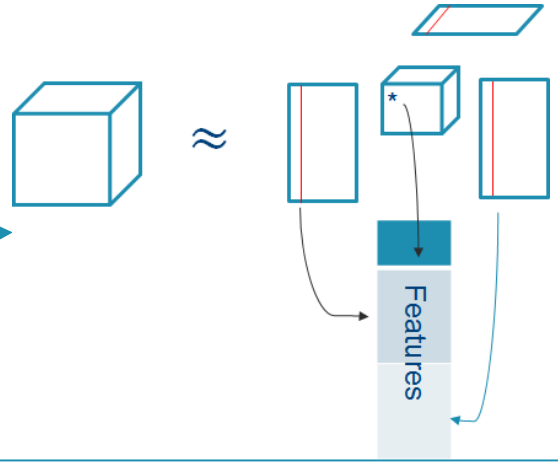
Super-pixel

- Each image is divided into smaller patches which are better aligned with intensity edges, called superpixels.
- The tissue assignment is done on superpixel-level instead of individual pixel, which helps to reduce computational cost and improve spatial smoothness.



MP-MRI: RF based tumor segmentation

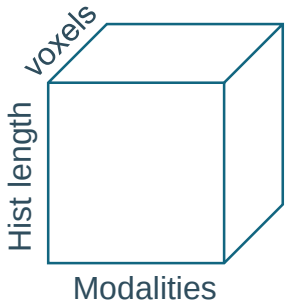




- Local histograms
- Intensities of the modalities
- Co-variance between modalities

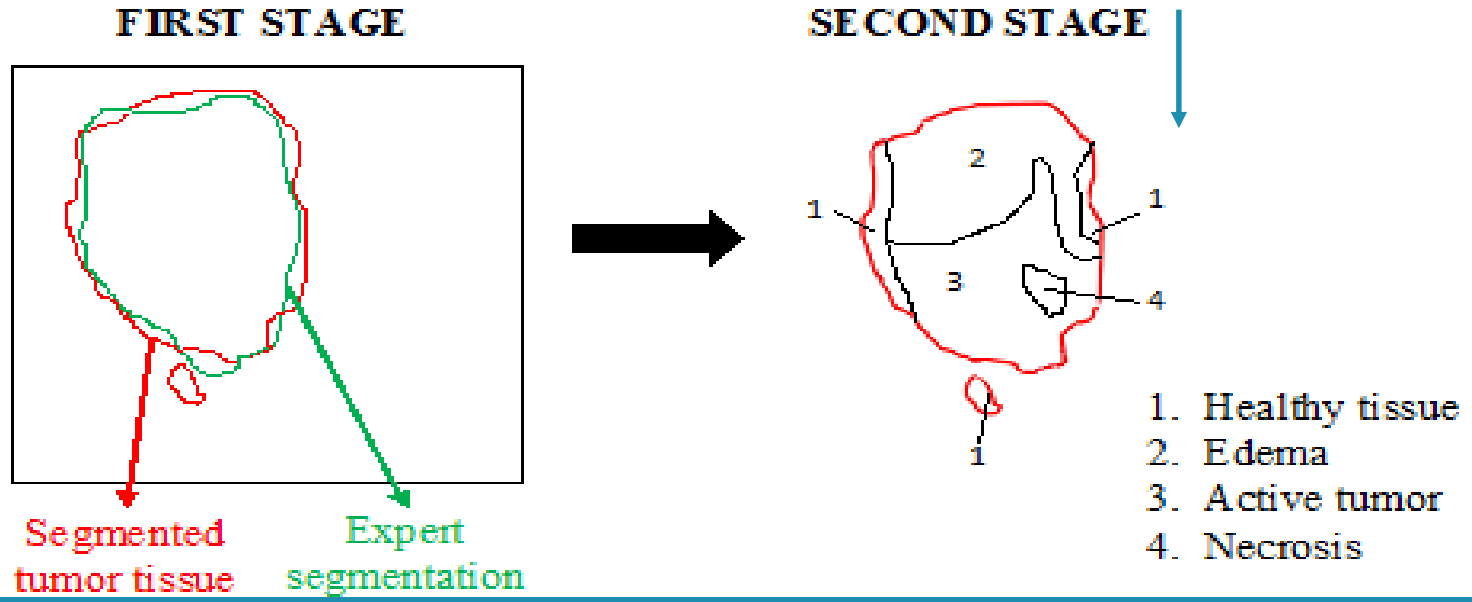
Others

- Average intensities
- Range values
- Standard deviation



MP-MRI: RF based tumor segmentation

- Two-stage Random forest classifier
  - First stage: whole tumor segmentation
  - Second stage: sub-region segmentation from the whole tumor
- Random forest classifier
  - 100 trees for first stage, 200 trees for second stage

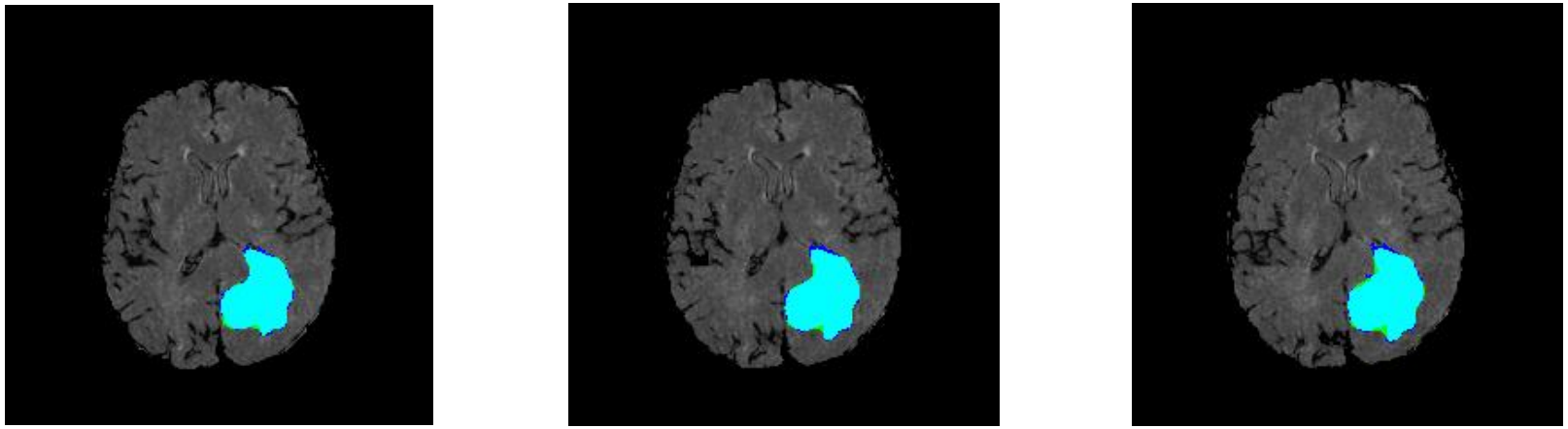


Tumor Segmentation from MP-MRI: Processing

- Pre-processing:
  - Histogram matching algorithm was applied to all the individual MRI data. Histogram from one of the patients was chosen as reference template histogram.
  - Each image data was scaled between 0-1.
  - Background is removed using Otsu's method.
  - Only the slices where tumor was present was considered for training.
- Post-processing:
  - Image-open followed by image-close operation is performed on the segmented image.

Tumor Segmentation from MP-MRI: Results

- Tested on BRATS 2017* validation dataset.
  - Tested on 46 patients
  - Average dice-score of 83.3% with standard deviation of 11.86% for total tumor.



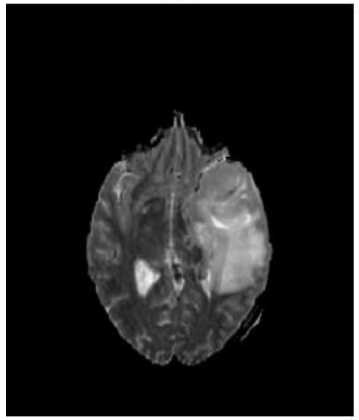
Blue:-Expert; Green:-MLSVD segmentation;
Cyan:- Overlap between expert and MLSVD

*Menze et al., The Multimodal Brain TumorImage Segmentation Benchmark (BRATS), IEEE Trans. Med. Imaging, 2015

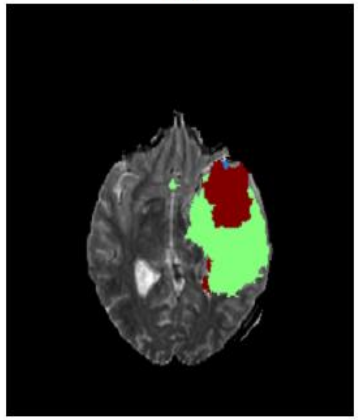
Results: High Grade Glioma

- 210 High Grade Glioma patients from BRATS 2017 dataset.
- Training – 70% (147)
- Testing --- 30% (63)
- Dice score
  - Enhanced tumor (Brown) - 76.1%
  - Tumor core (Brown + blue) - 78.3%
  - Whole Tumor (Brown + blue + green) - 83.3%

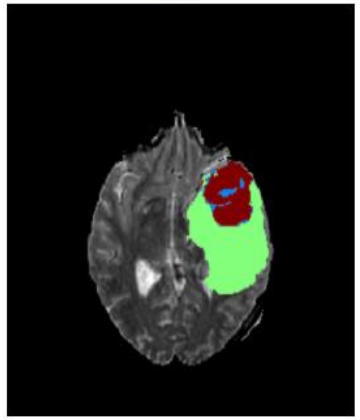
1:(a)



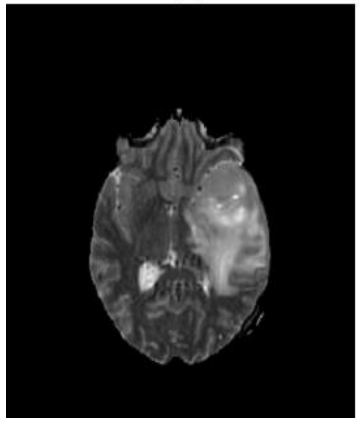
1:(b)



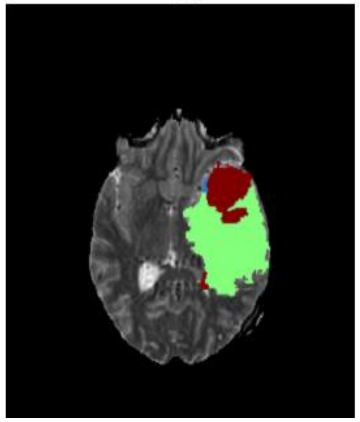
1:(c)



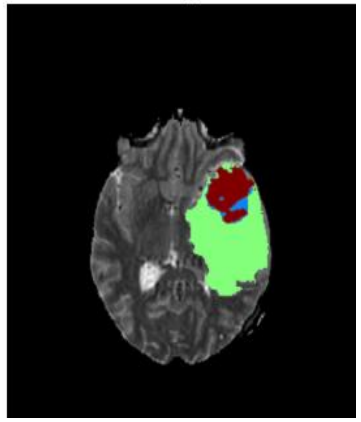
2:(a)



2:(b)



2:(c)



Results:

- BRATS 2017 Validation dataset: 46 patients

Tumor	Dice ET	Dice WT	Dice TC
Mean	61.25%	79.28%	67.34%
Std	30.13%	12.17%	22.15%
Median	74.19%	84.10%	72.91%

- BRATS 2017 test dataset: 146 patients

Tumor	Dice ET	Dice WT	Dice TC
Mean	50.32%	77.01%	61.05%
Std	30.54%	18.71%	29.54%
Median	63.76%	84.00%	74.16%

Conclusion

- Tensor based blind source separation methods has better performance than their matrix based counterpart.
  - Residual water suppression
  - Tissue type differentiation.
- Tensor tools can be used in supervised algorithms
  - Comparable results
  - Needs further exploration.

Thank you!