Brain Lesion Segmentation: A Bayesian Weighted EM Approach

Senan Doyle, Florence Forbes, Michel Dojat

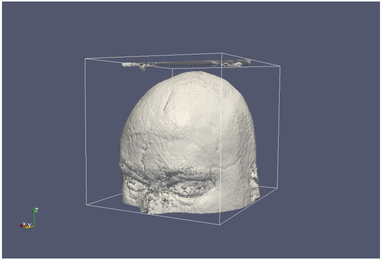
November 19, 2009

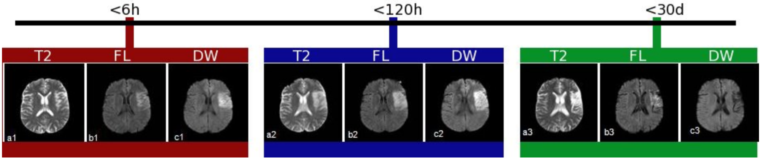
Table of contents

Goal

Delineation & Quantification of Lesions

- Establish Patient Prognosis
- Chart Development over Time

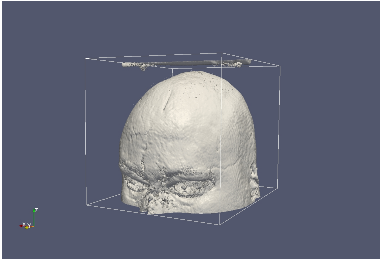


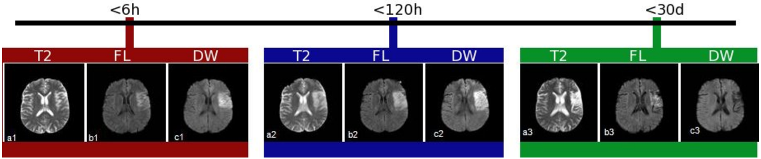


Goal

Delineation & Quantification of Lesions

- Establish Patient Prognosis
- Chart Development over Time
- Automatically!

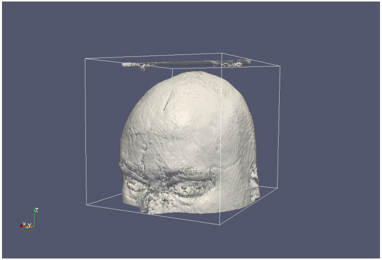


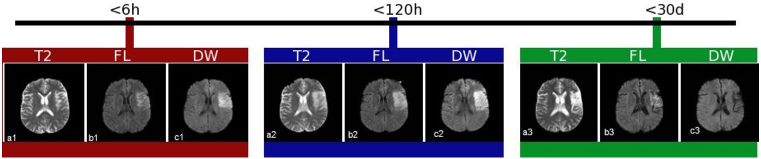


Goal

Delineation & Quantification of Lesions

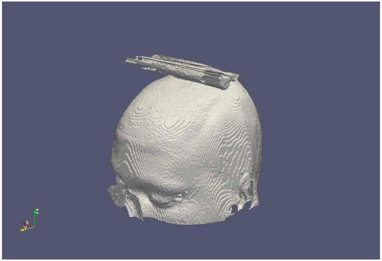
- Establish Patient Prognosis
- Chart Development over Time
- Automatically!

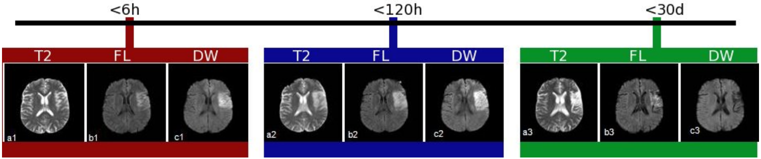




MRI Imaging

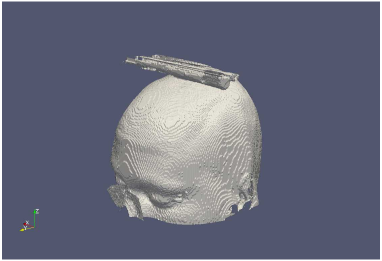
- 3D Volume of *Voxels*
- Different MRI *Sequences* *(T1, T2-w, FLAIR, Diffusion Weight . . .)*

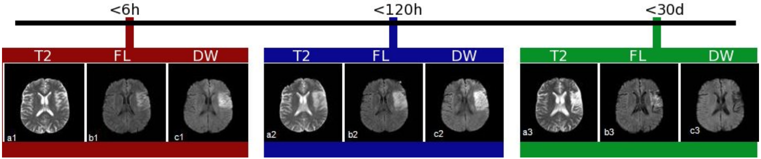




MRI Imaging

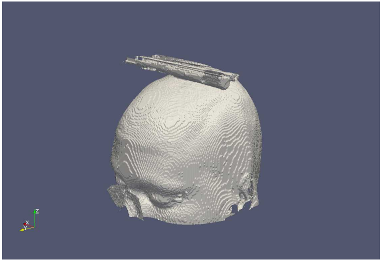
- 3D Volume of *Voxels*
- Different MRI *Sequences (T1, T2-w, FLAIR, Diffusion Weight . . .)*
- Cephalo Spinal Fluid (CSF), White Matter (WM), Grey Matter(GM)

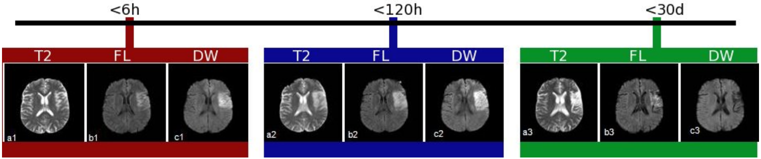




MRI Imaging

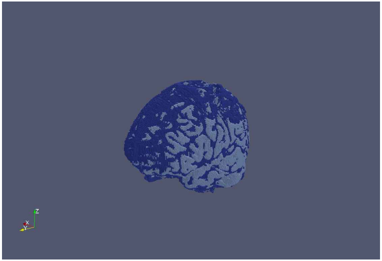
- 3D Volume of *Voxels*
- Different MRI *Sequences* *(T1, T2-w, FLAIR, Diffusion Weight . . .)*
- Cephalo Spinal Fluid (CSF), White Matter (WM), Grey Matter(GM)

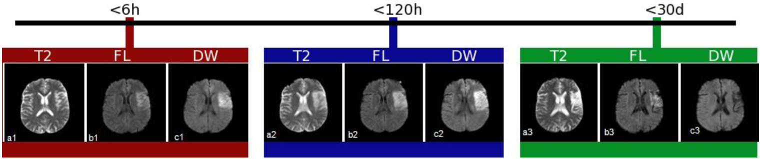




MRI Imaging

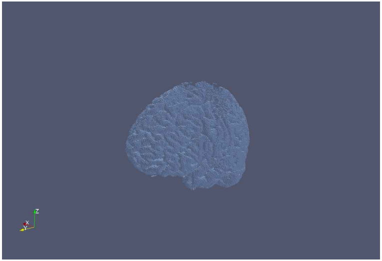
- 3D Volume of *Voxels*
- Different MRI *Sequences (T1, T2-w, FLAIR, Diffusion Weight . . .)*
- Cephalo Spinal Fluid (CSF), White Matter (WM), Grey Matter(GM)

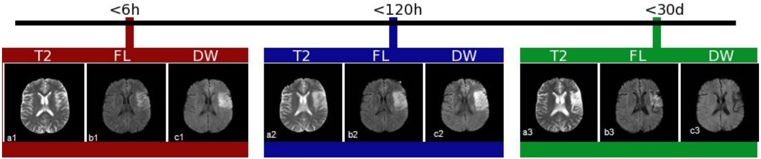




MRI Imaging

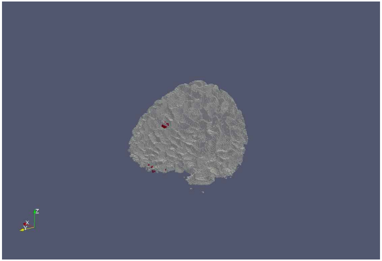
- 3D Volume of *Voxels*
- Different MRI *Sequences (T1, T2-w, FLAIR, Diffusion Weight . . .)*
- Cephalo Spinal Fluid (CSF), White Matter (WM), Grey Matter(GM)

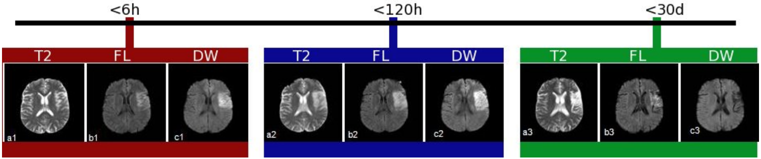




MRI Imaging

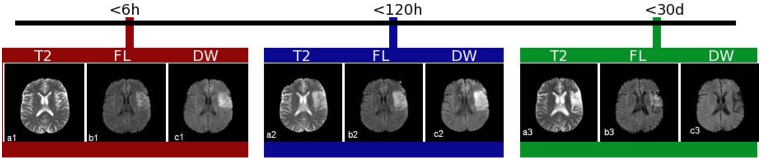
- 3D Volume of *Voxels*
- Different MRI *Sequences (T1, T2-w, FLAIR, Diffusion Weight . . .)*
- Cephalo Spinal Fluid (CSF), White Matter (WM), Grey Matter(GM)





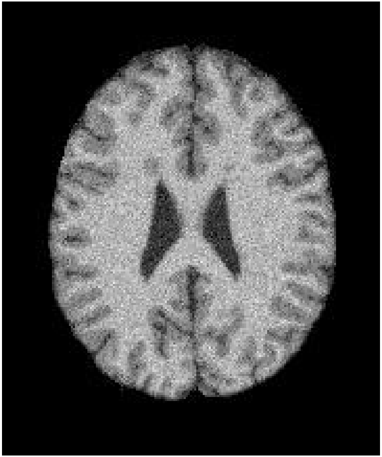
MRI Imaging

- 3D Volume of *Voxels*
- Different MRI *Sequences (T1, T2-w, FLAIR, Diffusion Weight . . .)*
- Cephalo Spinal Fluid (CSF), White Matter (WM), Grey Matter(GM)



Challenges

- Noise
- Intensity Inhomogeneity

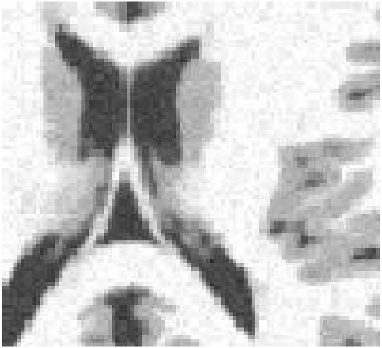


Challenges

- Noise
- Intensity Inhomogeneity
- Partial Volume Effects

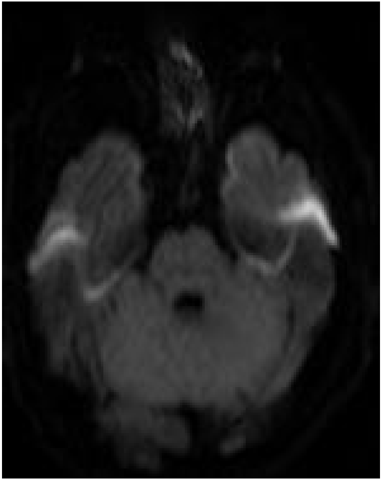
Challenges

- Noise
- Intensity Inhomogeneity
- Partial Volume Effects
- Other Artefacts



Challenges

- Noise
- Intensity Inhomogeneity
- Partial Volume Effects
- Other Artefacts
- Poor Contrast between Tissues



Challenges

- Noise
- Intensity Inhomogeneity
- Partial Volume Effects
- Other Artefacts
- Poor Contrast between Tissues
- Poor Representation of Lesion

Challenges

- Noise
- Intensity Inhomogeneity
- Partial Volume Effects
- Other Artefacts
- Poor Contrast between Tissues
- Poor Representation of Lesion

Inliers Versus Outliers

Lesions usually modelled as outliers.

- Widely Varying & Inhomogeneous Appearance

Inliers Versus Outliers

Lesions usually modelled as outliers.

- Widely Varying & Inhomogeneous Appearance
- Small in Size

Inliers Versus Outliers

Lesions usually modelled as outliers.

- Widely Varying & Inhomogeneous Appearance
- Small in Size

Our approach

Inliers Versus Outliers

Lesions usually modelled as outliers.

- Widely Varying & Inhomogeneous Appearance
- Small in Size

Our approach

- Adopt a Model so that Lesion Voxels become Inliers

Inliers Versus Outliers

Lesions usually modelled as outliers.

- Widely Varying & Inhomogeneous Appearance
- Small in Size

Our approach

- Adopt a Model so that Lesion Voxels become Inliers
- Introduce Weight Field

Inliers Versus Outliers

Lesions usually modelled as outliers.

- Widely Varying & Inhomogeneous Appearance
- Small in Size

Our approach

- Adopt a Model so that Lesion Voxels become Inliers
- Introduce Weight Field
- Bayesian Estimation of Weights

Inliers Versus Outliers

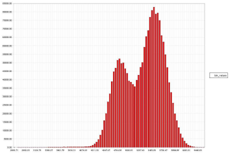
Lesions usually modelled as outliers.

- Widely Varying & Inhomogeneous Appearance
- Small in Size

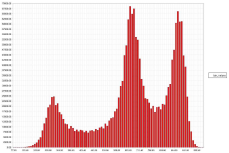
Our approach

- Adopt a Model so that Lesion Voxels become Inliers
- Introduce Weight Field
- Bayesian Estimation of Weights

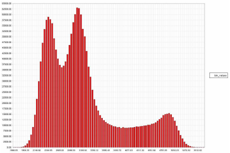
Multi-Sequence



PD



T1



T2

Multi-Sequence

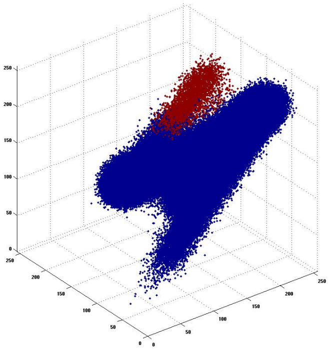


Figure: Feature Space, Lesion in Red

Multi-Sequence

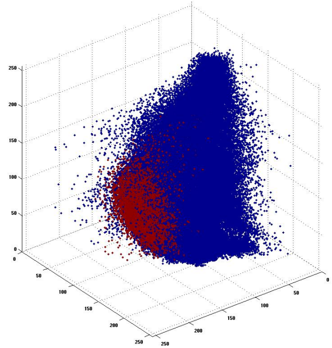


Figure: Feature Space, Lesion in Red

Additional Information

- Spatial Dependency
- Prior Probabilistic Tissue Atlas

Additional Information

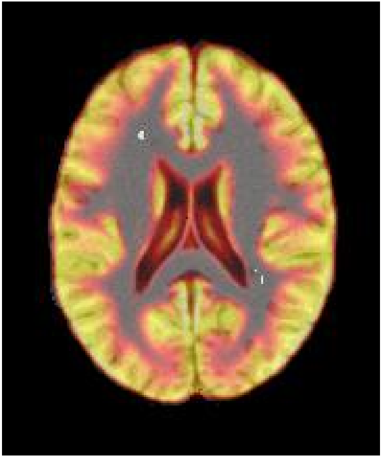
- Spatial Dependency
- Prior Probabilistic Tissue Atlas

Additional Information

- Spatial Dependency
- Prior Probabilistic Tissue Atlas

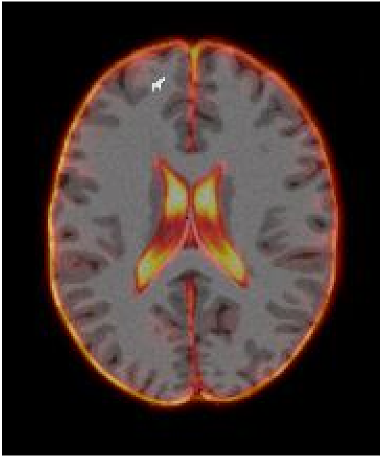
Additional Information

- Spatial Dependency
- Prior Probabilistic Tissue Atlas



Additional Information

- Spatial Dependency
- Prior Probabilistic Tissue Atlas



Definitions

- Regular 3D Grid, $i = 1, \ldots, N, i \in V$
- Observations $\mathbf{y}_i = \{y_{i1}, \ldots, y_{iM}\}$

Definitions

- Regular 3D Grid, $i = 1, \ldots, N, i \in V$
- Observations $\mathbf{y}_i = \{y_{i1}, \ldots, y_{iM}\}$
- Labelling $\mathbf{z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$, $\mathbf{z}_i \in \{1 \ldots K\}$, $\mathbf{z} \in \mathcal{Z}$.

Definitions

- Regular 3D Grid, $i = 1, \ldots, N, i \in V$
- Observations $\mathbf{y}_i = \{y_{i1}, \ldots, y_{iM}\}$
- Labelling $\mathbf{z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$, $\mathbf{z}_i \in \{1 \ldots K\}$, $\mathbf{z} \in \mathcal{Z}$.
- Nonnegative Weights $\omega = \{\omega_i, i \in V\}$, $\omega \in \mathcal{W}$, $\omega_i = \{\omega_{i1}, \ldots, \omega_{iM}\}$.

Definitions

- Regular 3D Grid, $i = 1, \ldots, N, i \in V$
- Observations $\mathbf{y}_i = \{y_{i1}, \ldots, y_{iM}\}$
- Labelling $\mathbf{z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$, $\mathbf{z}_i \in \{1 \ldots K\}$, $\mathbf{z} \in \mathcal{Z}$.
- Nonnegative Weights $\omega = \{\omega_i, i \in V\}$, $\omega \in \mathcal{W}$, $\omega_i = \{\omega_{i1}, \ldots, \omega_{iM}\}$.
- Joint distribution $p(\mathbf{y}, \mathbf{z}, \omega; \psi)$, $\psi = \{\beta, \phi\}$ with

$$H(\mathbf{y}, \mathbf{z}, \omega; \psi) = H_Z(\mathbf{z}; \beta) + H_W(\omega) + \sum_{i \in V} \log g(\mathbf{y}_i | \mathbf{z}_i, \omega_i; \phi)$$

- Missing Data term

Definitions

- Regular 3D Grid, $i = 1, \ldots, N, i \in V$
- Observations $\mathbf{y}_i = \{y_{i1}, \ldots, y_{iM}\}$
- Labelling $\mathbf{z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$, $\mathbf{z}_i \in \{1 \ldots K\}$, $\mathbf{z} \in \mathcal{Z}$.
- Nonnegative Weights $\omega = \{\omega_i, i \in V\}$, $\omega \in \mathcal{W}$, $\omega_i = \{\omega_{i1}, \ldots, \omega_{iM}\}$.
- Joint distribution $p(\mathbf{y}, \mathbf{z}, \omega; \psi)$, $\psi = \{\beta, \phi\}$ with

$$H(\mathbf{y}, \mathbf{z}, \omega; \psi) = \boxed{H_{\mathbf{Z}}(\mathbf{z}; \beta)} + H_W(\omega) + \boxed{\sum_{i \in V} \log g(\mathbf{y}_i | \mathbf{z}_i, \omega_i; \phi)}$$

- Missing Data term
- Parameter Prior term

Definitions

- Regular 3D Grid, $i = 1, \ldots, N, i \in V$
- Observations $\mathbf{y}_i = \{y_{i1}, \ldots, y_{iM}\}$
- Labelling $\mathbf{z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$, $\mathbf{z}_i \in \{1 \ldots K\}$, $\mathbf{z} \in \mathcal{Z}$.
- Nonnegative Weights $\omega = \{\omega_i, i \in V\}$, $\omega \in \mathcal{W}$, $\omega_i = \{\omega_{i1}, \ldots, \omega_{iM}\}$.
- Joint distribution $p(\mathbf{y}, \mathbf{z}, \omega; \psi)$, $\psi = \{\beta, \phi\}$ with

$$H(\mathbf{y}, \mathbf{z}, \omega; \psi) = H_{\mathbf{Z}}(\mathbf{z}; \beta) + H_W(\omega) + \sum_{i \in V} \log g(\mathbf{y}_i | \mathbf{z}_i, \omega_i; \phi)$$

- Missing Data term
- Parameter Prior term
- Data term

Definitions

- Regular 3D Grid, $i = 1, \ldots, N, i \in V$
- Observations $\mathbf{y}_i = \{y_{i1}, \ldots, y_{iM}\}$
- Labelling $\mathbf{z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$, $\mathbf{z}_i \in \{1 \ldots K\}$, $\mathbf{z} \in \mathcal{Z}$.
- Nonnegative Weights $\omega = \{\omega_i, i \in V\}$, $\omega \in \mathcal{W}$, $\omega_i = \{\omega_{i1}, \ldots, \omega_{iM}\}$.
- Joint distribution $p(\mathbf{y}, \mathbf{z}, \omega; \psi)$, $\psi = \{\beta, \phi\}$ with

$$H(\mathbf{y}, \mathbf{z}, \omega; \psi) = \boxed{H_{\mathbf{Z}}(\mathbf{z}; \beta)} + H_W(\omega) + \boxed{\sum_{i \in V} \log g(\mathbf{y}_i | \mathbf{z}_i, \omega_i; \phi)}$$

- Missing Data term
- Parameter Prior term
- Data term

Data term

- M-Dimensional Gaussians, Diagonal Covariance
- $$g(\mathbf{y}_i|\mathbf{z}_i, \omega_i; \phi) = \prod_{m=1}^{M} \mathcal{G}(y_{im}; \mu_{\mathbf{z}_i m}, \frac{s_{\mathbf{z}_i m}}{\omega_{im}})$$

Data term

- M-Dimensional Gaussians, Diagonal Covariance
- $$g(\mathbf{y}_i|\mathbf{z}_i, \omega_i; \phi) = \prod_{m=1}^{M} \mathcal{G}(y_{im}; \mu_{\mathbf{z}_i m}, \frac{s_{\mathbf{z}_i m}}{\omega_{im}})$$
- Related to *Logarithmic Pooling*

$$\left(\propto \prod_{m=1}^{M} \mathcal{G}(y_{im}; \mu_{\mathbf{z}_i m}, s_{\mathbf{z}_i m})^{\omega_{im}} . \right)$$

Data term

- M-Dimensional Gaussians, Diagonal Covariance
- $$g(\mathbf{y}_i|\mathbf{z}_i, \omega_i; \phi) = \prod_{m=1}^{M} \mathcal{G}(y_{im}; \mu_{\mathbf{z}_i m}, \frac{s_{\mathbf{z}_i m}}{\omega_{im}})$$
- Related to *Logarithmic Pooling*

$$\left(\propto \prod_{m=1}^{M} \mathcal{G}(y_{im}; \mu_{\mathbf{z}_i m}, s_{\mathbf{z}_i m})^{\omega_{im}} \; . \right)$$

Missing Data term

- Potts Model, external field ξ, interaction parameter η
- $$H_{\mathbf{Z}}(\mathbf{z};\beta) = \sum_{i \in V} \Big(\xi_{i\mathbf{z}_i} + \sum_{j \in \mathcal{N}(i)} \eta \, \langle \mathbf{z}_i, \mathbf{z}_j \rangle \Big)$$

Missing Data term

- Potts Model, external field ξ, interaction parameter η
- $$H_{\mathbf{Z}}(\mathbf{z};\beta) = \sum_{i \in V} (\xi_{i\mathbf{z}_i} + \sum_{j \in \mathcal{N}(i)} \eta \, \langle \mathbf{z}_i, \mathbf{z}_j \rangle)$$
- $\beta = \{\xi, \eta\}$ with $\xi = \{{}^t(\xi_{i1} \ldots \xi_{iK}), i \in V\}$ being a set of real-valued K-dimensional vectors and η a real positive value.

Missing Data term

- Potts Model, external field ξ, interaction parameter η
- $$H_{\mathbf{Z}}(\mathbf{z};\beta) = \sum_{i \in V} (\xi_{i\mathbf{z}_i} + \sum_{j \in \mathcal{N}(i)} \eta \, \langle \mathbf{z}_i, \mathbf{z}_j \rangle)$$
- $\beta = \{\xi, \eta\}$ with $\xi = \{{}^t(\xi_{i1} \ldots \xi_{iK}), i \in V\}$ being a set of real-valued K-dimensional vectors and η a real positive value.

Parameter Prior term

- Weights independent of ψ, across Sequences
- Define Conjugate Prior $p(\omega) = \prod_{m=1}^{M} \prod_{i \in V} p(\omega_{im})$

Parameter Prior term

- Weights independent of ψ, across Sequences
- Define Conjugate Prior $p(\omega) = \prod_{m=1}^{M} \prod_{i \in V} p(\omega_{im})$
- $p(\omega_{im})$ is a Gamma distribution with hyperparameters α_{im} (shape) and γ_{im} (inverse scale)

$$H_W(\omega) = \sum_{m=1}^{M} \sum_{i \in V} ((\alpha_{im} - 1) \log \omega_{im} - \gamma_{im}\, \omega_{im}).$$

Parameter Prior term

- Weights independent of ψ, across Sequences
- Define Conjugate Prior $p(\omega) = \prod_{m=1}^{M} \prod_{i \in V} p(\omega_{im})$
- $p(\omega_{im})$ is a Gamma distribution with hyperparameters α_{im} (shape) and γ_{im} (inverse scale)

$$H_W(\omega) = \sum_{m=1}^{M} \sum_{i \in V} ((\alpha_{im} - 1) \log \omega_{im} - \gamma_{im}\, \omega_{im}).$$

- Modes fixed to *expert weights* $\{\omega_{im}^{exp}, m = 1 \ldots M, i \in V\}$ by setting $\alpha_{im} = \gamma_{im}\, \omega_{im}^{exp} + 1$

Parameter Prior term

- Weights independent of ψ, across Sequences
- Define Conjugate Prior $p(\omega) = \prod_{m=1}^{M} \prod_{i \in V} p(\omega_{im})$
- $p(\omega_{im})$ is a Gamma distribution with hyperparameters α_{im} (shape) and γ_{im} (inverse scale)

$$H_W(\omega) = \sum_{m=1}^{M} \sum_{i \in V} ((\alpha_{im} - 1) \log \omega_{im} - \gamma_{im}\, \omega_{im}).$$

- Modes fixed to *expert weights* $\{\omega_{im}^{exp}, m = 1 \ldots M, i \in V\}$ by setting $\alpha_{im} = \gamma_{im}\, \omega_{im}^{exp} + 1$

Alternating Maximization

- Weights viewed as Missing Variables
- Let $\mathcal{D}$ be the set of all probability distributions on $\mathcal{Z} \times \mathcal{W}$.

Alternating Maximization

- Weights viewed as Missing Variables
- Let $\mathcal{D}$ be the set of all probability distributions on $\mathcal{Z} \times \mathcal{W}$.
- Alternating maximization of a function F s.t. for any $q \in \mathcal{D}$,

$$F(q, \psi) = E_q[\log p(\mathbf{y}, \mathbf{Z}, W \; ; \; \psi)] - E_q[\log q(\mathbf{Z}, W)]$$

Alternating Maximization

- Weights viewed as Missing Variables
- Let $\mathcal{D}$ be the set of all probability distributions on $\mathcal{Z} \times \mathcal{W}$.
- Alternating maximization of a function F s.t. for any $q \in \mathcal{D}$,

$$F(q, \psi) = E_q[\log p(\mathbf{y}, \mathbf{Z}, W \; ; \; \psi)] - E_q[\log q(\mathbf{Z}, W)]$$

-

$$\textbf{E-step: } q^{(r)} = \arg\max_{q \in \mathcal{D}} F(q, \psi^{(r)}) \tag{1}$$

$$\textbf{M-step: } \psi^{(r+1)} = \arg\max_{\psi \in \underline{\Psi}} F(q^{(r)}, \psi)$$

Alternating Maximization

- Weights viewed as Missing Variables
- Let $\mathcal{D}$ be the set of all probability distributions on $\mathcal{Z} \times \mathcal{W}$.
- Alternating maximization of a function F s.t. for any $q \in \mathcal{D}$,

$$F(q, \psi) = \boxed{E_q[\log p(\mathbf{y}, \mathbf{Z}, W\ ;\ \psi)]} - E_q[\log q(\mathbf{Z}, W)]$$

-

$$\textbf{E-step:}\ q^{(r)} = \arg\max_{q \in \mathcal{D}} F(q, \psi^{(r)}) \tag{1}$$

$$\textbf{M-step:}\ \psi^{(r+1)} = \arg\max_{\psi \in \underline{\Psi}} F(q^{(r)}, \psi)$$

Alternating Maximization

- Weights viewed as Missing Variables
- Let $\mathcal{D}$ be the set of all probability distributions on $\mathcal{Z} \times \mathcal{W}$.
- Alternating maximization of a function F s.t. for any $q \in \mathcal{D}$,

$$F(q, \psi) = E_q[\log p(\mathbf{y}, \mathbf{Z}, W \; ; \; \psi)] - E_q[\log q(\mathbf{Z}, W)]$$

-

$$\textbf{E-step:} \; q^{(r)} = \arg\max_{q \in \mathcal{D}} \; F(q, \psi^{(r)}) \tag{1}$$

$$\textbf{M-step:} \; \psi^{(r+1)} = \arg\max_{\psi \in \underline{\Psi}} \; F(q^{(r)}, \psi)$$

Alternating Maximization

- E-step solved over restricted class of probability distributions, $\tilde{\mathcal{D}}$, s.t. $q(\mathbf{z}, \omega) = q_Z(\mathbf{z})\, q_W(\omega)$ where $q_Z \in \mathcal{D}_Z$ and $q_W \in \mathcal{D}_W$
-

$$\textbf{E-Z-step:} \quad q_Z^{(r)} = \arg\max_{q_Z \in \mathcal{D}_Z} F(q_W^{(r-1)}\, q_Z; \psi^{(r)})$$

$$\textbf{E-W-step:} \quad q_W^{(r)} = \arg\max_{q_W \in \mathcal{D}_W} F(q_W\, q_Z^{(r)}; \psi^{(r)}) \, .$$

Alternating Maximization

- E-step solved over restricted class of probability distributions, $\tilde{\mathcal{D}}$, s.t. $q(\mathbf{z}, \omega) = q_Z(\mathbf{z})\, q_W(\omega)$ where $q_Z \in \mathcal{D}_Z$ and $q_W \in \mathcal{D}_W$
-

$$\textbf{E-Z-step:} \quad q_Z^{(r)} = \arg\max_{q_Z \in \mathcal{D}_Z} F(q_W^{(r-1)}\, q_Z; \psi^{(r)})$$

$$\textbf{E-W-step:} \quad q_W^{(r)} = \arg\max_{q_W \in \mathcal{D}_W} F(q_W\, q_Z^{(r)}; \psi^{(r)})\ .$$

Alternating Maximization

- Use K-L divergence properties to avoid taking derivatives
-

$$
\begin{aligned}
\textbf{E-Z:}\quad q_Z^{(r)} &\propto \exp\left(E_{q_W^{(r-1)}}[\log p(\mathbf{z}|\mathbf{y},\mathbf{W};\psi^{(r)})]\right) &(2)\\
\textbf{E-W:}\quad q_W^{(r)} &\propto \exp\left(E_{q_Z^{(r)}}[\log p(\omega|\mathbf{y},\mathbf{Z};\psi^{(r)})]\right) &(3)\\
\textbf{M:}\quad \psi^{(r+1)} &= \arg\max_{\psi\in\underline{\Psi}} E_{q_Z^{(r)} q_W^{(r)}}[\log p(\mathbf{y},\mathbf{Z},W;\psi)]\,. &(4)
\end{aligned}
$$

Alternating Maximization

- Use K-L divergence properties to avoid taking derivatives
-

$$
\begin{aligned}
\textbf{E-Z:}\;\; q_Z^{(r)} &\propto \exp\left(E_{q_W^{(r-1)}}[\log p(\mathbf{z}|\mathbf{y},\mathbf{W};\psi^{(r)}]\right) && (2)\\
\textbf{E-W:}\;\; q_W^{(r)} &\propto \exp\left(E_{q_Z^{(r)}}[\log p(\omega|\mathbf{y},\mathbf{Z};\psi^{(r)})]\right) && (3)\\
\textbf{M:}\;\; \psi^{(r+1)} &= \arg\max_{\psi\in\underline{\boldsymbol{\Psi}}} E_{q_Z^{(r)}q_W^{(r)}}[\log p(\mathbf{y},\mathbf{Z},W;\psi)]\,. && (4)
\end{aligned}
$$

E-Z

- Note $p(\mathbf{z}|\mathbf{y}, \omega; \psi)$ is Markovian with energy

$$H(\mathbf{z}|\mathbf{y}, \omega; \psi) \propto H_{\mathbf{Z}}(\mathbf{z}; \beta) + \sum_{i \in V} \log g(\mathbf{y}_i|\mathbf{z}_i, \omega_i; \phi), (\text{linear in } \omega) \tag{5}$$

- Therefore (2) becomes

$$\textbf{E-Z:} \quad q_Z^{(r)} \quad \propto \quad p(\mathbf{z}|\mathbf{y}, E_{q_W^{(r-1)}}[W]; \psi^{(r)}). \tag{6}$$

E-Z

- Note $p(\mathbf{z}|\mathbf{y}, \omega; \psi)$ is Markovian with energy

$$H(\mathbf{z}|\mathbf{y}, \omega; \psi) \propto H_{\mathbf{Z}}(\mathbf{z}; \beta) + \sum_{i \in V} \log g(\mathbf{y}_i|\mathbf{z}_i, \omega_i; \phi), \text{(linear in } \omega) \tag{5}$$

- Therefore (2) becomes

$$\textbf{E-Z:} \quad q_Z^{(r)} \quad \propto \quad p(\mathbf{z}|\mathbf{y}, E_{q_W^{(r-1)}}[W]; \psi^{(r)}). \tag{6}$$

- Mean-Field like algorithm provides MRF approximation and

$$\tilde{q}_Z^{(r)}(\mathbf{z}) \propto \prod_{i \in V} \prod_{m=1}^{M} \mathcal{G}(y_{im}; \mu_{\mathbf{z}_i m}^{(r)}, \frac{s_{\mathbf{z}_i m}^{(r)}}{\bar{\omega}_{im}^{(r-1)}}) p(\mathbf{z}_i|\tilde{\mathbf{z}}_{\mathcal{N}(i)}^{(r)}; \beta^{(r)}), \tag{7}$$

E-Z

- Note $p(\mathbf{z}|\mathbf{y},\omega;\psi)$ is Markovian with energy

$$H(\mathbf{z}|\mathbf{y},\omega;\psi) \propto H_{\mathbf{Z}}(\mathbf{z};\beta) + \sum_{i \in V} \log g(\mathbf{y}_i|\mathbf{z}_i,\omega_i;\phi), (\text{linear in } \omega) \tag{5}$$

- Therefore (2) becomes

$$\textbf{E-Z:} \quad q_Z^{(r)} \quad \propto \quad p(\mathbf{z}|\mathbf{y}, E_{q_W^{(r-1)}}[W];\psi^{(r)}). \tag{6}$$

- Mean-Field like algorithm provides MRF approximation and

$$\tilde{q}_Z^{(r)}(\mathbf{z}) \propto \prod_{i \in V} \prod_{m=1}^{M} \mathcal{G}(y_{im};\mu_{\mathbf{z}_i m}^{(r)}, \frac{s_{\mathbf{z}_i m}^{(r)}}{\bar{\omega}_{im}^{(r-1)}}) p(\mathbf{z}_i|\tilde{\mathbf{z}}_{\mathcal{N}(i)}^{(r)};\beta^{(r)}), \tag{7}$$

E-W

- Similarly $p(\omega|\mathbf{y},\mathbf{z};\psi)$ is Markovian with energy

$$H(\omega|\mathbf{y},\mathbf{z};\psi) \propto H_W(\omega) + \sum_{i \in V} \log g(\mathbf{y}_i|\mathbf{z}_i,\omega_i;\phi) \qquad (8)$$

- $E_{q_W^{(r)}}[W_{im}]$ denoted by $\bar{\omega}_{im}^{(r)}$ becomes

$$\bar{\omega}_{im}^{(r)} = \frac{\alpha_{im} + \frac{1}{2}}{\gamma_{im} + \frac{1}{2} \sum_{k=1}^{K} \delta(y_{im}, \mu_{km}^{(r)}, s_{km}^{(r)})\, q_{Z_i}^{(r)}(k)} \qquad (9)$$

where $\delta(y,\mu,s) = (y-\mu)^2/s$ is the squared Mahalanobis distance.

E-W

- Similarly $p(\omega|\mathbf{y}, \mathbf{z}; \psi)$ is Markovian with energy

$$H(\omega|\mathbf{y}, \mathbf{z}; \psi) \propto H_W(\omega) + \sum_{i \in V} \log g(\mathbf{y}_i|\mathbf{z}_i, \omega_i; \phi) \qquad (8)$$

- $E_{q_W^{(r)}}[W_{im}]$ denoted by $\bar{\omega}_{im}^{(r)}$ becomes

$$\bar{\omega}_{im}^{(r)} = \frac{\alpha_{im} + \frac{1}{2}}{\gamma_{im} + \frac{1}{2} \sum_{k=1}^{K} \delta(y_{im}, \mu_{km}^{(r)}, s_{km}^{(r)}) \; q_{Z_i}^{(r)}(k)} \qquad (9)$$

where $\delta(y, \mu, s) = (y - \mu)^2 / s$ is the squared Mahalanobis distance.

M

- β solved using mean field approximation
- ϕ updated using

$$\mu_{km}^{(r+1)} = \frac{\sum_{i=1}^{N} q_{Z_i}^{(r)}(k)\, \bar{\omega}_{im}^{(r)}\, y_{im}}{\sum_{i=1}^{N} q_{Z_i}^{(r)}(k)\, \bar{\omega}_{im}^{(r)}},$$

$$s_{km}^{(r+1)} = \frac{\sum_{i=1}^{N} q_{Z_i}^{(r)}(k)\, \bar{\omega}_{im}^{(r)}\, (y_{im} - \mu_{km}^{(r+1)})^2}{\sum_{i=1}^{N} q_{Z_i}^{(r)}(k)},$$

M

- β solved using mean field approximation
- ϕ updated using

$$\mu_{km}^{(r+1)} = \frac{\sum_{i=1}^{N} q_{Z_i}^{(r)}(k)\, \bar{\omega}_{im}^{(r)}\, y_{im}}{\sum_{i=1}^{N} q_{Z_i}^{(r)}(k)\, \bar{\omega}_{im}^{(r)}},$$

$$s_{km}^{(r+1)} = \frac{\sum_{i=1}^{N} q_{Z_i}^{(r)}(k)\, \bar{\omega}_{im}^{(r)}\, (y_{im} - \mu_{km}^{(r+1)})^2}{\sum_{i=1}^{N} q_{Z_i}^{(r)}(k)},$$

Expert Weighting

- Region $\mathcal{L}$ defines *candidate lesions*
- $\mathcal{L}$ weighted with $\omega_{\mathcal{L}} > 1$ ($\omega_{\setminus\mathcal{L}} = 1$)

Expert Weighting

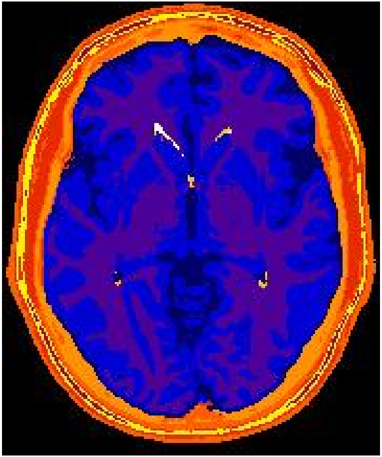
- Region $\mathcal{L}$ defines *candidate lesions*
- $\mathcal{L}$ weighted with $\omega_{\mathcal{L}} > 1$ ($\omega_{\backslash\mathcal{L}} = 1$)
- $K = 3$, $\omega_{im}^{exp} = 1$ and $\gamma_{im} = 1$ setting

Expert Weighting

- Region $\mathcal{L}$ defines *candidate lesions*
- $\mathcal{L}$ weighted with $\omega_{\mathcal{L}} > 1$ ($\omega_{\setminus\mathcal{L}} = 1$)
- $K = 3$, $\omega_{im}^{exp} = 1$ and $\gamma_{im} = 1$ setting
- Threshold Most Informative Weight Map (chi-square percentile)

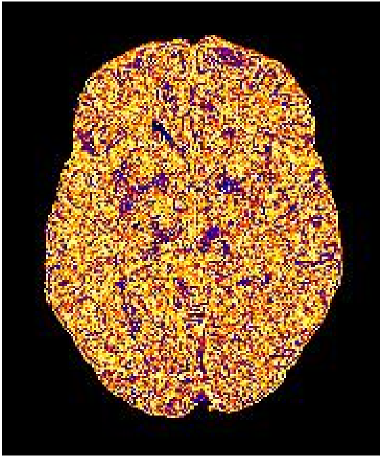
Expert Weighting

- Region $\mathcal{L}$ defines *candidate lesions*
- $\mathcal{L}$ weighted with $\omega_{\mathcal{L}} > 1$ ($\omega_{\setminus\mathcal{L}} = 1$)
- $K = 3$, $\omega_{im}^{exp} = 1$ and $\gamma_{im} = 1$ setting
- Threshold Most Informative Weight Map (chi-square percentile)



Expert Weighting

- Region $\mathcal{L}$ defines *candidate lesions*
- $\mathcal{L}$ weighted with $\omega_{\mathcal{L}} > 1$ $(\omega_{\setminus\mathcal{L}} = 1)$
- $K = 3$, $\omega_{im}^{exp} = 1$ and $\gamma_{im} = 1$ setting
- Threshold Most Informative Weight Map (chi-square percentile)



Inlier Setting

- $K = 4$, initialized with $K = 3$ result (+ thresholded lesions)
- $\gamma_{im} = \gamma_{\mathcal{L}} \forall i \in \mathcal{L}$ ($\gamma_{\mathcal{L}} = 10$)

Inlier Setting

- $K = 4$, initialized with $K = 3$ result (+ thresholded lesions)
- $\gamma_{im} = \gamma_{\mathcal{L}} \forall i \in \mathcal{L}$ ($\gamma_{\mathcal{L}} = 10$)
- $\gamma_{im} = \gamma_{\bar{\mathcal{L}}} \forall i \notin \mathcal{L}$. ($\gamma_{\bar{\mathcal{L}}} = 1000$)

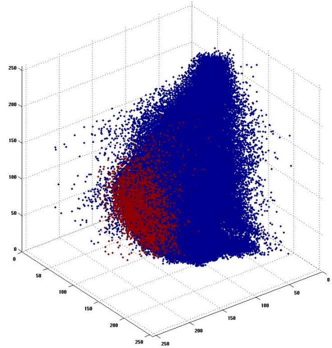
Inlier Setting

- $K = 4$, initialized with $K = 3$ result ($+$ thresholded lesions)
- $\gamma_{im} = \gamma_{\mathcal{L}} \forall i \in \mathcal{L}$ ($\gamma_{\mathcal{L}} = 10$)
- $\gamma_{im} = \gamma_{\bar{\mathcal{L}}} \forall i \notin \mathcal{L}$. ($\gamma_{\bar{\mathcal{L}}} = 1000$)
- $\omega_{\mathcal{L}}$ depends on number of voxels in $\mathcal{L}$

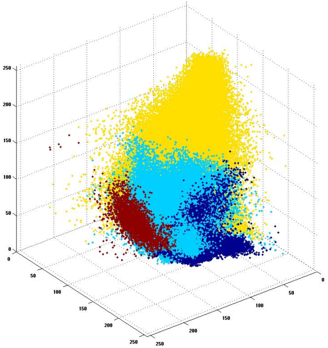
Inlier Setting

- $K = 4$, initialized with $K = 3$ result ($+$ thresholded lesions)
- $\gamma_{im} = \gamma_{\mathcal{L}} \forall i \in \mathcal{L}$ ($\gamma_{\mathcal{L}} = 10$)
- $\gamma_{im} = \gamma_{\bar{\mathcal{L}}} \forall i \notin \mathcal{L}$. ($\gamma_{\bar{\mathcal{L}}} = 1000$)
- $\omega_{\mathcal{L}}$ depends on number of voxels in $\mathcal{L}$

Inlier Setting



Inlier Setting



Simulated Data 0% IIH

Method	3%	5%	7%	9%
Mild lesions (0.02% of the voxels)				
AWEM	**68 (+1)**	49 (-21)	**36 (+2)**	**12 (+8)**
[1]	67	**70**	34	0
[3]	56	33	13	4
[2]	52	NA	NA	NA
Moderate lesions (0.18% of the voxels)				
AWEM	**86 (+7)**	80 (-1)	**73 (+14)**	**64 (+27)**
[1]	72	**81**	59	29
[3]	79	69	52	37
[2]	63	NA	NA	NA
Severe lesions (0.52% of the voxels)				
AWEM	**92 (+7)**	86 (-2)	**78 (+6)**	**68 (+27)**
[1]	79	**88**	72	41
[3]	85	72	56	41
[2]	82	NA	NA	NA

Simulated Data 40% IIH

Method	3%	5%	7%	9%
Mild lesions (0.02% of the voxels)				
AWEM	0 (-75)	0 (-65)	0 (-20)	0 (-30)
[1]	**75**	**65**	**20**	**30**
[3]	58	27	13	6
Moderate lesions (0.18% of the voxels)				
AWEM	52 (-24)	51 (-25)	52 (-15)	10 (-38)
[1]	75	**76**	**67**	**48**
[3]	**76**	64	47	31
Severe lesions (0.52% of the voxels)				
AWEM	**87 (+1)**	**84 (+1)**	**77 (+3)**	**66 (+8)**
[1]	75	83	74	58
[3]	86	74	62	45

Real Data (Rennes MS)

	LL	EMS	AWEM
Patient1	0.42	62	**82 (+20)**
Patient2	1.71	54	**56 (+2)**
Patient3	0.29	**47**	45 (-2)
Patient4	1.59	65	**72 (+7)**
Patient5	0.31	**47**	45 (-2)
Average		55 +/-8	60 +/-16

Results

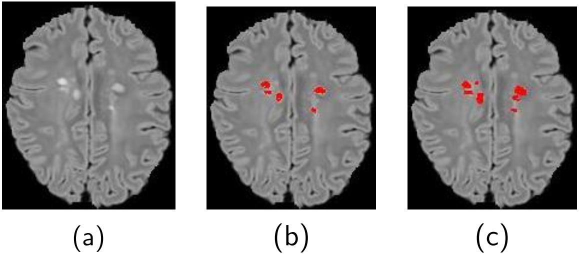
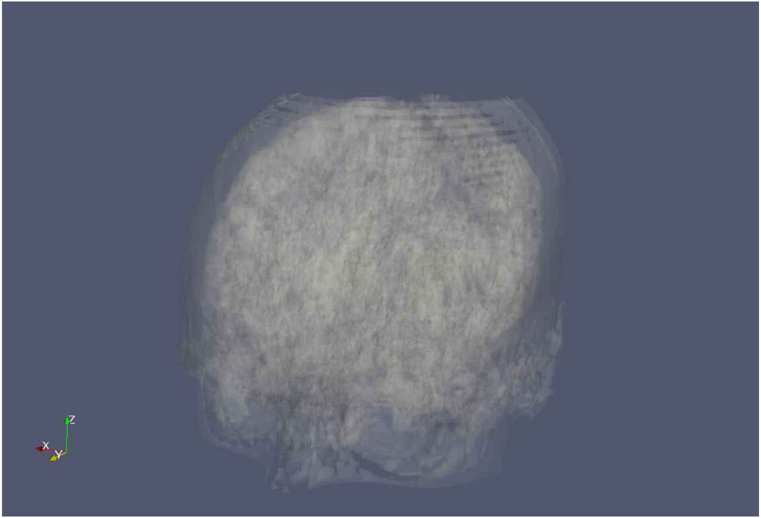
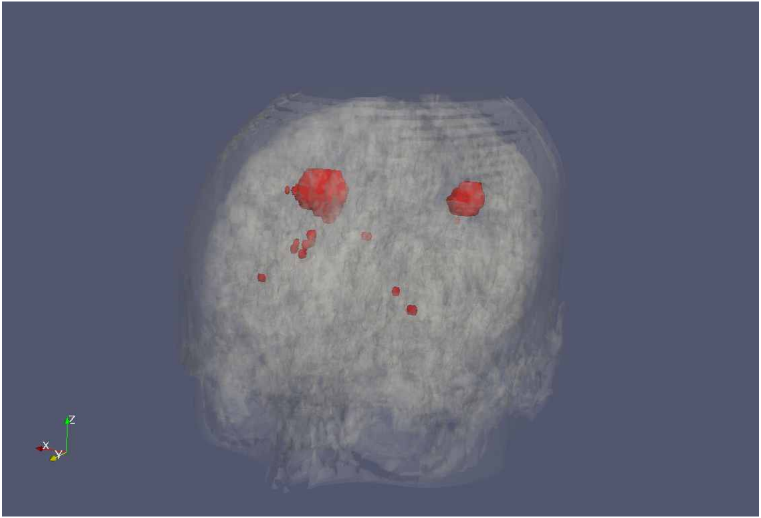


Figure: Real MS data, patient 3. (a): Flair image. (b): identified lesions with our approach (DSC 45%). (c): ground truth .

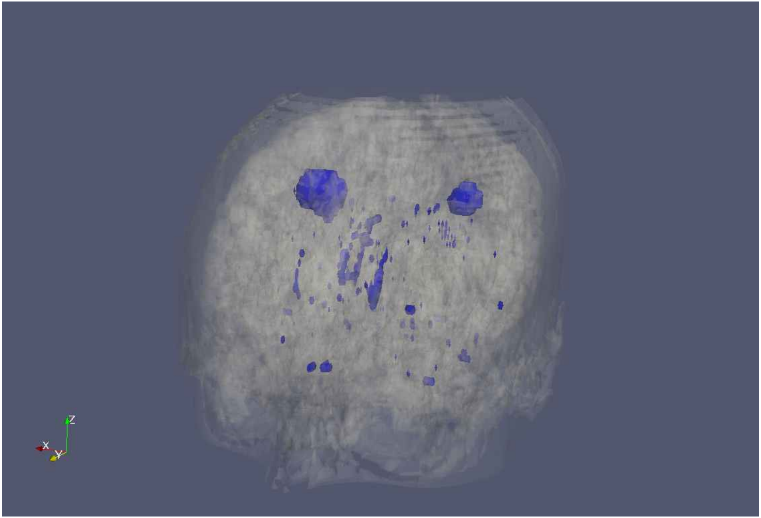
Results



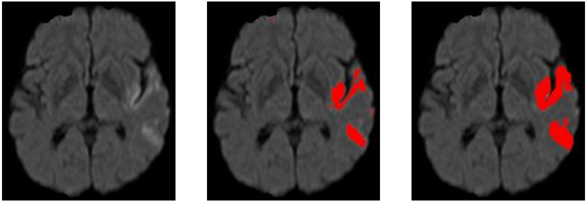
Results



Results



Results



(a) (b) (c)

Figure: Real stroke data. (a): DW image. (b): identified lesions with our approach (DSC 63%). (c): ground truth.

Discussion & Future Work

- Extension to full covariance matrices: temporal multi-sequence data, eg. patient follow-up
- Exploration of Markov Prior
- Other expert weighting schemes, possibly lesion specific
- Extension to handle intensity inhomogeneities
- Sensitivity analysis: initialization, parameter tuning etc. (Darren)
- Evaluation in a semi-supervised context

D. Garcia-Lorenzo, L. Lecoeur, D.L. Arnold, D. L. Collins, and C. Barillot.
Multiple Sclerosis lesion segmentation using an automatic multimodal graph cuts.
In *MICCAI*, pages 584–591, 2009.

F. Rousseau, F. Blanc, J. de Seze, L. Rumbac, and J.P. Armspach.
An a contrario approach for outliers segmentation: application to multiple sclerosis in MRI.
In *IEEE ISBI*, pages 9–12, 2008.

K. Van Leemput, F. Maes, D. Vandermeulen, A. Colchester, and P. Suetens.
Automated segmentation of multiple sclerosis lesions by model outlier detection.
IEEE trans. Med. Ima., 20(8):677–688, 2001.