# ECSE-626

# Statistical Computer Vision

Conditional Random Fields

# Classification/Labeling Problem

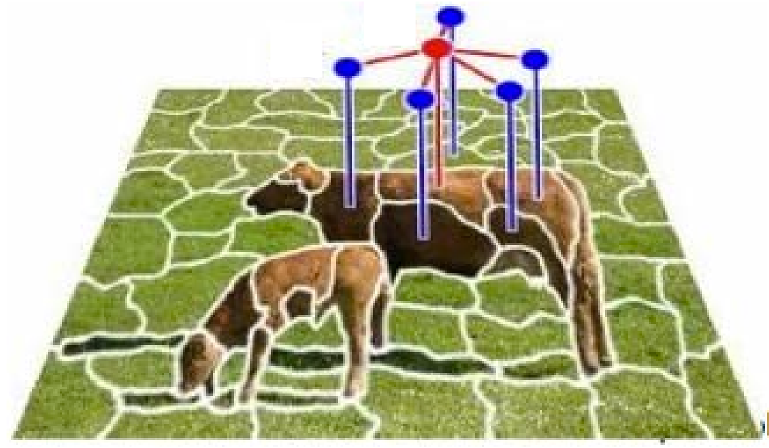
- Classification - Image segmentation
- Graphical models
- Generative Vs. Discriminative
- Markov random field (MRF)
- Conditional random field (CRF)
- Some examples

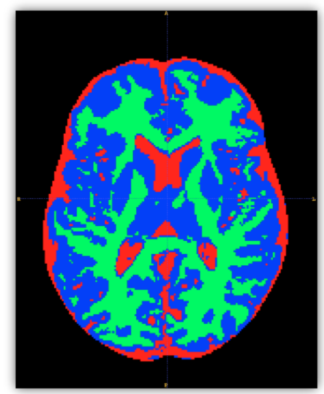
# Classification/Labeling Problem

- Given a set of input variables (features) denoted as X
- Estimate the probability of the output variable denoted by Y → P(Y|X)
- Where Y can take values from the finite set of possible classes: {0,1,…L}

# Image Segmentation/Classification

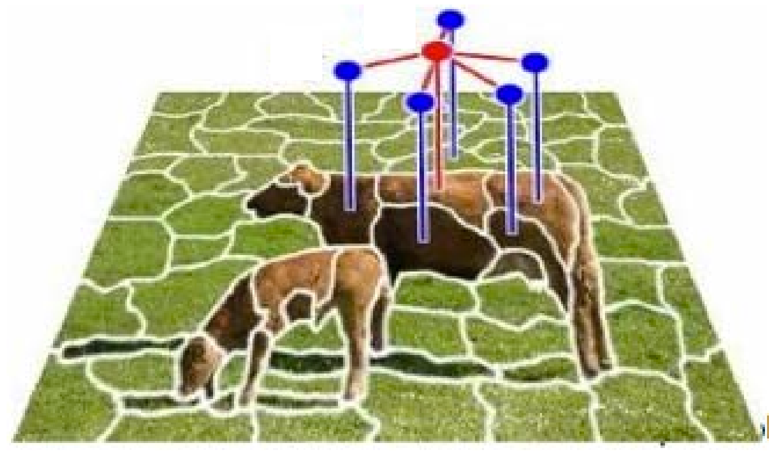
- The purpose of image segmentation is to partition an image into *meaningful* regions with respect to a particular application.
- Formally:
  - To assign each pixel/super pixel/region in the image a label from a set of possible labels

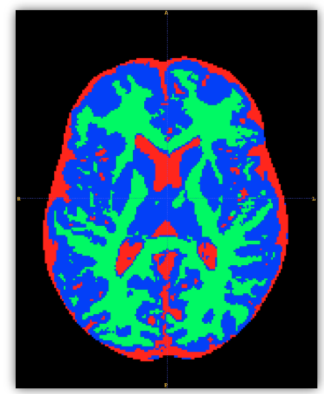




# Image Segmentation/Classification

- Let $x_i$ denote the observation (e.g. intensity at each pixel, extracted features)
- Let $y_i$ denote the label assigned to it.
- Let's define the *posterior probability density function* on the set of all possible labelings: $p(Y|X)$

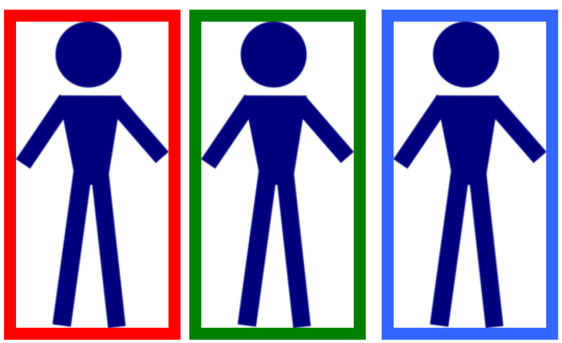




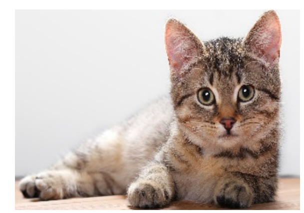
# Single Variable? Random field?

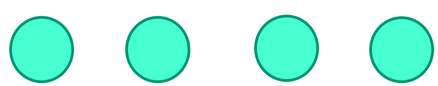
- Mainly depends on the nature of the problem:
- Examples of single variable:

Patient – Disease classification
Input – Symptom (cough)



Object classification
(for 1 detected object)





# Single Variable? Random field?

- Examples of dependent variables:

Part of speech tagging

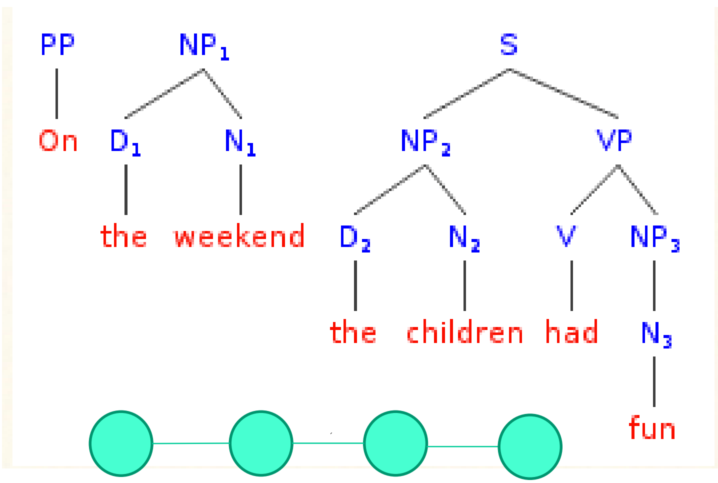
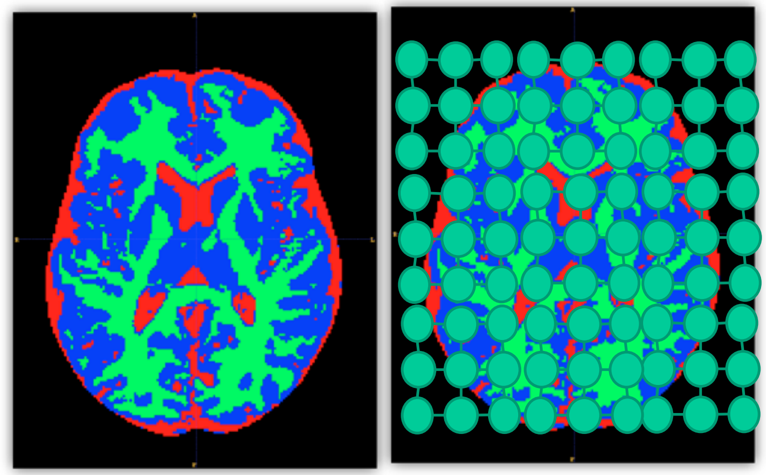



Image segmentation



When there is dependency between observation and/or labels, it is important to model these interaction instead of neglecting them.

# Random Fields

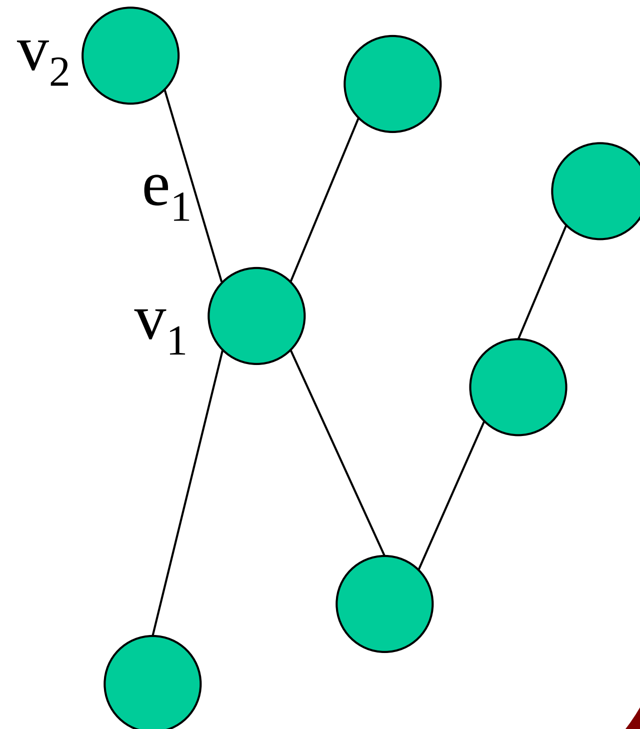
Probabilistic Graphical Models model joint probability of a set of random variables that are interdependent

In a graphical model G(V;E), random variables are shown with a set of nodes:

$$V= \{v_1,v_2,..,v_n\}$$

Their dependencies are presented by a set of edges:

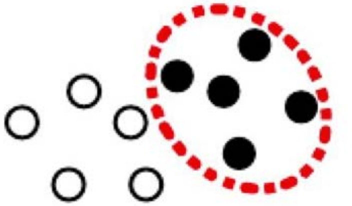
$$E=\{e_1,e_2,...,e_n\}$$



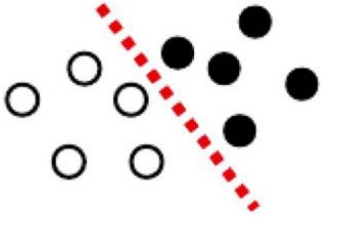
# Generative vs. Discriminative

X = observations Y = Labels

- Generative models:
  - Probabilistic model for each class
  - Good for when you can model the likelihood, prior probabilities



- Discriminative models:
  - Focus on the decision boundary
  - No attempt to model the underlying probability distributions



# Generative vs. Discriminative

X = observations Y = Labels

- Generative models:
  - Gaussians,
  - MRFs,
  - Naïve Bayes
- Discriminative models:
  - Logistic regression
  - SVM

# Generative vs. Discriminative

X = observations Y = Labels

- Generative models: modeling the joint distribution of X and Y → p(X,Y)

$$p(Y \mid X) = \frac{p(Y,X)}{p(X)}$$

- Discriminative models: model posterior probabilities P(Y|X) from the observations directly

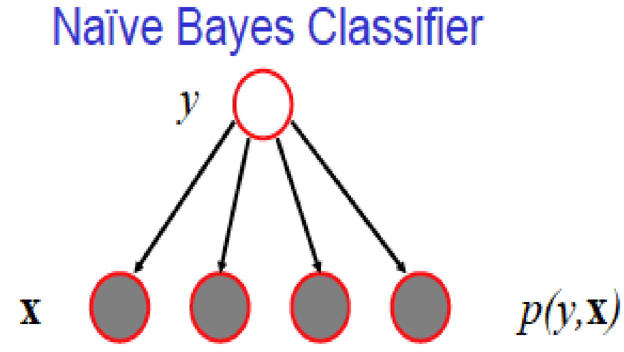
# Independent output variables

$$p(y|X) = \frac{p(y,X)}{p(X)}$$

Generative Model

$$p(X,y) = p(y)p(X|y)$$

$$= p(y)\prod_{k=1}^{K} p(x_k|y)$$

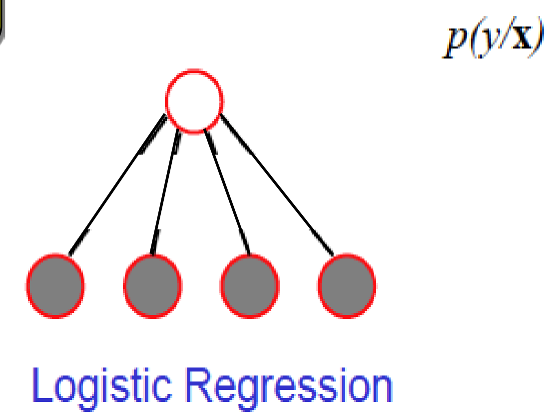



Discriminative Model

$$p(y|X) = \frac{1}{1+\exp(-W^T X)} = \sigma(W^T X)$$

$$\sigma(a) = \frac{1}{1+\exp(-a)}$$

$$a = W^T X$$



# Relationship between Logistic Regression and Naïve Bayes

$$y \in \{y1, y2\}$$

$$p(y = y1 \mid X)?$$

$$Naive\ Bayse:$$

$$p(y = y1 \mid X) = \frac{p(y = y1, X)}{p(X) = \sum_y p(y, X)}$$

$$= \frac{P(X \mid y = y1)p(y = y1)}{p(X \mid y = y1)p(y = y1) + p(X \mid y = y2)p(y = y2)}$$

$$Logistic\ regression:$$

$$p(y = y1 \mid X) = \frac{1}{1 + \exp(-a)}$$

$$a = \ln \frac{p(X \mid y = y1)p(y = y1)}{p(X \mid y = y2)p(y = y2)} = \ln \frac{p(y = y1 \mid X)}{p(y = y2 \mid X)}$$

$$a = W^T X$$

$$y \in \{y1, y2\}$$

$$p(y = y1 \mid X)?$$

$$Naive\ Bayse:$$

$$p(y = y1 \mid X) = \frac{p(y = y1, X)}{p(X) = \sum_y p(y, X)}$$

$$= \frac{P(X \mid y = y1)\, p(y = y1)}{p(X \mid y = y1)\, p(y = y1) + p(X \mid y = y2)\, p(y = y2)}$$

$$Logistic\ regression:$$

Modeling the likelihood is required

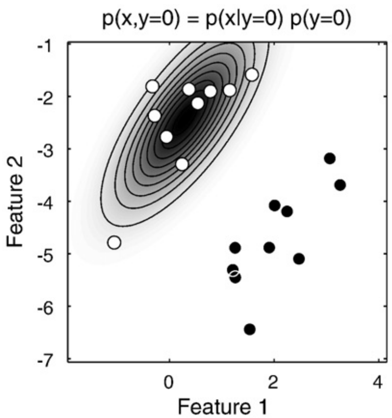
$$p(y = y1 \mid X) = \frac{1}{1 + \exp(-a)}$$

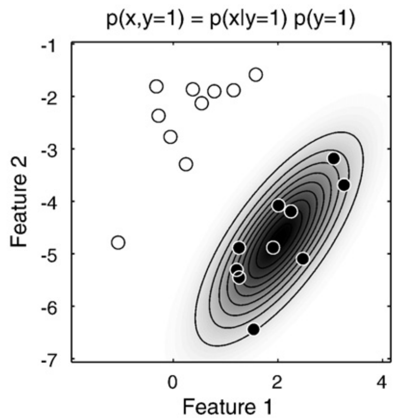
$$a = \ln \frac{\boxed{p(X \mid y = y1)}\, p(y = y1)}{\boxed{p(X \mid y = y2)}\, p(y = y2)} = \ln \frac{p(y = y1 \mid X)}{p(y = y2 \mid X)}$$

$$a = \boxed{W^T X}$$

No modeling

# Generative vs. Discriminative

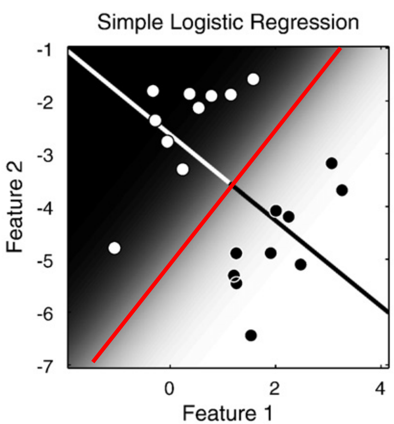



Generative Model

$$P(Y|X) = p(X,Y) / P(X)$$

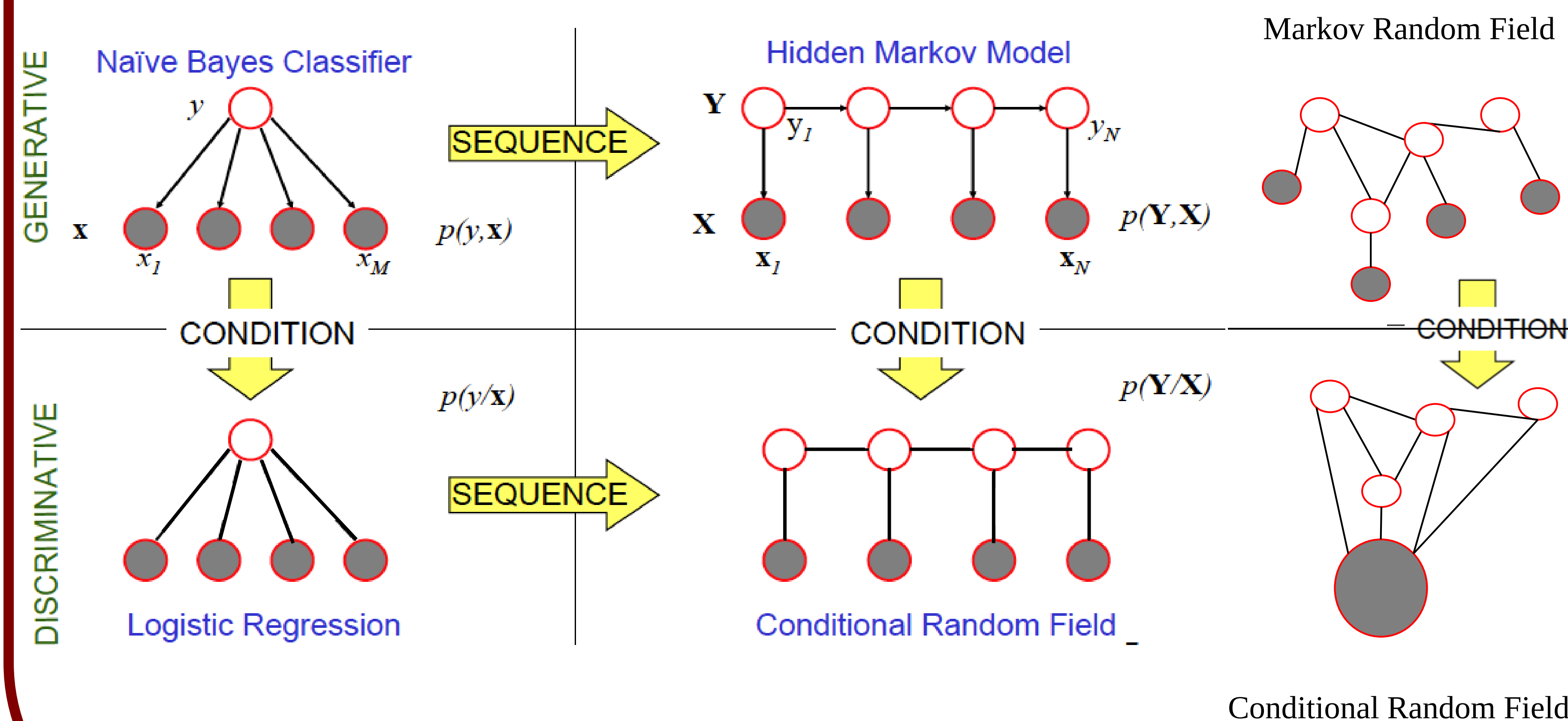
$$P(Y|X) = P(X|Y)P(Y)/P(X)$$



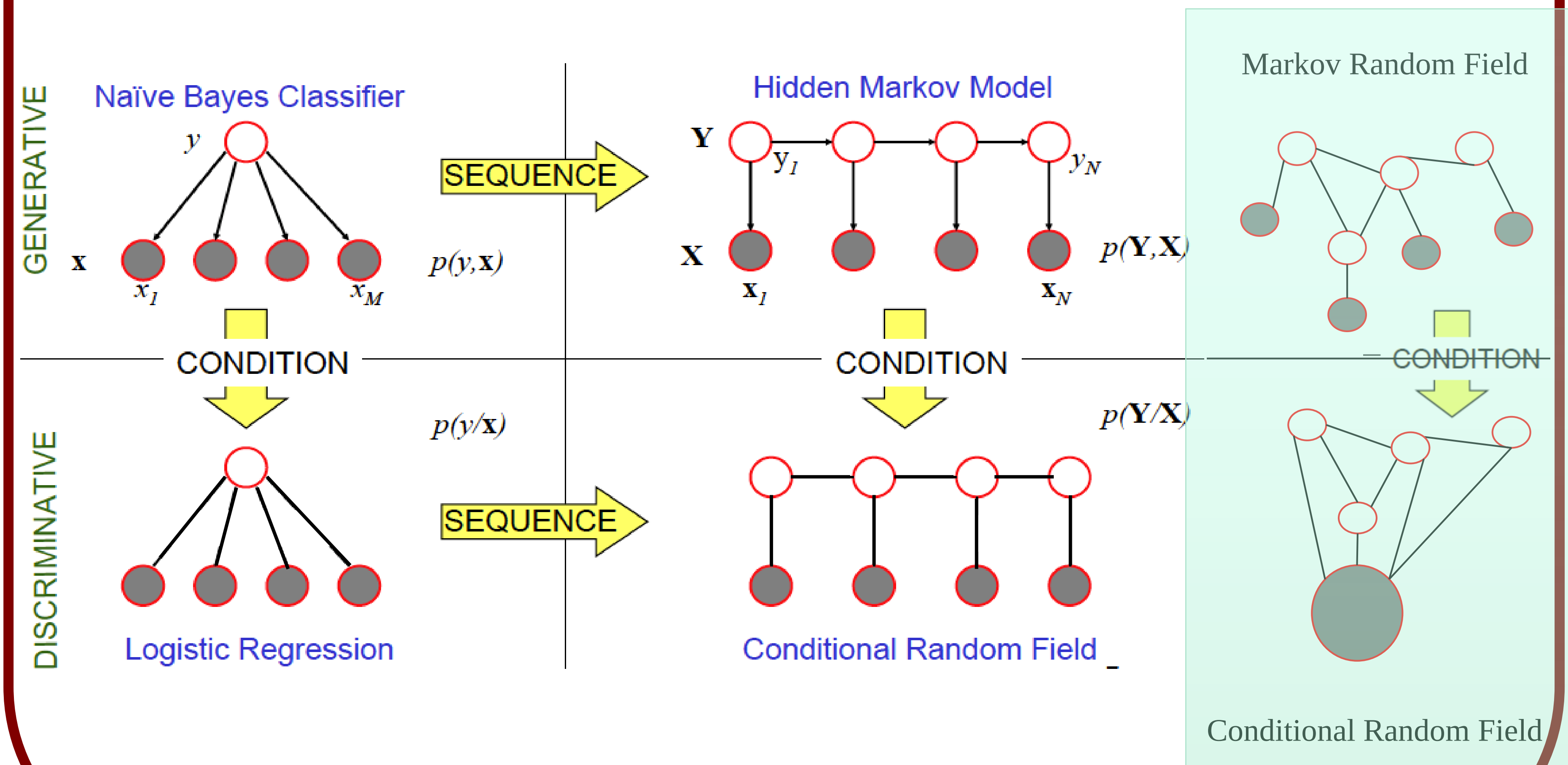

$$P(y=1|x) = \frac{1}{1 + exp(-\mathbf{w}^T \mathbf{x} \cdot)}$$

Discriminative Model

# Relationships between different graphical models



# Relationships between different graphical models



# Generative vs. Discriminative

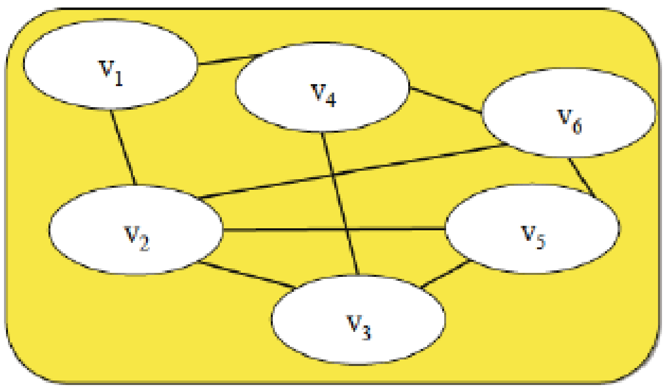
- Regions in an image are often homogenous; neighboring pixels usually have similar properties (intensity, color, texture, ...)
- A graphical random field lets us model the problem probabilistically and enables us to capture such contextual constraints
  1. Generative approach models joint distributions of X and Y
     → Markov random field (MRF)
  2. Discriminative approach models posterior distribution of Y given X
     → Conditional random field (CRF)

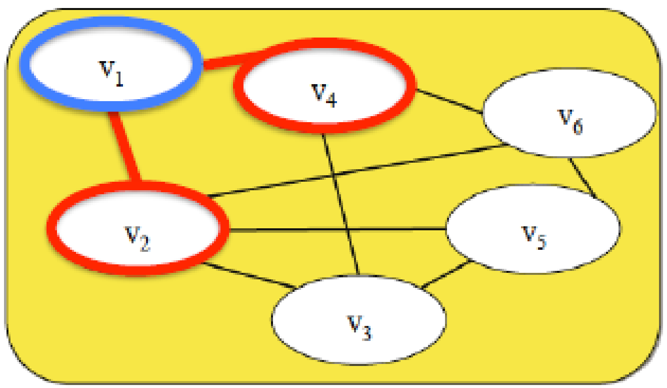
# Markovianity

- In a Markov random field each variable is independent of others given its neighbors, i.e. :

$$p(v_i|\{V - v_i\}) = p(v_i|V_{N_i})$$

$V_{Ni}$ = neighbours of node I
$N_i$= Markovian blanket of node i



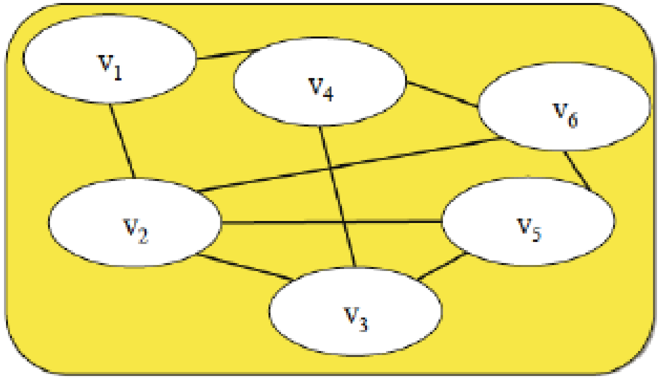

$P(v_1|V) = P(v_1|v_2, v_4)$

# Markov Random Fields

The joint probability distribution of all random variables can be written as product of its factors:

$$p(V) = \frac{1}{Z} \prod_{c_j \in C} \psi_{c_j}(V_{c_j}) \quad V=\{v_1,v_2,\ldots,v_n\}$$

$$Z = \sum_V \prod_{c_j \in C} \psi_{c_j}(V_{c_j})$$




**Clique:** A clique is a subset of random variables (i.e. nodes in the graph that are fully connected to each other.

**Potential function** (also called factor)**:** $\Psi_{cj}$ specifying the local interactions between the nodes in a clique $c_j$ (i.e. $V_{cj}$)
$\psi_{cj} : V_{cj} \rightarrow R^+$

Assuming pair-wise cliques, p(V) can be written as:

$$p(V) = \frac{1}{Z}\psi_1(v_1, v_2) \times \psi_2(v_1, v_4) \times \psi_3(v_2, v_6) \times \psi_4(v_2, v_5) \times$$
$$\psi_5(v_2, v_3) \times \psi_6(v_3, v_4) \times \psi_7(v_3, v_5) \times \psi_8(v_4, v_6) \times \psi_9(v_5, v_6)$$

# Hammersley-Clifford Theorem

- Hammersley-Clifford theorem: a random field is a MRF if and only if the joint distribution of the variables i.e. $P(\mathrm{V})$, follows a Gibbs distribution:

$$P(V) = \frac{1}{Z}\exp(-\overset{\text{Field Energy}}{E(V)}) = \frac{1}{Z}\exp(-\sum_{c_i \in C}\overset{\text{Clique Energy}}{E(V_{c_i})})$$

From before:

$$p(V) = \frac{1}{Z}\prod_{c_j \in C}\underbrace{\psi_{c_j}(V_{c_j})}_{\text{Clique potential}}$$

Therefore:

$$\psi_{c_j}(V_{c_j}) = \boxed{\phantom{xxxxxxxxxxxx}}$$

# Hammersley-Clifford Theorem

- Hammersley-Clifford theorem: a random field is a MRF if and only if the joint distribution of the variables i.e. $P$(V), follows a Gibbs distribution:

$$P(V) = \frac{1}{Z}\exp(-\underbrace{E(V)}_{\text{Field Energy}}) = \frac{1}{Z}\exp(-\sum_{c_i \in C}\underbrace{E(V_{c_i})}_{\text{Clique Energy}})$$

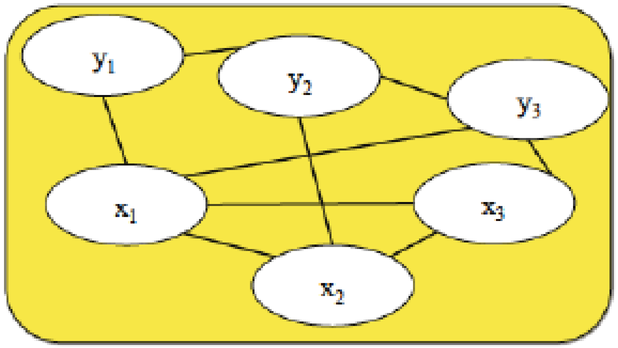
From before:

$$p(V) = \frac{1}{Z}\prod_{c_j \in C}\underbrace{\psi_{c_j}(V_{c_j})}_{\text{Clique potential}}$$

Therefore:

$$\psi_{c_j}(V_{c_j}) = \exp(-E(V_{c_j}))$$

# MRF for Classification




y: labels
x: observations

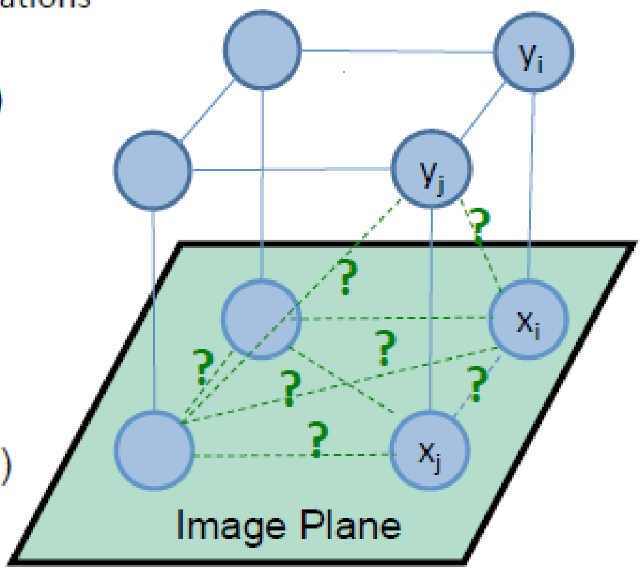
$$P(Y \mid X) = \frac{P(X,Y)}{P(X)} = \frac{P(X \mid Y)P(Y)}{P(X)}$$

$$p(X) = \sum_Y p(X,Y)$$

Generative

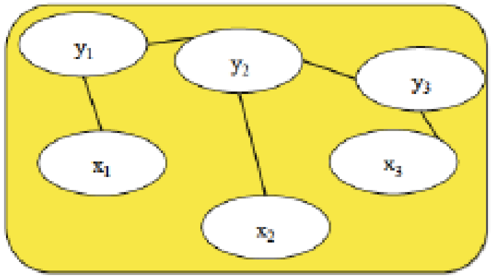
**Y** (labels)

**X** (pixels/ observations)

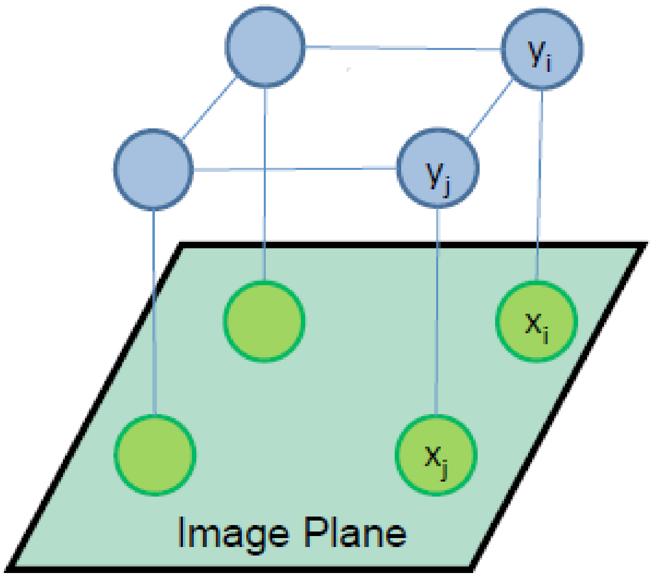



(1) Deciding how observation should be connected is not always obvious
(2) Modeling the joint likelihood over all of the observations makes the computation intractable.

# MRF for Classification




MRF assumption:
Observations are independent given their class labels




$$p(x_i|V) = p(x_i|y_i)$$

$$p(X|Y) = \prod_i p(x_i|y_i)$$

$$p(X,Y) \propto p(X|Y)p(Y) \propto \prod_{i\in V} p(x_i|y_i) \prod_{e\in E} p(y_e)$$

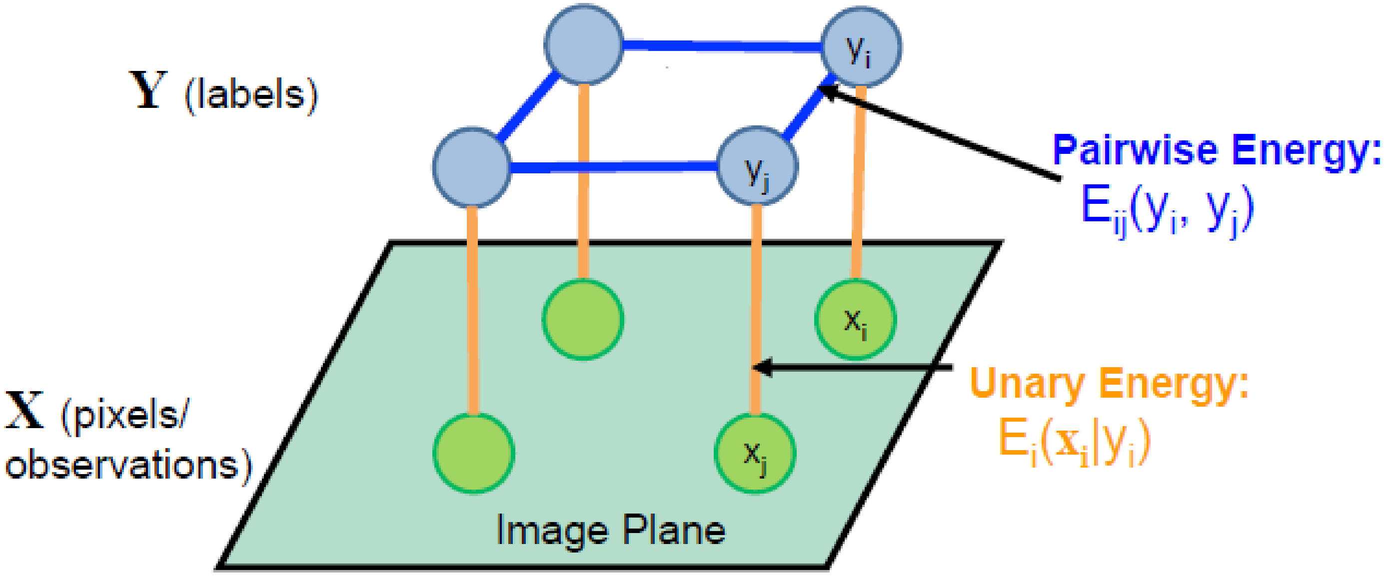
# MRF

Clique potential:

- Unary: proportional to the likelihood of features given labels: log(p(X|Y))
- Pairwise: favors similar labels at neighboring sites:

$$p(X,Y) \propto \boxed{p(X|Y)}\boxed{p(Y)} \propto \prod_{i\in V} p(x_i|y_i) \prod_{e\in E} p(y_e)$$

$$p(X,Y) = \frac{1}{Z}\exp(-E(Y,X)) = \frac{1}{Z}\exp(-(\sum_i \underbrace{E_i(x_i\,|\,y_i)}_{\text{Likelihood}} + \sum_{i,j} \underbrace{E_{i,j}(y_i,y_j)}_{\text{MRF Prior}}))$$

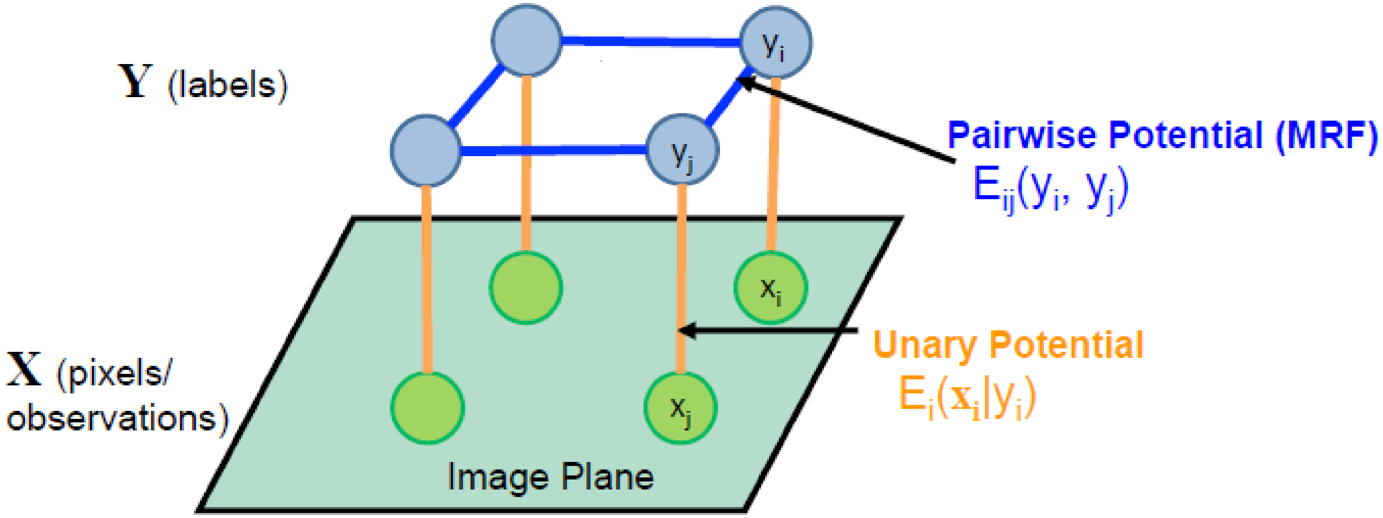


# MRF

$$\exp(-(\sum_i E(x_i \mid y_i)) = \prod_{i \in V} p(x_i | y_i) \longrightarrow E(x_i \mid y_i) = -\log p(x_i \mid y_i)$$

$$p(x_i \mid y_i) = \frac{1}{\sqrt{2\pi\sigma_y}} \exp(-\frac{1}{2\sigma_y^2}(x_i - \mu_y))$$

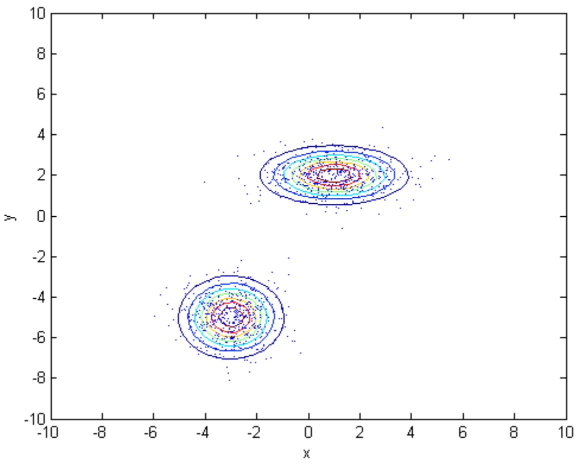
$$\exp(-\sum_{i,j} E(y_i, y_j)) = \prod_{e \in E} p(y_e) \longrightarrow E(y_i, y_j) = \begin{cases} \beta & y_i \neq y_j \\ 0 & y_i = y_j \end{cases}$$

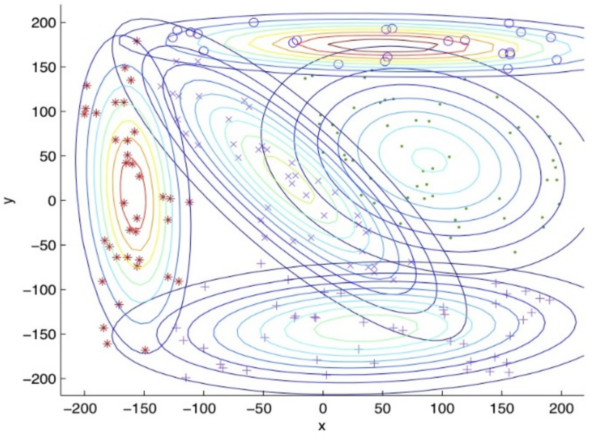


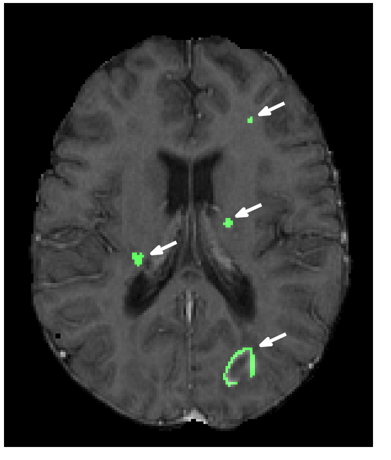
# Shortcomings of MRF for segmentation

- Modeling complicated interactions between observations and labels (which is often required for most vision tasks including the segmentation) is not always straightforward.

- Likelihood term is specifically challenging in binary classification
- Modeling the background class is difficult
- Needs complex models and many training samples



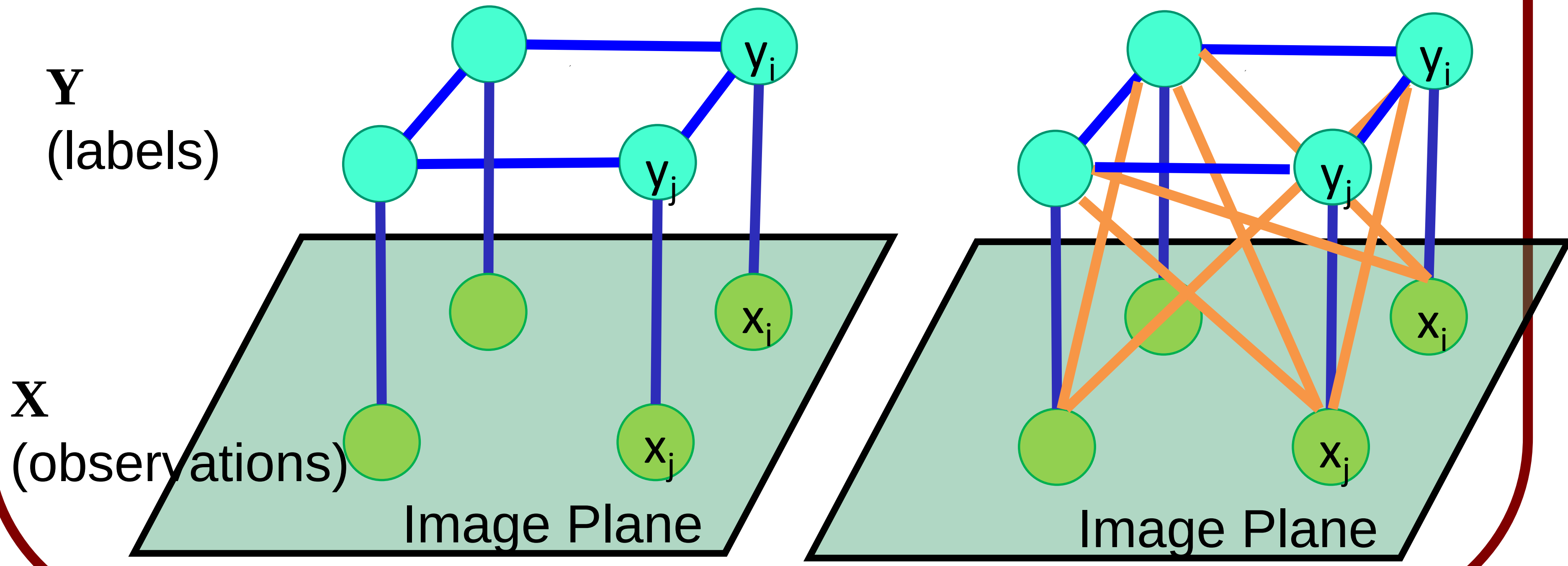



- Long-range complex interactions cannot be modeled so typically restricted to a series of small, local neighbourhoods.
- Another disadvantage of the traditional MRF is the typical assumption made regarding conditional independence of the observations given labels, which leads to over-smoothing of small regions.

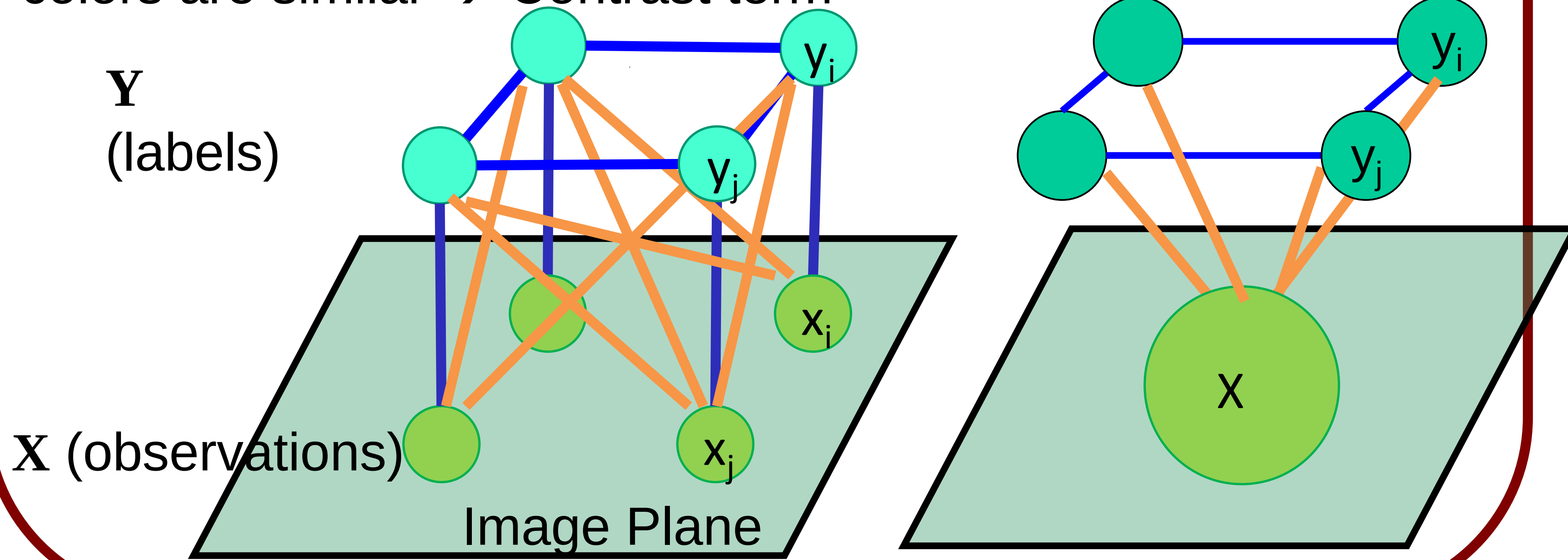
# Conditional Random Fields

- Features can be very correlated with each other. For example: textures, mean value, etc.
- Not modeling the correlation may cause wrong results



# Conditional Random Fields

- No simplifying assumption
- Dependency on X allows introduction of pairwise (or higher order) terms that make use of image.
- For example, neighboring labels should be similar only if pixel colors are similar → Contrast term



# Conditional Random Fields

$$P(Y) = \frac{1}{Z} \exp(-E(Y))$$

$$P(Y \mid X) = \frac{1}{Z(X)} \exp(-E(Y \mid X))$$

$$\frac{1}{Z(X)} \exp(-(\sum_i E_i(y_i \mid X) + \sum_{i,j} E_{ij}(y_i, y_j \mid X))$$

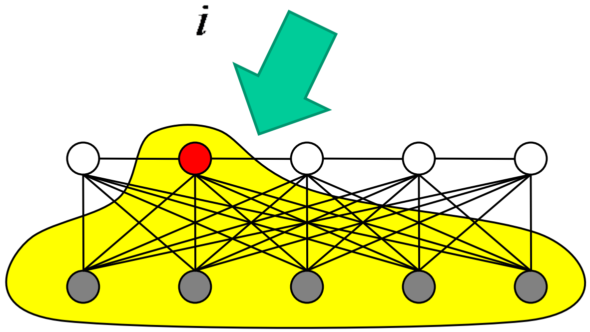
# Conditional Random Fields

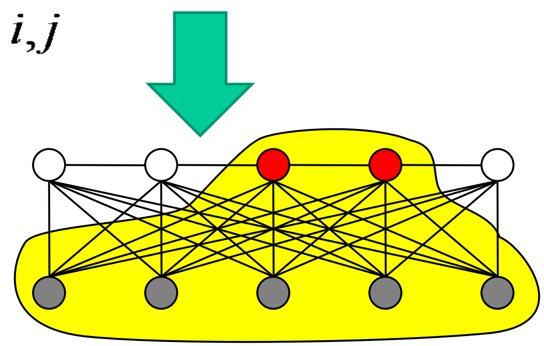
$$P(Y) = \frac{1}{Z}\exp(-E(Y))$$

$$P(Y \mid X) = \frac{1}{Z(X)}\exp(-E(Y \mid X))$$

$$\frac{1}{Z(X)}\exp\Big(-\Big(\sum_i \overset{\text{unary}}{\boxed{E_i(y_i \mid X)}} + \sum_{i,j} \overset{\text{pairwise}}{\boxed{E_{ij}(y_i, y_j}} \mid X)\Big)\Big)$$

Any discriminative classifier such as SVM/RVM/RF/LR





# Conditional Random Fields

- Long-range complex interactions are modelled as ALL the observations are taken intro account
- No restrictions to small neighbourhoods
- Relationships between observations and labels can be complex

# Conditional Random Fields

| | Observed variable x | Target variable y |
|---|---|---|
| Image Segmentation | Pixels/features extracted | Pixel labels |
| Language processing | Words in a sentence | Parts of the speech labeling |

# Other properties…

- Parameter learning:

$$W^* = \arg\max_w \sum_t \log p(Y_t \mid X_t, W)$$

- Inference:

$$Y^* = \arg\max_Y p(Y \mid X, W)$$

# Other properties…

- Parameter learning:

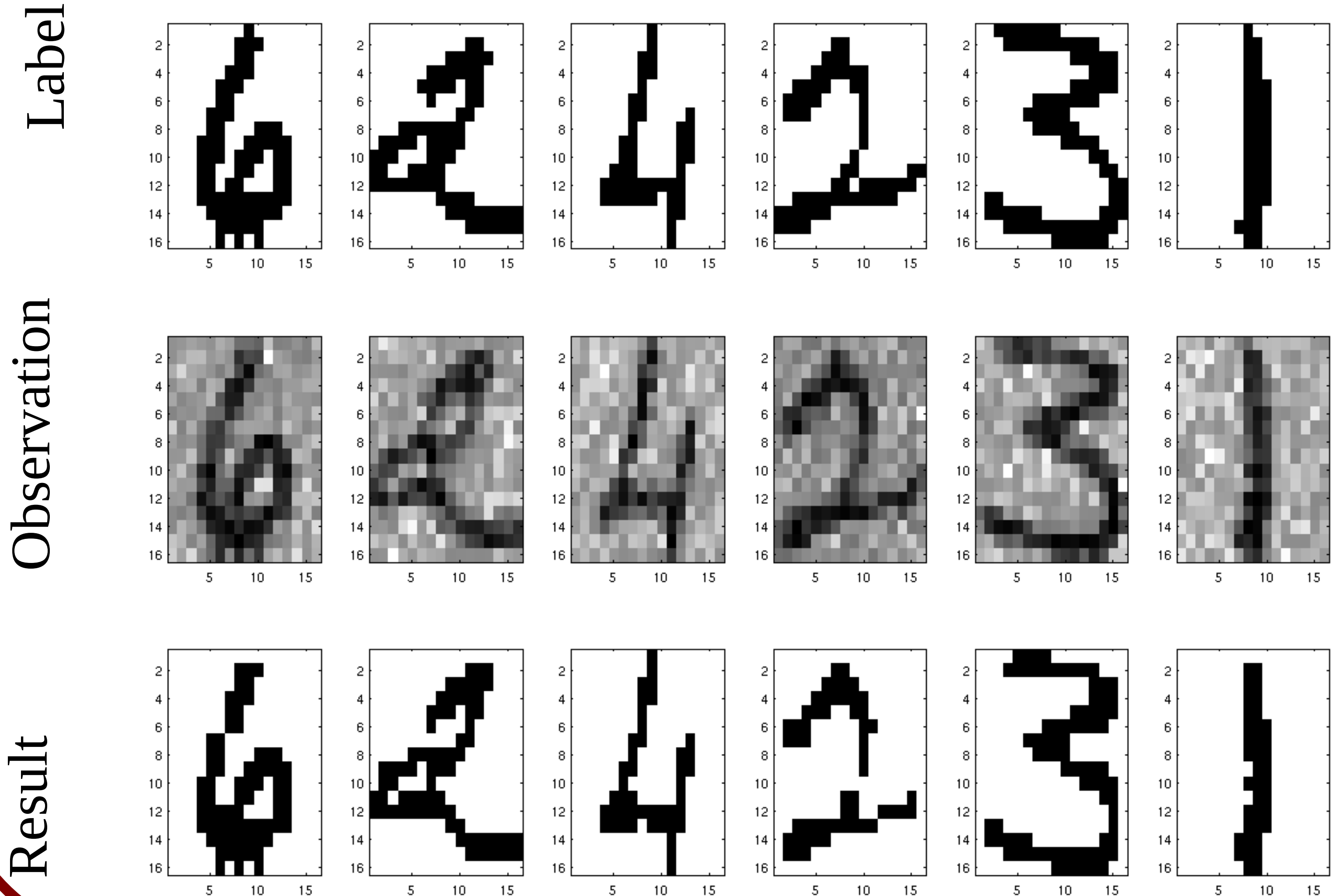
$$W^* = \arg\max_w \sum_t \log p(Y_t \mid X_t, W)$$

- Parameter learning in CRF can be hard.
-This is because the partition function Z, is a constant in MRF but it is a function of observation Z(X) in CRF.

# Some examples

# Toy example

$$p(Y|X) = \frac{1}{Z}\, exp(\sum_{i \in V} U_V(y_i|X) + \sum_{i,j \in E, i \neq j} U_E(y_i, y_j|X))$$



# Sign Detection in Natural Images

Weinman, J. Hanson, A. McCallum, A. Sign detection in natural images with conditional random fields. 2004.


Calculates a joint labeling of image patches, rather than labeling patches independently and use CRF to learn the characteristics of regions that contain text.

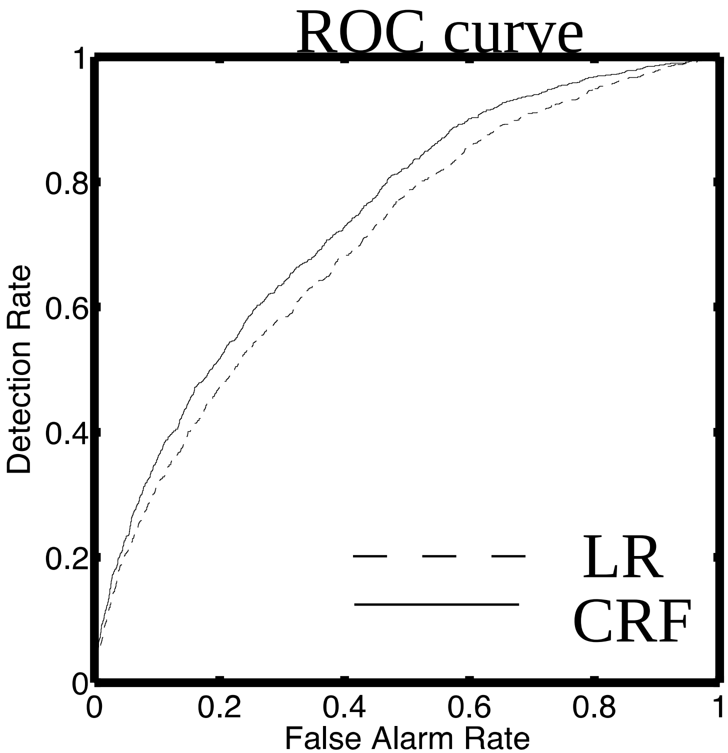
# Sign Detection in Natural Images

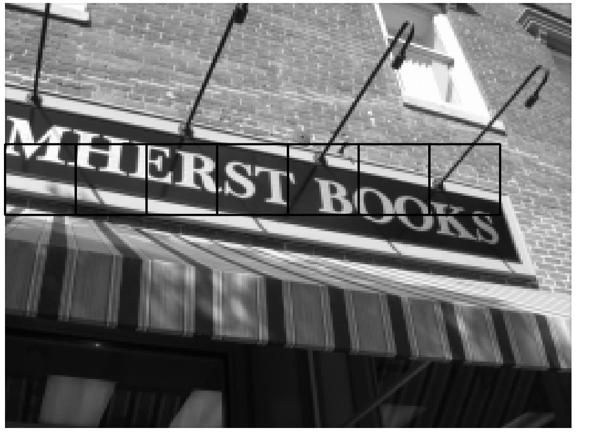
$$
\begin{aligned}
p(\mathbf{y} \mid \mathbf{x}) &= \frac{1}{Z(\mathbf{x})} \prod_{v \in V} \psi_V(y_v, \mathbf{x}) \prod_{(u,v) \in E} \psi_E(y_u, y_v, \mathbf{x}) \\
&= \frac{1}{Z(\mathbf{x})} \exp\left( \sum_{v \in V} \boldsymbol{\lambda} \cdot F(y_v, \mathbf{x}) + \sum_{(u,v) \in E} \boldsymbol{\mu} \cdot G(y_u, y_v, \mathbf{x}) \right)
\end{aligned}
$$

$$
\begin{aligned}
f_y^k(y_v, \mathbf{x}) &= \delta(y, y_v) f^k(\mathbf{x}) \\
g_{y,y'}^j(y_u, y_v, \mathbf{x}) &= \delta(y, y_u) \delta(y', y_v) g^j(\mathbf{x})
\end{aligned}
$$

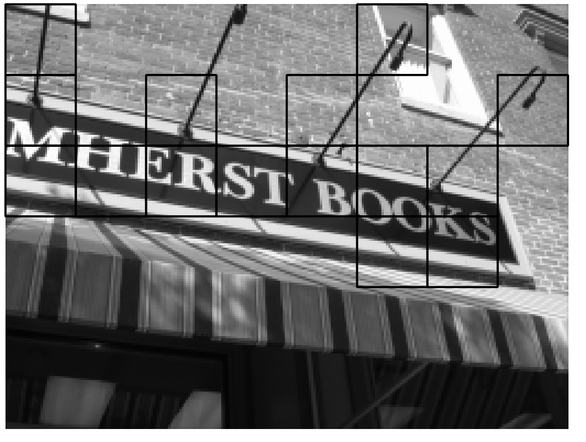
where $\mathbf{f} = \left(f^k\right)_{k=1\ldots K}$ is a vector of node features (i.e., texture statistics of a region) and $\mathbf{g} = \left(g^j\right)_{j=1\ldots J}$ is a vector of edge features (i.e., differences between statistics of neighboring regions), so that $F = \left(f_y^k\right)_{k=1\ldots K, y \in \mathcal{Y}}$ and $G = \left(g_{y,y'}^j\right)_{j=1\ldots J, y,y' \in \mathcal{Y} \times \mathcal{Y}}$. Thus $\boldsymbol{\lambda} \in \mathbb{R}^{K|\mathcal{Y}|}$ and $\boldsymbol{\mu} \in \mathbb{R}^{J|\mathcal{Y}|^2}$.

# Sign Detection in Natural Images




CRF




LR

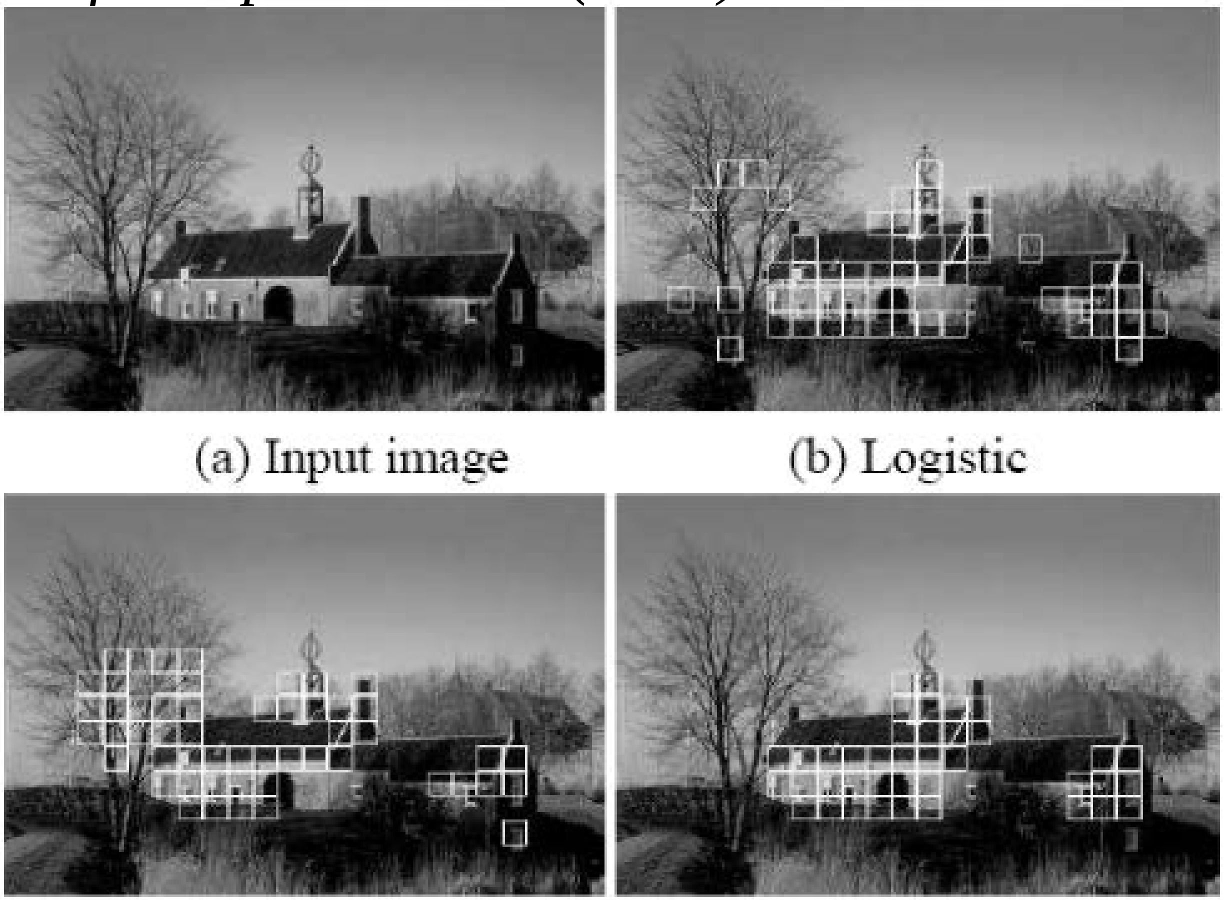
# Structure detection

*Kumar, S., Hebert, M.: Discriminative Random Fields. International Journal of Computer Vision (2006) 179–201*

- DRF: discriminative model used for classifying patterns (e.g. target vs. non-target)
- DRF is a generally thought of as the first CRF where:
  - Unary (association) and pairwise (interaction) potentials designed using local discriminative classifiers

# Structure detection

*Kumar, S., Hebert, M.: Discriminative Random Fields. International Journal of Computer Vision (2006) 179–201*



(a) Input image (b) Logistic

(c) MRF (d) DRF

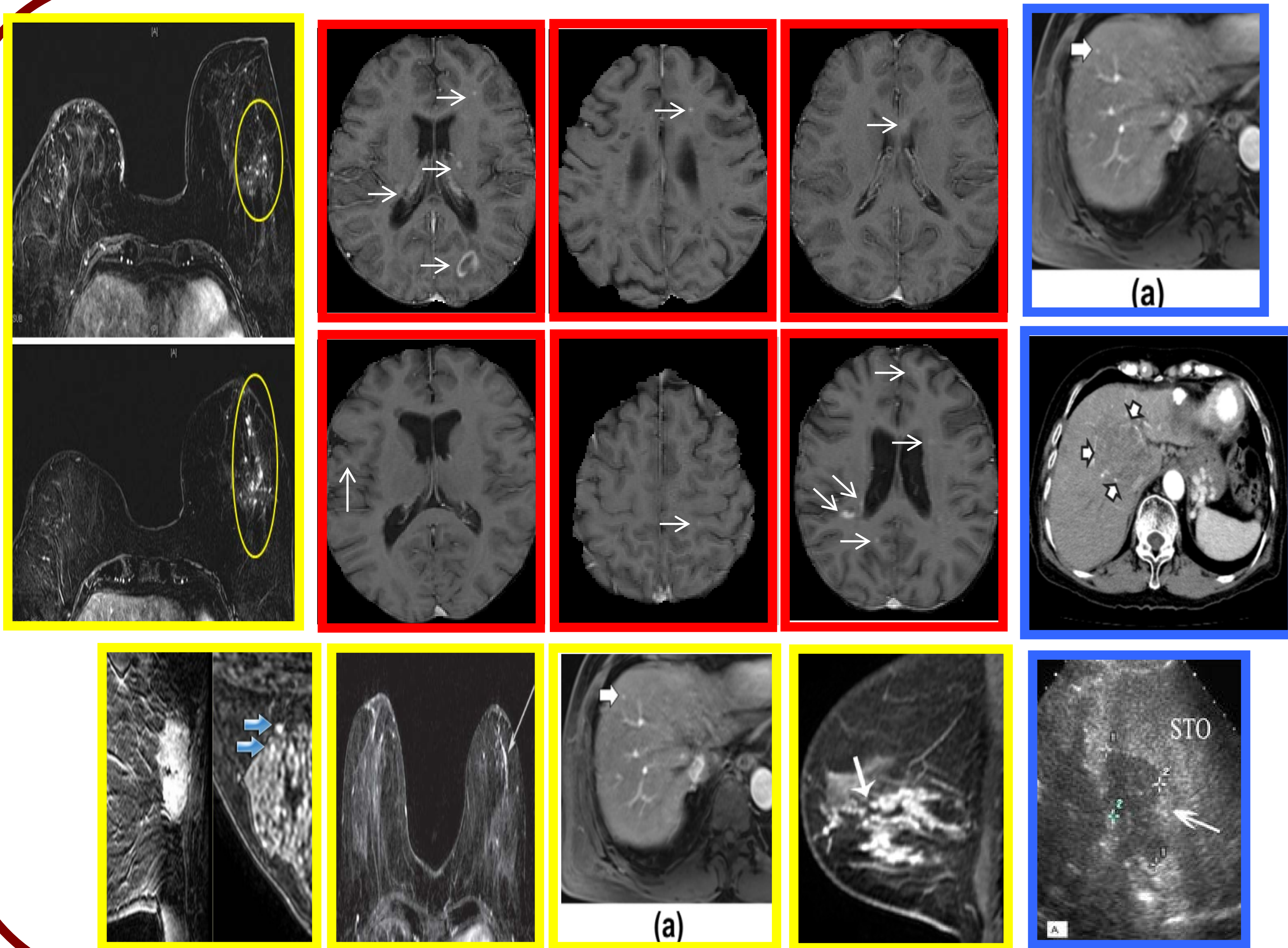
structure detection by DRF

(label each image site as *structured* or *nonstructured*)

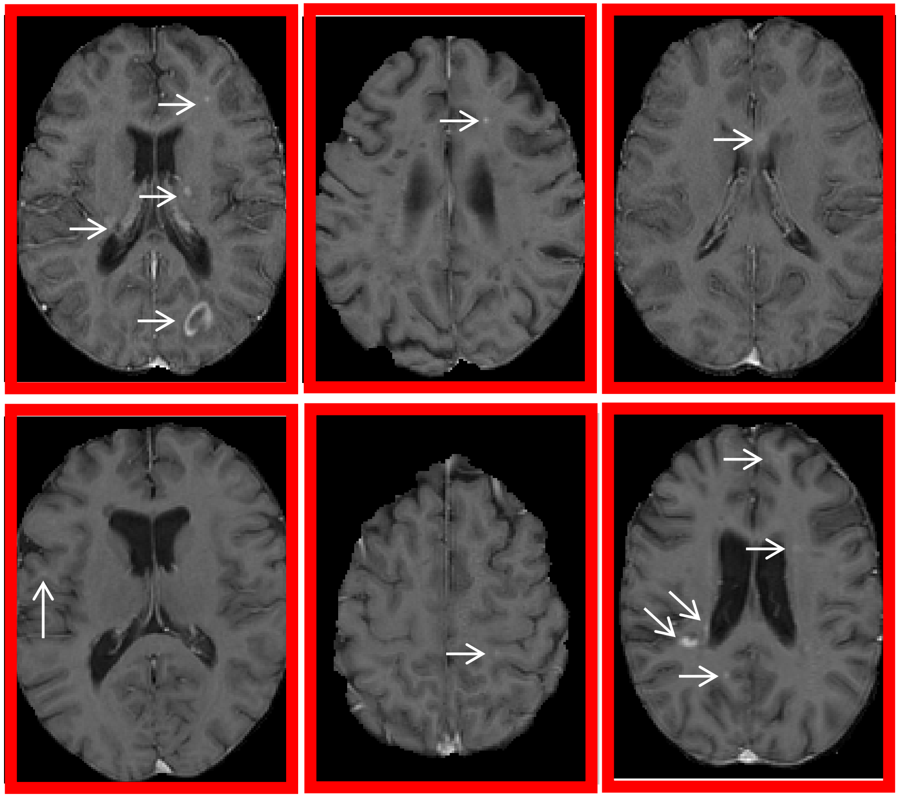
For similar detection rates, DRF reduces the false positives considerably.

McG Arbel

# Detection of Small Enhanced Pathologies in Medical Image Analysis


(a)
STO
(a)

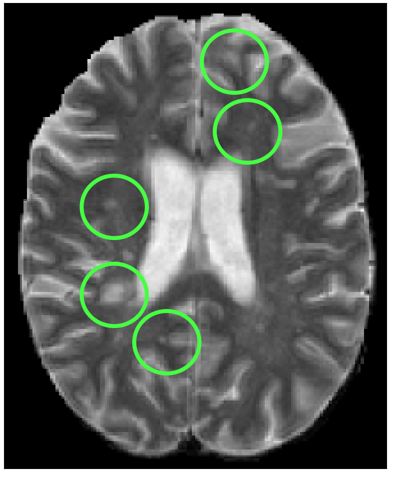
# Detection of enhancing lesions in MS



# Multiple Sclerosis (MS)

- The **most common** neurological disorder in Canada and around the world, affecting **young adults**.
- There is **no cure**. It results in **disability or death**.
- Characterized by lesions appearing in brain MRI.

lesions

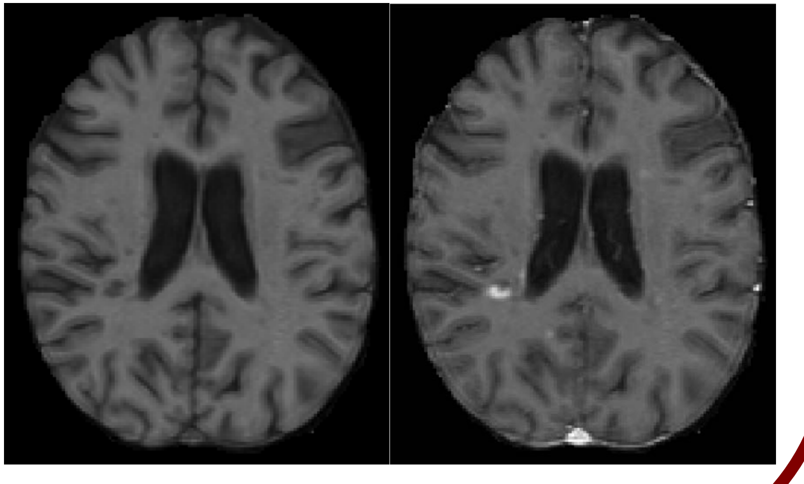


# Multiple Sclerosis (MS)

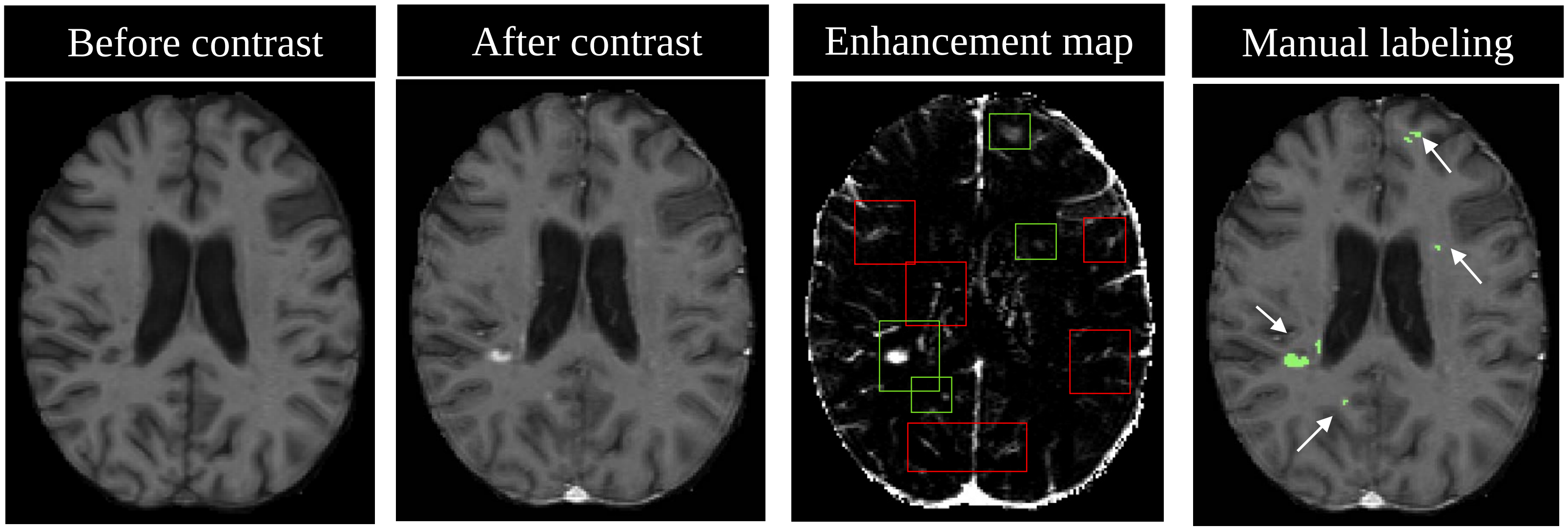
- **Enhancing lesions**: enhanced after injection of a contrast agent
- **Biomarkers of disease activity**
- **Early detection is crucial**
- **Drug efficacy**
- **History and evolution**

**of the disease**

Enhancing lesions



# Enhancing MS Lesions

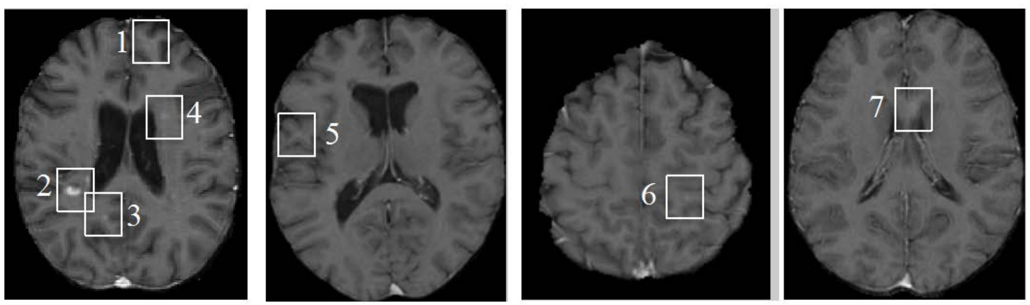


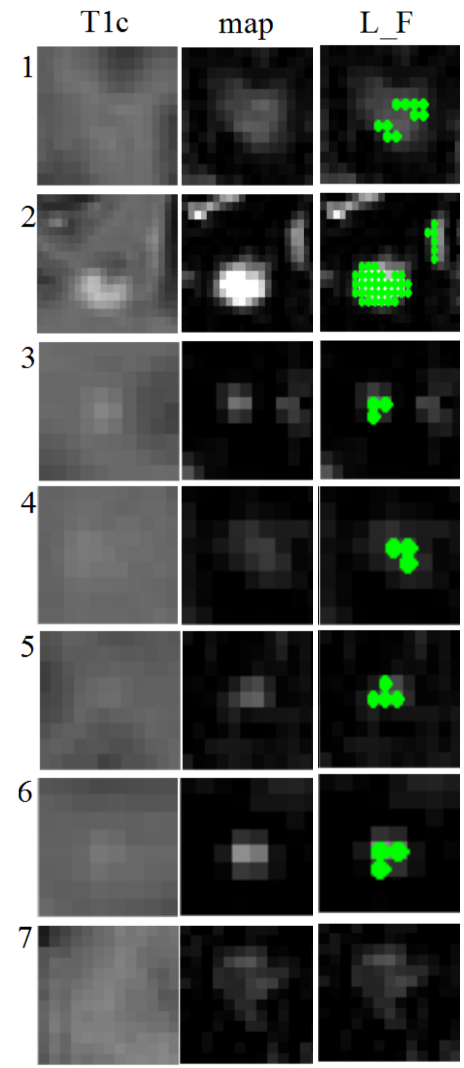

**Challenges:**

- Various shapes, location and intensity patterns,
- Small sizes,
- Large amount of non-lesional enhancements

→ Current standard: **manual labeling**

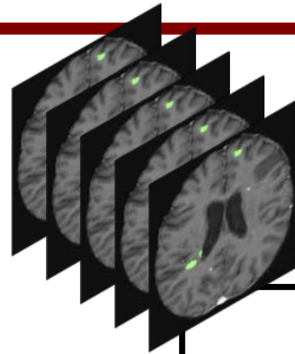
**Automatic technique needed!**


1
2
3
4
5
6
7


T1c
map
L_F
1
2
3
4
5
6
7

# Hierarchical Adaptive Texture CRF : HAT- CRF

# HAT-CRF



output

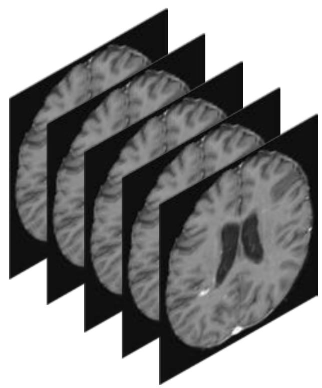
**Adaptive lesion level**

$$\mathbb{E}(Y^{\mathsf{p}}, X, \boldsymbol{\lambda}^{\mathsf{p}}) = \sum_{c\in C^{\mathsf{p}}} \lambda^{\mathsf{p}}_{\mathbf{y}_c} E^v_c(\mathbf{y}_c, X) + \sum_k \lambda^{\mathsf{p}}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}) + \sum_k \sum_i \lambda^{\mathsf{p}}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}).$$

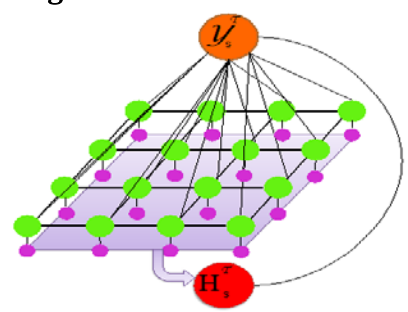
4

4

**Voxel-wise interactions:**
taken from image volume Time = t



**Higher order textural feature per each candidate**



Stationary Textures:
1- RIFT$^t$
2- spin$^t$
3- LBP$^t$

3

Candidate Lesion Detection

2

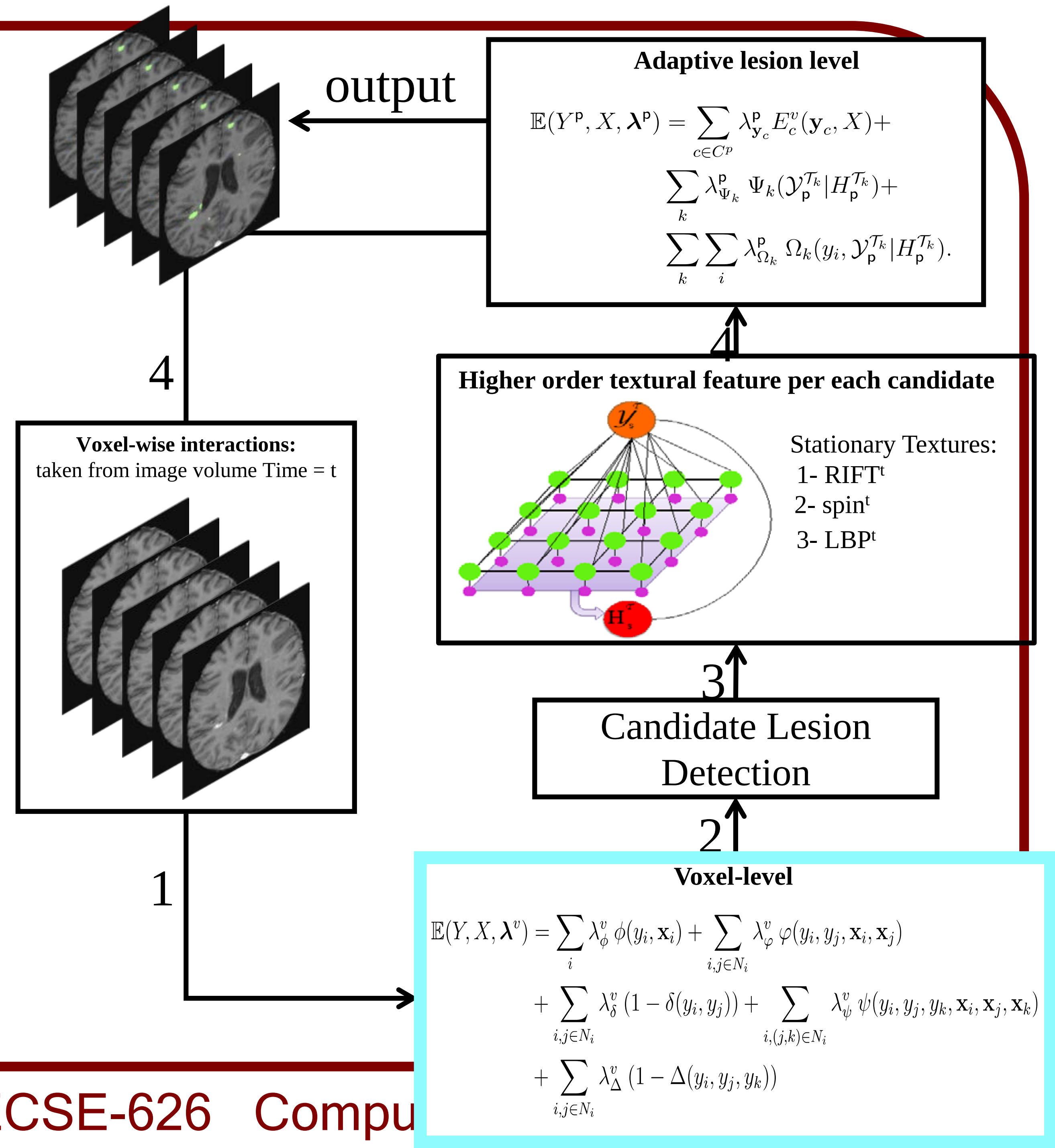
**Voxel-level**

$$\mathbb{E}(Y, X, \boldsymbol{\lambda}^v) = \sum_i \lambda^v_\phi \, \phi(y_i, \mathbf{x}_i) + \sum_{i,j\in N_i} \lambda^v_\varphi \, \varphi(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) + \sum_{i,j\in N_i} \lambda^v_\delta \, (1-\delta(y_i, y_j)) + \sum_{i,(j,k)\in N_i} \lambda^v_\psi \, \psi(y_i, y_j, y_k, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) + \sum_{i,j\in N_i} \lambda^v_\Delta \, (1-\Delta(y_i, y_j, y_k))$$

1

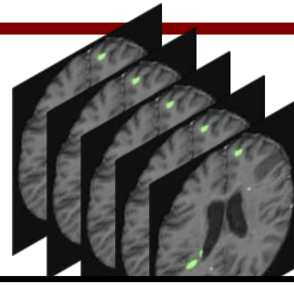
- **Karimaghaloo**, *et al.* “Adaptive Multi-level Conditional Random Fields for Small Enhanced Pathology Detection”, submitted to MIA’14.
- **Karimaghaloo**, *et al.* “Hierarchical Conditional Random Fields for Detection of Gad-enhancing Lesions in Multiple Sclerosis”, MICCAI ’12.

# HAT-CRF



output
Adaptive lesion level
$\mathbb{E}(Y^{\mathsf{p}}, X, \boldsymbol{\lambda}^{\mathsf{p}}) = \sum_{c \in C^{\mathsf{p}}} \lambda^{\mathsf{p}}_{\mathbf{y}_c} E^{v}_{c}(\mathbf{y}_c, X) + \sum_{k} \lambda^{\mathsf{p}}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}) + \sum_{k} \sum_{i} \lambda^{\mathsf{p}}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}).$
4
4
Higher order textural feature per each candidate
Stationary Textures:
1- RIFT^t
2- spin^t
3- LBP^t
Voxel-wise interactions:
taken from image volume Time = t
3
Candidate Lesion Detection
2
1
Voxel-level
$\mathbb{E}(Y, X, \boldsymbol{\lambda}^{v}) = \sum_{i} \lambda^{v}_{\phi} \phi(y_i, \mathbf{x}_i) + \sum_{i,j \in N_i} \lambda^{v}_{\varphi} \varphi(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) + \sum_{i,j \in N_i} \lambda^{v}_{\delta} (1 - \delta(y_i, y_j)) + \sum_{i,(j,k) \in N_i} \lambda^{v}_{\psi} \psi(y_i, y_j, y_k, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) + \sum_{i,j \in N_i} \lambda^{v}_{\Delta} (1 - \Delta(y_i, y_j, y_k))$

# HAT-CRF



output

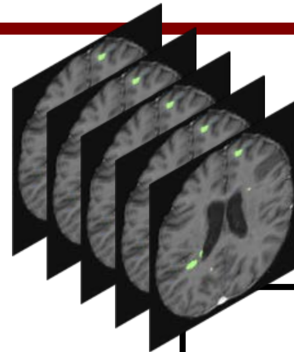
**Adaptive lesion level**

$$\mathbb{E}(Y^{\mathsf{p}}, X, \boldsymbol{\lambda}^{\mathsf{p}}) = \sum_{c \in C^{p}} \lambda^{\mathsf{p}}_{\mathbf{y}_c} E^{v}_{c}(\mathbf{y}_c, X) + \sum_{k} \lambda^{\mathsf{p}}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}) +$$

ate

res:

$$\mathbb{E}(Y, X, \boldsymbol{\lambda}^{v}) = \sum_{i} \lambda^{v}_{\phi}\, \phi(y_i, \mathbf{x}_i) + \sum_{i,j \in N_i} \lambda^{v}_{\varphi}\, \varphi(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j)$$
$$+ \sum_{i,j \in N_i} \lambda^{v}_{\delta}\, (1 - \delta(y_i, y_j)) + \sum_{i,(j,k) \in N_i} \lambda^{v}_{\psi}\, \psi(y_i, y_j, y_k, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)$$
$$+ \sum_{i,j \in N_i} \lambda^{v}_{\Delta}\, (1 - \Delta(y_i, y_j, y_k))$$



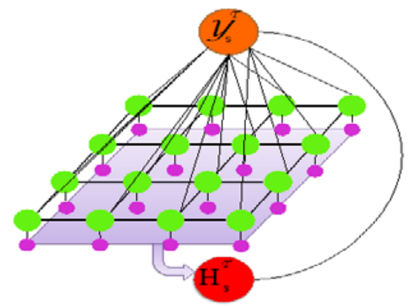
output

**Adaptive lesion level**

$$\mathbb{E}(Y^{\mathsf{p}}, X, \boldsymbol{\lambda}^{\mathsf{p}}) = \sum_{c \in C^{p}} \lambda^{\mathsf{p}}_{\mathbf{y}_c} E^{v}_{c}(\mathbf{y}_c, X) + \sum_{k} \lambda^{\mathsf{p}}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}) + \sum_{k} \sum_{i} \lambda^{\mathsf{p}}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}).$$

4

**Higher order textural feature per each candidate**

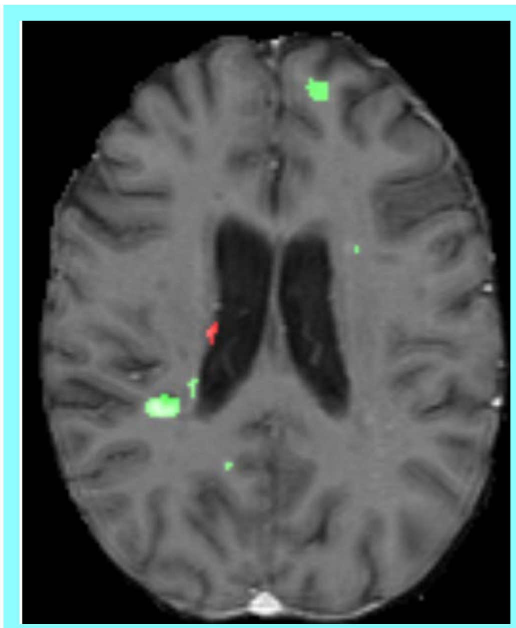


Stationary Textures:
1- RIFT[t]
2- spin[t]
3- LBP[t]

3

Candidate Lesion Detection

2



True Positive

False Positive

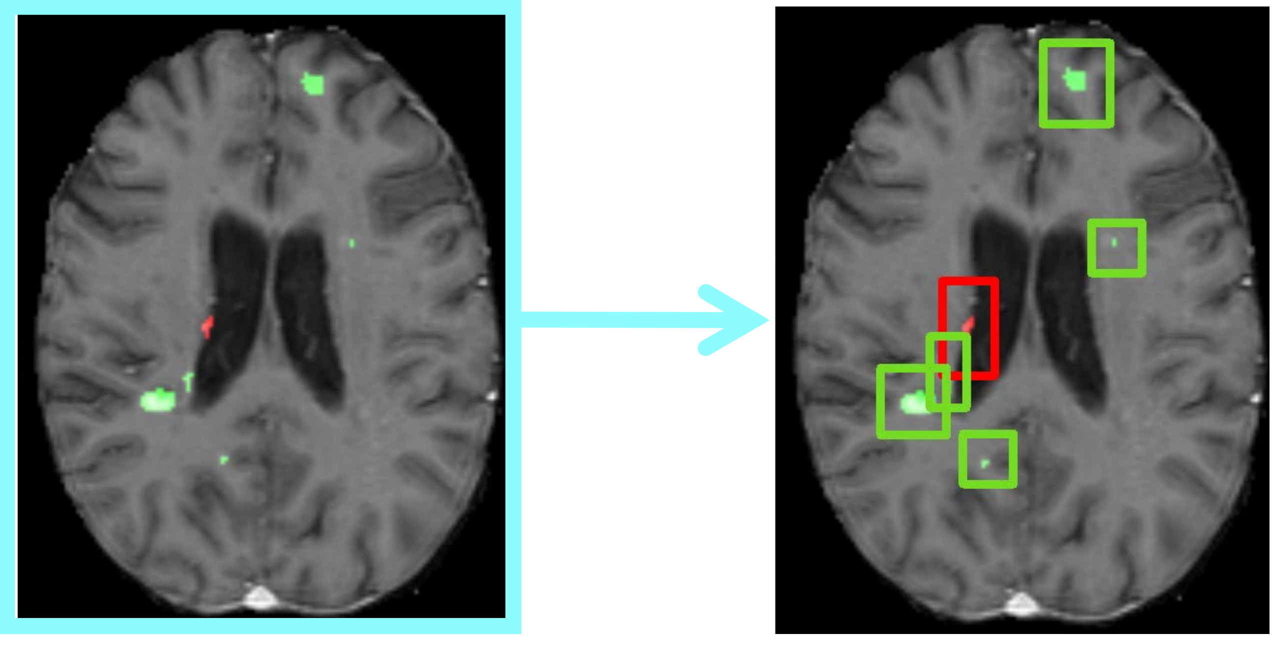
**Voxel-level**

$$\mathbb{E}(Y, X, \boldsymbol{\lambda}^{v}) = \sum_{i} \lambda^{v}_{\phi} \phi(y_i, \mathbf{x}_i) + \sum_{i,j \in N_i} \lambda^{v}_{\varphi} \varphi(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) + \sum_{i,j \in N_i} \lambda^{v}_{\delta} (1 - \delta(y_i, y_j)) + \sum_{i,(j,k) \in N_i} \lambda^{v}_{\psi} \psi(y_i, y_j, y_k, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) + \sum_{i,j \in N_i} \lambda^{v}_{\Delta} (1 - \Delta(y_i, y_j, y_k))$$

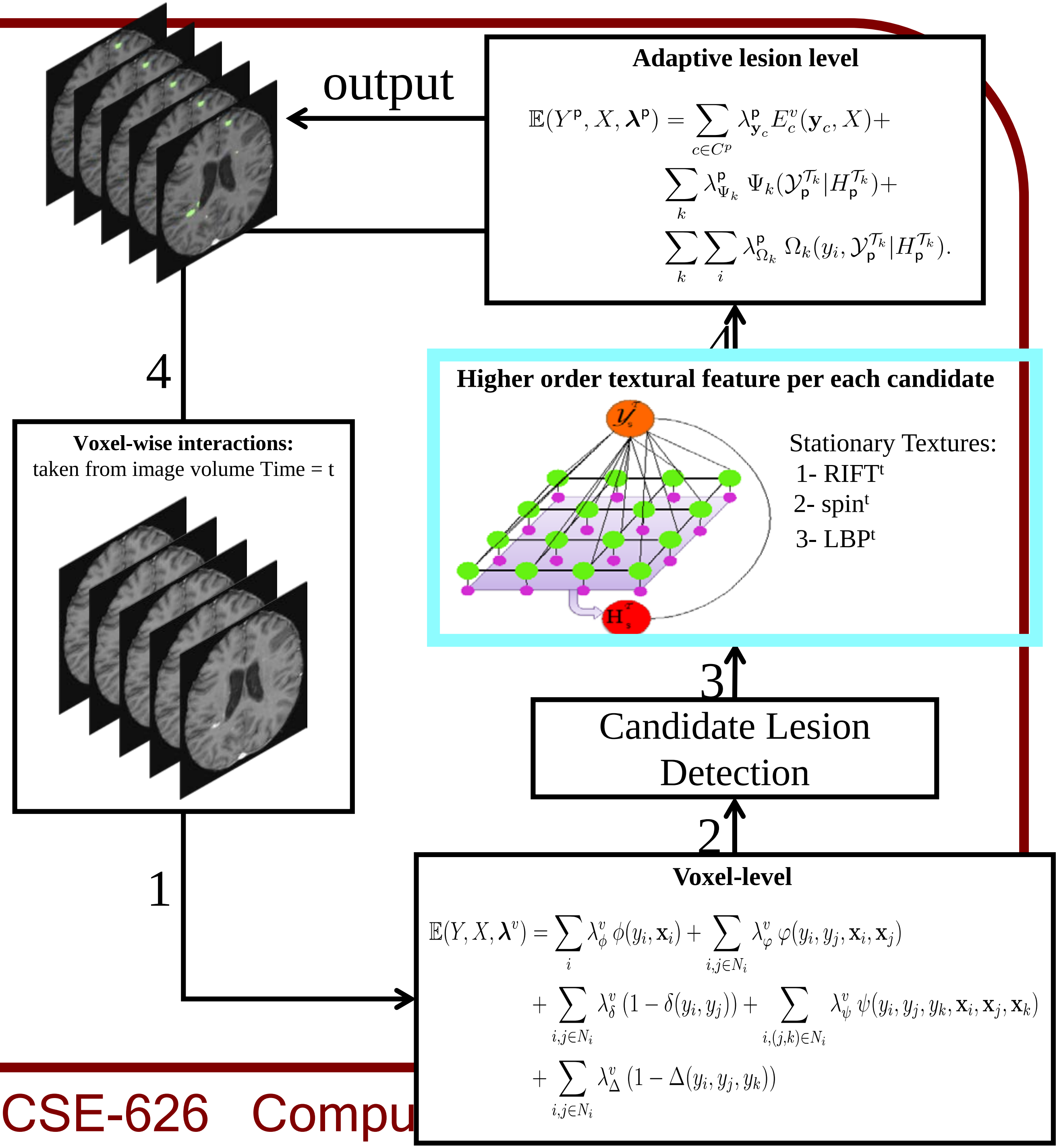
# Candidate lesions

True Positive

False Positive

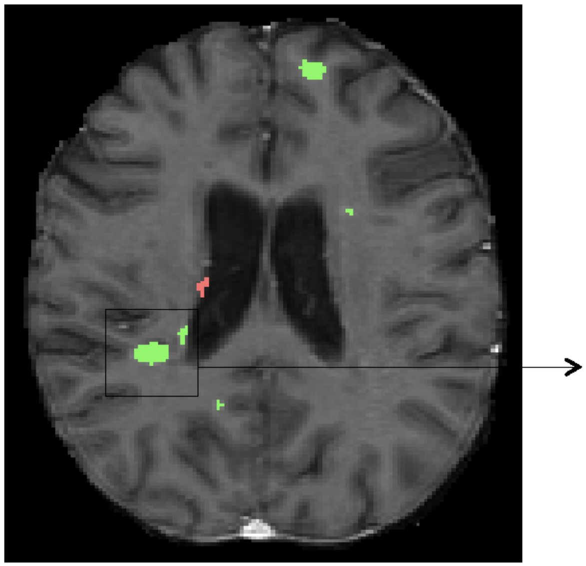


# HAT-CRF

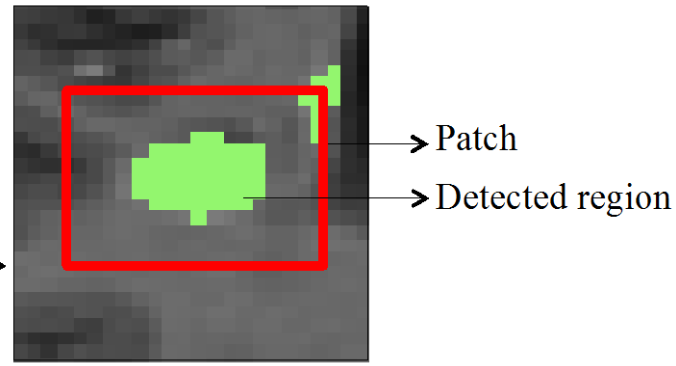


output
Adaptive lesion level
$$\mathbb{E}(Y^{\mathsf{p}}, X, \boldsymbol{\lambda}^{\mathsf{p}}) = \sum_{c \in C^{p}} \lambda^{\mathsf{p}}_{\mathbf{y}_c} E^{v}_{c}(\mathbf{y}_c, X) + \sum_{k} \lambda^{\mathsf{p}}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}) + \sum_{k} \sum_{i} \lambda^{\mathsf{p}}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}).$$
4
Higher order textural feature per each candidate
Stationary Textures:
1- RIFT$^{t}$
2- spin$^{t}$
3- LBP$^{t}$
Voxel-wise interactions:
taken from image volume Time = t
3
Candidate Lesion Detection
2
1
Voxel-level
$$\mathbb{E}(Y, X, \boldsymbol{\lambda}^{v}) = \sum_{i} \lambda^{v}_{\phi} \phi(y_i, \mathbf{x}_i) + \sum_{i,j \in N_i} \lambda^{v}_{\varphi} \varphi(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) + \sum_{i,j \in N_i} \lambda^{v}_{\delta} (1 - \delta(y_i, y_j)) + \sum_{i,(j,k) \in N_i} \lambda^{v}_{\psi} \psi(y_i, y_j, y_k, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) + \sum_{i,j \in N_i} \lambda^{v}_{\Delta} (1 - \Delta(y_i, y_j, y_k))$$

# Higher Order Textures



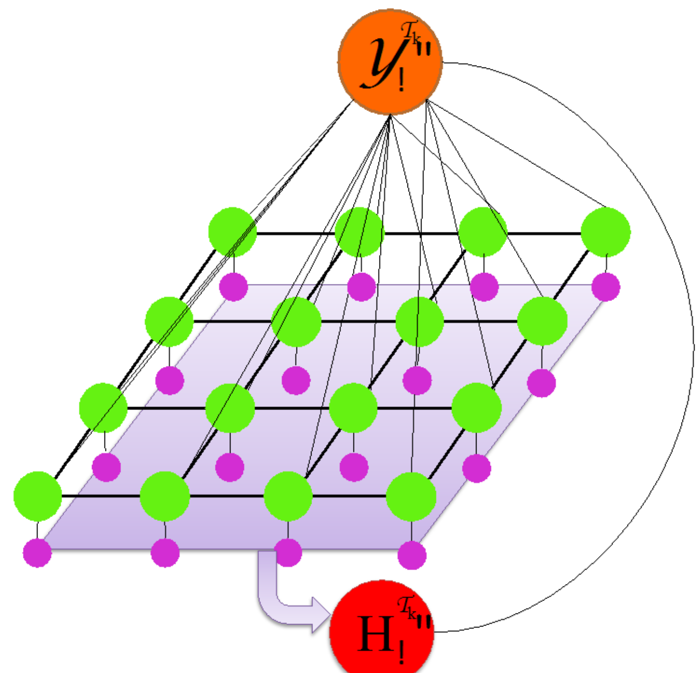
(a)




(b)

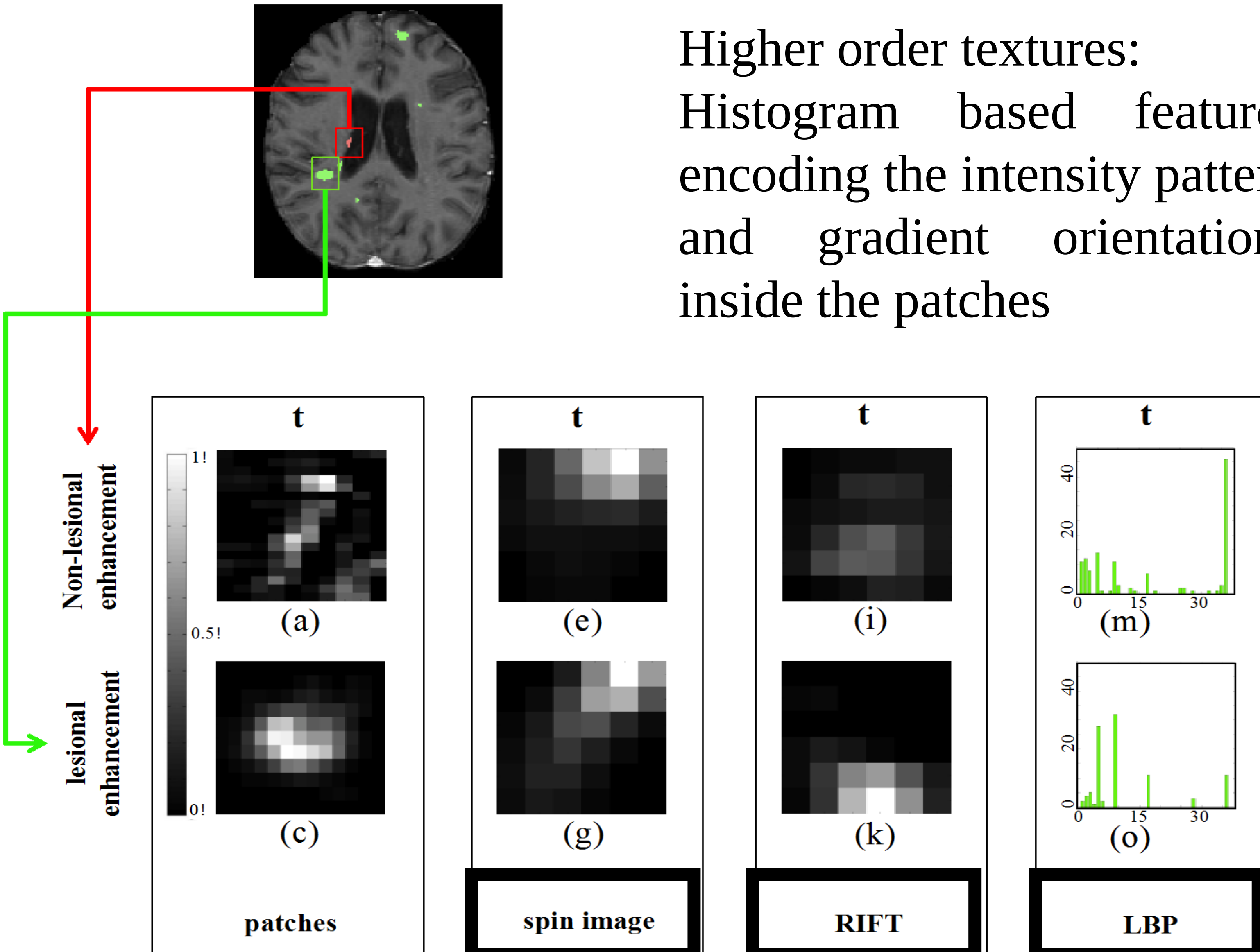
$H_p^{Tk}$ : $k^{th}$ texture extracted for the region inside the patch p

Higher order graphical model

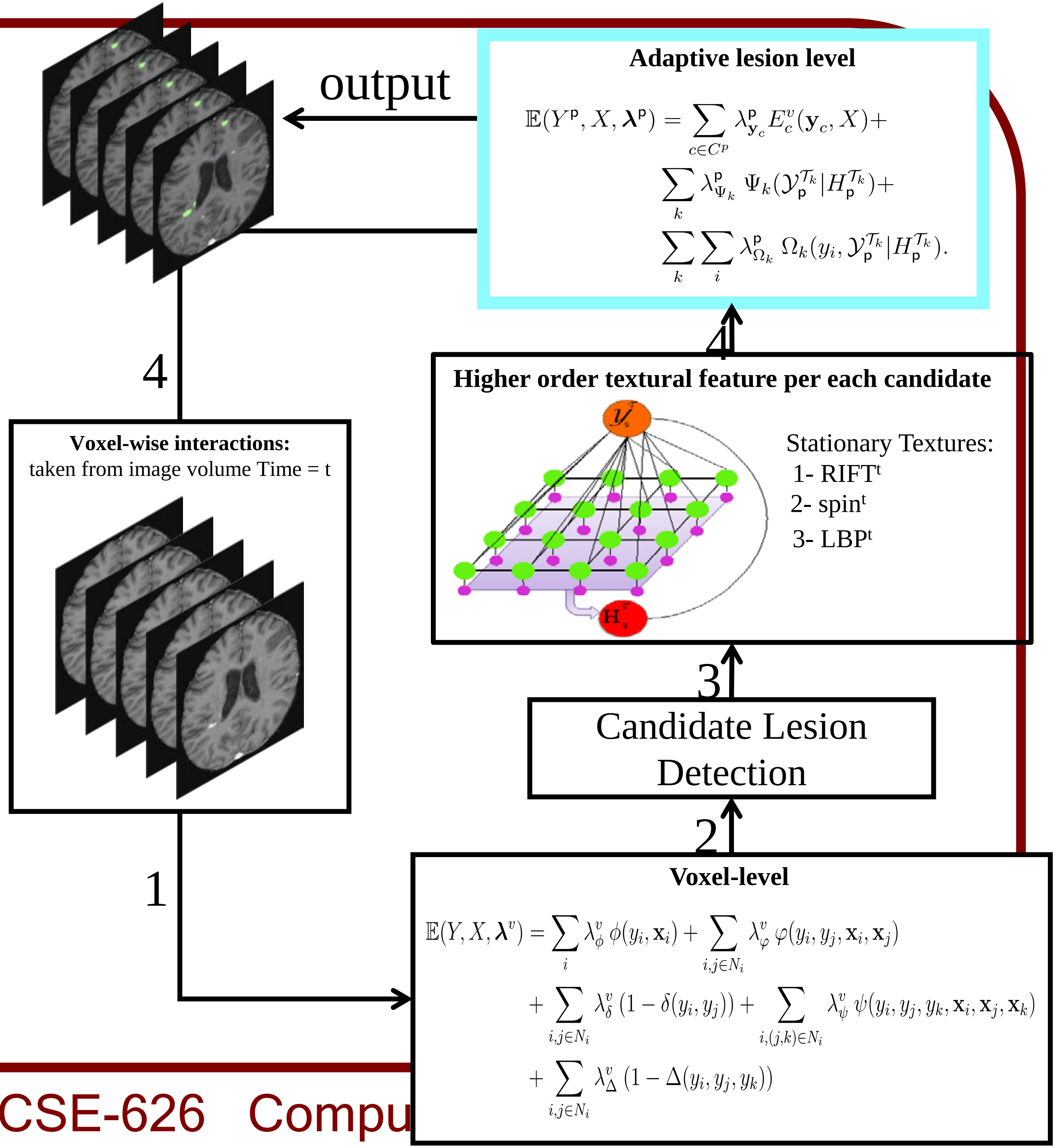


# Stationary Textures

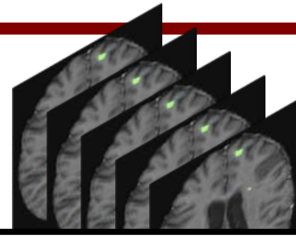
Higher order textures:
Histogram based features encoding the intensity pattern and gradient orientations inside the patches



# HAT-CRF



output
Adaptive lesion level
$$\mathbb{E}(Y^{\mathsf{p}}, X, \boldsymbol{\lambda}^{\mathsf{p}}) = \sum_{c \in C^{p}} \lambda^{\mathsf{p}}_{\mathbf{y}_c} E^{v}_{c}(\mathbf{y}_c, X) + \sum_{k} \lambda^{\mathsf{p}}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}) + \sum_{k} \sum_{i} \lambda^{\mathsf{p}}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}).$$
4
4
Higher order textural feature per each candidate
Stationary Textures:
1- RIFT^t
2- spin^t
3- LBP^t
H
Voxel-wise interactions:
taken from image volume Time = t
3
Candidate Lesion Detection
2
1
Voxel-level
$$\mathbb{E}(Y, X, \boldsymbol{\lambda}^{v}) = \sum_{i} \lambda^{v}_{\phi} \phi(y_i, \mathbf{x}_i) + \sum_{i,j \in N_i} \lambda^{v}_{\varphi} \varphi(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) + \sum_{i,j \in N_i} \lambda^{v}_{\delta} (1 - \delta(y_i, y_j)) + \sum_{i,(j,k) \in N_i} \lambda^{v}_{\psi} \psi(y_i, y_j, y_k, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) + \sum_{i,j \in N_i} \lambda^{v}_{\Delta} (1 - \Delta(y_i, y_j, y_k))$$



output

**Adaptive lesion level**

$$
\begin{aligned}
\mathbb{E}(Y^{\mathsf{p}}, X, \boldsymbol{\lambda}^{\mathsf{p}}) = \sum_{c \in C^p} \lambda^{\mathsf{p}}_{\mathbf{y}_c} E^v_c(\mathbf{y}_c, X) + \\
\sum_k \lambda^{\mathsf{p}}_{\Psi_k} \, \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}) + \\
\sum_k \sum_i \lambda^{\mathsf{p}}_{\Omega_k} \, \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}).
\end{aligned}
$$

**ch candidate**

ary Textures:

FT$^t$

h$^t$

P$^t$

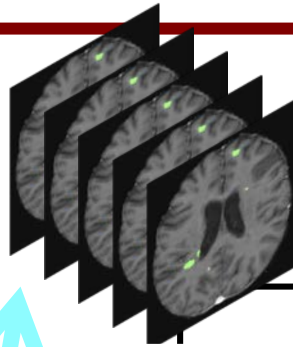
h

1

**Voxel-level**

$$
\begin{aligned}
\mathbb{E}(Y, X, \boldsymbol{\lambda}^v) = \sum_i \lambda^v_\phi \, \phi(y_i, \mathbf{x}_i) + \sum_{i,j \in N_i} \lambda^v_\varphi \, \varphi(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) \\
+ \sum_{i,j \in N_i} \lambda^v_\delta \, (1 - \delta(y_i, y_j)) + \sum_{i,(j,k) \in N_i} \lambda^v_\psi \, \psi(y_i, y_j, y_k, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) \\
+ \sum_{i,j \in N_i} \lambda^v_\Delta \, (1 - \Delta(y_i, y_j, y_k))
\end{aligned}
$$

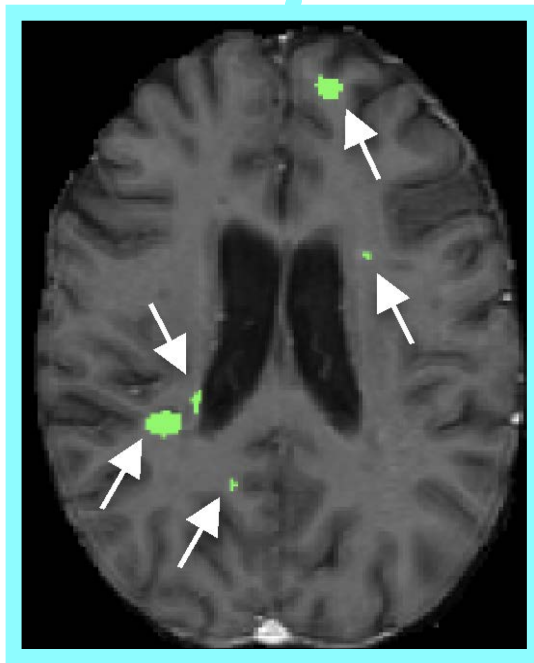
# HAT-CRF



output

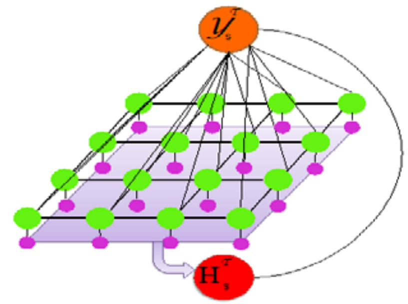
**Adaptive lesion level**

$$\mathbb{E}(Y^{\mathsf{p}}, X, \boldsymbol{\lambda}^{\mathsf{p}}) = \sum_{c\in C^{\mathsf{p}}} \lambda^{\mathsf{p}}_{\mathbf{y}_c} E^{v}_{c}(\mathbf{y}_c, X) + \sum_{k} \lambda^{\mathsf{p}}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}) + \sum_{k}\sum_{i} \lambda^{\mathsf{p}}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}).$$



4

**Higher order textural feature per each candidate**



Stationary Textures:
1- RIFT$^{t}$
2- spin$^{t}$
3- LBP$^{t}$

3

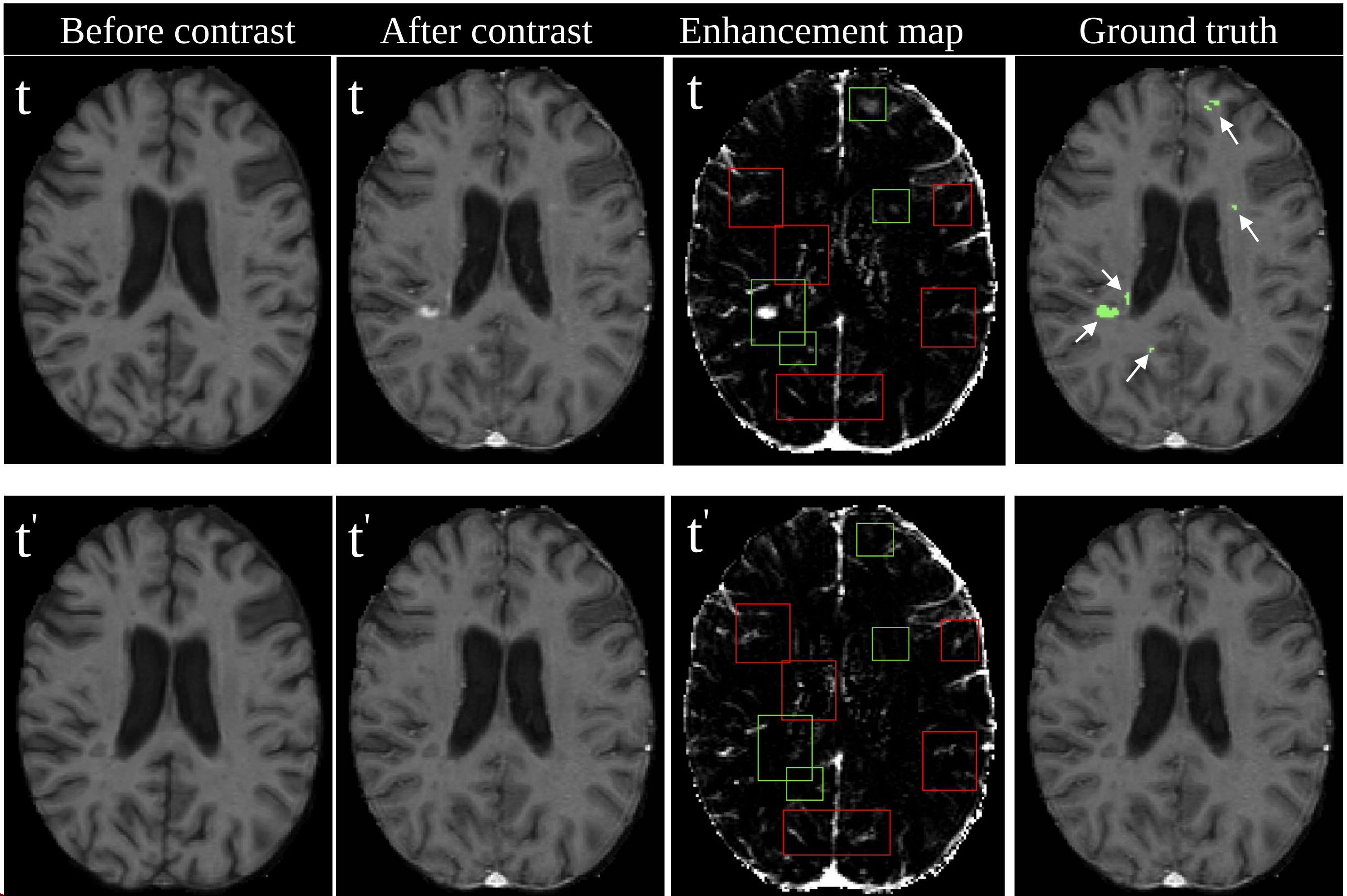
Candidate Lesion Detection

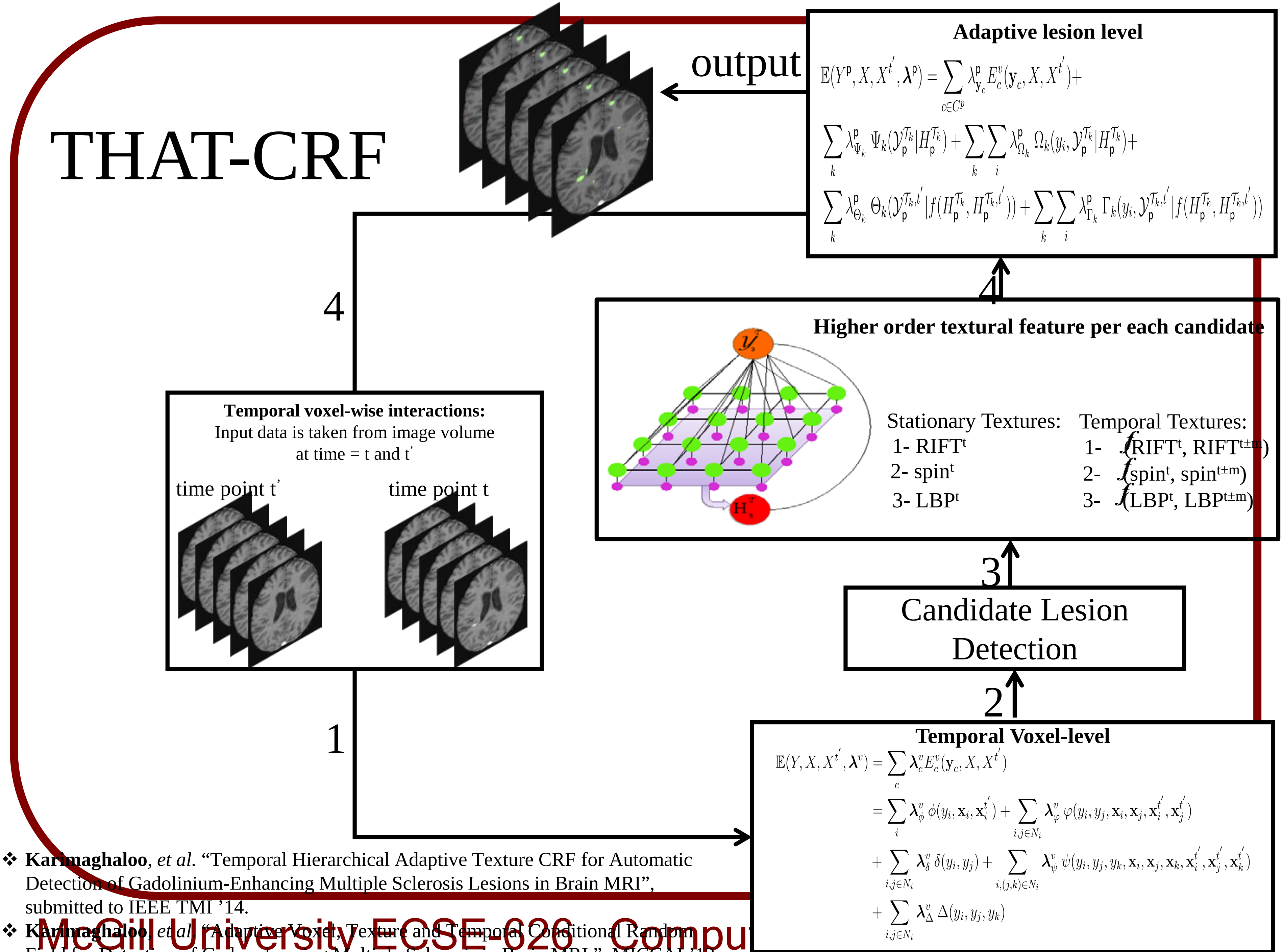
2

**Voxel-level**

$$\mathbb{E}(Y, X, \boldsymbol{\lambda}^{v}) = \sum_{i} \lambda^{v}_{\phi}\, \phi(y_i, \mathbf{x}_i) + \sum_{i,j\in N_i} \lambda^{v}_{\varphi}\, \varphi(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) + \sum_{i,j\in N_i} \lambda^{v}_{\delta}\, (1-\delta(y_i, y_j)) + \sum_{i,(j,k)\in N_i} \lambda^{v}_{\psi}\, \psi(y_i, y_j, y_k, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) + \sum_{i,j\in N_i} \lambda^{v}_{\Delta}\, (1-\Delta(y_i, y_j, y_k))$$

# Temporal Hierarchical Adaptive Texture CRF : THAT - CRF

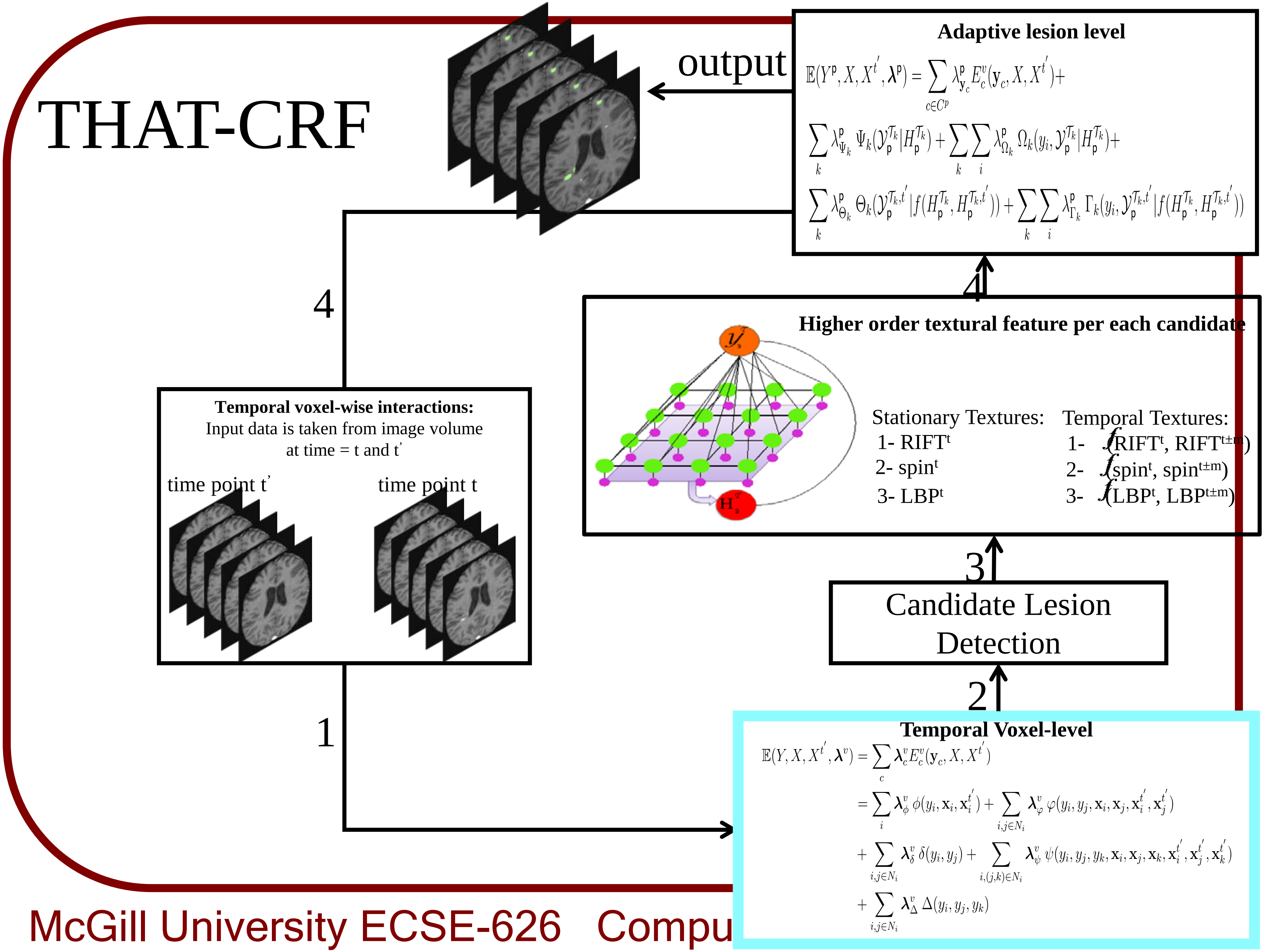




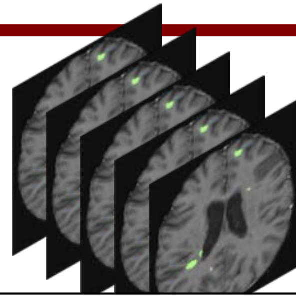


❖ **Karimaghaloo**, *et al.* “Temporal Hierarchical Adaptive Texture CRF for Automatic Detection of Gadolinium-Enhancing Multiple Sclerosis Lesions in Brain MRI”, submitted to IEEE TMI ’14.

❖ **Karimaghaloo**, *et al.* “Adaptive Voxel, Texture and Temporal Conditional Random Field for Detection of Gad- enhancing Multiple Sclerosis Lesions in Brain MRI ”, MICCAI ’13.


THAT-CRF
output
Adaptive lesion level
$\mathbb{E}(Y^{\mathrm{p}}, X, X^{t'}, \lambda^{\mathrm{p}}) = \sum_{c \in C^p} \lambda^{\mathrm{p}}_{\mathrm{y}_c} E^v_c(\mathrm{y}_c, X, X^{t'}) +$
$\sum_k \lambda^{\mathrm{p}}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathrm{p}} | H^{\mathcal{T}_k}_{\mathrm{p}}) + \sum_k \sum_i \lambda^{\mathrm{p}}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{\mathrm{p}} | H^{\mathcal{T}_k}_{\mathrm{p}}) +$
$\sum_k \lambda^{\mathrm{p}}_{\Theta_k} \Theta_k(\mathcal{Y}^{\mathcal{T}_k, t'}_{\mathrm{p}} | f(H^{\mathcal{T}_k}_{\mathrm{p}}, H^{\mathcal{T}_k, t'}_{\mathrm{p}})) + \sum_k \sum_i \lambda^{\mathrm{p}}_{\Gamma_k} \Gamma_k(y_i, \mathcal{Y}^{\mathcal{T}_k, t'}_{\mathrm{p}} | f(H^{\mathcal{T}_k}_{\mathrm{p}}, H^{\mathcal{T}_k, t'}_{\mathrm{p}}))$
4
4
Higher order textural feature per each candidate
$y^{\tau}_s$
$H^{\tau}_s$
Stationary Textures:
1- RIFT$^t$
2- spin$^t$
3- LBP$^t$
Temporal Textures:
1- $f$(RIFT$^t$, RIFT$^{t\pm m}$)
2- $f$(spin$^t$, spin$^{t\pm m}$)
3- $f$(LBP$^t$, LBP$^{t\pm m}$)
3
Temporal voxel-wise interactions:
Input data is taken from image volume
at time = t and t'
time point t'
time point t
Candidate Lesion Detection
2
1
Temporal Voxel-level
$\mathbb{E}(Y, X, X^{t'}, \lambda^v) = \sum_c \lambda^v_c E^v_c(\mathrm{y}_c, X, X^{t'})$
$= \sum_i \lambda^v_\phi \phi(y_i, \mathrm{x}_i, \mathrm{x}^{t'}_i) + \sum_{i,j \in N_i} \lambda^v_\varphi \varphi(y_i, y_j, \mathrm{x}_i, \mathrm{x}_j, \mathrm{x}^{t'}_i, \mathrm{x}^{t'}_j)$
$+ \sum_{i,j \in N_i} \lambda^v_\delta \delta(y_i, y_j) + \sum_{i,(j,k) \in N_i} \lambda^v_\psi \psi(y_i, y_j, y_k, \mathrm{x}_i, \mathrm{x}_j, \mathrm{x}_k, \mathrm{x}^{t'}_i, \mathrm{x}^{t'}_j, \mathrm{x}^{t'}_k)$
$+ \sum_{i,j \in N_i} \lambda^v_\Delta \Delta(y_i, y_j, y_k)$

# THAT-CRF



output

**Adaptive lesion level**

$$\mathbb{E}(Y^{\mathsf{p}}, X, X^{t'}, \boldsymbol{\lambda}^{\mathsf{p}}) = \sum_{c \in C^{p}} \lambda^{\mathsf{p}}_{\mathbf{y}_c} E^{v}_{c}(\mathbf{y}_c, X, X^{t'}) +$$

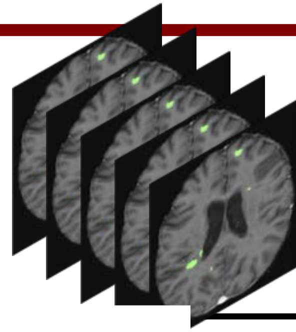
$$\sum_{k} \lambda^{\mathsf{p}}_{\Psi_k} \Psi_k(\mathcal{Y}_{\mathsf{p}}^{\mathcal{T}_k} | H_{\mathsf{p}}^{\mathcal{T}_k}) + \sum_{k} \sum_{i} \lambda^{\mathsf{p}}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}_{\mathsf{p}}^{\mathcal{T}_k} | H_{\mathsf{p}}^{\mathcal{T}_k}) +$$

$$\begin{aligned}
\mathbb{E}(Y, X, X^{t'}, \boldsymbol{\lambda}^{v}) &= \sum_{c} \boldsymbol{\lambda}^{v}_{c} E^{v}_{c}(\mathbf{y}_c, X, X^{t'}) \\
&= \sum_{i} \boldsymbol{\lambda}^{v}_{\phi}\, \phi(y_i, \mathbf{x}_i, \mathbf{x}_i^{t'}) + \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\varphi}\, \varphi(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_i^{t'}, \mathbf{x}_j^{t'}) \\
&+ \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\delta}\, \delta(y_i, y_j) + \sum_{i,(j,k) \in N_i} \boldsymbol{\lambda}^{v}_{\psi}\, \psi(y_i, y_j, y_k, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_i^{t'}, \mathbf{x}_j^{t'}, \mathbf{x}_k^{t'}) \\
&+ \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\Delta}\, \Delta(y_i, y_j, y_k)
\end{aligned}$$

1

**Temporal Voxel-level**

# THAT-CRF



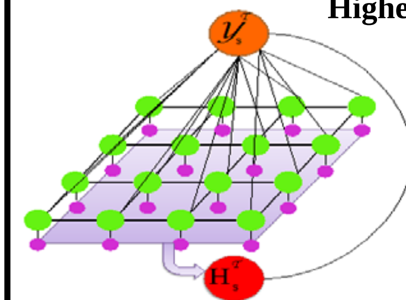
output

**Adaptive lesion level**

$$\mathbb{E}(Y^{p}, X, X^{t'}, \lambda^{p}) = \sum_{c\in C^{p}} \lambda^{p}_{y_c} E^{v}_{c}(y_c, X, X^{t'}) +$$
$$\sum_{k} \lambda^{p}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{p} | H^{\mathcal{T}_k}_{p}) + \sum_{k}\sum_{i} \lambda^{p}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{p} | H^{\mathcal{T}_k}_{p}) +$$
$$\sum_{k} \lambda^{p}_{\Theta_k} \Theta_k(\mathcal{Y}^{\mathcal{T}_k,t'}_{p} | f(H^{\mathcal{T}_k}_{p}, H^{\mathcal{T}_k,t'}_{p})) + \sum_{k}\sum_{i} \lambda^{p}_{\Gamma_k} \Gamma_k(y_i, \mathcal{Y}^{\mathcal{T}_k,t'}_{p} | f(H^{\mathcal{T}_k}_{p}, H^{\mathcal{T}_k,t'}_{p}))$$

4

**Higher order textural feature per each candidate**



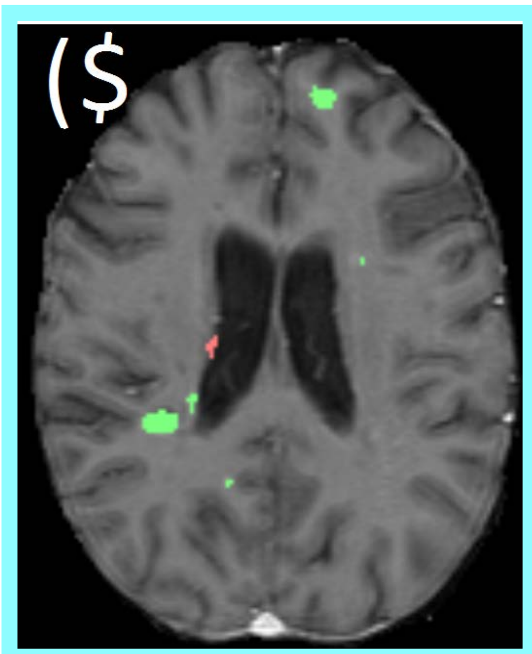

Stationary Textures:
1- $RIFT^{t}$
2- $spin^{t}$
3- $LBP^{t}$

Temporal Textures:
1- $f(RIFT^{t}, RIFT^{t\pm m})$
2- $f(spin^{t}, spin^{t\pm m})$
3- $f(LBP^{t}, LBP^{t\pm m})$

3

Candidate Lesion Detection

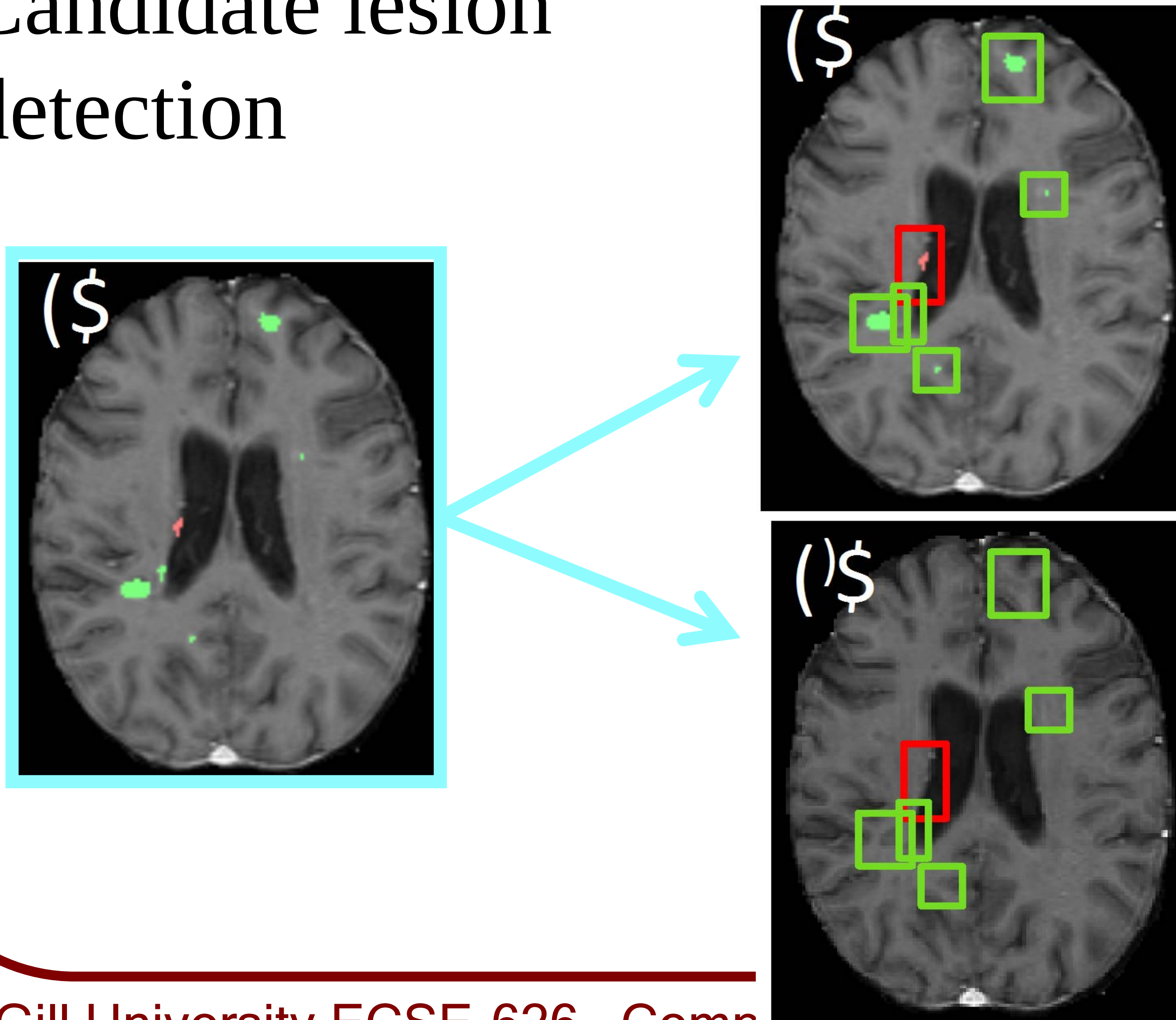
2



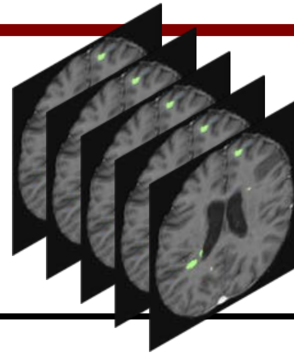
**Temporal Voxel-level**

$$\mathbb{E}(Y, X, X^{t'}, \lambda^{v}) = \sum_{c} \lambda^{v}_{c} E^{v}_{c}(y_c, X, X^{t'})$$
$$= \sum_{i} \lambda^{v}_{\phi}\, \phi(y_i, x_i, x^{t'}_{i}) + \sum_{i,j\in N_i} \lambda^{v}_{\varphi}\, \varphi(y_i, y_j, x_i, x_j, x^{t'}_{i}, x^{t'}_{j})$$
$$+ \sum_{i,j\in N_i} \lambda^{v}_{\delta}\, \delta(y_i, y_j) + \sum_{i,(j,k)\in N_i} \lambda^{v}_{\psi}\, \psi(y_i, y_j, y_k, x_i, x_j, x_k, x^{t'}_{i}, x^{t'}_{j}, x^{t'}_{k})$$
$$+ \sum_{i,j\in N_i} \lambda^{v}_{\Delta}\, \Delta(y_i, y_j, y_k)$$

# Candidate lesion detection



# THAT-CRF



output

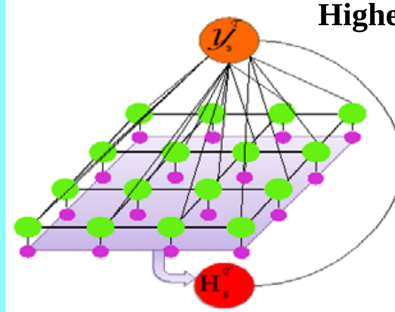
**Adaptive lesion level**

$$\mathbb{E}(Y^{\mathrm{p}}, X, X^{t'}, \lambda^{\mathrm{p}}) = \sum_{c \in C^{p}} \lambda^{\mathrm{p}}_{\mathrm{y}_c} E^{v}_{c}(\mathrm{y}_c, X, X^{t'}) +$$
$$\sum_{k} \lambda^{\mathrm{p}}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathrm{p}} | H^{\mathcal{T}_k}_{\mathrm{p}}) + \sum_{k} \sum_{i} \lambda^{\mathrm{p}}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{\mathrm{p}} | H^{\mathcal{T}_k}_{\mathrm{p}}) +$$
$$\sum_{k} \lambda^{\mathrm{p}}_{\Theta_k} \Theta_k(\mathcal{Y}^{\mathcal{T}_k,t'}_{\mathrm{p}} | f(H^{\mathcal{T}_k}_{\mathrm{p}}, H^{\mathcal{T}_k,t'}_{\mathrm{p}})) + \sum_{k} \sum_{i} \lambda^{\mathrm{p}}_{\Gamma_k} \Gamma_k(y_i, \mathcal{Y}^{\mathcal{T}_k,t'}_{\mathrm{p}} | f(H^{\mathcal{T}_k}_{\mathrm{p}}, H^{\mathcal{T}_k,t'}_{\mathrm{p}}))$$

4

4

**Higher order textural feature per each candidate**



Stationary Textures:
1- $\mathrm{RIFT}^{t}$
2- $\mathrm{spin}^{t}$
3- $\mathrm{LBP}^{t}$

Temporal Textures:
1- $f(\mathrm{RIFT}^{t}, \mathrm{RIFT}^{t\pm m})$
2- $f(\mathrm{spin}^{t}, \mathrm{spin}^{t\pm m})$
3- $f(\mathrm{LBP}^{t}, \mathrm{LBP}^{t\pm m})$

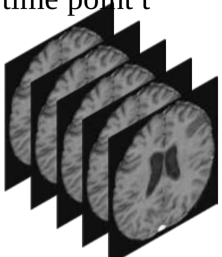
3

Candidate Lesion Detection

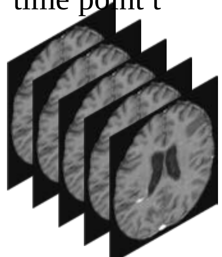
2

**Temporal voxel-wise interactions:**
Input data is taken from image volume at time = t and t'

time point t'



time point t

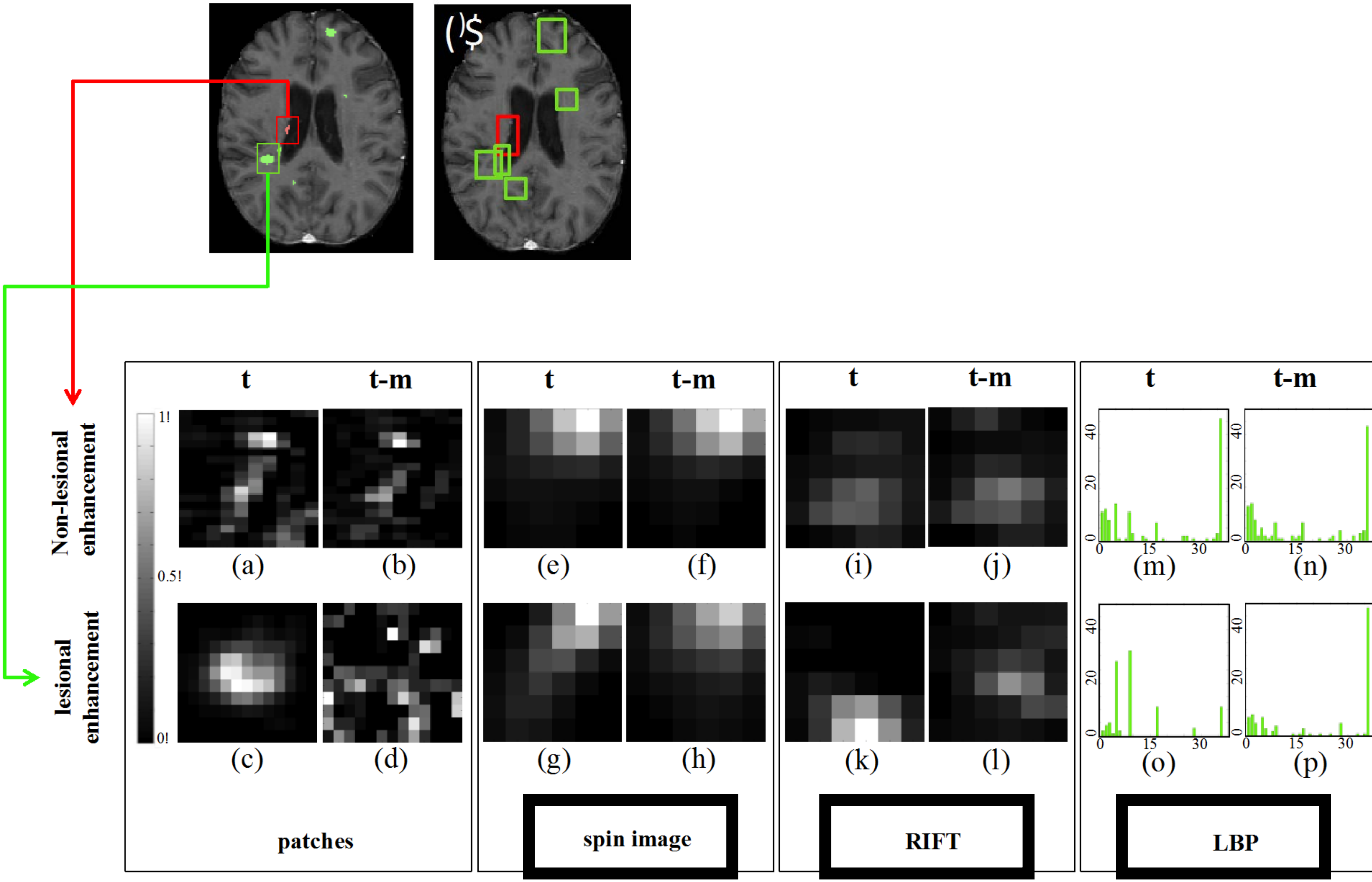


1

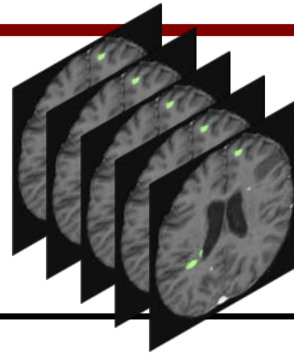
**Temporal Voxel-level**

$$\mathbb{E}(Y, X, X^{t'}, \boldsymbol{\lambda}^{v}) = \sum_{c} \boldsymbol{\lambda}^{v}_{c} E^{v}_{c}(\mathrm{y}_c, X, X^{t'})$$
$$= \sum_{i} \boldsymbol{\lambda}^{v}_{\phi} \phi(y_i, \mathrm{x}_i, \mathrm{x}^{t'}_i) + \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\varphi} \varphi(y_i, y_j, \mathrm{x}_i, \mathrm{x}_j, \mathrm{x}^{t'}_i, \mathrm{x}^{t'}_j)$$
$$+ \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\delta} \delta(y_i, y_j) + \sum_{i,(j,k) \in N_i} \boldsymbol{\lambda}^{v}_{\psi} \psi(y_i, y_j, y_k, \mathrm{x}_i, \mathrm{x}_j, \mathrm{x}_k, \mathrm{x}^{t'}_i, \mathrm{x}^{t'}_j, \mathrm{x}^{t'}_k)$$
$$+ \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\Delta} \Delta(y_i, y_j, y_k)$$

# Temporal Textures



# THAT-CRF



output

**Adaptive lesion level**

$$\mathbb{E}(Y^{\mathsf{p}}, X, X^{t'}, \boldsymbol{\lambda}^{\mathsf{p}}) = \sum_{c \in C^{p}} \lambda^{\mathsf{p}}_{\mathbf{y}_c} E^{v}_{c}(\mathbf{y}_c, X, X^{t'}) +$$

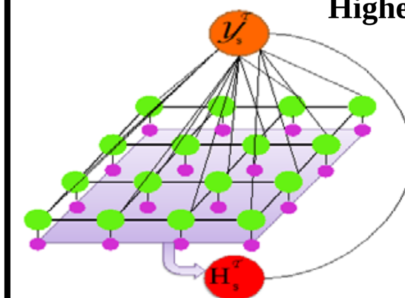
$$\sum_{k} \lambda^{\mathsf{p}}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}) + \sum_{k} \sum_{i} \lambda^{\mathsf{p}}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}) +$$

$$\sum_{k} \lambda^{\mathsf{p}}_{\Theta_k} \Theta_k(\mathcal{Y}^{\mathcal{T}_k, t'}_{\mathsf{p}} | f(H^{\mathcal{T}_k}_{\mathsf{p}}, H^{\mathcal{T}_k, t'}_{\mathsf{p}})) + \sum_{k} \sum_{i} \lambda^{\mathsf{p}}_{\Gamma_k} \Gamma_k(y_i, \mathcal{Y}^{\mathcal{T}_k, t'}_{\mathsf{p}} | f(H^{\mathcal{T}_k}_{\mathsf{p}}, H^{\mathcal{T}_k, t'}_{\mathsf{p}}))$$

4

4

**Higher order textural feature per each candidate**



Stationary Textures:

1- RIFT$^{t}$

2- spin$^{t}$

3- LBP$^{t}$

Temporal Textures:

1- $f$(RIFT$^{t}$, RIFT$^{t\pm m}$)

2- $f$(spin$^{t}$, spin$^{t\pm m}$)

3- $f$(LBP$^{t}$, LBP$^{t\pm m}$)

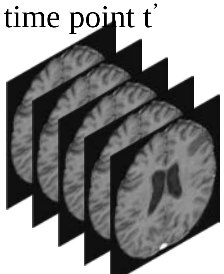
3

Candidate Lesion Detection

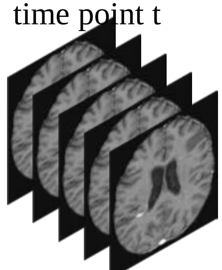
2

**Temporal voxel-wise interactions:**
Input data is taken from image volume at time = t and t'

time point t'



time point t



1

**Temporal Voxel-level**

$$\mathbb{E}(Y, X, X^{t'}, \boldsymbol{\lambda}^{v}) = \sum_{c} \boldsymbol{\lambda}^{v}_{c} E^{v}_{c}(\mathbf{y}_c, X, X^{t'})$$

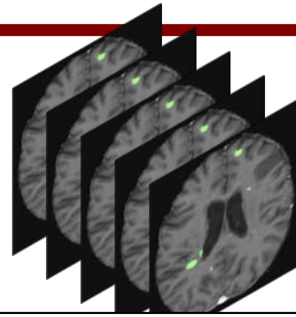
$$= \sum_{i} \boldsymbol{\lambda}^{v}_{\phi} \phi(y_i, \mathbf{x}_i, \mathbf{x}^{t'}_i) + \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\varphi} \varphi(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}^{t'}_i, \mathbf{x}^{t'}_j)$$

$$+ \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\delta} \delta(y_i, y_j) + \sum_{i,(j,k) \in N_i} \boldsymbol{\lambda}^{v}_{\psi} \psi(y_i, y_j, y_k, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k, \mathbf{x}^{t'}_i, \mathbf{x}^{t'}_j, \mathbf{x}^{t'}_k)$$

$$+ \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\Delta} \Delta(y_i, y_j, y_k)$$

**Adaptive lesion level**

# THAT-CRF



output

$$\mathbb{E}(Y^{\mathsf{p}}, X, X^{t'}, \boldsymbol{\lambda}^{\mathsf{p}}) = \sum_{c \in C^{p}} \lambda^{\mathsf{p}}_{\mathbf{y}_c} E^{v}_{c}(\mathbf{y}_c, X, X^{t'}) +$$

$$\sum_{k} \lambda^{\mathsf{p}}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}) + \sum_{k} \sum_{i} \lambda^{\mathsf{p}}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{\mathsf{p}} | H^{\mathcal{T}_k}_{\mathsf{p}}) +$$

$$\sum_{k} \lambda^{\mathsf{p}}_{\Theta_k} \Theta_k(\mathcal{Y}^{\mathcal{T}_k, t'}_{\mathsf{p}} | f(H^{\mathcal{T}_k}_{\mathsf{p}}, H^{\mathcal{T}_k, t'}_{\mathsf{p}})) + \sum_{k} \sum_{i} \lambda^{\mathsf{p}}_{\Gamma_k} \Gamma_k(y_i, \mathcal{Y}^{\mathcal{T}_k, t'}_{\mathsf{p}} | f(H^{\mathcal{T}_k}_{\mathsf{p}}, H^{\mathcal{T}_k, t'}_{\mathsf{p}}))$$

idate

es:

$T^{t\pm m})$

$\pm m)$

$t\pm m)$

1

2

**Temporal Voxel-level**

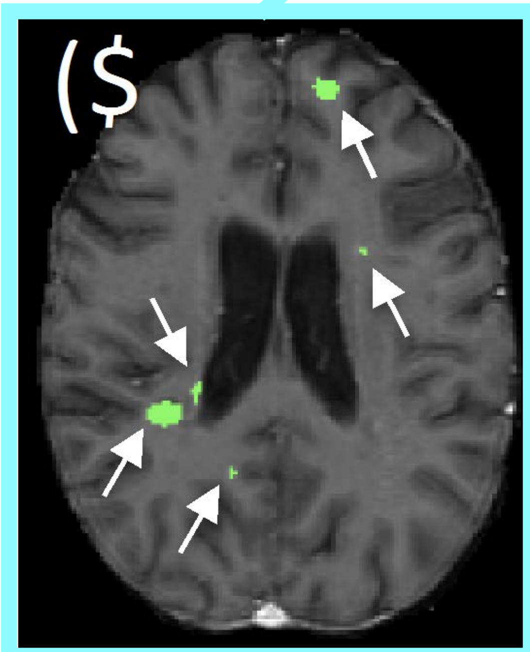
$$\mathbb{E}(Y, X, X^{t'}, \boldsymbol{\lambda}^{v}) = \sum_{c} \boldsymbol{\lambda}^{v}_{c} E^{v}_{c}(\mathbf{y}_c, X, X^{t'})$$

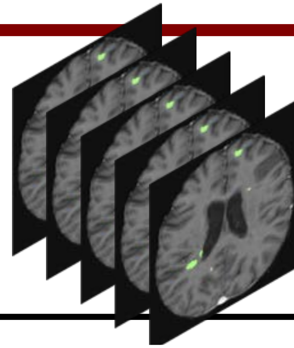
$$= \sum_{i} \boldsymbol{\lambda}^{v}_{\phi}\, \phi(y_i, \mathbf{x}_i, \mathbf{x}^{t'}_i) + \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\varphi}\, \varphi(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}^{t'}_i, \mathbf{x}^{t'}_j)$$

$$+ \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\delta}\, \delta(y_i, y_j) + \sum_{i,(j,k) \in N_i} \boldsymbol{\lambda}^{v}_{\psi}\, \psi(y_i, y_j, y_k, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k, \mathbf{x}^{t'}_i, \mathbf{x}^{t'}_j, \mathbf{x}^{t'}_k)$$

$$+ \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\Delta}\, \Delta(y_i, y_j, y_k)$$

# THAT-CRF



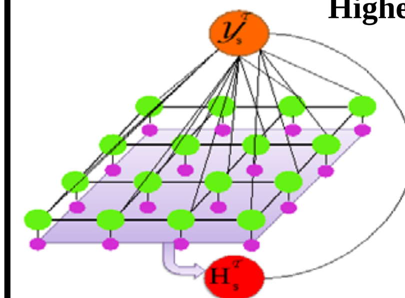
output

**Adaptive lesion level**

$$\mathbb{E}(Y^{\mathrm{p}}, X, X^{t'}, \lambda^{\mathrm{p}}) = \sum_{c \in C^{p}} \lambda^{\mathrm{p}}_{\mathbf{y}_c} E^{v}_{c}(\mathbf{y}_c, X, X^{t'}) +$$
$$\sum_{k} \lambda^{\mathrm{p}}_{\Psi_k} \Psi_k(\mathcal{Y}^{\mathcal{T}_k}_{\mathrm{p}} | H^{\mathcal{T}_k}_{\mathrm{p}}) + \sum_{k}\sum_{i} \lambda^{\mathrm{p}}_{\Omega_k} \Omega_k(y_i, \mathcal{Y}^{\mathcal{T}_k}_{\mathrm{p}} | H^{\mathcal{T}_k}_{\mathrm{p}}) +$$
$$\sum_{k} \lambda^{\mathrm{p}}_{\Theta_k} \Theta_k(\mathcal{Y}^{\mathcal{T}_k, t'}_{\mathrm{p}} | f(H^{\mathcal{T}_k}_{\mathrm{p}}, H^{\mathcal{T}_k, t'}_{\mathrm{p}})) + \sum_{k}\sum_{i} \lambda^{\mathrm{p}}_{\Gamma_k} \Gamma_k(y_i, \mathcal{Y}^{\mathcal{T}_k, t'}_{\mathrm{p}} | f(H^{\mathcal{T}_k}_{\mathrm{p}}, H^{\mathcal{T}_k, t'}_{\mathrm{p}}))$$

4

**Higher order textural feature per each candidate**



Stationary Textures:
1- $\mathrm{RIFT}^{t}$
2- $\mathrm{spin}^{t}$
3- $\mathrm{LBP}^{t}$

Temporal Textures:
1- $f(\mathrm{RIFT}^{t}, \mathrm{RIFT}^{t\pm m})$
2- $f(\mathrm{spin}^{t}, \mathrm{spin}^{t\pm m})$
3- $f(\mathrm{LBP}^{t}, \mathrm{LBP}^{t\pm m})$

3

Candidate Lesion Detection

2

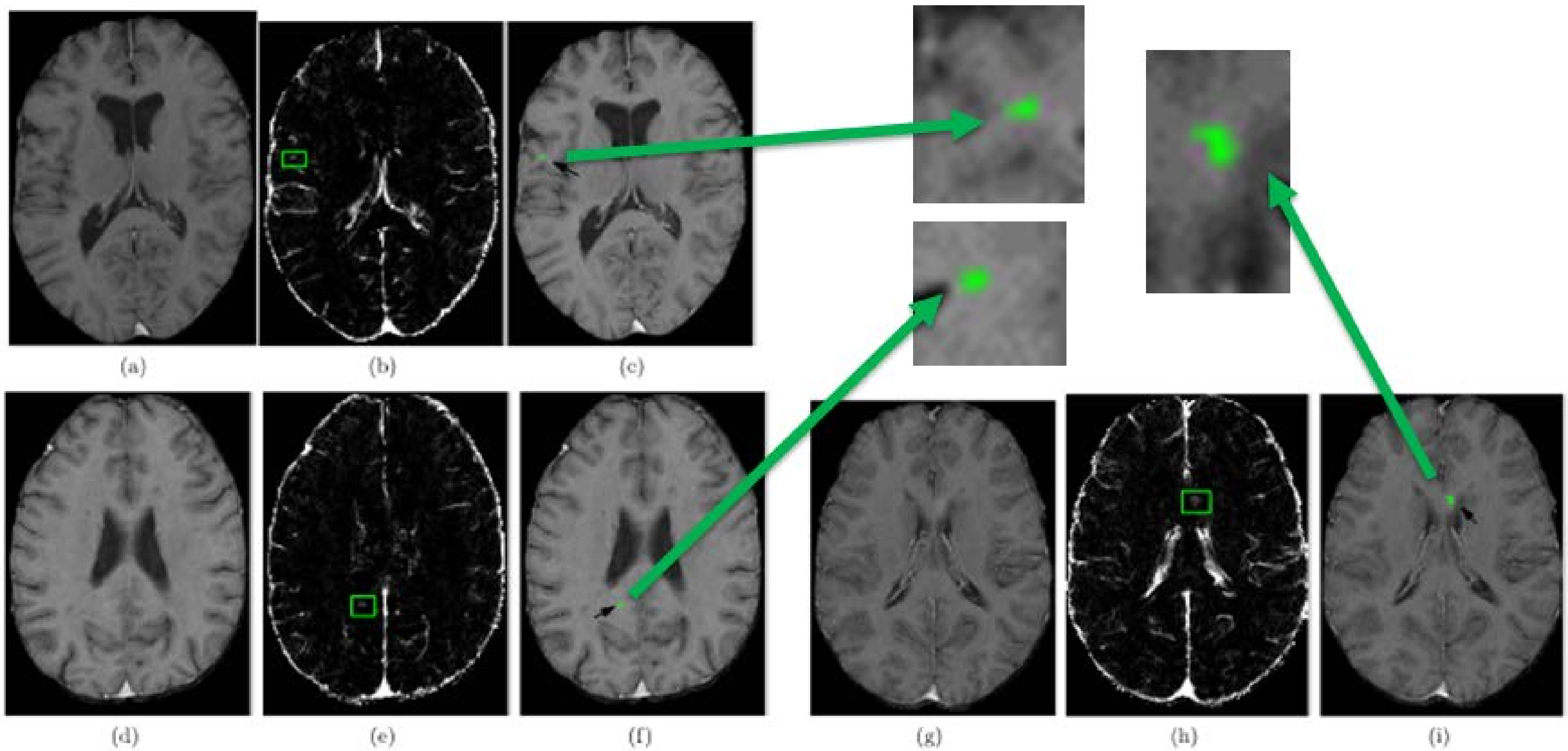
**Temporal Voxel-level**

$$\mathbb{E}(Y, X, X^{t'}, \boldsymbol{\lambda}^{v}) = \sum_{c} \boldsymbol{\lambda}^{v}_{c} E^{v}_{c}(\mathbf{y}_c, X, X^{t'})$$
$$= \sum_{i} \boldsymbol{\lambda}^{v}_{\phi}\, \phi(y_i, \mathbf{x}_i, \mathbf{x}^{t'}_i) + \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\varphi}\, \varphi(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}^{t'}_i, \mathbf{x}^{t'}_j)$$
$$+ \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\delta}\, \delta(y_i, y_j) + \sum_{i,(j,k) \in N_i} \boldsymbol{\lambda}^{v}_{\psi}\, \psi(y_i, y_j, y_k, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k, \mathbf{x}^{t'}_i, \mathbf{x}^{t'}_j, \mathbf{x}^{t'}_k)$$
$$+ \sum_{i,j \in N_i} \boldsymbol{\lambda}^{v}_{\Delta}\, \Delta(y_i, y_j, y_k)$$

# “Big Clinical Data” and Experiments

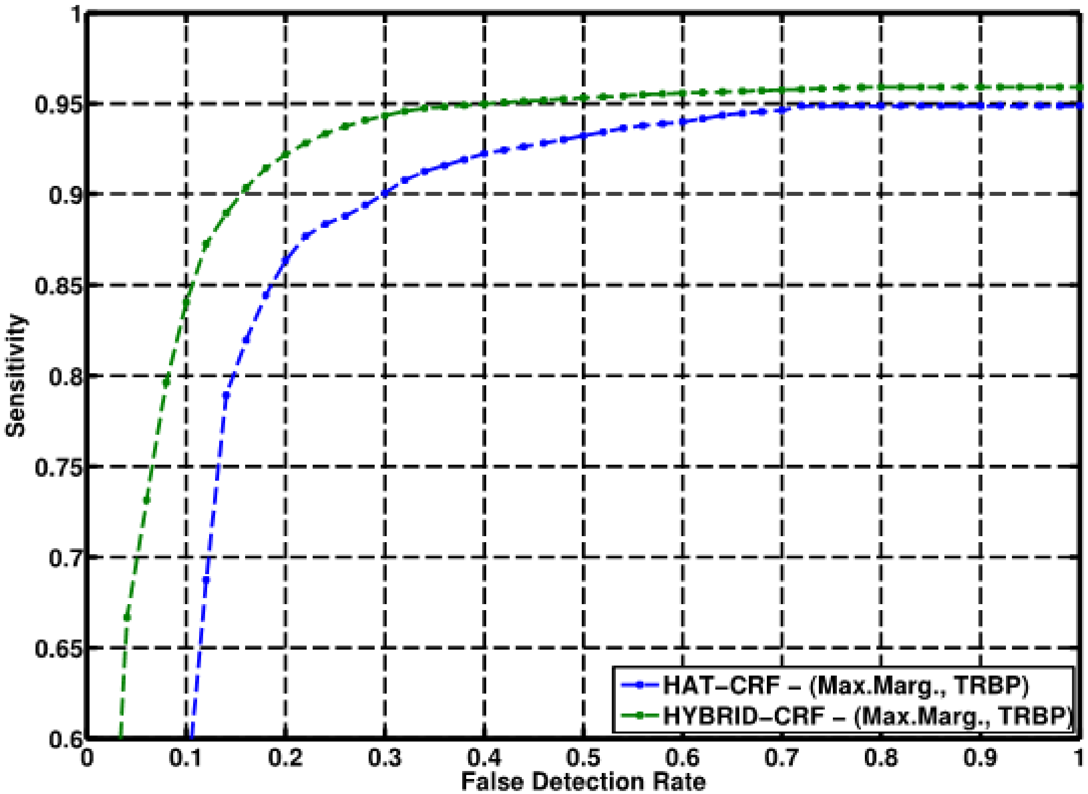
- Trained on a large multi-center clinical trial consisting of 2380 scans from 247 different centers.
- Tested on two separate clinical trials consisting of 813 scans from 27 centers and 2770 scans from 142 centers.
- Various number of lesions per scan (0 to $100^{+}$) and with different sizes (3 to $100^{+}$) at various stage of the disease.
- Ground truth: manual labeling from two trained experts with consensus (might contain errors).
- Further assessment is performed in some cases by an expert MS neuroradiologist.

# Qualitative Results



# Quantitative Results

THAT vs HAT ROC Curve




Sensitivity:

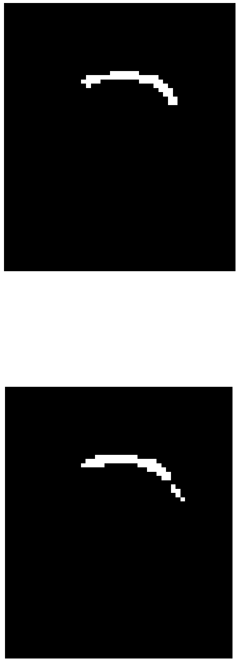
TP/(TP+FN)

False Detection Rate:

FP/(TP+FP)

# Myocardiom segmentation in heart

Z. Karimaghaloo, et. al. “Hierarchical Conditional Random Fields for Myocardium Detection”. MICCAI Workshop 2012, Nice, France.



(h) Complete HCRF

**Fig. 2.** The performance of the different components of the proposed HCRF classifier shown for two example slices from pig data. The shown images are: the delayed MRI (a), the delayed MRI only showing the myocardium tissue (b), Nyul intensity normalization of the myocardium (c), the manual labels of myocardium infarction (d), the classification results using only the unary potential (e), the classification result using the complete CRF model (Unary, pairwise and triplet potentials) at the first level (f), the classification result of the second level with using only the unary potential (g), the classification result of the complete HCRF model (h).