Forests and Jungles for Medical Image Analysis

A. Criminisi

Talk overview

- A brief introduction to machine learning
- Decision forests and jungles
- Applications in medical image analysis
  - Anatomy localization
  - Spine detection
  - Brain tumour segmentation
  - Learned image super-resolution
  - Quantifying progression of multiple sclerosis

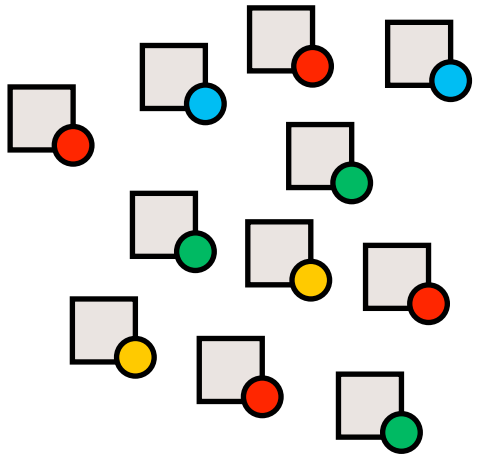
Talk overview

- **A brief introduction to machine learning**
- Decision forests and jungles
- Applications in medical image analysis
  - Anatomy localization
  - Spine detection
  - Brain tumour segmentation
  - Learned image super-resolution
  - Quantifying progression of multiple sclerosis

Supervised machine learning (classification)

Training phase (usually offline)

Training data set

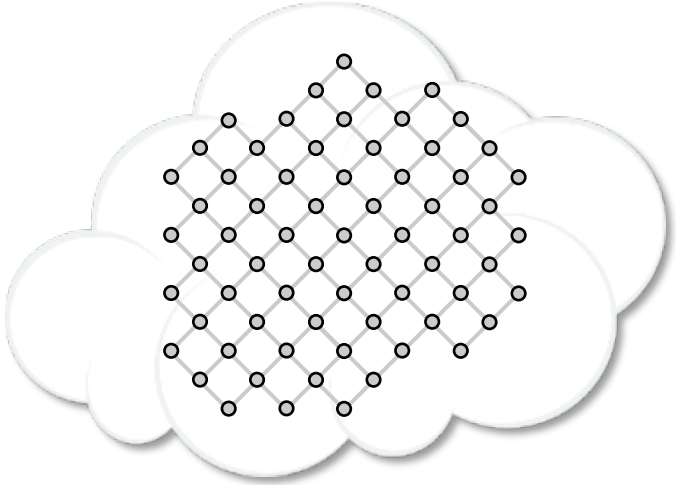


measurements (features)
&
associated ‘class’ labels

(colors denote class labels)

Training algorithm

Learned model (classifier)

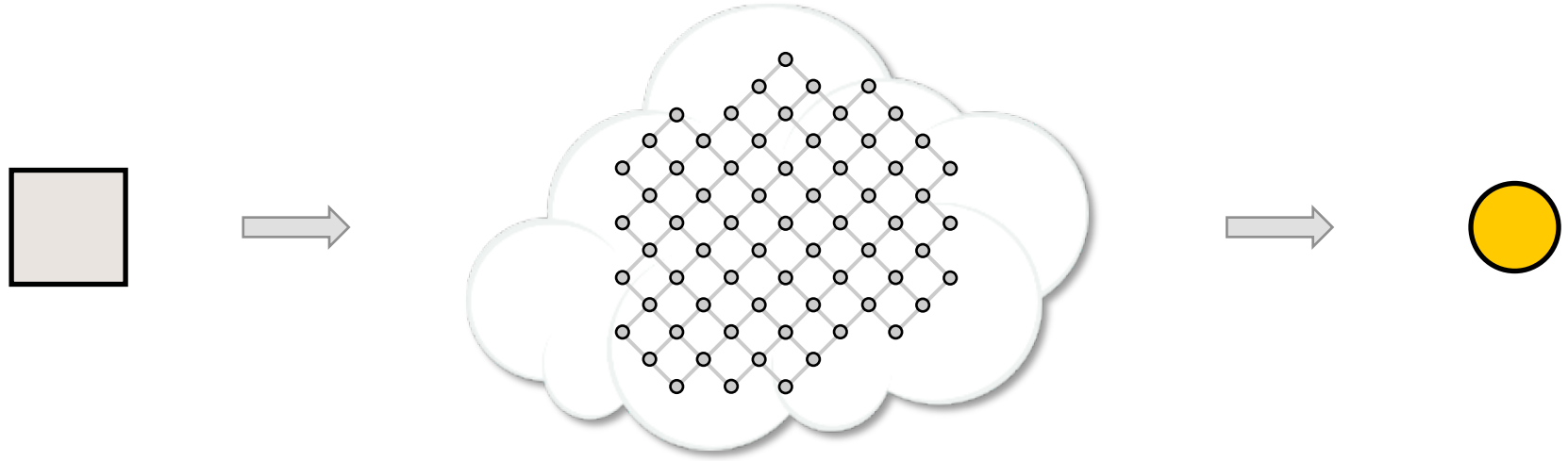


structure & parameters

Supervised machine learning (classification)

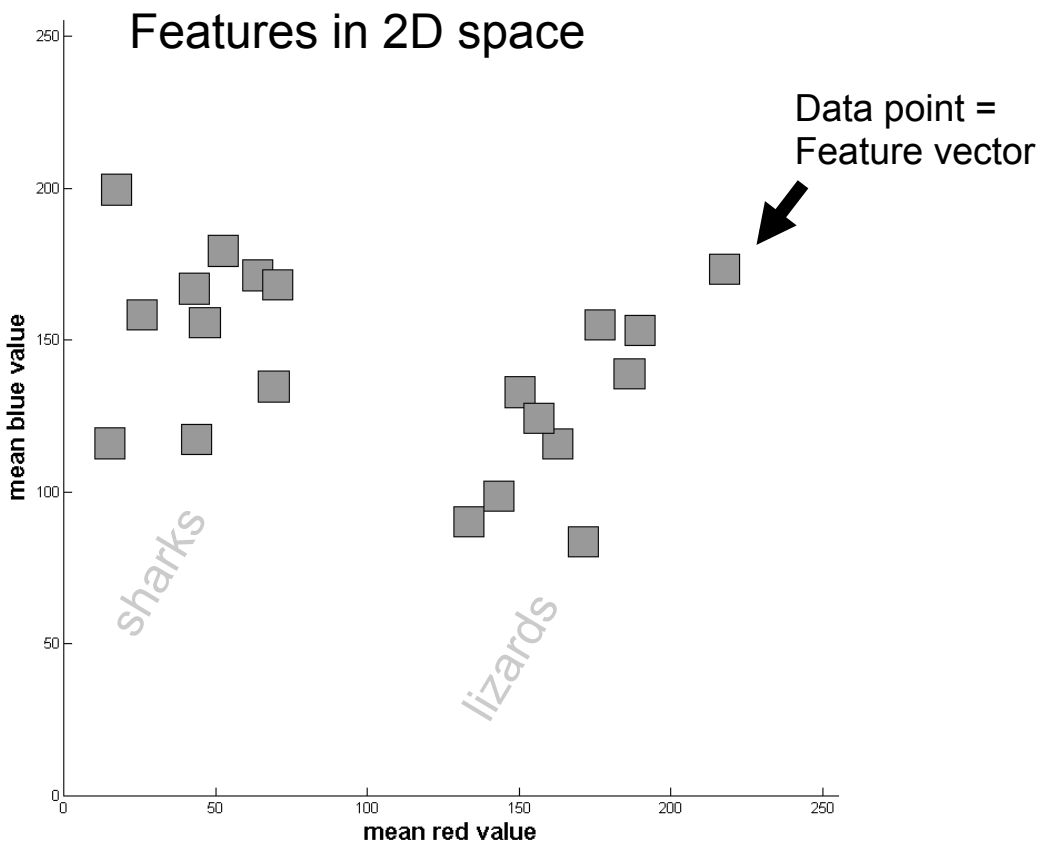
Test phase (run time, online)

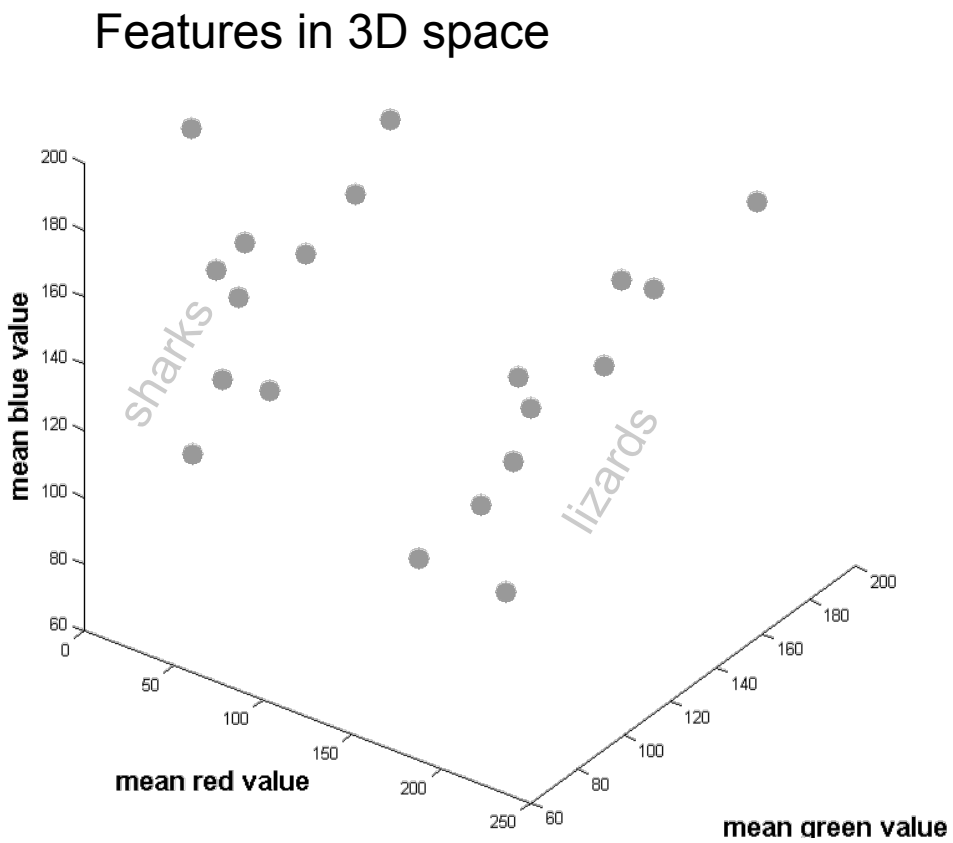
Input test data point ***Learned model (classifier)*** ***Output***

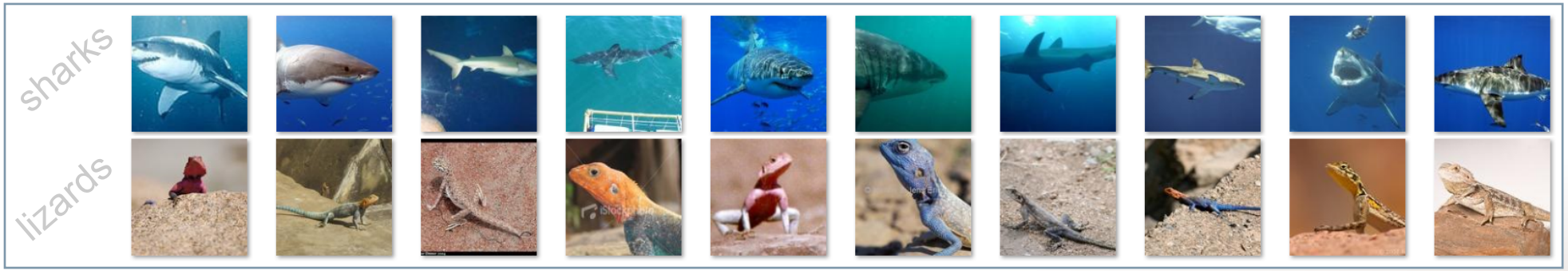


measurements (features) only structure + parameters predicted class label

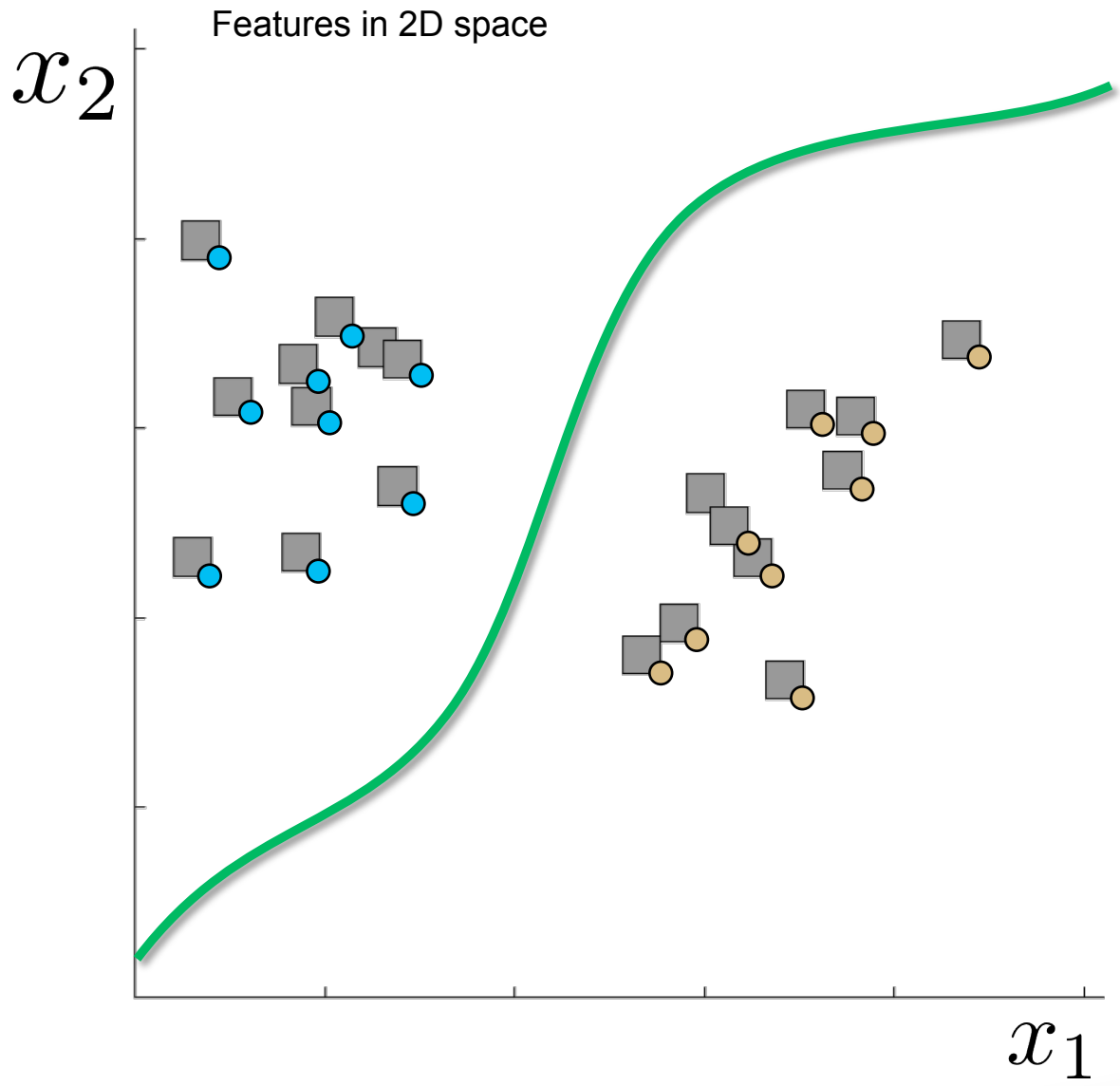
Data representation, feature vectors and data points







Data representation, feature vectors and data points



induction

Application: Kinect body part recognition

Task: assigning body part labels to each pixel in Kinect-acquired depth videos

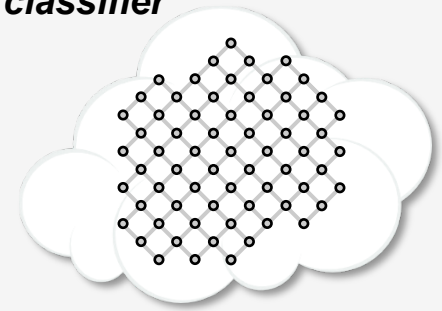
Input test depth image

Body part segmentation

image measurements made relative to pixel

e.g. depth, color, neighbors

classifier



per-pixel prediction of class label

J. Shotton, R. Girshick, A. Fitzgibbon, T. Sharp, M. Cook, M. Finocchio, R. Moore, P. Kohli, A. Criminisi, A. Kipman, and A. Blake, **Efficient Human Pose Estimation from Single Depth Images**, in *Trans. PAMI*, IEEE, 2012

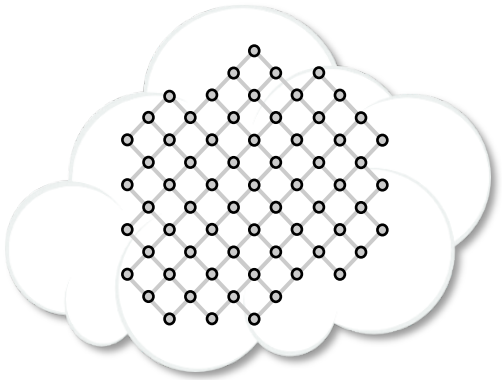
Talk overview

- A brief introduction to machine learning
- **Decision forests** and jungles
- Applications in medical image analysis
  - Anatomy localization
  - Spine detection
  - Brain tumour segmentation
  - Learned image super-resolution
  - Quantifying progression of multiple sclerosis

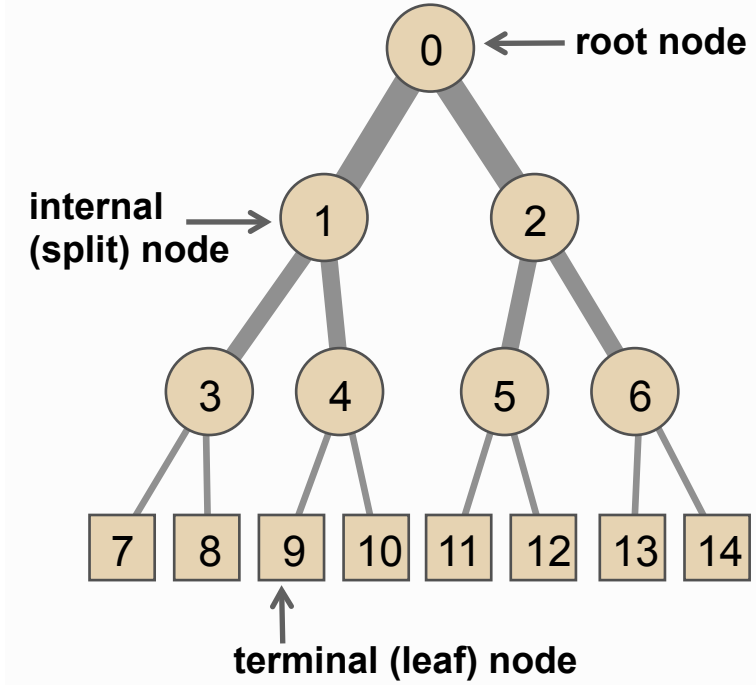
A. Criminisi and J. Shotton, **Decision Forests for Computer Vision and Medical Image Analysis**, Springer, February 2013

Decision trees

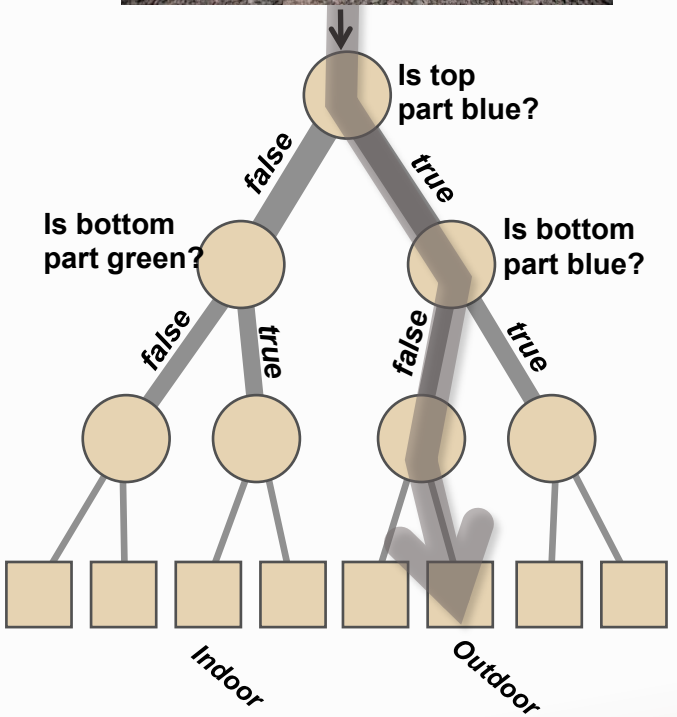
A general learned predictor



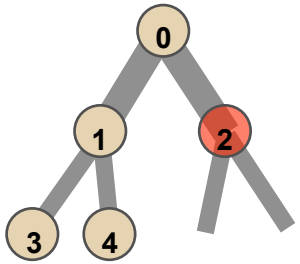
A general (binary) tree structure



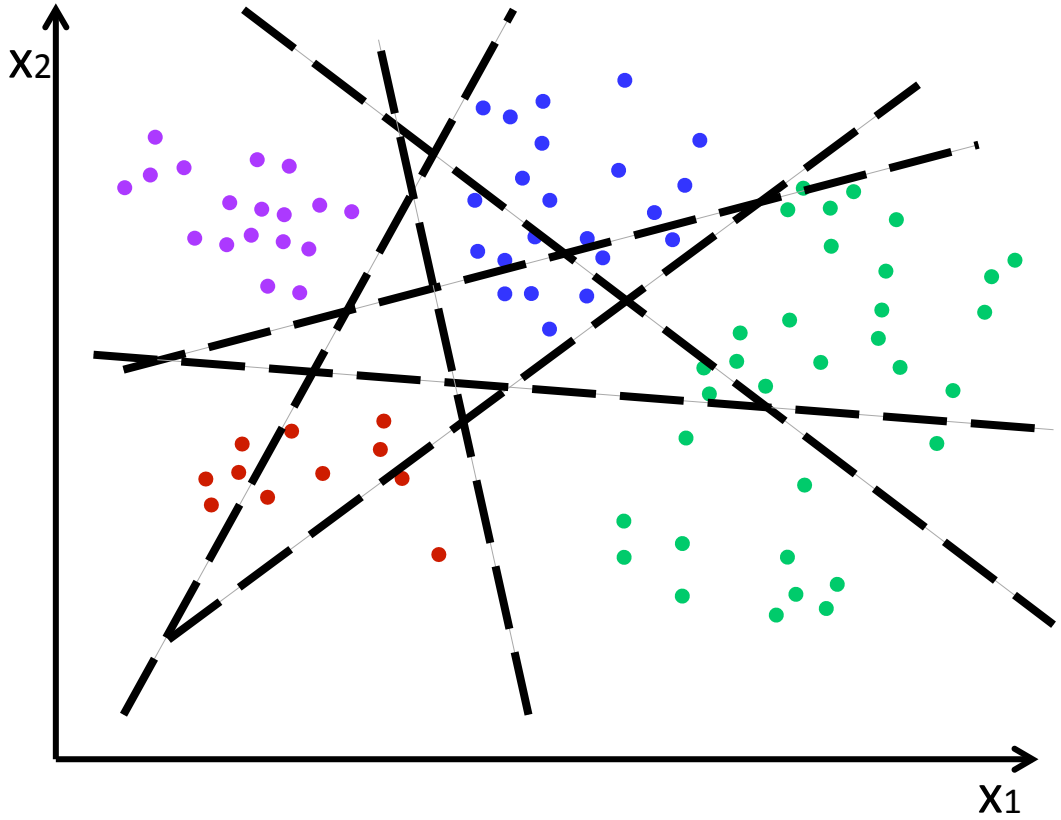
A decision tree



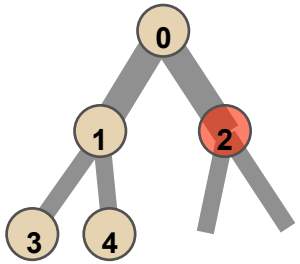
Toy tree learning example



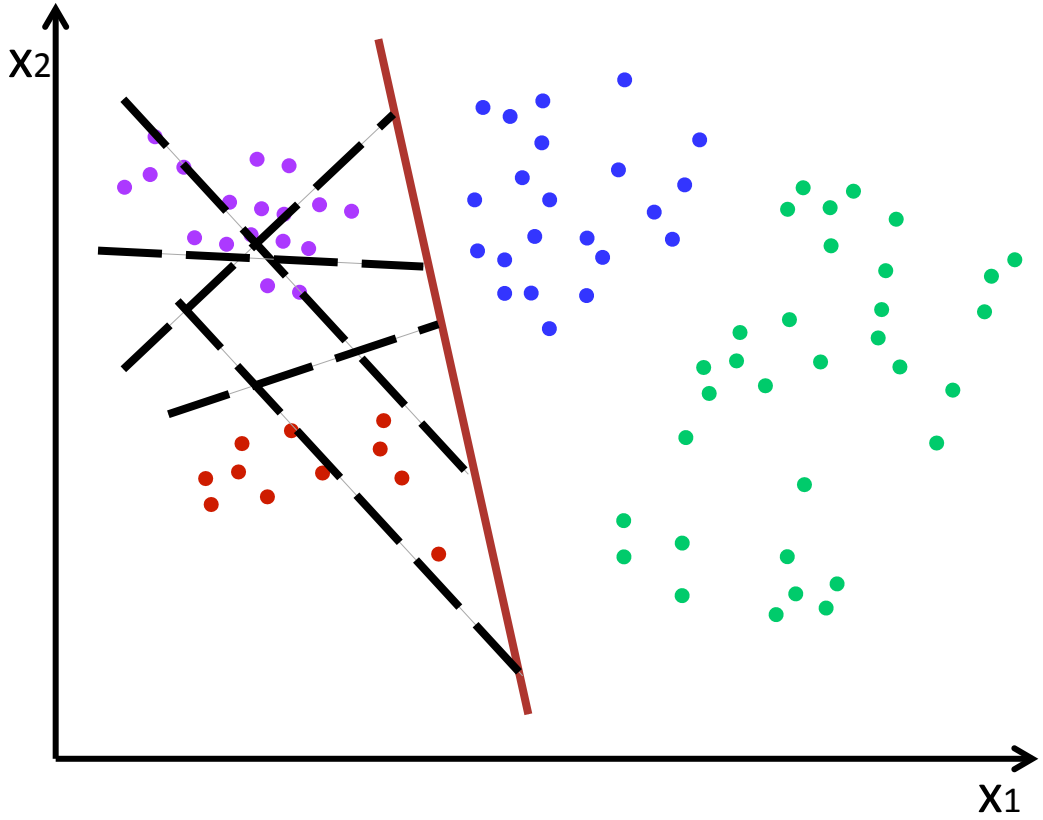
- Try several lines, chosen at random
- Keep line that best separates data
  - information gain
- Recurse



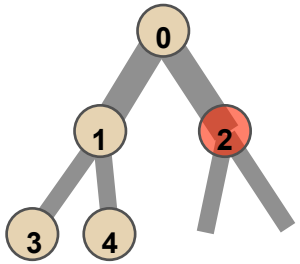
Toy tree learning example



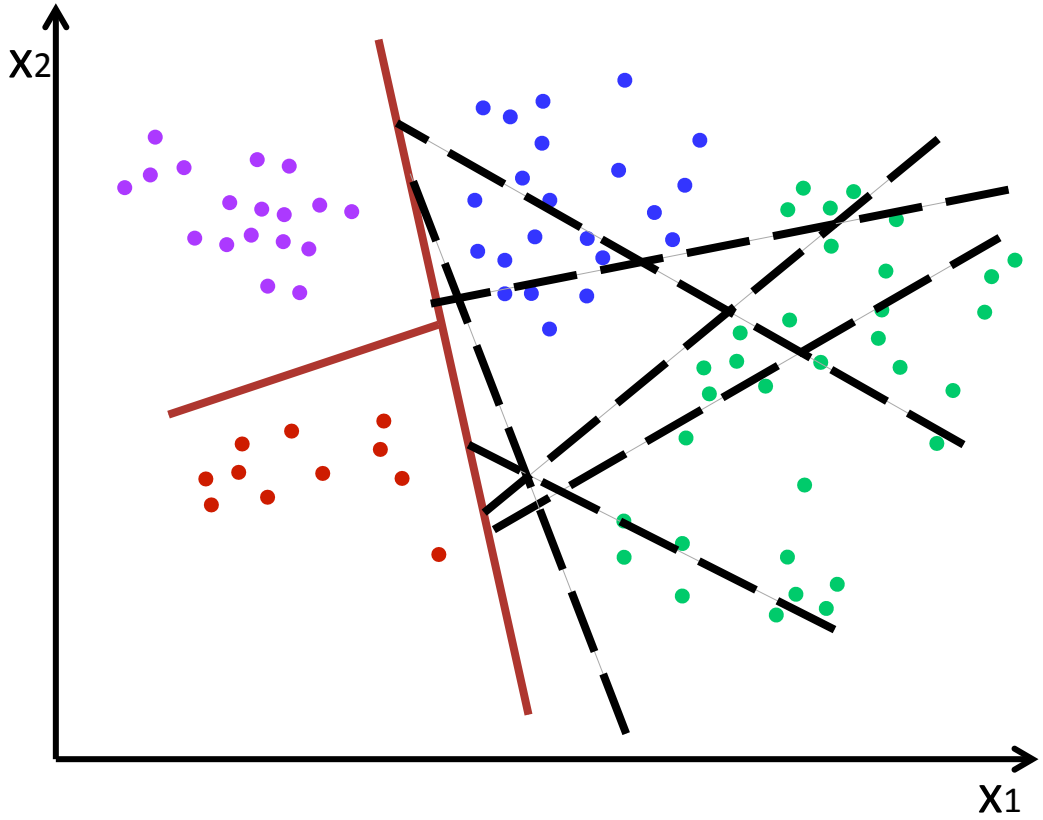
- Try several lines, chosen at random
- Keep line that best separates data
  - information gain
- Recurse



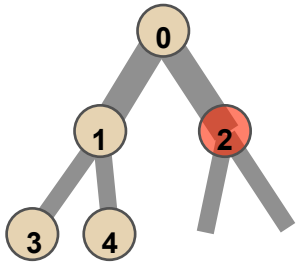
Toy tree learning example



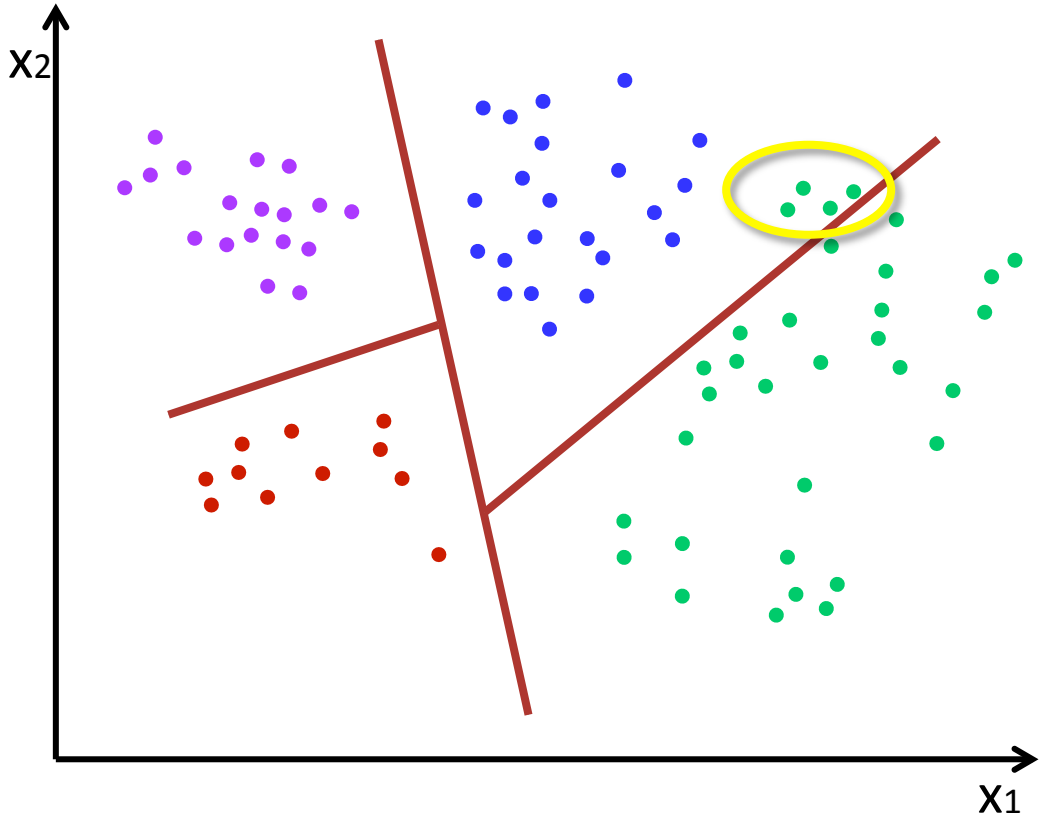
- Try several lines, chosen at random
- Keep line that best separates data
  - information gain
- Recurse



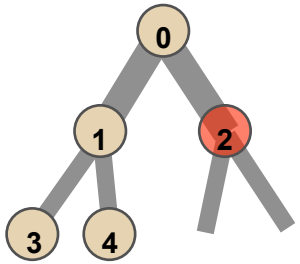
Toy tree learning example



- Try several lines, chosen at random
- Keep line that best separates data
  - information gain
- Recurse



Training objective function

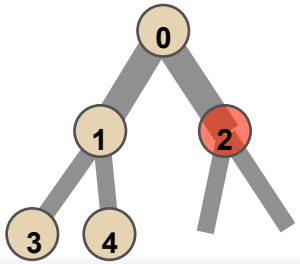


- Used to decide which candidate split function is best

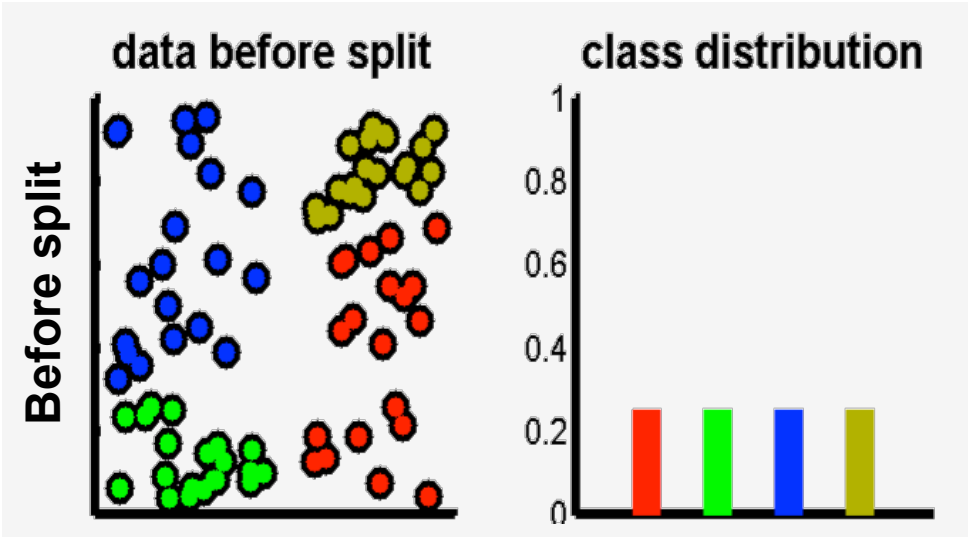
- Typically an “information gain” – a very general and flexible formulation

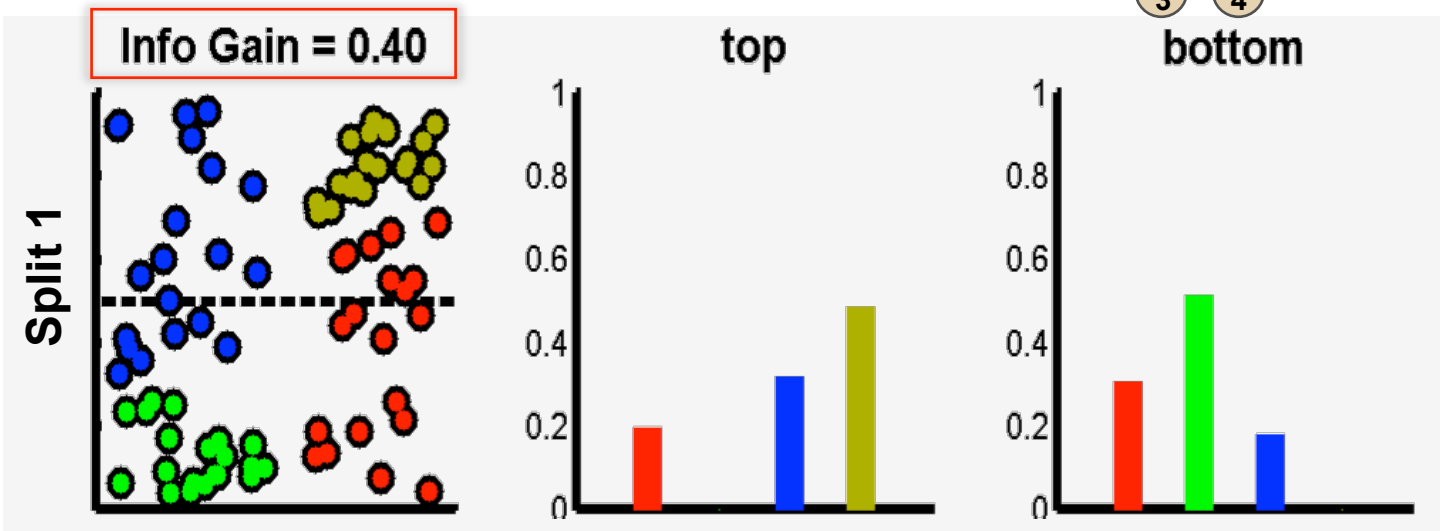
$$I = \underbrace{H(\mathcal{S}_j)}_{\text{entropy of examples at parent node}} - \sum_{i=\mathrm{L,R}} \underbrace{\frac{|\mathcal{S}_j^i|}{|\mathcal{S}_j|}}_{\text{weighting left/right children}} \underbrace{H(\mathcal{S}_j^i)}_{\text{entropy of examples at child nodes}}$$

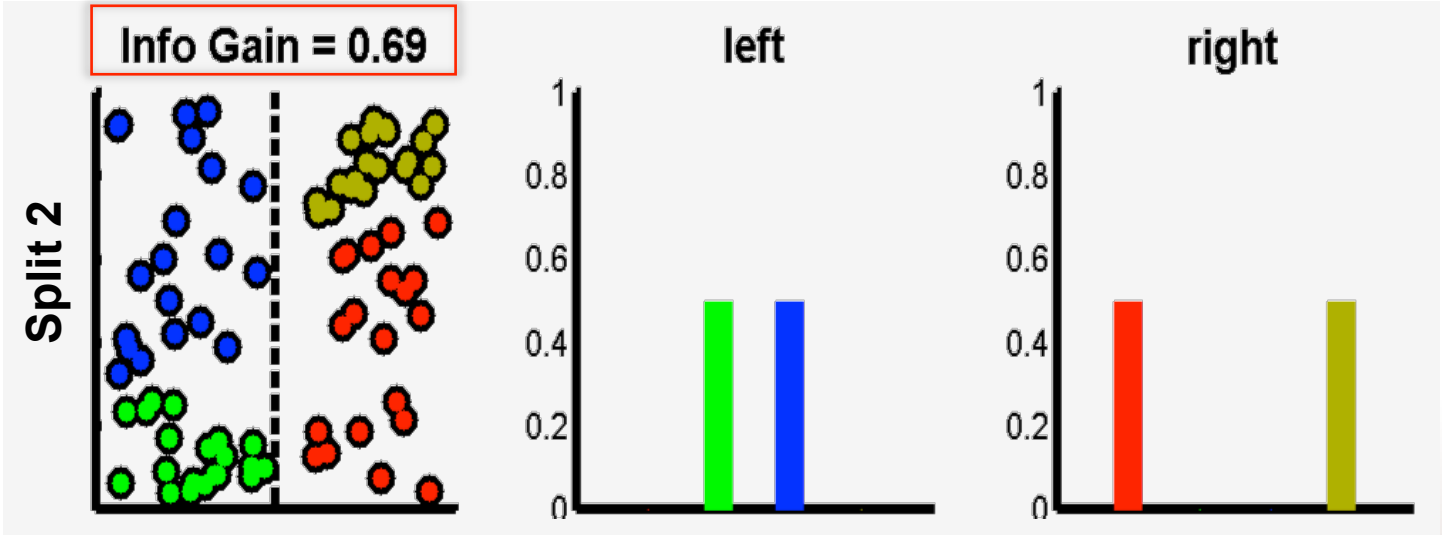
Training objective function



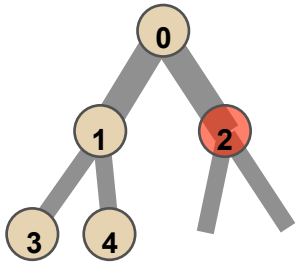
Information gain

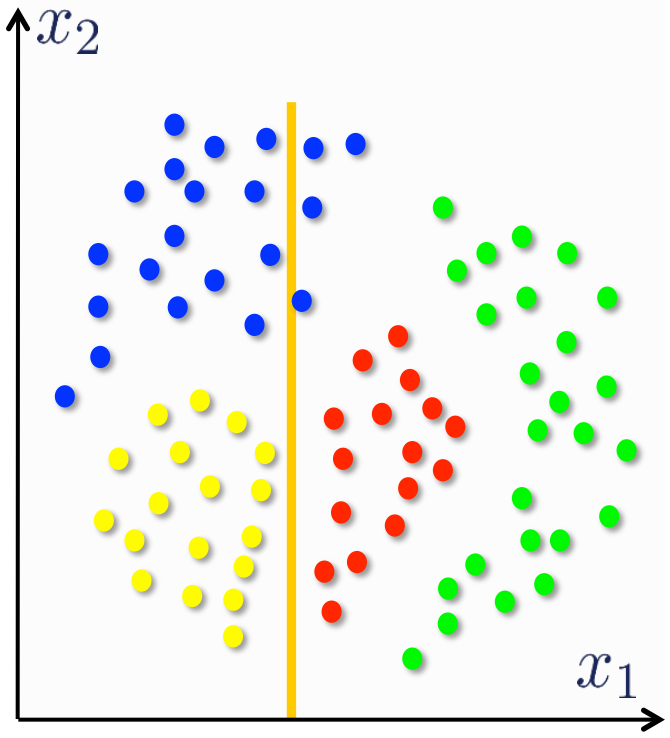




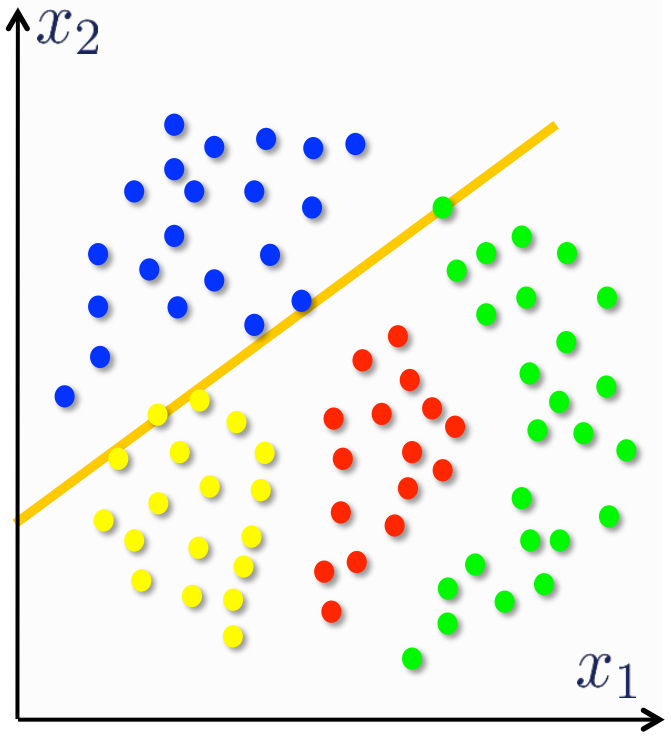


Examples of split functions

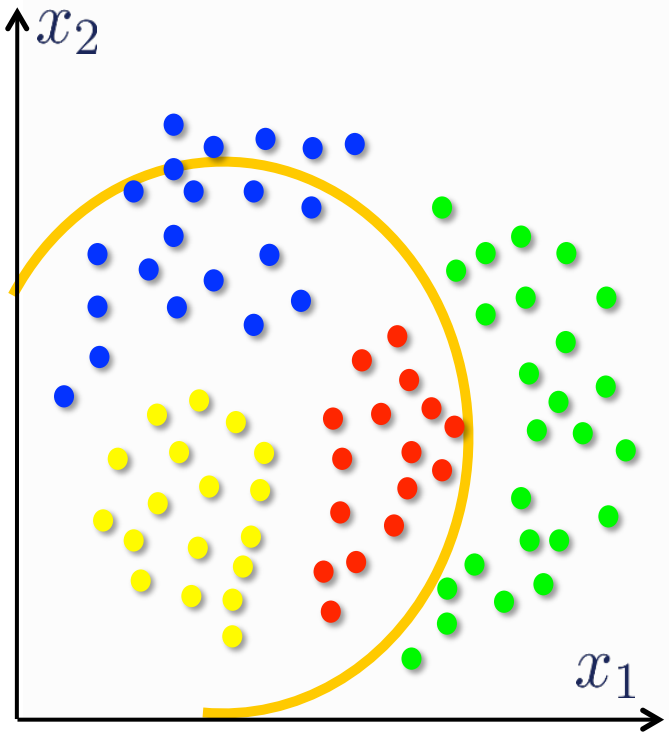




“Axis aligned”



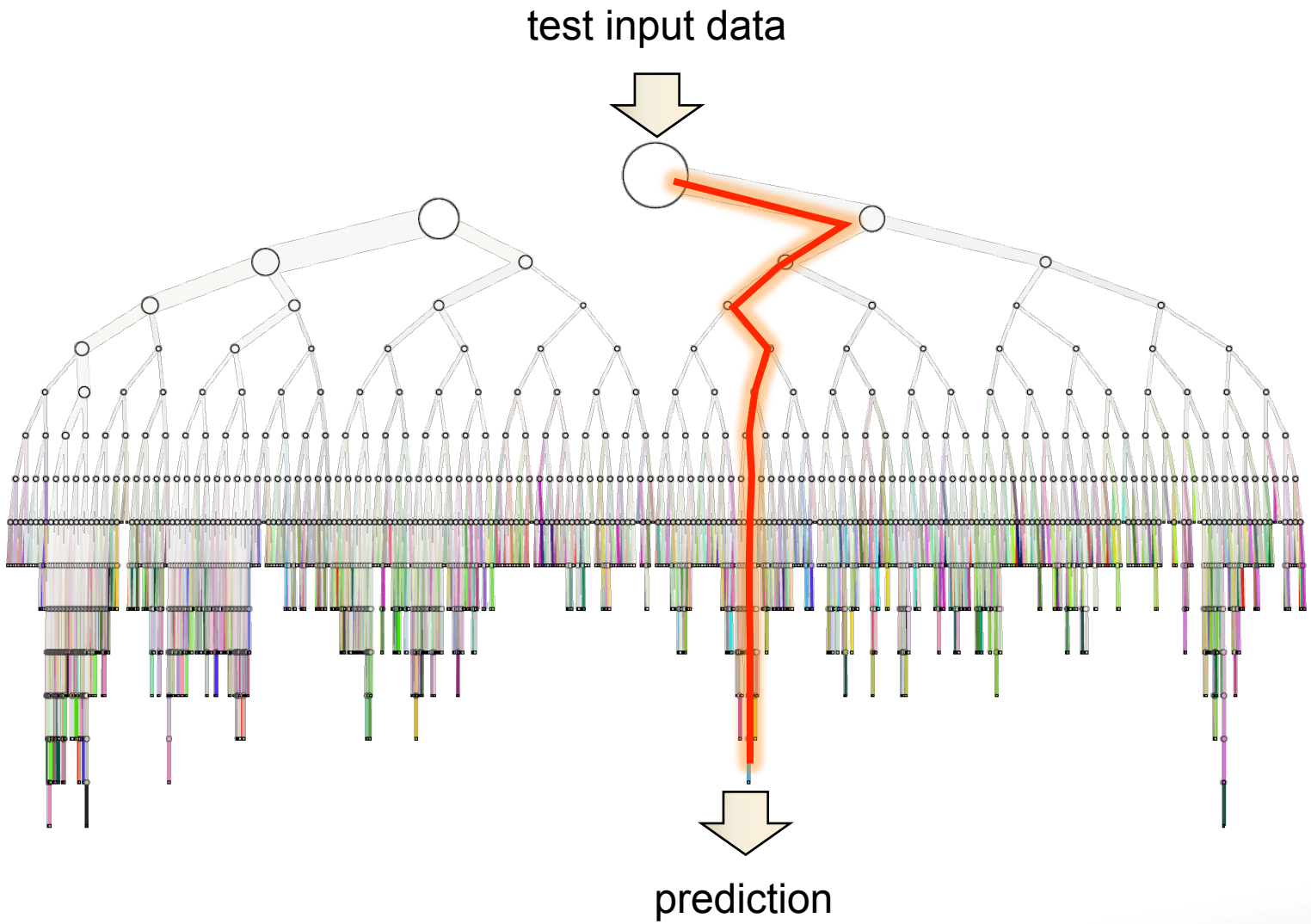
“Oriented hyper-plane”



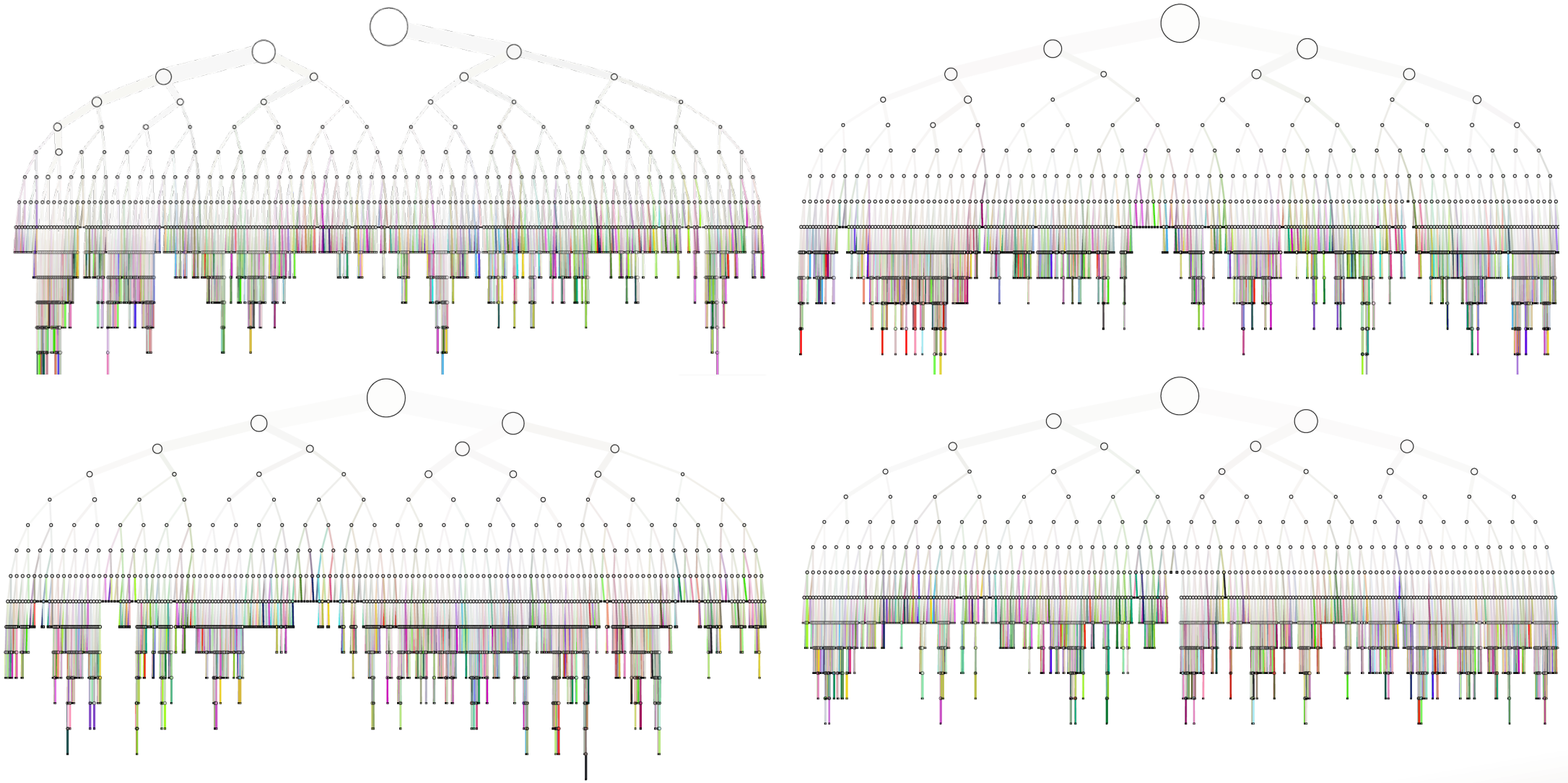
“Conic section”

Particularly efficient

Decision trees: test time prediction

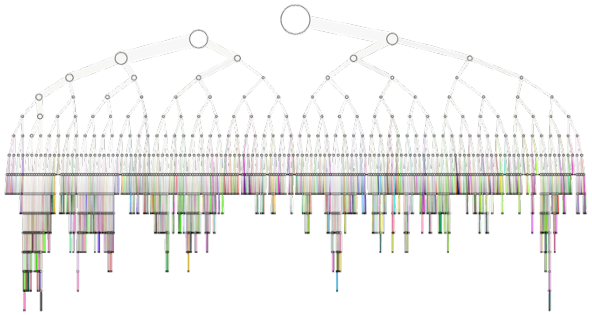


Decision forests

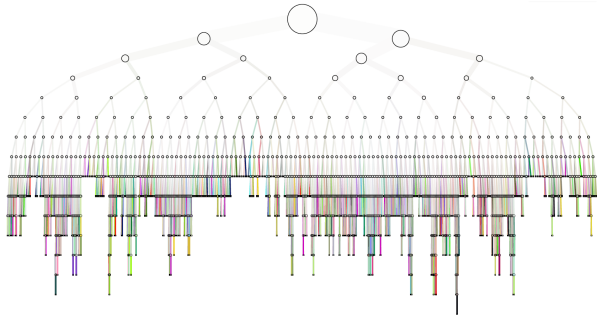


Forest prediction is an aggregate of the predictions across all trees (*e.g.* average probability)

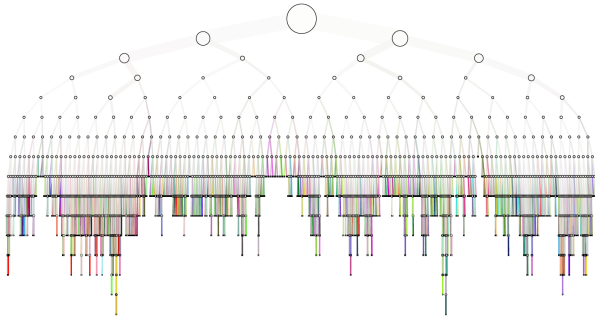
Aggregating tree predictions



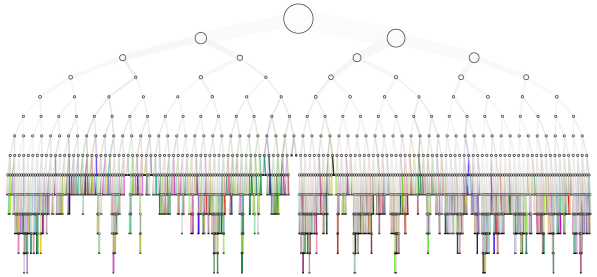
$p_{t=1}(y|\mathbf{v})$



$p_{t=2}(y|\mathbf{v})$

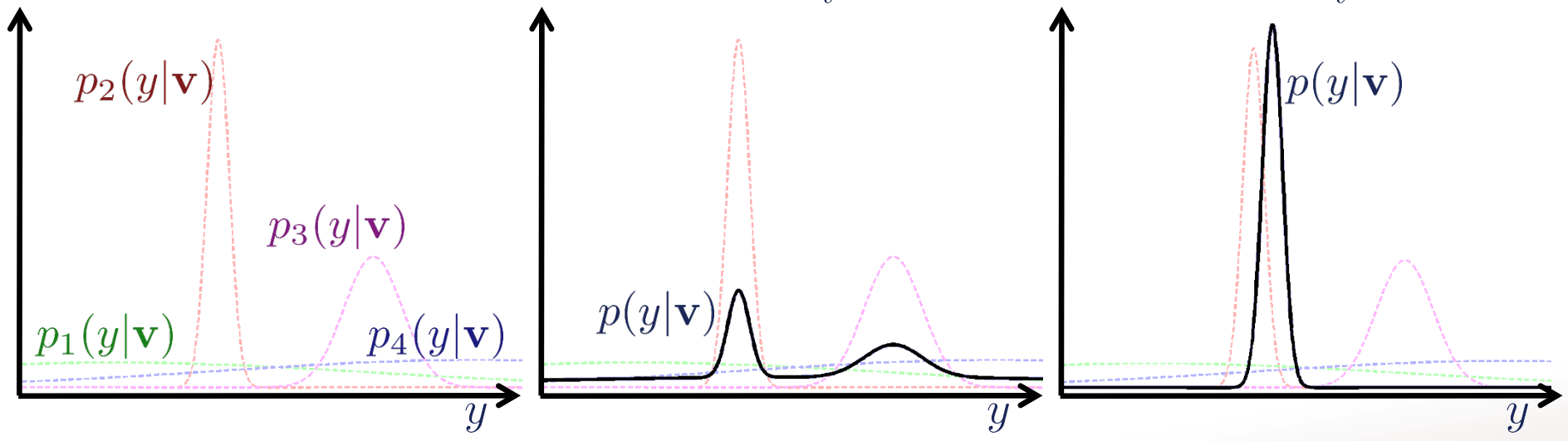


$p_{t=3}(y|\mathbf{v})$

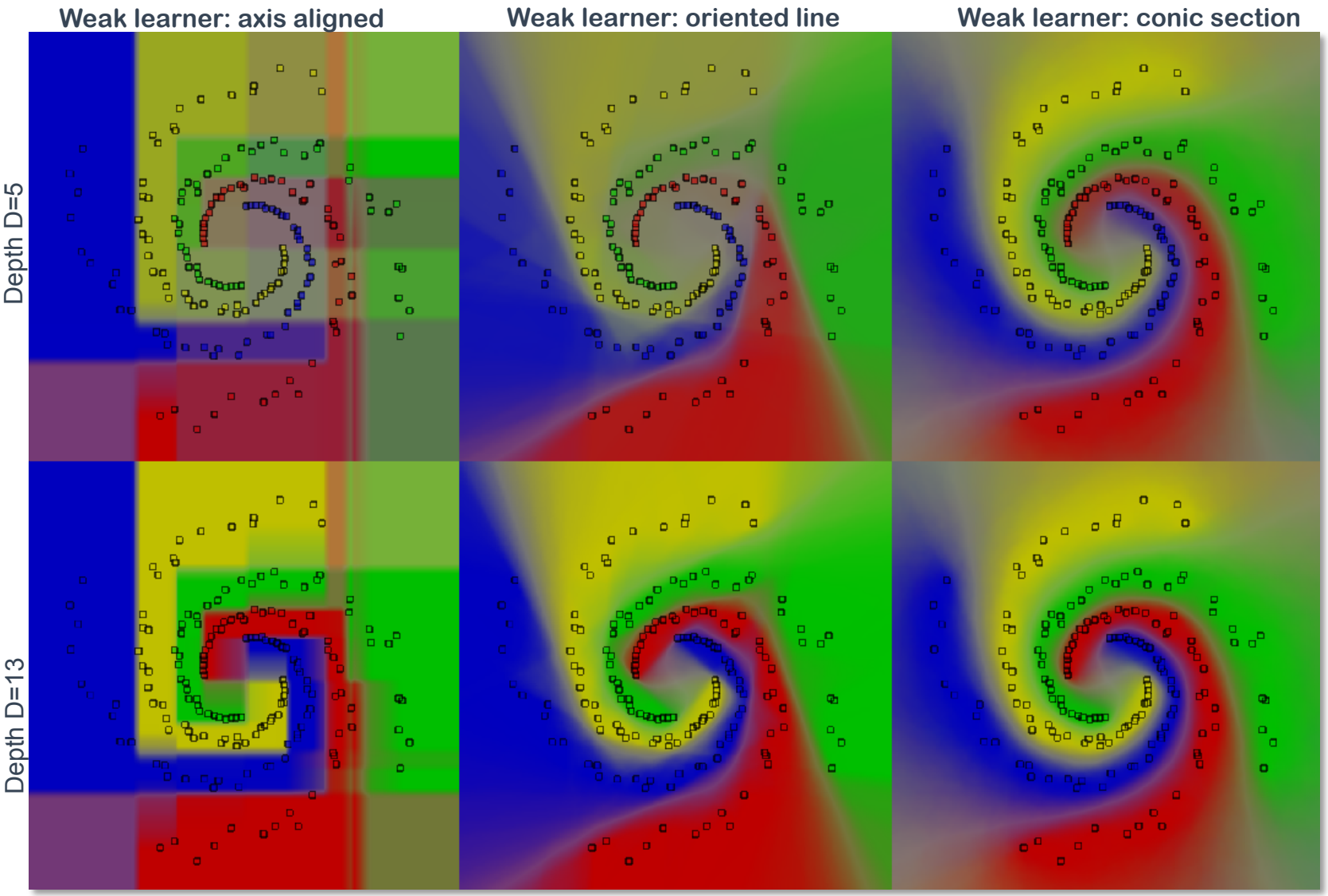


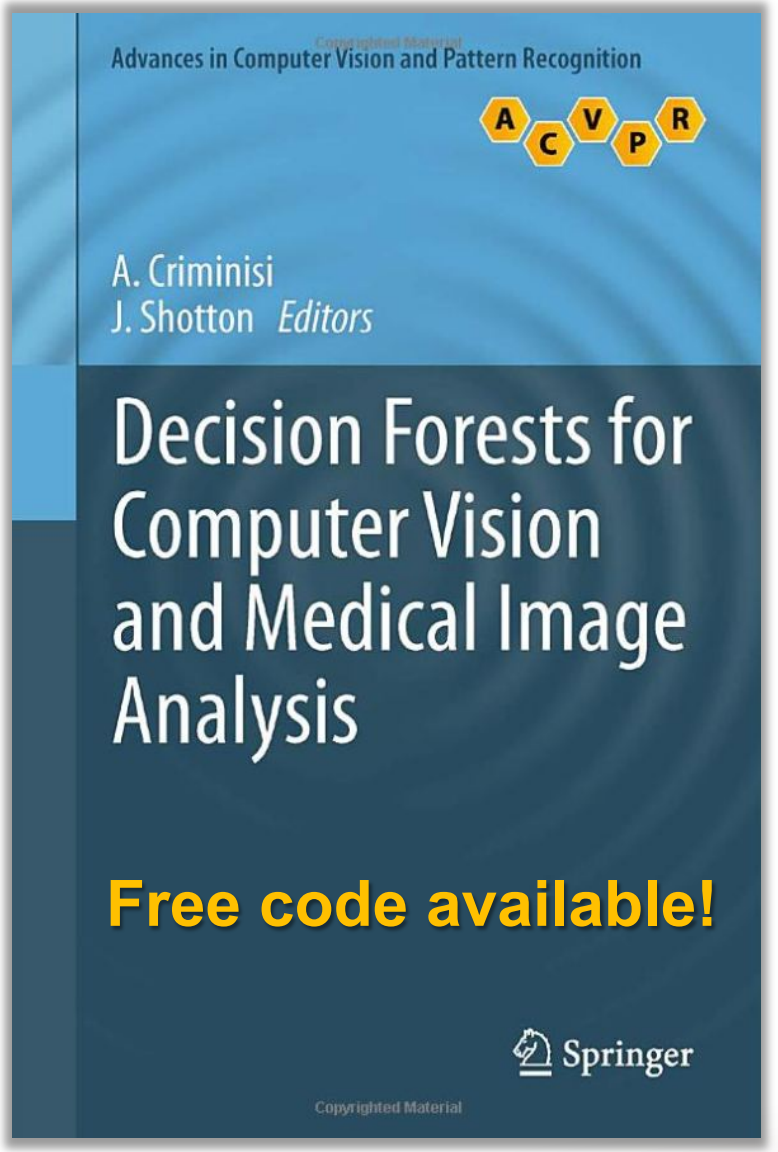
$p_{t=4}(y|\mathbf{v})$

$$p(y|\mathbf{v}) = \frac{1}{T}\sum_{t}^{T} p_t(y|\mathbf{v}) \qquad p(y|\mathbf{v}) = \frac{1}{Z}\prod_{t}^{T} p_t(y|\mathbf{v})$$



Effect of tree depth and randomness





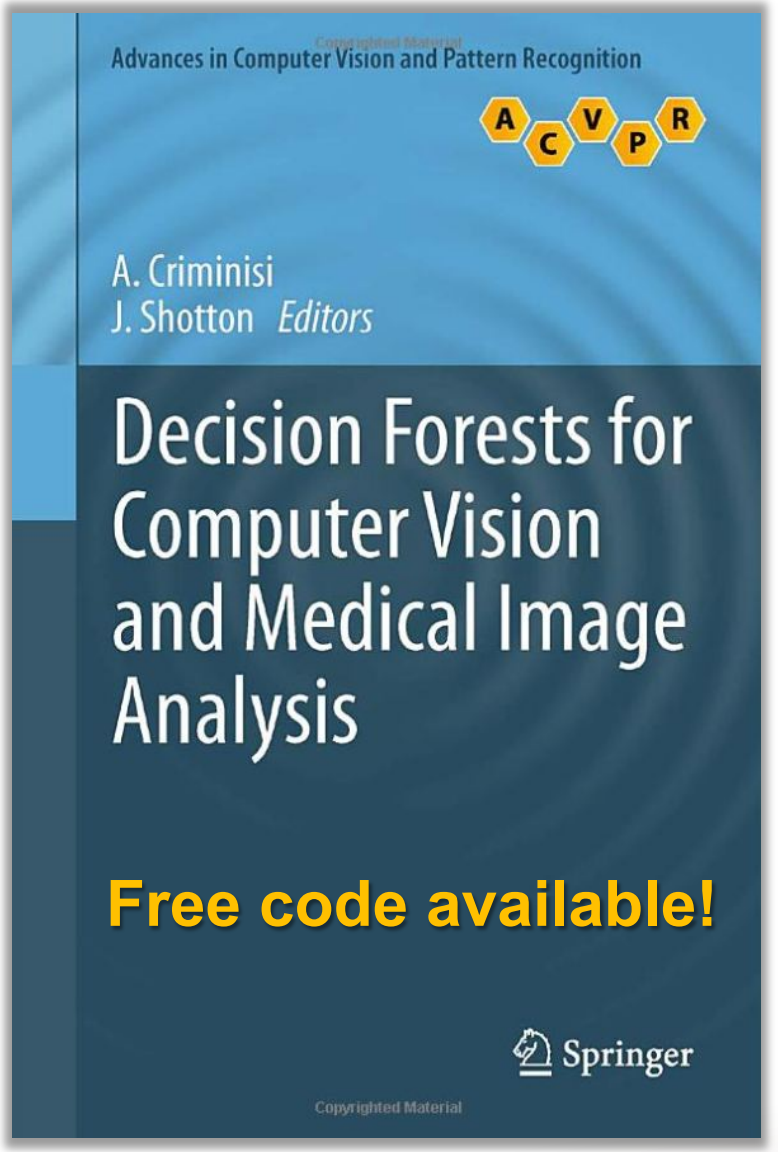
A. Criminisi and J. Shotton, **Decision Forests for Computer Vision and Medical Image Analysis**, Springer, February 2013

The Sherwood free software library

- **Properties**
  - In C++ and C#
  - Can be called from Matlab
  - Source code is available
  - Easy to read and understand.
  - Follows the book naming convention and structure

- **Classification examples**
  - ./sw clas /d 15 /t 100 /split linear exp3_n4.txt
  - ./sw clas /d 15 /t 100 /split linear exp7_n4.txt

- **Regression examples**
  - ./sw regression /d 4 /t 100 exp8.txt
  - ./sw regression /d 4 /t 100 exp10.txt



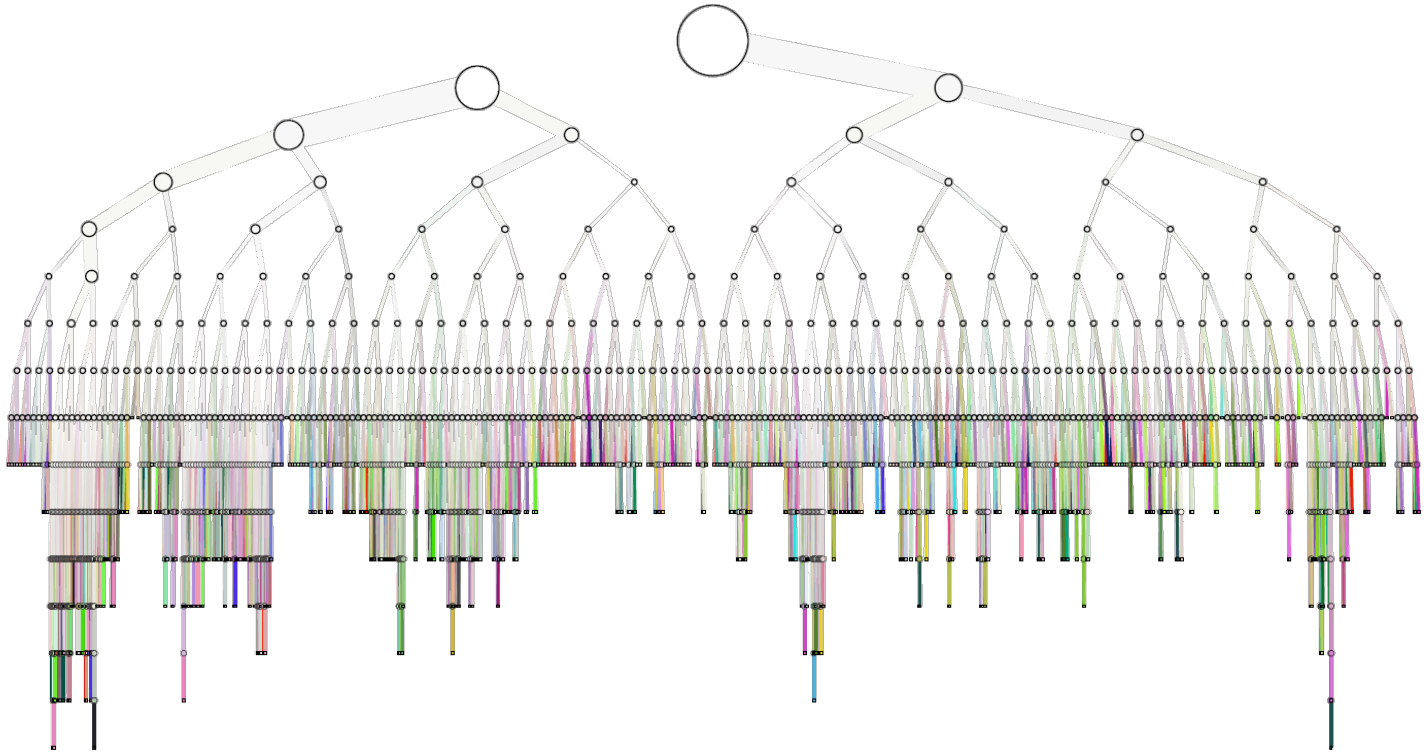
A. Criminisi and J. Shotton, **Decision Forests for Computer Vision and Medical Image Analysis**, Springer, February 2013

Talk overview

- A brief introduction to machine learning
- **Decision** forests and **jungles**
- Applications in medical image analysis
  - Anatomy localization
  - Spine detection
  - Brain tumour segmentation
  - Learned image super-resolution
  - Quantifying progression of multiple sclerosis

J. Shotton, T. Sharp, P. Kohli, S. Nowozin, J. Winn, and A. Criminisi, **Decision Jungles: Compact and Rich Models for Classification**, in *Proc. NIPS*, 2013

Are forests sufficient?



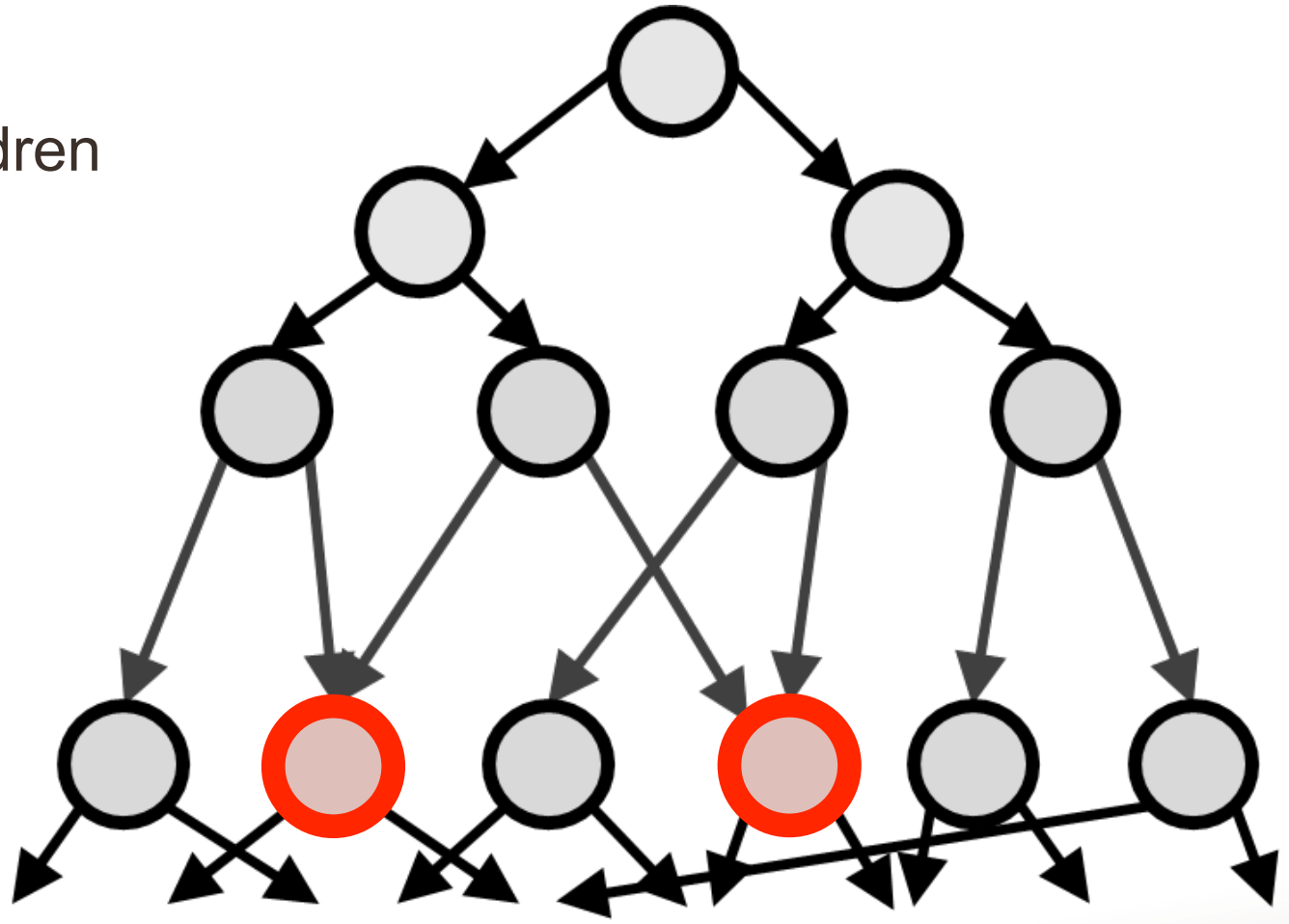
- Memory issues:
  - Number of nodes in trees grows exponentially with depth

- Amount of training data
  - Training data is quickly diluted with depth
  - Yet, training deeper trees (on enough data) yields highest test accuracy

    (several real applications, e.g. Kinect, have “infinite” data available)

From trees to DAGs: node merging

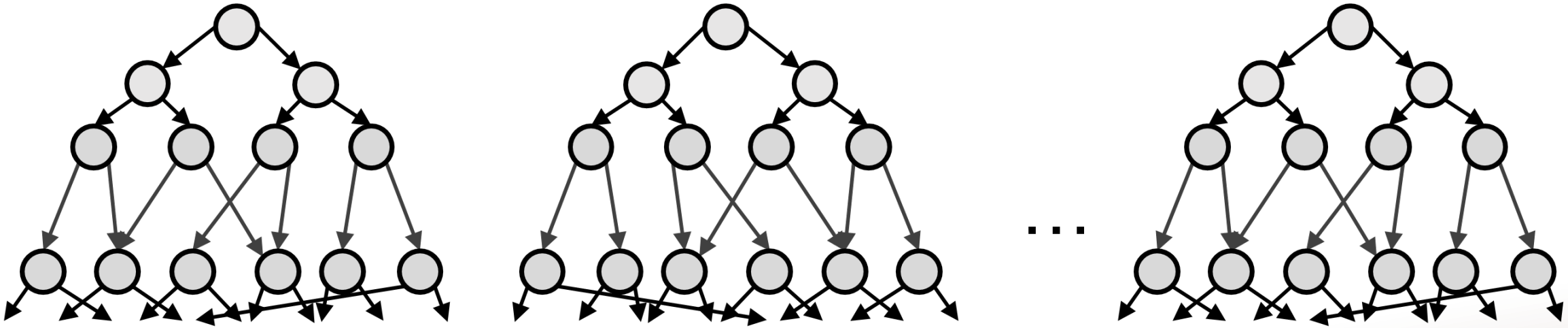
- Each internal node has 2 children (like in binary trees)
- Each non-root node can have more than 1 parent



Decision jungles

- A “jungle” is an ensemble of *rooted* decision DAGs

- We train each DAG layer by layer, jointly optimizing both
  - the structure of the DAG
  - the split node features



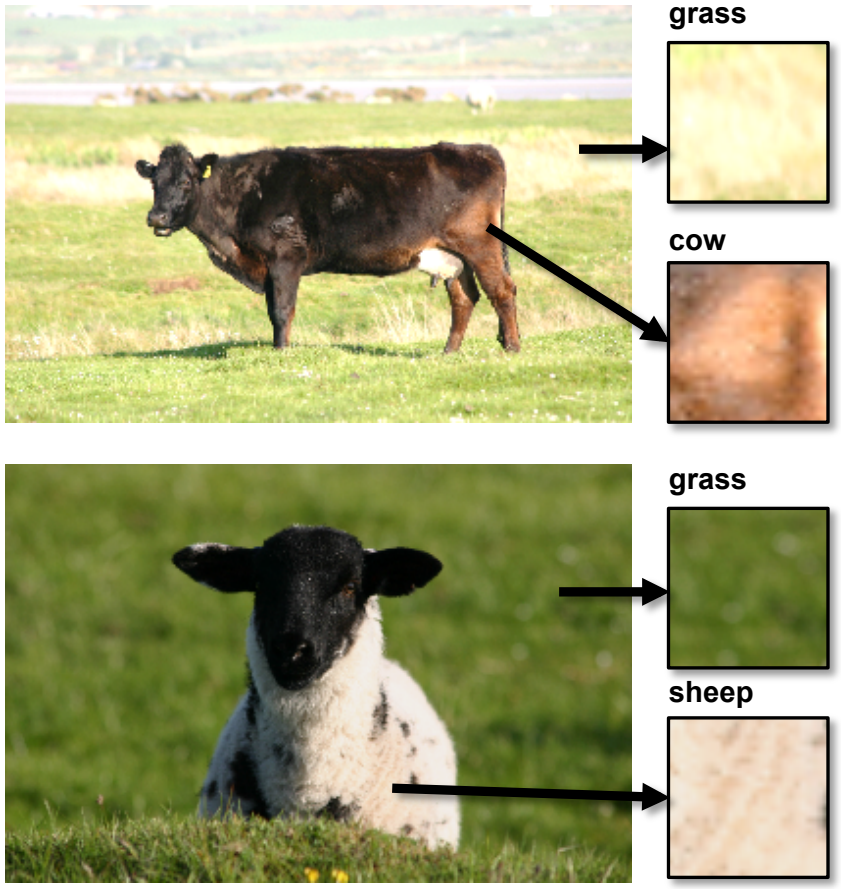
Properties of jungles

- **Limited memory consumption**
  - e.g. by specifying a width at each layer in the DAG
- Potentially improved **generalization**
  - fewer parameters
  - **less “dilution”** of training data

How do DAGs help in practice?

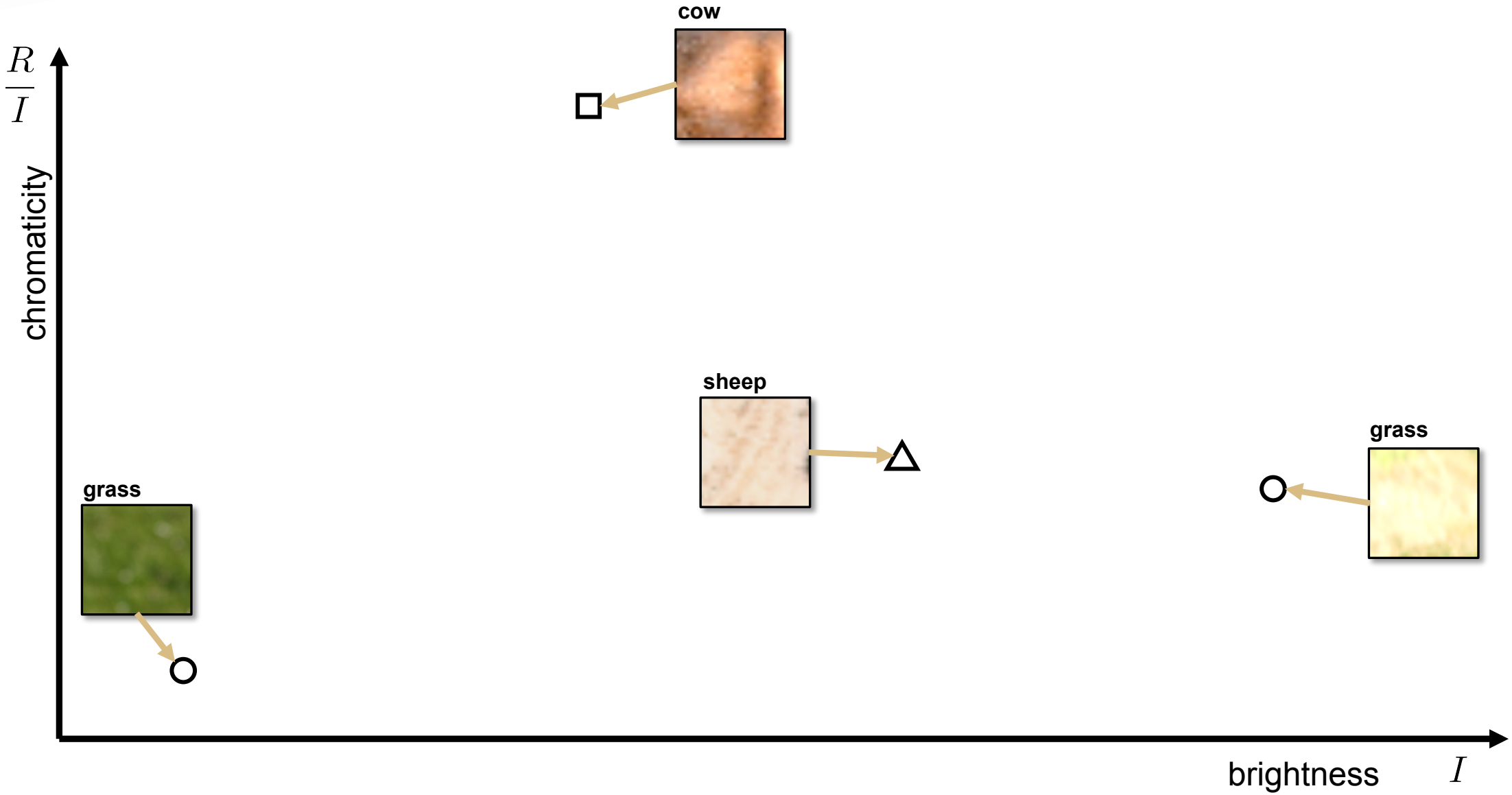
A toy example on classifying images of cows, sheep and grass

Training data

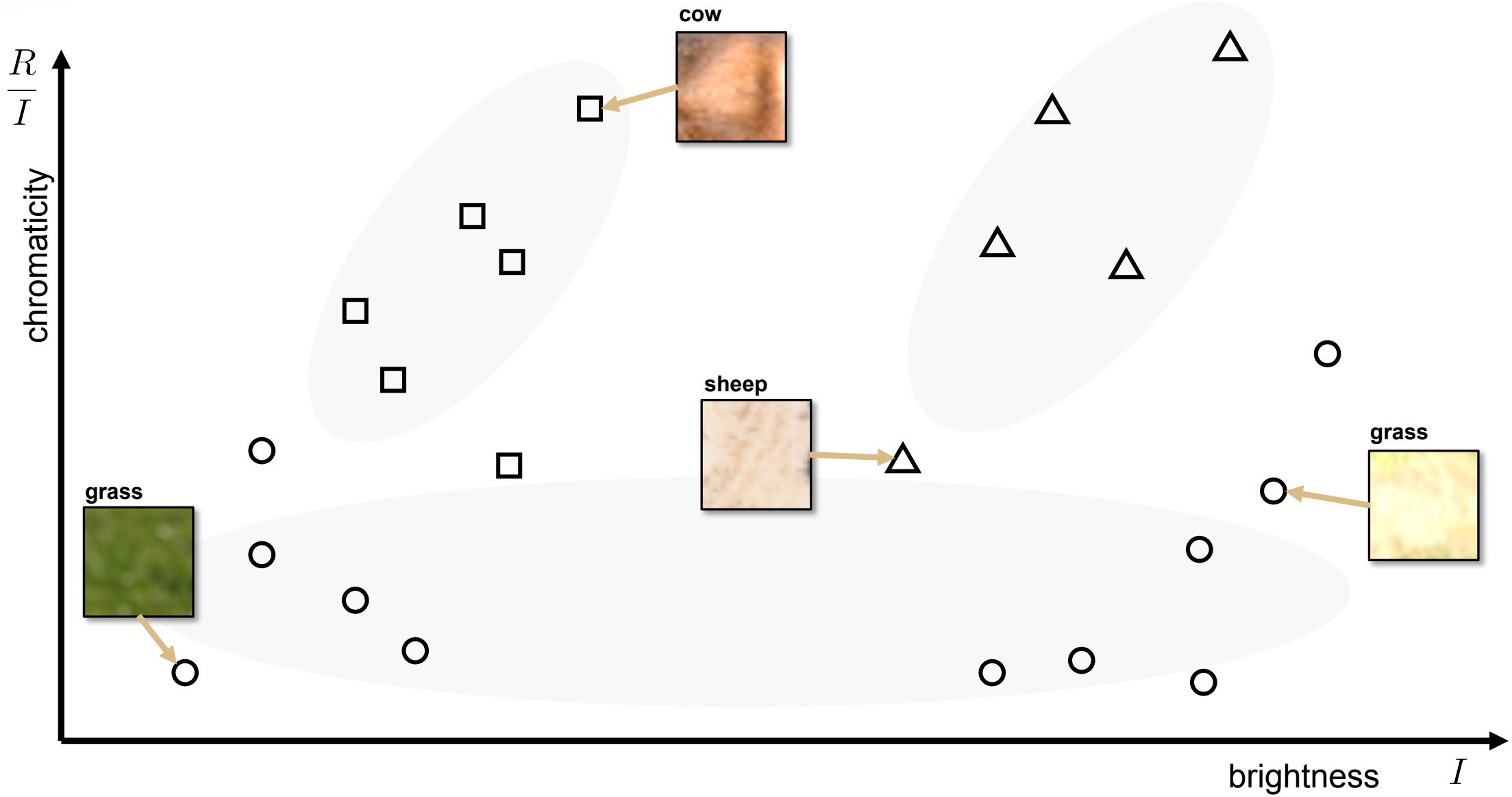


...

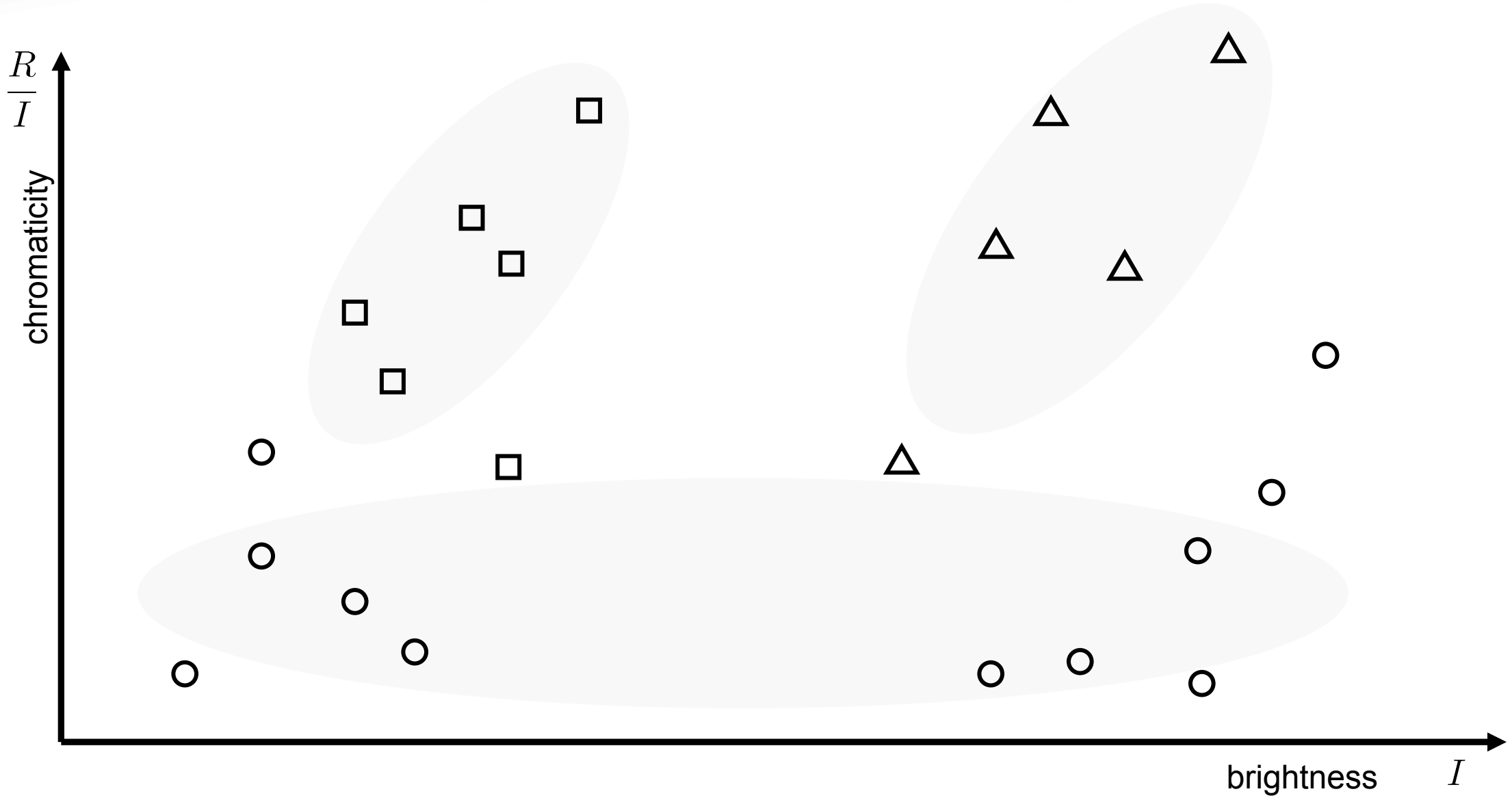
A toy example on classifying images of cows, sheep and grass



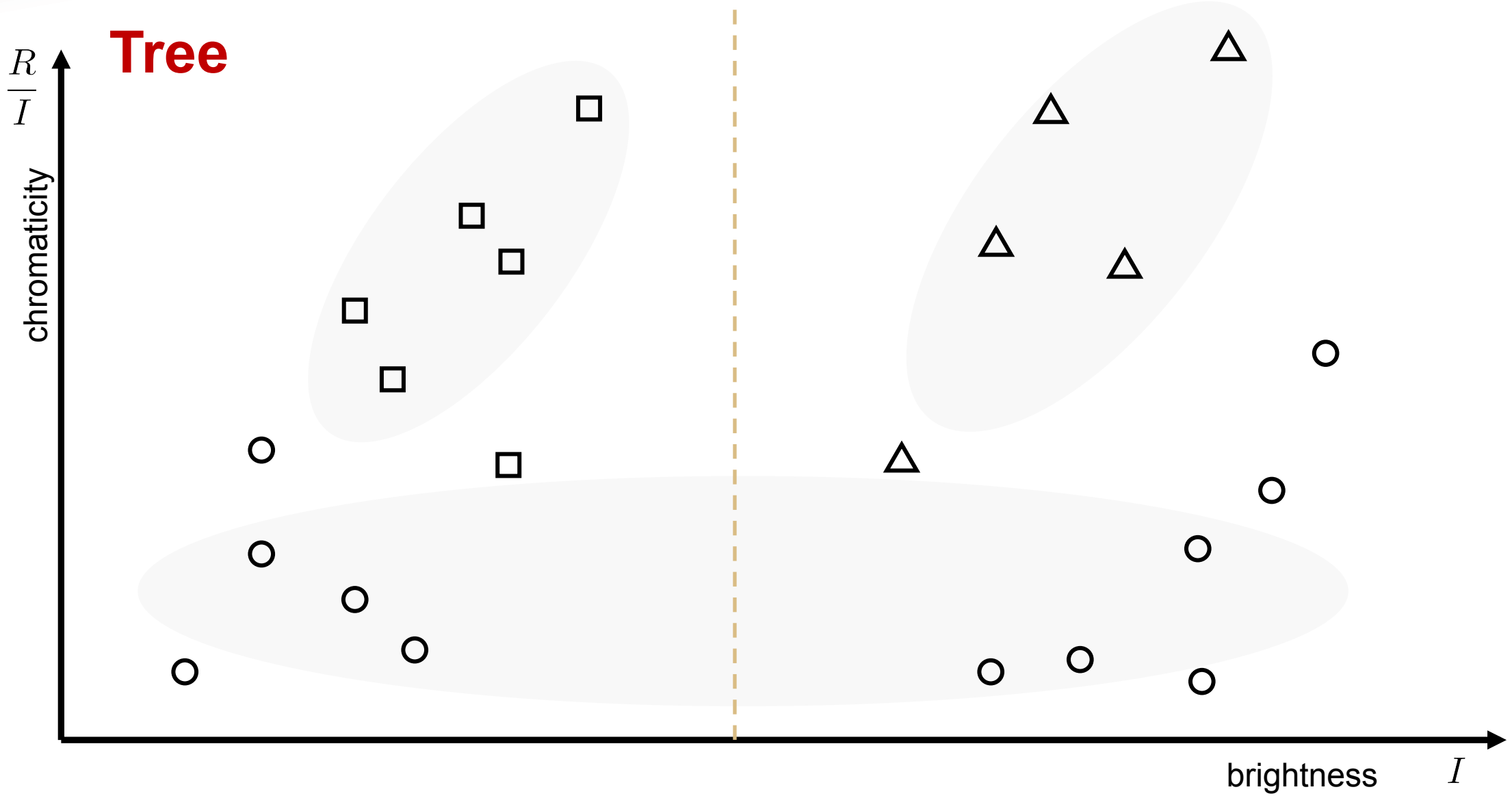
A toy example on classifying images of cows, sheep and grass



A toy example on classifying images of cows, sheep and grass

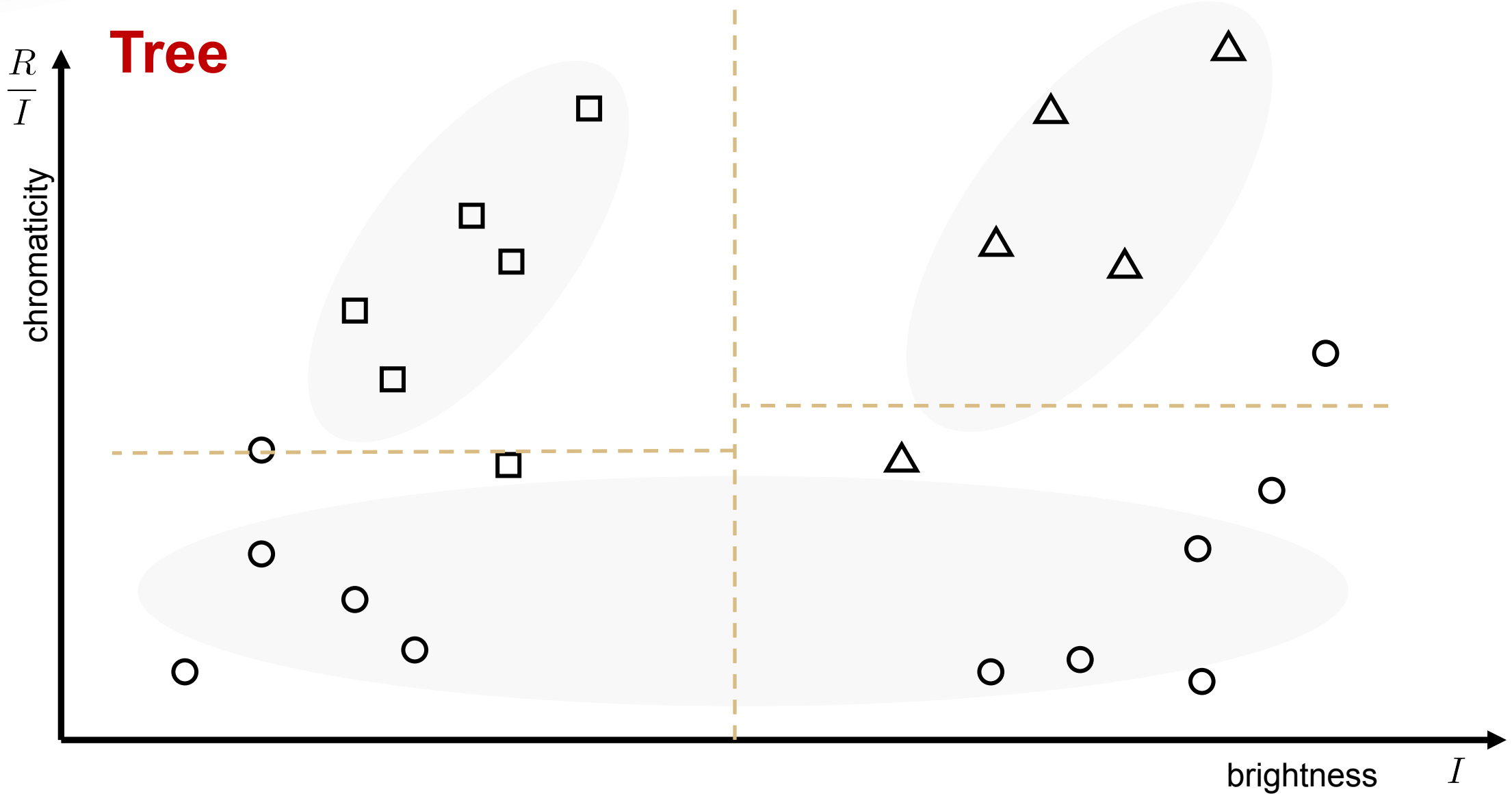


A toy example on classifying images of cows, sheep and grass

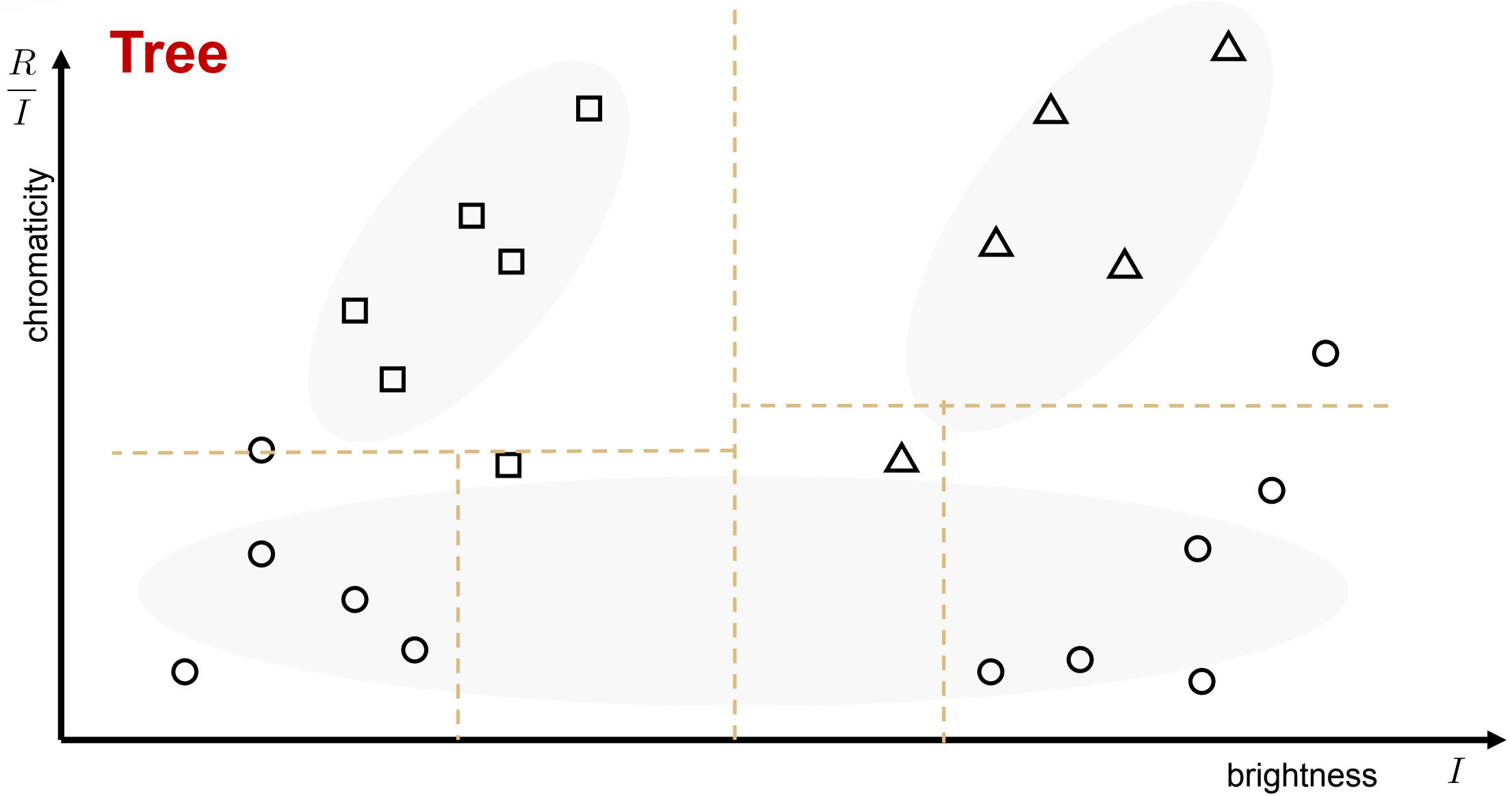


Axis-aligned splits only

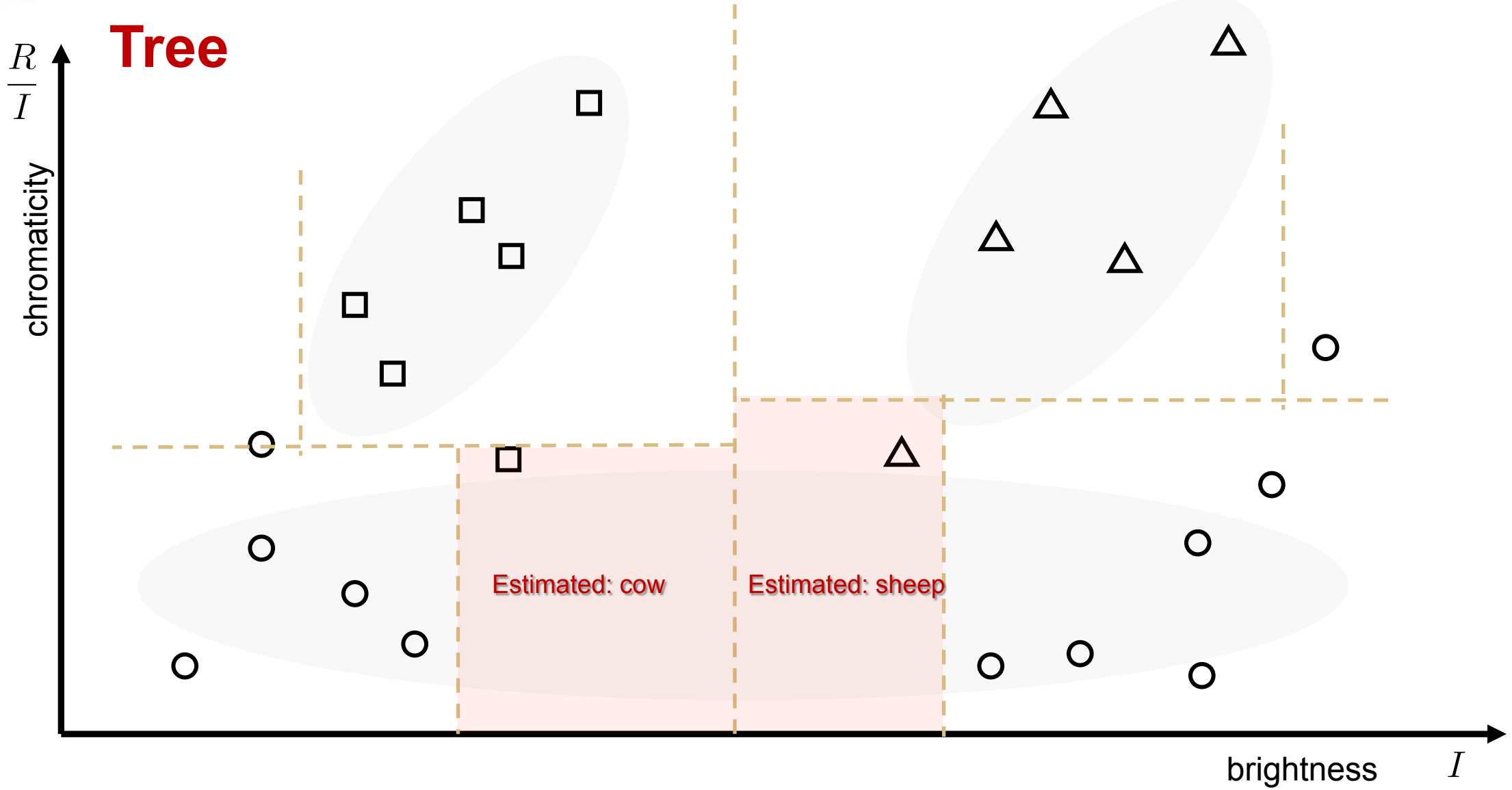
A toy example on classifying images of cows, sheep and grass



A toy example on classifying images of cows, sheep and grass

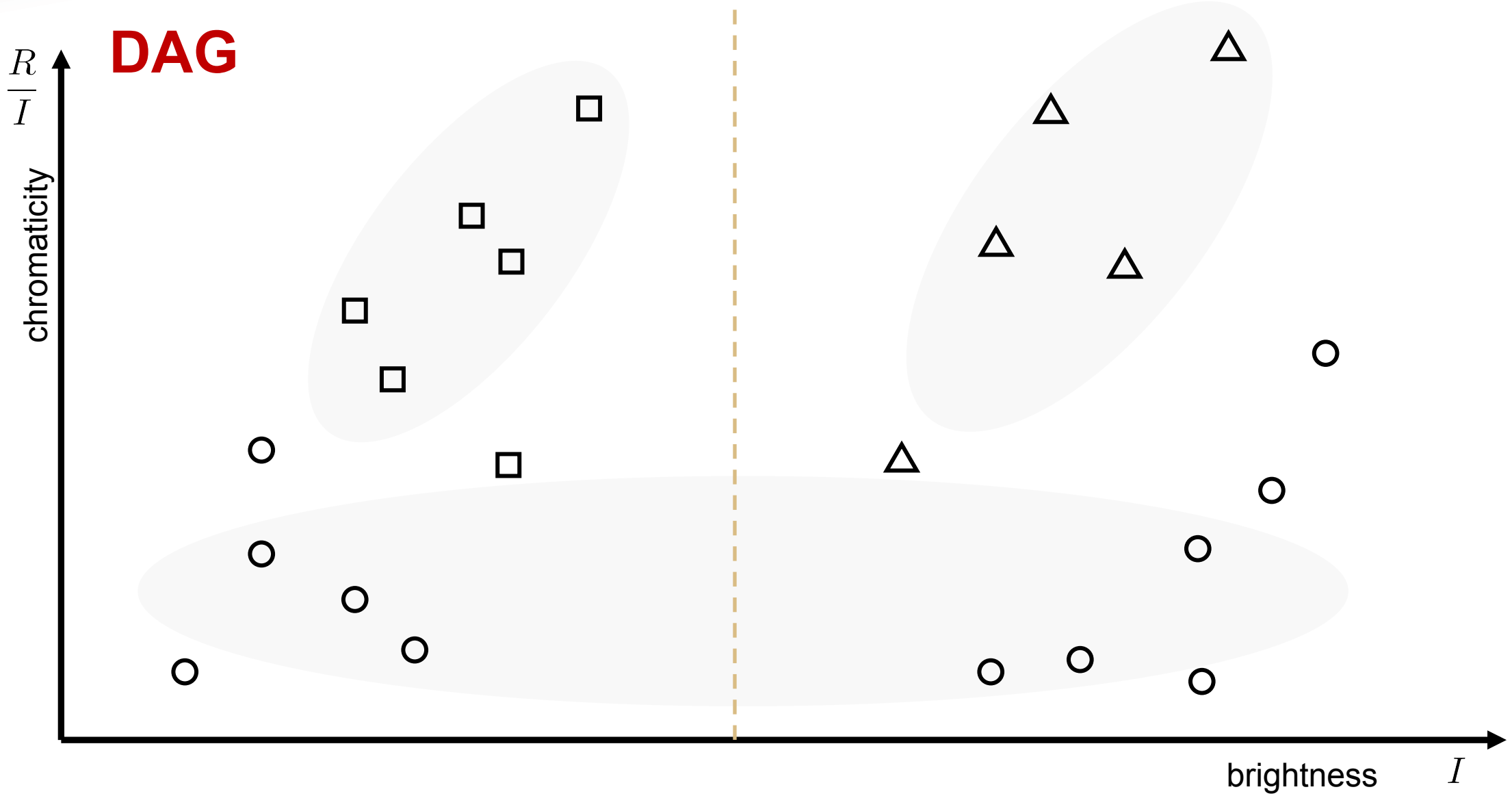


A toy example on classifying images of cows, sheep and grass

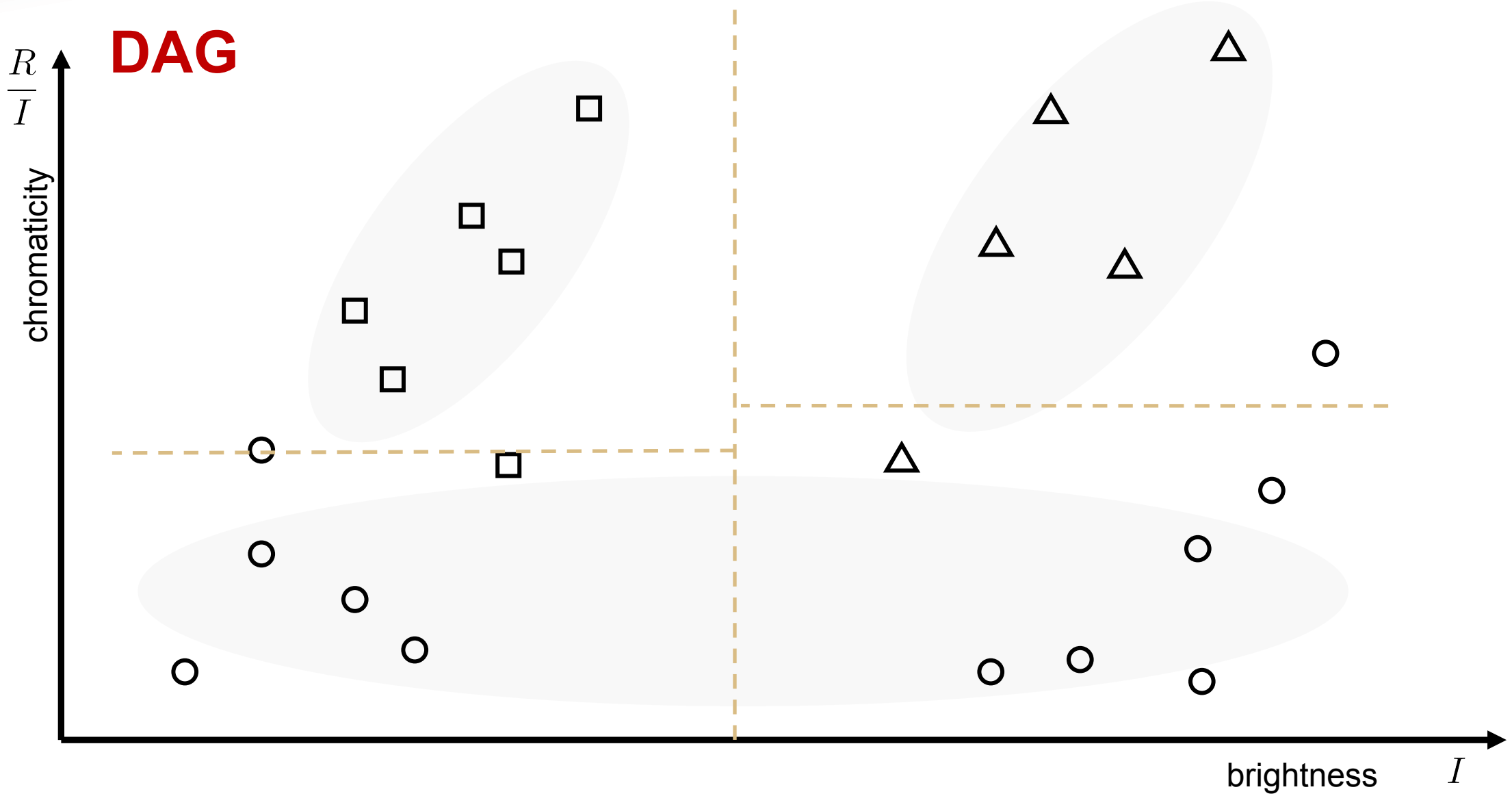


Too many model parameters: overfitting

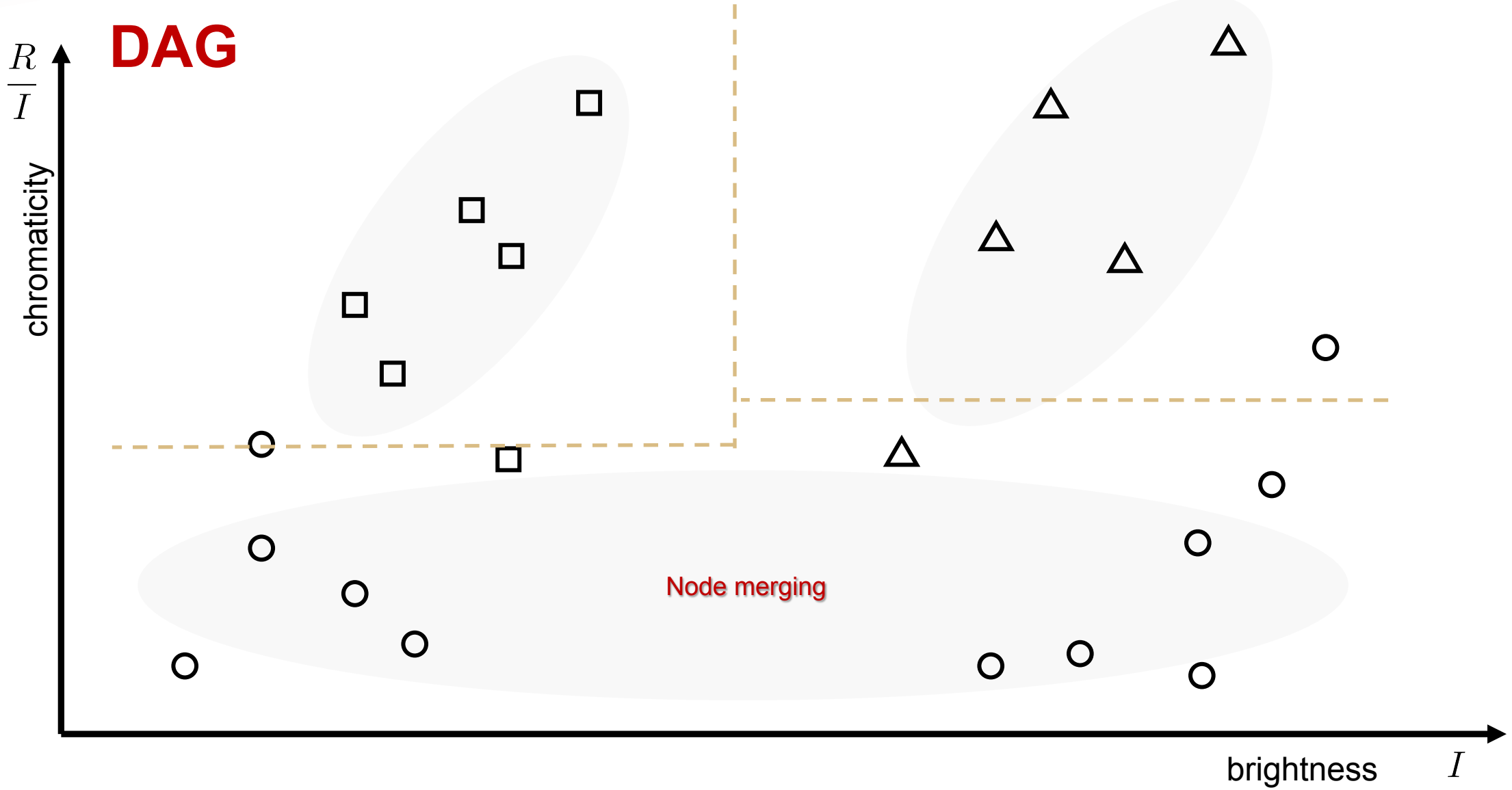
A toy example on classifying images of cows, sheep and grass



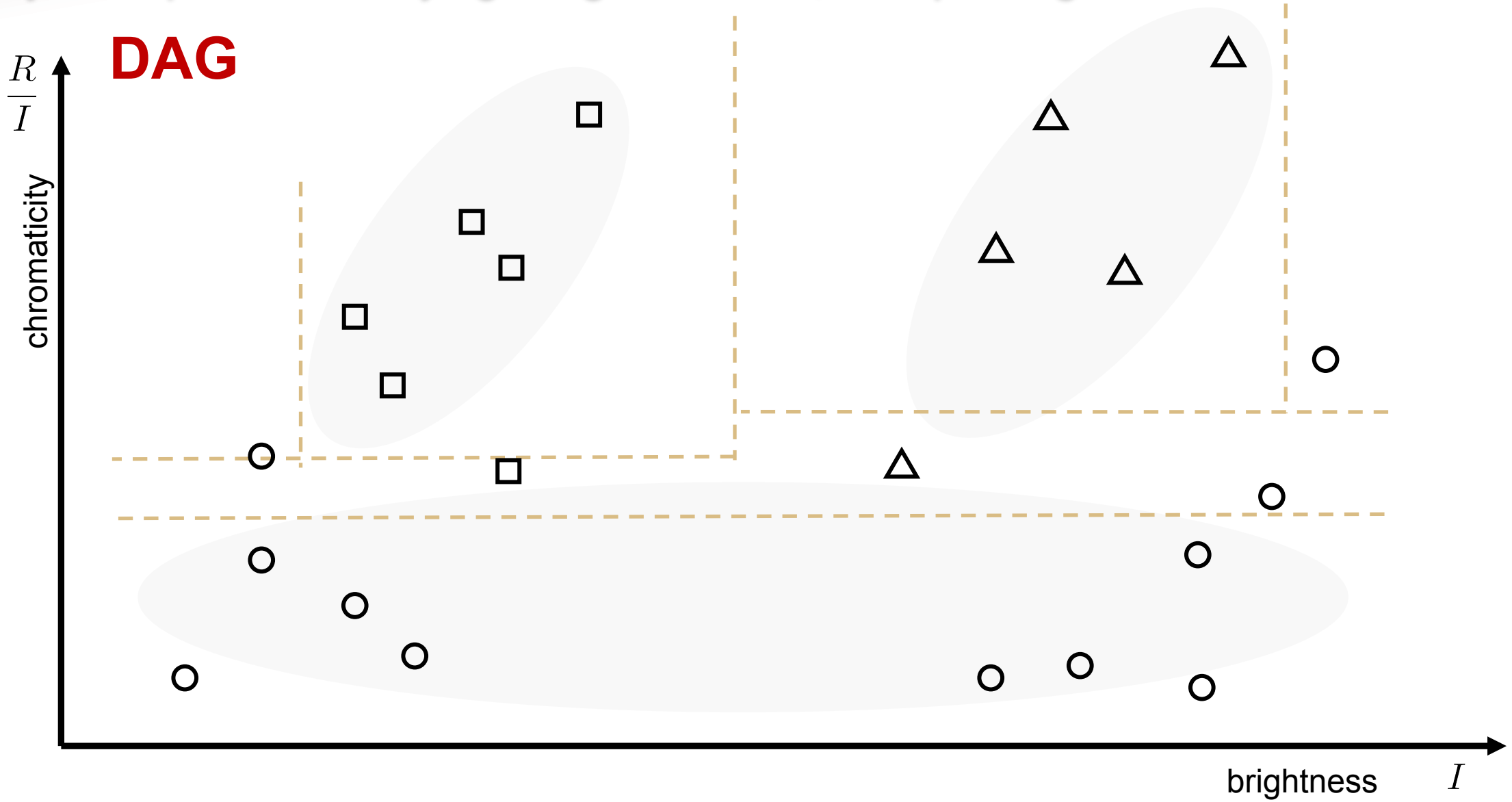
A toy example on classifying images of cows, sheep and grass



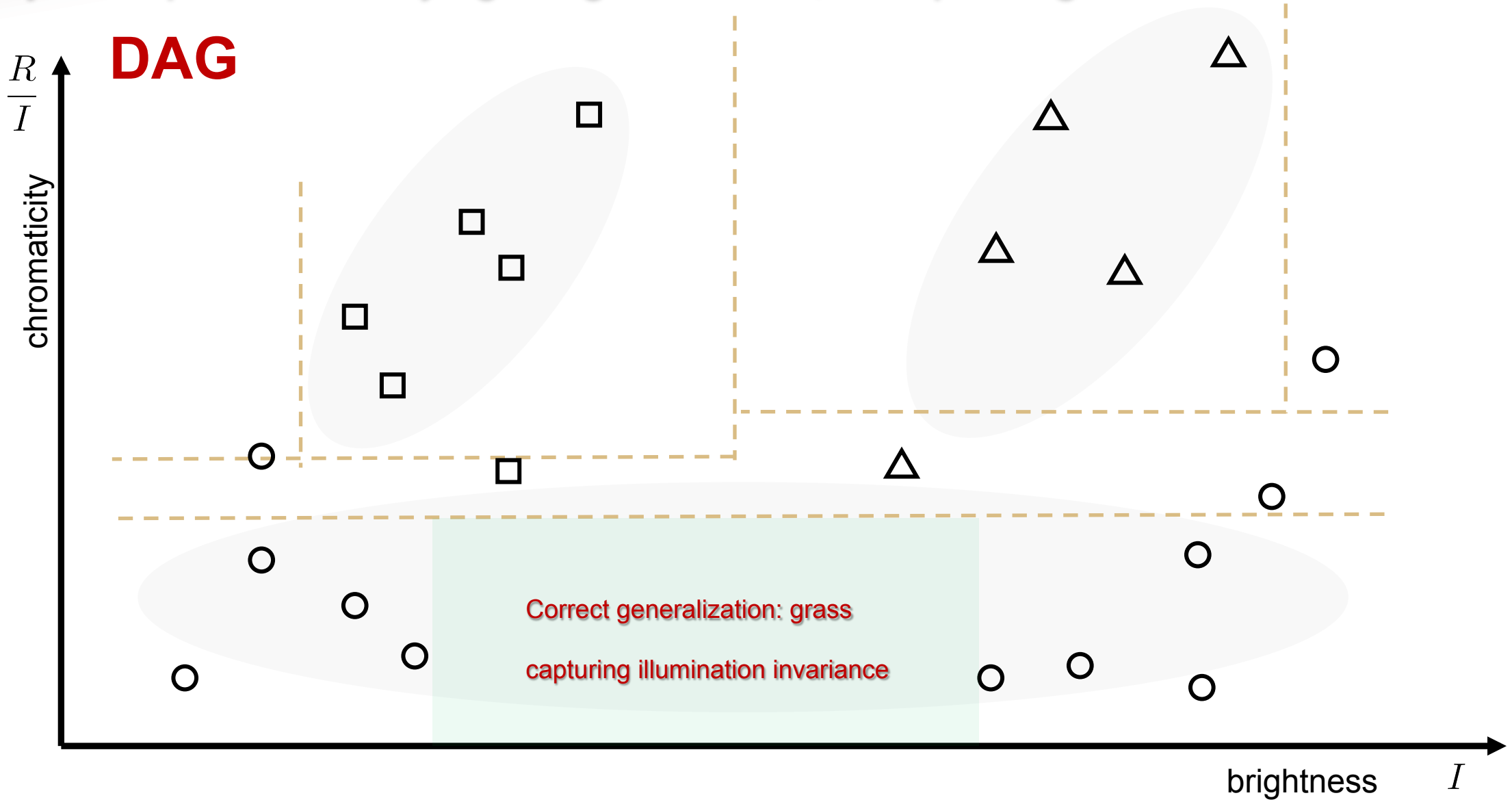
A toy example on classifying images of cows, sheep and grass



A toy example on classifying images of cows, sheep and grass



A toy example on classifying images of cows, sheep and grass

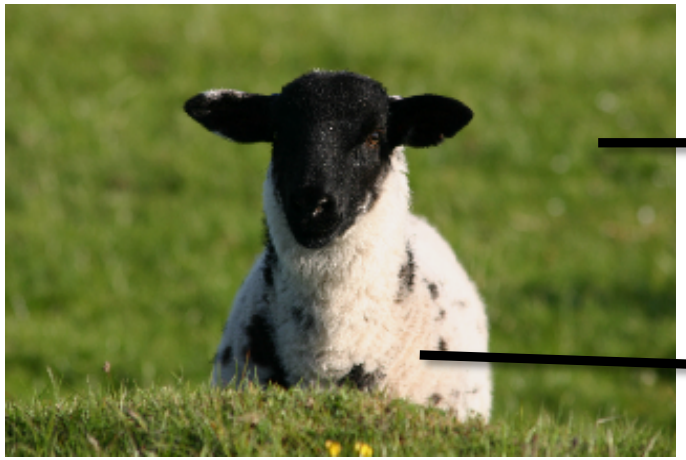


A toy example on classifying images of cows, sheep and grass

Trees vs DAGs

grass

cow

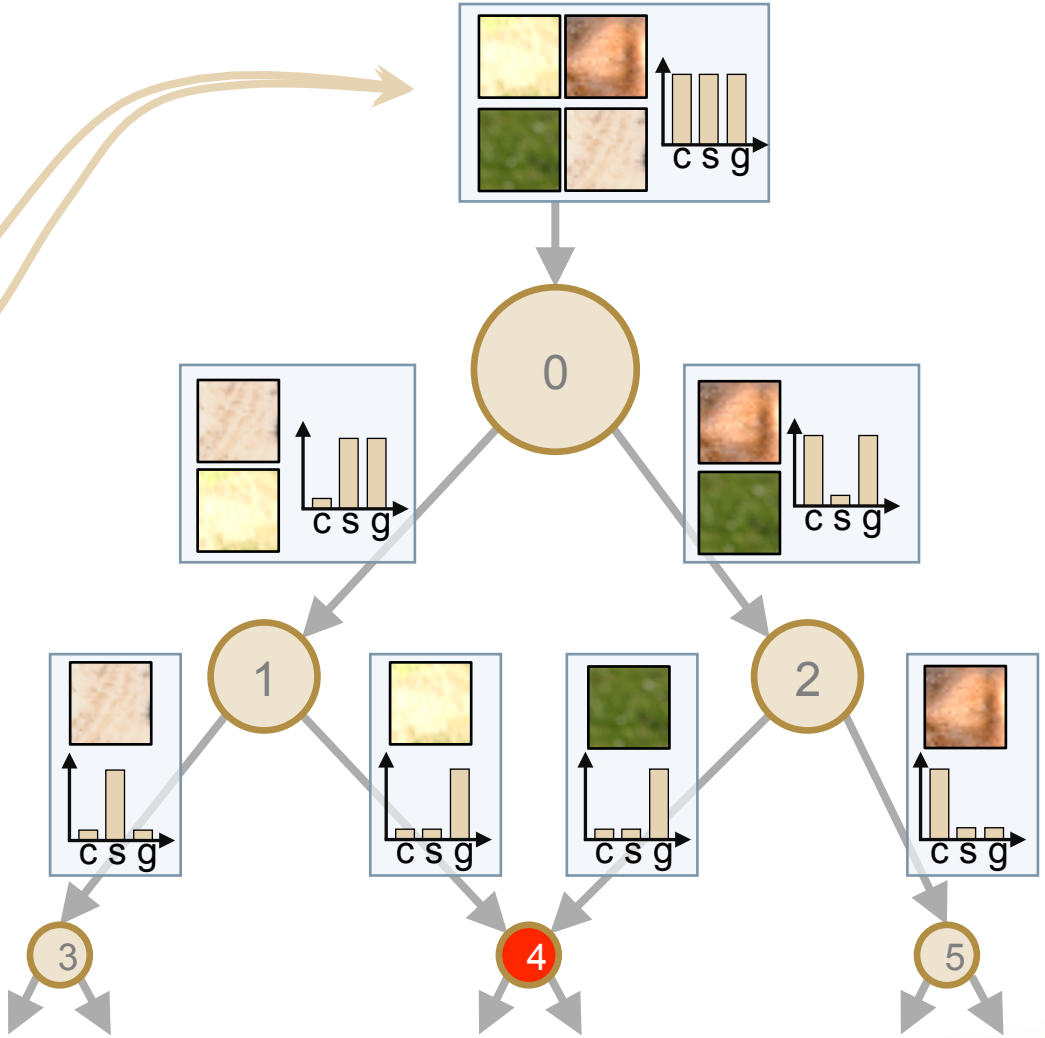


grass

sheep

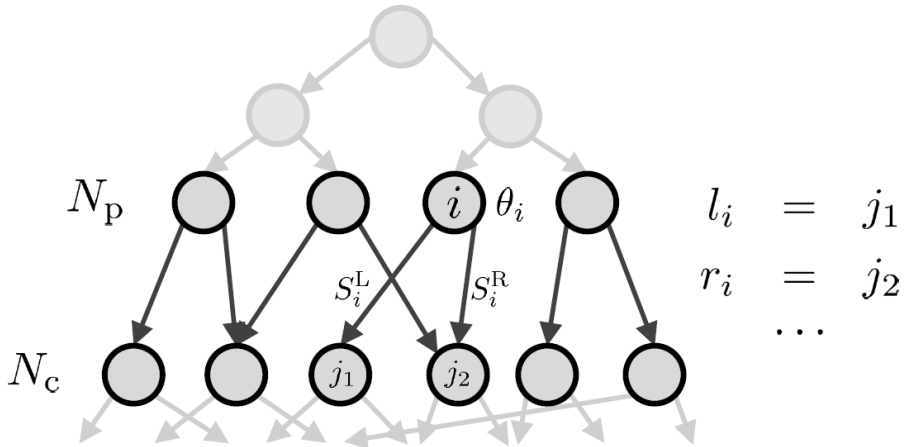
...

Training patches



Merged nodes help capture appearance invariance

Jungles: training objective



$$E(\{\theta_i\}, \{l_i\}, \{r_i\}) = \sum_{j \in N_{\mathrm{c}}} |S_j| \, H(S_j)$$

features and branches for all parent nodes i

sum over child nodes j

number of examples at j

entropy of examples that reach child node j

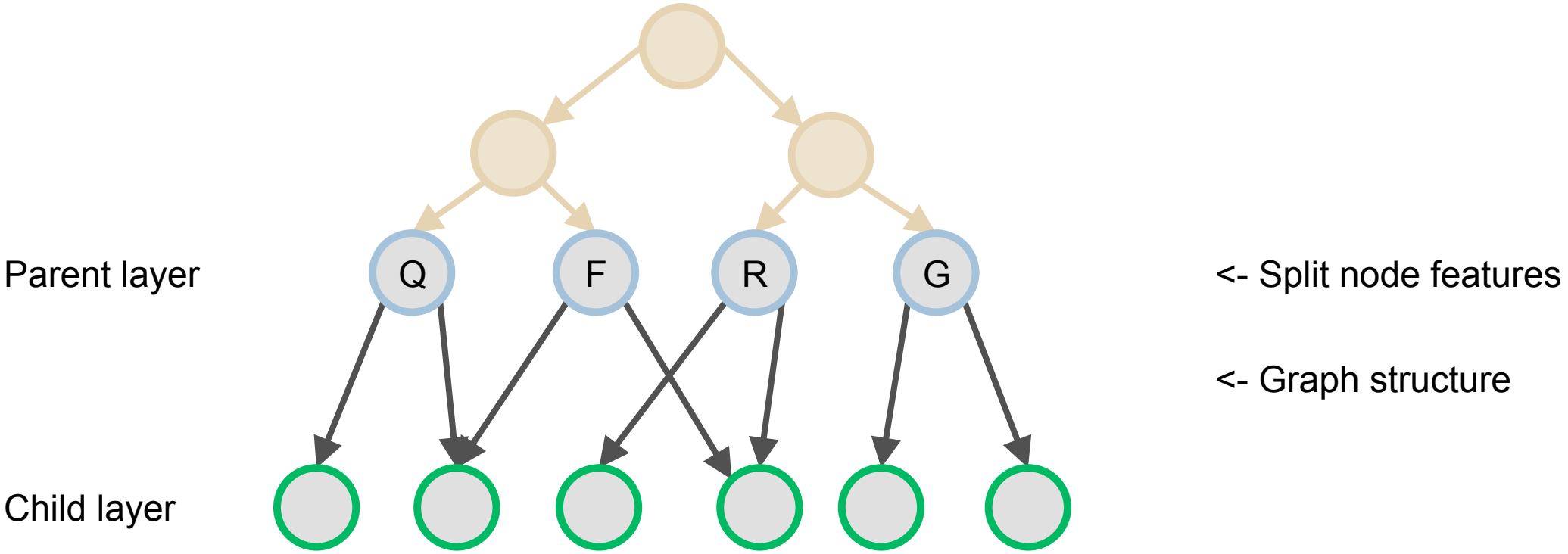
Jungles: optimization algorithm

- Allocate a maximum of $M = \lfloor N \downarrow c \rfloor$ nodes per level
  - allows us to fix memory budget

- Simple “move-making” optimization algorithm
  - start from “feasible” initialization
  - randomly choose a parent node
  - either update its split function (given fixed DAG structure)
  - or update its left or right branch (given fixed split function)

Algorithm overview

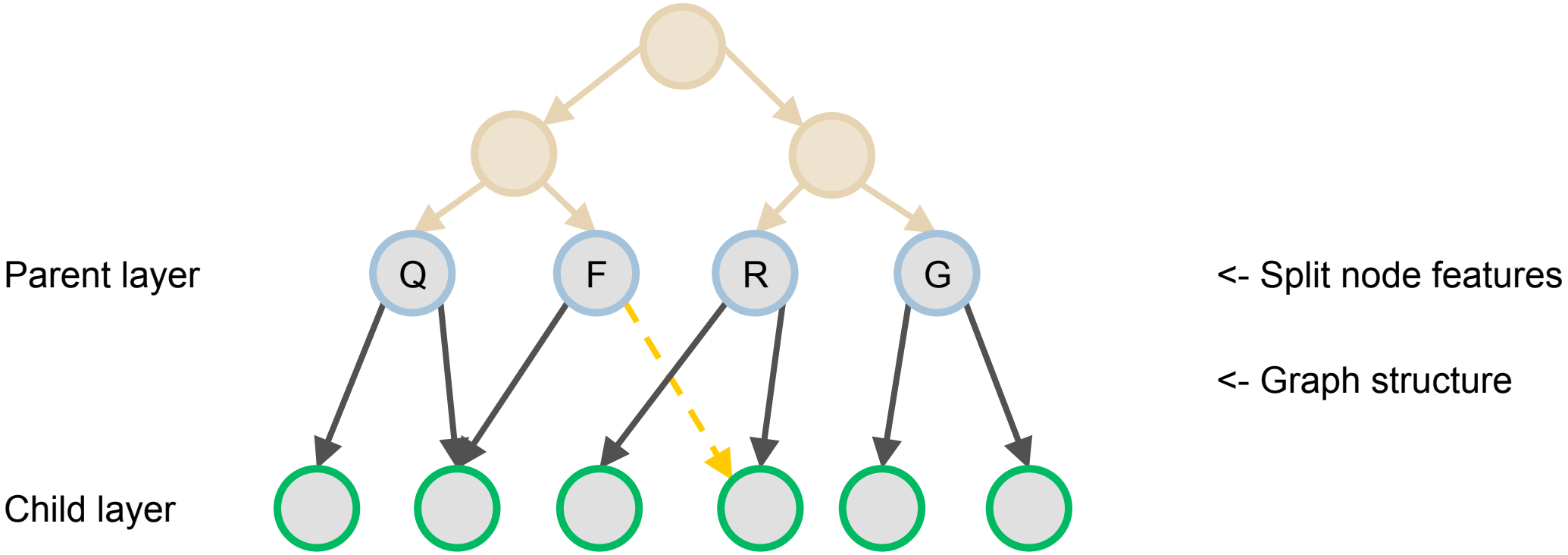
- **Train each DAG layer by layer**, jointly optimizing both
  - the structure of the DAG
  - the split node features themselves



Current objective score = 2.87

Algorithm overview

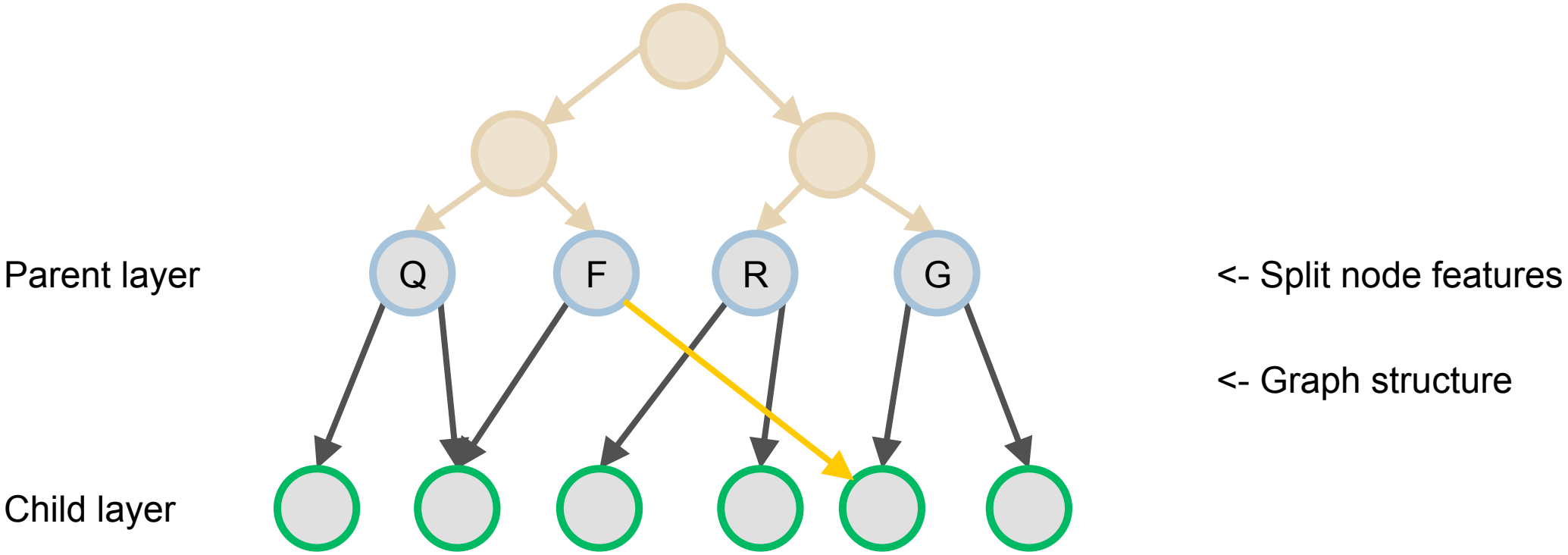
- **Train each DAG layer by layer**, jointly optimizing both
  - the structure of the DAG
  - the split node features themselves



Current objective score = 2.87

Algorithm overview

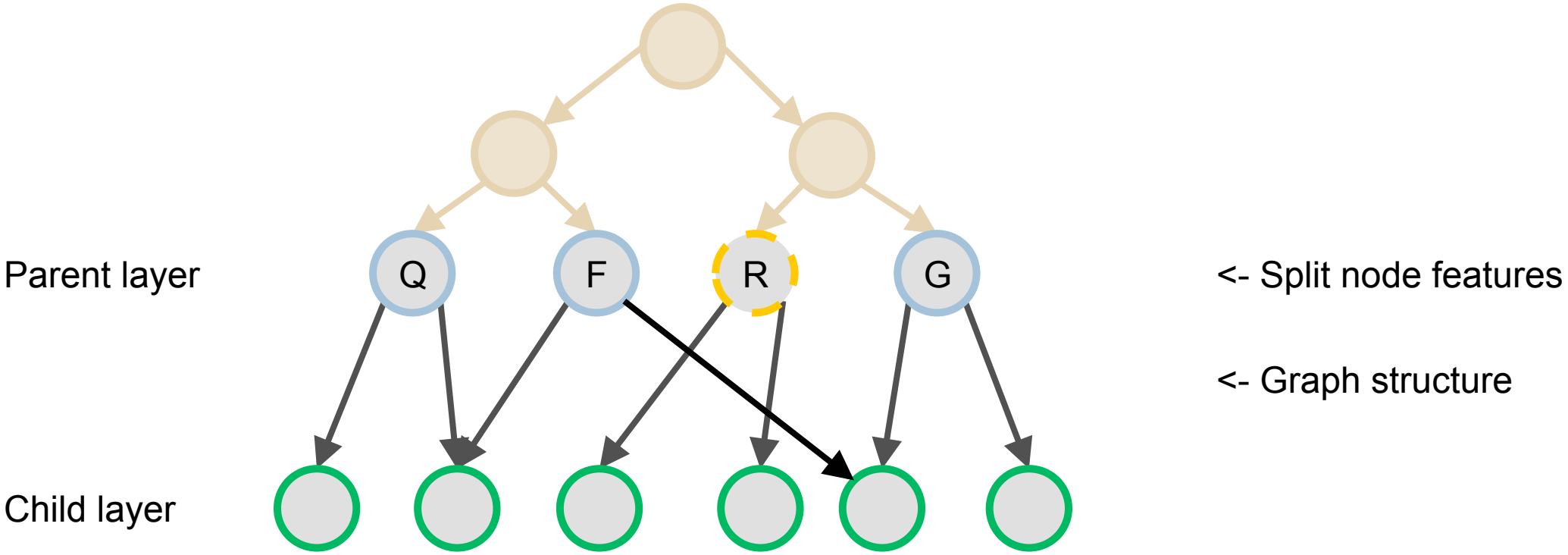
- **Train each DAG layer by layer**, jointly optimizing both
  - the structure of the DAG
  - the split node features themselves



Current objective score = 2.52

Algorithm overview

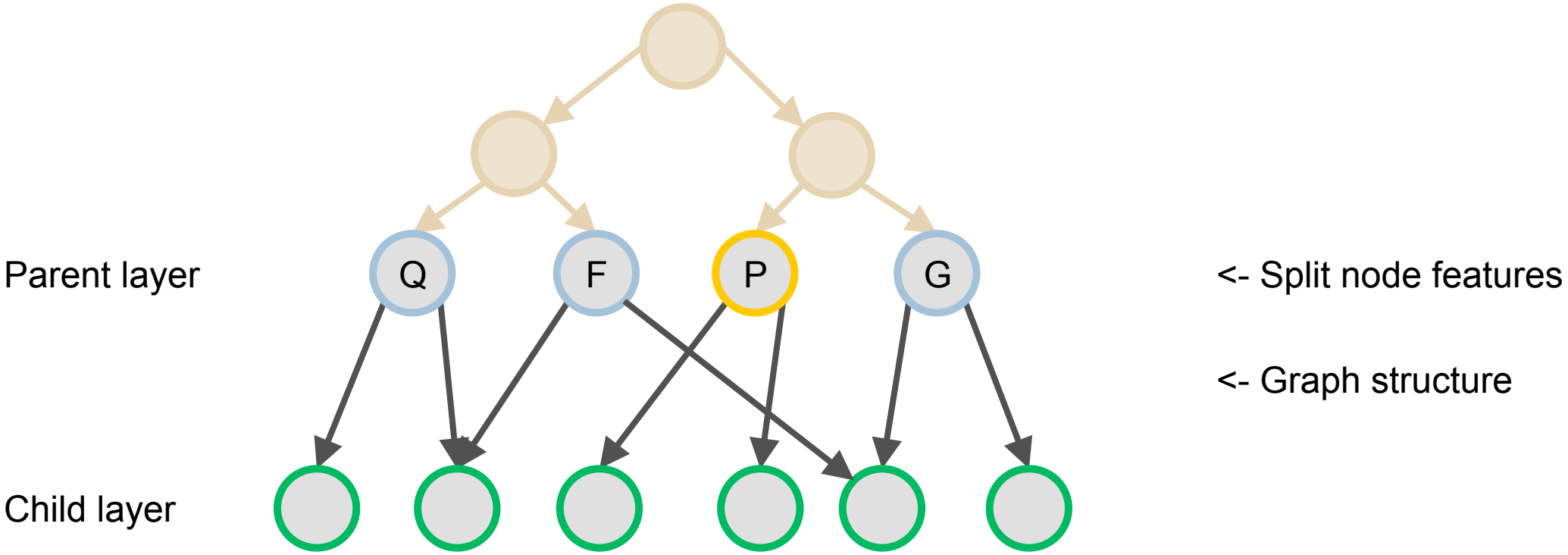
- **Train each DAG layer by layer**, jointly optimizing both
  - the structure of the DAG
  - the split node features themselves



Current objective score = 2.52

Algorithm overview

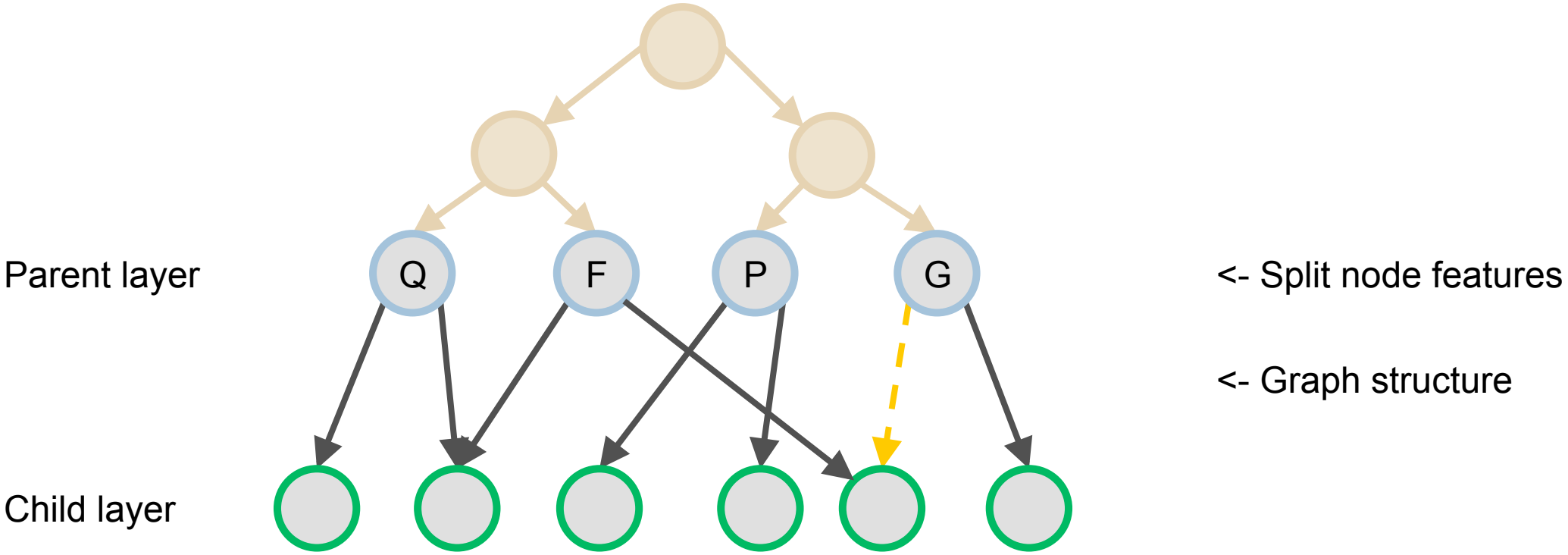
- **Train each DAG layer by layer**, jointly optimizing both
  - the structure of the DAG
  - the split node features themselves



Current objective score = 2.12

Algorithm overview

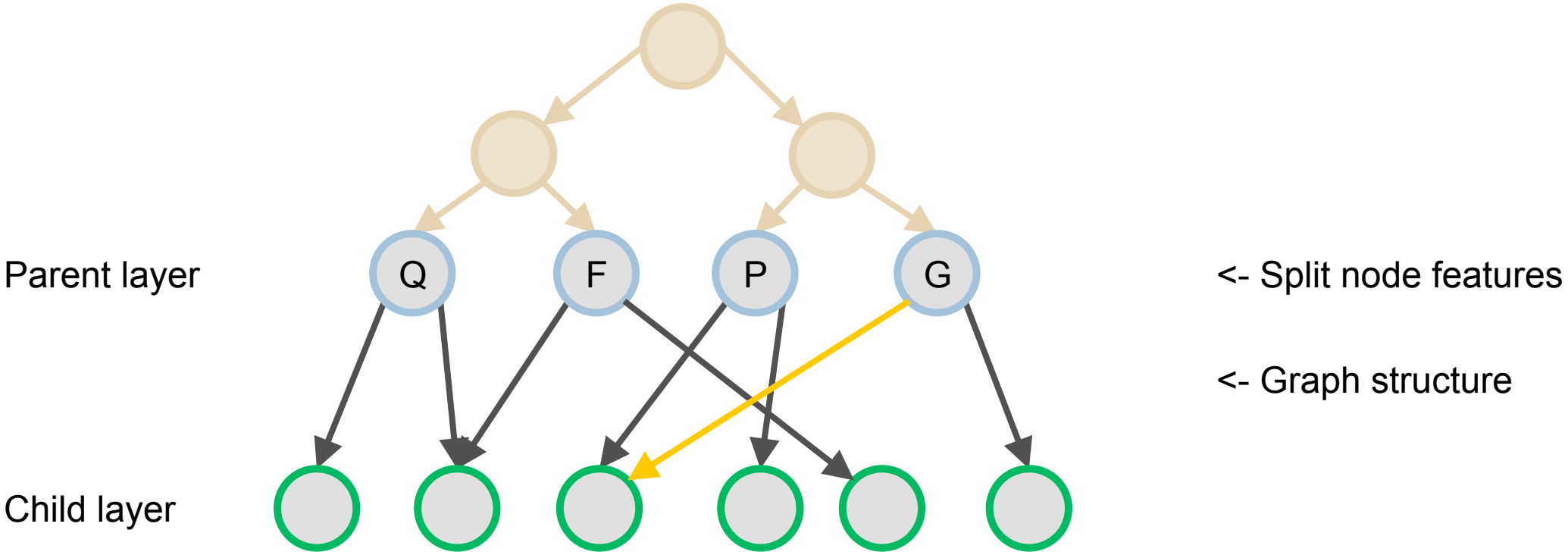
- **Train each DAG layer by layer**, jointly optimizing both
  - the structure of the DAG
  - the split node features themselves



Current objective score = 2.12

Algorithm overview

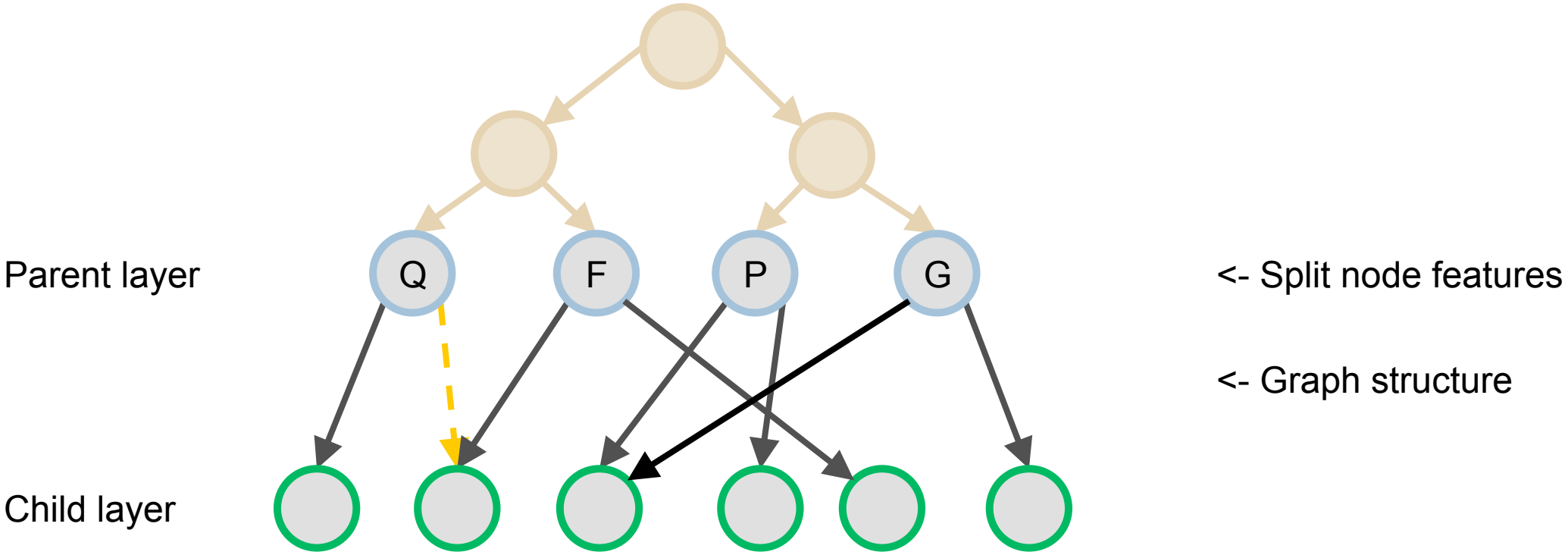
- **Train each DAG layer by layer**, jointly optimizing both
  - the structure of the DAG
  - the split node features themselves



Current objective score = 1.72

Algorithm overview

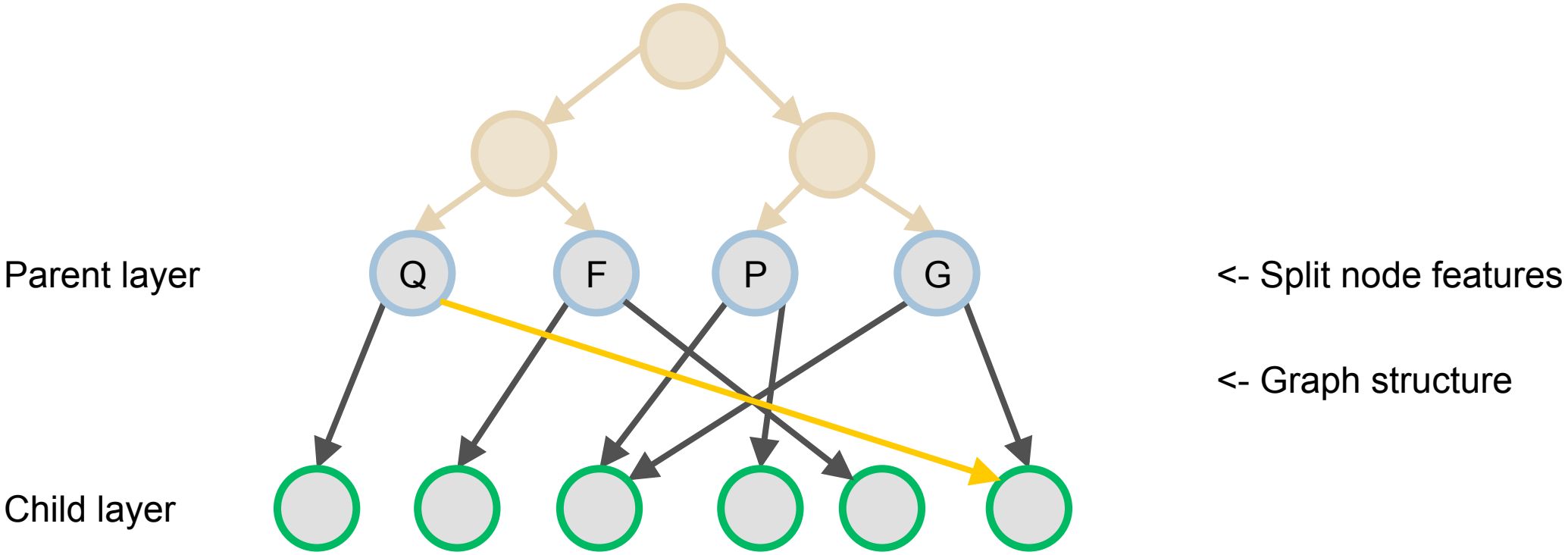
- **Train each DAG layer by layer**, jointly optimizing both
  - the structure of the DAG
  - the split node features themselves



Current objective score = 1.72

Algorithm overview

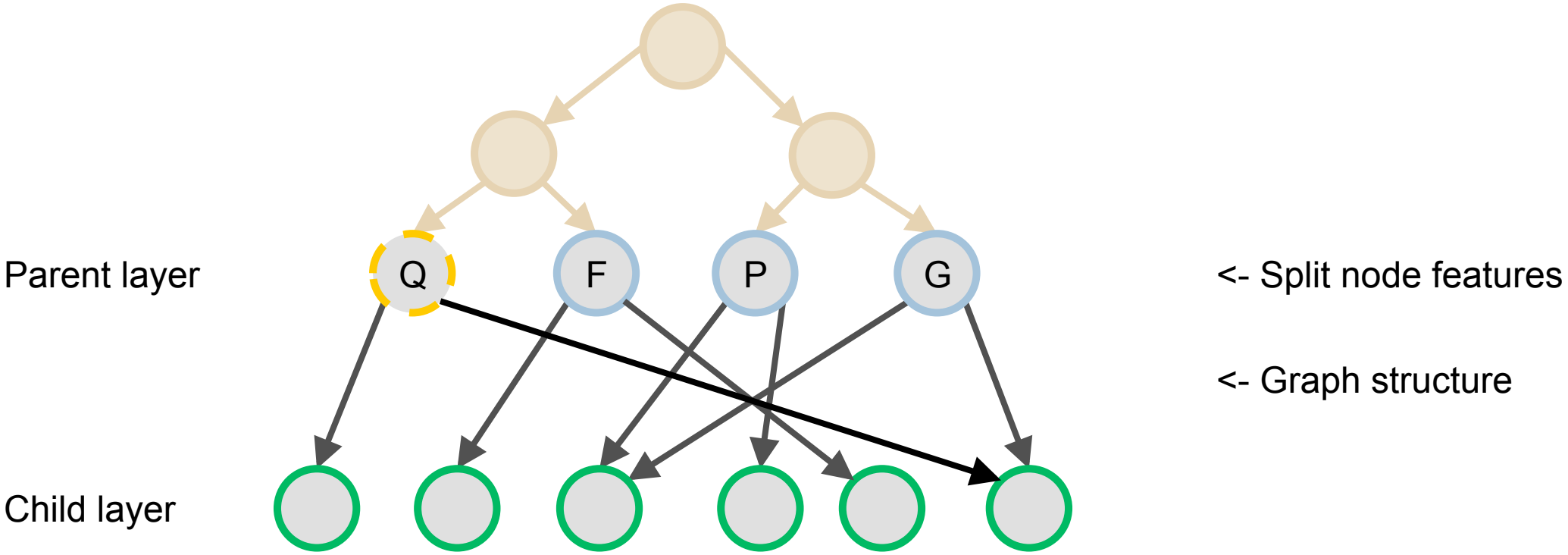
- **Train each DAG layer by layer**, jointly optimizing both
  - the structure of the DAG
  - the split node features themselves



Current objective score = 1.59

Algorithm overview

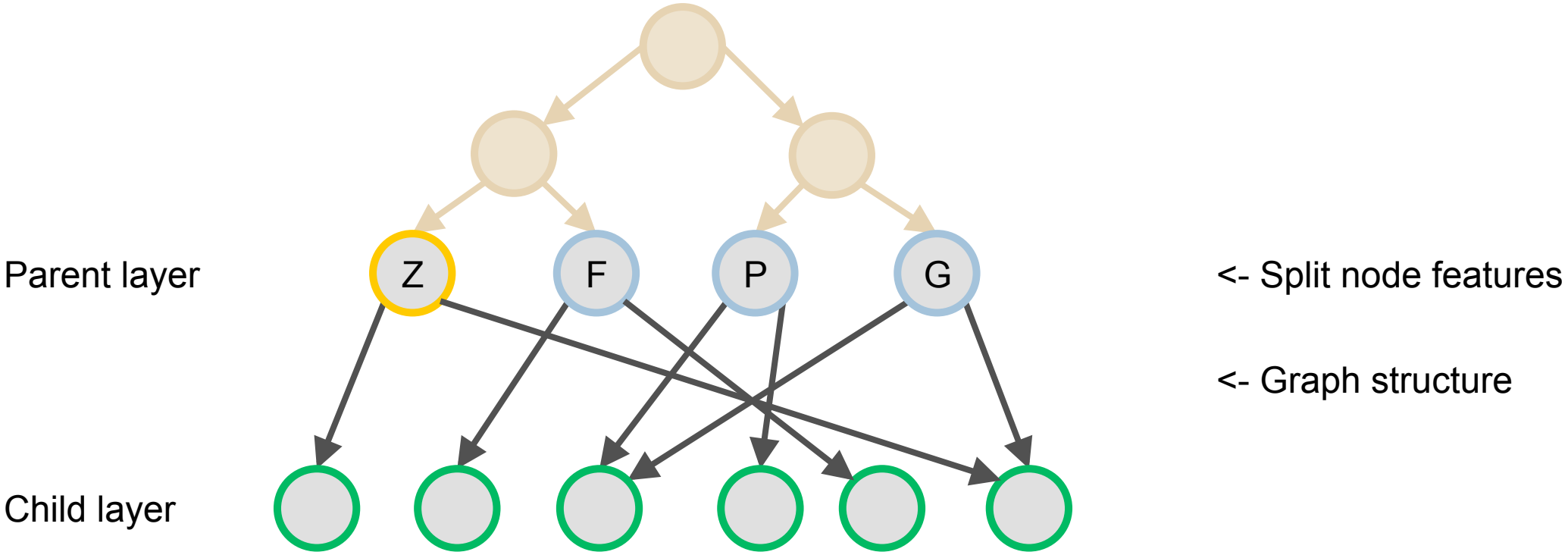
- **Train each DAG layer by layer**, jointly optimizing both
  - the structure of the DAG
  - the split node features themselves



Current objective score = 1.59

Algorithm overview

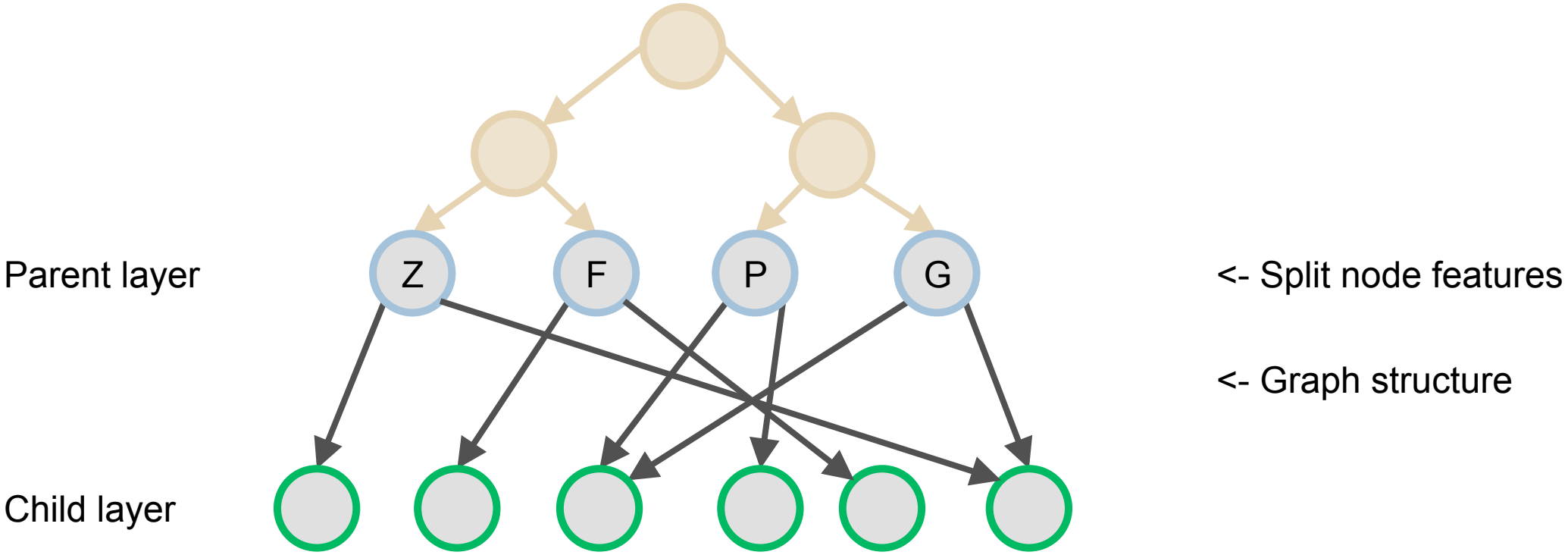
- **Train each DAG layer by layer**, jointly optimizing both
  - the structure of the DAG
  - the split node features themselves



Current objective score = 1.44

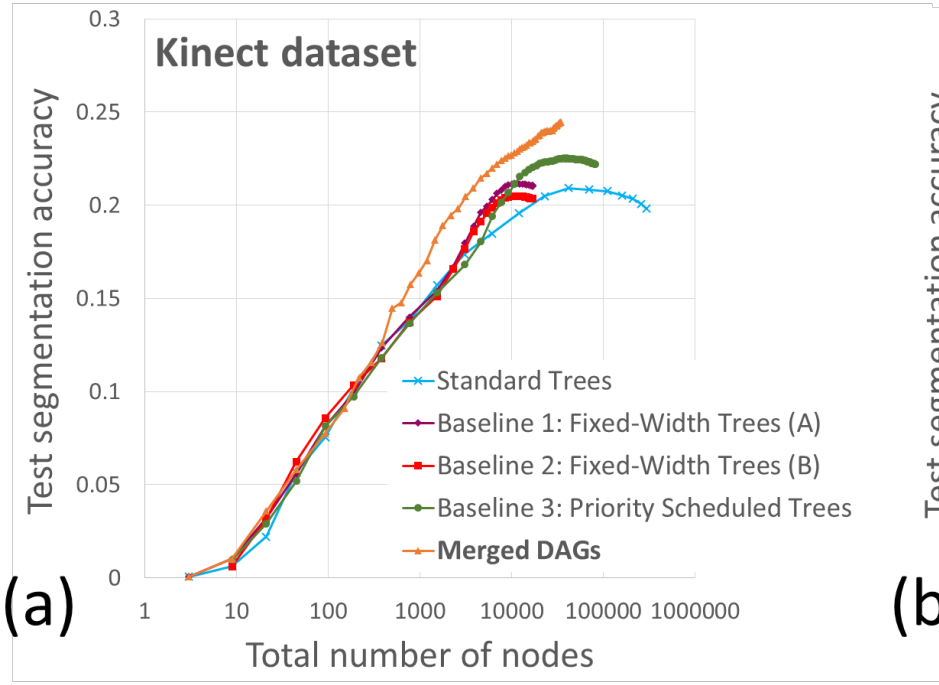
Algorithm overview

- **Train each DAG layer by layer**, jointly optimizing both
  - the structure of the DAG
  - the split node features themselves

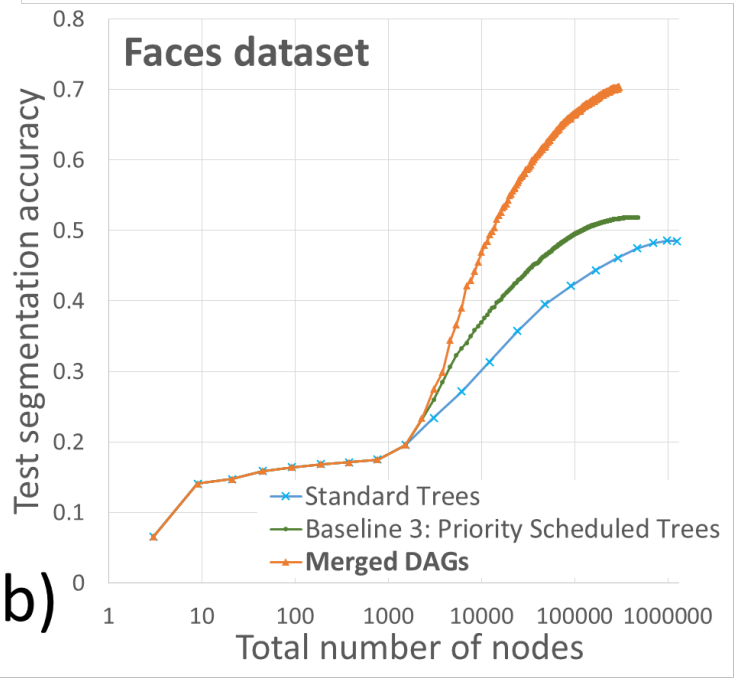


Current objective score = 1.44

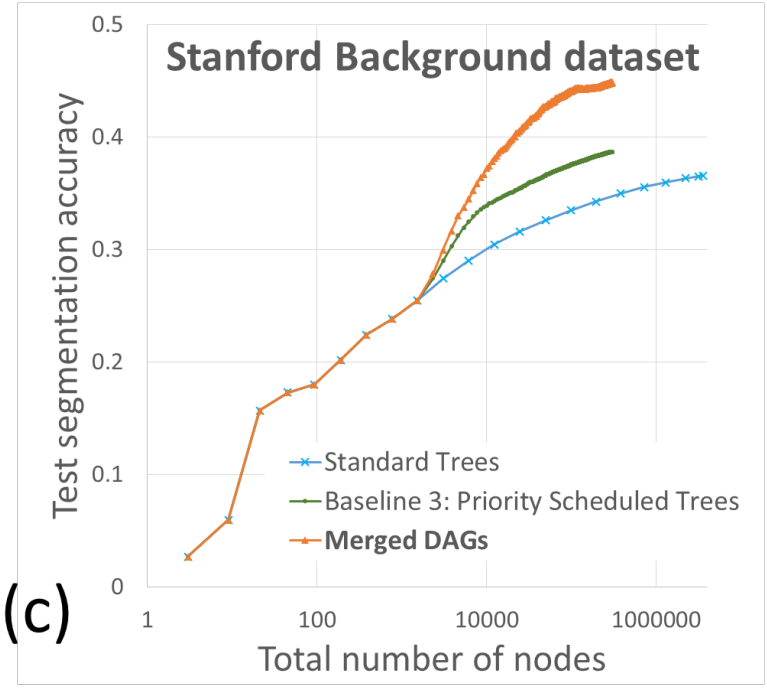
Jungles: results



(a)

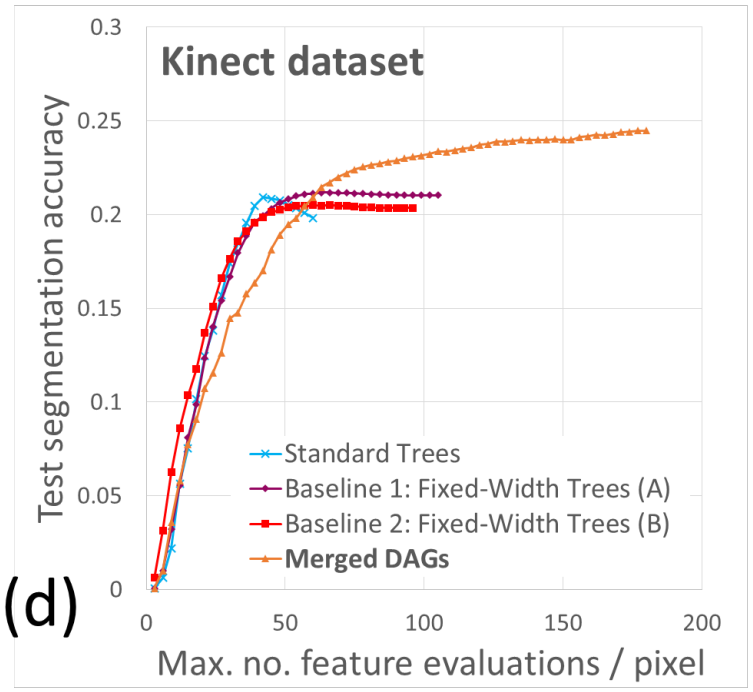


(b)

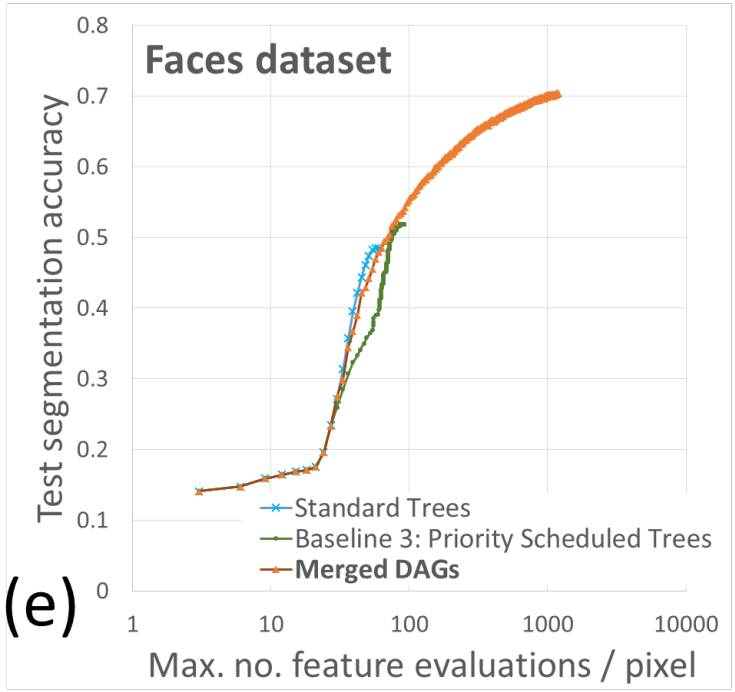


(c)

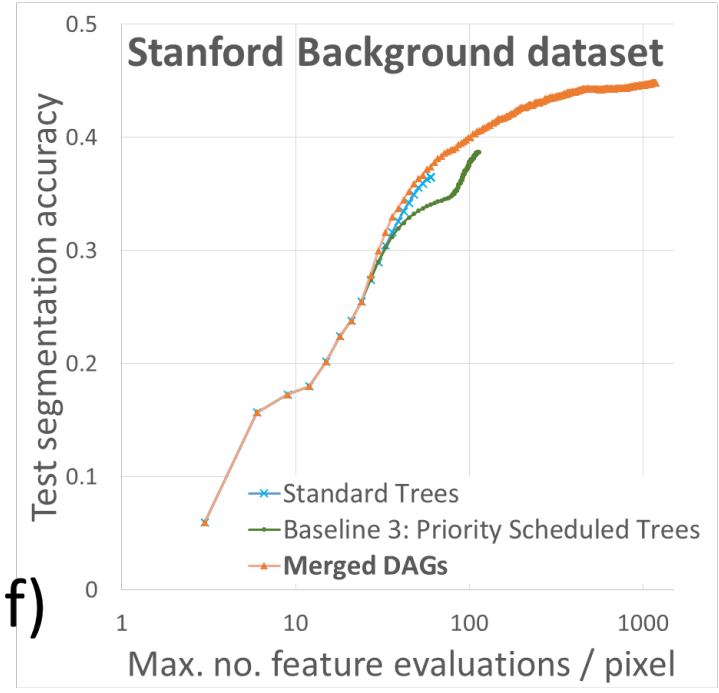
Jungles: accuracy *vs.* compute time



(d)

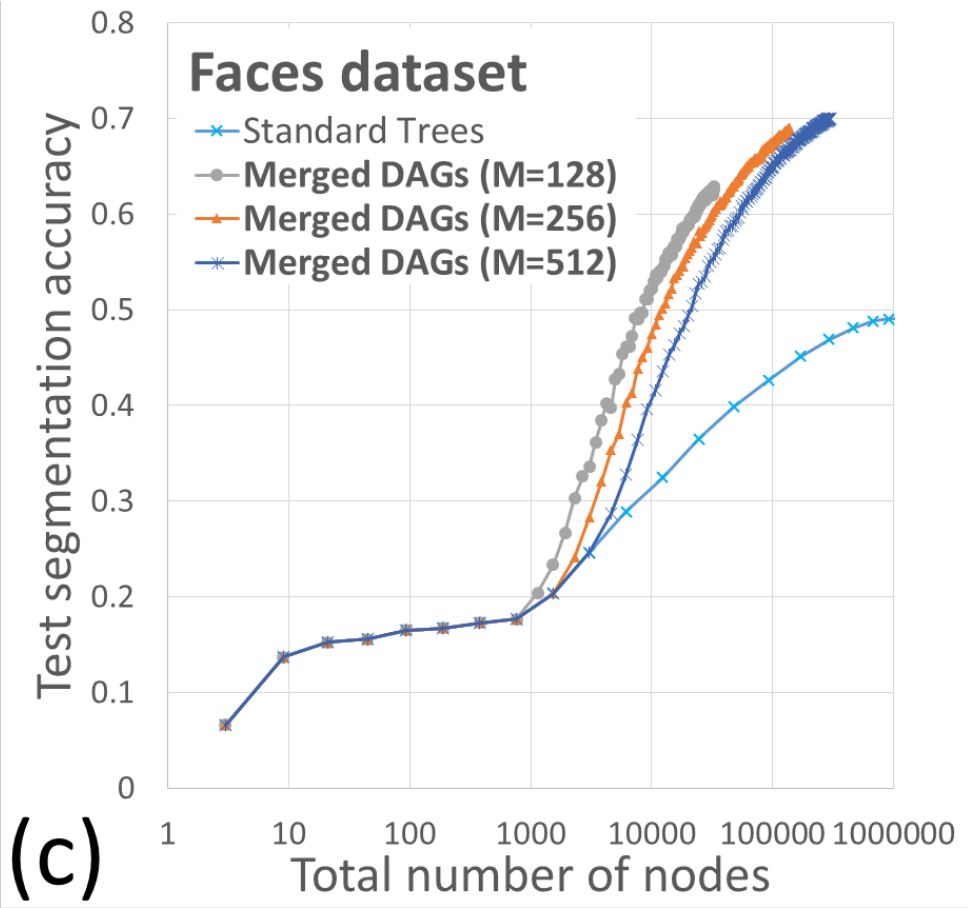


(e)



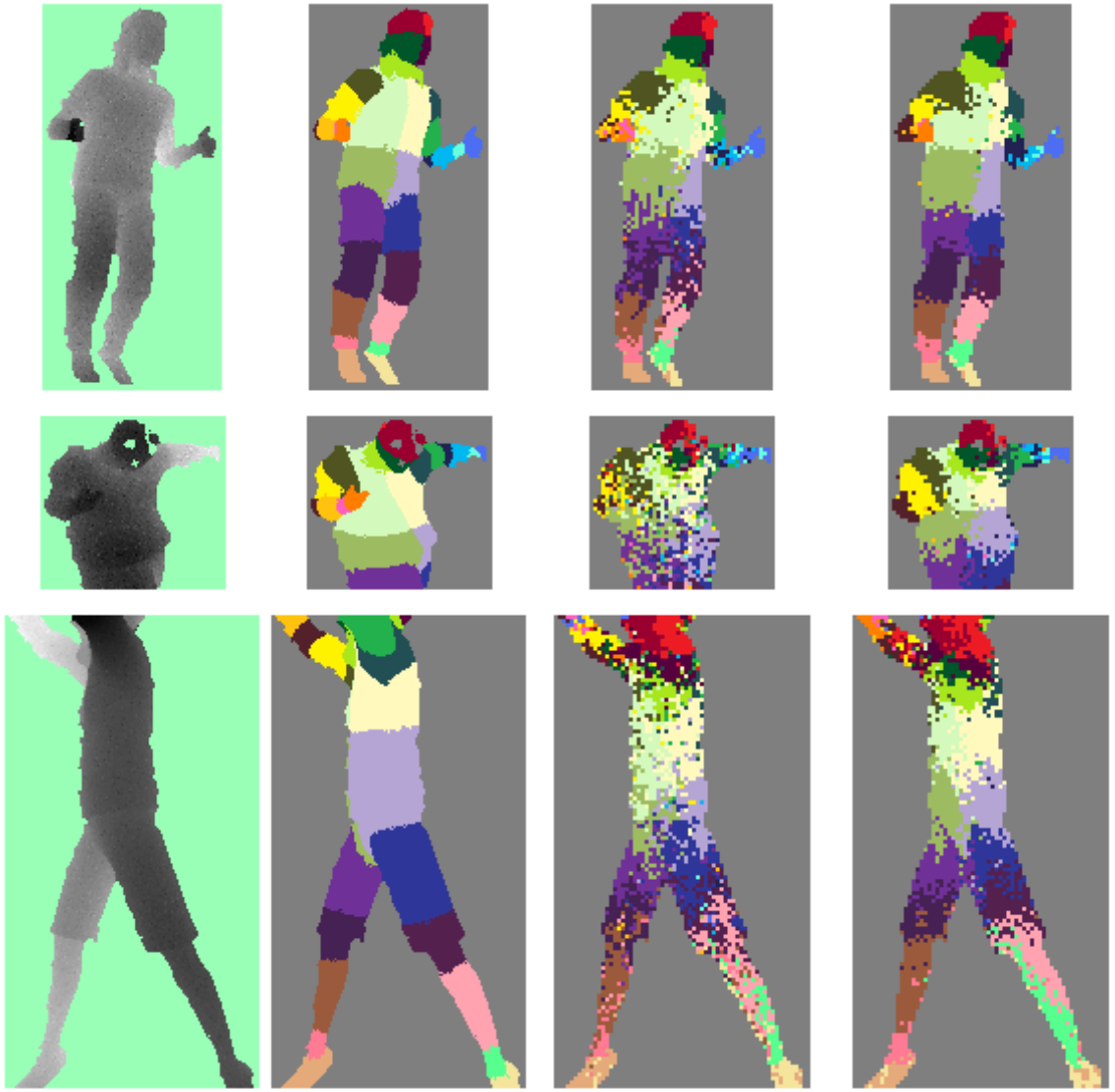
(f)

Jungles: node budget *M*

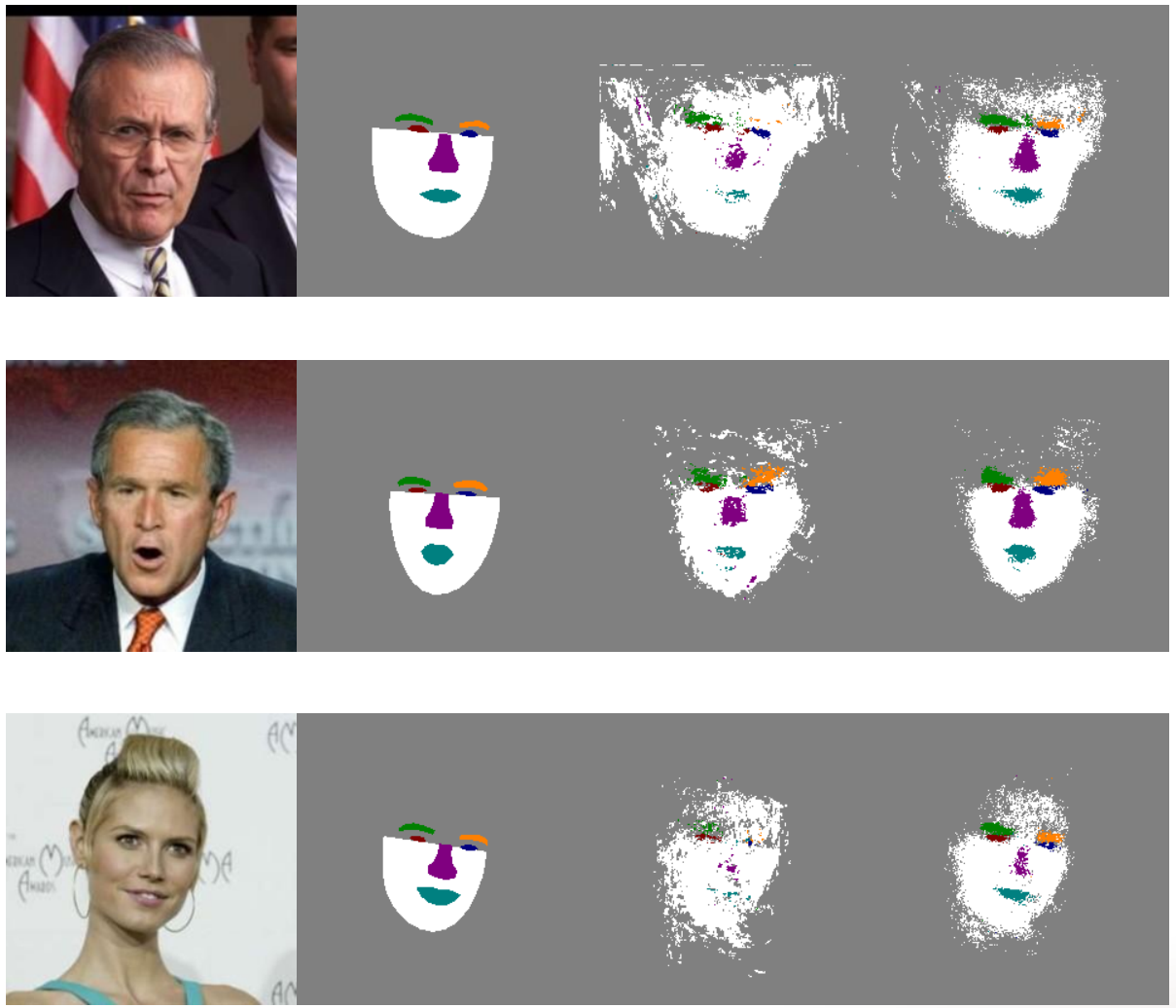


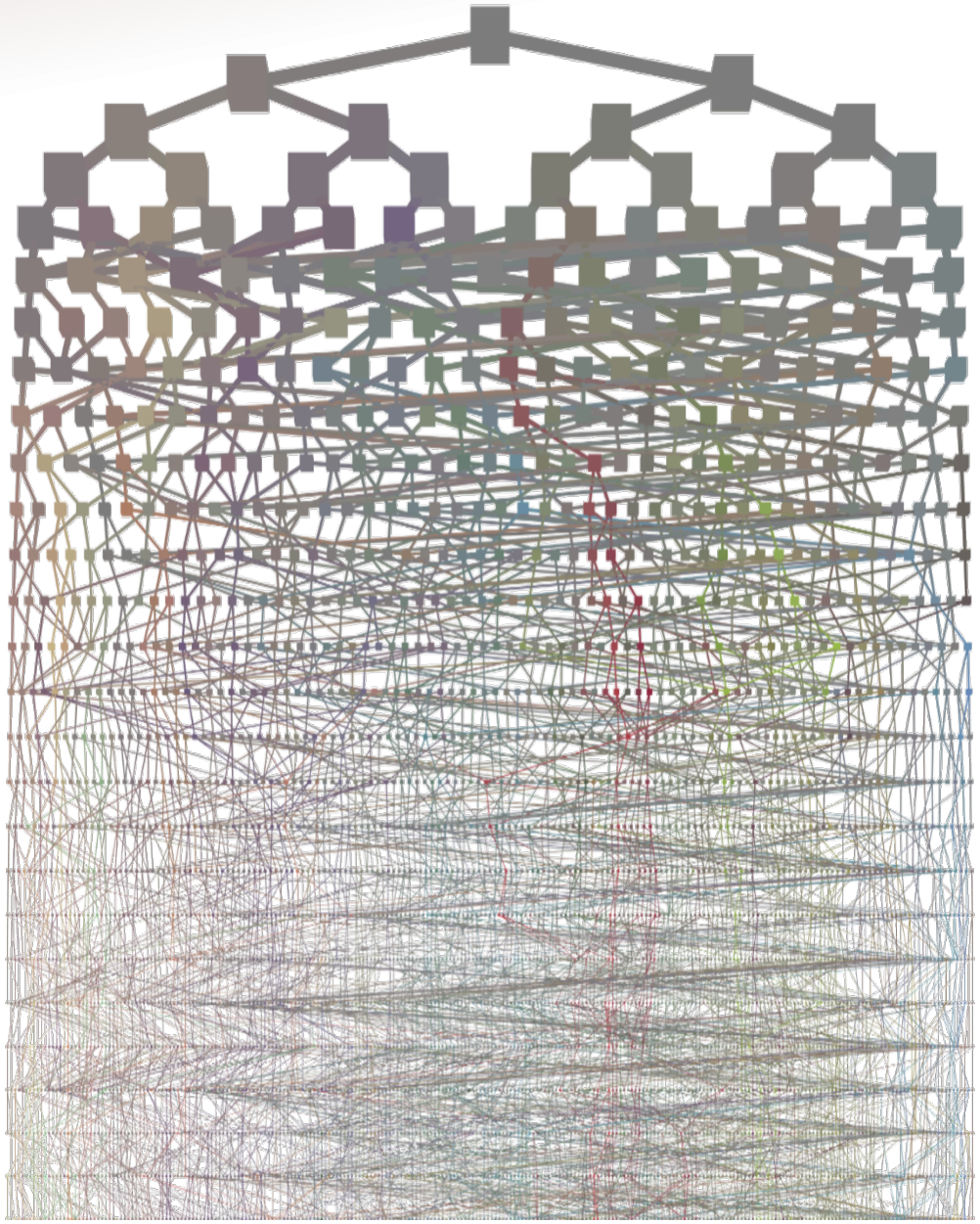
Jungles: results

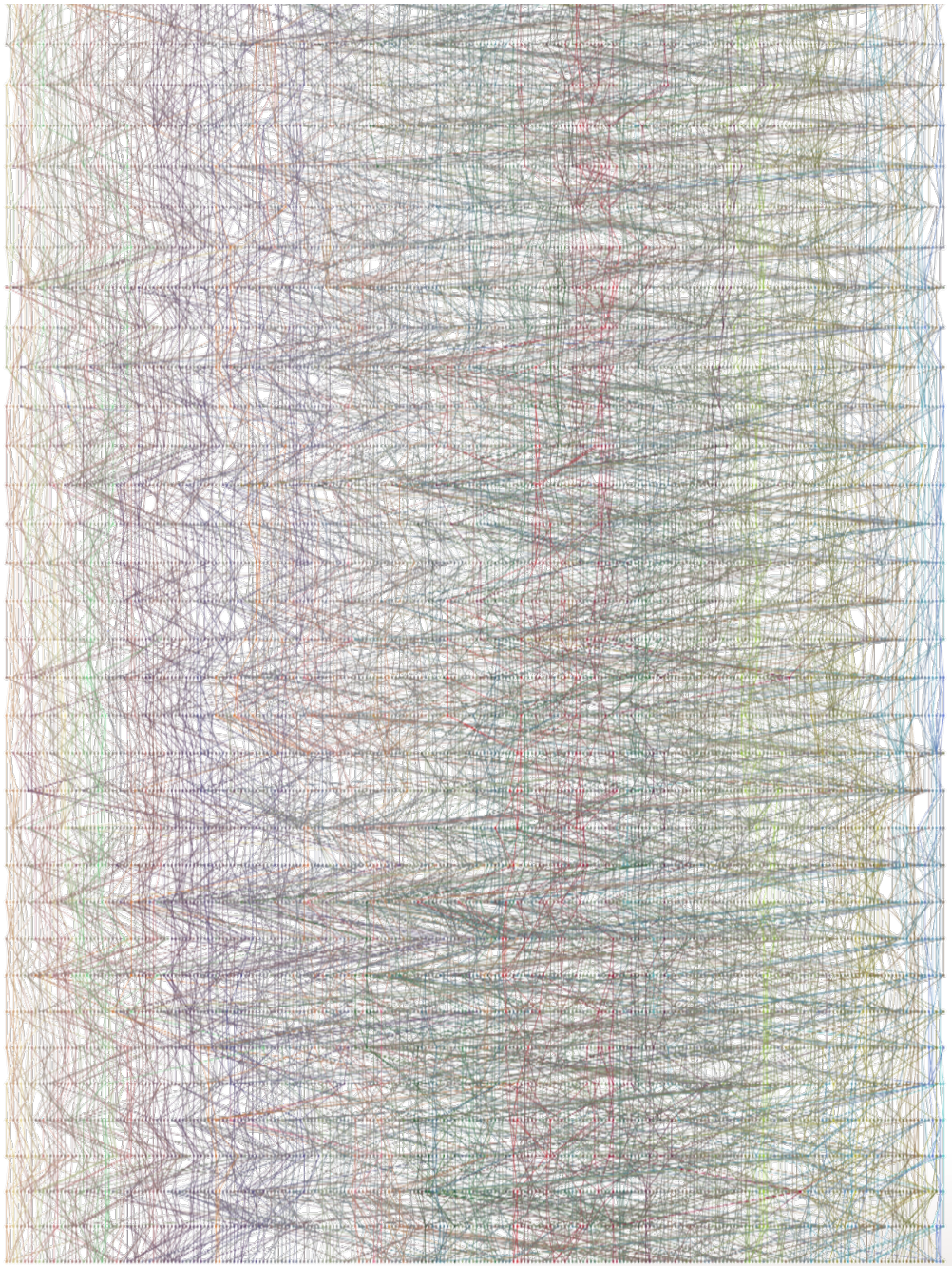
Input Image	Ground Truth	**Standard Trees** Segmentation	**Merged DAGs** Segmentation



Input Image	Ground Truth	**Standard Trees** Segmentation	**Merged DAGs** Segmentation





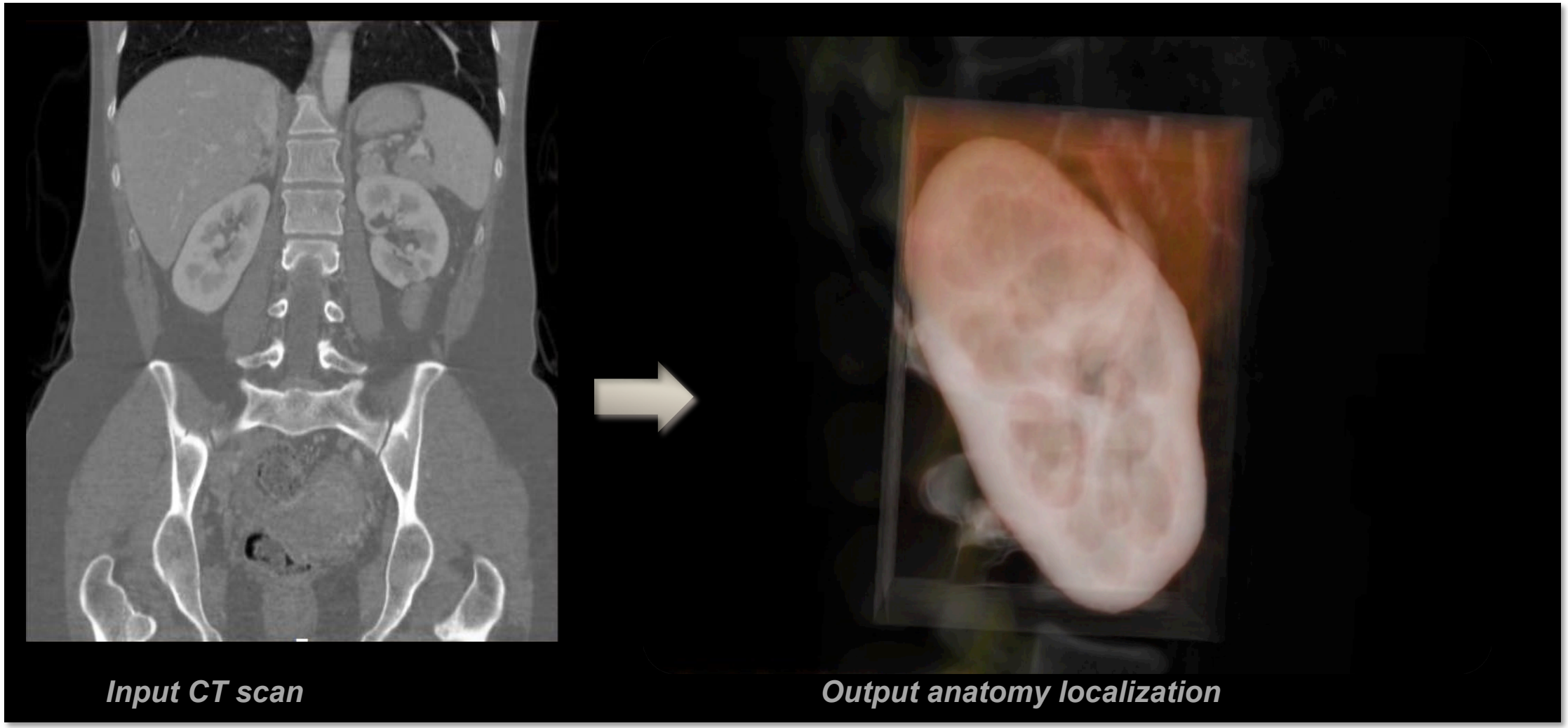


Talk overview

- A brief introduction to machine learning
- Decision forests and jungles
- **Applications in medical image analysis**
  - **Anatomy localization**
  - Spine detection
  - Brain tumour segmentation
  - Learned image super-resolution
  - Quantifying progression of multiple sclerosis

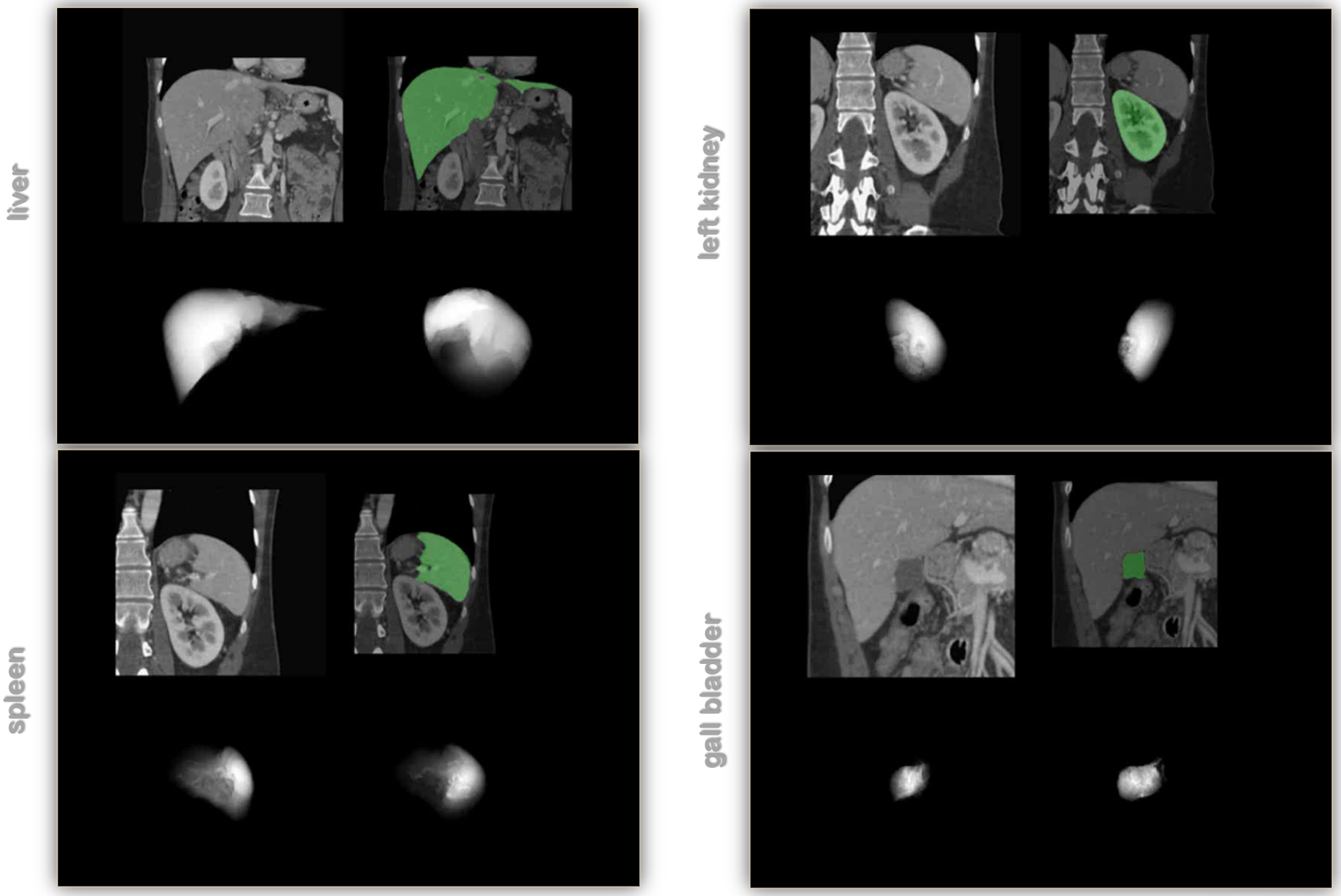
A. Criminisi, D. Robertson, E. Konukoglu, J. Shotton, S. Pathak, S. White, and K. Siddiqui, **Regression Forests for Efficient Anatomy Detection and Localization in Computed Tomography Scans**, in *Medical Image Analysis (MedIA)*, Elsevier, 2013

Anatomy Localization in 3D Computed Tomography Scans



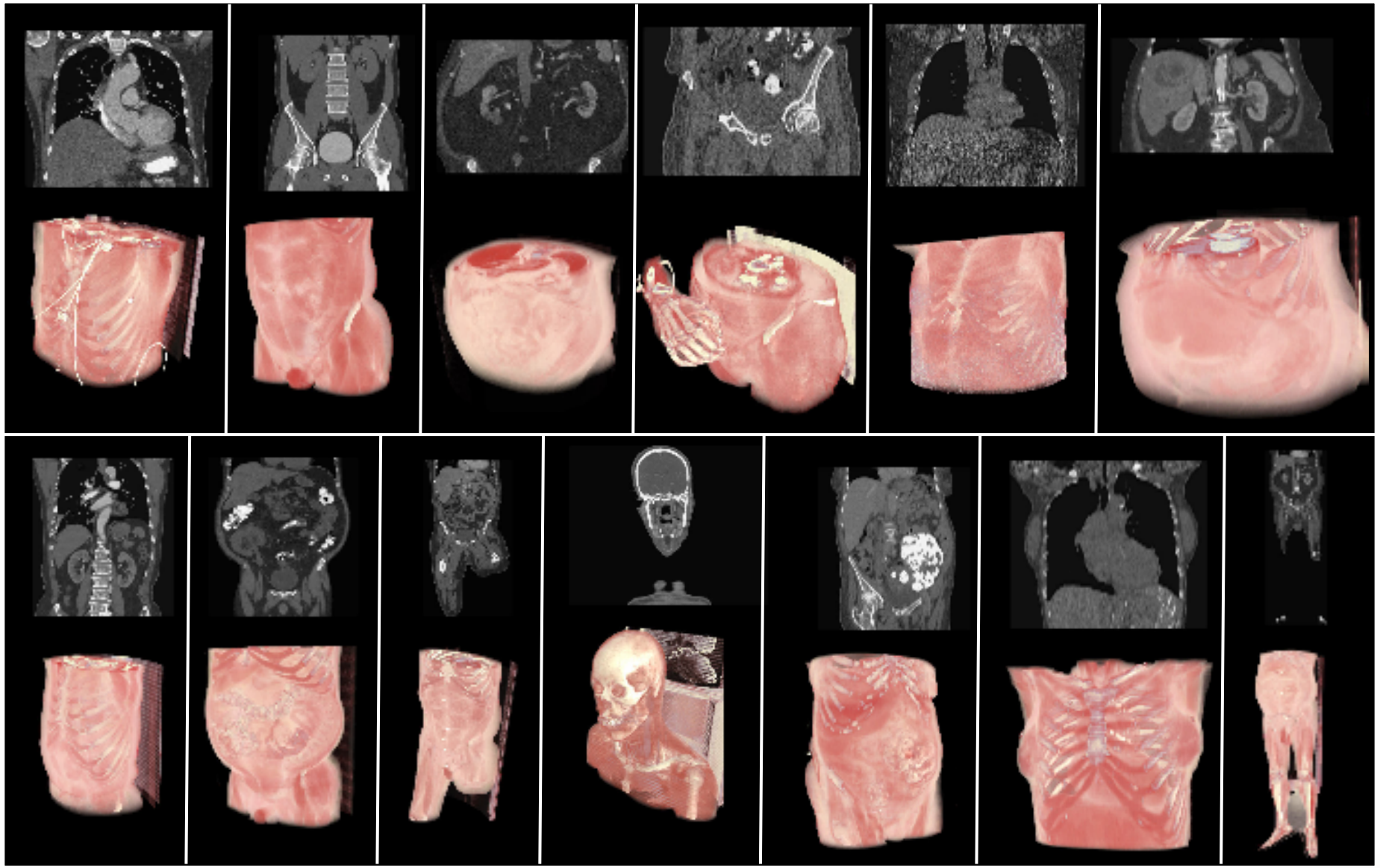
- Direct mapping of voxels to organ bounding boxes.
- No search, no sliding window.
- No need for registration. No other pre-processing steps.

Anatomy localization: why is it hard?



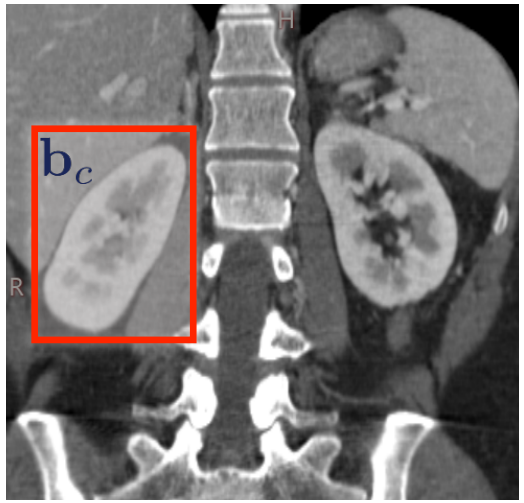
High variability in appearance, shape, location, resolution, noise, pathologies …

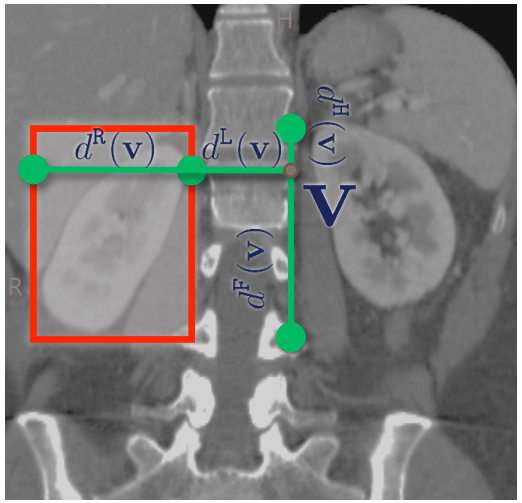
Anatomy localization: the ground-truth database

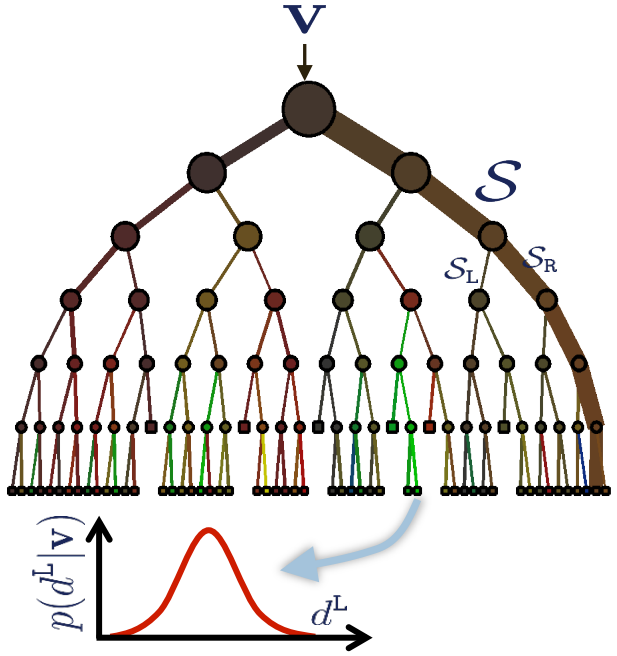


Different image cropping, noise, contrast/no-contrast, resolution, scanners, body shapes/sizes, patient position…

Anatomy localization: regression forest







- **Each voxel** in the volume votes for the position of the 6 box sides
- We wish to learn a set of **discriminative points** (landmarks, clusters) which can predict the kidney position with **high confidence.**

Input data point $\mathbf{v} = (v_x \; v_y \; v_z)$ *(voxel position in volume)*

Output $\mathbf{b}_c = (b_c^{\mathrm{L}} \; b_c^{\mathrm{R}} \; b_c^{\mathrm{A}} \; b_c^{\mathrm{P}} \; b_c^{\mathrm{H}} \; b_c^{\mathrm{F}})$ *(bound. box continuous pos.)*

Multiple organs $c \in \{\mathtt{liver}, \mathtt{spleen}, \mathtt{l.-kidney}, \mathtt{r.-kidney}...\}$

Node split function $\xi_j > f(\mathbf{v}; \theta_j) > \tau_j$

Node optimization $IG = \frac{1}{2}\left(\sum_c p(c;\mathcal{S}) \log|\Lambda_c(\mathcal{S})| - \sum_{i\in\{\mathrm{L},\mathrm{R}\}} \omega_i \sum_c p(c;\mathcal{S}_i) \log|\Lambda_c(\mathcal{S}_i)|\right)$

Node training $(\theta^*, \xi^*, \tau^*) = \max_{\theta,\xi,\tau} IG$

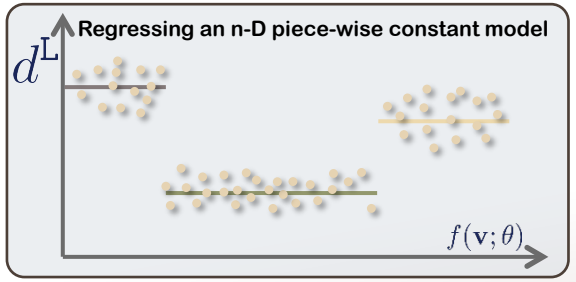
Feature response $f(\mathbf{v}, \theta_j) = \frac{1}{|B_j|} \sum_{\mathbf{p}\in B_j} I(\mathbf{p})$ *(mean over displaced 3D boxes)*

Error in model fit

$\sum_c p(c;\mathcal{S}) \log|\Lambda_c(\mathcal{S})|$ *(weighted uncertainty for all organs)*

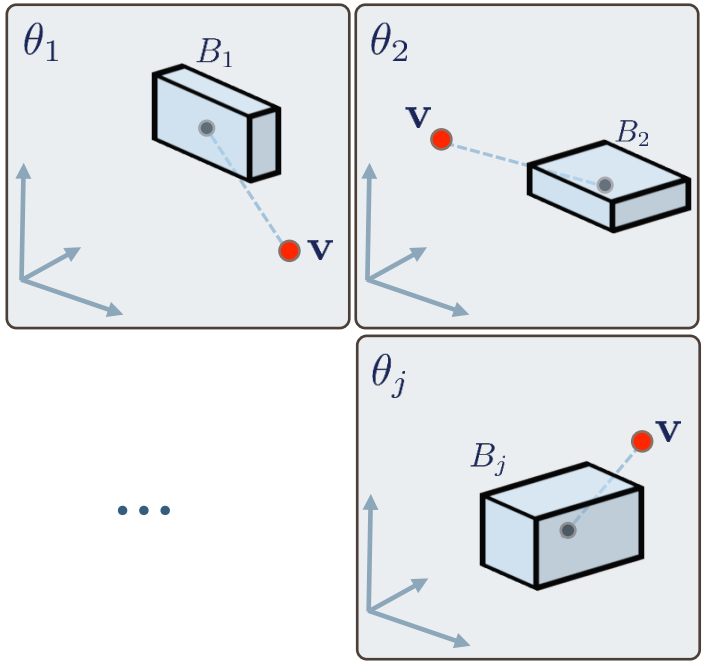
$\mathbf{d}_c(\mathbf{v}) = (d_c^{\mathrm{L}} \; d_c^{\mathrm{R}} \; d_c^{\mathrm{A}}; d_c^{\mathrm{P}} \; d_c^{\mathrm{H}}; d_c^{\mathrm{F}})$ *(relative displacement)*

$p(\mathbf{d}) = \mathcal{N}(\mathbf{d}; \overline{\mathbf{d}}, \Lambda)$ *(Gaussian repres. of distribs)*

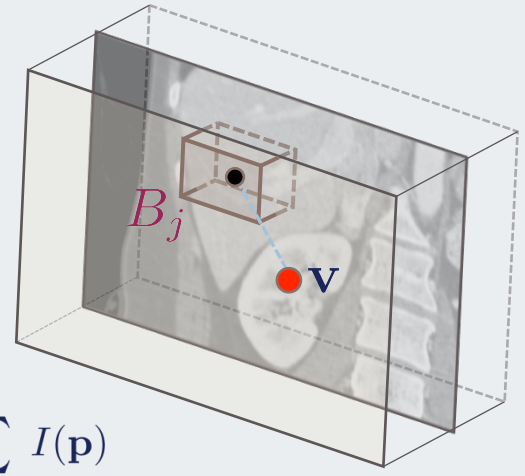


Anatomy localization: context-rich visual features

Possible visual features



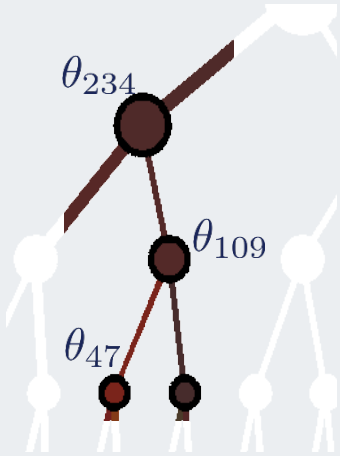
Computing the feature response

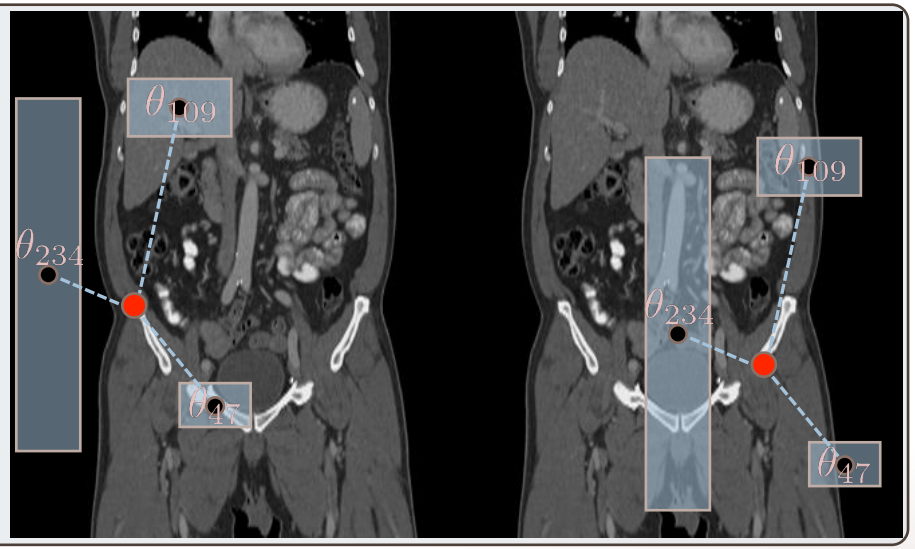


$$\mathbf{v} = (x\ y\ z)$$

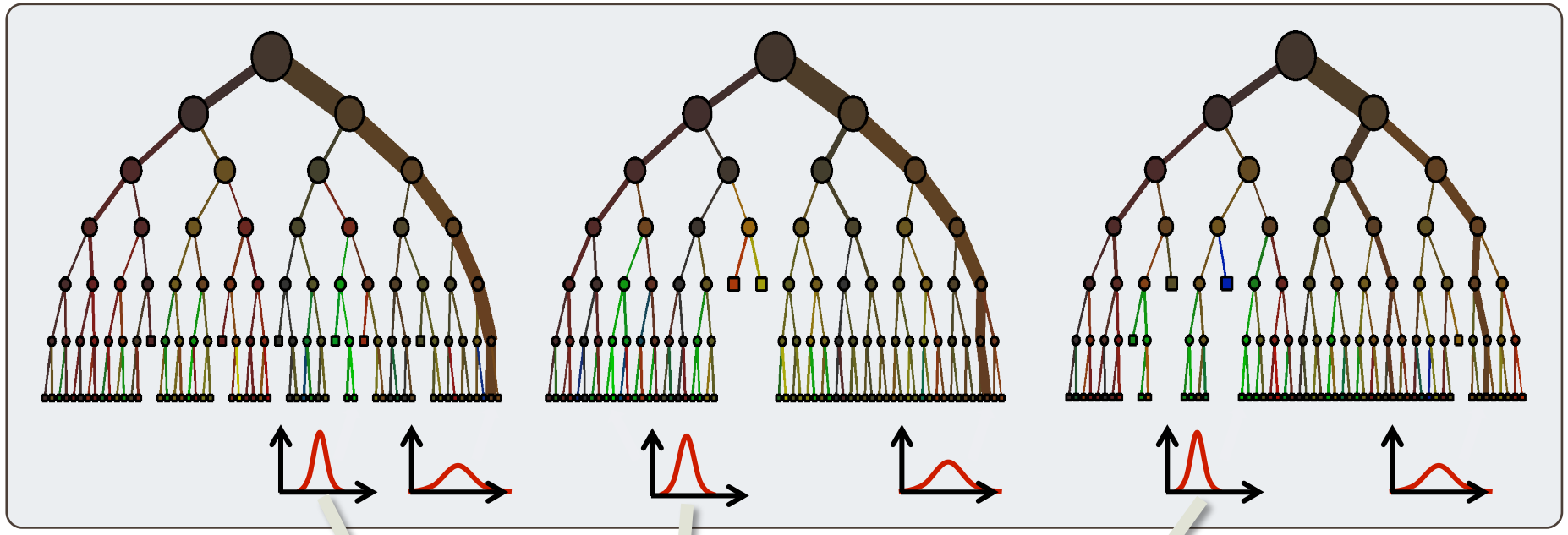
$$f(\mathbf{v}, \theta_j) = \frac{1}{|B_j|} \sum_{\mathbf{p} \in B_j} I(\mathbf{p})$$

Capturing spatial context





Anatomy localization: automatic landmark discovery



Input CT scan and detected landmark regions

Here the system is trained to detect left and right kidneys.

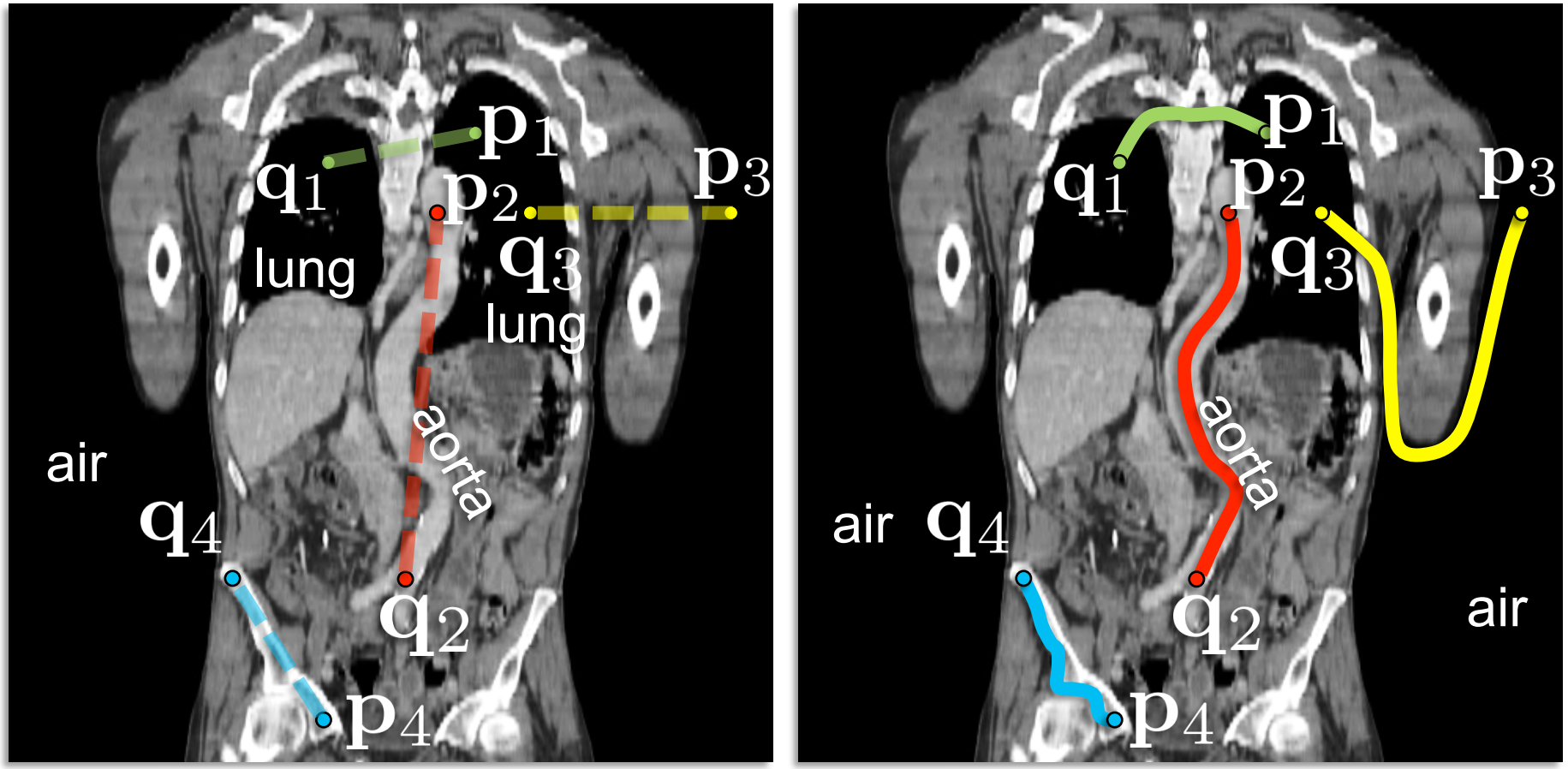
The system learns to use bottom of lung and top of pelvis to localize kidneys with highest confidence.

Talk overview

- A brief introduction to machine learning
- Decision forests and jungles
- **Applications in medical image analysis**
  - Anatomy localization
  - **Anatomy segmentation**
  - Spine detection
  - Brain tumour segmentation
  - Learned image super-resolution
  - Quantifying progression of multiple sclerosis

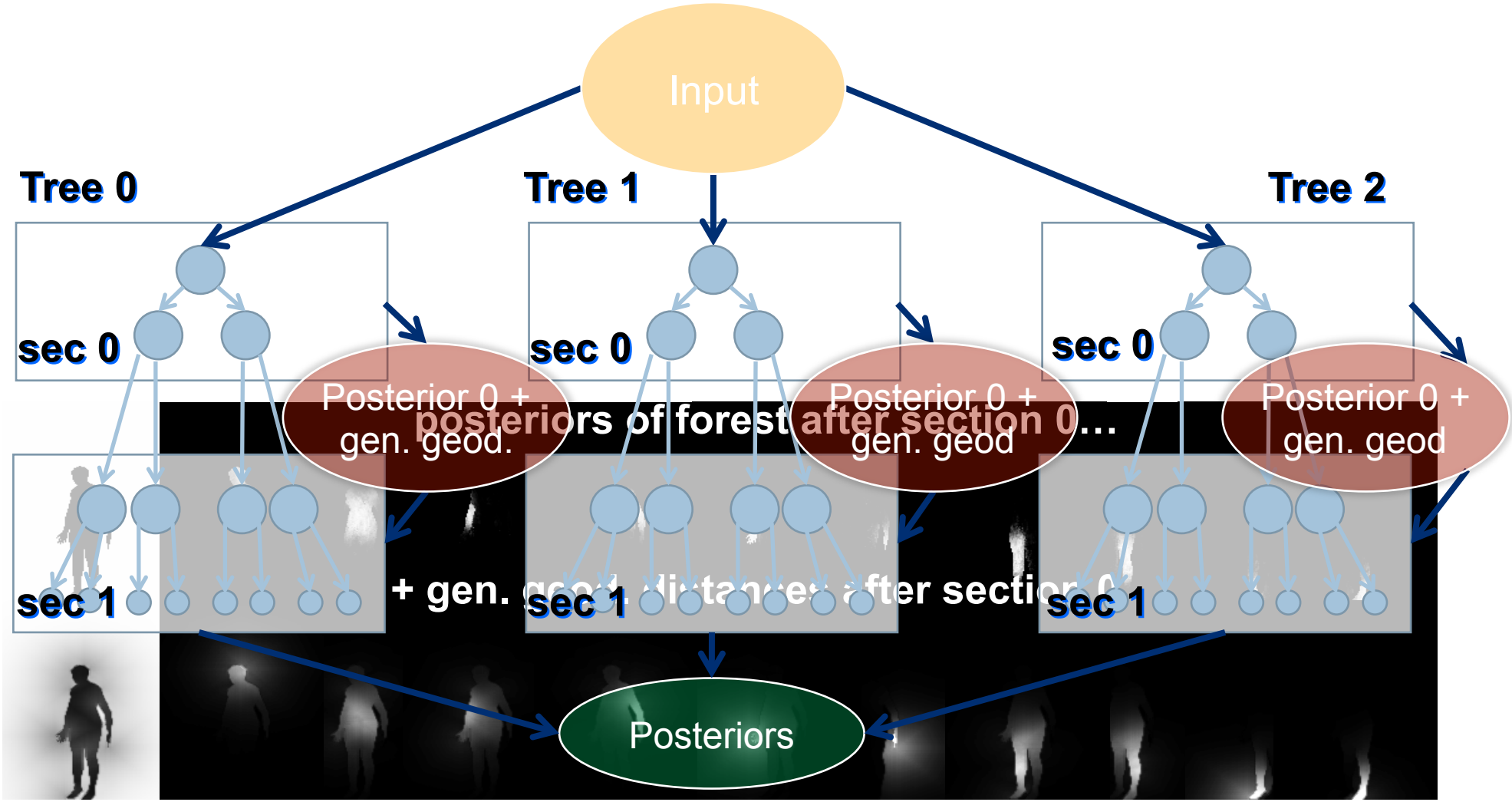
P. Kontschieder, P. Kohli, J. Shotton, and A. Criminisi, **GeoF: Geodesic Forests for Learning Coupled Predictors,** in *Proc. Computer Vision and Pattern Recognition (CVPR)*, IEEE, June 2013

Entangled geodesic forests for semantic segmentation



- Using soft connectivity features efficiently
- Capturing semantic context
- No need for Markov-, Conditional Random Field post-processinf

Entangled geodesic forests for semantic segmentation



Entangled geodesic forests for semantic segmentation

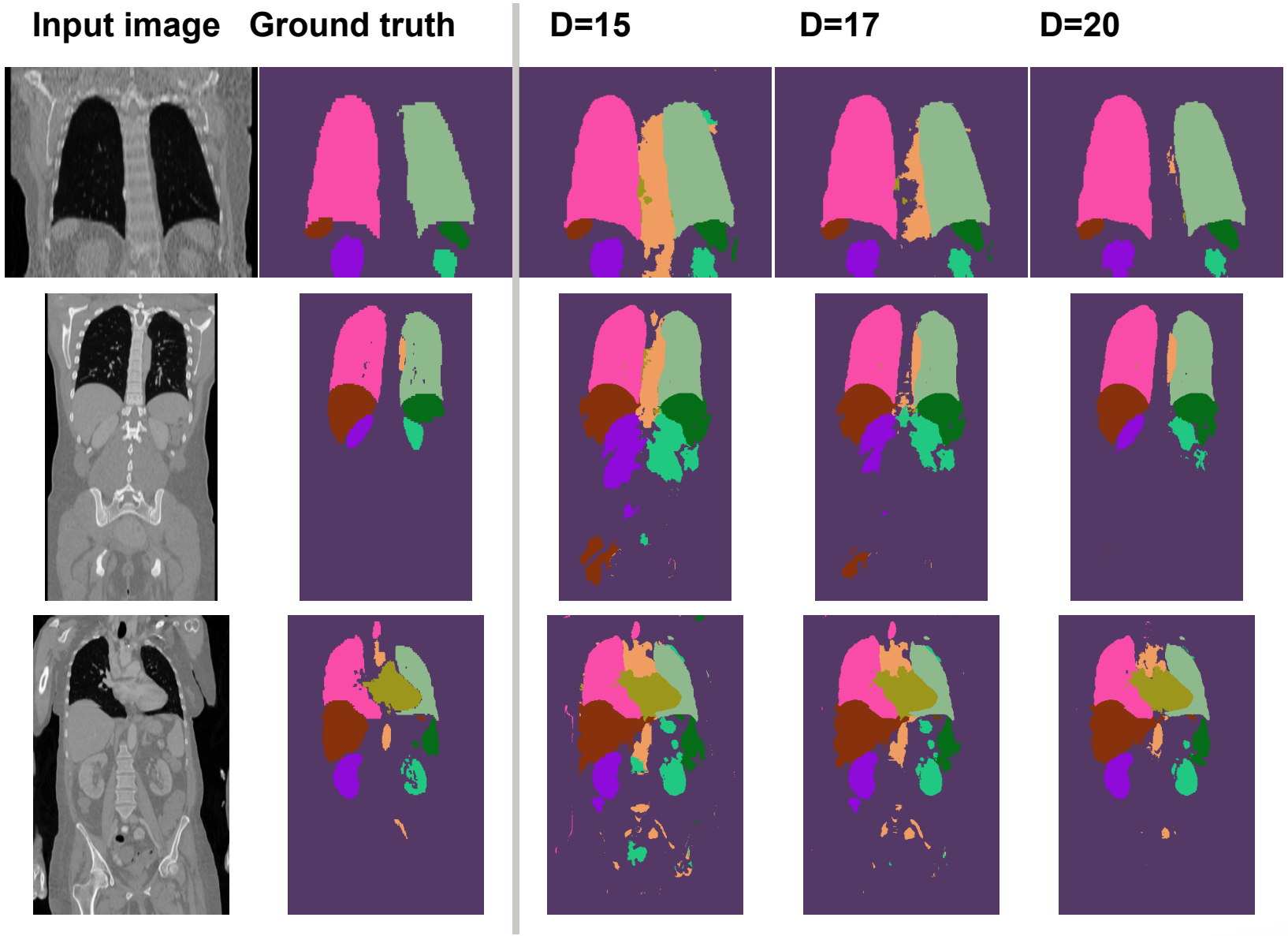
Algorithm	Jaccard
Conventional Classification Forest	53.2
Classification forest + (CRF)	68.3
Auto-context classification forest	65.9
Entangled classification forest	58.3
Auto-context geodesic forests	69.2
Entangled geodesic forests	**72.3**

Input

Ground truth

Our results

Entangled geodesic forests for semantic segmentation



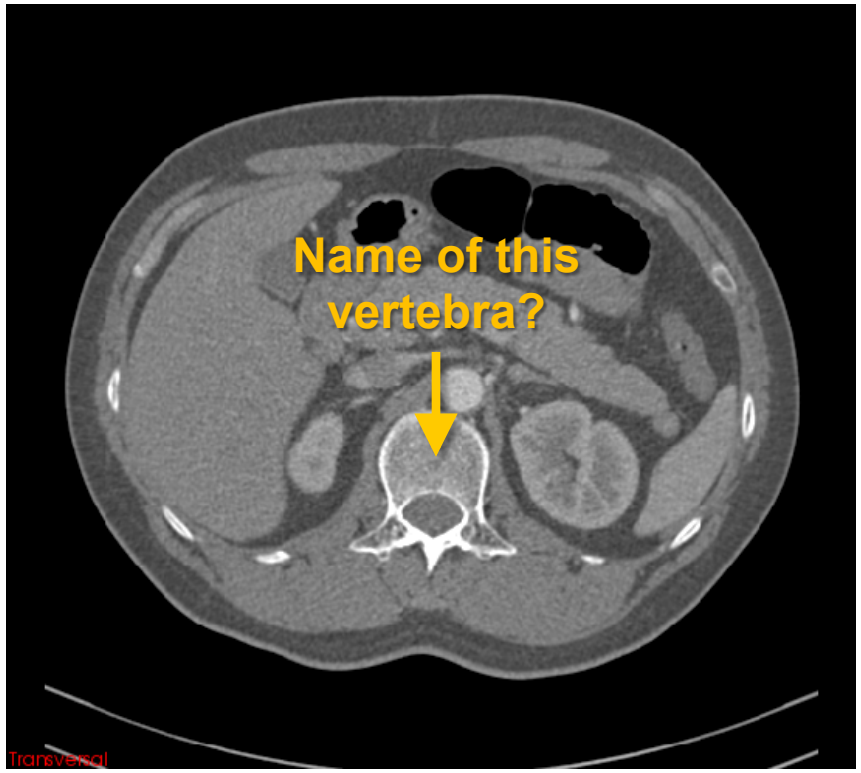
Automatically correcting for over-use of context

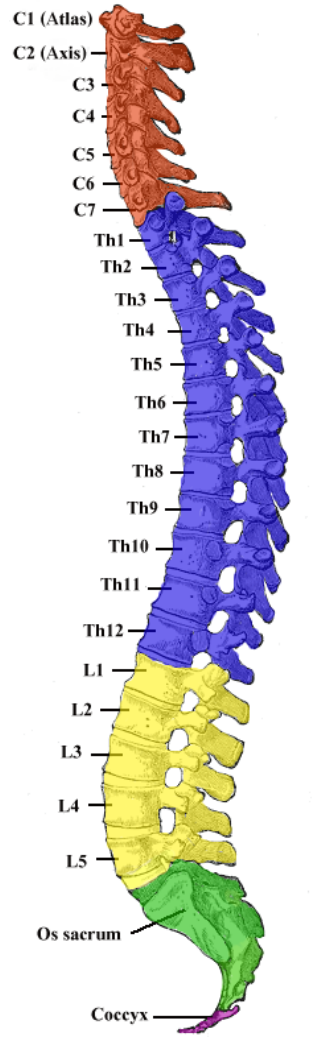
Talk overview

- A brief introduction to machine learning
- Decision forests and jungles
- **Applications in medical image analysis**
  - Anatomy localization
  - **Spine detection**
  - Brain tumour segmentation
  - Learned image super-resolution
  - Quantifying progression of multiple sclerosis

B. Glocker, D. Zikic, E. Konukoglu, D. R. Haynor, and A. Criminisi, **Vertebrae Localization in Pathological Spine CT via Dense Classification from Sparse Annotations**, in *MICCAI 2013*, Springer, September 2013

Vertebrae localization and classification

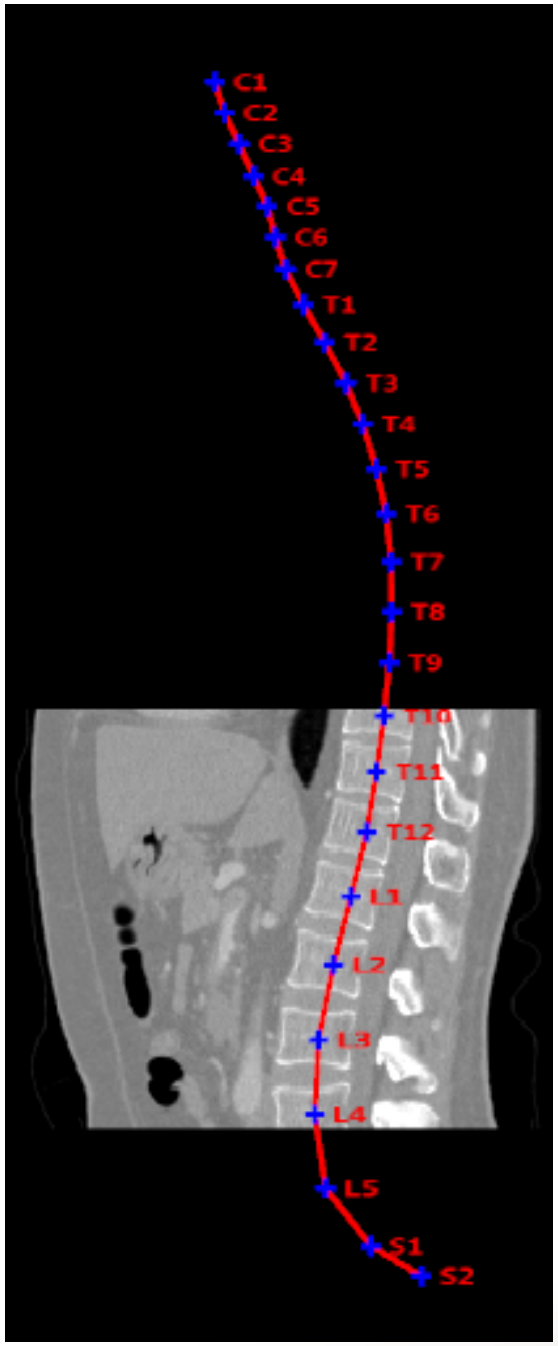




Clinical motivation

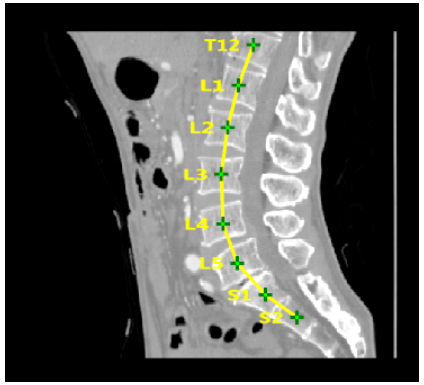
Extracting a patient-specific coordinate system

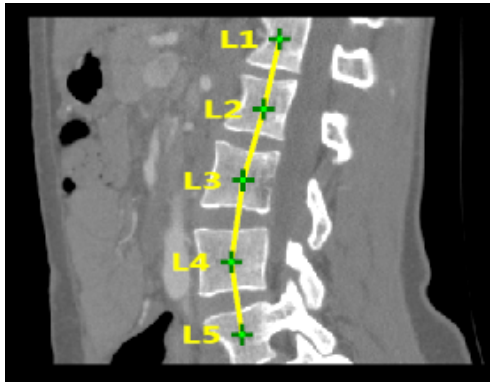
- Guided visualization/navigation in diagnostic tools
- Longitudinal assessment after surgical Intervention
- Shape/population analysis for disease modelling

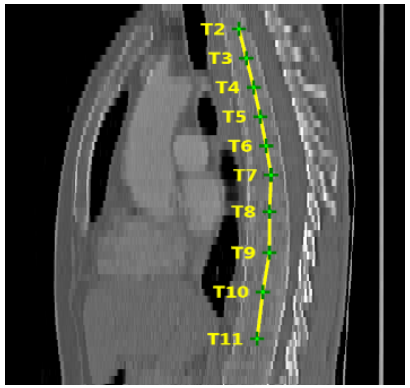


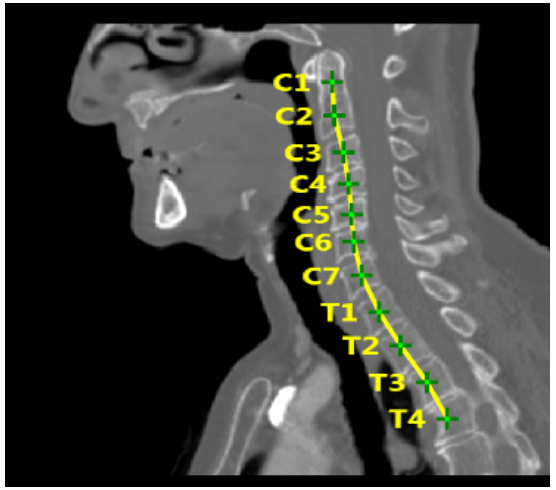
Challenges

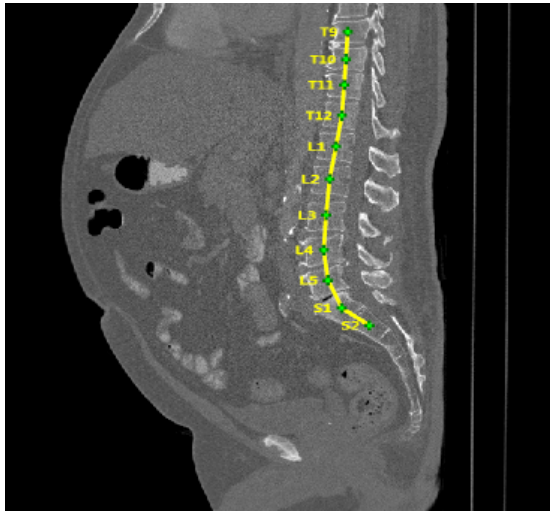
- Repetitive nature of structures
- Variability from normal anatomy
- Presence of pathologies and/or external objects
- Varying image acquisition
  (FOV, noise level, resolution, …)

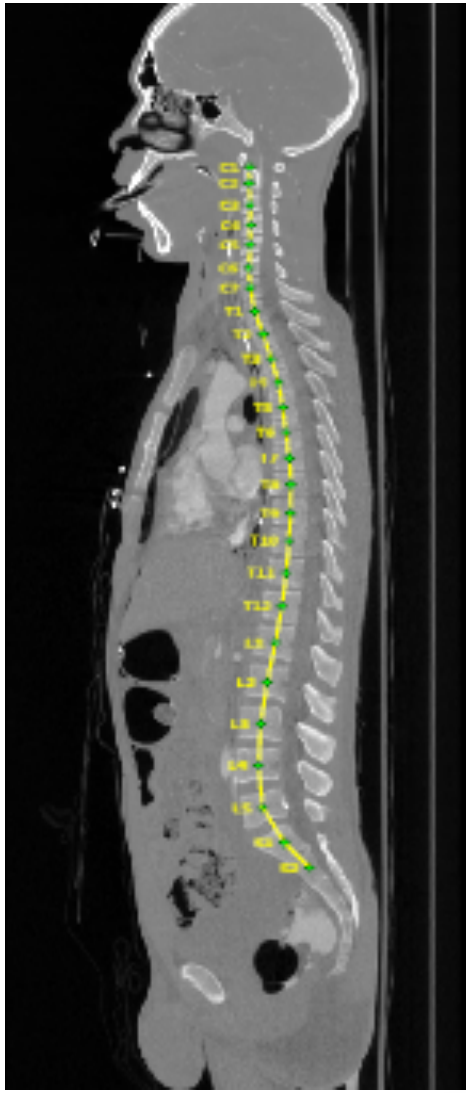




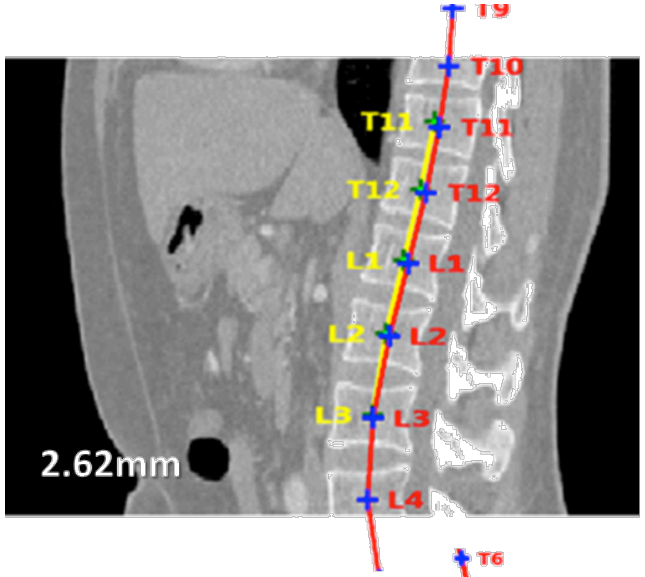


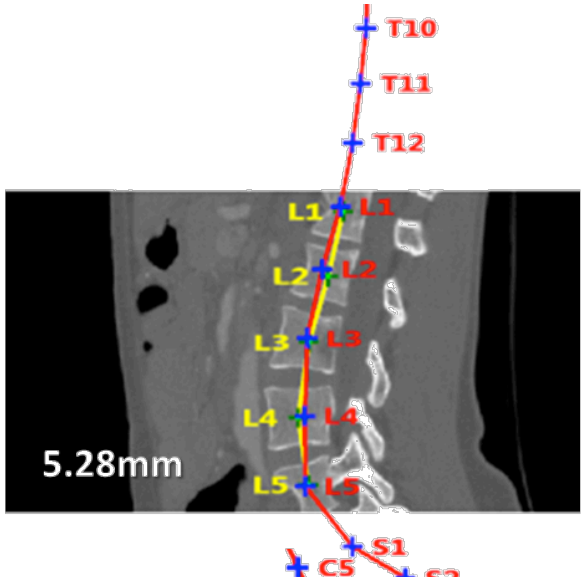


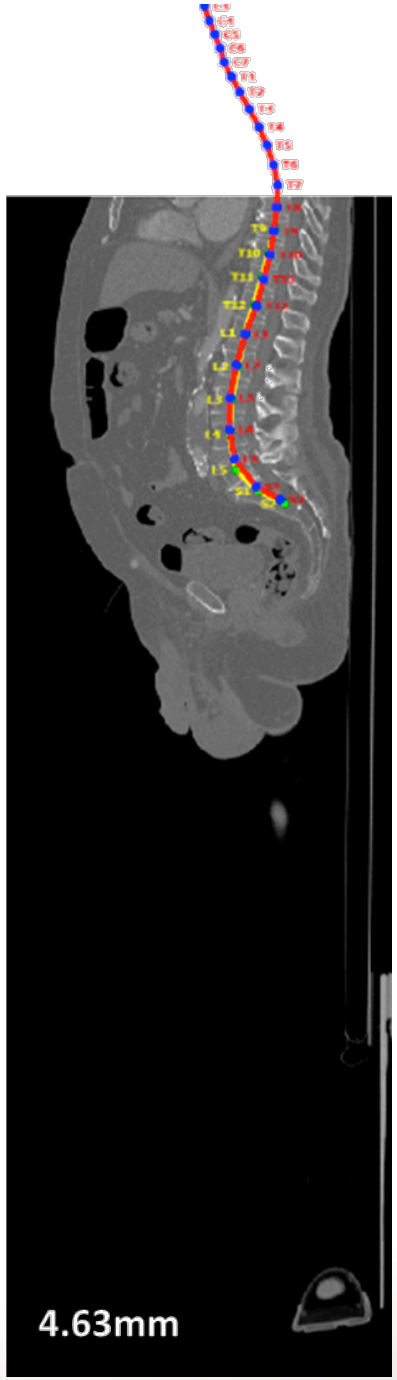


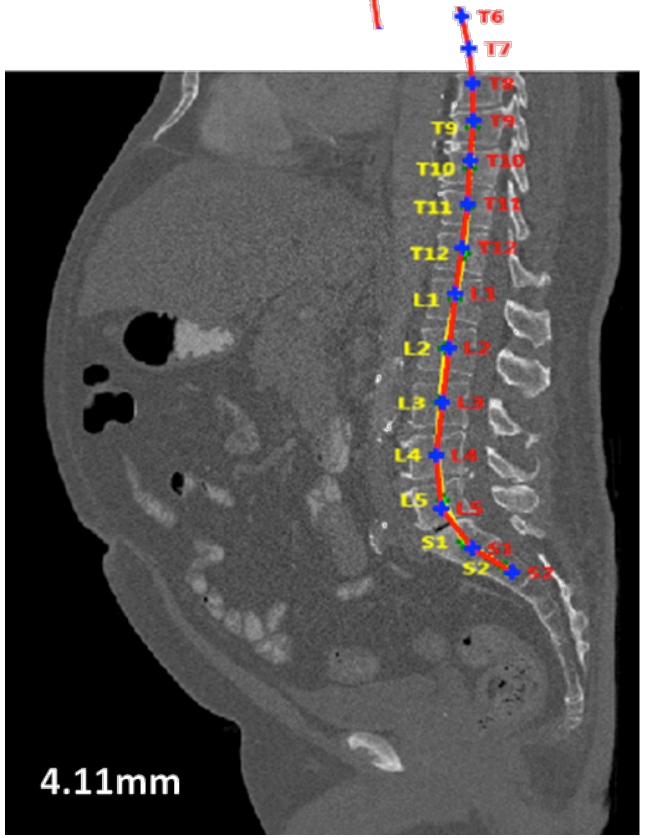


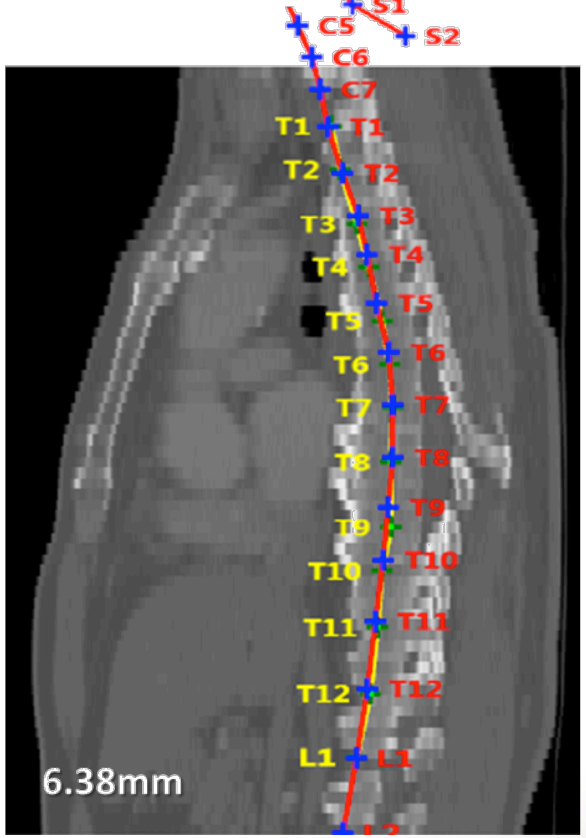
Some results



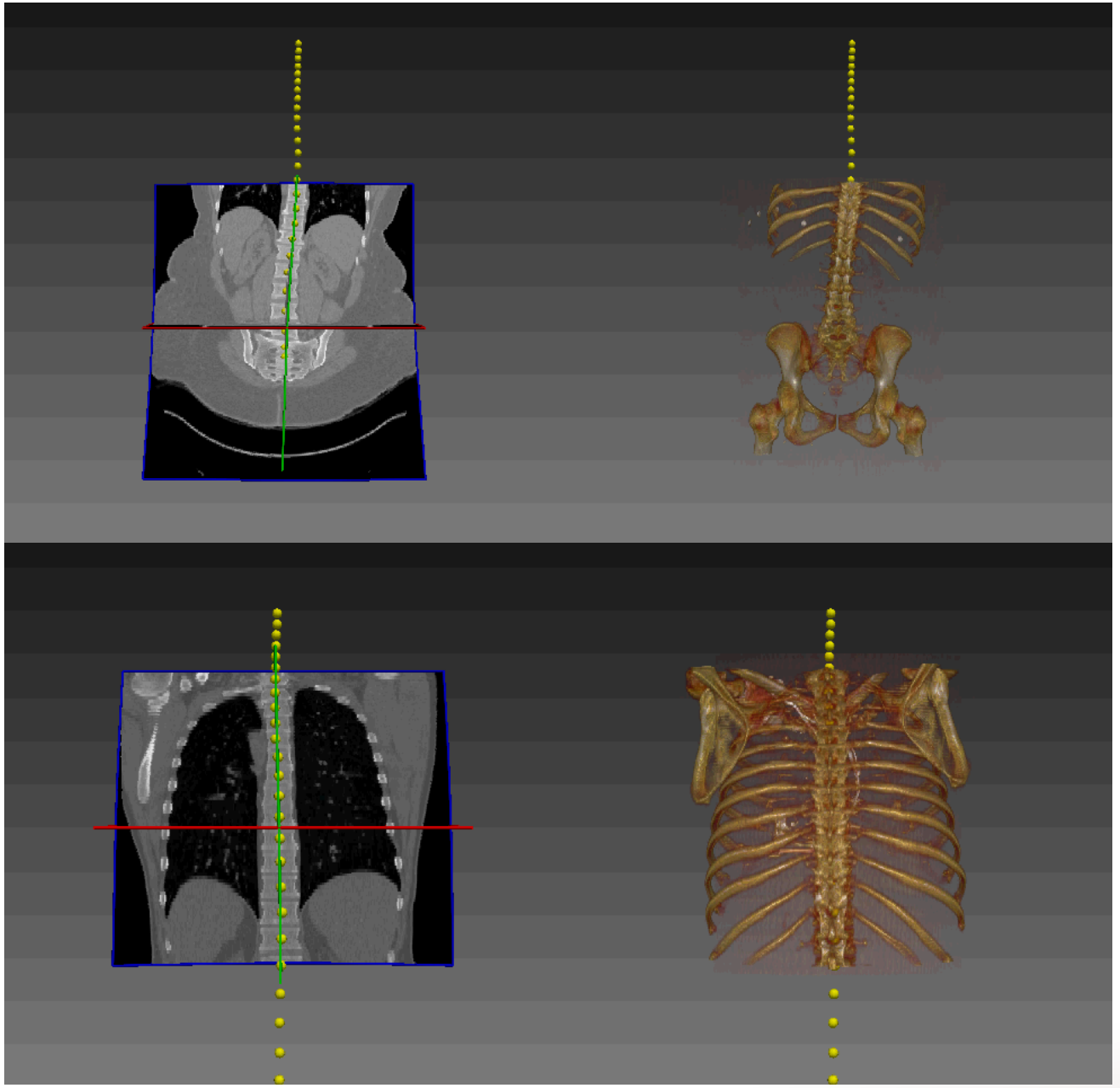








Some results

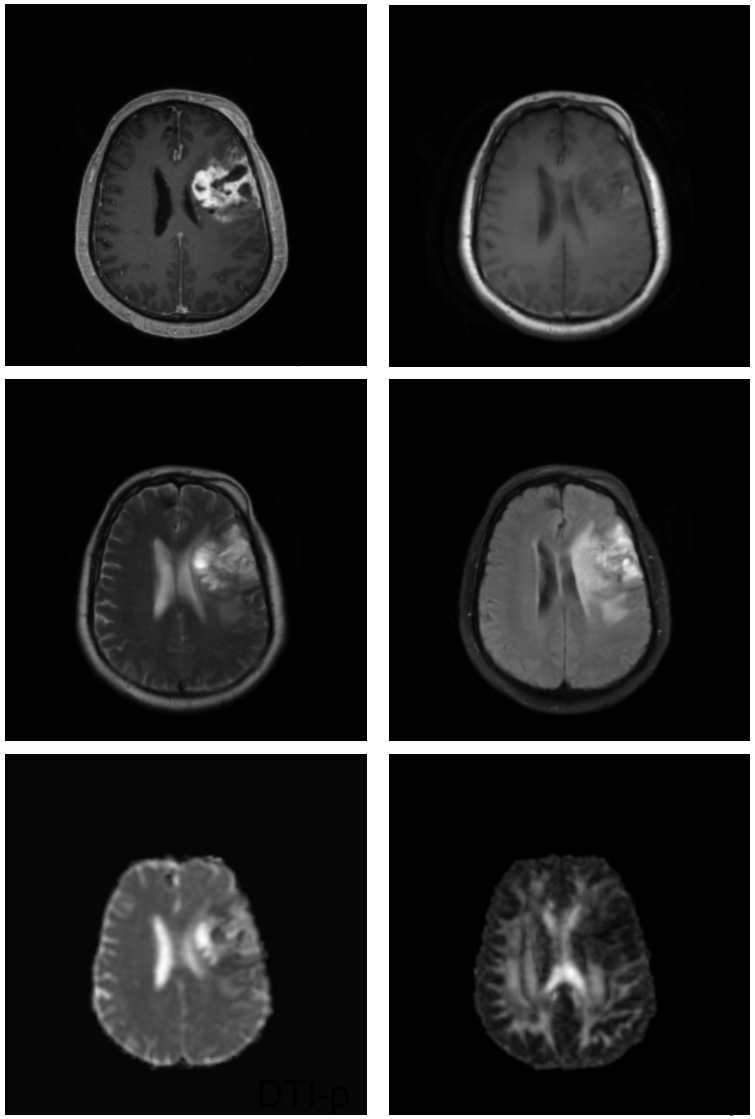


Talk overview

- A brief introduction to machine learning
- Decision forests and jungles
- **Applications in medical image analysis**
  - Anatomy localization
  - Spine detection
  - **Brain tumour segmentation**
  - Learned image super-resolution
  - Quantifying progression of multiple sclerosis

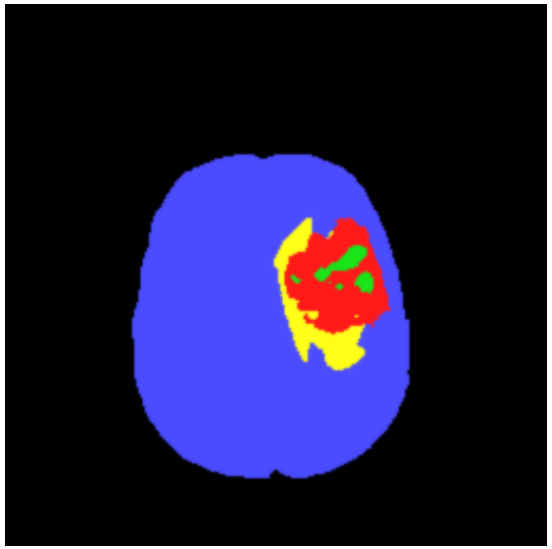
D. Zikic, B. Glocker, E. Konukoglu, A. Criminisi, J. Shotton, C. Demiralp, O. Thomas, T. Das, R. Jena, and S. Price, **Decision Forests for Tissue-specific Segmentation of High-grade Gliomas in Multi-channel MR**, in *MICCAI 2012*, Springer, October 2012

Automatic segmentation of brain tumour



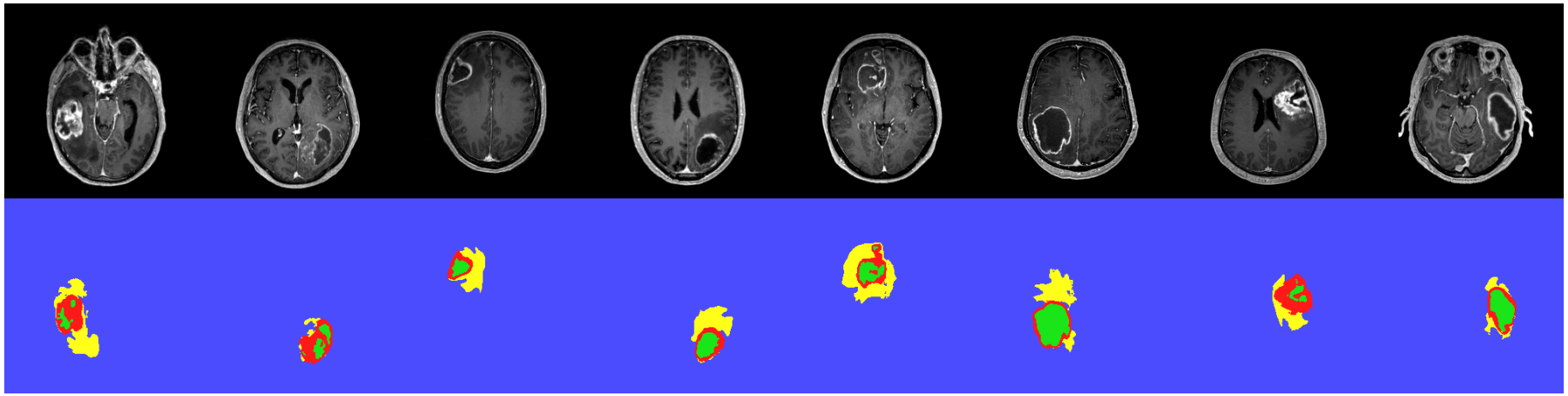
3D MRI input data

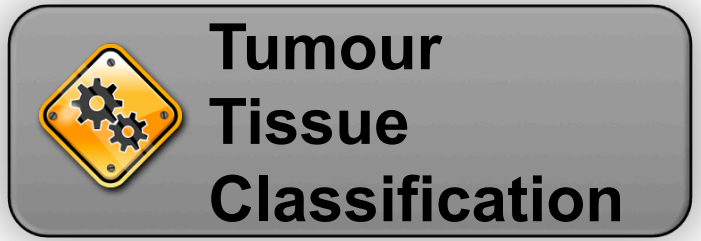
Segmentation of tumorous tissues:



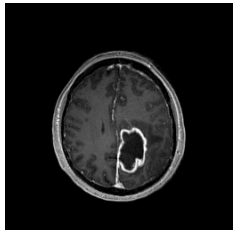
---- Active cells
---- Necrotic core
---- Edema
---- Background

Training a voxel-wise forest classifier

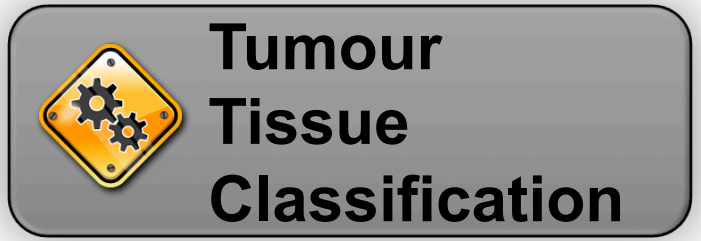




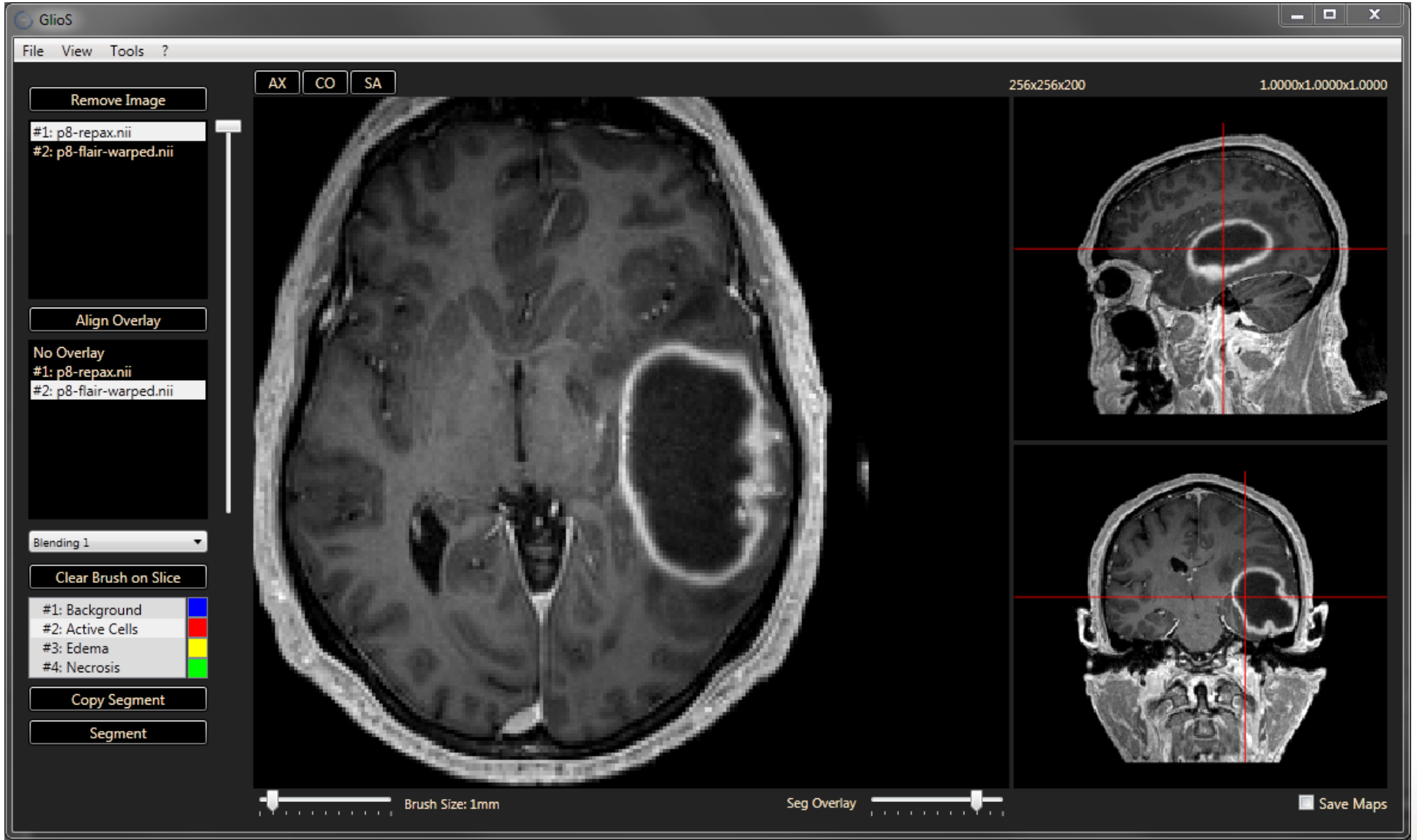
Testing the voxel-wise forest classifier



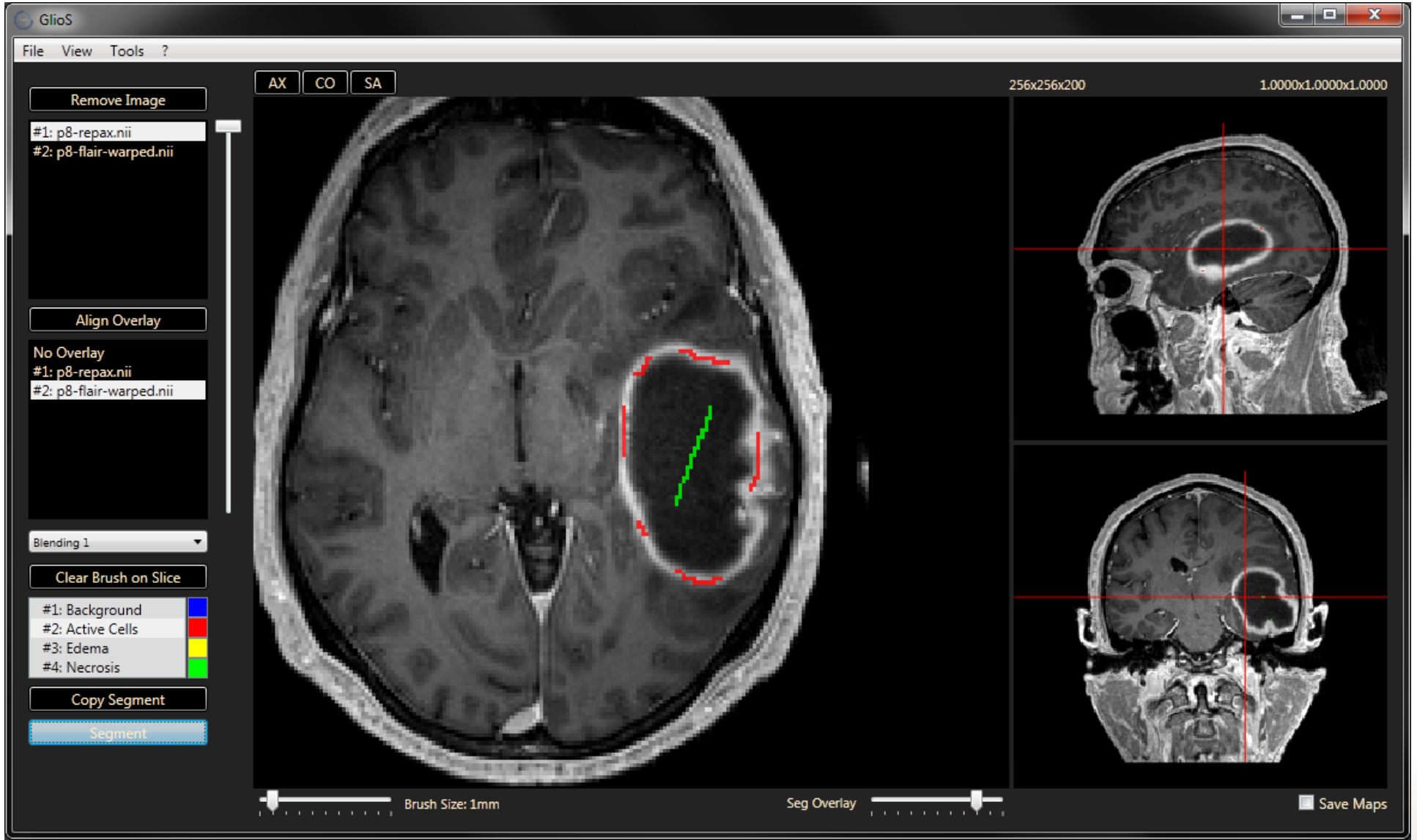
New Patient,
previously unseen



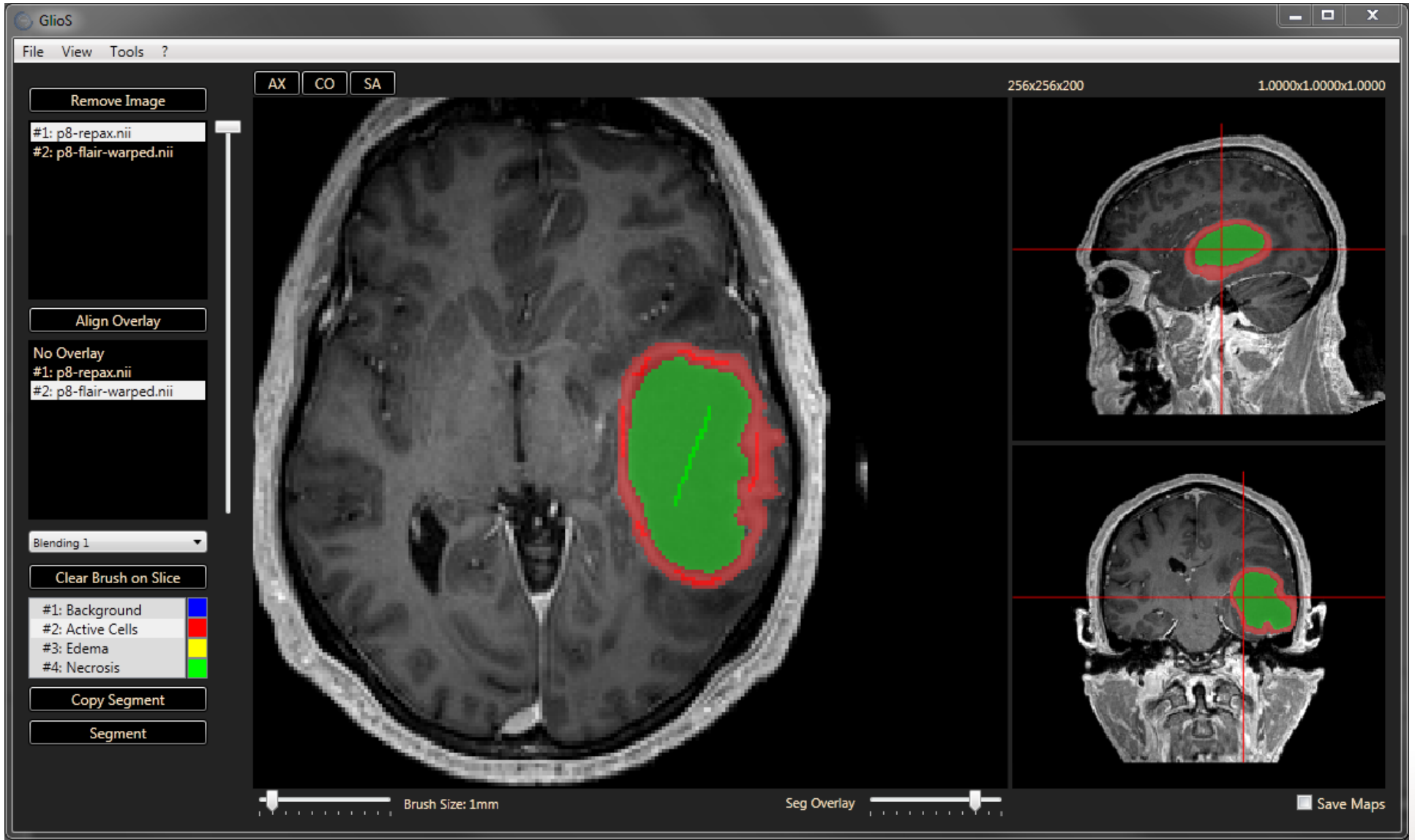
Building the training database



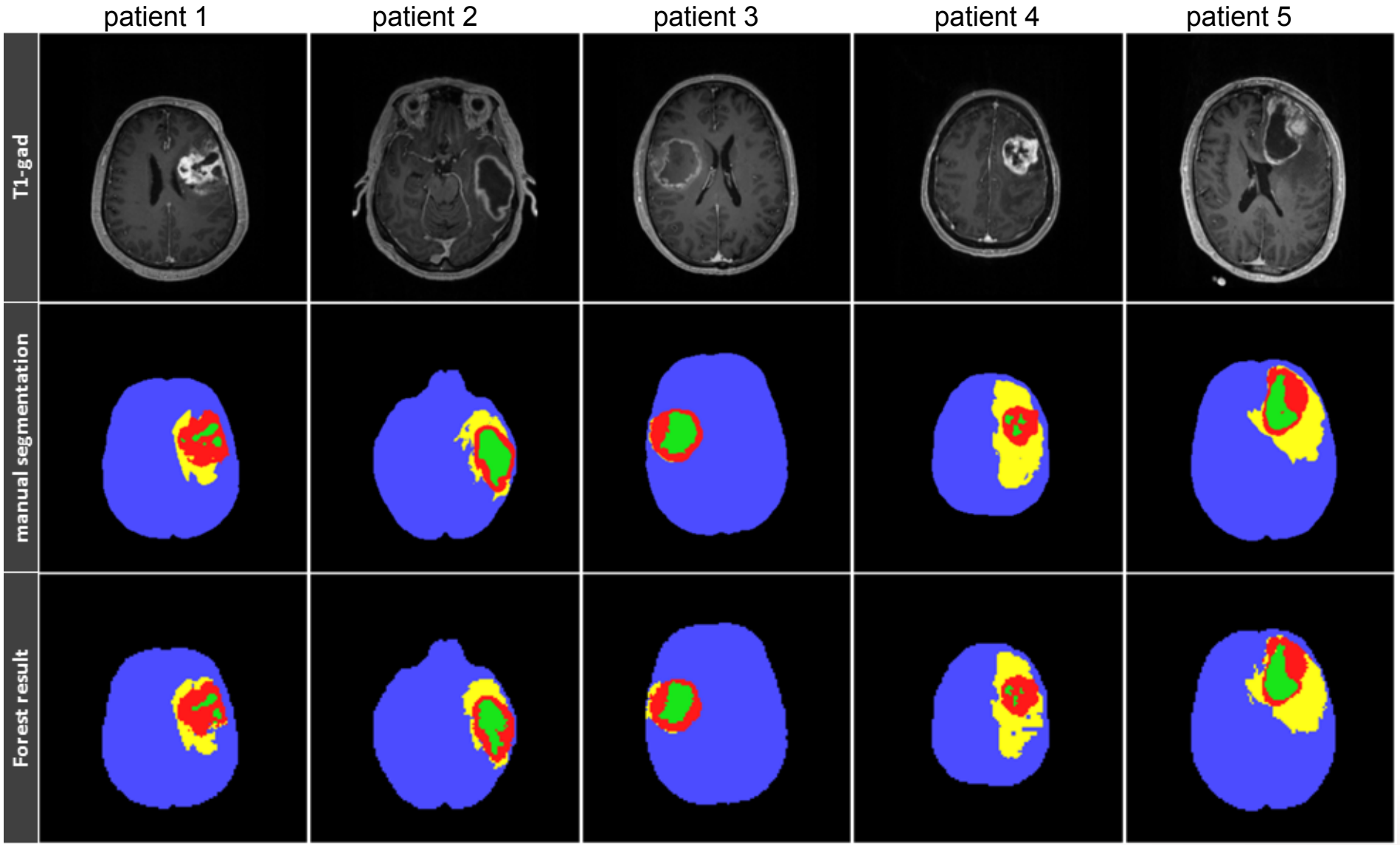
Building the training database



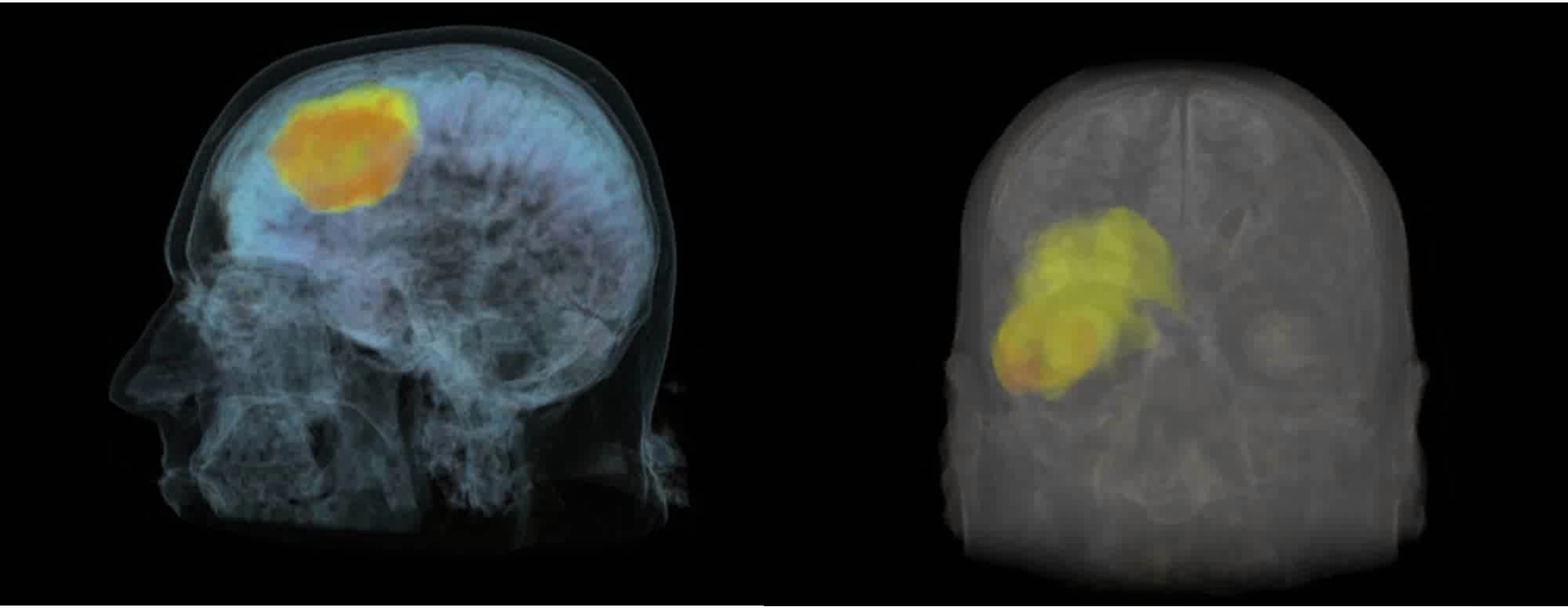
Building the training database



Glioblastoma segmentation results



Glioblastoma segmentation results

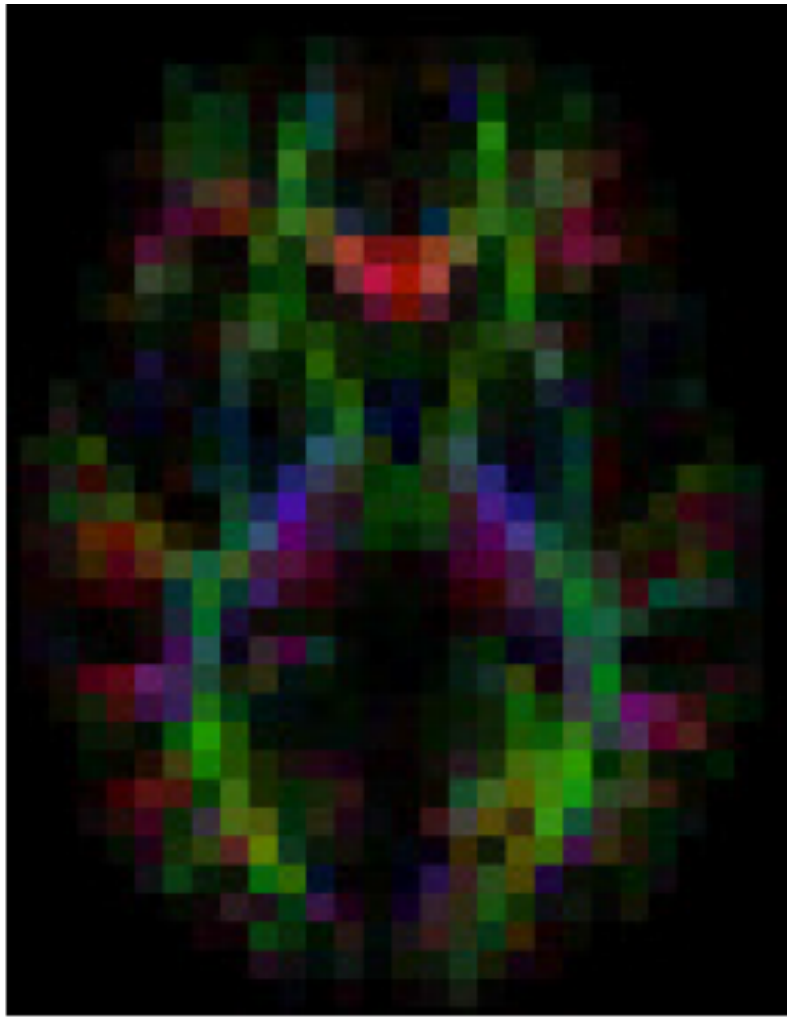


Talk overview

- A brief introduction to machine learning
- Decision forests and jungles
- **Applications in medical image analysis**
  - Anatomy localization
  - Spine detection
  - Brain tumour segmentation
  - **Learned image super-resolution**
  - Quantifying progression of multiple sclerosis

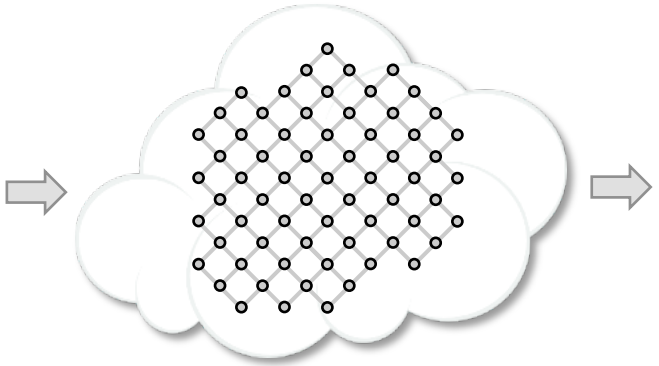
D. Alexander, D. Zikic, J. Zhang, H. Zhang, and A. Criminisi, **Image Quality Transfer via Random Forest Regression: Applications in Diffusion MRI**, in *MICCAI*, Springer, 2014

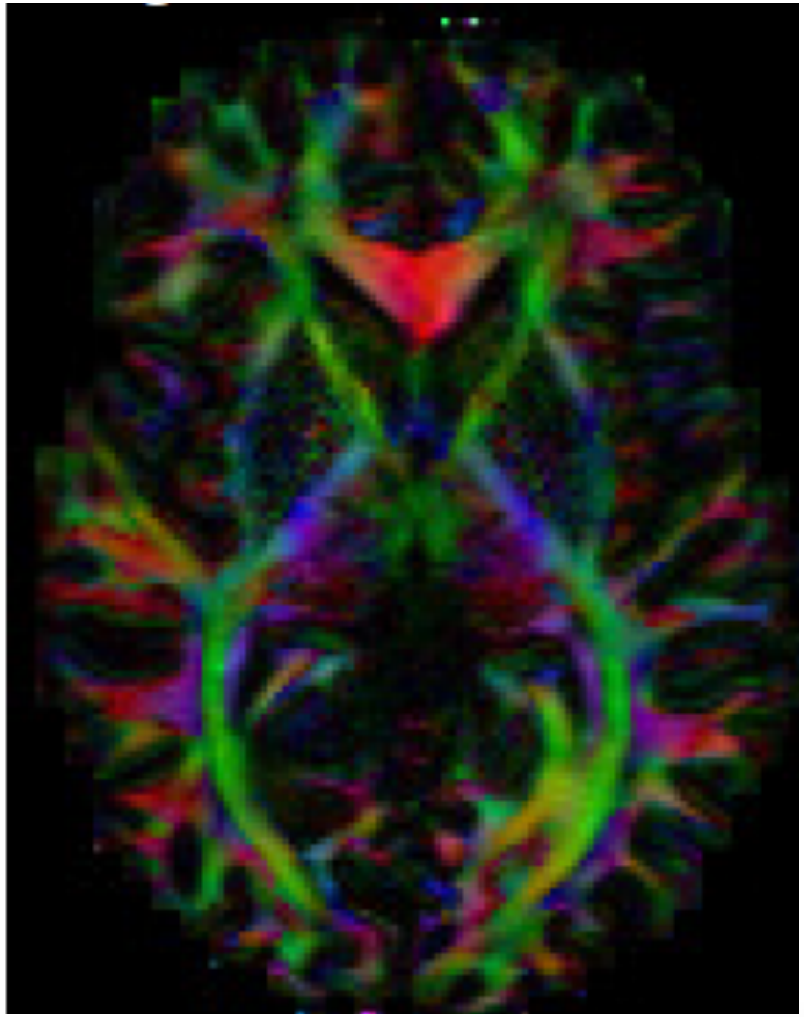
Learned image super-resolution



Low-res diffusion MRI
(faster acquisition, cheaper)

Learned voxel predictor





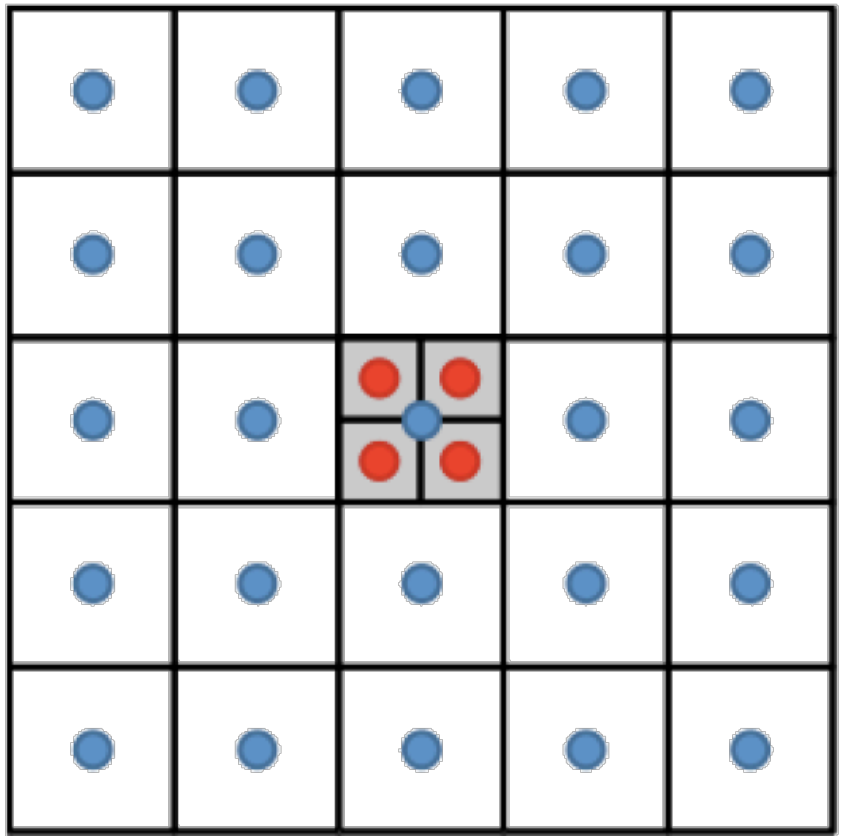
High-res diffusion MRI

Learned image super-resolution

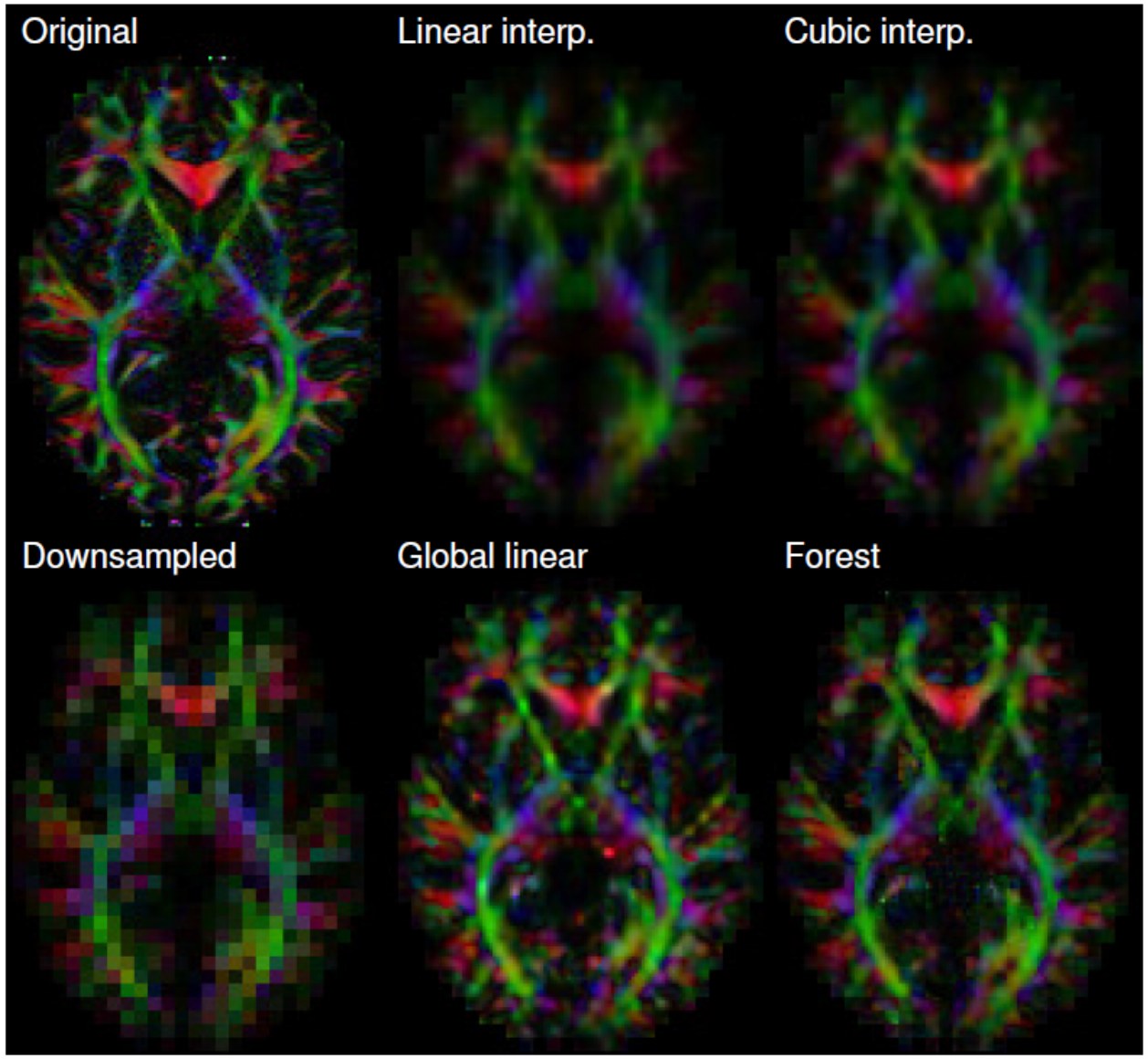
Problem statement

learning to predict the value of the **high-res voxels** from the **low-res voxels.**

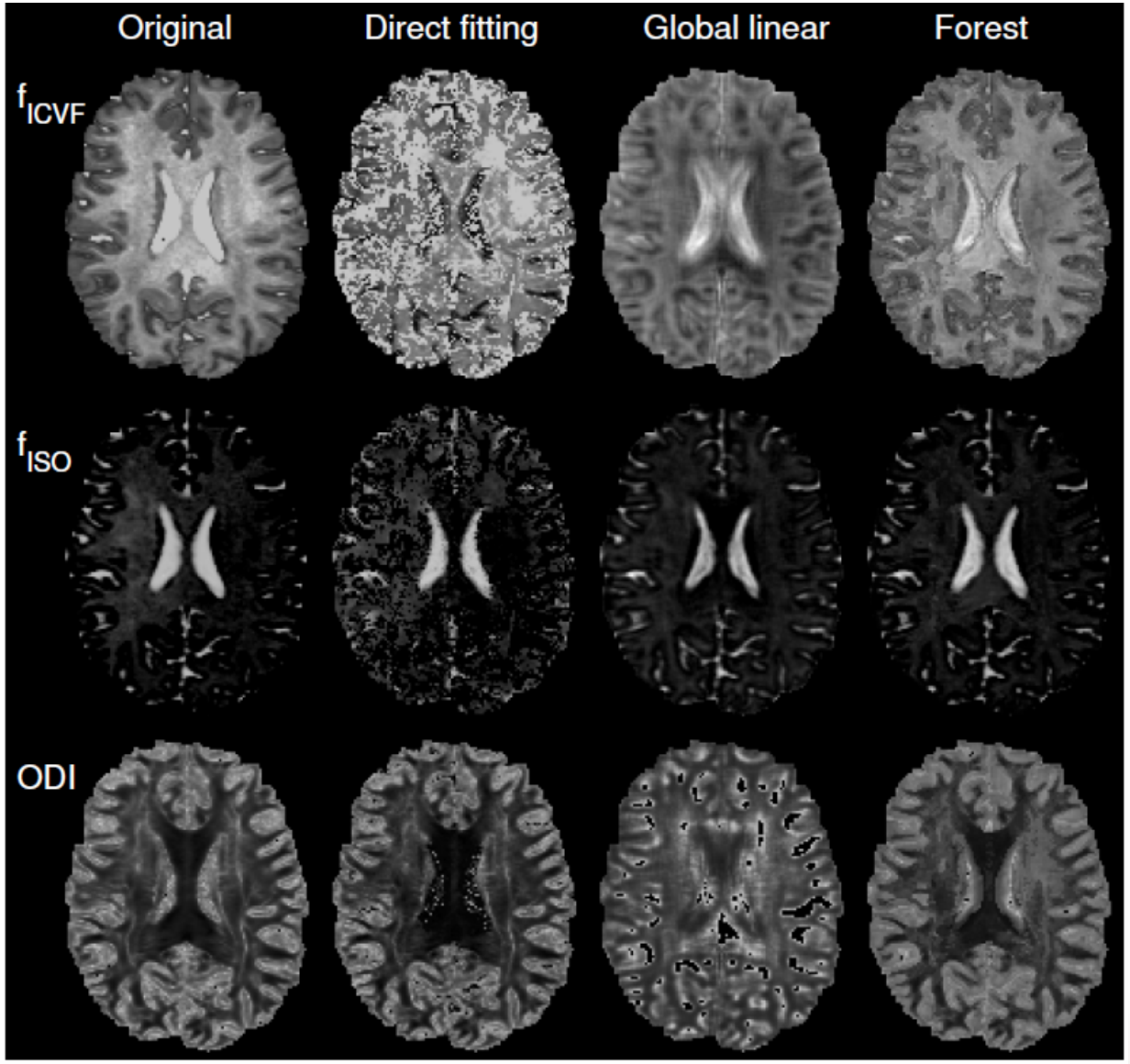
- Training data can be easily obtained
- Well defined accuracy measure



Learned image super-resolution

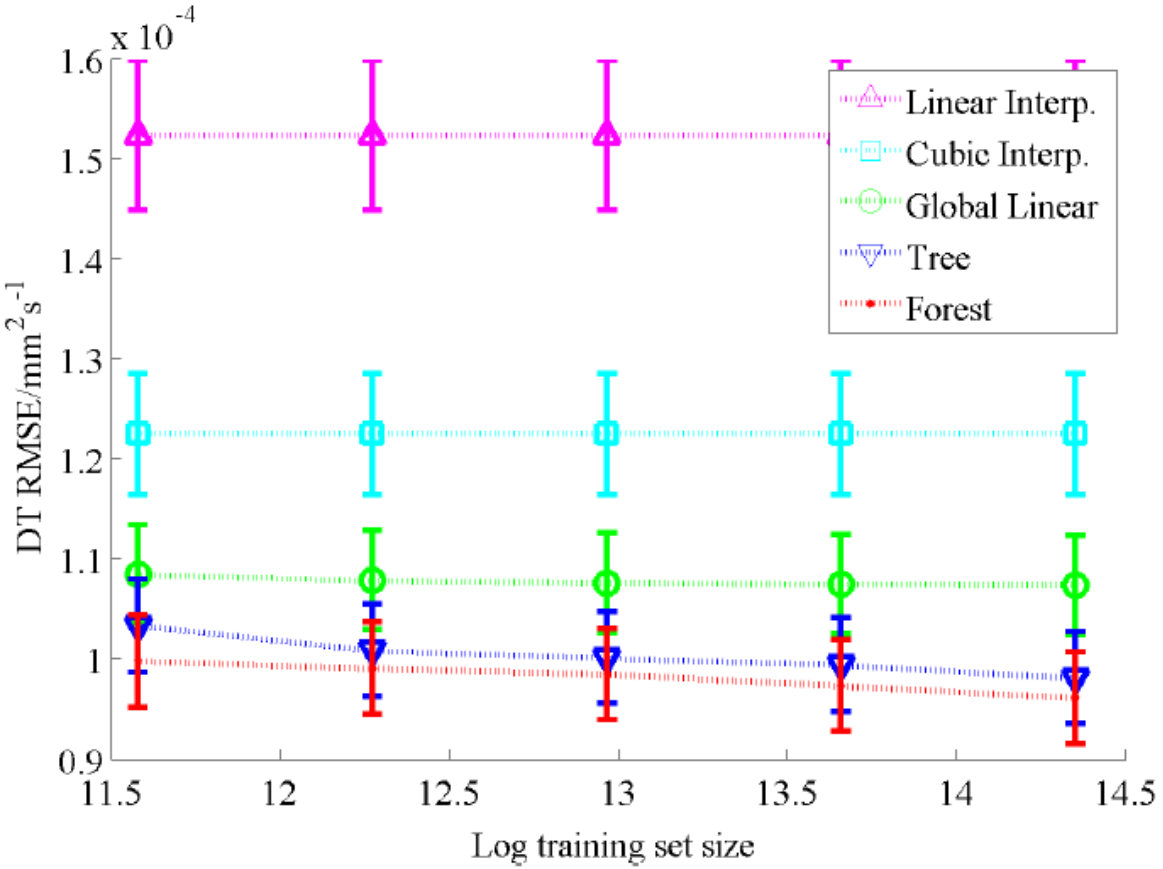


Direction-encoded colour FA maps for various reconstructed DTIs

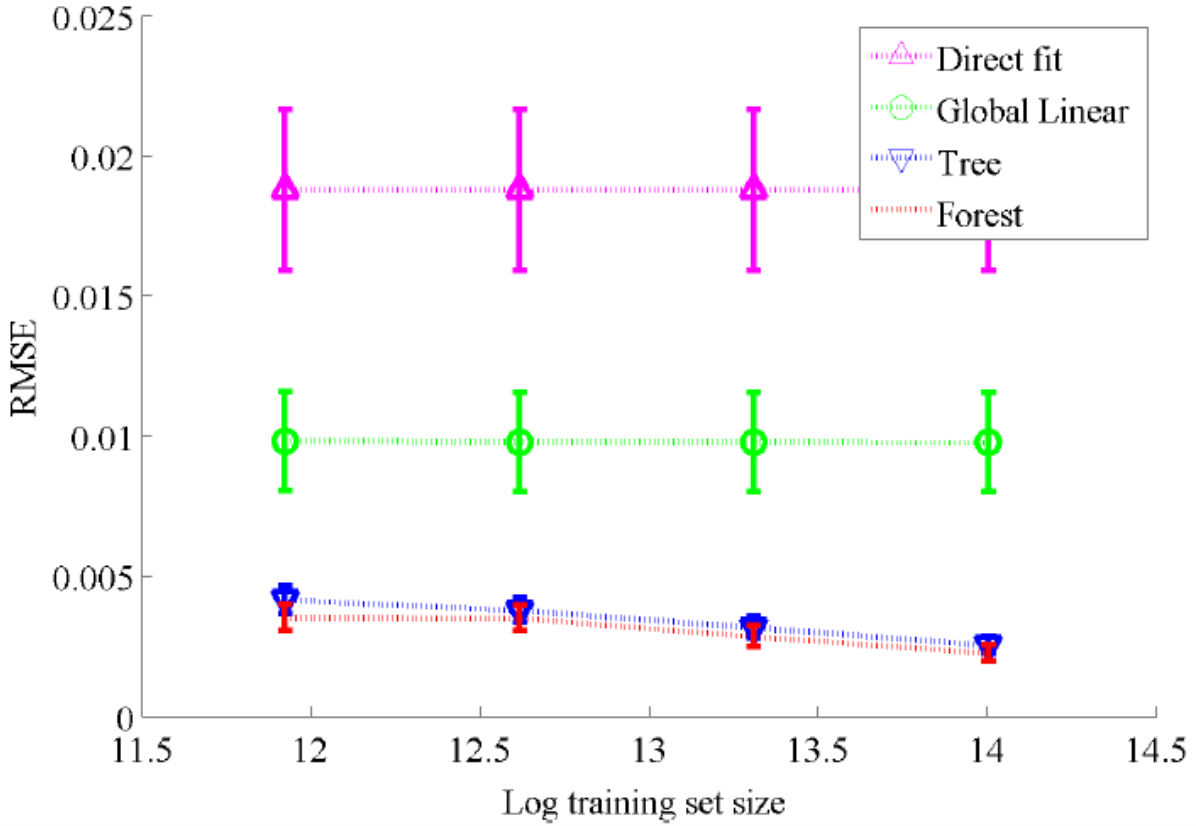


Comparison of ground truth NODDI parameter maps with various fitting techniques

Learned image super-resolution



Reconstruction errors for DT maps



Reconstruction errors for NODDI parameter maps

D. Alexander, D. Zikic, J. Zhang, H. Zhang, and A. Criminisi, **Image Quality Transfer via Random Forest Regression: Applications in Diffusion MRI**, in *MICCAI*, Springer, 2014

Talk overview

- A brief introduction to machine learning
- Decision forests and jungles
- **Applications in medical image analysis**
  - Anatomy localization
  - Spine detection
  - Brain tumour segmentation
  - Learned image super-resolution
  - **Quantifying progression of multiple sclerosis**

Prize for largest number of authors ever?

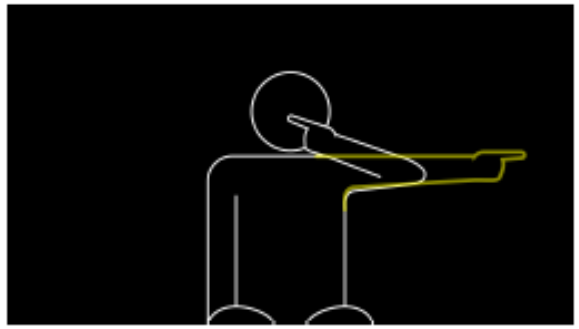
P. Kontschieder, B. Glocker, J.F. Dorn, C. Morrison, R. Corish, D. Zikic, A. Sellen, M. DSouza, C. P. Kamm, J. Burggraaff, P. Tewarie, T. Vogel, M. Azzarito, P. Chin, F. Dahlke, C. Polman, L. Kappos, B. Uitdehaag, and A. Criminisi, **Quantifying Progression of Multiple Sclerosis via Classification of Depth Videos**, in *MICCAI 2014, Boston*, Springer, 2014

Quantifying progression of Multiple Sclerosis

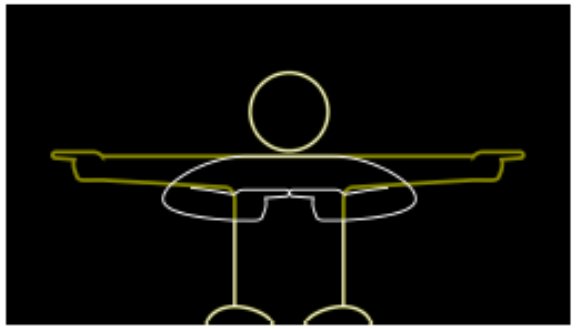
- Using Kinect to assess MS patients
- No use of MR imaging
- In the hospital or at home
- Through games and physical exercises
- Cannot use the available Kinect SDK
  (designed for gamers rather than patients)

Quantifying progression of Multiple Sclerosis

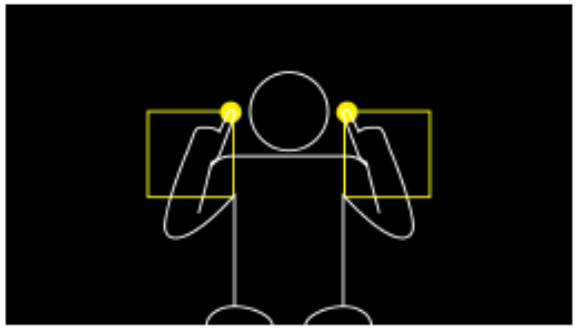
- 9 movements/tests designed by the radiologists in our team to tease out various disfunctions.



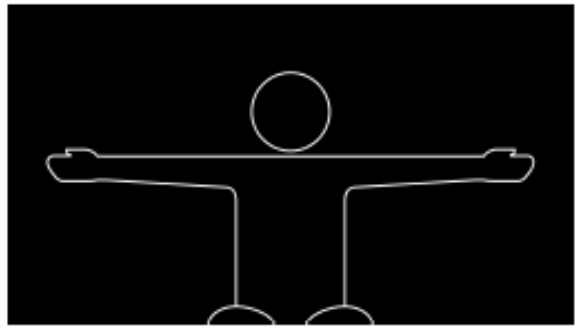
Finger to nose test



Finger to finger test



Drawing square test



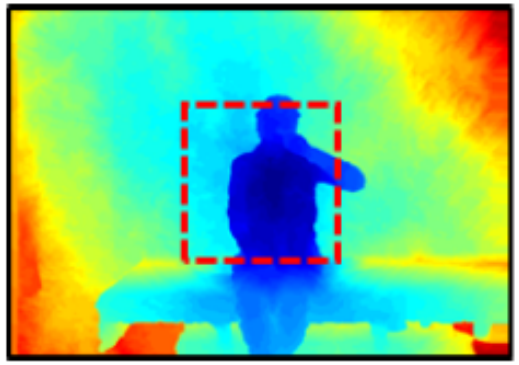
Truncal ataxia test

- Hundreds of depth videos from dozens of patients in 5 different European medical centres

a

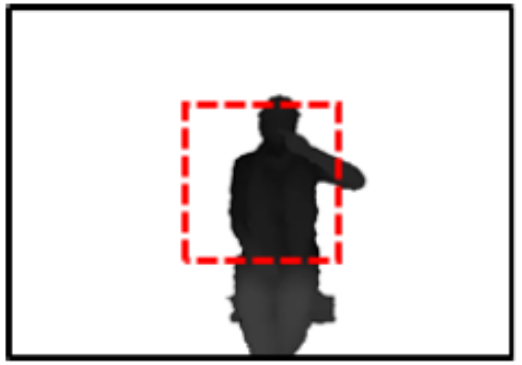
Input depth video

b



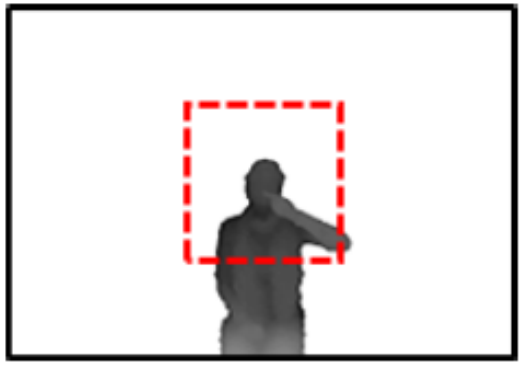
Geodesic segmentation in depth space

c



Extracted foreground

d



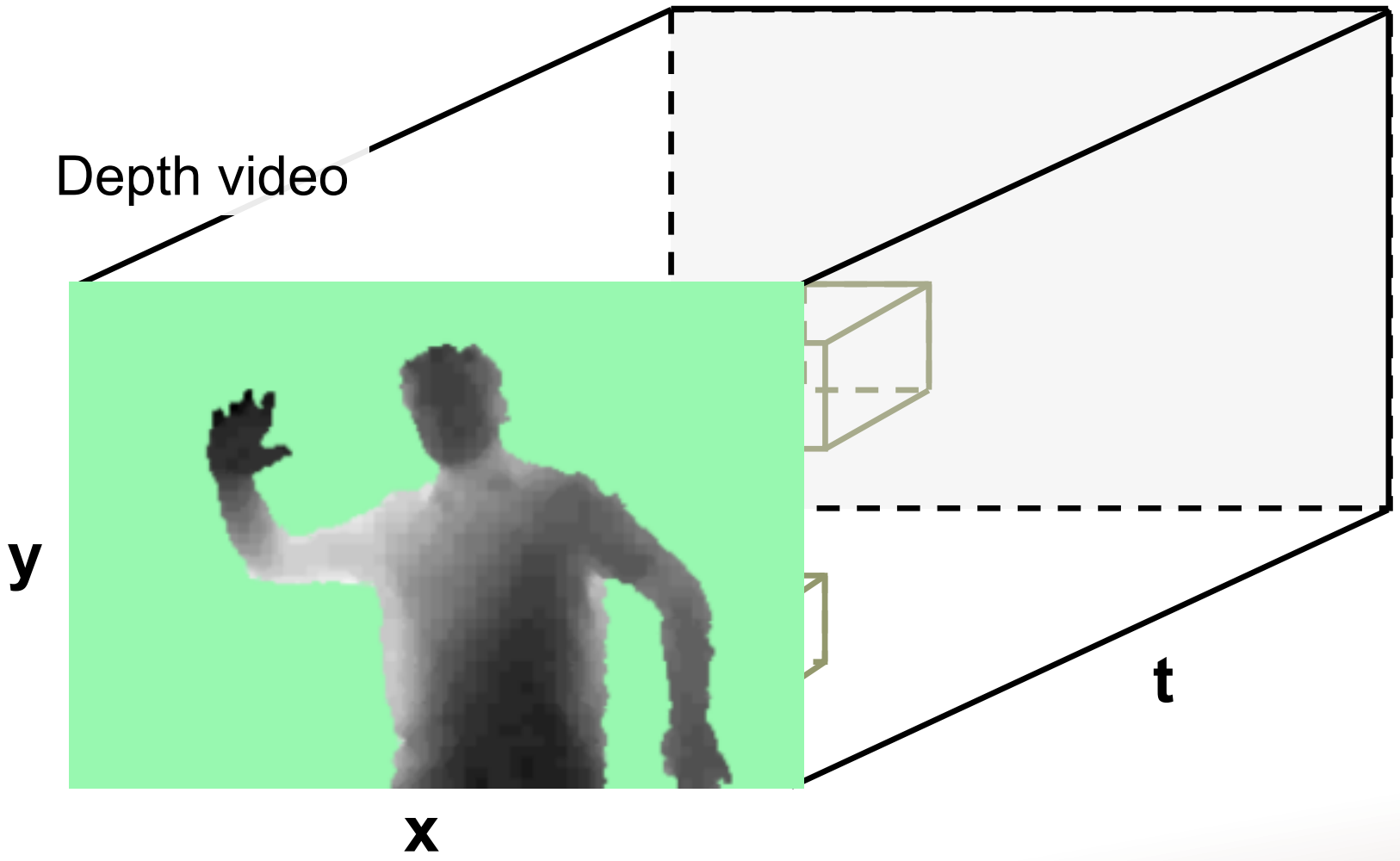
Normalization in x,y and depth

Quantifying progression of Multiple Sclerosis

Spatio-temporal features

Channels:
- Optical flow
- Depth values
- Gradients
- Colour

$$f(\mathbf{v}, \theta_j) = \frac{1}{|B_j|} \sum_{\mathbf{p} \in B_j} I(\mathbf{p})$$



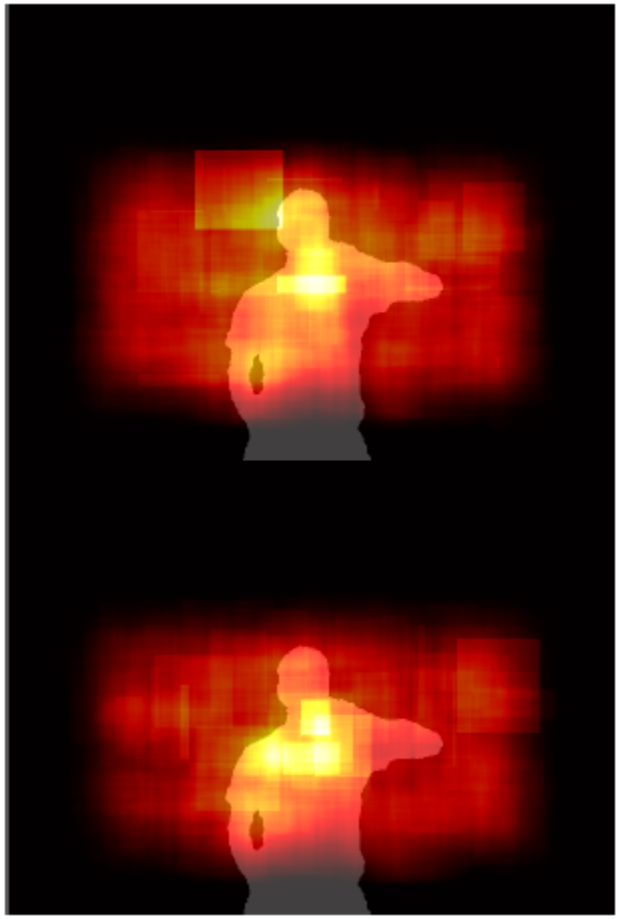
Quantifying progression of Multiple Sclerosis

Method	FNT			FFT			DRS			TAT		
	$\overline{D}$	D_{PAT}	D_{HS}	$\overline{D}$	D_{PAT}	D_{HS}	$\overline{D}$	D_{PAT}	D_{HS}	$\overline{D}$	D_{PAT}	D_{HS}
Forest	84.3 ±4.4	79.4 ±5.9	89.2 ±2.9	74.9 ±6.3	58.1 ±11.0	91.7 ±1.7	81.1 ±3.8	70.3 ±5.1	92.1 ±2.6	74.3 ±5.7	57.8 ±8.8	90.9 ±2.6
Linear SVM	80.9 ±3.0	75.3 ±5.0	86.4 ±1.1	79.1 ±4.4	65.6 ±7.6	92.7 ±1.2	84.5 ±2.8	75.1 ±4.0	93.9 ±1.5	73.4 ±3.7	56.5 ±5.1	90.3 ±2.3
Kernel SVM	85.2 ±4.0	80.5 ±5.7	89.9 ±2.3	81.0 ±3.6	68.3 ±6.6	93.7 ±0.6	81.4 ±2.6	70.6 ±3.4	92.2 ±1.8	66.2 ±4.0	45.8 ±7.2	86.7 ±0.8
S_{SENS} S_{SPEC} S_{GLOB}	78.3 ±10.1	91.4 ±2.2	86.7 ±3.2	79.3 ±17.4	91.1 ±2.9	89.5 ±0.9	89.8 ±7.1	90.3 ±3.0	90.2 ±2.3	74.3 ±11.9	86.8 ±3.5	85.1 ±3.9
Avg. #Train/fold PAT, HS	103.2	186.4		52.8	96.6		48.0	87.0		49.6	89.4	
Avg. #Test/fold PAT, HS	25.8	46.4		13.2	76.6		12.0	61.6		12.4	78.4	
Segmentation S_{SENS} S_{SPEC}	99.9	97.9		98.2	99.9		99.9	99.8		99.8	99.8	

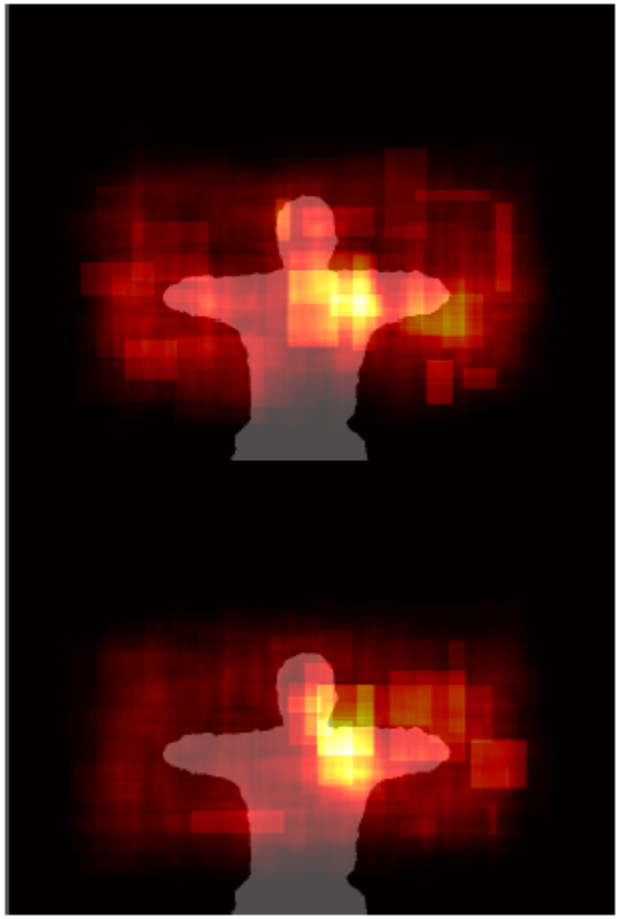
Table 1. Quantitative results for all experiments, all D, S in %. See text.

Quantifying progression of Multiple Sclerosis

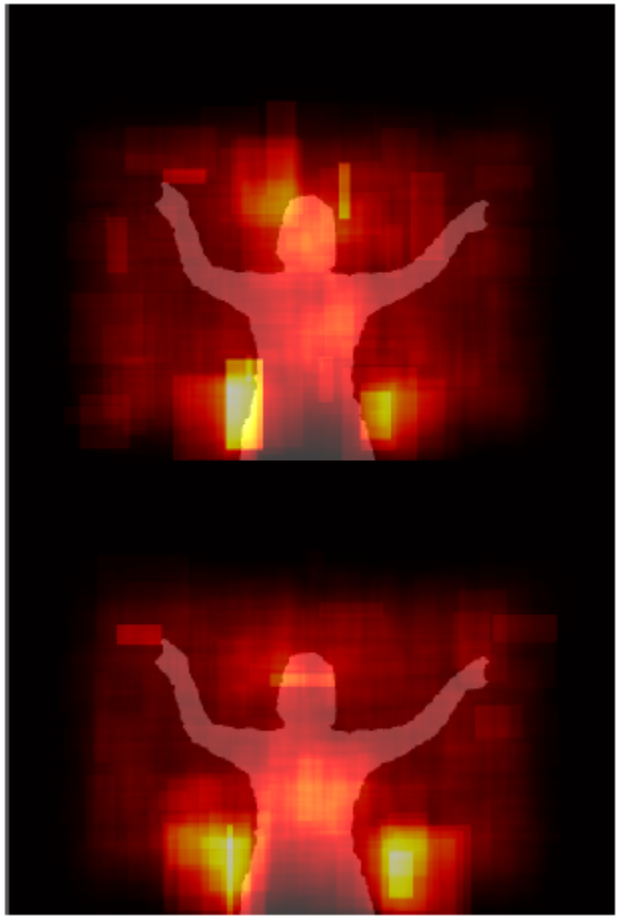
Interpretability of results (forests). Automatic discovery of **discriminative spatio-temporal landmarks**.



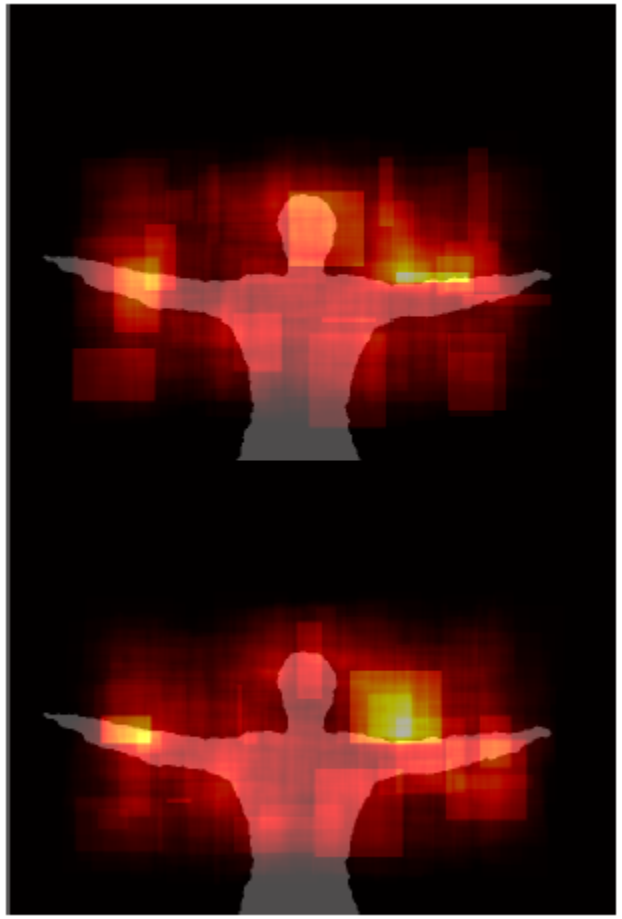
Finger to nose test



Finger to finger test



Drawing square test



Truncal ataxia test

Modern, efficient machine learning has the potential to revolutionize medicine!

Microsoft Research Bright Minds Competition

research.microsoft.com/undergrad