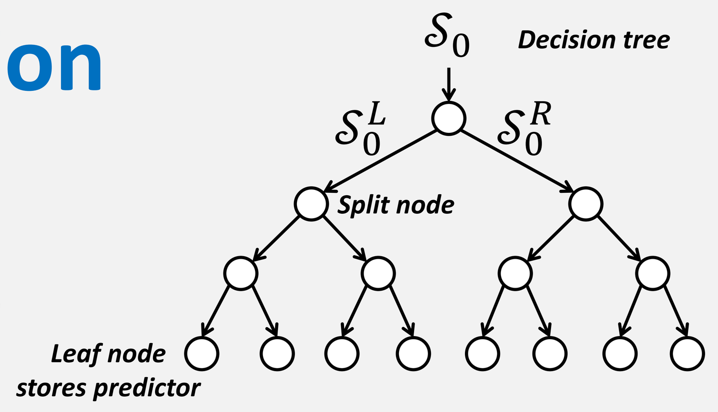
Joint Classification-Regression Forests for Spatially Structured Multi-Object Segmentation

Ben Glocker[1], Olivier Pauly[2,3], Ender Konukoglu[1], Antonio Criminisi[1]

[1]Microsoft Research Cambridge, UK [2]Helmholtz Zentrum München, Germany [3]Technische Universität München, Germany

Microsoft Research

Joint Classification-Regression



Learning Objective

$p(\boldsymbol{y}|\boldsymbol{x})$

$\boldsymbol{x} \in \mathbb{R}^m$: feature space

$\boldsymbol{y} \in \mathcal{O}$: prediction space

Categorical-Continuous Prediction

$$\boldsymbol{y} = (\boldsymbol{c}, \boldsymbol{r})$$

$\boldsymbol{c} \in \mathcal{C}$: classification space

$\boldsymbol{r} \in \mathbb{R}^n$: regression space

$$\mathcal{O} = \mathcal{C} \times \mathbb{R}^n$$

Tree Training: Information Gain

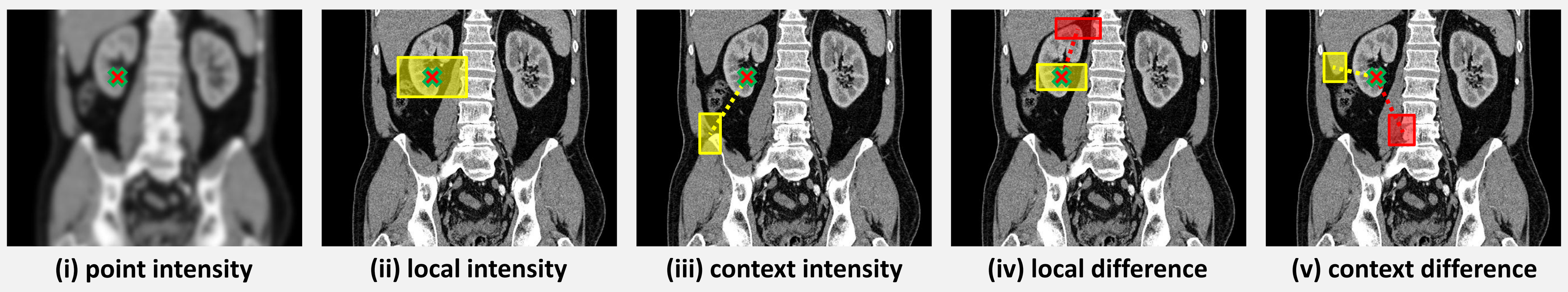
$$\mathcal{S} = \{(\boldsymbol{x}_k, \boldsymbol{y}_k)\}_{k=1}^K \qquad I(\mathcal{S}_i, \mathcal{S}_i^L, \mathcal{S}_i^R) = H(\mathcal{S}_i) - \sum_{j\in\{L,R\}} \frac{|\mathcal{S}_i^j|}{|\mathcal{S}_i|} H(\mathcal{S}_i^j)$$

Joint Entropy

$$H(\mathcal{S}) = -\sum_{c\in\mathcal{C}} \int_{r\in\mathbb{R}^n} p(\boldsymbol{c}, \boldsymbol{r}|\boldsymbol{x}) \log p(\boldsymbol{c}, \boldsymbol{r}|\boldsymbol{x})\, d\boldsymbol{r}$$

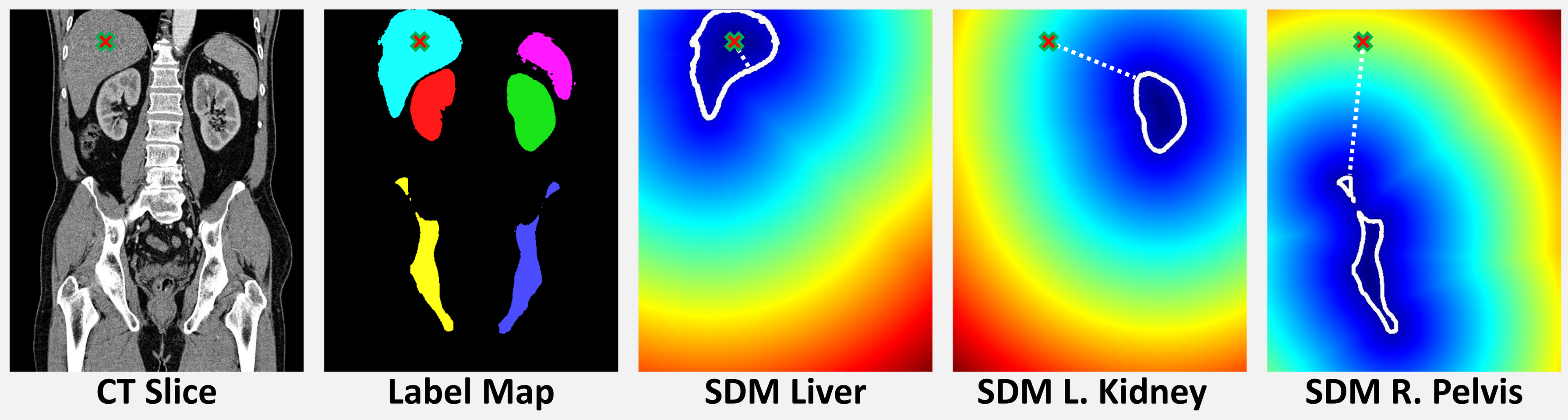
$$= \underbrace{-\sum_{c\in\mathcal{C}} p(\boldsymbol{c}|\boldsymbol{x}) \log p(\boldsymbol{c}|\boldsymbol{x})}_{\text{Shannon Entropy: } H_c} + \underbrace{\sum_{c\in\mathcal{C}} p(\boldsymbol{c}|\boldsymbol{x}) \left(-\int_{r\in\mathbb{R}^n} p(\boldsymbol{r}|\boldsymbol{c}, \boldsymbol{x}) \log p(\boldsymbol{r}|\boldsymbol{c}, \boldsymbol{x}) d\boldsymbol{r}\right)}_{\text{Weighted Differential Entropy: } H_{r|c}}$$

Features: Capture local and contextual appearance



Motivation: Class + Spatial Consistency

Exploit rich nature of label maps



Spatial Consistency via Distance Regression

Predictor

$$p(\boldsymbol{r}|\boldsymbol{c}, \boldsymbol{x}) \triangleq \mathcal{N}(\boldsymbol{r}; \mu_{r|c}, \Sigma_{r|c}|\boldsymbol{c}, \boldsymbol{x})$$

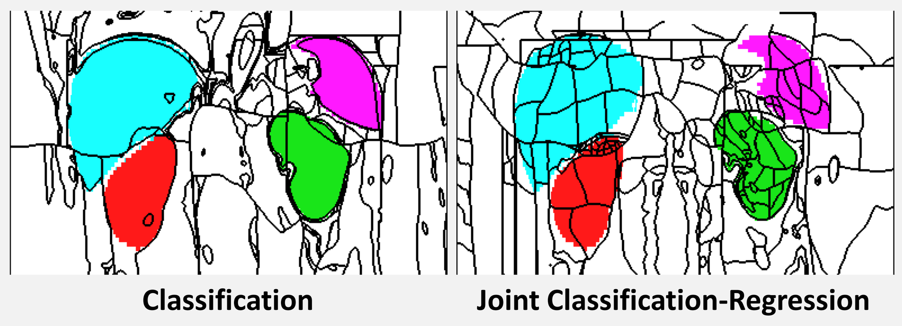
Entropy

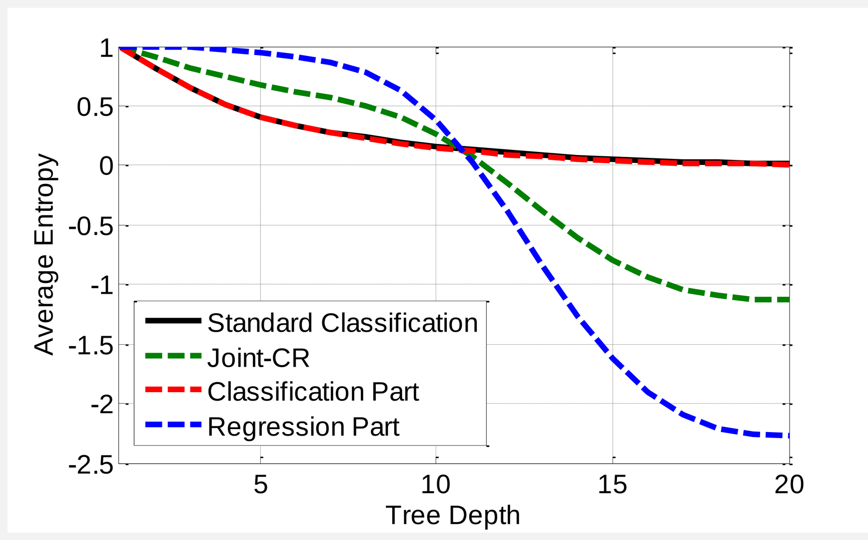
$$H_{r|c} = \sum_{c\in\mathcal{C}} p(\boldsymbol{c}|\boldsymbol{x}) \left(\frac{1}{2}\log[(2\pi e)^n |\Sigma_{r|c}|]\right)$$

Entropy Normalization

$$H(\mathcal{S}) = \frac{1}{2}\left(\frac{H_c(\mathcal{S})}{H_c(\mathcal{S}_0)} + \frac{H_{r|c}(\mathcal{S})}{H_{r|c}(\mathcal{S}_0)}\right)$$

Spatial Regularization





Robust Parameter Estimation

Parent nodes provide prior on Gaussian parameters

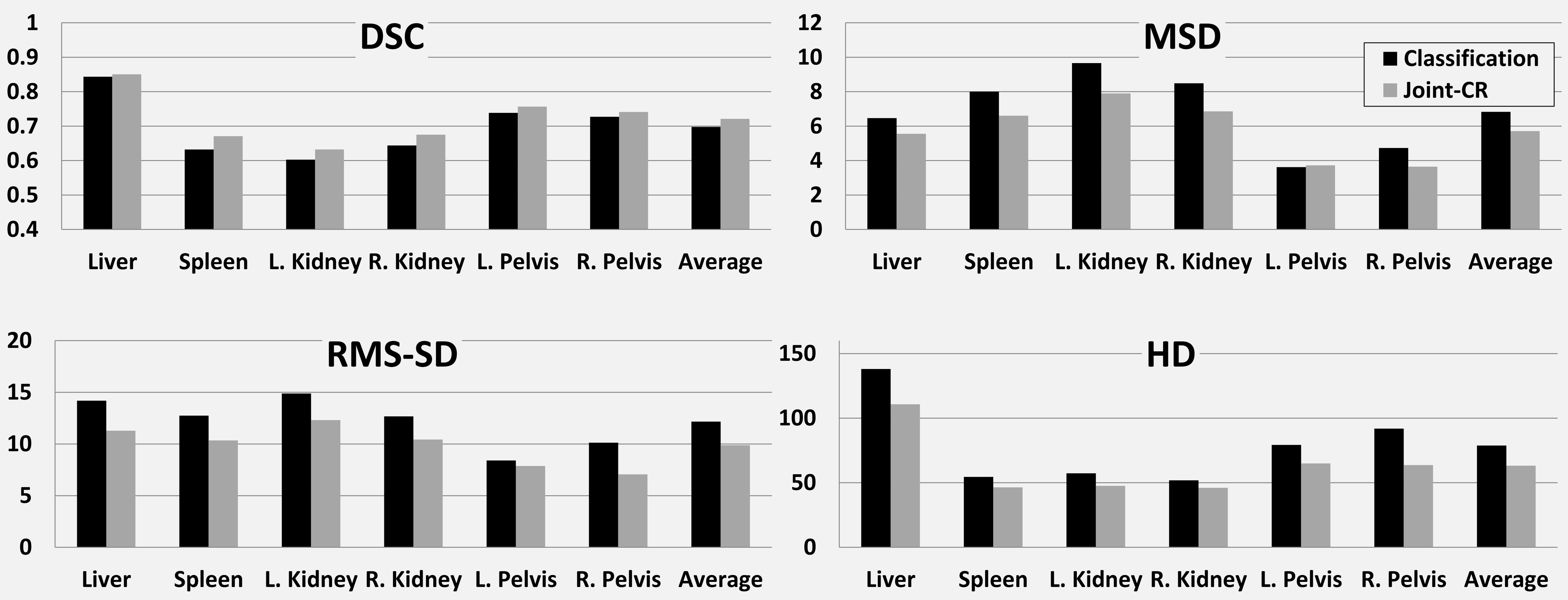
$$\mu_{r|c}^{\text{child}} = \frac{|\mathcal{S}_{r|c}^{\text{child}}|}{\kappa + |\mathcal{S}_{r|c}^{\text{child}}|} \bar{\mu}_{r|c}^{\text{child}} + \frac{\kappa}{\kappa + |\mathcal{S}_{r|c}^{\text{child}}|} \mu_{r|c}^{\text{parent}} \qquad \Sigma_{r|c}^{\text{child}} = \frac{|\mathcal{S}_{r|c}^{\text{child}}|}{Z} \bar{\Sigma}_{r|c}^{\text{child}} + \frac{\nu + n - 1}{Z} \Sigma_{r|c}^{\text{parent}} + \text{mean correction term}$$

Experimental Evaluation

Setup

- **80 3D-CT scans, 6 major organs**
- 2-fold cross-validation (40/40 train/test split)
- 50 trees, depth 20: trained on 10% of image points
- 100 features per node from a pool of 1000 features
- Greedy optimization over 10 uniform thresholds

Quantitative Results



Forest Predictions

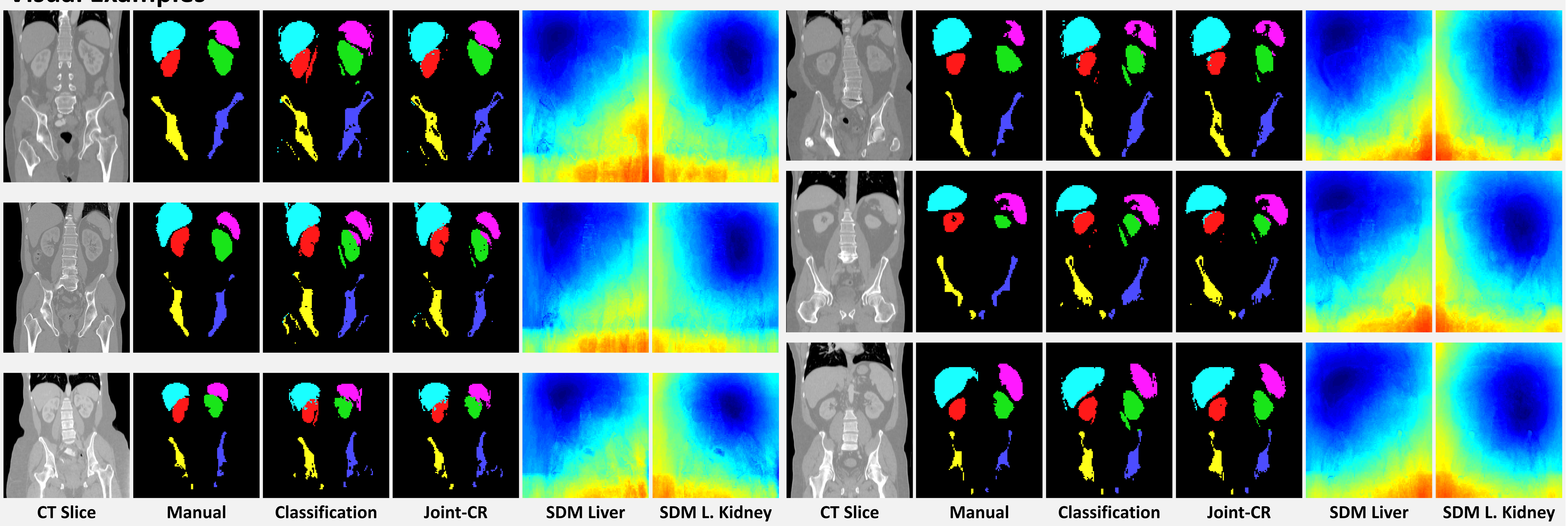
Classification MAP estimate

$$\hat{\boldsymbol{c}} = \arg\max_{c\in\mathcal{C}} p(\boldsymbol{c}|\boldsymbol{x})$$

Regression mixture mean estimate

$$\tilde{\boldsymbol{r}} = \sum_{c\in\mathcal{C}} p(\boldsymbol{c}|\boldsymbol{x}) \mu_{r|c}$$

Visual Examples



Joint classification-regression yields class and spatial consistency