Automatic Localization and Identification of Vertebrae in Arbitrary Field-of-View CT Scans

Ben Glocker[1], Johannes Feulner[2], Antonio Criminisi[1], David R. Haynor[3], Ender Konukoglu[1]
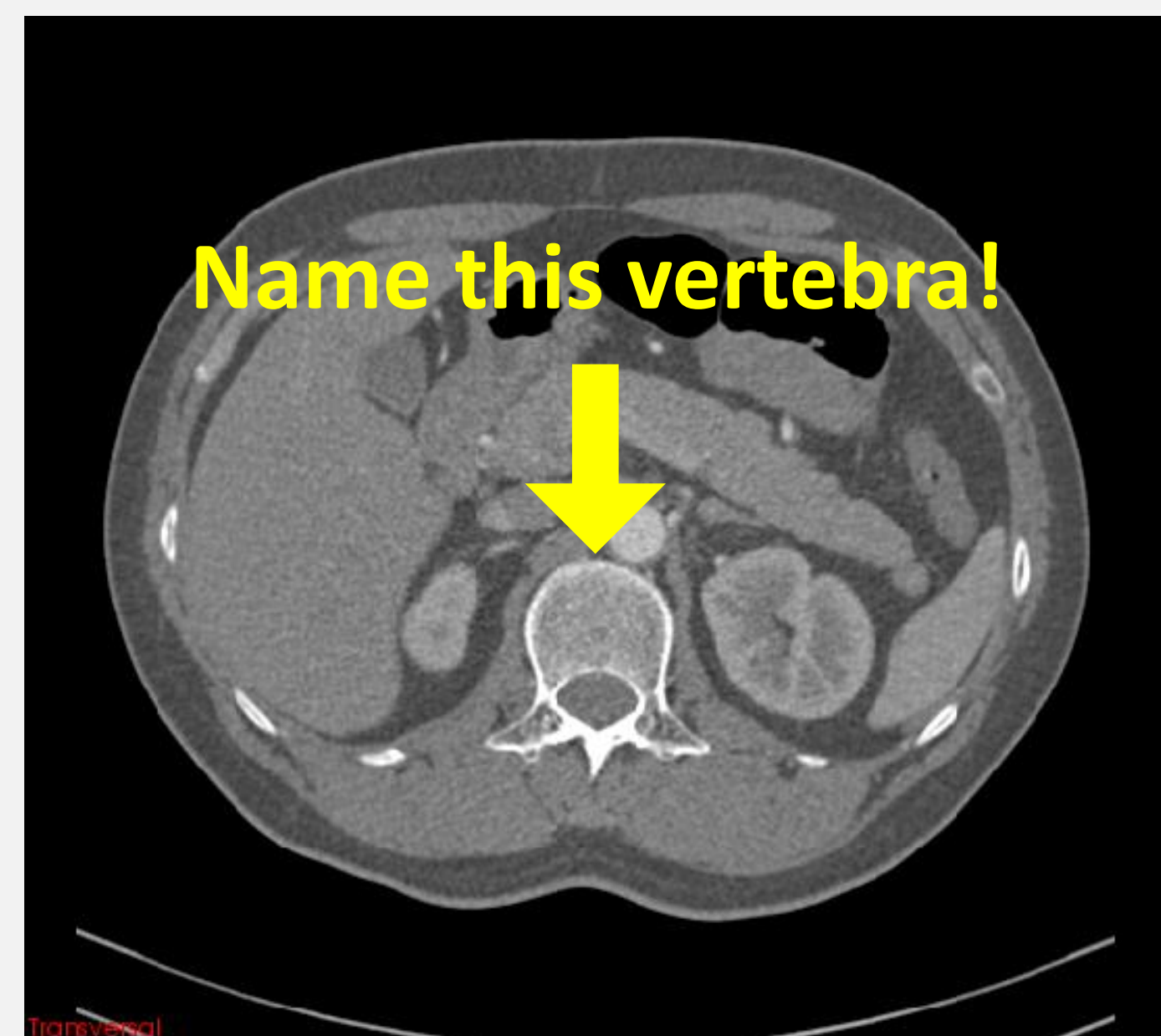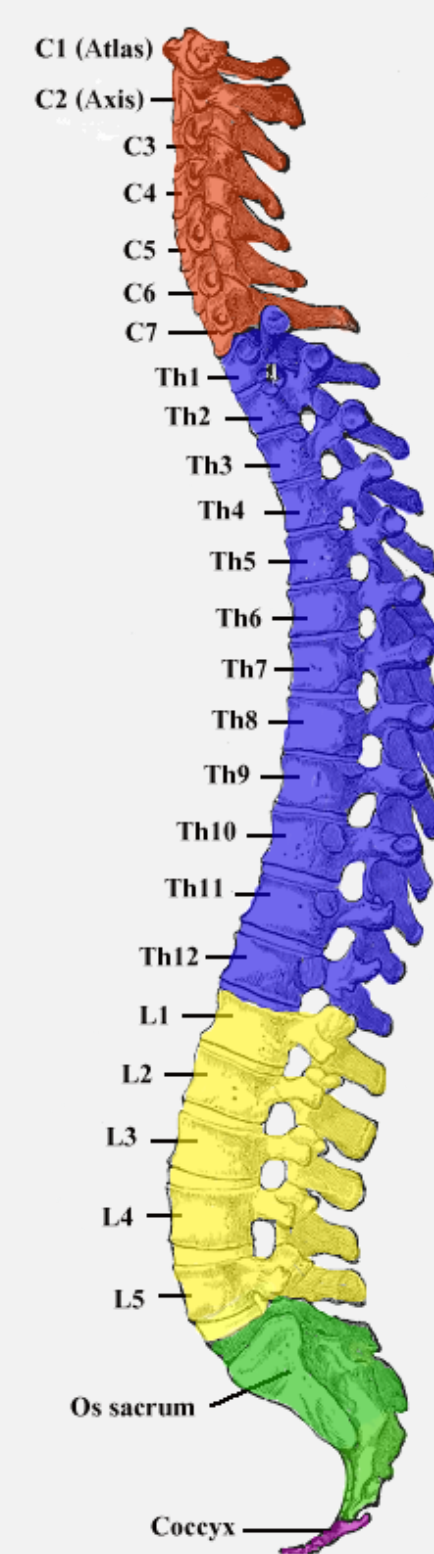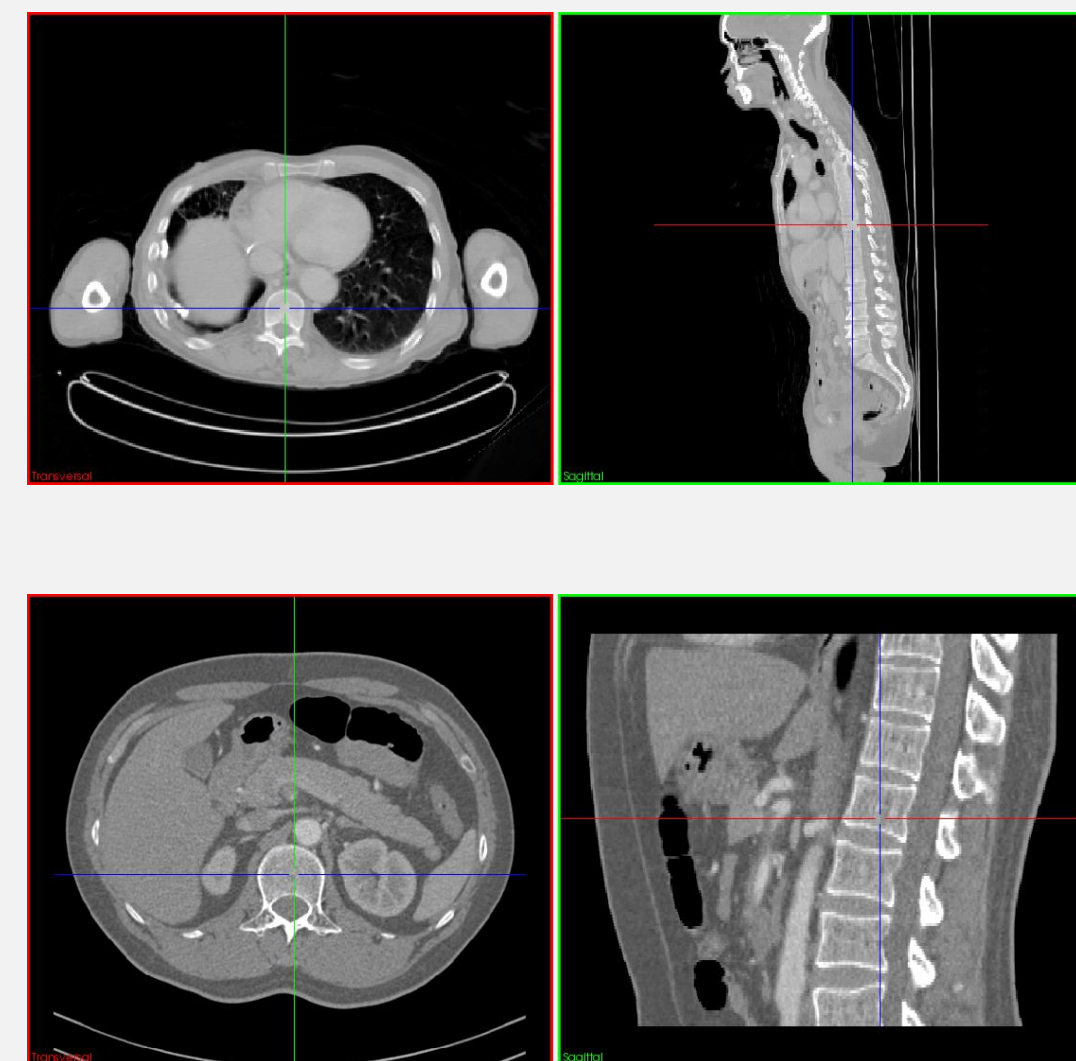
[1]Microsoft Research Cambridge, UK   [2]University of Erlangen-Nuremberg, Germany   [3]University of Washington, Seattle, USA

Microsoft Research

Problem Statement

Name this vertebra!

The Difficulty of Counting

Motivation: Patient-specific coordinate system

- guided visualization/navigation
- shape & population analysis
- registration & segmentation

Challenges

- repetitive appearance
- anatomical variability
- varying imaging parameters
- presence of pathologies

Our Two-Stage Approach

1. Regression Forests: Discriminative

Rough localization via centroid regression

Learn conditional distribution of centroids given feature points
Output: $\mathcal{C} = \{c_i\}$ with $c_i \in \mathbb{R}^3$ and $|\mathcal{C}| = 26$ (centroids)
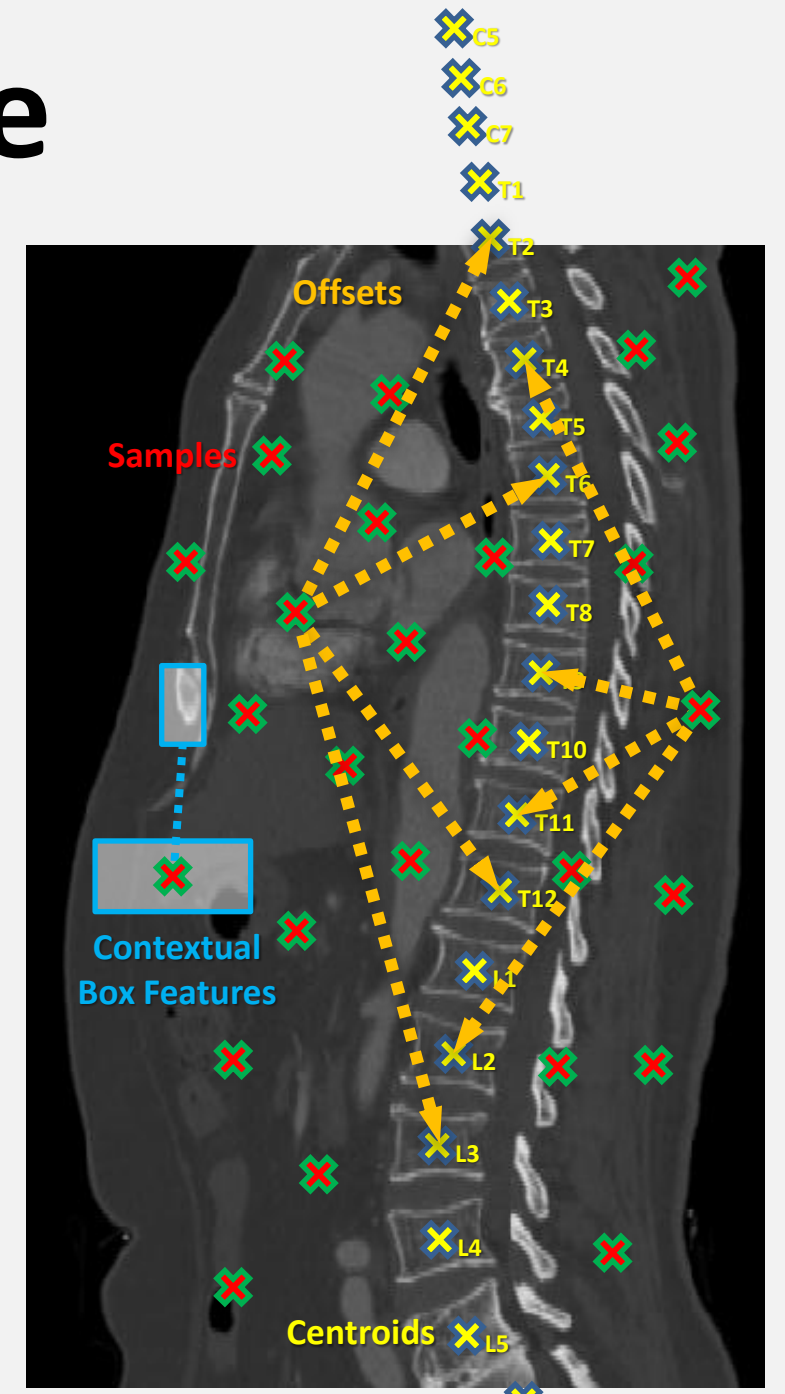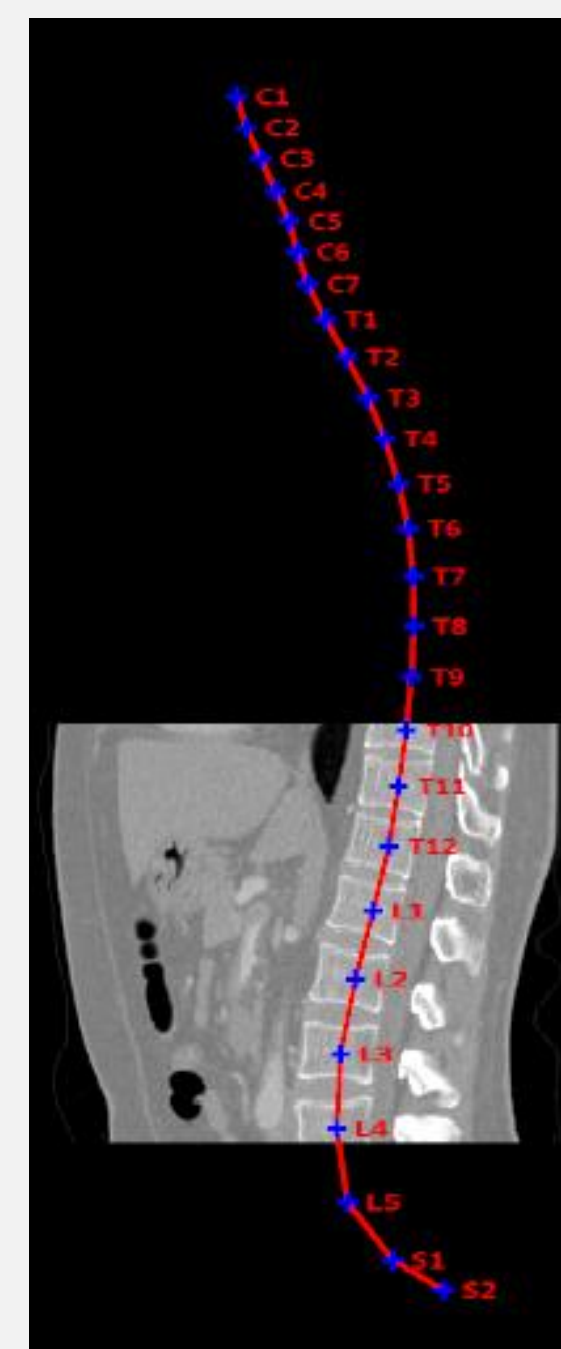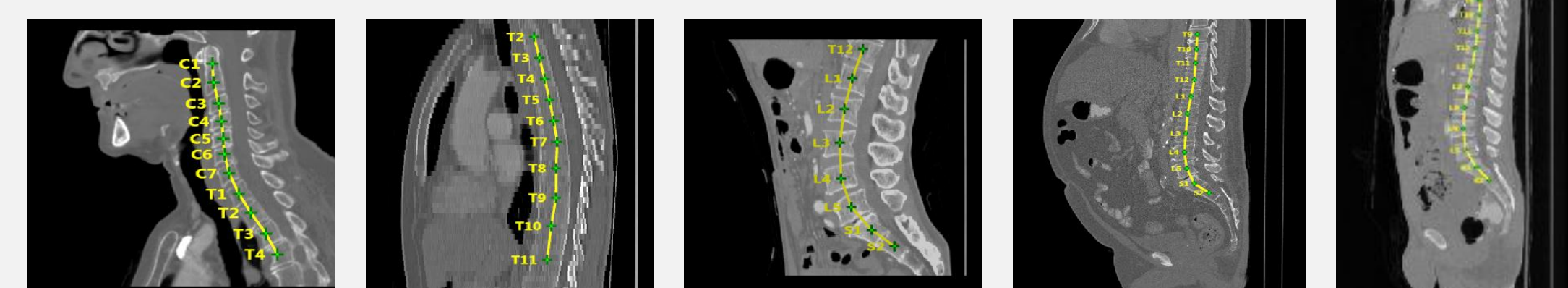Input: $\mathcal{X} = \{(x_j, f_j)\}$ and $x_j \in \mathbb{R}^3$ and $f_j \in \mathbb{R}^m$ (feature points)

Features: Intensity-based box features
Capture local and contextual appearance

Split objective: minimize variance over offsets
$\xi = tr(\Sigma_L) + tr(\Sigma_R)$

Centroid prediction
$\hat{\mathcal{C}} = \arg\max_{\mathcal{C}} p(\mathcal{C}|\mathcal{X})$
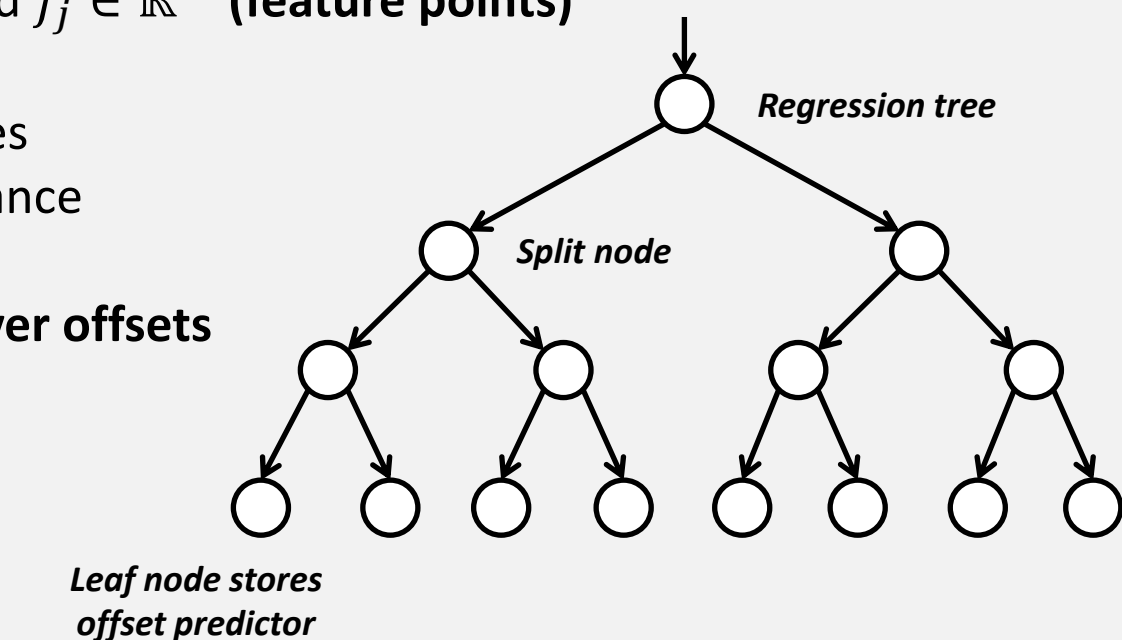
Cluster image points in leaf nodes having similar features and offsets

2. Hidden Markov Model: Generative

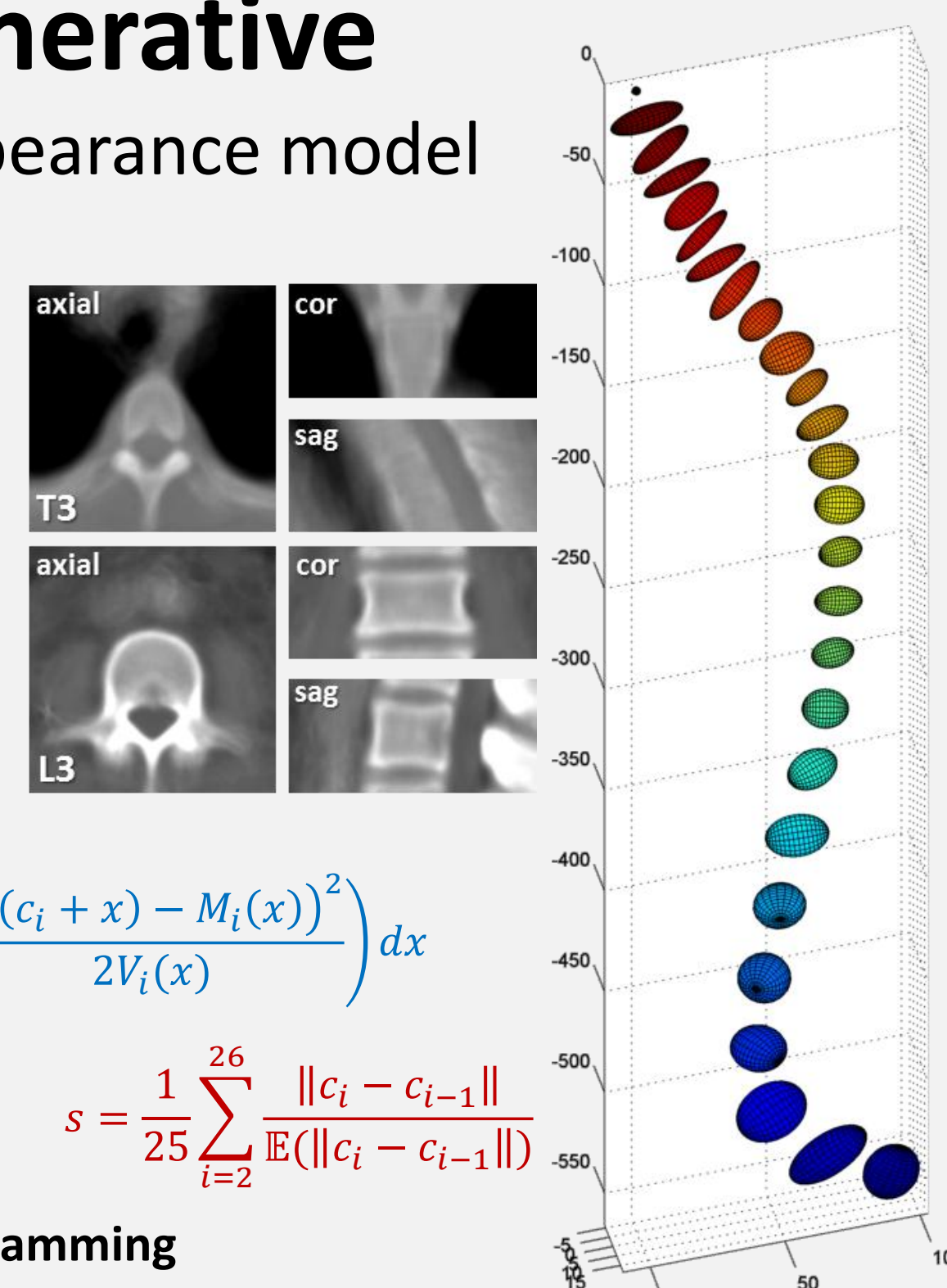Accurate refinement using shape and appearance model
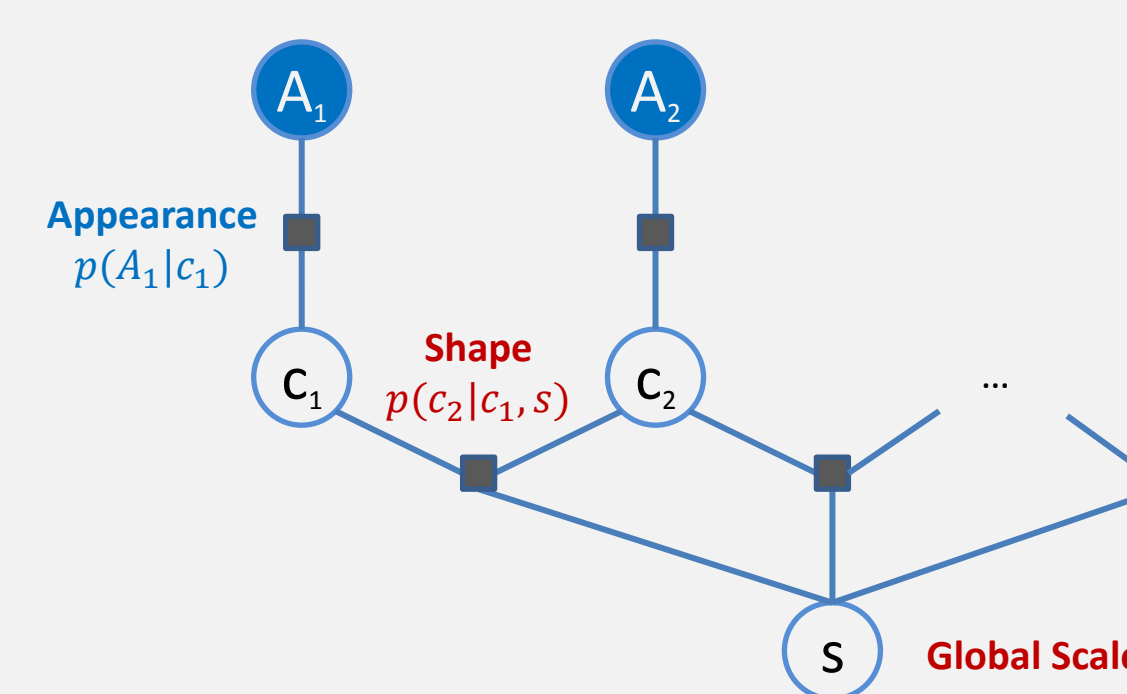
$$p(\mathcal{C}|\mathcal{A},\mathcal{S},s) = \prod_{i=1}^{26} p(A_i|c_i) \prod_{i=2}^{26} p(c_i|c_{i-1},s)$$

Appearance   Shape

$$p(A_i|c_i) = \int_{\Omega_i} \frac{1}{\sqrt{2\pi V_i(x)}} \exp\left(-\frac{(I(c_i+x) - M_i(x))^2}{2V_i(x)}\right) dx$$

$$p(c_i|c_{i-1},s) = \mathcal{N}\left(\frac{c_i - c_{i-1}}{s} \mid \mu_i, \Sigma_i\right) \qquad s = \frac{1}{25}\sum_{i=2}^{26} \frac{\|c_i - c_{i-1}\|}{\mathbb{E}(\|c_i - c_{i-1}\|)}$$

Initialized with forest output, optimized using dynamic programming

Quantitative Evaluation

Clinical Data

- **200 CT scans**, mostly trauma patients: **2595 vertebrae**
- Slice distances between [0.5, 6.5]mm
- Number of slices between [51, 2058]
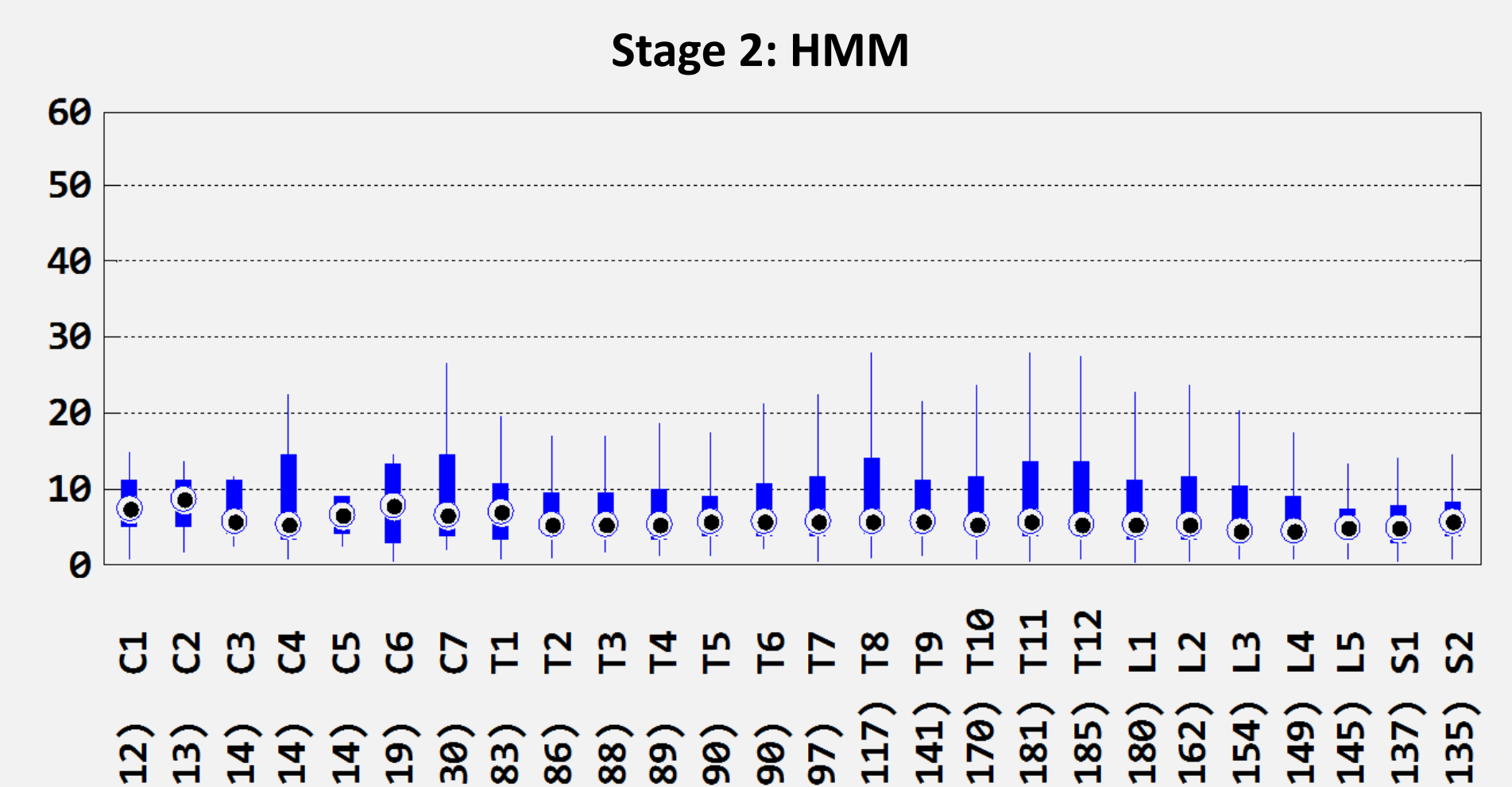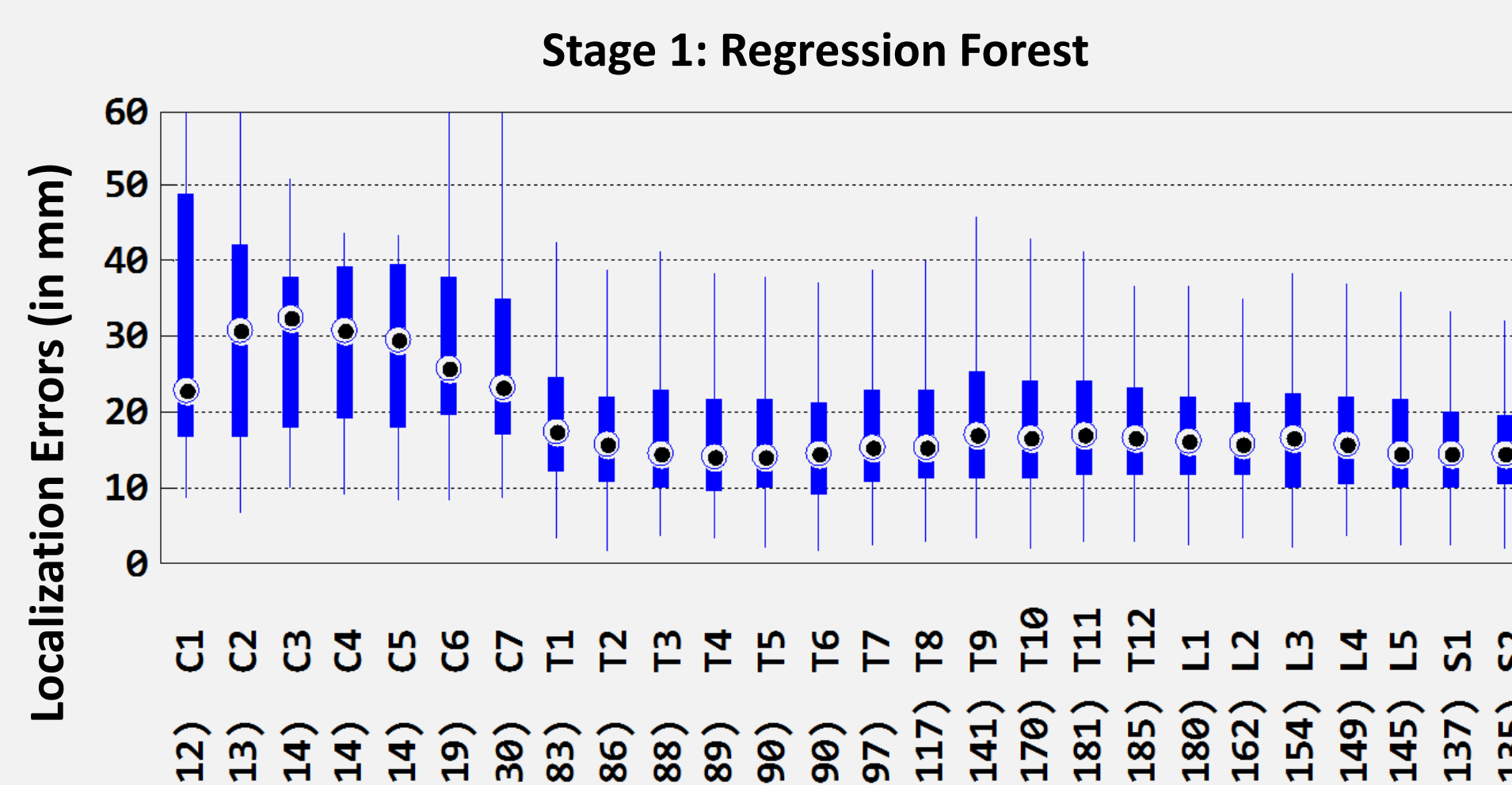- From only 4 vertebrae up to whole-body scans

Experimental Setup

- 2-fold cross-validation (100/100 train/test split)
- 50 trees, depth 20: trained on 5% of image points
- HMM: search over 7 scales between [0.85,1.15]

Computational Efficiency

- Stage 1: about 1 second
- Stage 2: 5-15 seconds per scale
- **Localization of all vertebrae in less than 2 minutes**

Localization Errors & Identification Rates

Vertebrae		Stage 1: Regression Forest			Stage 2: HMM			Distance to Closest			Identification	
Region	Counts	Median	Mean	Std	Median	Mean	Std	Median	Mean	Std	Correct	Rate
All	2595	15.91	18.35	11.32	5.31	9.50	10.55	4.79	6.10	5.53	2089	81%
Cervical	116	25.97	30.74	18.64	6.87	10.85	12.49	6.14	8.53	9.05	84	72%
Thoracic	1417	15.79	18.20	10.81	5.51	9.83	10.44	4.91	5.94	4.84	1100	78%
Lumbar	1062	15.40	17.20	10.07	4.88	8.92	10.45	4.59	6.06	5.82	905	85%

Stage 1: Regression Forest

Stage 2: HMM

Visual Examples