CRF segmentation of cardiac MRIs

J.F. Dreijer
B.M. Herbst
J.A. Du Preez

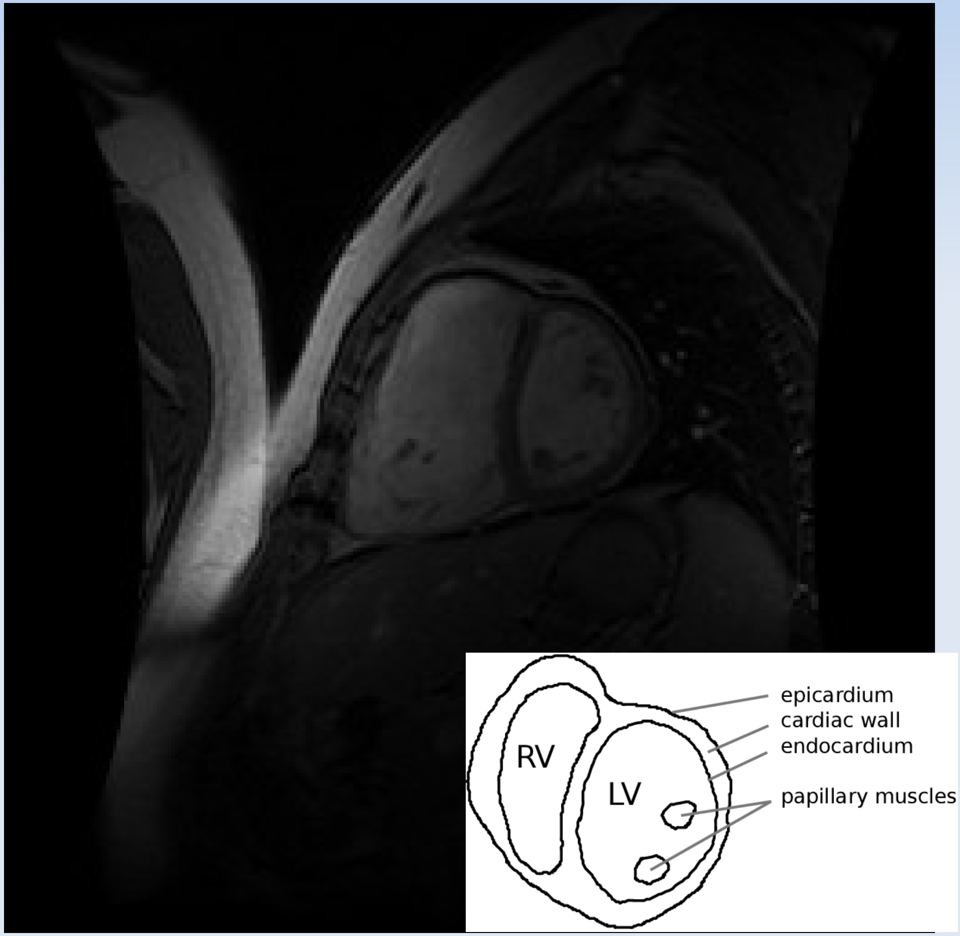
Stellenbosch University
Applied Mathematics / Electronic Engineering

CRF segmentation of cardiac MRIs

- Cardiac MRIs
  - Papillary muscles
- Our approach
  - Problem formulation
  - Feature functions
  - Conditional Random Fields
  - Inference
  - Parameter estimation
- Videos

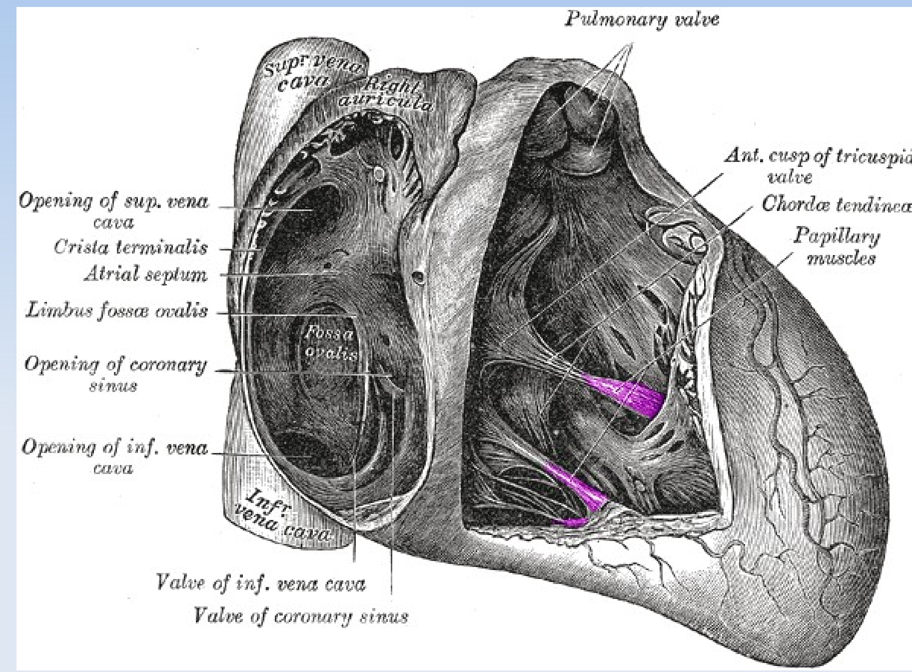
Cardiac MRI segmentation

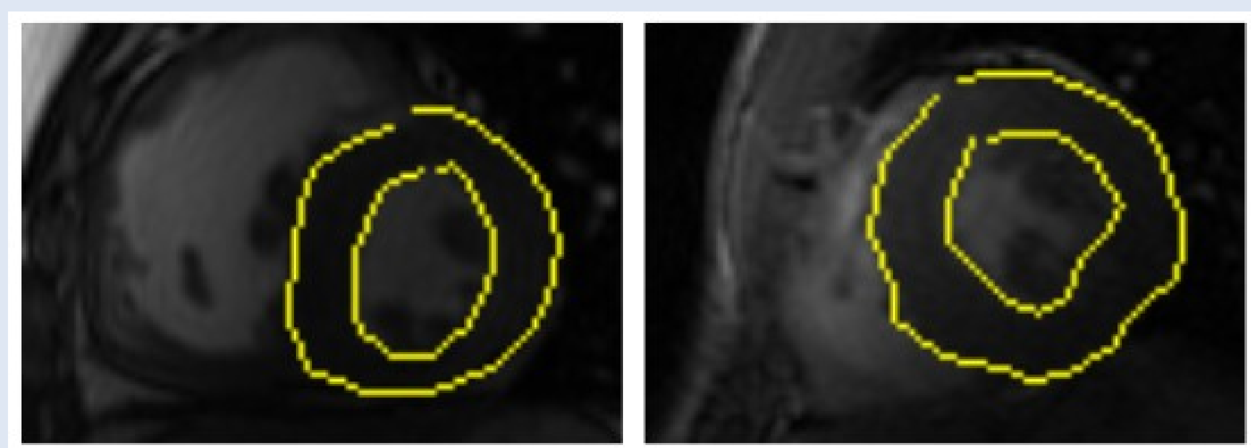
- Magnetic Resonance Imaging
  - Can visualize detailed internal structures
  - Relatively high contrast (compared to ultrasound)
  - Interested in identifying the left ventricle's inner/outer contour



Why is this difficult?

- Papillary muscles
  - Inside ventricles
  - Prevents inversion of heart valves
  - Contrast problems at border of inner contour
  - Esp. at t=T/2





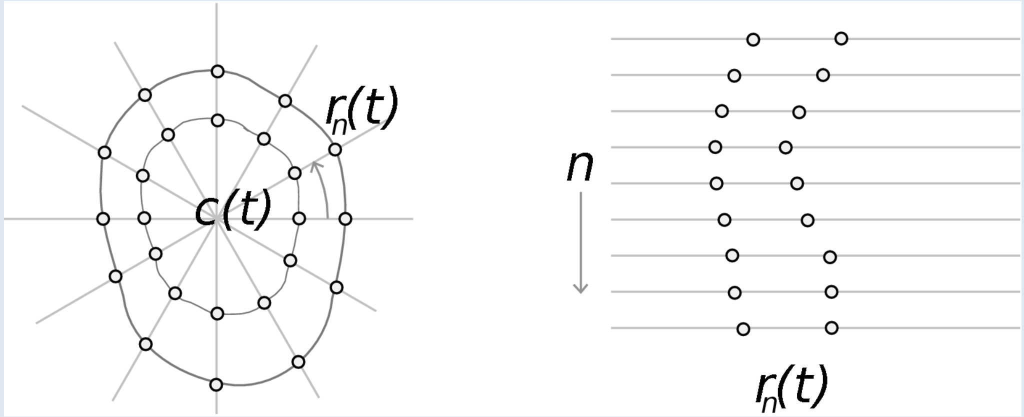
Video

Problem formulation

- Log-polar transform
- Advantage of Polar
  - Simple representation for annular shapes
- Adv. of Log radii
  - Higher resolution of small radii
  - Somewhat linear relationship of inner/outer ratio

$$\rho^{\text{in}} = \left\{\rho_n^{\text{in}}(t)\right\}_{n=0,\dots,N-1,\, t=0,\dots,T-1}$$
$$\rho^{\text{out}} = \left\{\rho_n^{\text{out}}(t)\right\}_{n=0,\dots,N-1,\, t=0,\dots,T-1}$$

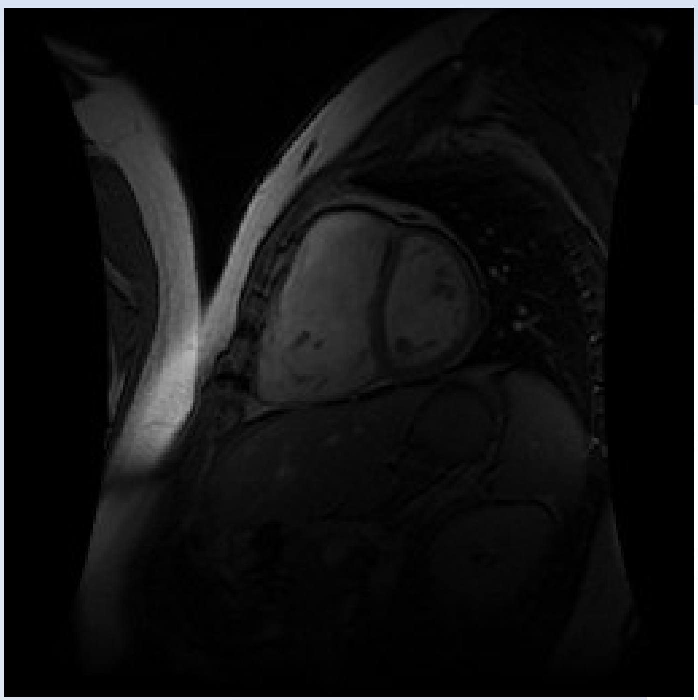
$$c = \{c(t)\}_{t=0,\dots,T-1}$$

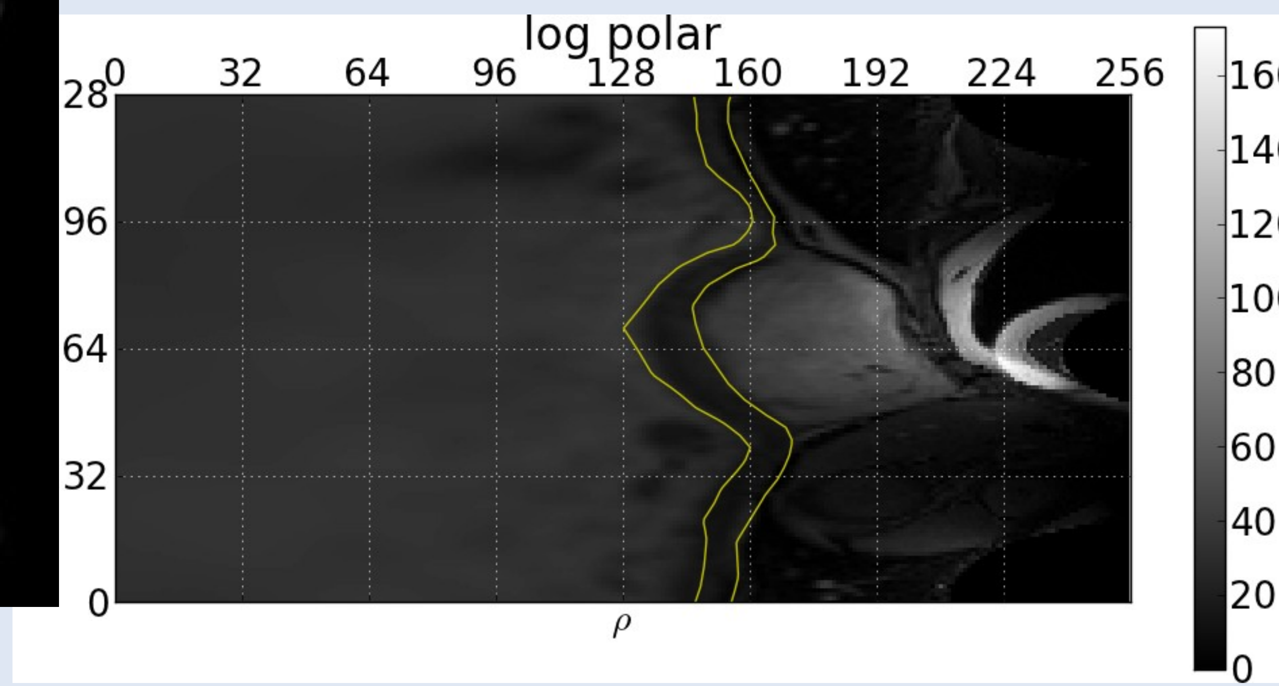


Problem formulation

- Log-polar transform
  - Note papillary

$$
\begin{aligned}
\rho^{\text{in}} &= \left\{\rho_n^{\text{in}}(t)\right\}_{n=0,\dots,N-1,\,t=0,\dots,T-1} \\
\rho^{\text{out}} &= \left\{\rho_n^{\text{out}}(t)\right\}_{n=0,\dots,N-1,\,t=0,\dots,T-1}
\end{aligned}
$$



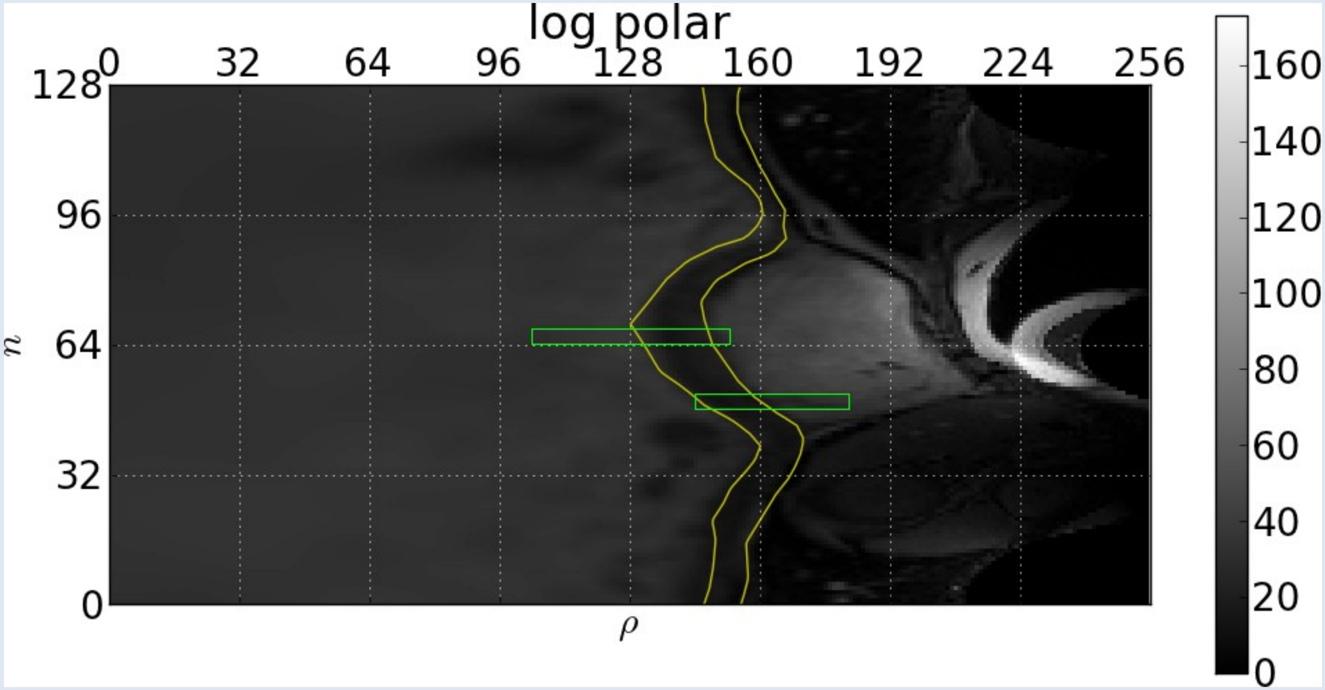


Edge classifier

- Discriminative edge classifier (logit regression)

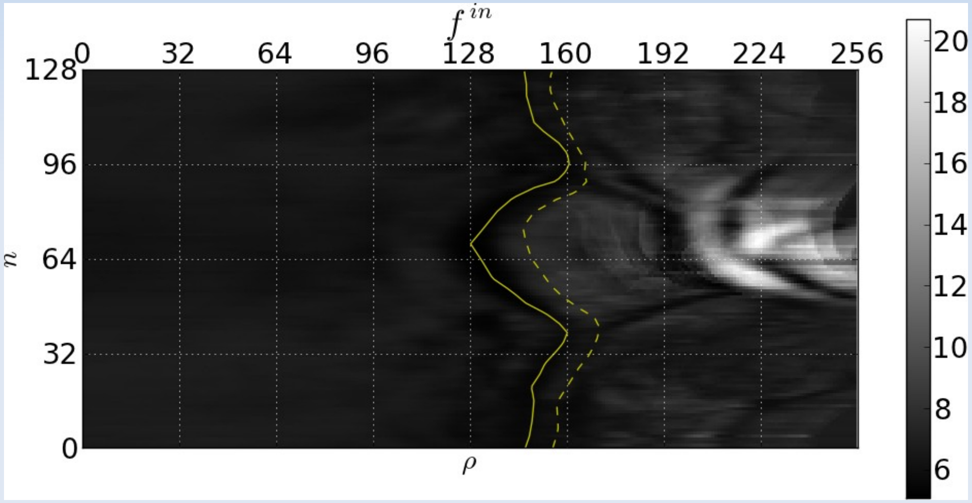
$$P\left(e \,\middle|\, \kappa\left(v_\rho\right), \rho\right) = \frac{1}{1 + \exp\left(-\boldsymbol{\beta} \cdot \kappa\left(v_\rho\right)\right)},$$

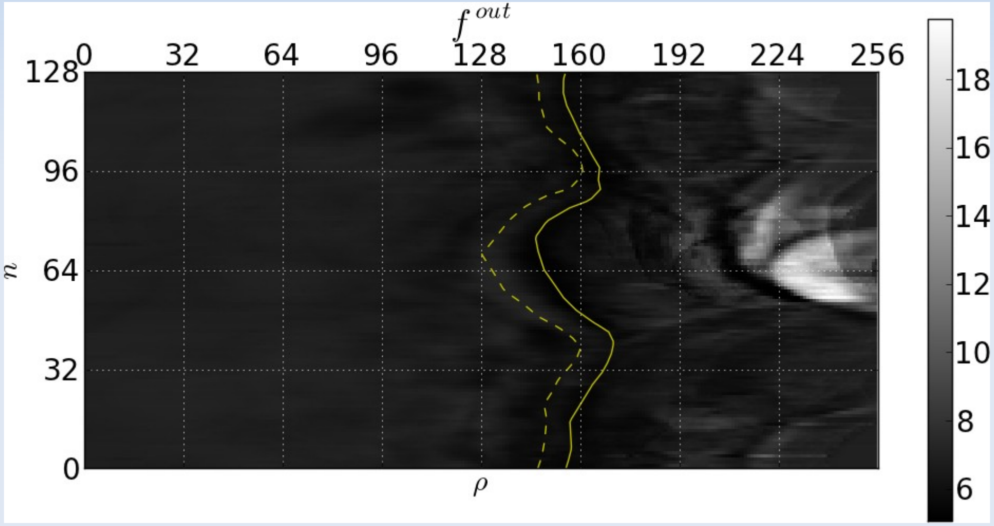
- Features = image gradient in window



Edgeness

- We train different edge a classifier for inner and outer contours





- Not perfect
  - Edges are everywhere (local info only)
  - GT used for training is inconsistent (papillary, no temporal)

Exploit structure of segmentation

- Improve results by incorporating contextual cost functions
- We have a sequence of images
- Relationship between radii
  - Spatial
  - Temporal (dynamic)
  - Inner / outer ratio

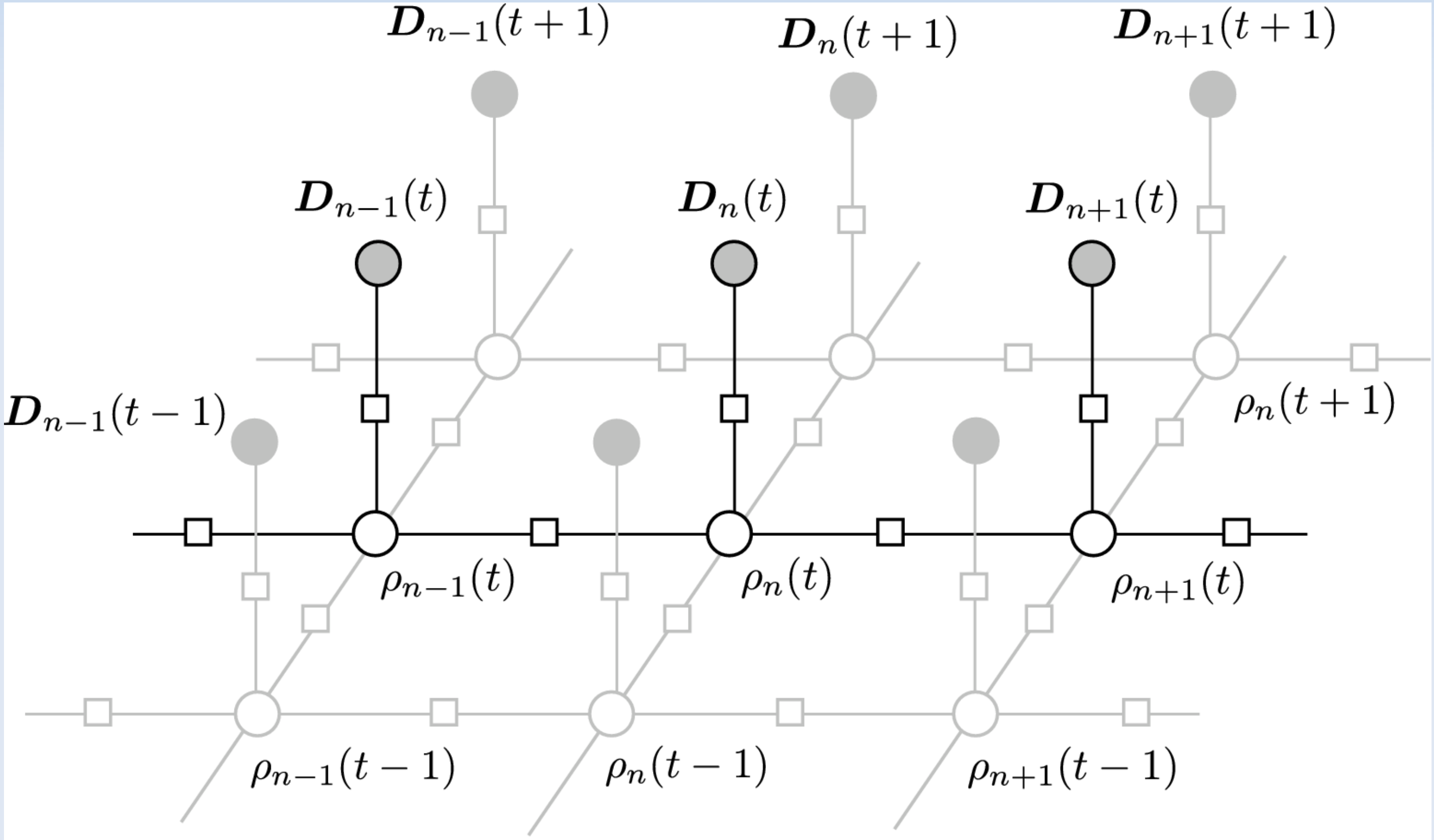
$$f_r\left(\rho_n(t), \rho_{n-1}(t)\right) = \left(\frac{\rho_n(t) - \rho_{n-1}(t)}{M}\right)^2 \quad (13)$$

$$f_t\left(\rho_n(t), \rho_n(t-1)\right) = \left(\frac{\rho_n(t) - \rho_n(t-1)}{M}\right)^2 \quad (14)$$

$$f_t'\left(\rho_n(t), \rho_n(t-1)\right) = \begin{cases} [\rho_n(t-1) < \rho_n(t)] & \text{if } t < t_{ES} \\ [\rho_n(t) < \rho_n(t-1)] & \text{otherwise.} \end{cases} \quad (15)$$

Exploit structure of segmentation

- Improve results by incorporating contextual cost functions



Log-linear Conditional Random Field

- Combine separate cost functions
- Log-linear combination of weighted functions

$$E\left(\boldsymbol{\rho}|\boldsymbol{\theta},\boldsymbol{D}\right)=\sum_{c}\theta_c f_c\left(\boldsymbol{\rho}_c,\boldsymbol{D}_c\right)$$

- CRF model yields probabilistic interpretations

$$P\left(\boldsymbol{\rho}|\boldsymbol{\theta},\boldsymbol{D}\right)=\frac{1}{Z\left(\boldsymbol{\theta},\boldsymbol{D}\right)}\exp\left(-E(\boldsymbol{\rho}|\boldsymbol{\theta},\boldsymbol{D})\right)$$

$$Z\left(\boldsymbol{\theta},\boldsymbol{D}\right)=\sum_{\boldsymbol{\rho}}\exp\left(-E\left(\boldsymbol{\rho}|\boldsymbol{\theta},\boldsymbol{D}\right)\right)$$

Why Conditional Random Fields?

- From Bayesian interpretation of probability
  - Probability is correct way to manage **uncertainty** (anything else is either inconsistent or effectively doing the same thing)
  - Can make (inference) problem tractable through Markov assumptions (i.e. conditional independence)
- We use “Conditional” model (i.e. discriminative as opposed to generative)
  - Not really interested in the joint distribution of the data (can be difficult to simplify and keep accuracy)
  - Only what separates good segmentations from bad

Inference / segmentation

- When provided sequence of images and suitable parameters we want to find segmentation (radii) $\rho^\star = \arg\max_{\rho} P\left(\rho \middle| \theta, D\right)$
- NP-complete problem (in general)
- Approx through Loopy belief propagation
  - Algorithmically similar to dynamic programming
  - If run on problem where functions dependencies form chain or tree will give exact results (=DP)
  - Otherwise approximate solution

Parameter estimation

- **Maximum likelihood** is popular
  - Maximize likelihood of ground truth

$$\theta^\star = \arg\max_{\theta} \prod_{i}^{\text{train}} P\left(\rho^{(i)} | D^{(i)}, \theta\right),$$

  - However partition function is intractable in general
    - 256^(2*128*20) terms

$$Z(\theta, D) = \sum_{\rho} \exp\left(-E\left(\rho | \theta, D\right)\right)$$

  - Biggest challenge in application of RF
  - Common is Pseudo likelihood to estimate
    - Together with approx inference does what?

Parameter estimation

- Rather **minimize the error** between inferred segmentations and ground truth

$$\boldsymbol{\theta}^{\star} = \arg\min_{\boldsymbol{\theta}} \left(e\left(\boldsymbol{\theta}\right) \right)$$

$$e\left(\boldsymbol{\theta}\right) = \sum_{i} e_{\text{dice}} \left(\boldsymbol{\rho}^{(i)}, \arg\max_{\boldsymbol{\rho}} P\left(\boldsymbol{\rho} | \boldsymbol{\theta}, \boldsymbol{D}^{(i)}\right) \right)$$

  - Gradient-free (Powell's method works well)
  - Can use complex loss function

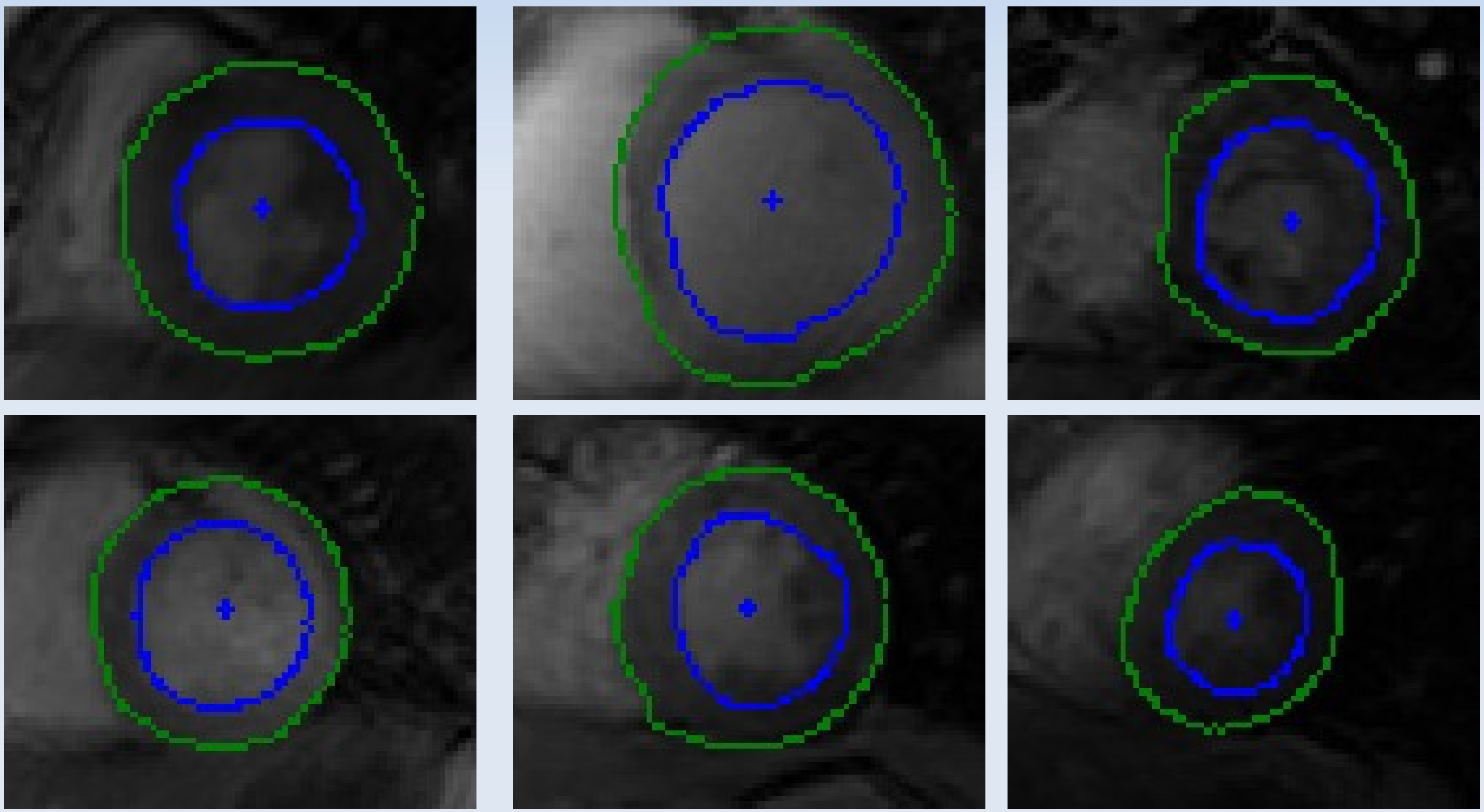
$$e_{\text{dice}}\left(A, B\right) = 1 - \frac{2 \left\| A \cap B \right\|}{\left\| A \right\| + \left\| B \right\|}$$

Things we skipped

- Centre Points
  - C(0) initiated from ground truth, tracked over frames
- Badly labelled ground truth
  - Human annotator looks at one slide at a time, ignores dynamic nature
  - So can easily miss papillary muscle obscuring edge
  - Automatically inferred segmentation can look better than GT, more consistent

Results

- Selection of frames



Videos