Anatomically constrained CNNs

Oktay et al.: IEEE TMI 2018

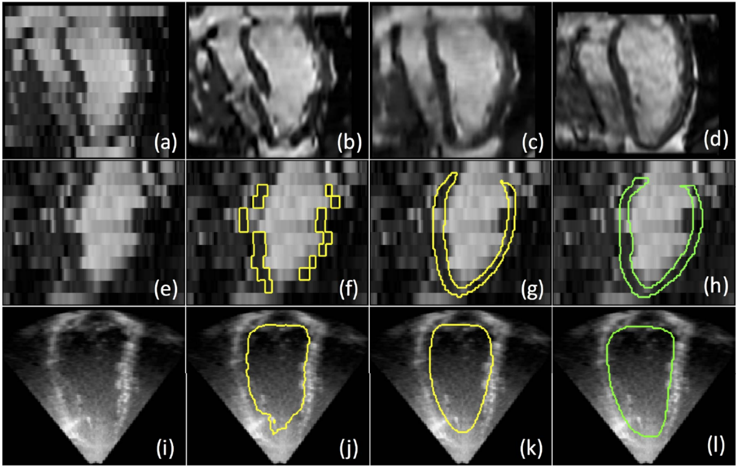
Jan Kybic

2020-2022

ACNNs

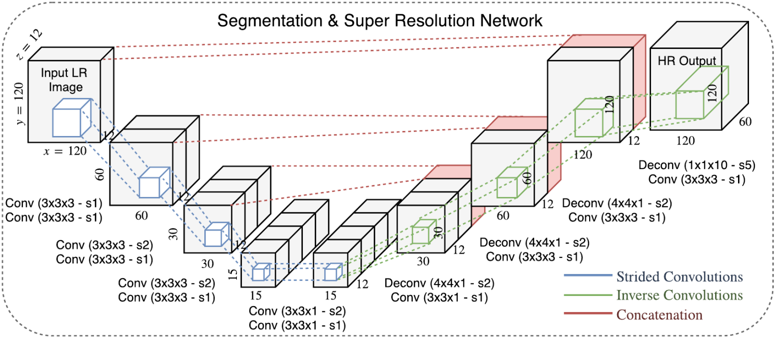
- 3D
- prior shape model
- encoder-decoder to represent shape
- two tasks - segmentation, superresolution
- aim for sub-pixel accuracy
- segmentation uses softmax + cross-entropy cost

Results on heart ultrasound

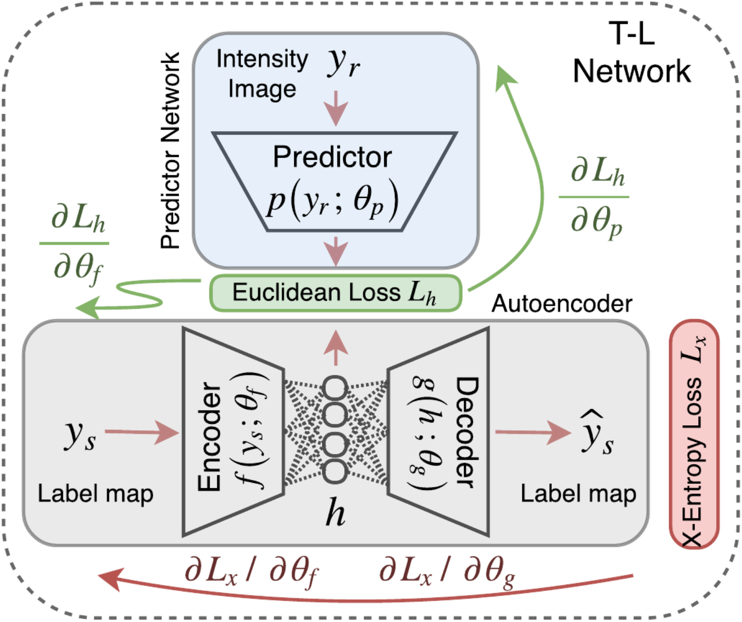


(a) input MR, (b) standard superresolution, (c) this work, (d) reference;
(e) input MR, (f) 2D segmentation, (g) this work, (h) manual;
(i) input 3D US, (j) FCN, (k) this work, (l) reference

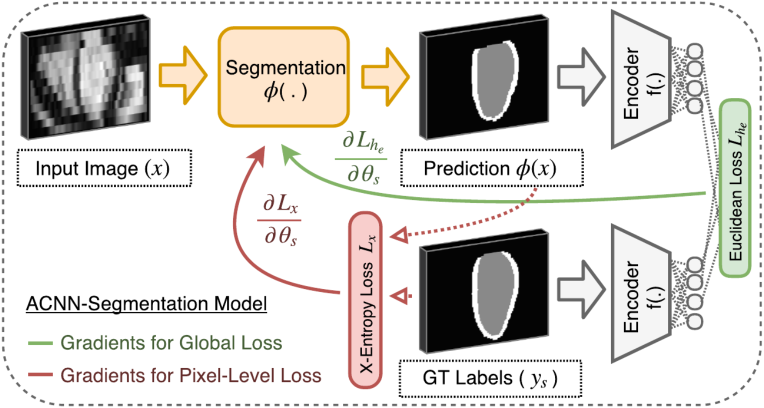
Segmentation & SR network



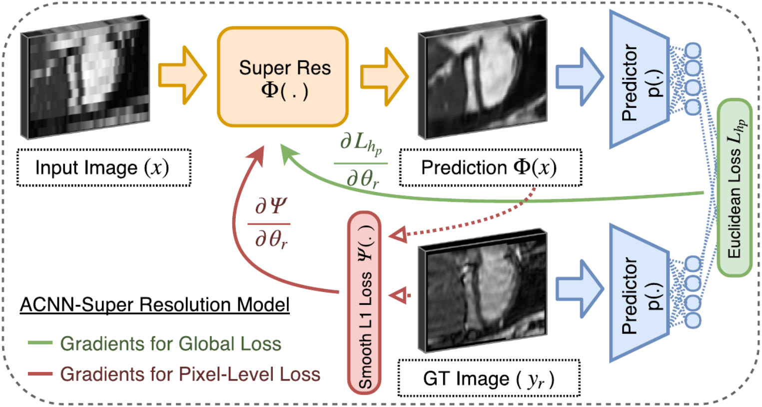
Autoencoder



Segmentation ACNN



Superresolution ACNN



Cost function

ACNN-Seg training objective function though a linear combination of cross-entropy (L_x), shape regularisation loss (L_{h_e}), and weight decay terms as follows:

$$L_{h_e} = \| f(\phi(\boldsymbol{x}); \boldsymbol{\theta}_f) - f(\boldsymbol{y}; \boldsymbol{\theta}_f) \|_2^2$$

$$\min_{\boldsymbol{\theta}_s} \left(L_x \left(\phi(\boldsymbol{x}; \boldsymbol{\theta}_s), \boldsymbol{y}\right) + \lambda_1 \cdot L_{h_e} + \frac{\lambda_2}{2} ||\boldsymbol{w}||_2^2 \right) \quad (1)$$

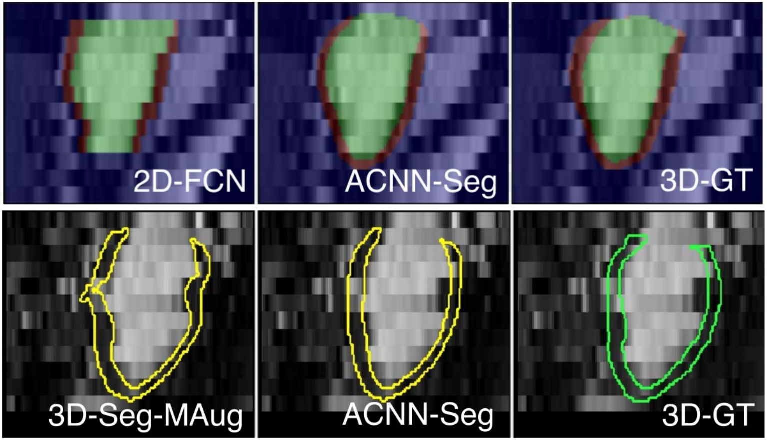
Here $\boldsymbol{w}$ corresponds to weights of the convolution filters, and $\boldsymbol{\theta}_s$ denotes all trainable parameters of the segmentation

imising the smooth ℓ_1 loss, also known as Huber loss, between the ground-truth high resolution image and the corresponding prediction. The smooth ℓ_1 norm is defined as $\Psi_{\ell_1}(k) = \{0.5\, k^2 \text{ if } |k| < 1\,,\ |k| - 0.5 \text{ otherwise}\}$ and the SR training objective becomes $\min_{\theta} \sum_{i \in \mathcal{S}} \Psi_{\ell_1} \left(\Phi\ (\boldsymbol{x}_i\,; \boldsymbol{\theta}_r) - \boldsymbol{y}_i \right)$

Augmentation

- Spatial transformation
- Gaussian noise
- neighborhood label swapping

MR heart example



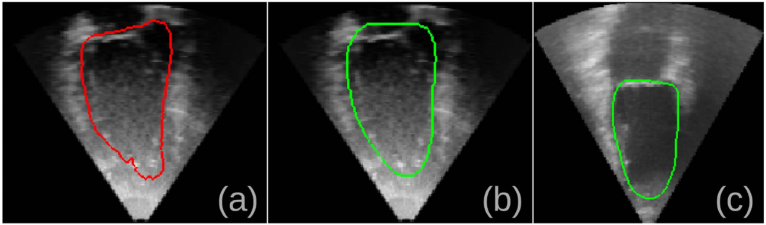


Fig. 7. (a) Cavity noise limits accurate delineation of the LV cavity in apical areas. (b) The segmentation model can be guided through learnt shape priors to output anatomically correct delineations. (c) Similarly, it can make accurate predictions even when the ventricle boundaries are occluded.

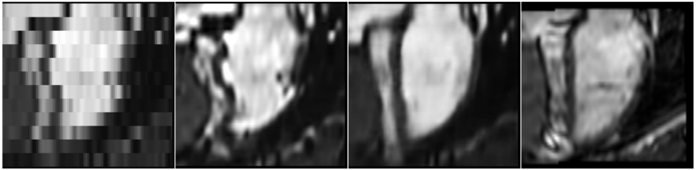


Fig. 8. Image super-resolution (SR) results. From left to right, input low resolution MR image, baseline SR approach [34] (no global loss), the proposed anatomically constrained SR model, and the ground-truth high resolution acquisition.

Classifications

- use autoencoder values as features
- random forest to classify healthy vs. dilated and hypertrophic cardiomyopathy patients (92% accuracy)