Interpretable Medical Image Classification with Self-Supervised Anatomical Embedding and Prior Knowledge

Ke Yan[1], Youbao Tang[1], Adam P. Harrison[1], Jinzheng Cai[1], Le Lu[1], Jingjing Lu[2]

[1] PAII Inc., Bethesda, MD, USA [2] Beijing United Family Hospital, Beijing ,China

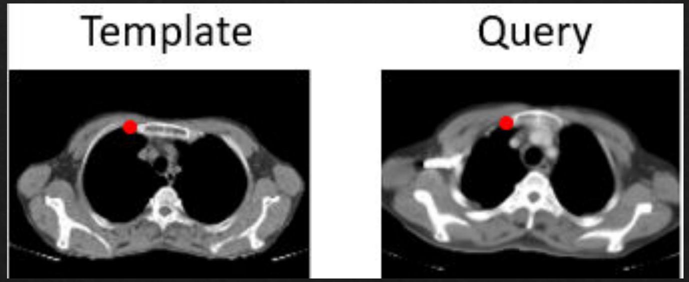
yankethu@gmail.com

Motivation

- CNN classifiers need interpretability
  - The network's attention should be on the correct body part
- Network visualization algorithms (e.g., CAM) are
  - Post-hoc, implicit, uncontrollable
- Our method: explicitly inject clinical prior knowledge
  - Manually specify body parts of interest
  - Guide the network to directly learn from these regions
- Challenge
  - Cost many labeled images to train organ segmentation or landmark detection models

Related Work

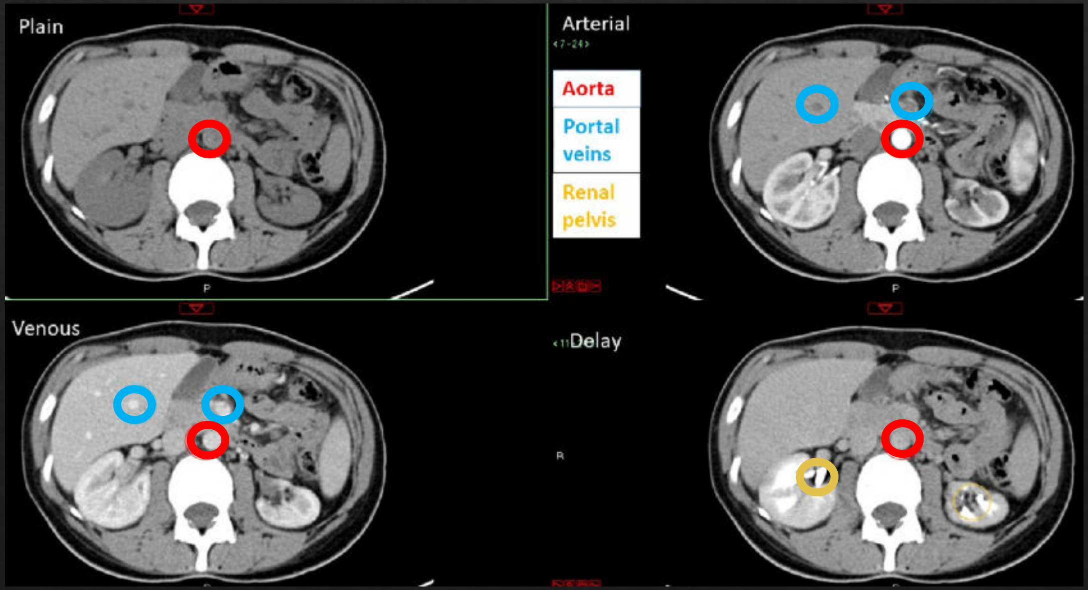
- Self-Supervised Anatomical Embedding (SAM)
  - A convenient tool to detect arbitrary anatomical landmarks in radiological images
  - Low-cost, only need one labeled image as template
  - Has been employed in one-shot landmark detection and lesion matching



Ke Yan et al., Self-supervised learning of pixel-wise anatomical embeddings in radiological images, 2020. https://arxiv.org/abs/2012.02383 .

Problem

- An example of the SAM-based image classification
- Task: classify four phases in a dynamic liver CT [1]
  - Non-contrast / Arterial / Venous / Delay
- Data: 1000 training volumes, 491 testing
- Baseline: 3DSE [1] (train CNN with whole 3D image as input)
- Prior knowledge
  - The Hounsfield unit (HU) values of certain anatomical landmarks correlate with contrast phases



[1] Bo Zhou et al., "CT data curation for liver patients: Phase recognition in dynamic contrast-enhanced CT". In MICCAI - Med. Image Learn. with Less Labels Imperfect Data, 2019.

Method

1. Manually label 32 phase-related landmarks on one random patient as guidance
2. Use SAM to detect these landmarks on all unlabeled images
3. Crop a $3 \times 3 \times 3$ patch around each landmark
4. Extract the maximum HU value on each patch, get a 32D feature vector for each image
5. Train a linear SVM on the z-scored features for phase classification

Results

Table 1: Comparison of F1 scores for contrast phase classification, trained on 1,000 volumes and tested on 491 volumes.

Method	Non-contrast	Arterial	Venous	Delay	Mean
3DSE (Zhou et al., 2019)	**98.5**	**97.4**	91.8	90.4	94.5
SAM + SVM (proposed)	98.4	**97.4**	**94.4**	**93.4**	**95.9**

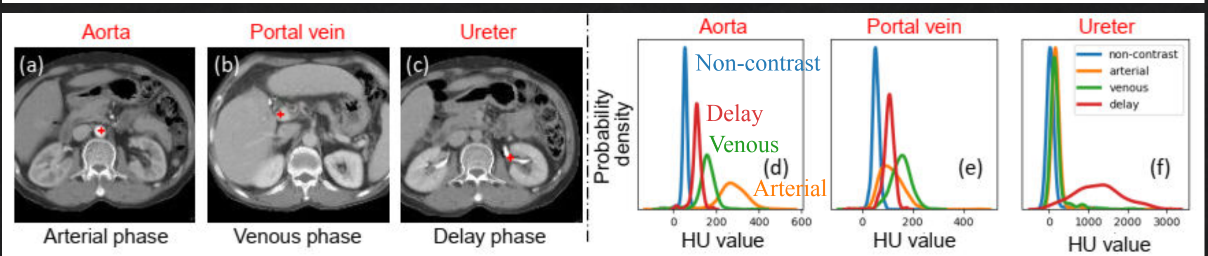


Figure 1: (a-c): Examples of phase-related anatomical landmarks detected by SAM. The phase label of each image is also shown. (d-f): The probability distribution of HU values of different phases on the three anatomical landmarks.