HKU Med LKS Faculty of Medicine The University of Hong Kong 香港大學李嘉誠醫學院
DEPARTMENT OF DIAGNOSTIC RADIOLOGY

3D-RADNet: Extracting labels from DICOM metadata for training general medical domain deep 3D convolution neural networks

Richard Du, MSc & Varut Vardhanabhuti, FRCR, PhD

Purpose

Training deep convolution neural network requires a large amount of data

Transfer learning from pre-existing large datasets are typically to improve performance.

Labelling large amount of medical images maybe difficult due to expert domain knowledge

Lack of large datasets for 3D medical images

Related Work

Tencent MedicalNet (Chen et al. 2019)

3D CNN (3D ResNet) trained on segmentation data

Large dataset involving mainly public dataset and private

Segmentation data require expert and time consuming annotations

Network	Pretrain	LungSeg(Dice)
3D-ResNet10	Train from scratch	69.31%
	MedicalNet	96.56%
3D-ResNet18	Train from scratch	70.89%
	MedicalNet	94.68%
3D-ResNet34	Train from scratch	75.25%
	MedicalNet	94.14%
3D-ResNet50	Train from scratch	52.94%
	MedicalNet	89.25%

References:
Sihong Chen, Kai Ma, and Yefeng Zheng. Med3D: Transfer Learning for 3D Medical Image Analysis. CoRR, abs/1904.00625, 2019. URL http://arxiv.org/abs/1904.00625.

Our Approach

We examined the use of DICOM metadata and also heuristics of the datasets to semi-automatically labelled large amount of medical images.

DICOM tags such as modality, sequences and orientations may provide descriptive features of the images to train a network.

We applied our method to large cancer dataset (TCIA) consisting of over 60,000+ imaging series.

Image Metadata:

Key	Value
Focal Spot(s)	1.200000
Convolution Kernel	SOFT
Patient Position	HFS
Revolution Time	0.9
Single Collimation Width	0.625
Total Collimation Width	10
Study Instance UID	1.2.276.0.7230010.3.1.2.296485376.1.1521714307.2031999
Series Instance UID	1.2.276.0.7230010.3.1.3.296485376.1.1521714314.2034075
Study ID	
Series Number	2
Acquisition Number	1
Instance Number	1
Image Position (Patient)	-120.200\-104.478\-11.702
Image Orientation (Patient)	1.000000\0.000000\0.000000\0.000000\0.987688\-0.156434
Position Reference Indicator	OM
Slice Location	-28.250
Samples per Pixel	1
Photometric Interpretation	MONOCHROME2
Rows	512
Columns	512
Pixel Spacing	0.451172\0.451172
Bits Allocated	16
Bits Stored	16
High Bit	15
Pixel Representation	1
Pixel Padding Value	63536
Window Center	30

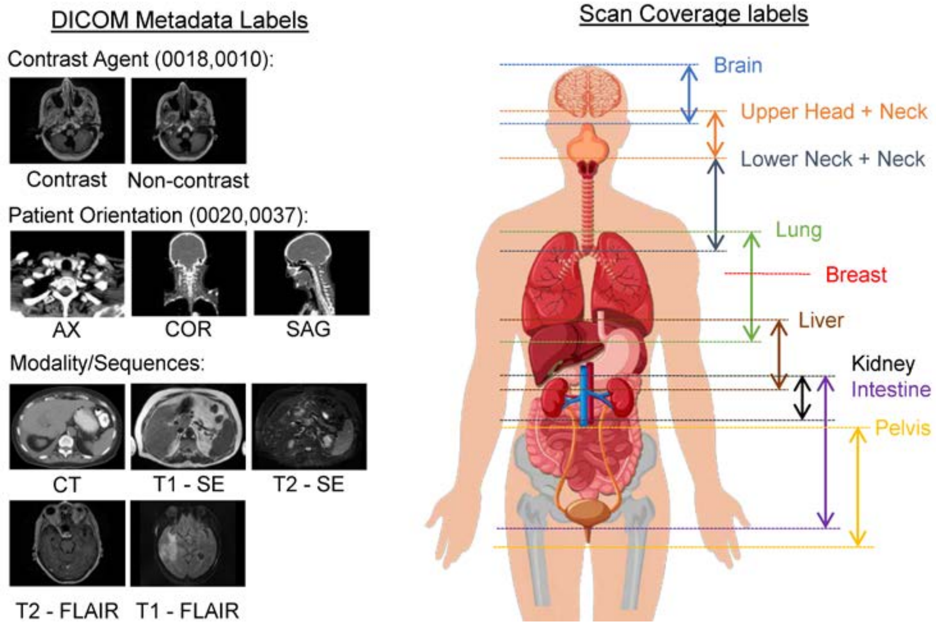
Network and Training

Adopted ResNet to 3D ResNet

Trained to classify the modality/sequences, anatomical planes, presence of contrast agent and body coverage

A total of 15305 image series were used for training

Hold out test set of 316 series from individual patients were used for evaluation.



Results

	Samples	Accuracy	AUC
Modality/Sequence			
CT	117	100%	100%
T1 - SE	51	97.8%	98.1%
T2 - SE	101	96.8%	98.8%
T1 - FLAIR	27	99.7%	100%
T2 - FLAIR	20	97.5%	96.5%
Total	316	95.9%	
View			
Axial	267	99.4%	100.0%
Coronal	32	100%	100%
Sagittal	17	99.4%	100.0%
Total	316	99.4%	
Contrast			
Contrast	100	84.8%	91.7%
No Contrast	216	84.8%	91.7%

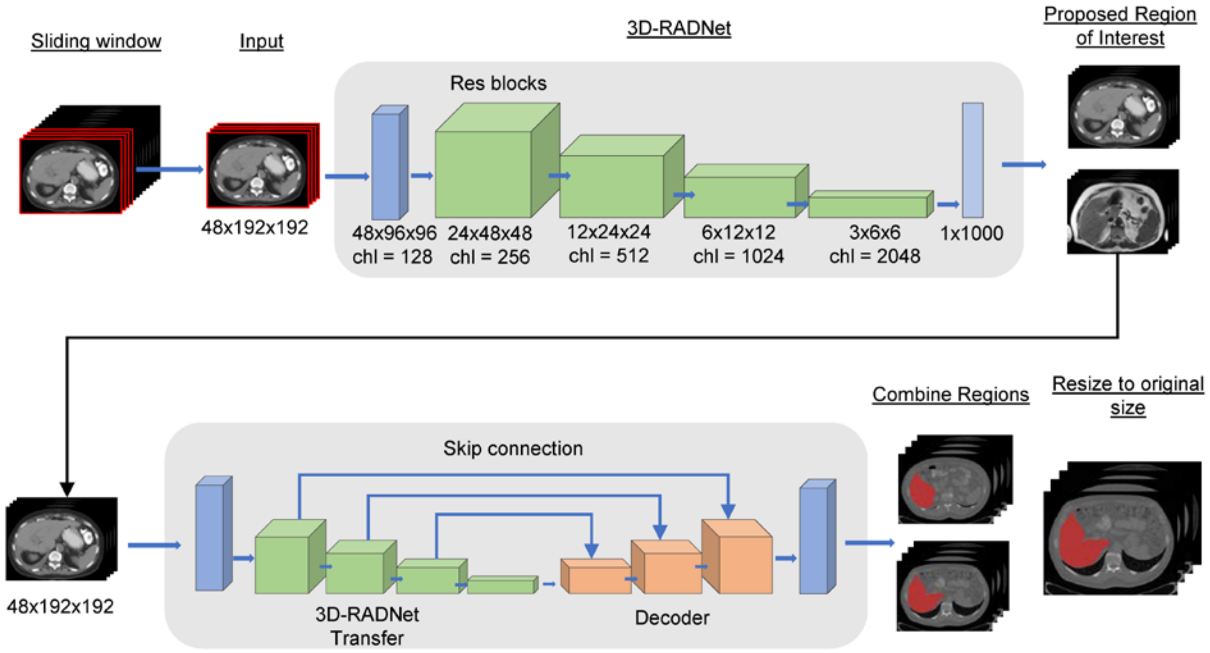
	Samples	Accuracy	AUC
Scan Coverage			
Brain	200	98.4%	99.9%
Upper Head-Neck	42	98.7%	99.7%
Lower Head-Neck	38	98.7%	98.7%
Lung	64	98.4%	98.8%
Breast	82	98.7%	99.6%
Liver	52	99.1%	99.7%
Kidney	52	98.7%	99.8%
Intestine	46	98.1%	100.0%
Pelvis	55	98.1%	99.9%
Total	316	91.5%	

Experiment

Transfer learning to a liver segmentation task (LITS dataset)

Standard VNet Architecture with skip connections

Region of proposal of liver slices by using a sliding window approach



Results

Table 2: Impact of liver segmentation performance by freezing different layers of the network and comparison of performance to no transfer learning is also given.

Modification	Mean DICE	Mean IOU	Median DICE	Median IOU
Froze all layers	90.0%	80.7%	90.0%	81.8%
Froze to block 3	72.1%	59.7%	81.0%	68.1%
Froze to block 2	84.3%	76.8%	90.3%	82.3%
Froze to block 1	77.6%	65.1%	81.8%	69.3%
Weights only	44.8%	38.1%	43.0%	27.4%
No transfer learning	41.8%	38.0%	43.0%	27.3%

Results

Table 3: Impact of liver segmentation performance by percentage of training samples

% Training Data	Mean DICE	Mean IOU	Median DICE	Median IOU
100% (n=92)	90.0%	81.6%	90.1%	82.0%
80% (n=74)	84.7%	75.0%	89.0%	80.0%
60% (n=55)	84.2%	73.9%	87.2%	77.3%
40% (n=37)	87.0%	77.9%	90.9%	83.4%
20% (n=18)	80.0%	69.30%	86.5%	76.1%

Summary

We devised a strategy to extract useful labels from large amount of medical images to train medical images

Network trained from public TCIA datasets can be used for transfer learning onto other tasks.

Good segmentation performance can be achieved even with low number of training data

Thank you very much!