Identifying Carotid Plaque Composition in MRI with Convolutional Neural Networks

Author: Yuxi Dong

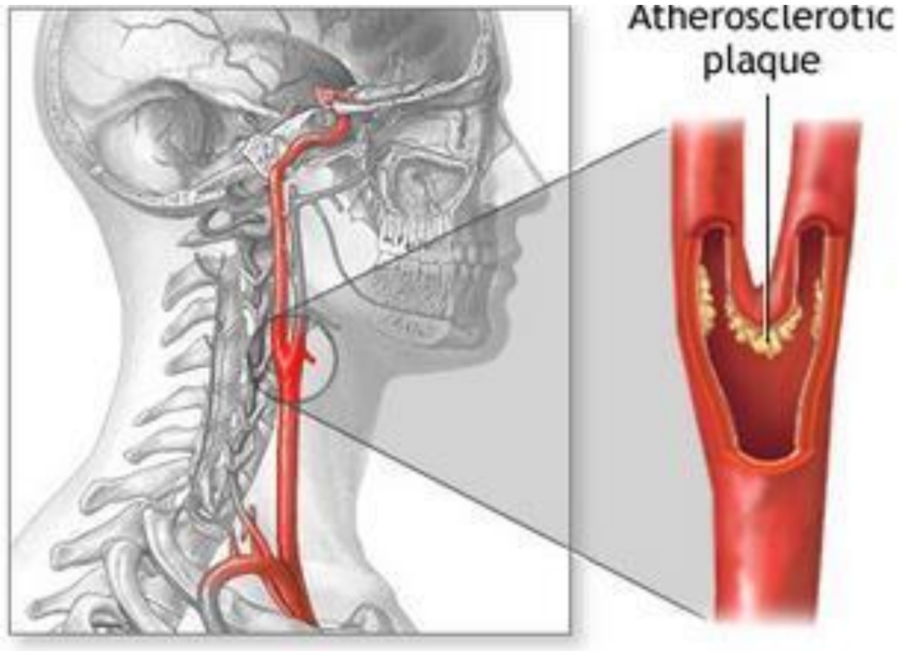
Director: Wei Xu

清華大學
Tsinghua University

交叉信息研究院
Institute for Interdisciplinary Information Sciences

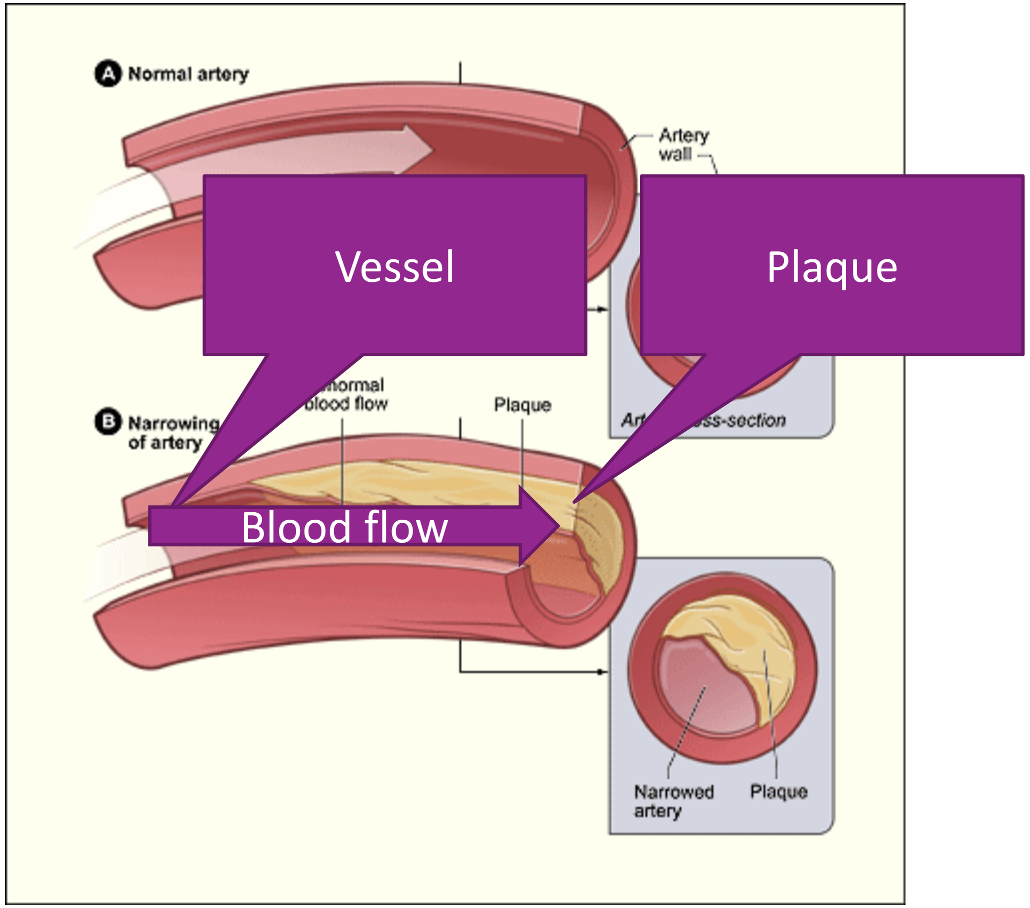
Background: Atherosclerosis

- Caused by accumulation of substances in arteries
- Cause stroke, the second place in global death ranks from 1990 to 2010



Background: Dangerous of carotid plaques

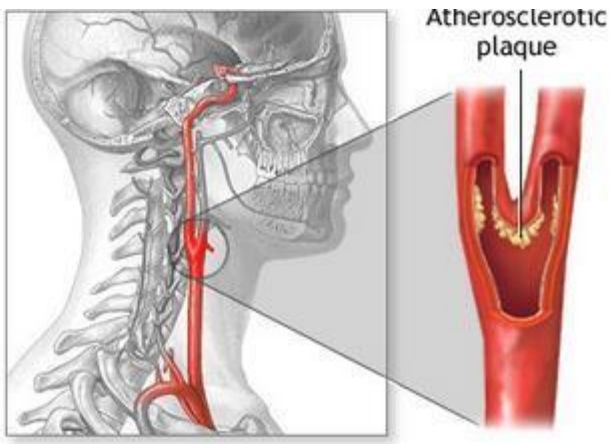
- What we see: Plaque
  - Reduced or blocked blood flow
- When a plaque breaks up
  - Rupture from vessel
  - Flow with blood to other parts of body
  - May block the vessel somewhere
- Composition of the plaque => different risk level



Atherosclerosis[1]

[1] What Is Atherosclerosis?, http://www.nhlbi.nih.gov/health/health-topics/topics/atherosclerosis/

Our goal: Identify composition of plaques



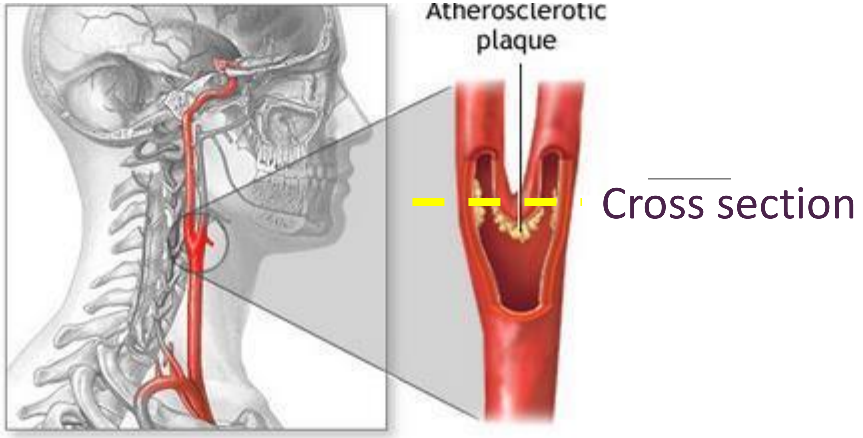
- We focus on carotid vessels (arteries on the neck)
- Traditional method: MRI + Trained radiologist
  - Time consuming
  - Requires expertise
  - Inter-reviewer variability
- We want to identify the composition of carotid plaques in MRI automatically

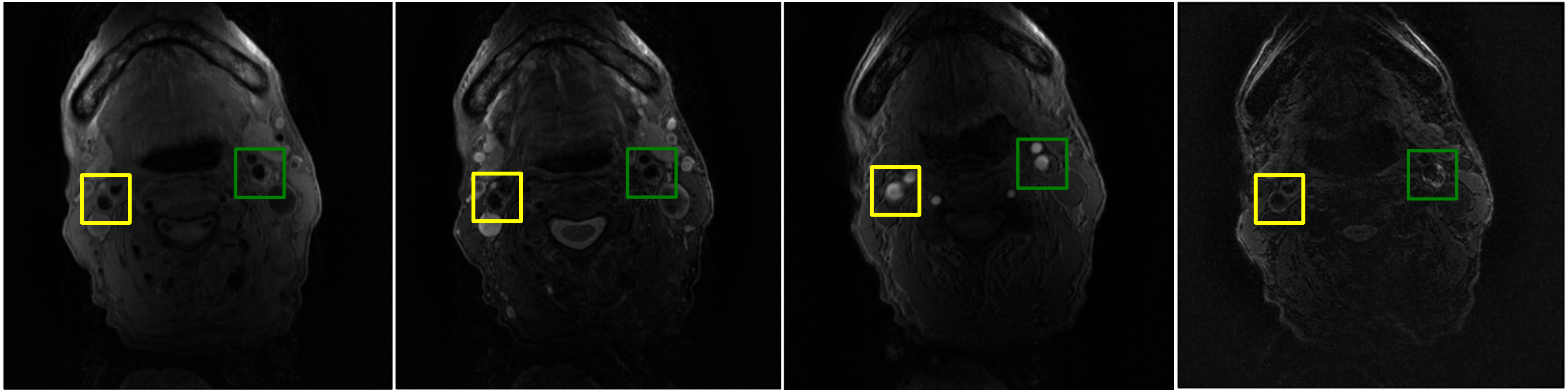
Outline

- Background of MRI and plaques
- Dataset and preprocessing
- Our model
- Evaluation

MRI produces multi-contrast images

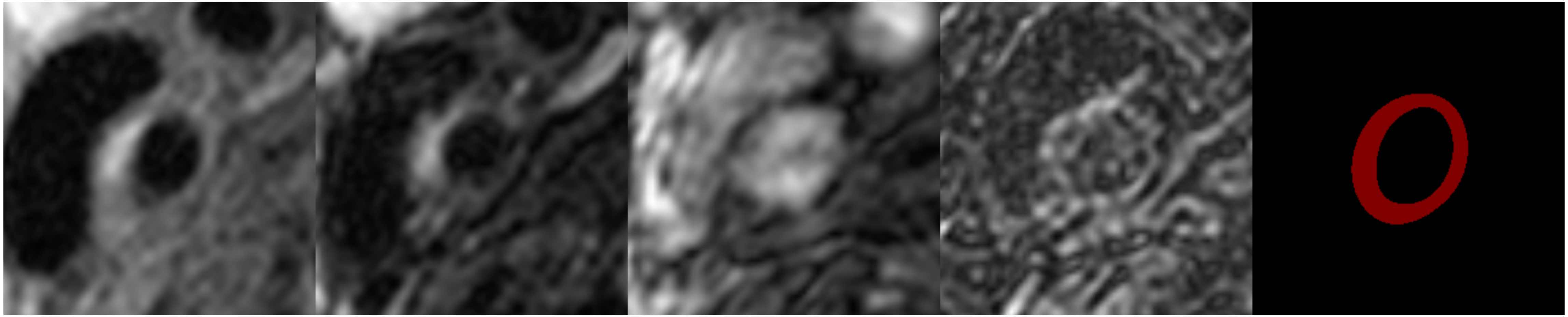
- 4 *contrast weightings:* T1W, T2W, TOF, MP-RAGE
- Each from a different physical scanning method





T1W | T2W | TOF | MP-RAGE

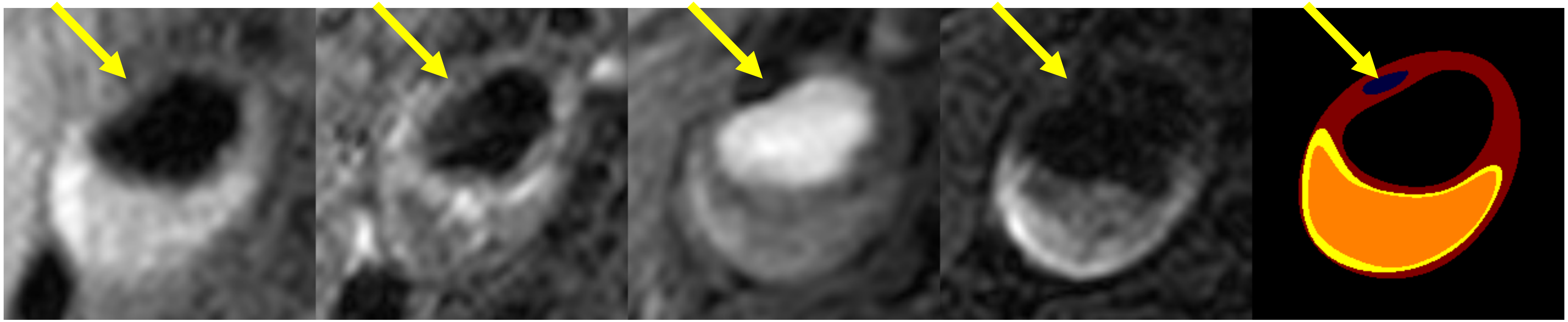
The vessel: when it is normal



T1W T2W TOF MP-RAGE Manual

When there is a plaque

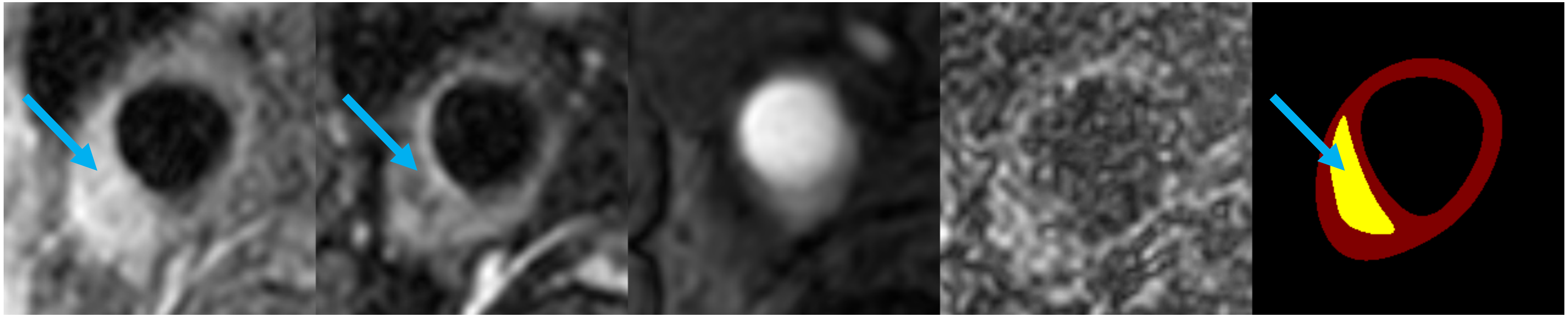
- Calcification: calcium builds up in blood vessels



T1W T2W TOF MP-RAGE Manual

When there is a plaque

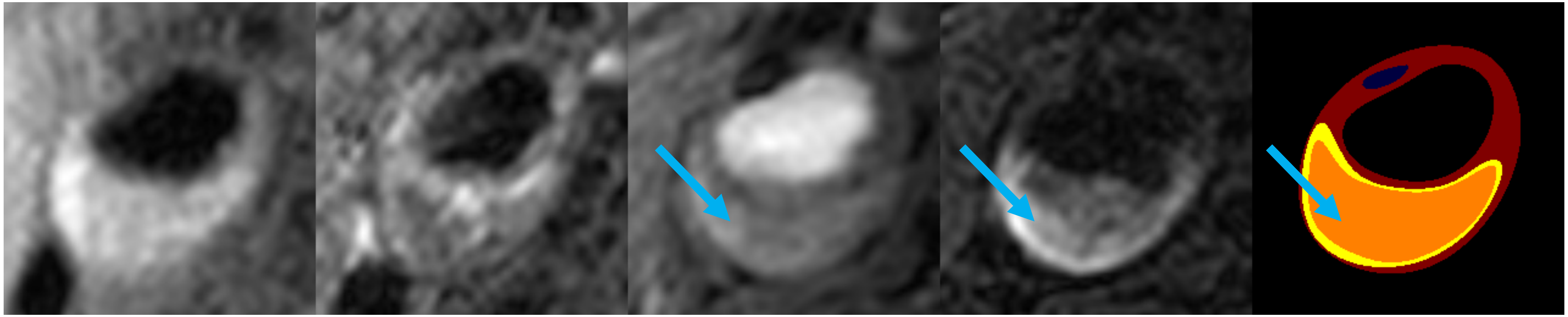
- Calcification: calcium builds up in blood vessels
- Lipid-rich/necrotic core (LR/NC): extracellular mass in the intima



T1W T2W TOF MP-RAGE Manual

When there is a plaque

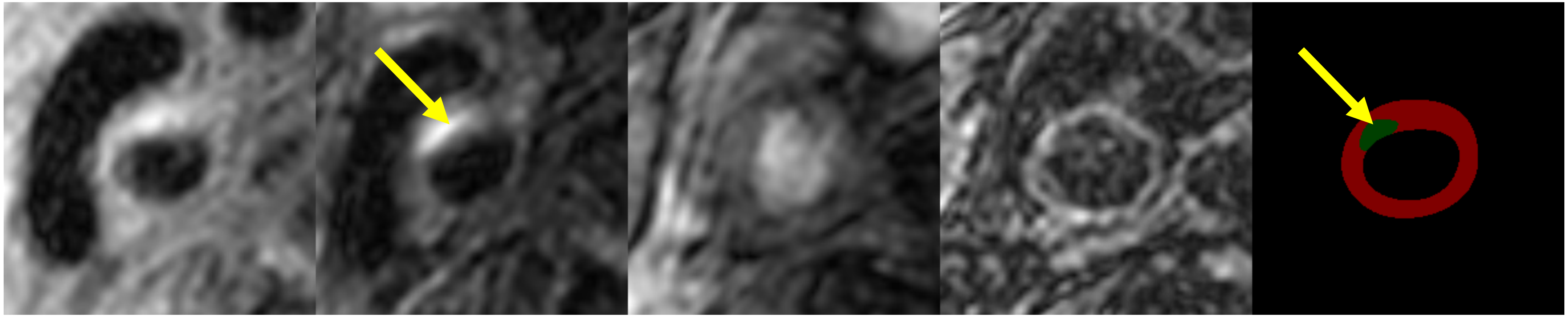
- Calcification: calcium builds up in blood vessels
- Lipid-rich/necrotic core (LR/NC): extracellular mass in the intima
- Hemorrhage: liquid plaque component



T1W T2W TOF MP-RAGE Manual

When there is a plaque

- Calcification: calcium builds up in blood vessels
- Lipid-rich/necrotic core (LR/NC): extracellular mass in the intima
- Hemorrhage: liquid plaque component
- Loose matrix: tissues that are loosely woven



Previous work requires hand-crafted features, yet not achieving usable accuracy

- MEPPS
  - Morphology-enhanced probability map
  - Intensity + morphology information
- Van *et al.*
  - Bayes classifier
  - Intensity + zero-, first and second derivatives

- Using deep learning, we can improve the performance up to 2x compared to MEPPS
- Do not need ad hoc features

Outline

- Background
- Dataset and preprocessing
- Our model
- Evaluation

Dataset: Chinese Atherosclerosis Risk Evaluation study (CARE II)

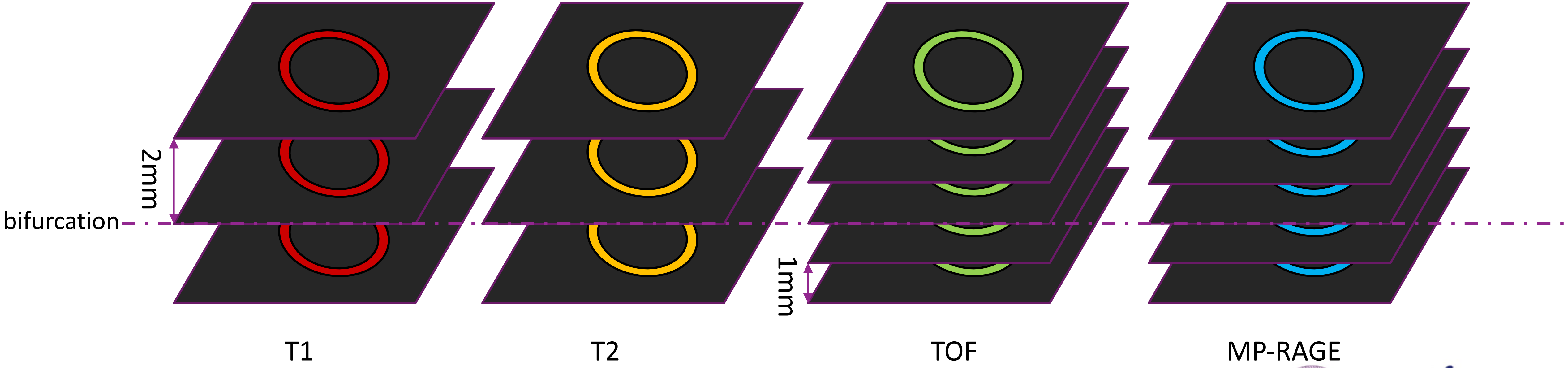
- Collected 13 medical centers and hospitals all over China
- Over 1000 patients, we used ~580, age between 18 and 80
- All patients have stroke or transient ischemic attack within two weeks after onsets of symptoms

- Professionally labeled to identify all plaques.

清華大学 Tsinghua University 交叉信息研究院 Institute for Interdisciplinary Information Sciences

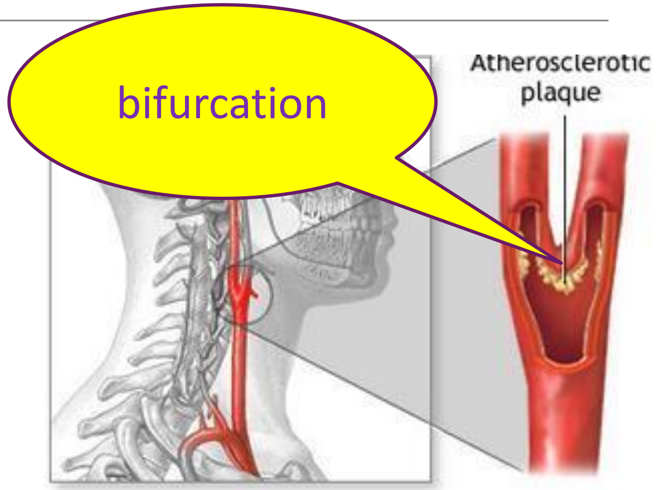
Dataset labeling: Alignment of different contrasts

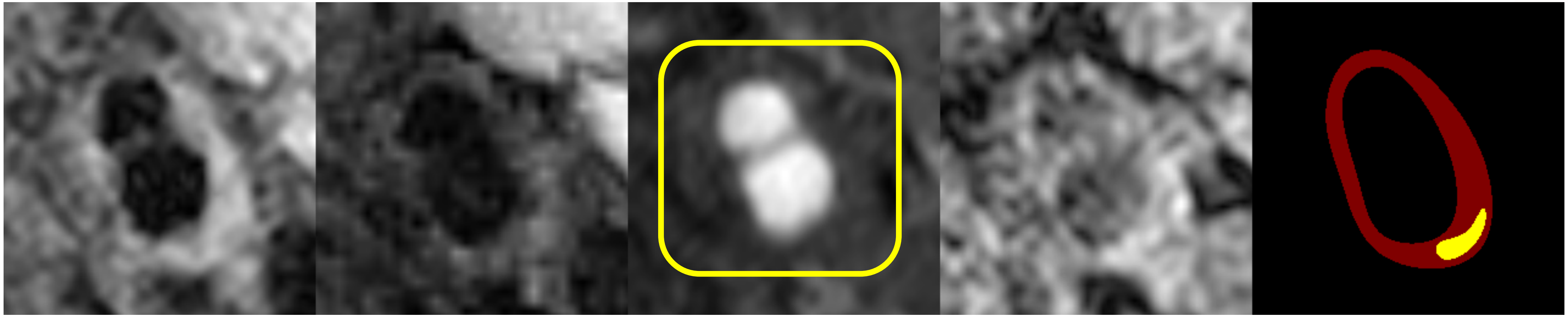
- Each case has 16 slices with 4 *contrast weightings*
- *Different slice thickness => requires an alignment*



Dataset labeling: Alignment of different contrasts

- Each case has 16 slices with 4 *contrast weightings*
- *Different slice thickness => requires an alignment*

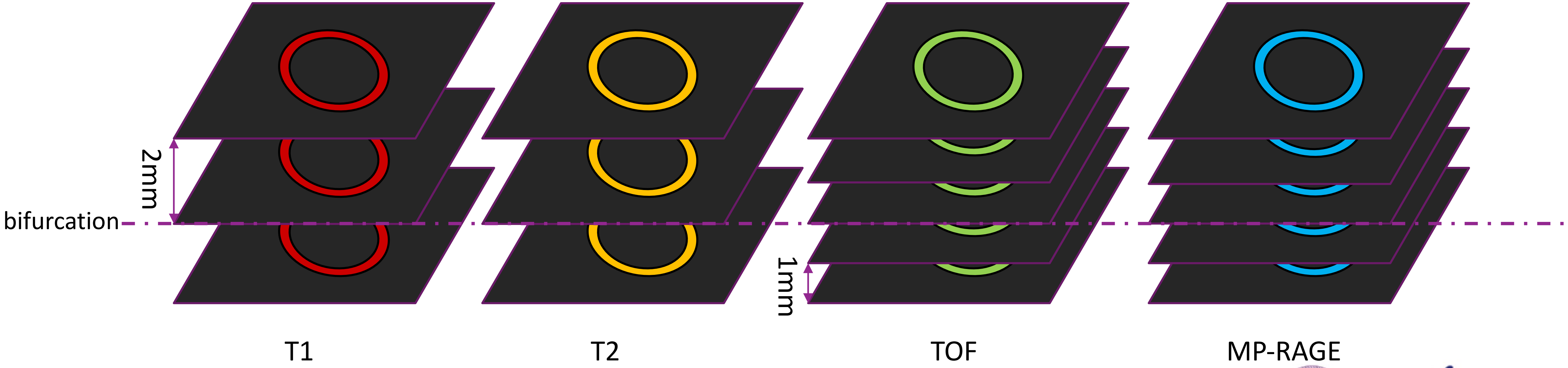




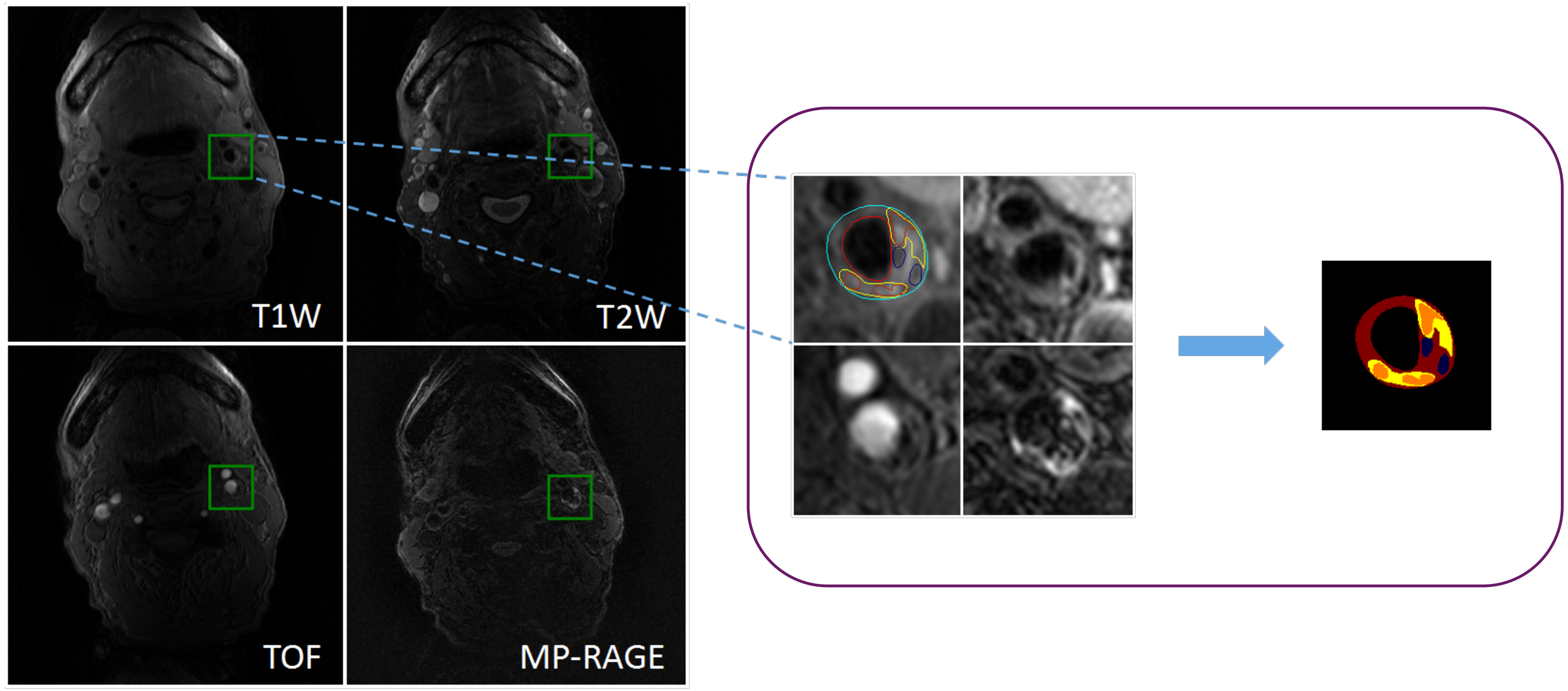
T1W T2W TOF MP-RAGE Manual

Dataset labeling: Alignment of different contrasts

- Each case has 16 slices with 4 *contrast weightings*
- *Different slice thickness => requires an alignment*



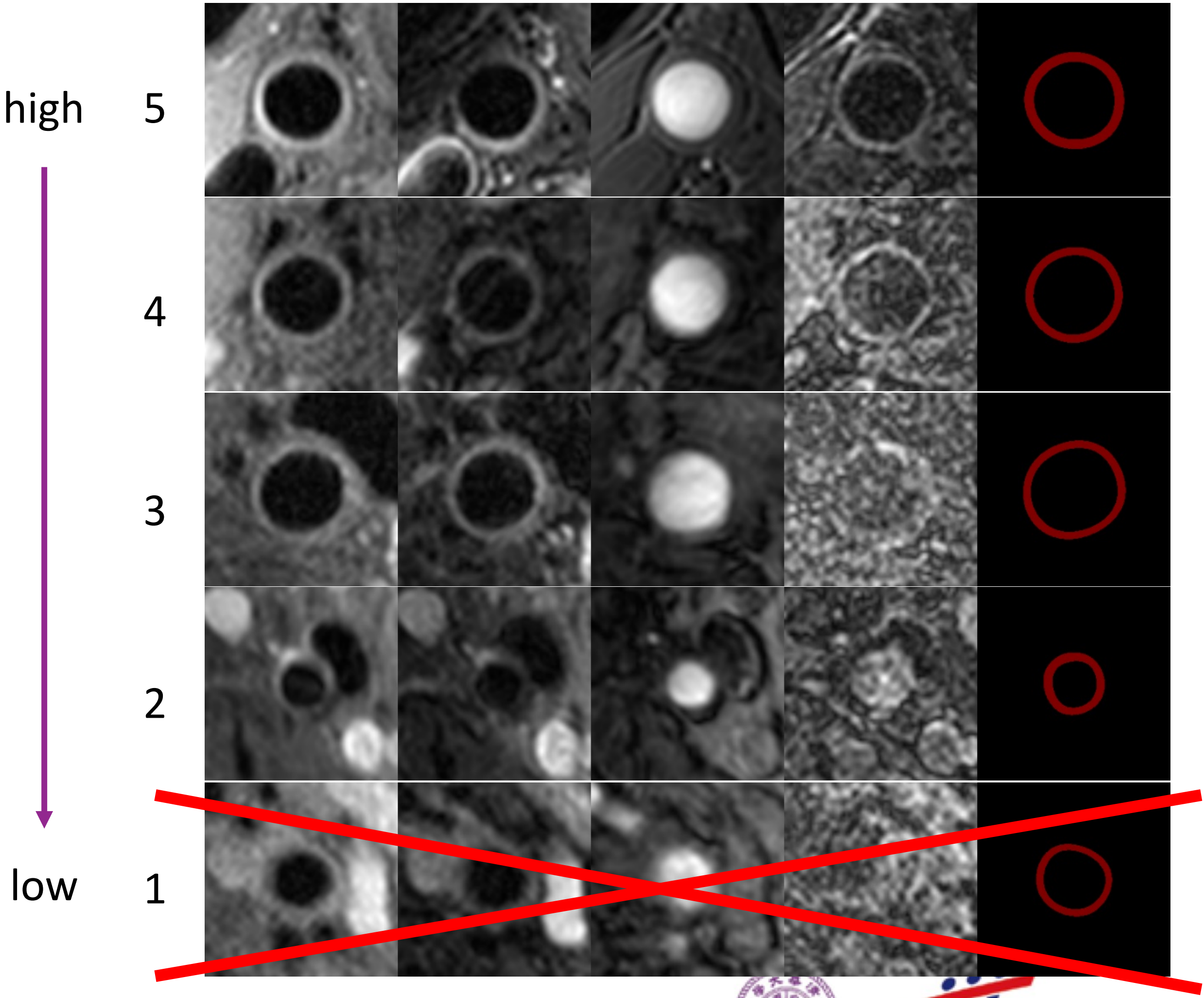
Dataset Labeling: segment all the component => pixel level labeling



(a) (b) (c)

Dataset labeling: Image quality filtering

- Reviewers provide a 5-level quality score
- We ignore the lowest quality ones



Our training / testing set selection

- We choose 1098 vessels (16 slices each), from ~580 people

- 20% test set
- 80% training + validation

Outline

- Background
- Dataset and preprocessing
- Our model
- Evaluation

Our Approach

- We use convolutional neural networks (CNN) to learn the input
  - Base models: VGG-16[1], GoogLeNet[2], ResNet-101[3]

- Key questions:
  - Still not enough training data.
    - Natural image datasets, e.g. ImageNet, 1.26 million
  - Does ImageNet pre-trained models help?
  - How to adapt the multi-contrast images to a pre-trained model?
  - Plaques is very small in the image, pretrained CNN does not offer not enough resolution

[1] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.
[2] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[J]. 2015:1-9.
[3] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. 2015:770-778.

Our Key Ideas

- We fine-tune base models pre-trained with ImageNet
- Allow inputting the 4 contrast weightings with reasonable overhead
- Maintaining high resolution by reducing the down-sampling factor from 32x to 8x

Key idea 1: Fine-tuning a Pre-trained Model

- Low-level features of pretrained model contains texture information
  - Similar for natural images vs. medical ??
- We can re-use them through fine-tuning

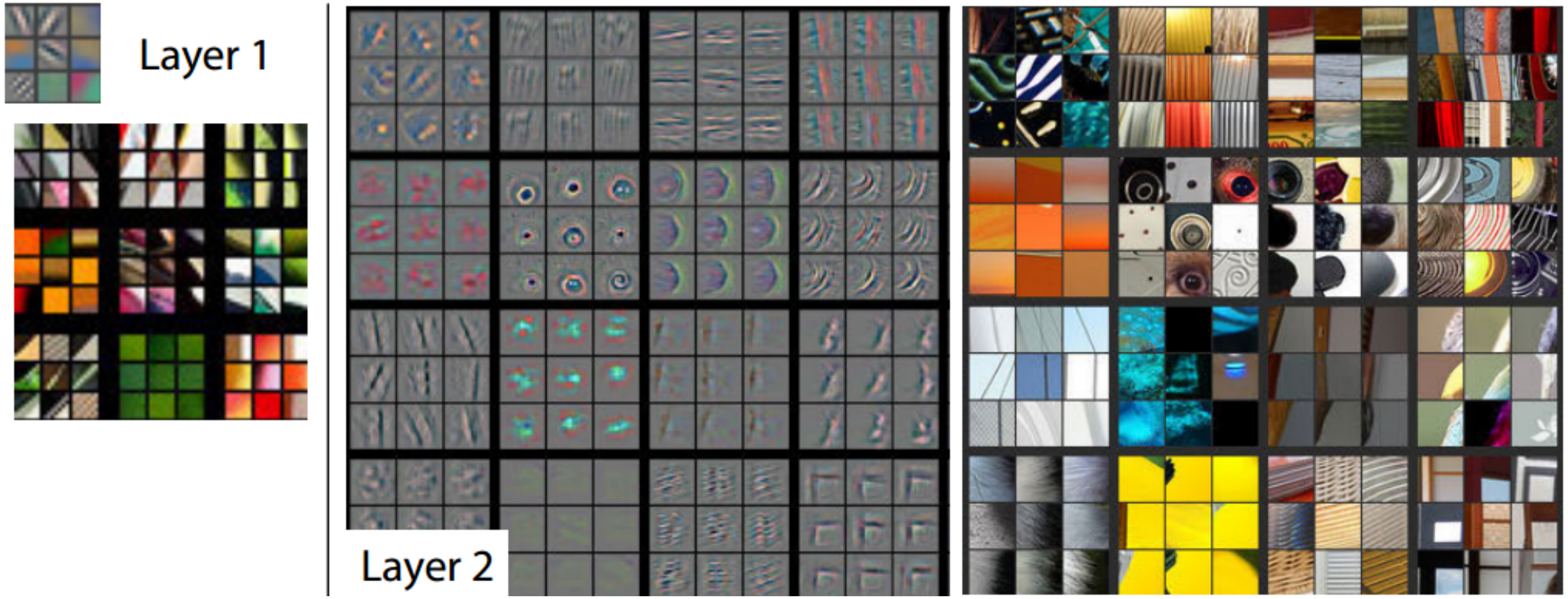
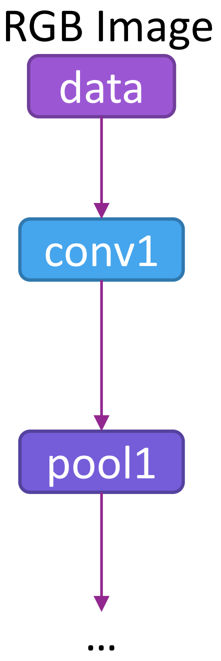


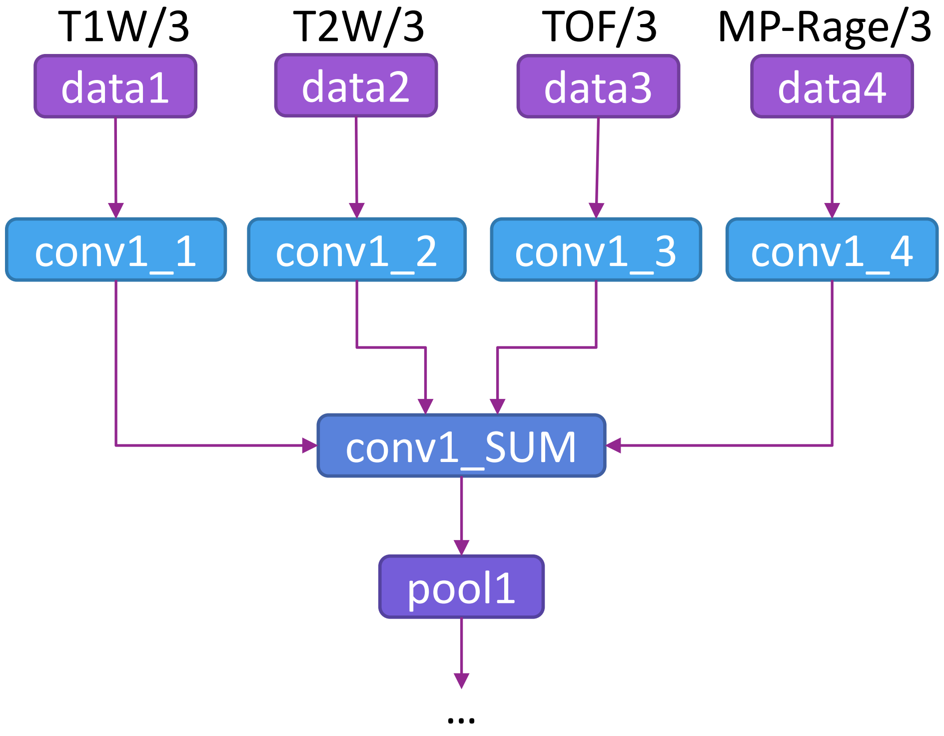
Figure from: Zeiler M D, Fergus R. Visualizing and understanding convolutional networks.

Key idea 2: Adapting multi-contrast images into the pre-trained network (VGG, GoogLeNet, ResNet…)

- Input: RGB Image (3 input channels) -> Multi-contrast MR Images (4 input channels)



(a). The original structure



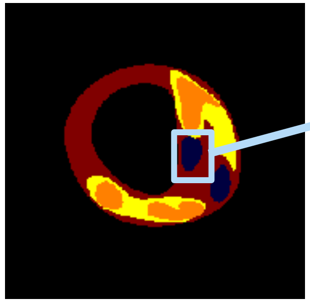
(b). Modified Structure

Key Idea 3: Maintaining high resolution

- Pretrained model has 32x reduction on input images

- Our input size is 320x320
- Plaque composition may be less than 32x32, => less than 1 pixel

- Thus: 8x reduction
- Modify two strides of 2 to 1, and add dilation kernels[1]



Less than 32x32

[1] Chen L C, Papandreou G, Kokkinos I, et al. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs[J]. Computer Science, 2014(4):357-361.

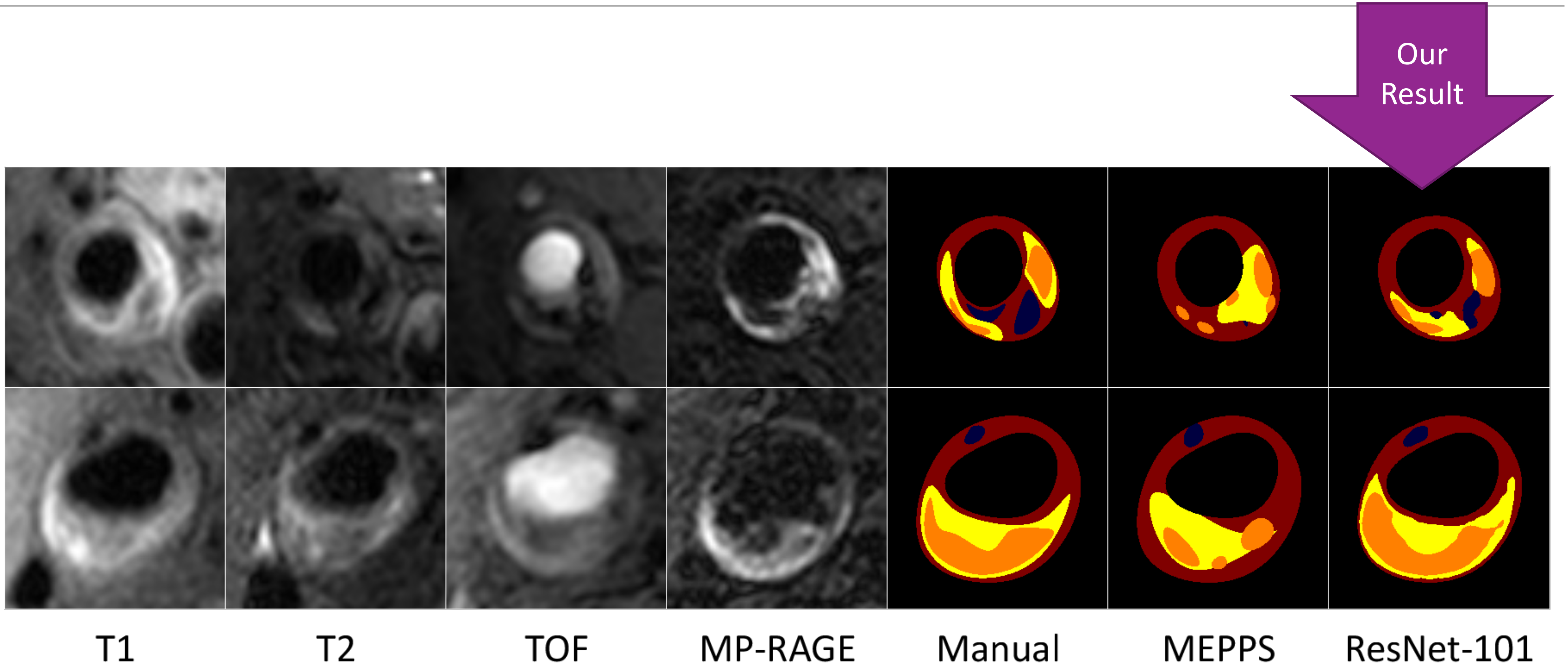
Implementation Details

- Imbalance Data
  - Many slices only have normal tissues
  - Features of normal tissues can also be learned from other slices, while learning abnormal classes simultaneously
  - Thus we throw normal slices away
- Data augmentation
  - Since there isn't any known impact of the position (left or right) of the carotid on plaques
  - We flip the image horizontally with ½ probability
  - With an input of 320*320, we randomly rescale the image to 1x~1.25x, and randomly crop 320*320 on the rescaled image and put into the net

Outline

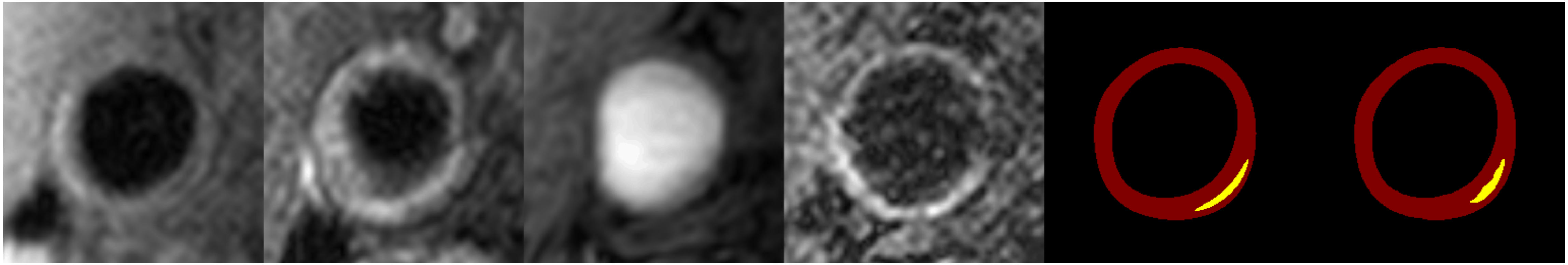
- Background
- Dataset and preprocessing
- Our model
- Evaluation

Results



Results

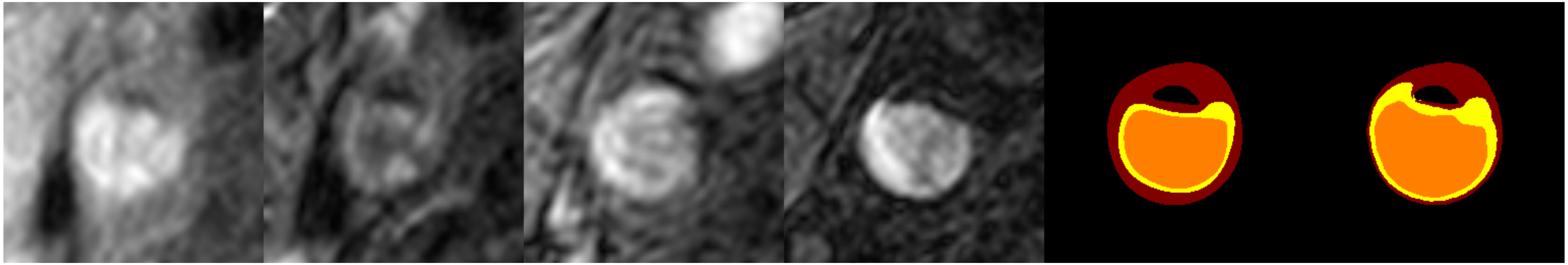
Our Result



T1W T2W TOF MP-RAGE Manual ResNet-101

Results: Metrics

- Pixel-wise accuracy
  - Recall, precision and f-score
- Pixel-wise accuracy is strict
  - Recall: 0.967; Precision: 0.789; F-score: 0.869



T1W T2W TOF MP-RAGE Manual ResNet-101

Results

- Metric
  - Precision
  - Recall
  - F-measure

Precision/Recall

	MEPPS	GoogLeNet	VGG-16	ResNet-101
Calcification	0.698/0.457	0.673/0.446	0.663/0.481	0.704/0.492
Lipid Core	0.373/0.273	0.533/0.419	0.536/0.372	0.576/0.474
Hemorrhage	0.526/0.299	0.710/0.499	0.717/0.487	0.729/0.622
Loose Matrix	0.103/0.253	0.422/0.091	0.522/0.138	0.488/0.246

F-measure

	MEPPS	GoogLeNet	VGG-16	ResNet-101
Calcification	0.552	0.536	0.557	**0.580**
Lipid Core	0.315	0.469	0.439	**0.520**
Hemorrhage	0.382	0.586	0.580	**0.671**
Loose Matrix	0.146	0.150	0.218	**0.327**

Results: Comparing to MEPPS

- Metric
  - Precision
  - Recall
  - F-measure

Precision/Recall

	MEPPS	GoogLeNet	VGG-16	ResNet-101
Calcification	0.698/0.457	0.673/0.446	0.663/0.481	0.704/0.492
Lipid Core	0.373/0.273	0.533/0.419	0.536/0.372	0.576/0.474
Hemorrhage	0.526/0.299	0.710/0.499	0.717/0.487	0.729/0.622
Loose Matrix	0.103/0.253	0.422/0.091	0.522/0.138	0.488/0.246

F-measure

	MEPPS	GoogLeNet	VGG-16	ResNet-101
Calcification	0.552	0.536	0.557	**0.580**
Lipid Core	0.315	0.469	0.439	**0.520**
Hemorrhage	0.382	0.586	0.580	**0.671**
Loose Matrix	0.146	0.150	0.218	**0.327**

清華大学 Tsinghua University
交叉信息研究院 Institute for Interdisciplinary Information Sciences

Results: Different CNNs

- Metric
  - Precision
  - Recall
  - F-measure

Precision/Recall

	MEPPS	GoogLeNet	VGG-16	ResNet-101
Calcification	0.698/0.457	0.673/0.446	0.663/0.481	0.704/0.492
Lipid Core	0.373/0.273	0.533/0.419	0.536/0.372	0.576/0.474
Hemorrhage	0.526/0.299	0.710/0.499	0.717/0.487	0.729/0.622
Loose Matrix	0.103/0.253	0.422/0.091	0.522/0.138	0.488/0.246

F-measure

	MEPPS	GoogLeNet	VGG-16	ResNet-101
Calcification	0.552	0.536	0.557	**0.580**
Lipid Core	0.315	0.469	0.439	**0.520**
Hemorrhage	0.382	0.586	0.580	**0.671**
Loose Matrix	0.146	0.150	0.218	**0.327**

Results: Different accuracy on different compositions

- Metric
  - Precision
  - Recall
  - F-measure

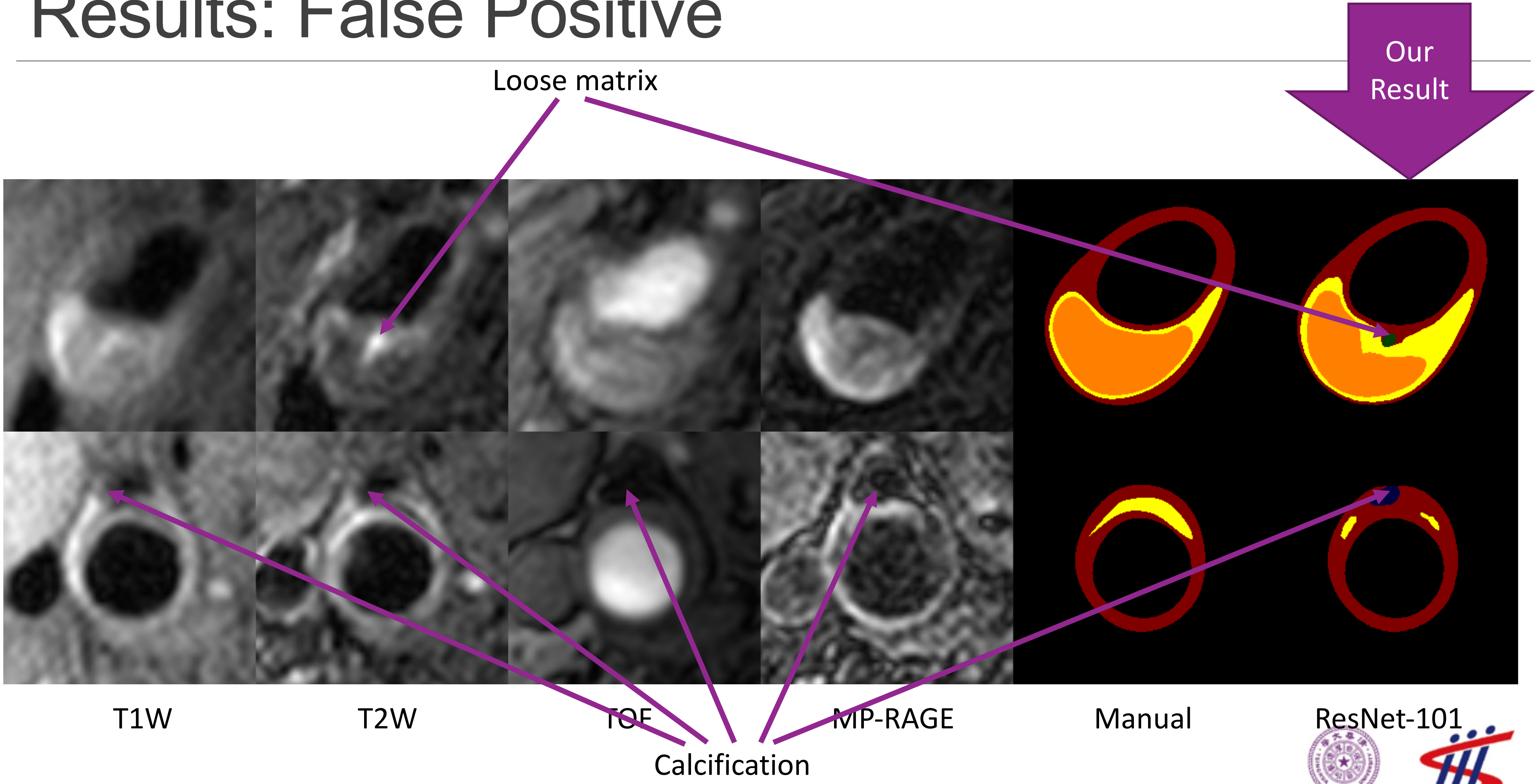
Precision/Recall

	MEPPS	GoogLeNet	VGG-16	ResNet-101
Calcification	0.698/0.457	0.673/0.446	0.663/0.481	0.704/0.492
Lipid Core	0.373/0.273	0.533/0.419	0.536/0.372	0.576/0.474
Hemorrhage	0.526/0.299	0.710/0.499	0.717/0.487	0.729/0.622
Loose Matrix	0.103/0.253	0.422/0.091	0.522/0.138	0.488/0.246

F-measure

	MEPPS	GoogLeNet	VGG-16	ResNet-101
Calcification	0.552	0.536	0.557	**0.580**
Lipid Core	0.315	0.469	0.439	**0.520**
Hemorrhage	0.382	0.586	0.580	**0.671**
Loose Matrix	0.146	0.150	0.218	**0.327**

Results: False Positive



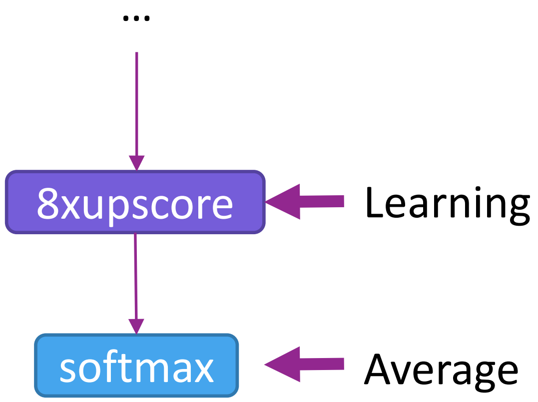
Contributions of Each Contrast Weighting

- Use each contrast weighting to train separate models
- F-measure of each tissue class

Contrast Weighting	Calcification	Lipid/Necrotic Core	Hemorrhage	Loose Matrix
T1W	**0.538**	0.496	0.443	0.020
T2W	0.494	**0.515**	0.323	**0.387**
TOF	0.468	0.465	0.487	0.080
MP-RAGE	0.337	0.437	**0.681**	0.015
ALL	0.580	0.520	0.671	0.327

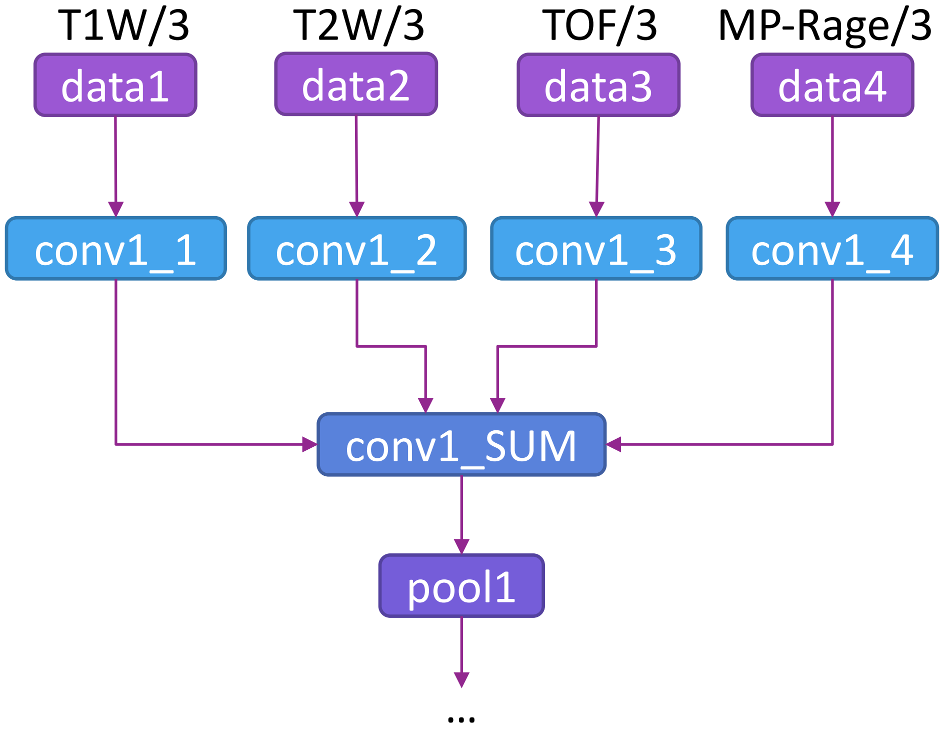
Model Ensemble

- Average: average over the softmax layer of four models
- Learning: learn the weights of feature maps of upscore layer

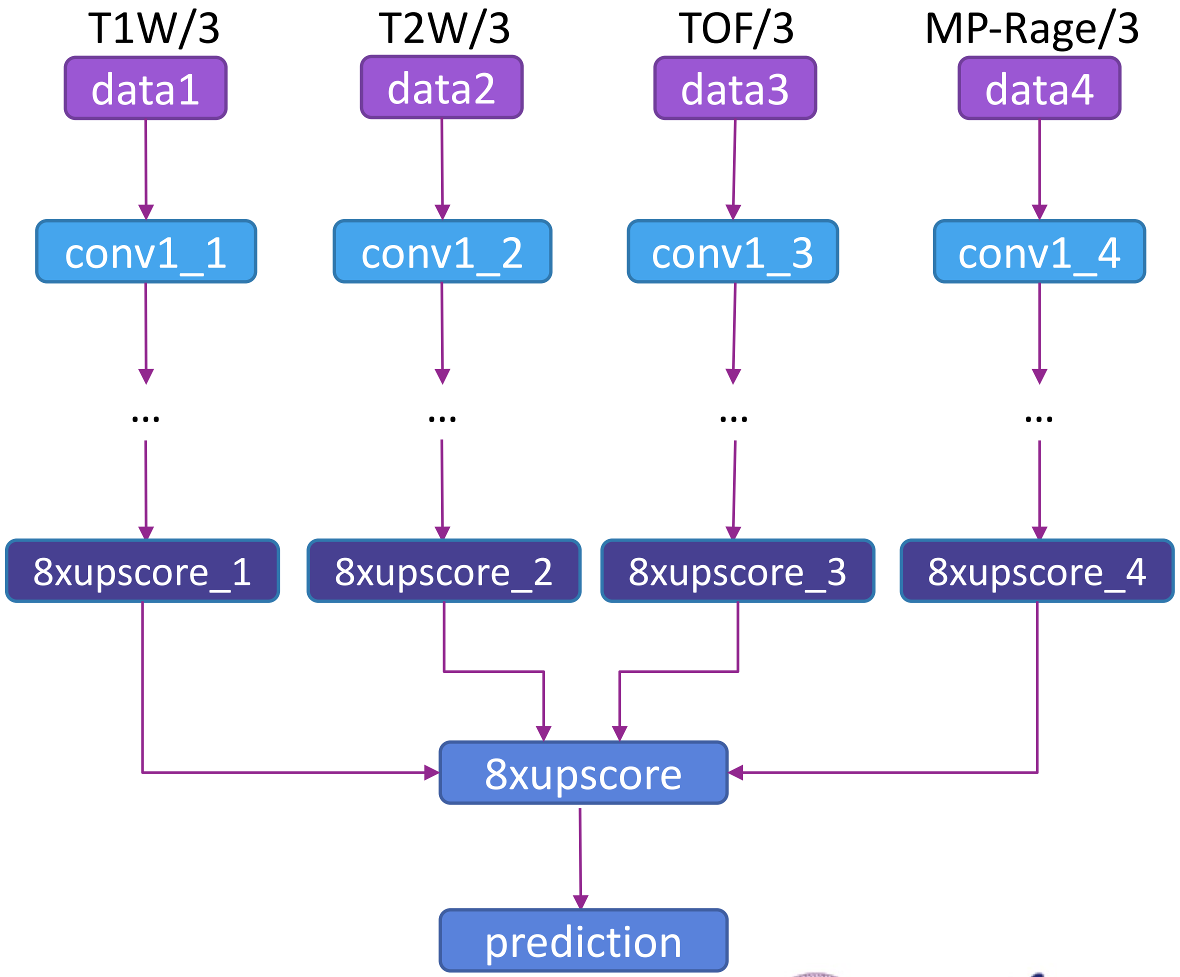


Model	**Fibrous Tissue**	**Calcification**	**Lipid/Necrotic Core**	**Hemorrhage**	**Loose Matrix**
Average	0.963	0.518	0.522	0.608	0.009
Learning	**0.963**	**0.585**	**0.557**	**0.691**	**0.335**
ResNet-101	0.962	0.580	0.520	0.671	0.327

Model Ensemble: Learning



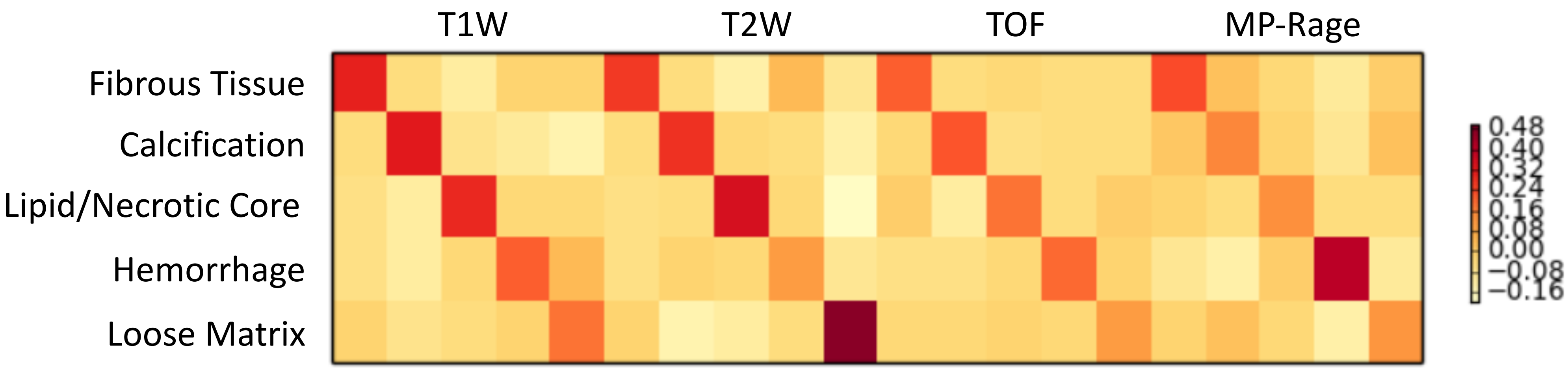
(a). Aggregate early



(b). Aggregate later

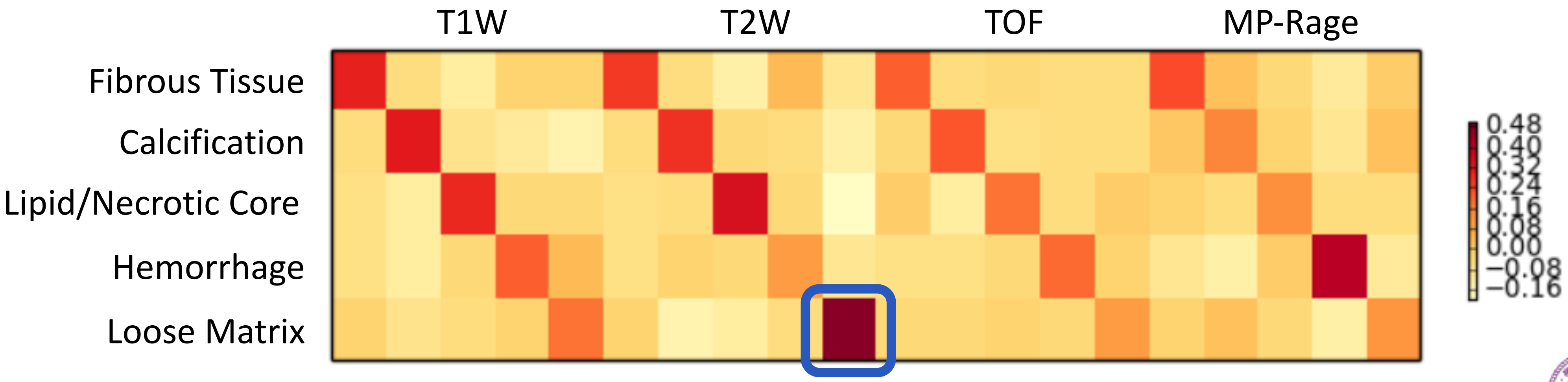
Model Ensemble: weights of each feature map

- 4 models trained with 4 channels separately
- 5 score maps for each contrast weighting in each model
- 20 feature maps



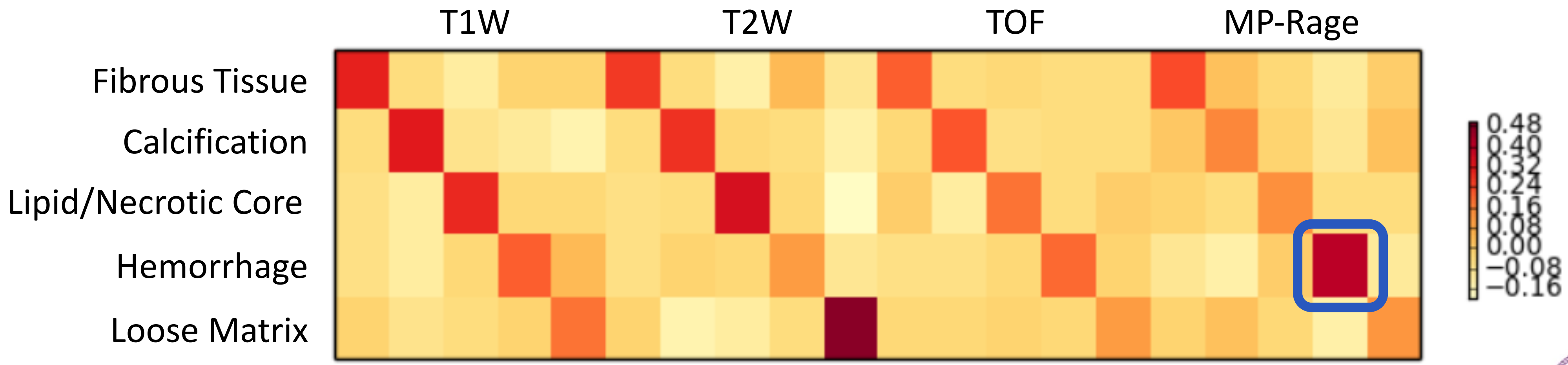
Model Ensemble: weights of each feature map

Contrast Weighting	Calcification	Lipid/Necrotic Core	Hemorrhage	Loose Matrix
T1W	**0.538**	0.496	0.443	0.020
T2W	0.494	**0.515**	0.323	**0.387**
TOF	0.468	0.465	0.487	0.080
MP-RAGE	0.337	0.437	**0.681**	0.015
ALL	0.580	0.520	0.671	0.327



Model Ensemble: weights of each feature map

Contrast Weighting	Calcification	Lipid/Necrotic Core	Hemorrhage	Loose Matrix
T1W	**0.538**	0.496	0.443	0.020
T2W	0.494	**0.515**	0.323	**0.387**
TOF	0.468	0.465	0.487	0.080
MP-RAGE	0.337	0.437	**0.681**	0.015
ALL	0.580	0.520	0.671	0.327



Confusion Matrix

TP	FP
FN	TN
sensitivity	specificity

MEPPS

Calcification

202	14
139	3085
0.592	0.995

Lipid Core

306	69
379	2686
0.447	0.975

Hemorrhage

81	98
19	3242
0.810	0.971

Loose Matrix

72	634
85	2649
0.459	0.807

ResNet

Calcification

270	42
71	3057
0.792	0.986

Lipid Core

567	169
118	2586
0.827	0.939

Hemorrhage

84	34
16	3306
0.840	0.990

Loose Matrix

83	91
74	3192
0.529	0.972

Results: Running time

- Takes ~11s on a Titan X GPU
- For whole-slice (16 slices) prediction

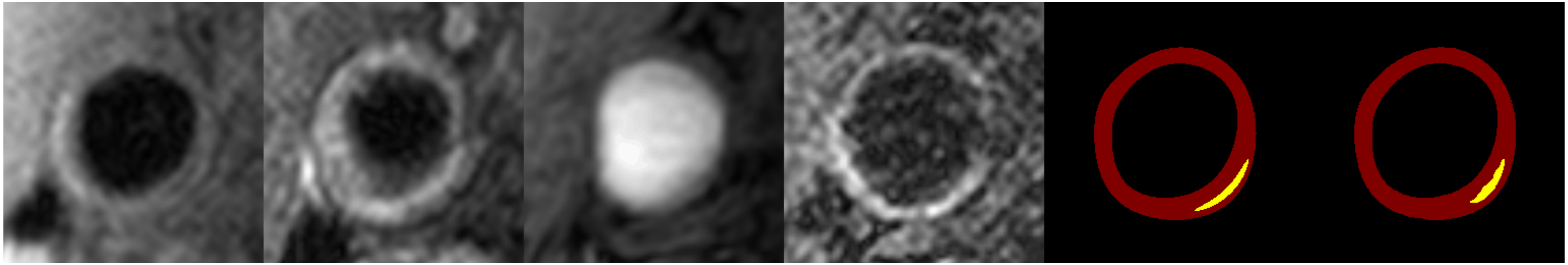
	MEPPS	GoogLeNet	VGG-16	ResNet-101
Time (sec)	10.0	9.1	8.9	11.4

Conclusion

- We apply CNNs to automatically recognize carotid plaque components
- Modify the network to receive multi-contrast input
- Lower the down sampling ratio to maintain high resolution
- CNNs achieve better accuracy than traditional Bayesian methods while running in acceptable time

Final Remarks

- CNNs can replace many traditional methods in medical image processing
- Key challenge: **labeled** data
- esp. high quality label for CNN training ≠ medical report
- Future direction: reducing the labeling requirement, transfer? Active learning?



T1W | T2W | TOF | MP-RAGE | Manual | ResNet-101

Q&A