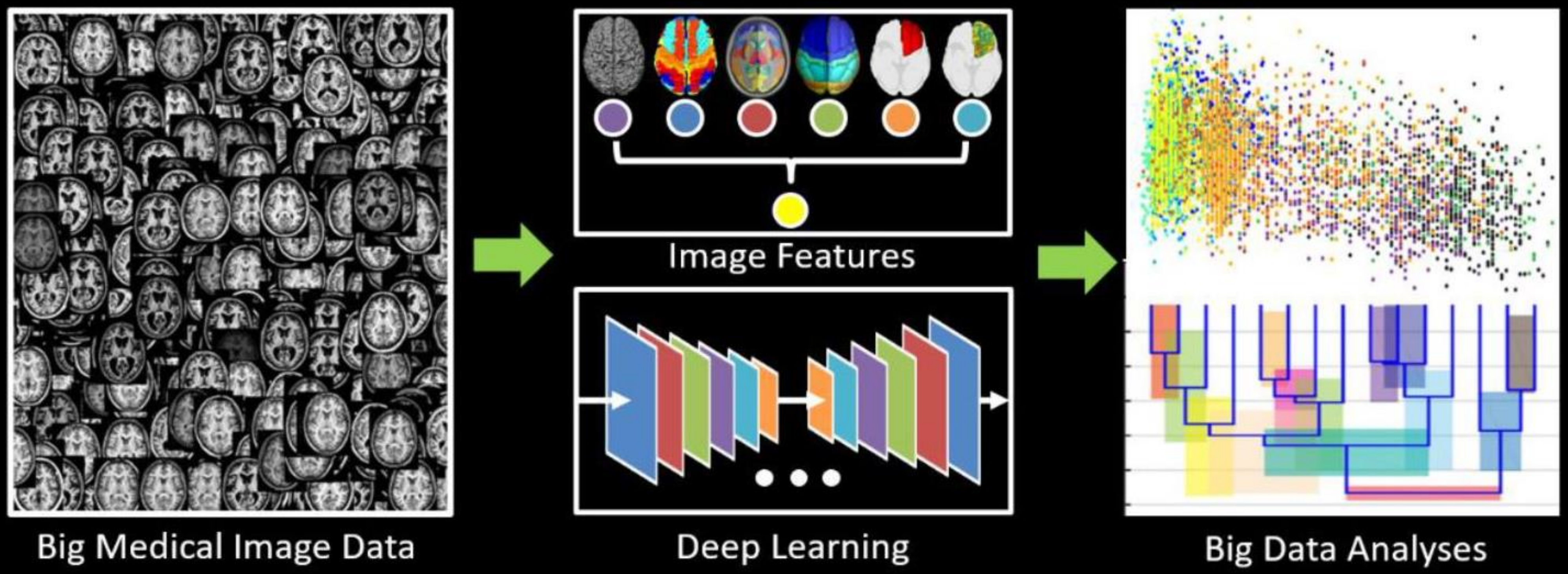


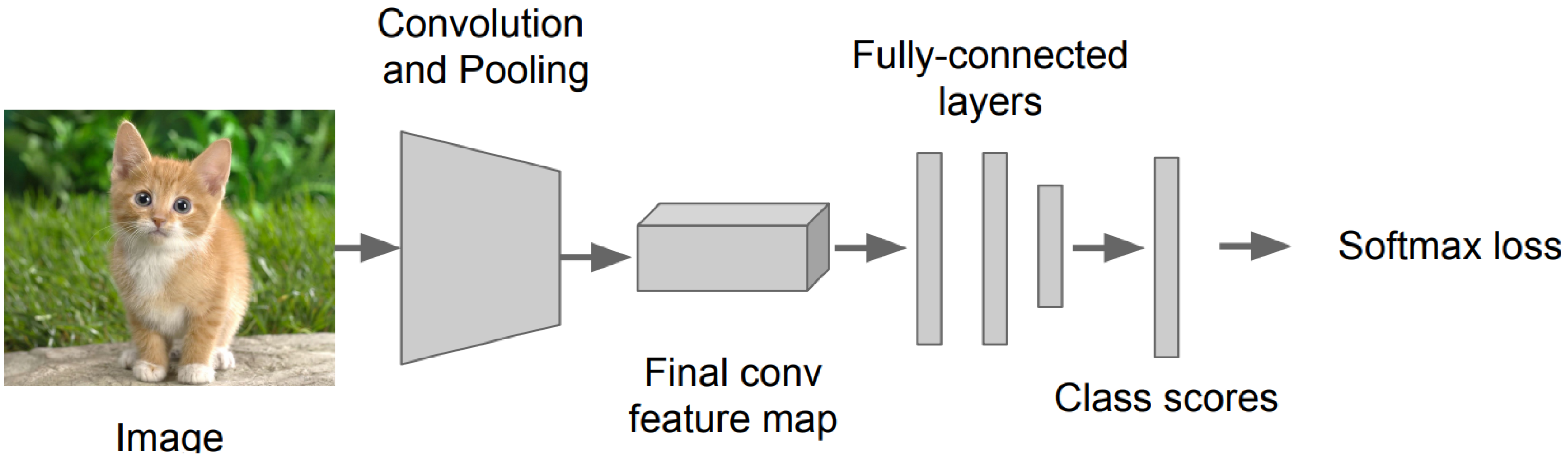
Co-learning and Self-supervised Learning

[Spring 2020 CS-8395 Deep Learning in Medical Image Computing]

Instructor: Yuankai Huo, Ph.D.
Department of Electrical Engineering and Computer Science
Vanderbilt University

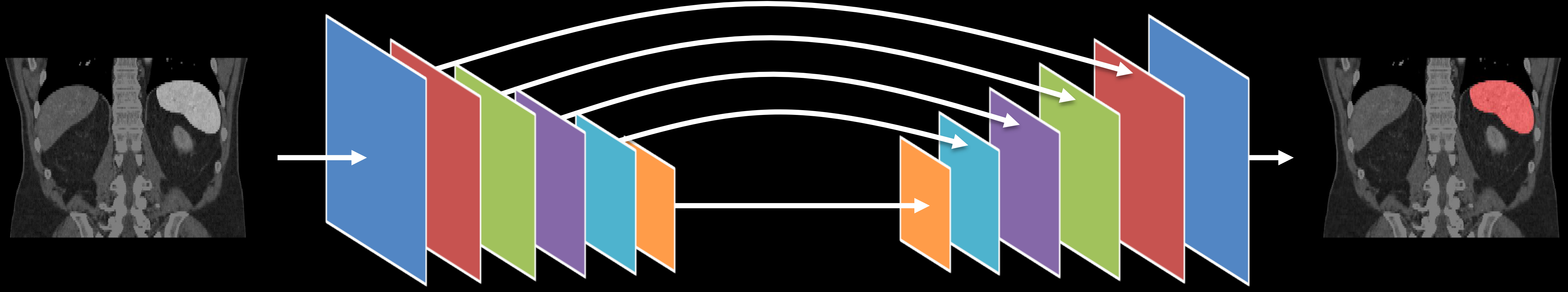
Simple Recipe for Classification

"CONV-POOL-FC"



http://cs231n.stanford.edu/slides/2016/winter1516_lecture8.pdf

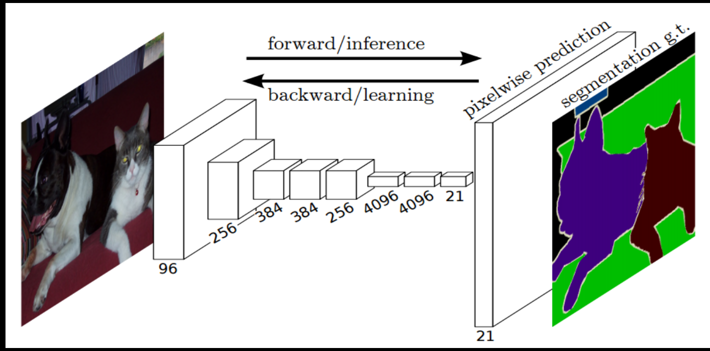
Fully Convolutional Network (FCN)



Encoder

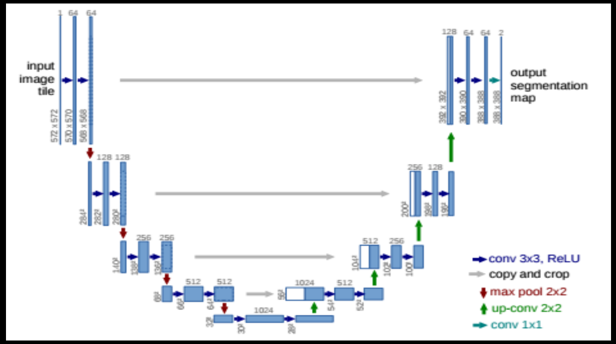
Decoder

FCN:



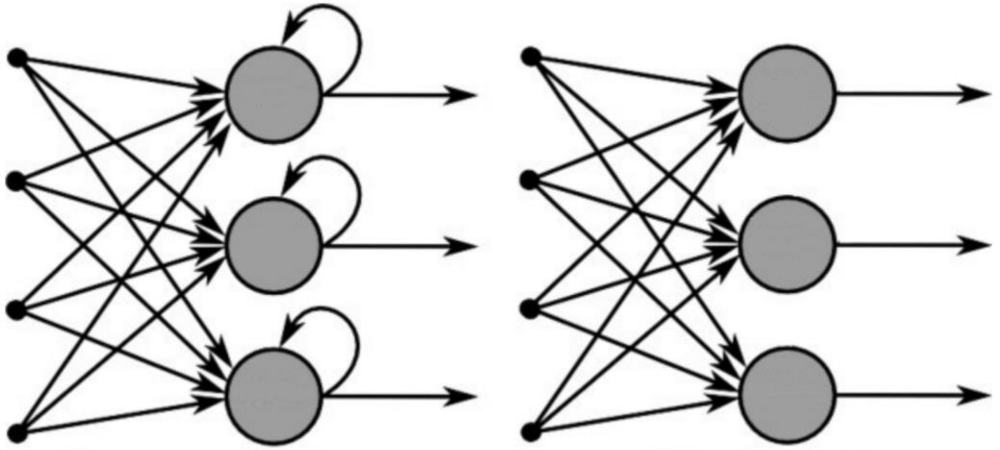
Long, et al. CVPR (2015)

Unet:

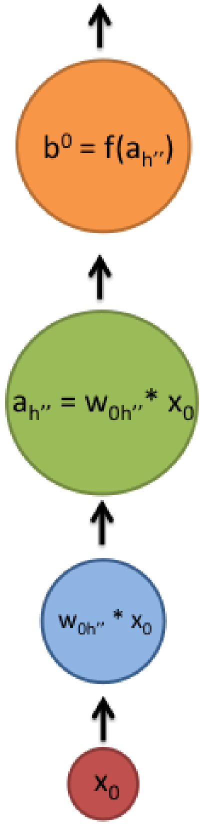


Ronneberger, et al. MICCAI (2015)

Feed-forward and Recurrent Network



Recurrent Neural Network   Feed-Forward Neural Network



https://towardsdatascience.com/recurrent-neural-networks-and-lstm-4b601dd822a5   https://www.leiphone.com/news/201608/syAwLNx4bGPuFYI1.html   https://slideplayer.com/slide/3383596/

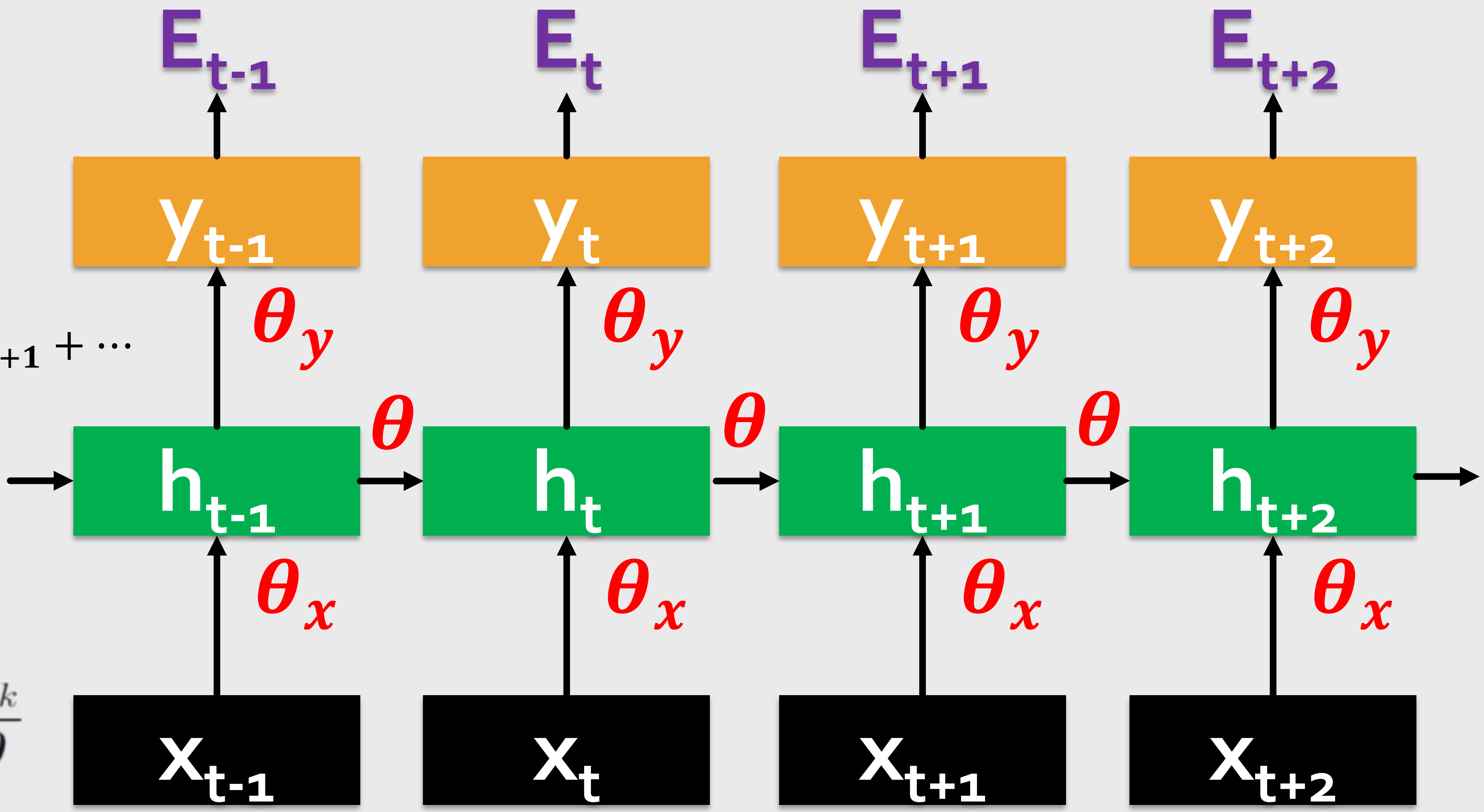
RNN

$$\mathbf{h}_t = \boldsymbol{\theta}\phi(\mathbf{h}_{t-1}) + \boldsymbol{\theta}_x \mathbf{x}_t$$
$$\mathbf{y}_t = \boldsymbol{\theta}_y \phi(\mathbf{h}_t)$$

$$\mathbf{E} = \ldots + \mathbf{E}_{\mathbf{t}-\mathbf{1}} + \mathbf{E}_{\mathbf{t}} + \mathbf{E}_{\mathbf{t}+\mathbf{1}} + \mathbf{E}_{\mathbf{t}+\mathbf{1}} + \cdots$$

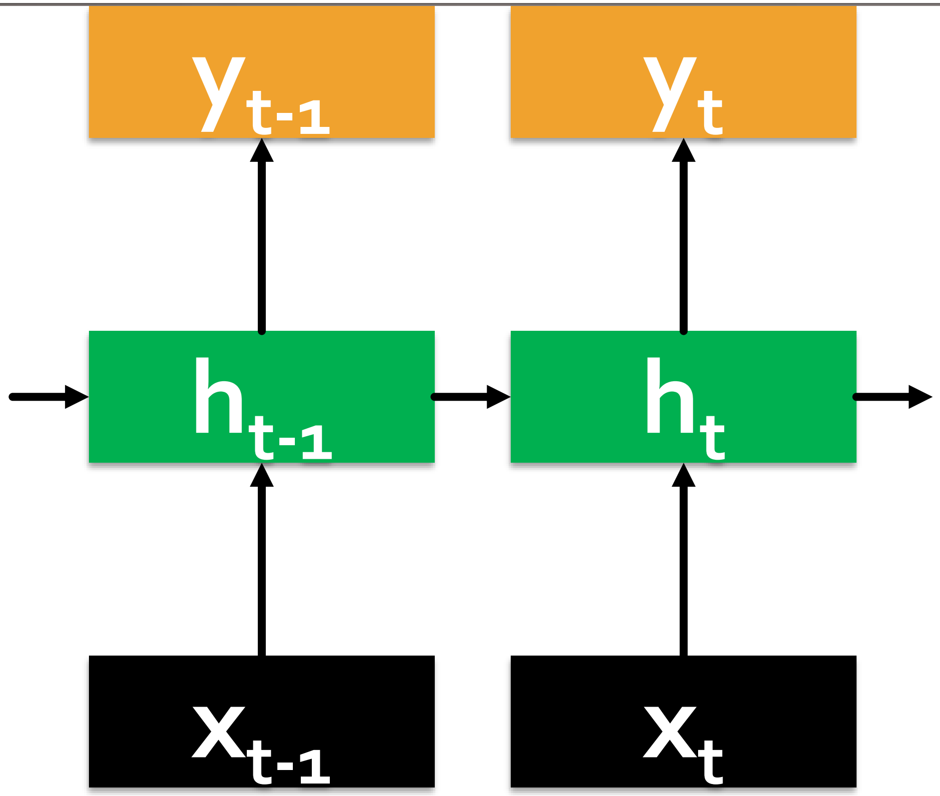
$$\frac{\partial E}{\partial \theta} = \sum_{t=1}^{S} \frac{\partial E_t}{\partial \theta}$$

$$\frac{\partial E_t}{\partial \theta} = \sum_{k=1}^{t} \frac{\partial E_t}{\partial \mathbf{y}_t} \frac{\partial \mathbf{y}_t}{\partial \mathbf{h}_t} \frac{\partial \mathbf{h}_t}{\partial \mathbf{h}_k} \frac{\partial \mathbf{h}_k}{\partial \theta}$$



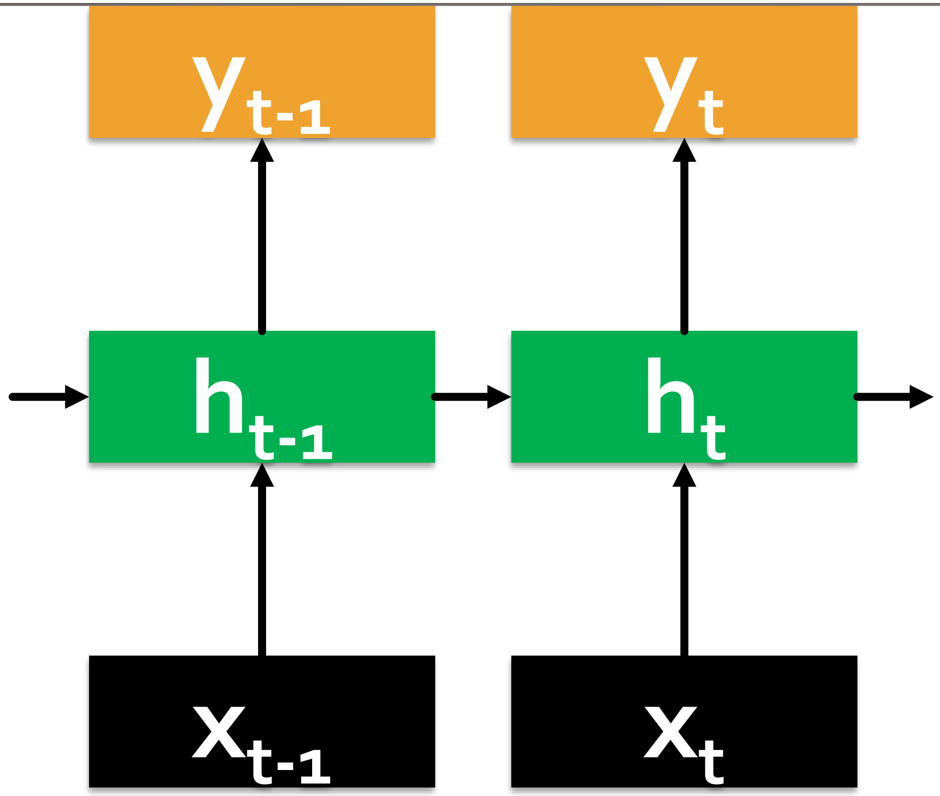
https://www.cs.ox.ac.uk/people/nando.defreitas/machinelearning/

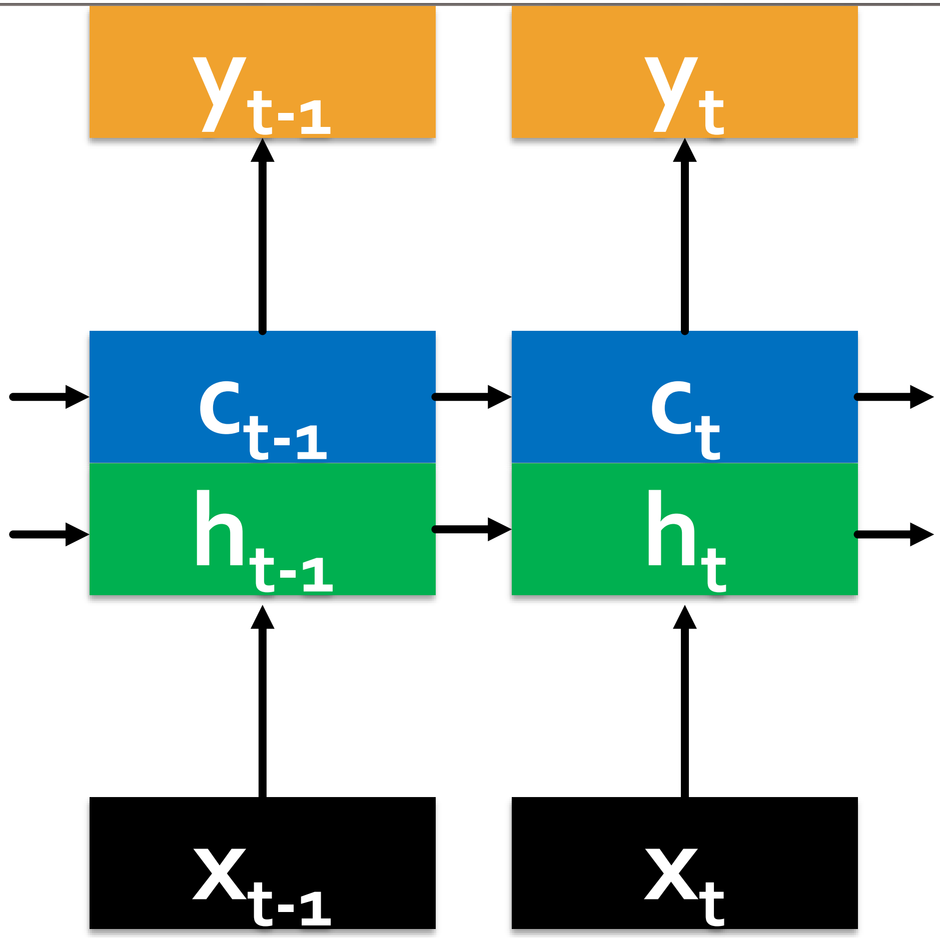
RNN



http://colah.github.io/posts/2015-08-Understanding-LSTMs/

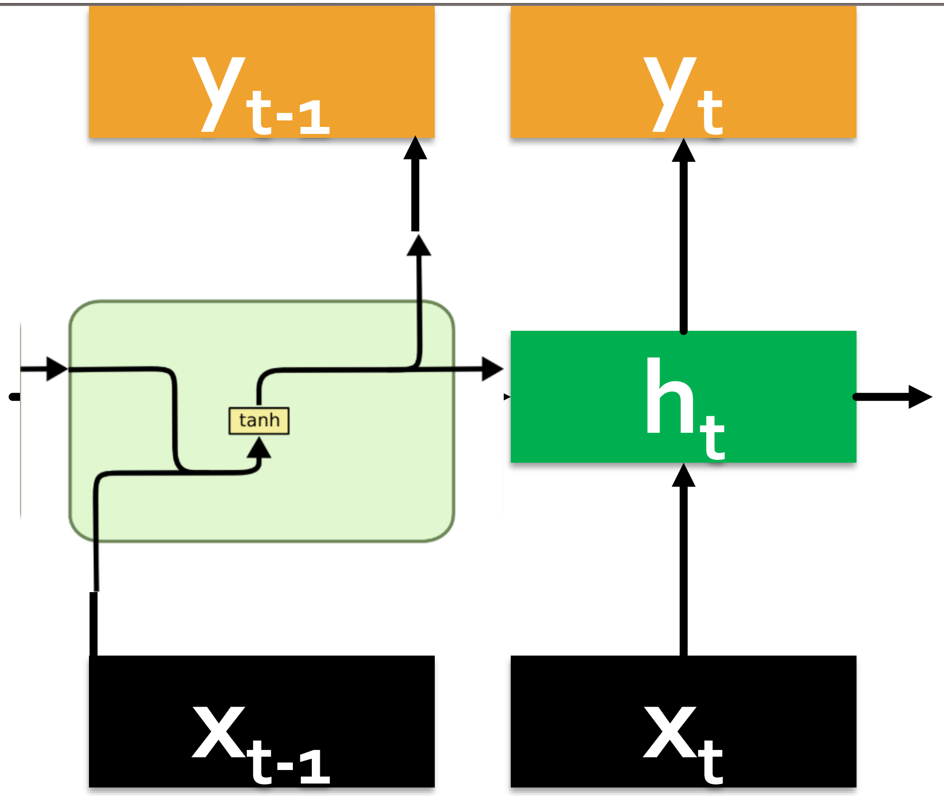
LSTM

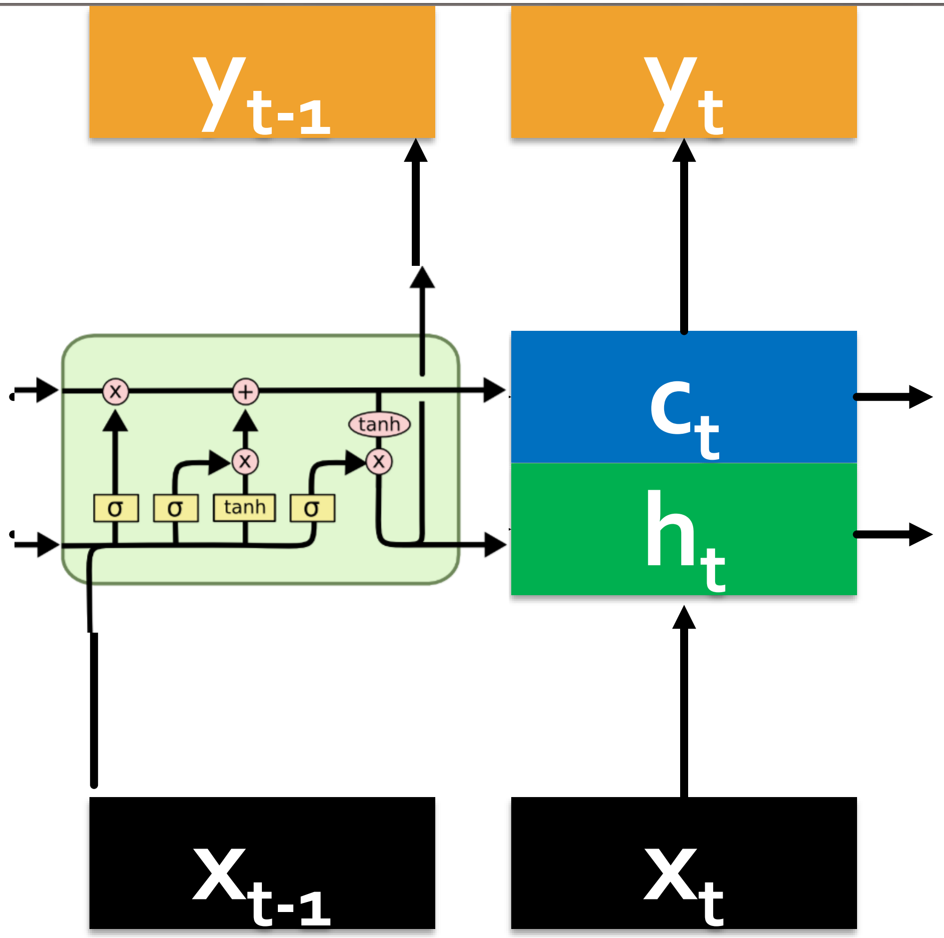




http://colah.github.io/posts/2015-08-Understanding-LSTMs/

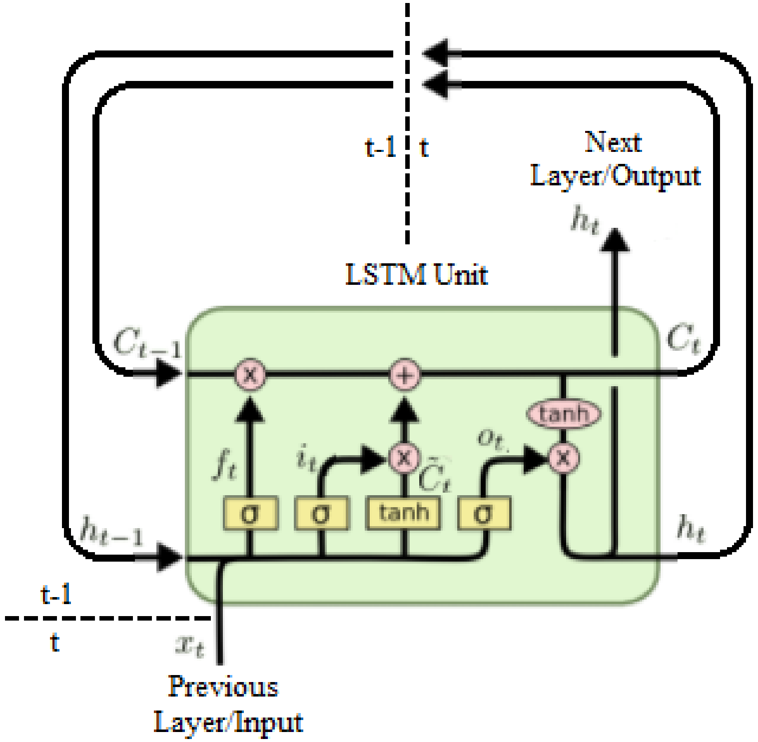
LSTM



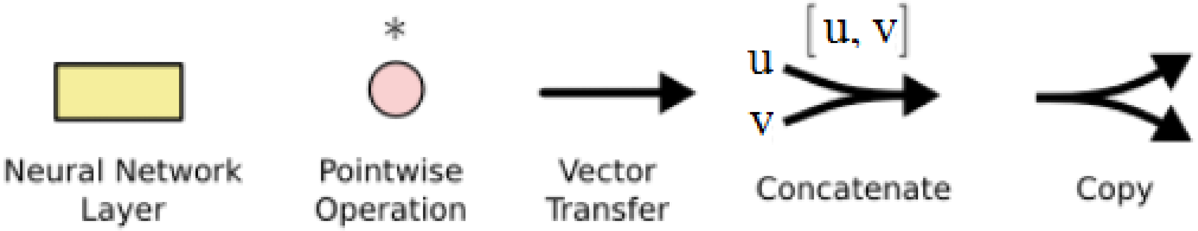


http://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTM

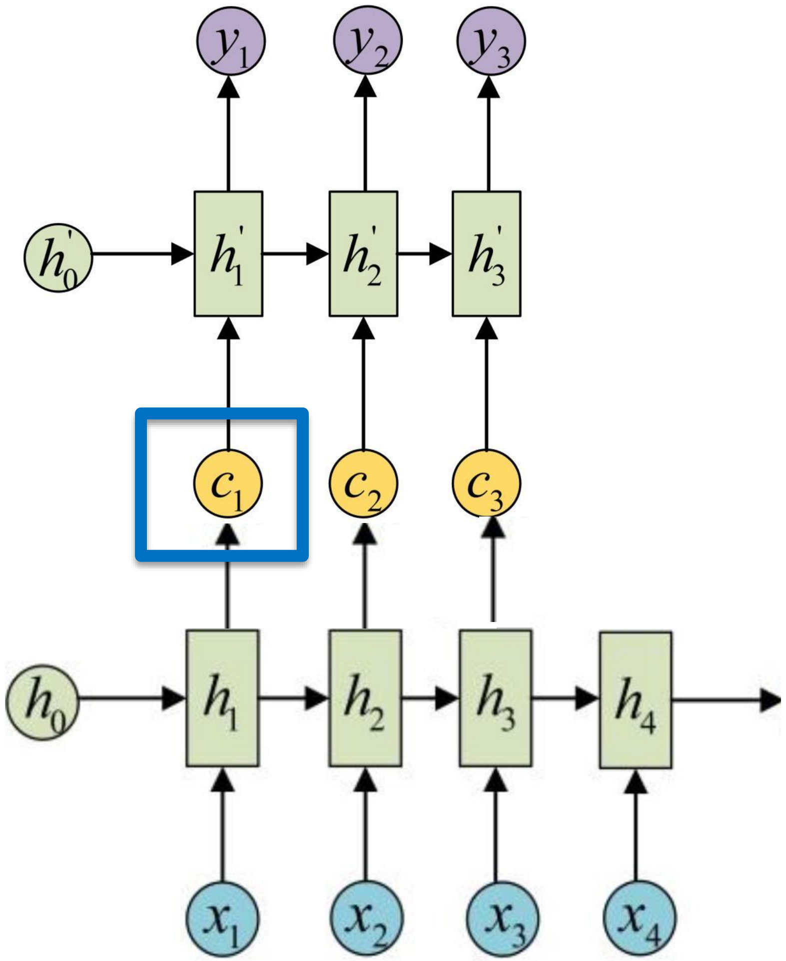


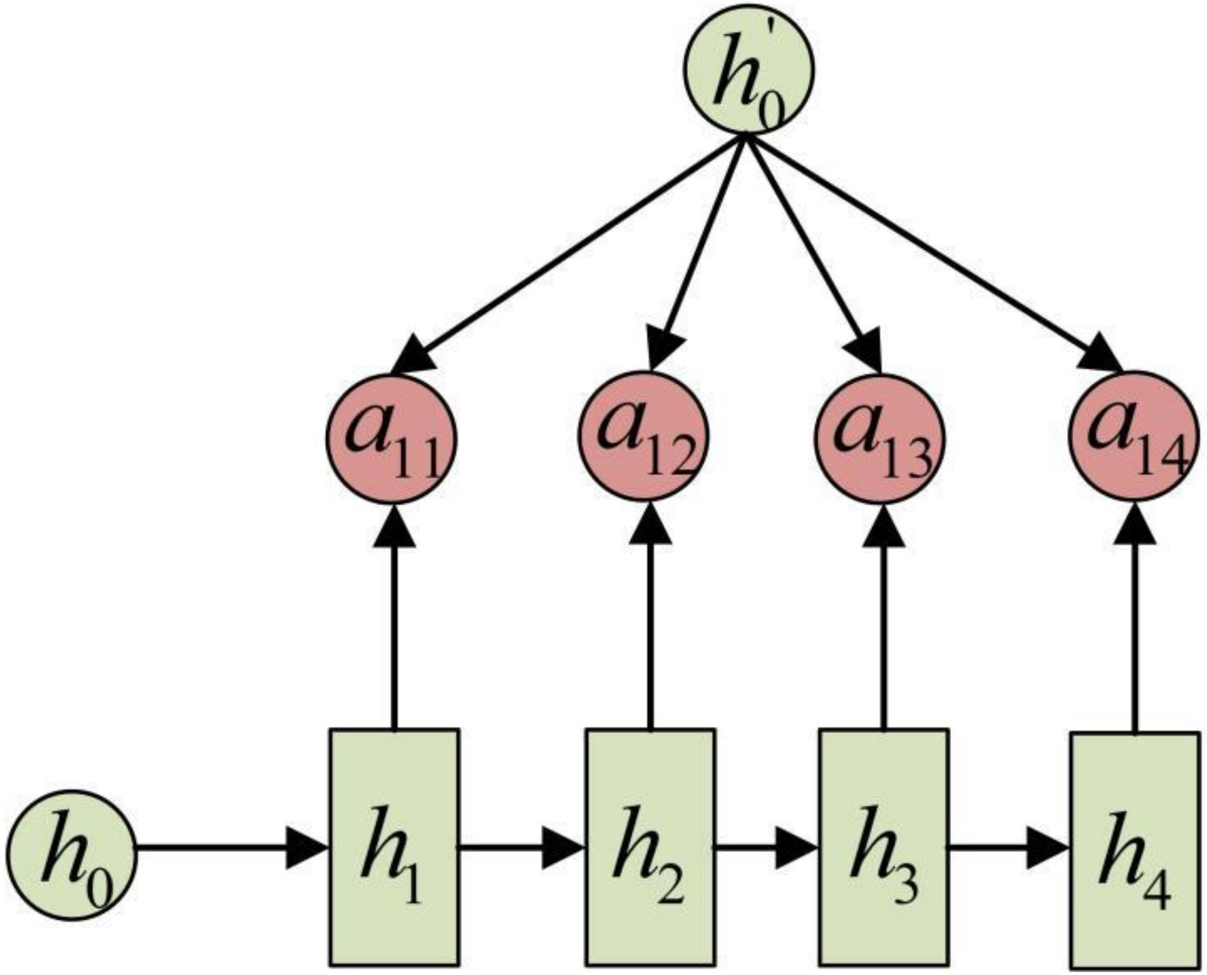
$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \ + \ b_f\right)$$
$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \ + \ b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \ + \ b_C)$$
$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$
$$o_t = \sigma\left(W_o \ [h_{t-1}, x_t] \ + \ b_o\right)$$
$$h_t = o_t * \tanh\left(C_t\right)$$



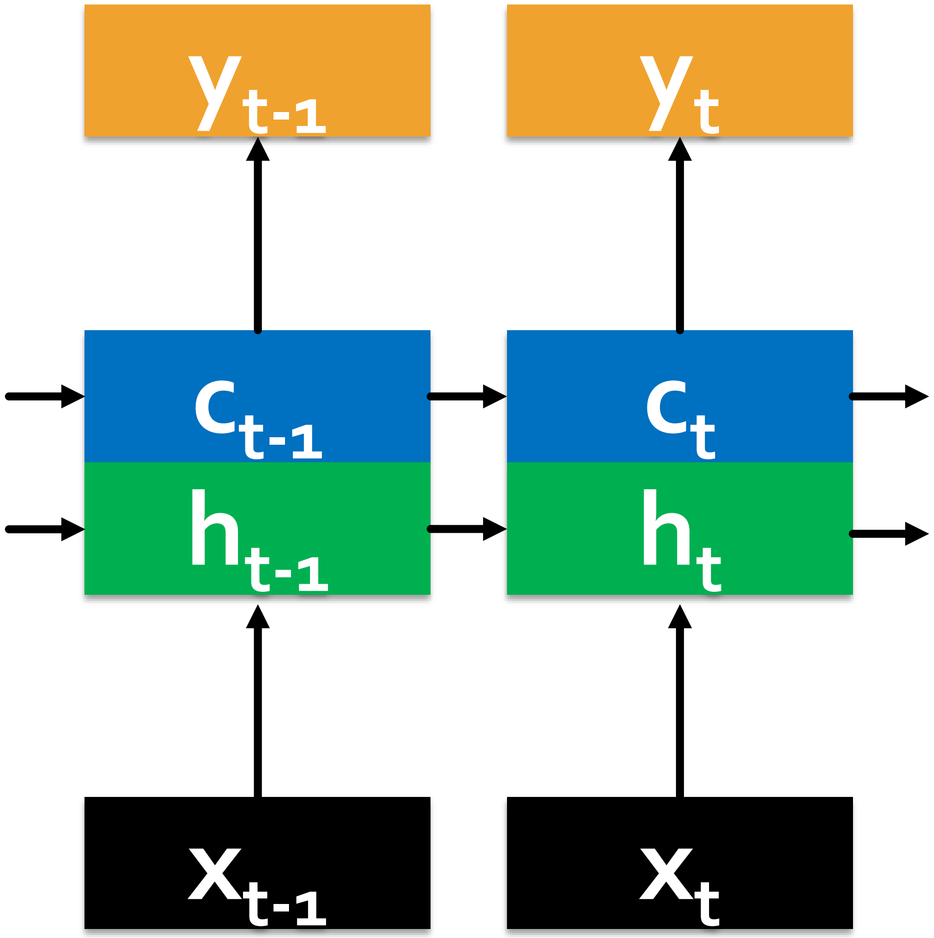
http://colah.github.io/posts/2015-08-Understanding-LSTMs/

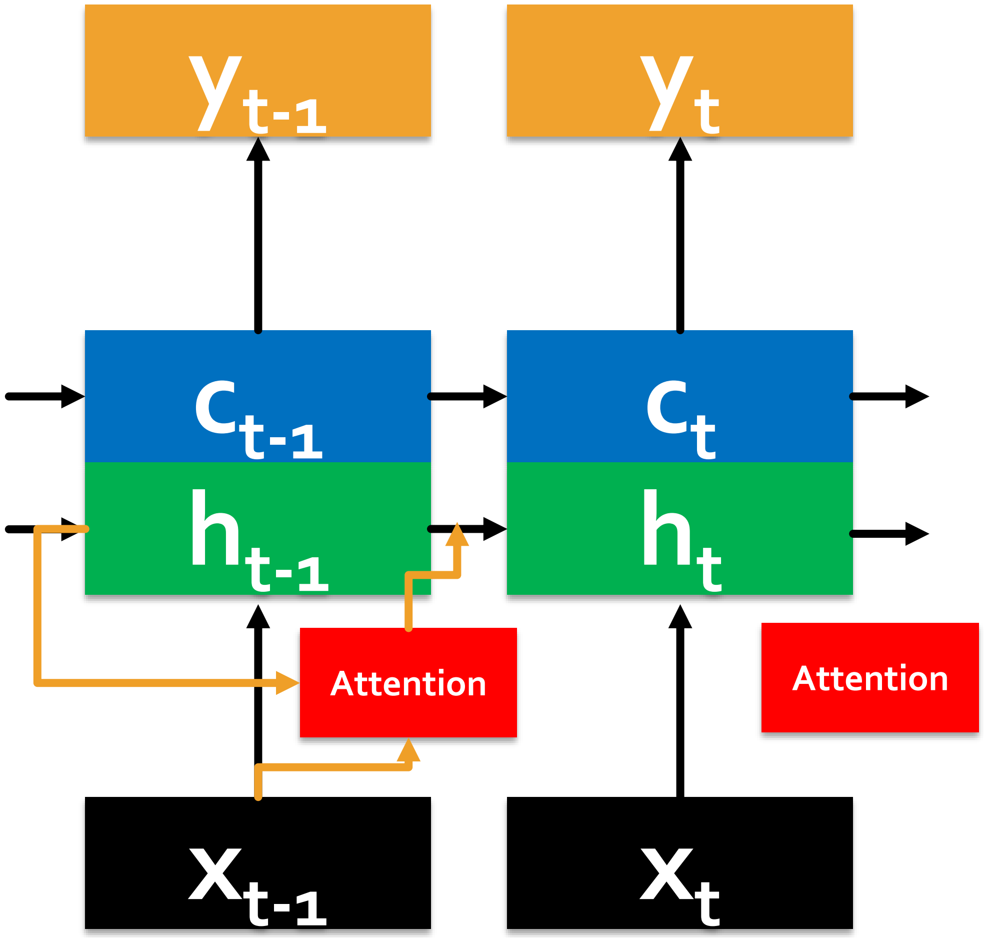
Attention Model



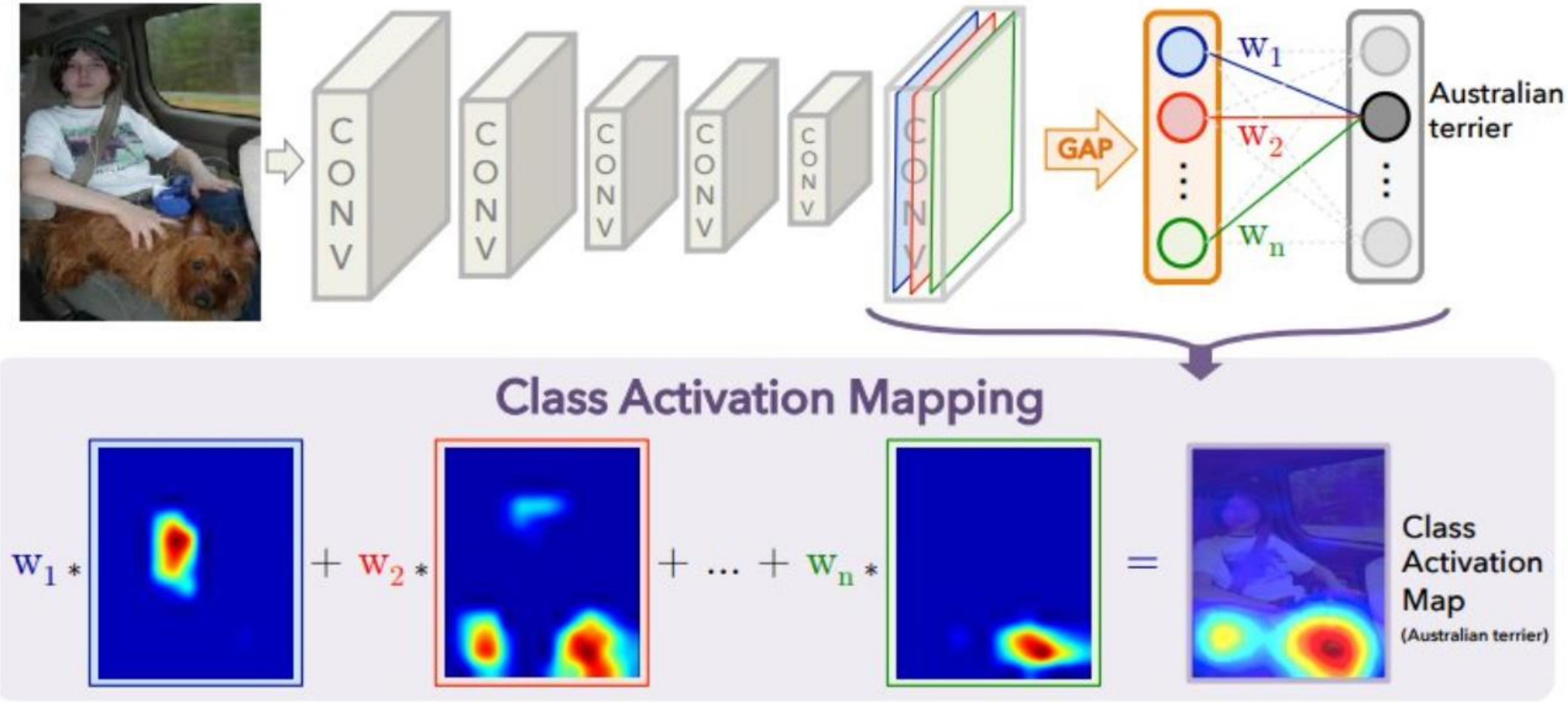


Attention Model with LSTM

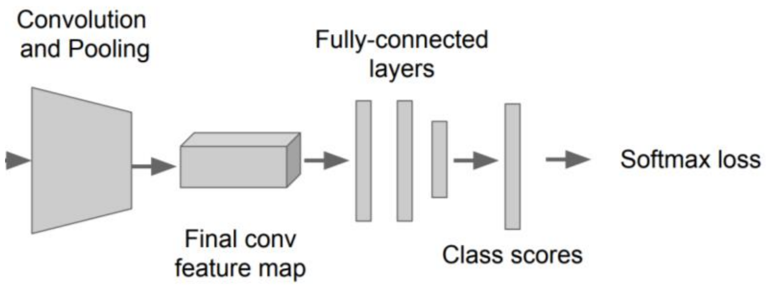




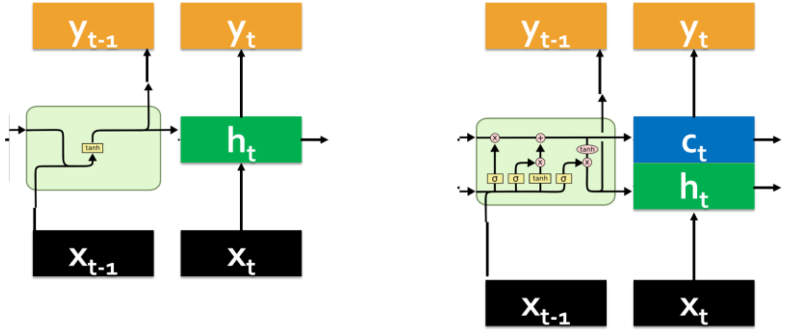
Class Activation Mapping



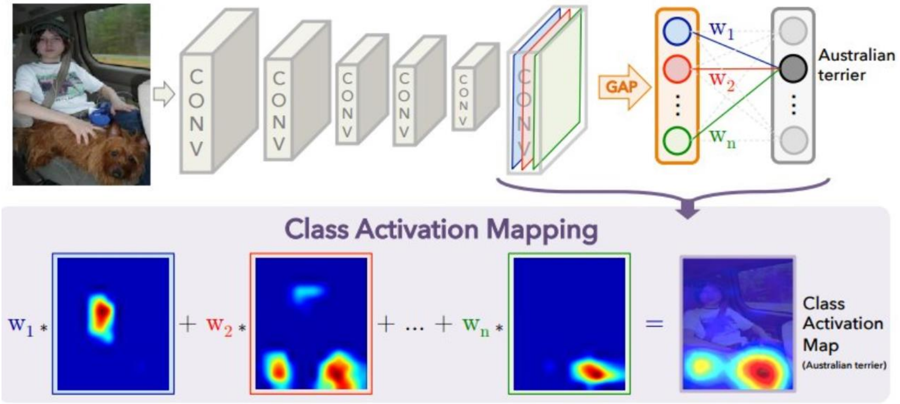
Knowledge For Today's Class



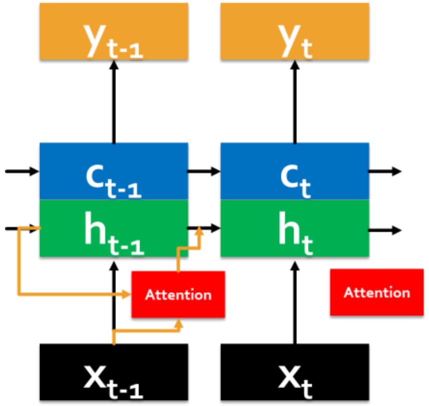
CNN



RNN (LSTM)



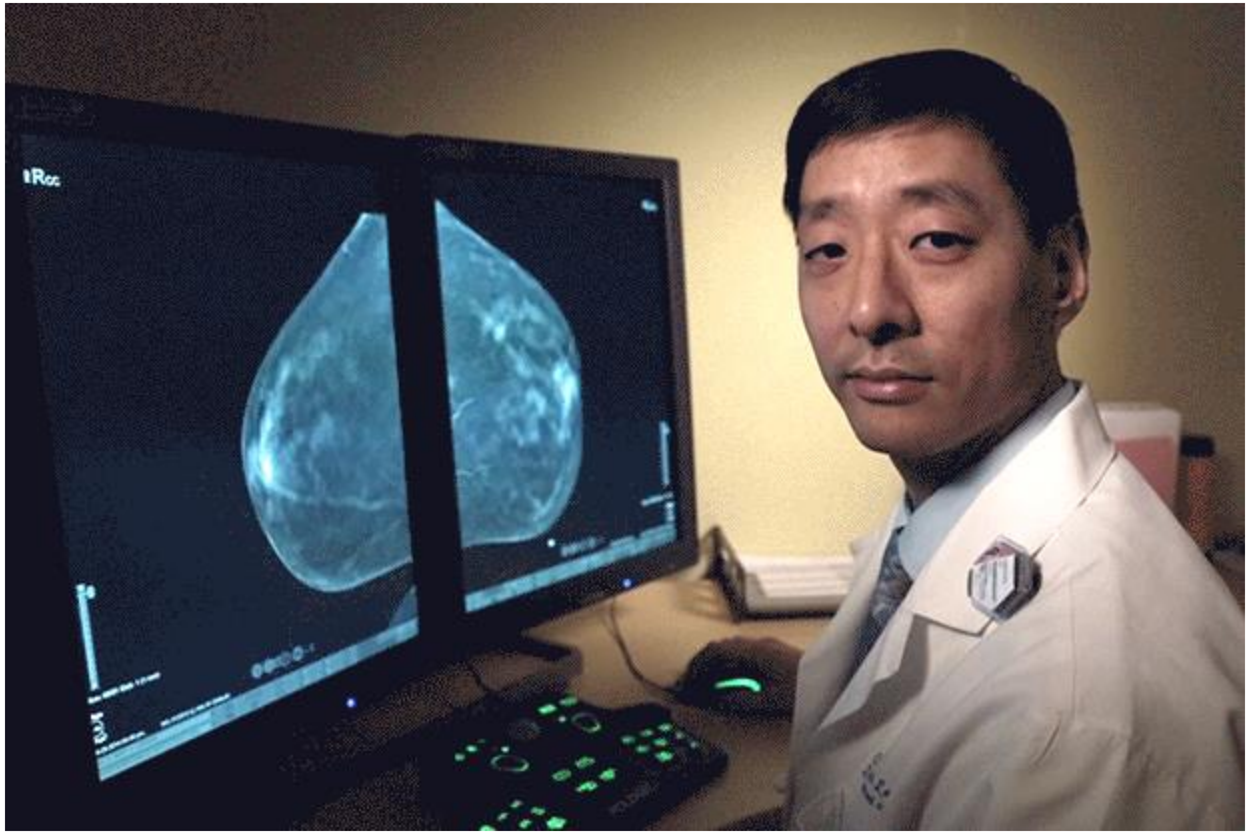
CNN with Attention



RNN with Attention

CS-8395 Deep Learning in Medical Image Computing

What do doctor do?

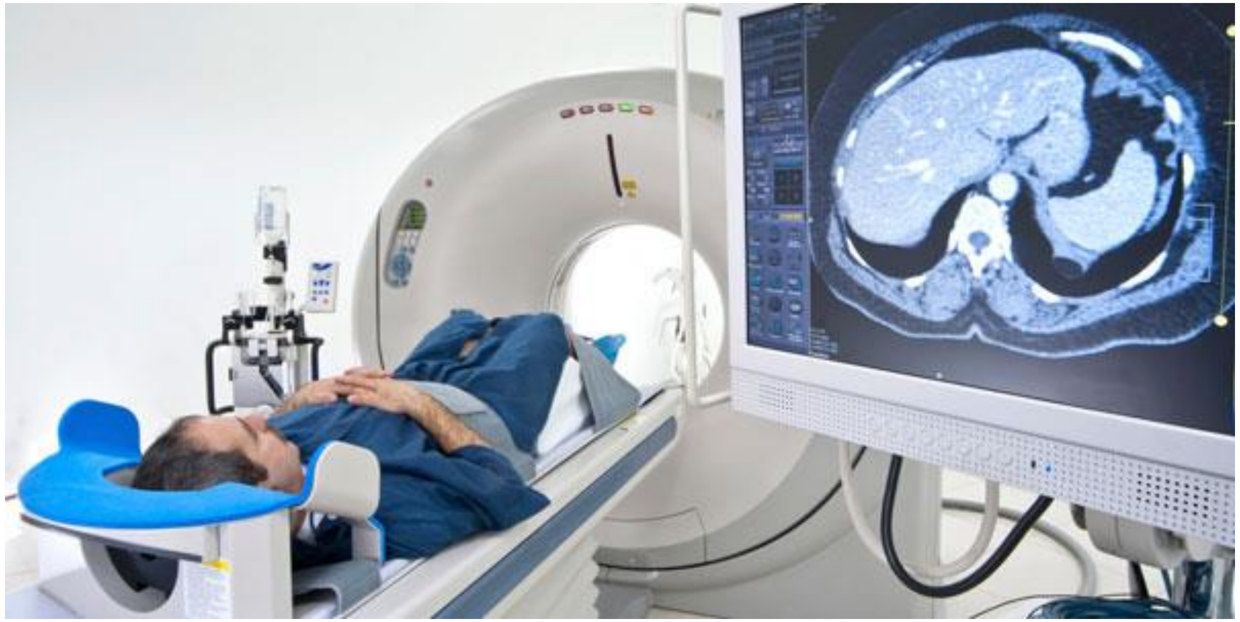


https://gfycat.com/uk/gifs/search/dr+z

In Practice

https://thealthtech.com/oscar-emr/

How?

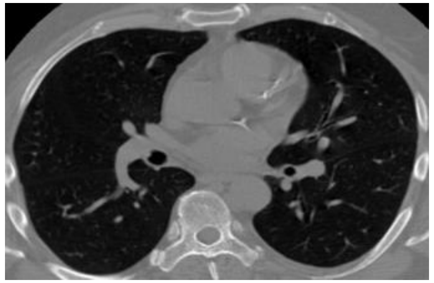


?

https://www.cdc.gov/features/medical-imaging-procedures/index.html

https://www.healthcare-informatics.com/news-item/report-emr-market-hit-25b-2014

Summarize



MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

CONFIDENTIAL

VA Medication History

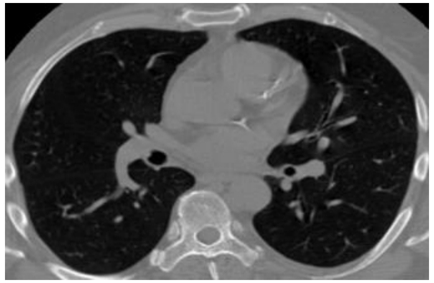
Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

Medication:	DOCUSATE NA 100MG CAP		
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY		
Status:	Active		
Refills Remaining:	10		
Last Filled On:	12 Apr 2012		
Initially Ordered On:	12 Apr 2012		
Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3

Summarize



MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

VA Medication History

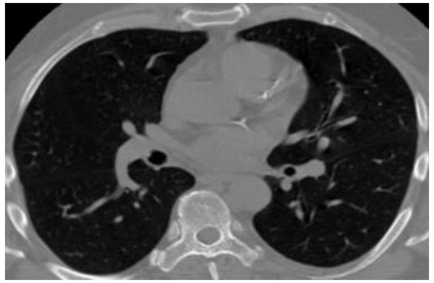
Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

Medication:	DOCUSATE NA 100MG CAP
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY
Status:	Active
Refills Remaining:	10
Last Filled On:	12 Apr 2012
Initially Ordered On:	12 Apr 2012

Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3



MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

VA Medication History

Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

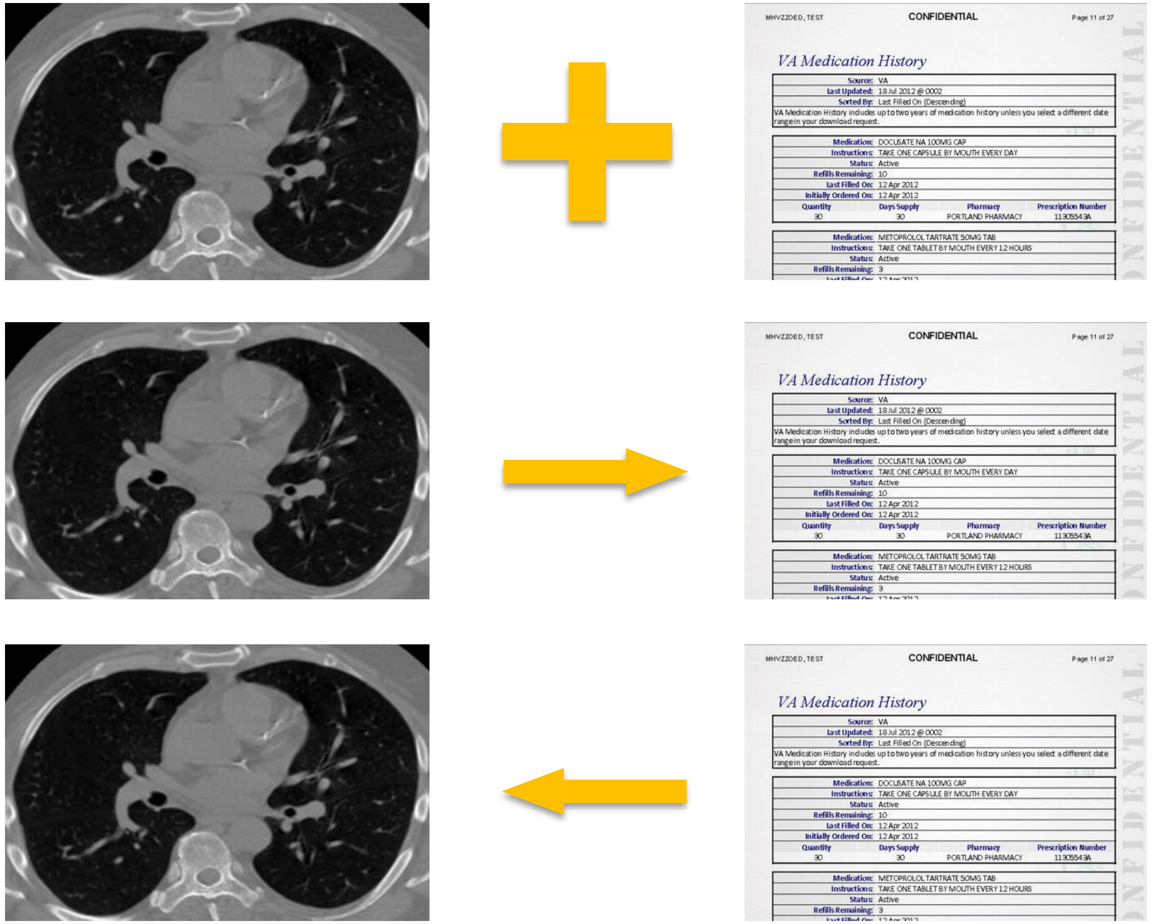
VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

Medication:	DOCUSATE NA 100MG CAP
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY
Status:	Active
Refills Remaining:	10
Last Filled On:	12 Apr 2012
Initially Ordered On:	12 Apr 2012

Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

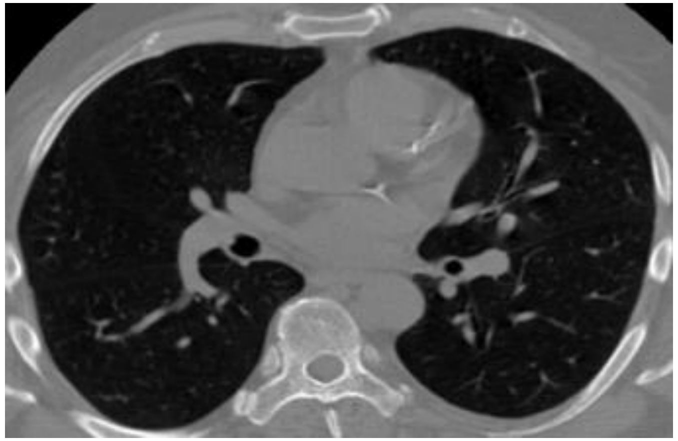
Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3

Summarize



MHVZZOED, TEST
CONFIDENTIAL
Page 11 of 27
VA Medication History
Source: VA
Last Updated: 18 Jul 2012 @ 0002
Sorted By: Last Filled On (Descending)
VA Medication History includes up to two years of medication history unless you select a different date range in your download request.
Medication: DOCUSATE NA 100MG CAP
Instructions: TAKE ONE CAPSULE BY MOUTH EVERY DAY
Status: Active
Refills Remaining: 10
Last Filled On: 12 Apr 2012
Initially Ordered On: 12 Apr 2012
Quantity 30
Days Supply 30
Pharmacy PORTLAND PHARMACY
Prescription Number 11305543A
Medication: METOPROLOL TARTRATE 50MG TAB
Instructions: TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status: Active
Refills Remaining: 3
CONFIDENTIAL

Strategy



MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

VA Medication History

Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

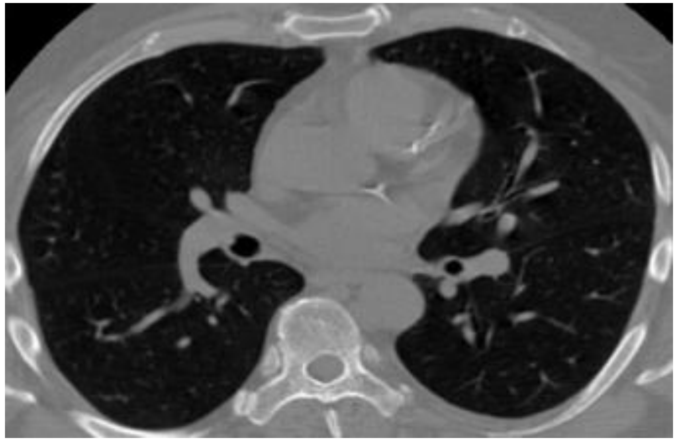
VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

Medication:	DOCUSATE NA 100MG CAP
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY
Status:	Active
Refills Remaining:	10
Last Filled On:	12 Apr 2012
Initially Ordered On:	12 Apr 2012

Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3

Strategy



MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

VA Medication History

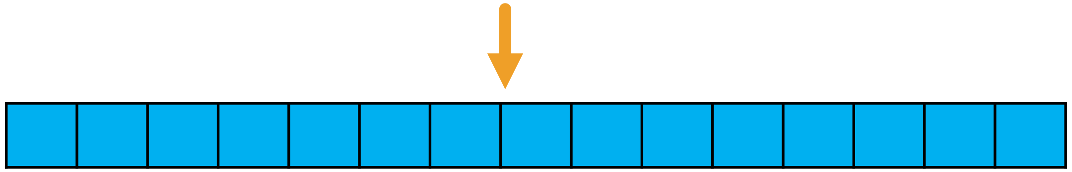
Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

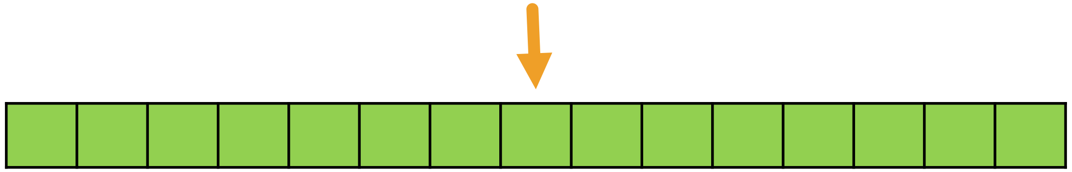
VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

Medication:	DOCUSATE NA 100MG CAP
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY
Status:	Active
Refills Remaining:	10
Last Filled On:	12 Apr 2012
Initially Ordered On:	12 Apr 2012

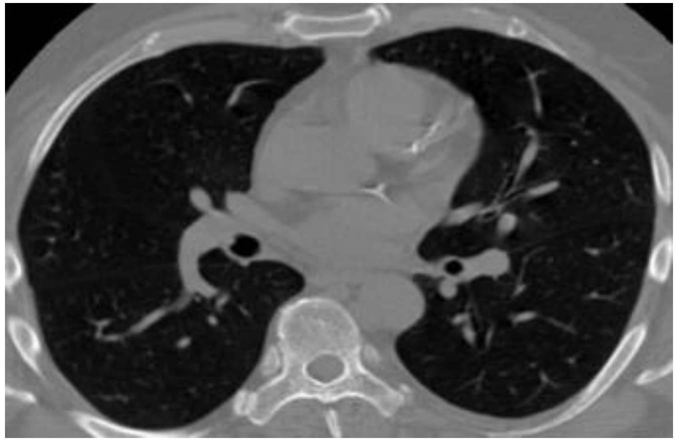
Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3





Strategy



MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

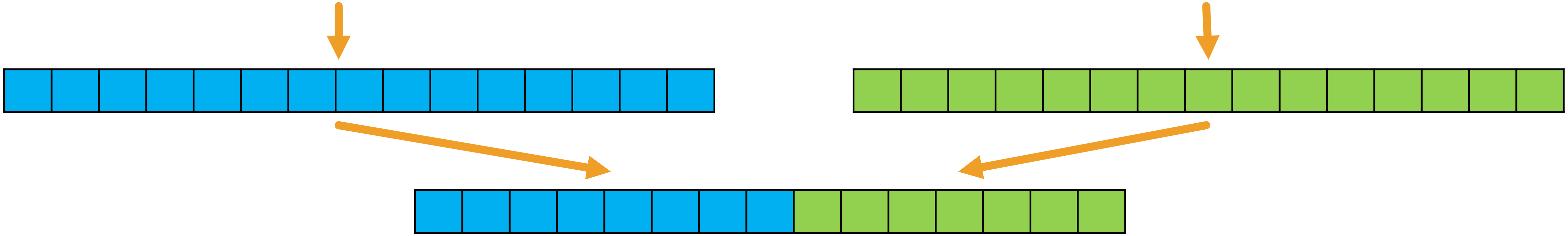
VA Medication History

Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)
VA Medication History includes up to two years of medication history unless you select a different date range in your download request.	

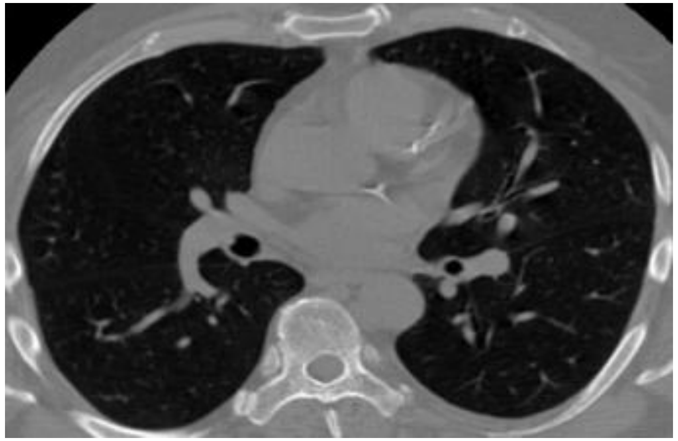
Medication:	DOCUSATE NA 100MG CAP		
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY		
Status:	Active		
Refills Remaining:	10		
Last Filled On:	12 Apr 2012		
Initially Ordered On:	12 Apr 2012		
Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3

CONFIDENTIAL



Strategy



MHVZZOED, TEST　　CONFIDENTIAL　　Page 11 of 27

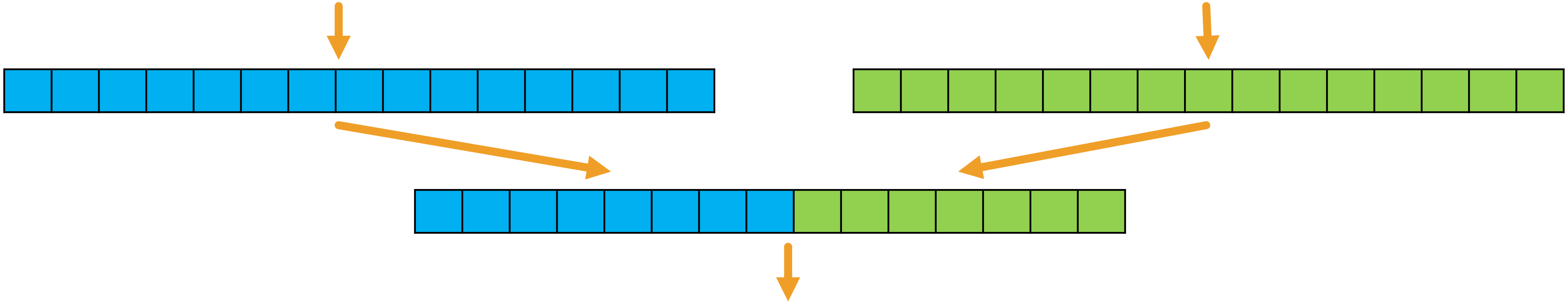
VA Medication History

Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

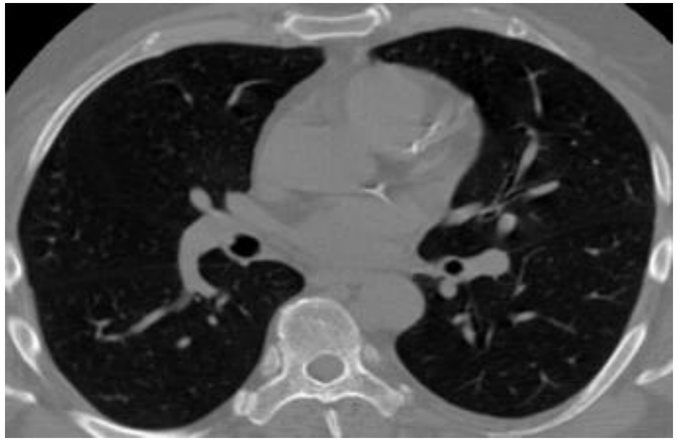
Medication:	DOCUSATE NA 100MG CAP		
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY		
Status:	Active		
Refills Remaining:	10		
Last Filled On:	12 Apr 2012		
Initially Ordered On:	12 Apr 2012		
Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3



Cancer?

Strategy



MHVZZOED, TEST — CONFIDENTIAL — Page 11 of 27

VA Medication History

Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

Medication:	DOCUSATE NA 100MG CAP
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY
Status:	Active
Refills Remaining:	10
Last Filled On:	12 Apr 2012
Initially Ordered On:	12 Apr 2012

Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3

CNN

DNN or RNN

DNN or RNN

Cancer?

Image + Text

TandemNet: Distilling Knowledge from Medical Images Using Diagnostic Reports as Optional Semantic References

Zizhao Zhang, Pingjun Chen, Manish Sapkota, Lin Yang

University of Florida

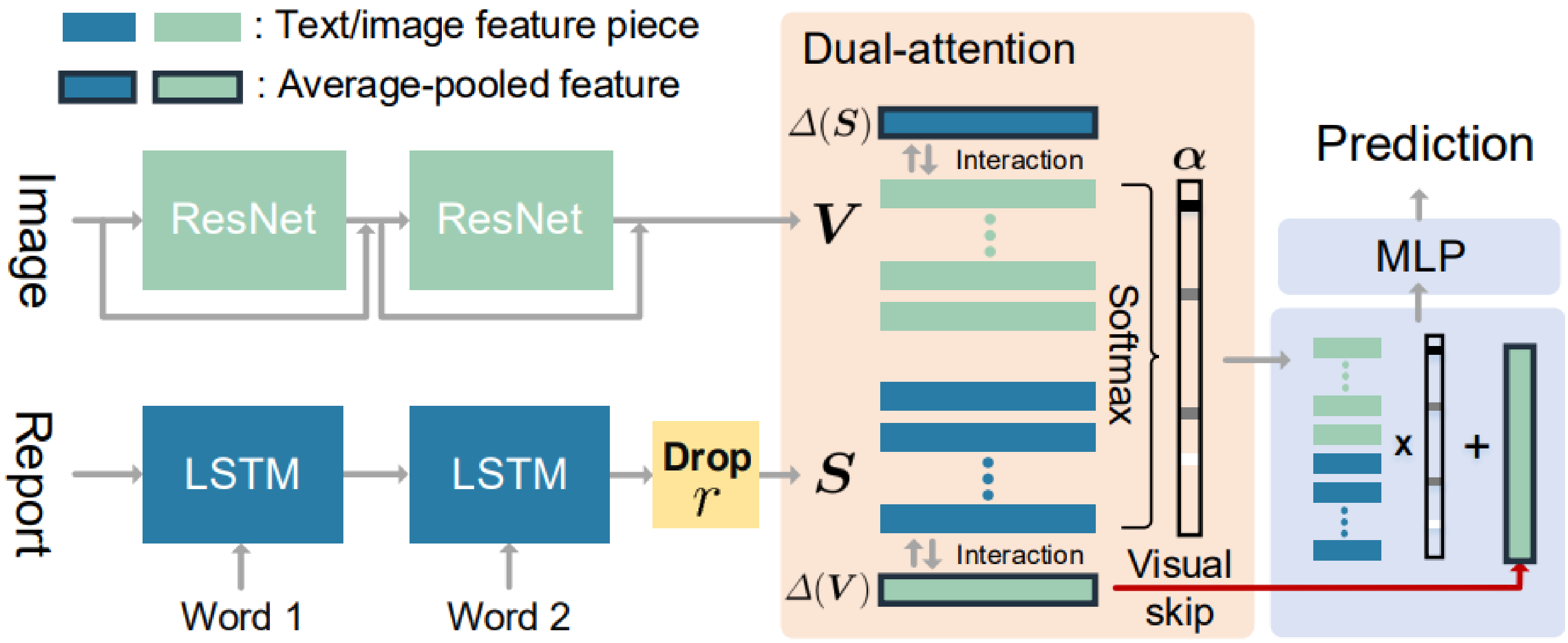


Fig. 1: The illustration of the TandemNet.

Image + Text

Method	Accuracy (%)	
	w/o text	w/ text
WRN16-4	75.4	-
ResNet18-TL	79.4	-
TandemNet-WVS	79.4	85.6
TandemNet	82.4	89.9
TandemNet-TL	84.9	88.6

Table 1: The quantitative evaluation (averaged on 3 trials). The first block shows standard CNNs so text is irrelevent.

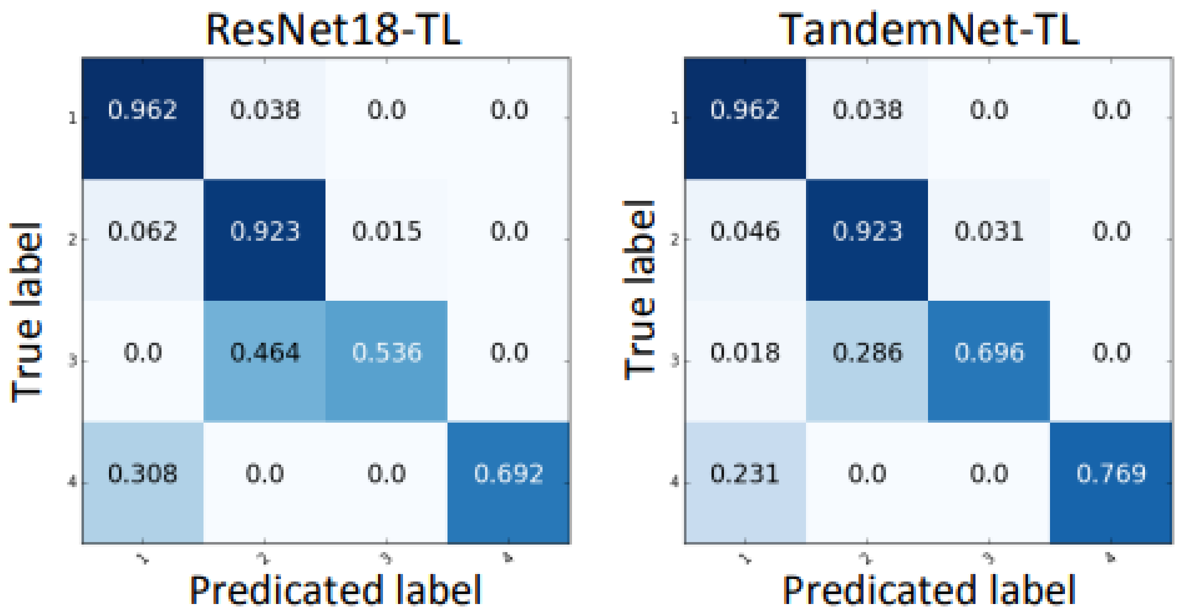
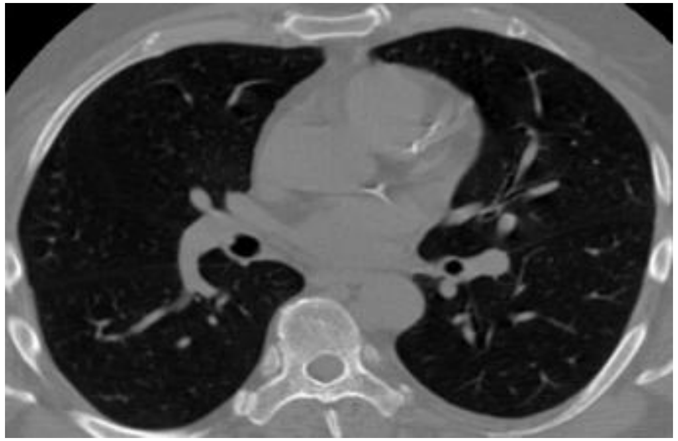


Fig. 2: The confusion matrices of two compared methods ResNet18-TL and TandemNet-TL (w/o text) in Table 1.

Strategy



MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

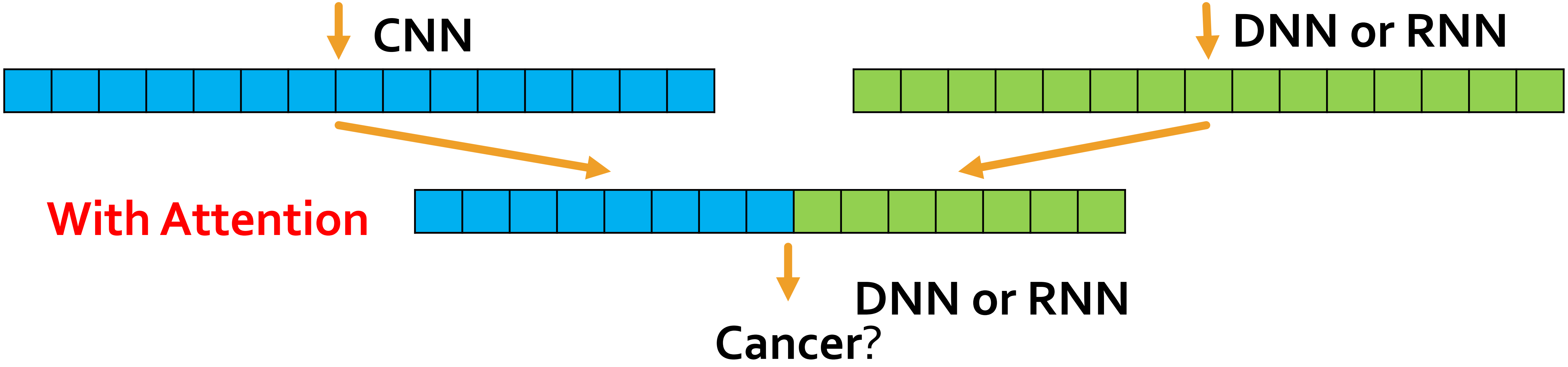
VA Medication History

Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

Medication:	DOCUSATE NA 100MG CAP		
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY		
Status:	Active		
Refills Remaining:	10		
Last Filled On:	12 Apr 2012		
Initially Ordered On:	12 Apr 2012		
Quantity	**Days Supply**	**Pharmacy**	**Prescription Number**
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3



Be Careful

Are we fooling ourselves?
• E.g. action recognition
– Very hard to improve on single frame classifiers
– Consider "opening fridge" action
example by David Fouhey

Image + Text

MDNet: A Semantically and Visually Interpretable Medical Image Diagnosis Network

Zizhao Zhang, Yuanpu Xie, Fuyong Xing, Mason McGough, Lin Yang
University of Florida
zizhao@cise.ufl.edu

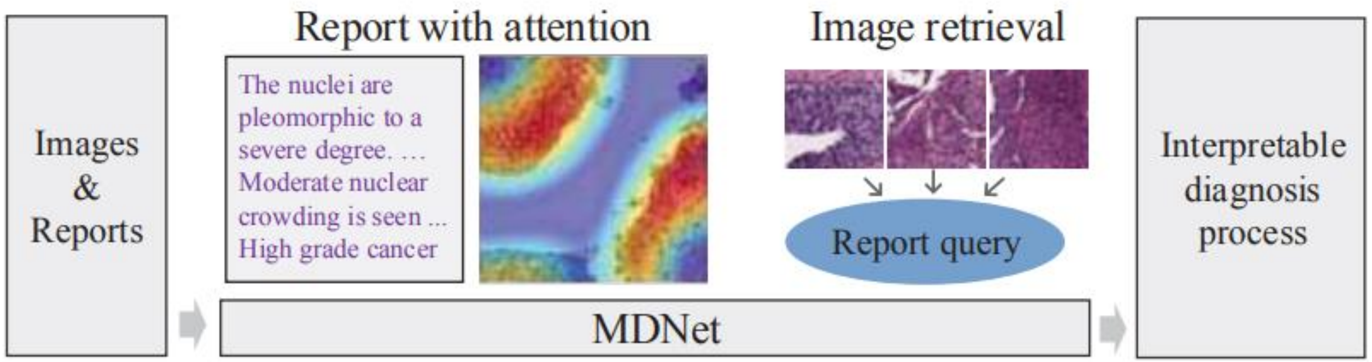


Figure 1: Overview of our medical image diagnosis network (MDNet) for interpretable diagnosis process.

Image + Text

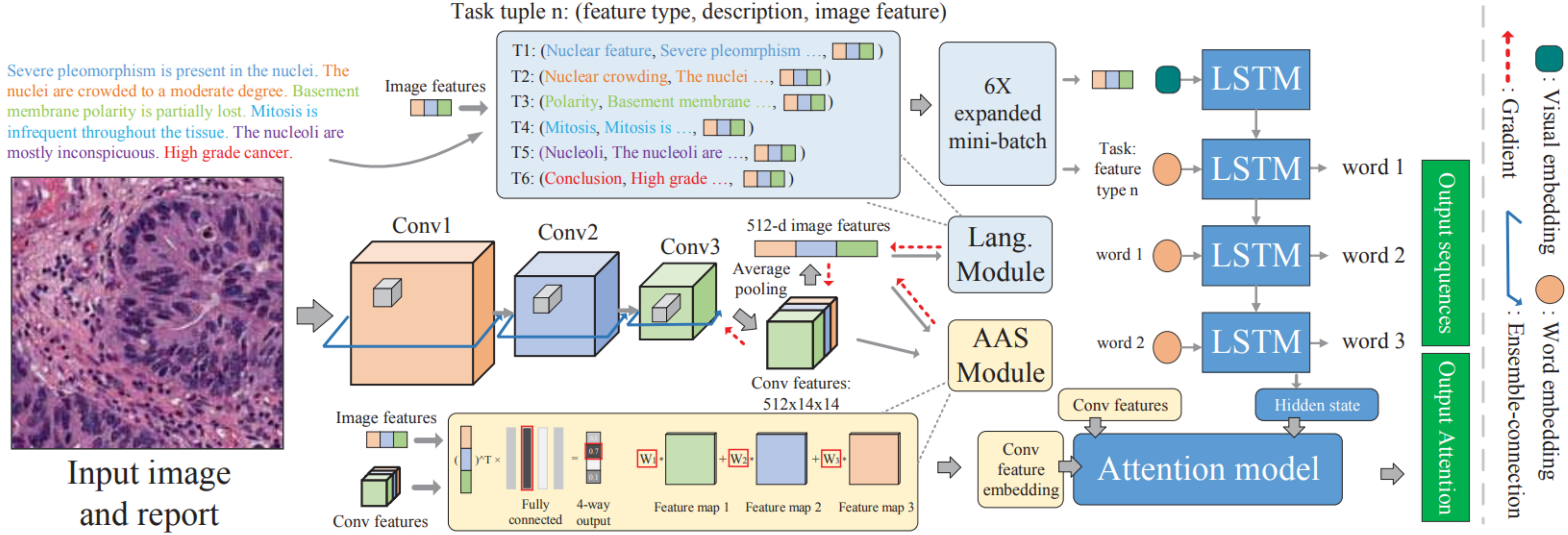


Figure 2: Overall illustration of MDNet. We use a bladder image with its diagnostic report as an example. The image model generates an image feature to pass to LSTM in the form of a task tuple and a Conv feature embedding (for the attention model) computed by the AAS module (defined in the method). LSTM executes prediction tasks according to the specified image feature type (best viewed in color).

Image + Text

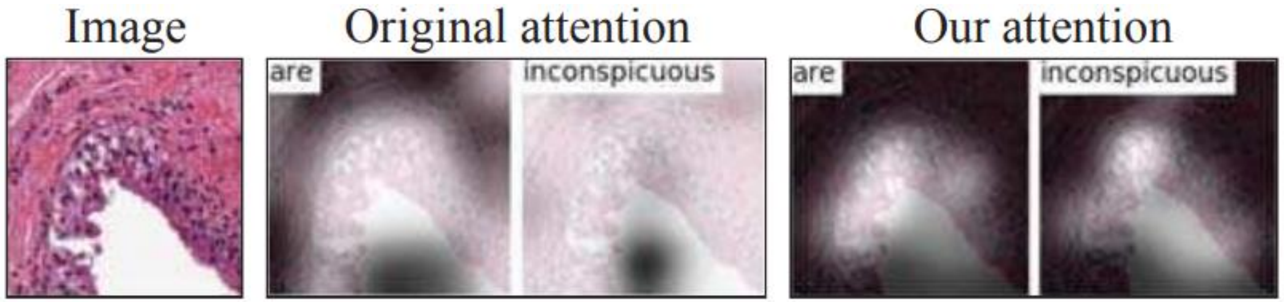


Figure 3: The attention maps of the original method (middle) and our method (right). Our method generates more focal attention on informative (urothelial) regions.

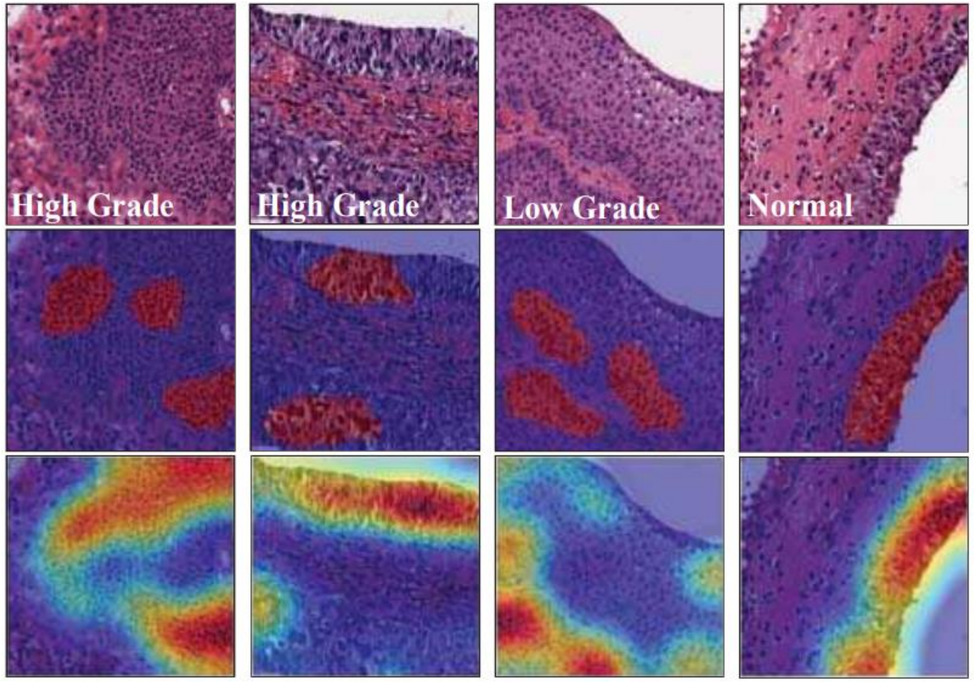
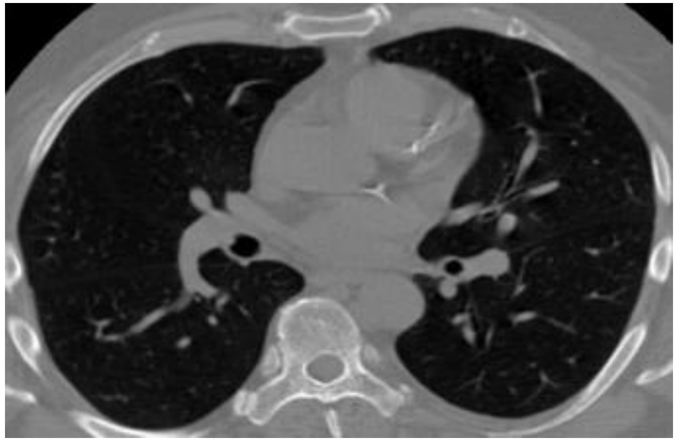


Figure 6: The illustration of class-specific attention. From top to bottom, test images, pathologist annotations, and class attention maps. Like the pathologist annotations, the attention maps are most activated in urothelial regions, largely ignoring stromal or background regions. Best viewed in color.

Strategy



MHVZZOED, TEST | CONFIDENTIAL | Page 11 of 27

VA Medication History

Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

Medication:	DOCUSATE NA 100MG CAP
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY
Status:	Active
Refills Remaining:	10
Last Filled On:	12 Apr 2012
Initially Ordered On:	12 Apr 2012

Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3

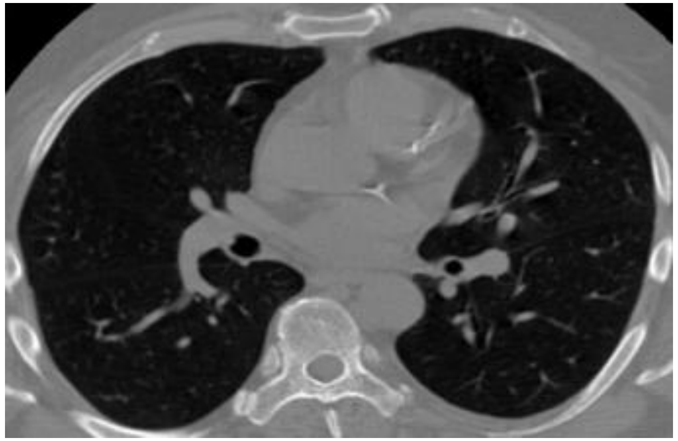
CNN

DNN or RNN

DNN or RNN

Cancer?

Strategy



MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

VA Medication History

Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)
VA Medication History includes up to two years of medication history unless you select a different date range in your download request.	

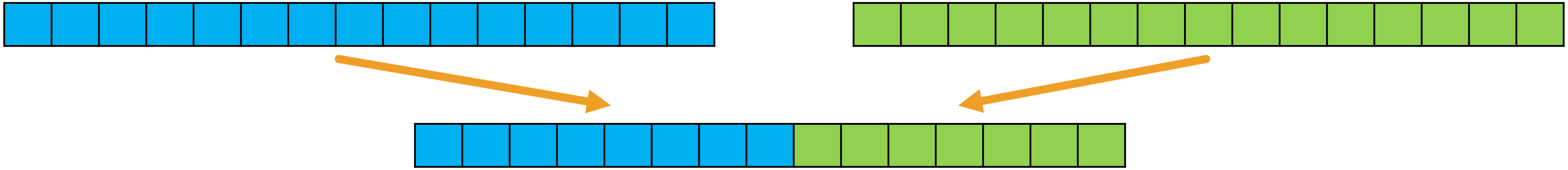
Medication:	DOCUSATE NA 100MG CAP
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY
Status:	Active
Refills Remaining:	10
Last Filled On:	12 Apr 2012
Initially Ordered On:	12 Apr 2012

Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3

CNN

DNN or RNN



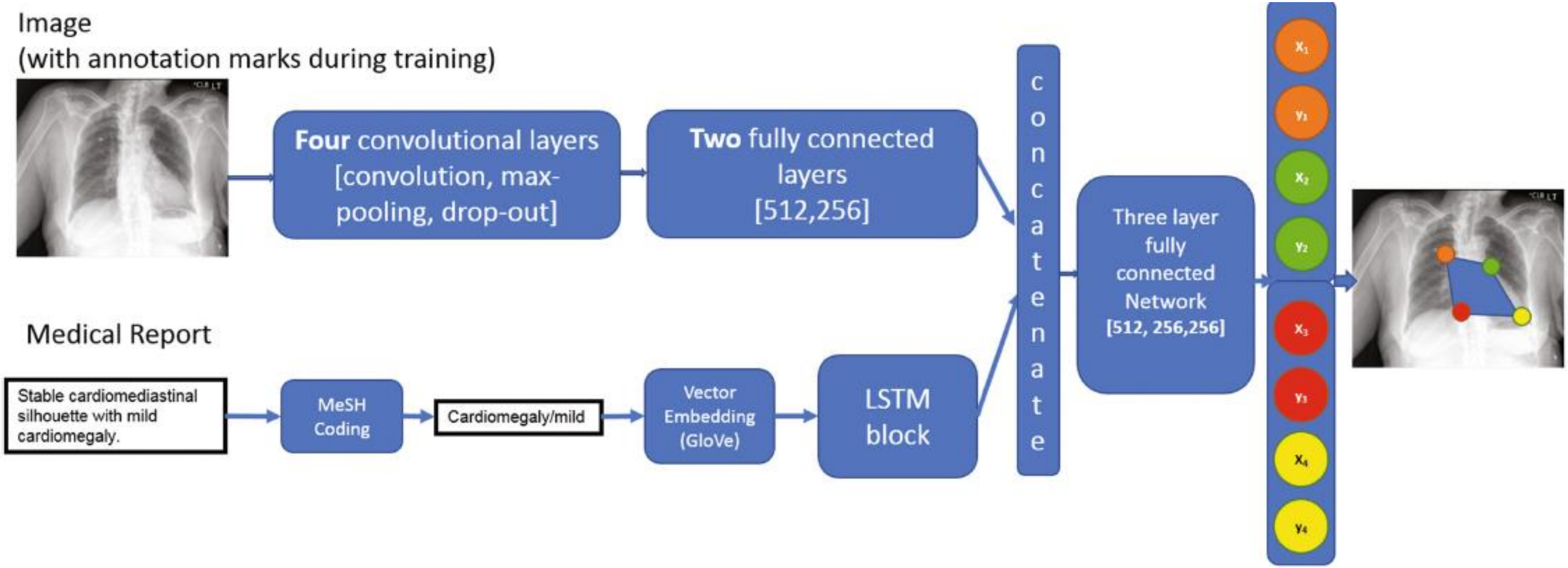
DNN or RNN

Segmentation

Bimodal Network Architectures for Automatic Generation of Image Annotation from Text

Mehdi Moradi(✉), Ali Madani, Yaniv Gur, Yufan Guo, and Tanveer Syeda-Mahmood

IBM Research - Almaden Research Center, San Jose, USA
mmoradi@us.ibm.com



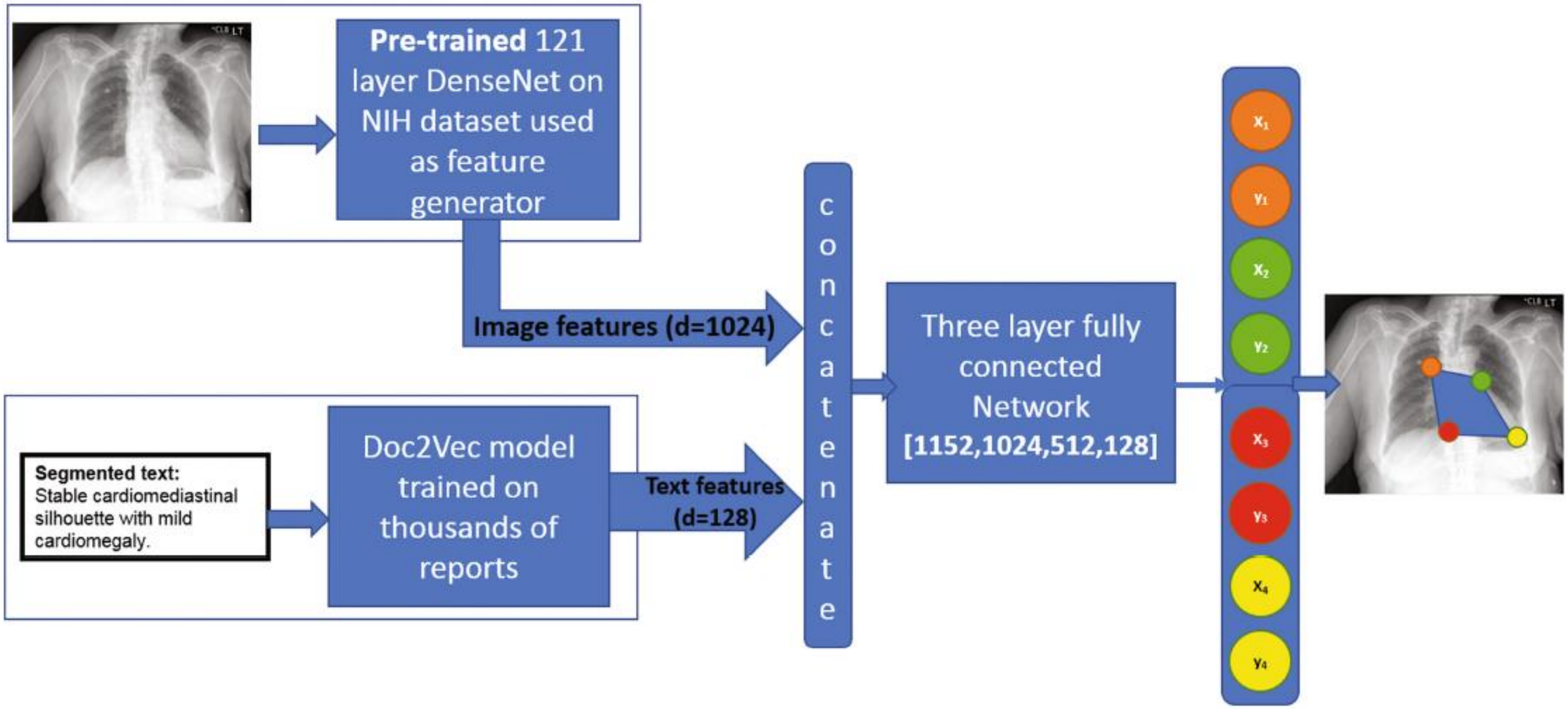


Fig. 2. Architecture 2: Image and text networks are trained separately. The resulting feature vectors are used to train a network that estimates ROI coordinates.

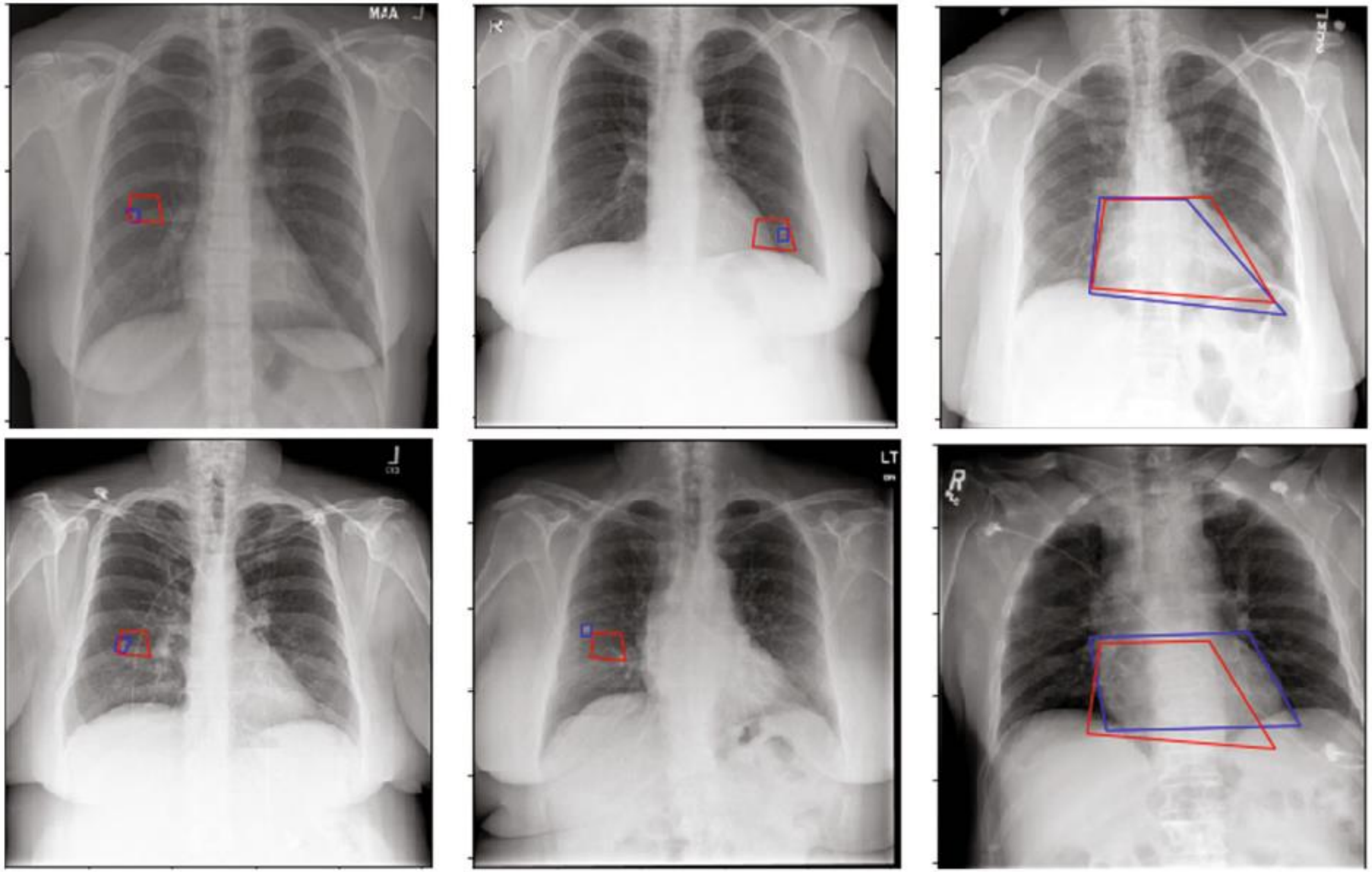
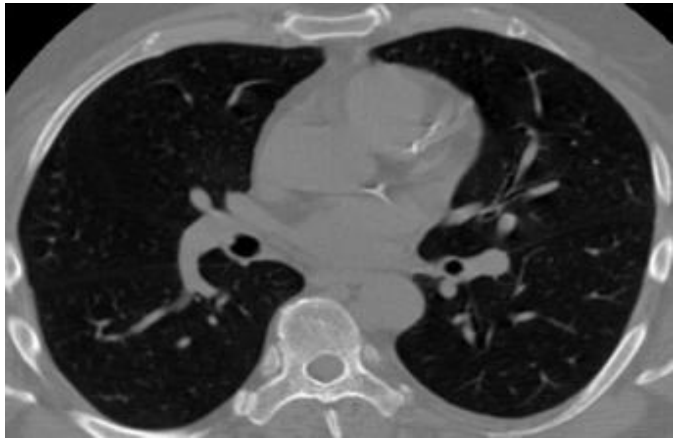


Fig. 3. Sample results from Architecture 2. Red is the estimated quadrilateral, blue is the one marked by a radiologist.

Strategy




MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

VA Medication History

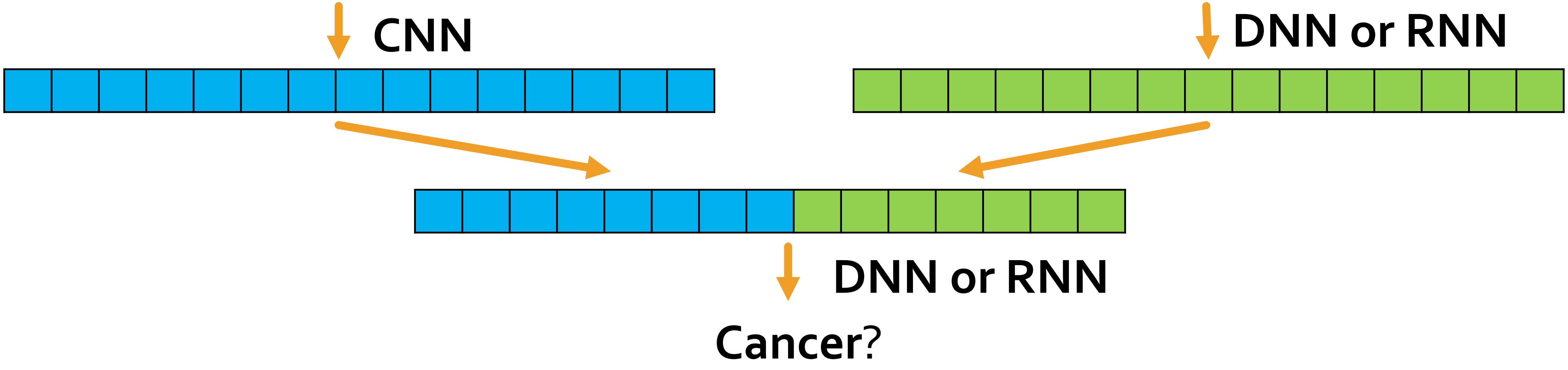
Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

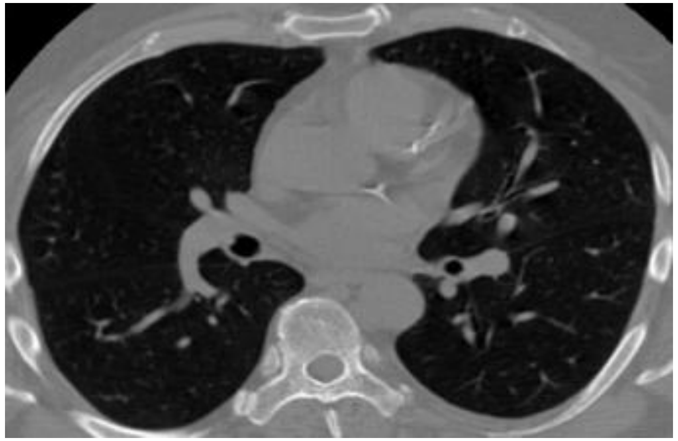
Medication:	DOCUSATE NA 100MG CAP
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY
Status:	Active
Refills Remaining:	10
Last Filled On:	12 Apr 2012
Initially Ordered On:	12 Apr 2012

Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3



Strategy



CNN, DNN, RNN

MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

VA Medication History

Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

Medication:	DOCUSATE NA 100MG CAP		
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY		
Status:	Active		
Refills Remaining:	10		
Last Filled On:	12 Apr 2012		
Initially Ordered On:	12 Apr 2012		
Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3

Image to Text

On the Automatic Generation of Medical Imaging Reports

Baoyu Jing[†*] **Pengtao Xie**[†*] **Eric P. Xing**[†]
[†]Petuum Inc, USA
[*]School of Computer Science, Carnegie Mellon University, USA
{baoyu.jing, pengtao.xie, eric.xing}@petuum.com

Impression: No acute cardiopulmonary abnormality.

Findings: There are no focal areas of consolidation. No suspicious pulmonary opacities. Heart size within normal limits. No pleural effusions. There is no evidence of pneumothorax. Degenerative changes of the thoracic spine.

MTI Tags: degenerative change

Image to Text

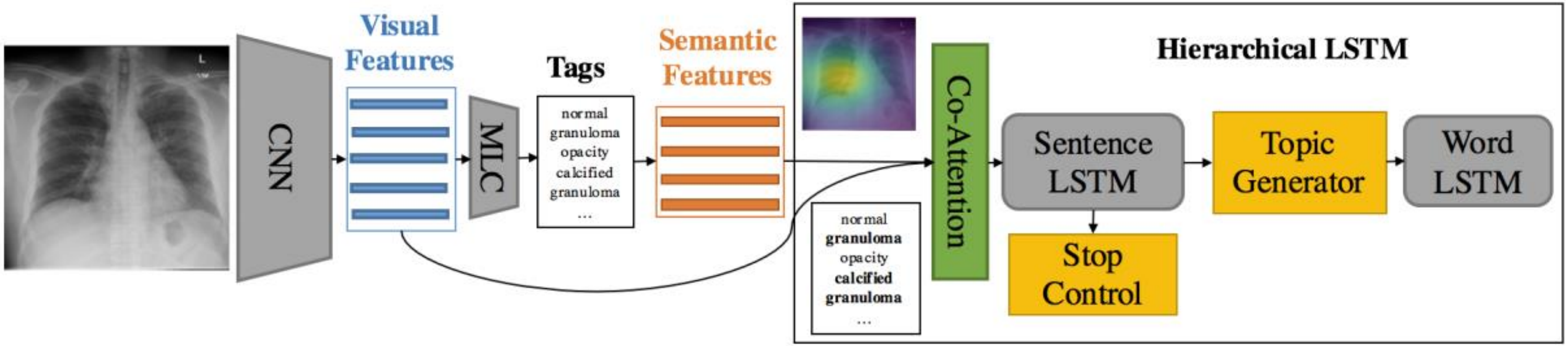


Figure 2: Illustration of the proposed model. MLC denotes a *multi-label classification* network. Semantic features are the word embeddings of the predicted tags. The boldfaced tags "calcified granuloma" and "granuloma" are attended by the co-attention network.

Image to Text

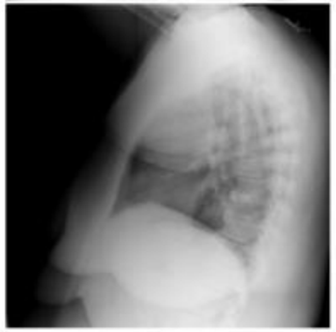
	Ground Truth	Ours-CoAttention	Ours-no-Attention	Soft Attention
	No active disease. The heart and lungs have in the interval. Both lungs are clear and expanded. Heart and mediastinum normal.	No active disease. The heart and lungs have in the interval. Lungs are clear and expanded. Cardiomediastinal silhouette is within normal limits. No pleural effusion or pneumothorax is seen. No pleural effusion. No cavitary or pneumothorax.	The lungs are clear bilaterally. The are grossly normal. No focal lung consolidation. No acute bony abnormality. cm nodule within the right lower lobe on the lateral view. No pneumothorax or pleural effusion. No acute bony abnormality. The heart is not enlarged. The lungs are clear. No acute bony abnormality .	No acute cardiopulmonary abnormality. The lungs are clear bilaterally. Specifically no evidence of focal airspace consolidation pleural effusion or pneumothorax. Cardio mediastinal silhouette is unremarkable. Visualized osseous structures of the thorax are without acute abnormality.
	No evidence of active disease. The lungs are clear. There is no focal airspace consolidation. No pleural effusion or pneumothorax. Heart size and mediastinal contour are within normal limits. There are multilevel degenerative changes of the spine.	No acute cardiopulmonary findings. Heart size is not enlarged. No focal airspace consolidation suspicious pulmonary opacity large pleural effusion or pneumothorax. No focal areas of consolidation. Degenerative changes of the spine. This is moderate exam of the hydropneumothorax. Lungs are clear. There is no focal airspace consolidation pleural effusion or pneumothorax.	The lungs are clear bilaterally. The are grossly normal. No pleural effusion. The heart is normal in size and contour. The lungs are clear. There are no acute bony findings.	No acute cardiopulmonary abnormality. The lungs are clear bilaterally. There is no pleural effusion or pneumothorax. The heart and mediastinum are normal. There is no focal air space opacity to suggest a pneumonia.
	No acute cardiopulmonary abnormality. Normal heart size mediastinal contours. Eventration of the right hemidiaphragm. No focal airspace consolidation. No pleural effusion or pneumothorax.	No acute cardiopulmonary abnormality. Stable appearance of the thoracic aorta. The right lateral lower lobe is noted in the right lower right midlung. No large pleural effusion or focal airspace disease. Mild interstitial opacities. Atherosclerotic calcifications bony structures bilaterally. There is no pleural effusion or pneumothorax developed in the right lower lobe.	The lungs are clear bilaterally. The are grossly normal. No acute bony abnormality. The lungs are otherwise clear. No acute osseous abnormality. No acute osseous abnormality. The heart and mediastinum are normal. There is no focal air space opacity.	No acute cardiopulmonary abnormality. The lungs are clear bilaterally. There is no focal airspace consolidation. No pleural effusion or pneumothorax. Heart size and pulmonary vascularity appear within normal limits.

Image to Text

TieNet: Text-Image Embedding Network for Common Thorax Disease Classification and Reporting in Chest X-rays

Xiaosong Wang*[1], Yifan Peng*[2], Le Lu[1], Zhiyong Lu[2], Ronald M. Summers[1]
[1]Department of Radiology and Imaging Sciences, Clinical Center,
[2] National Center for Biotechnology Information, National Library of Medicine,
National Institutes of Health, Bethesda, MD 20892
{xiaosong.wang,yifan.peng,luzh,rms}@nih.gov, lel@nvidia.com

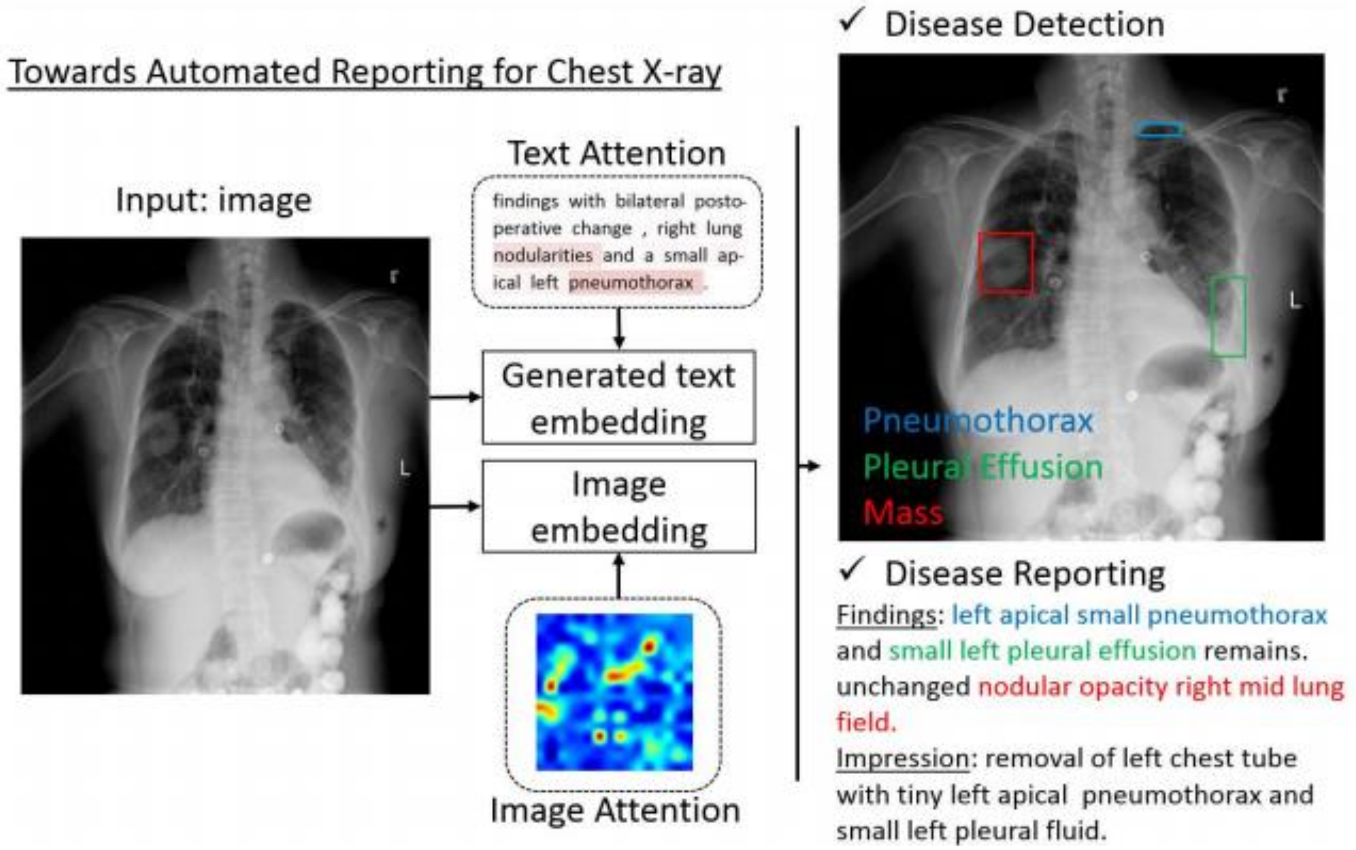


Figure 1. Overview of the proposed automated chest X-ray reporting framework. A multi-level attention model is introduced.

Image to Text

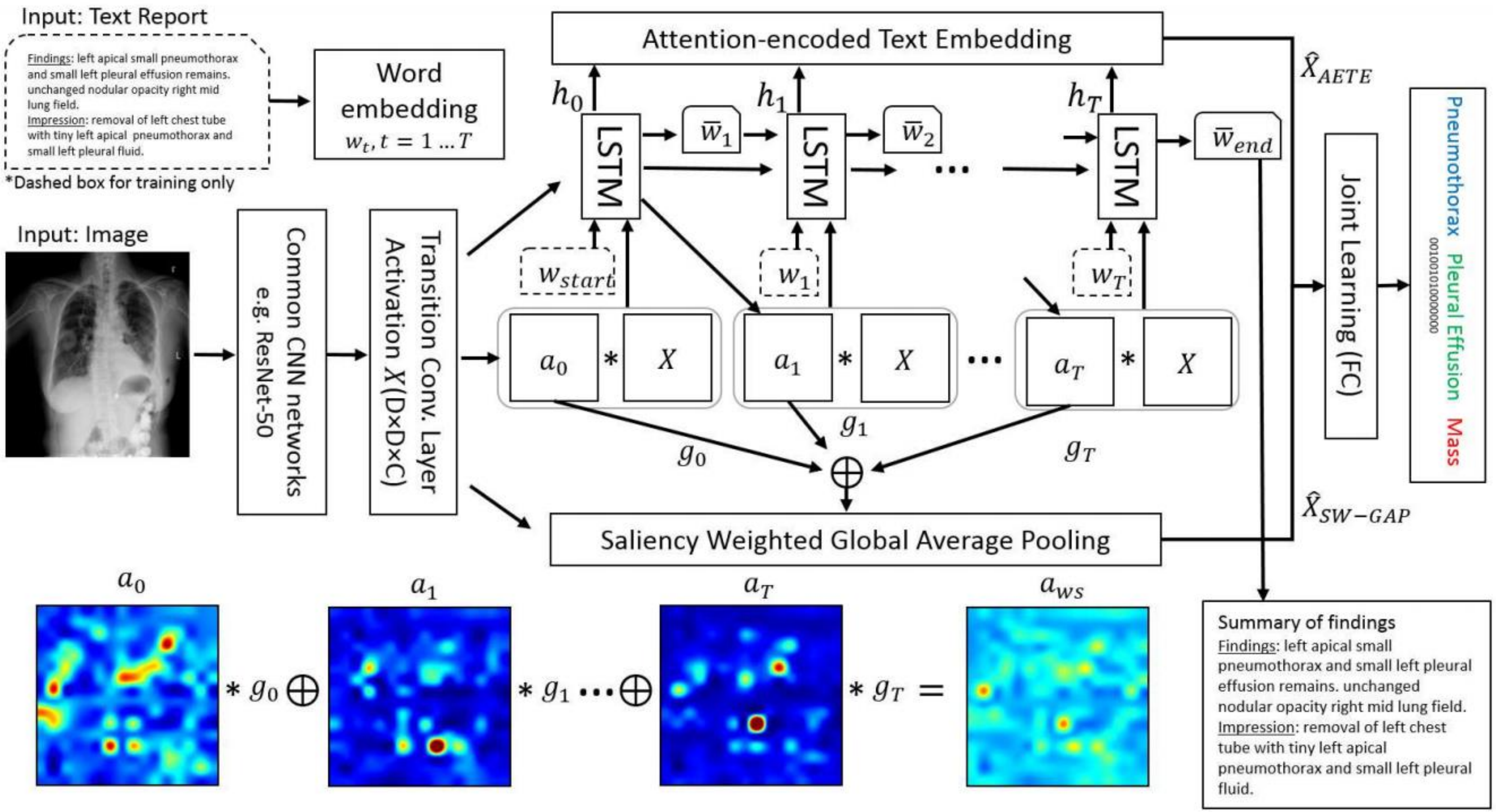
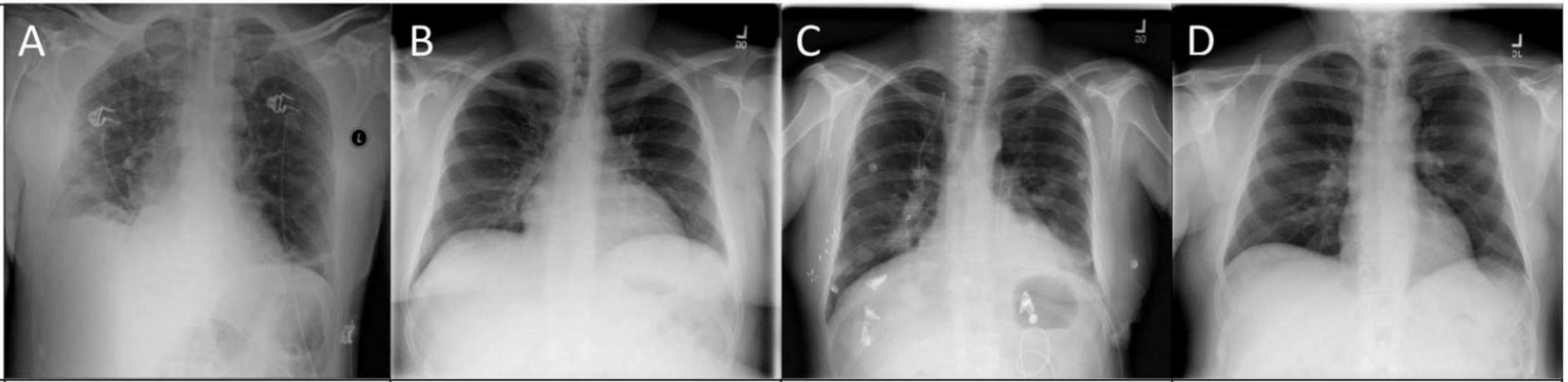


Image to Text

Image Sample cases



	A	B	C	D
P	Atelectasis Effusion	No finding	Nodule Pneumothorax Mass Consolidation	Mass
Original report	findings : a single ap view of the chest demonstrates increasing bibasilar interstitial opacities with decreased overall aeration . increasing blunting of right costophrenic angle. … impression : increasing bibasilar atelectasis with possible development of right pleural effusion .	Normal no evidence of lung infiltrate .	findings : heart and mediastinum unchanged . multiple lung nodules . evidence of recent left chest surgery with left chest tube in place . very small left apical pneumothorax . lungs unchanged , no evidence of acute infiltrates . impression : stable chest .	findings : large left suprahilar and infrahilar masses as well as the well circumscribed nodule the level of the aortic knob . the right infrahilar mass as well . no effusion . impression : metastatic lung disease .
Generated Report	findings : a single ap view of the chest demonstrates unchanged bilateral reticular opacities , consider atelectasis . continued left basilar atelectasis . no evidence of developing infiltrate . the cardiac and mediastinal contours are stable . impression : no evidence of developing infiltrate .	findings : pa and lateral views of the chest demonstrate lungs that are clear without focal mass , infiltrate or effusion . cardiomediastinal silhouette is normal size and contour . pulmonary vasculartity is normal in caliber and distribution . impression : no evidence of acute pulmonary pathology	findings : pa and lateral views of the chest demonstrate unchanged bilateral chest tubes . again pulmonary nodules are seen on the right and cardiac silhouette unchanged . the cardiac and mediastinal contours are stable . impression : 1. bilateral masses and left lower lung field consolidation . 2.new bilateral lung masses .	comparison is to previous upright study of no significant interval change is seen in the appearance of the chest . the mediastinal soft tissue and pulmonary vascularity are stable . there are blastic bone lesions in the chest . bones , soft tissues are normal . the lung fields are clear . there are calcified lymph nodes in the left lower lung . impression : . sclerotic lesions in the left humeral , consistent with metastasis.

Image to Text

Multimodal Recurrent Model with Attention for Automated Radiology Report Generation

Yuan Xue[1], Tao Xu[2], L. Rodney Long[3], Zhiyun Xue[3], Sameer Antani[3], George R. Thoma[3], and Xiaolei Huang[1(✉)]

[1] College of Information Sciences and Technology, Penn State University, University Park, PA, USA
sharon.x.huang@gmail.com
[2] Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA, USA
[3] National Library of Medicine, National Institutes of Health, Bethesda, MD, USA

Abstract. Radiologists routinely examine medical images such as X-Ray, CT, or MRI and write reports summarizing their descriptive findings and conclusive impressions. A computer-aided radiology report generation system can lighten the workload for radiologists considerably and assist them in decision making. Although the rapid development of deep learning technology makes the generation of a single conclusive sentence possible, results produced by existing methods are not sufficiently reliable due to the complexity of medical images. Furthermore, generating detailed paragraph descriptions for medical images remains a challenging problem. To tackle this problem, we propose a novel generative model which generates a complete radiology report automatically. The proposed model incorporates the Convolutional Neural Networks (CNNs) with the Long Short-Term Memory (LSTM) in a recurrent way. It is capable of not only generating high-level conclusive impressions, but also generating detailed descriptive findings sentence by sentence to support the conclusion. Furthermore, our multimodal model combines the encoding of the image and one generated sentence to construct an attention input to guide the generation of the next sentence, and henceforth maintains coherence among generated sentences. Experimental results on the publicly available Indiana U. Chest X-rays from the Open-i image collection show that our proposed recurrent attention model achieves significant improvements over baseline models according to multiple evaluation metrics.

Image to Text

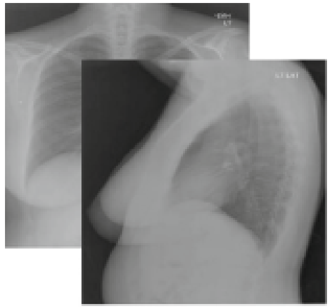
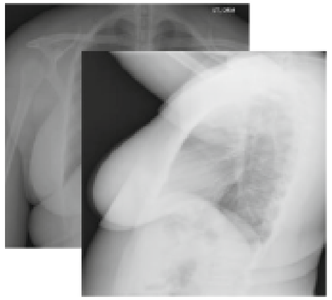
Input Image	Recurrent Attention	Ground Truth
	**Findings**: The heart size and mediastinal contours appear within normal limits. No focal airspace consolidation , pleural effusion or pneumothorax. No acute bony abnormalities. **Impression**: No acute cardiopulmonary finding.	**Findings**: The heart size and mediastinal silhouette are within normal limits for contour. The lungs are clear. No pneumothorax or pleural effusions. The XXXX are intact. **Impression**: No acute cardiopulmonary abnormalities.
	**Findings**: The heart size and mediastinal silhouette are within normal limits for contour. The lungs are clear. No focal airspace consolidation. No pleural effusion or pneumothorax. Normal cardiomediastinal silhouette. Heart size is normal. **Impression**: Clear lungs. No acute cardiopulmonary abnormality.	**Findings**: Mediastinal contours are within normal limits. Heart size is within normal limits. No focal consolidation, pneumothorax or pleural effusion. No bony abnormality. Vague density in right mid lung, XXXX related to scapular tip and superimposed ribs. Not visualized on lateral exam. **Impression**: Vague density in right XXXX, XXXX related to scapular tip and superimposed ribs. Consider oblique images to exclude true nodule. 2. No acute cardiopulmonary abnormality.

Fig. 1. Examples of original reports vs. reports generated by our recurrent attention model. Note that, Findings is a paragraph containing some descriptive sentences; Impression is a conclusive sentence. XXXXs are wrongly removed keywords due to de-identification.

Image to Text

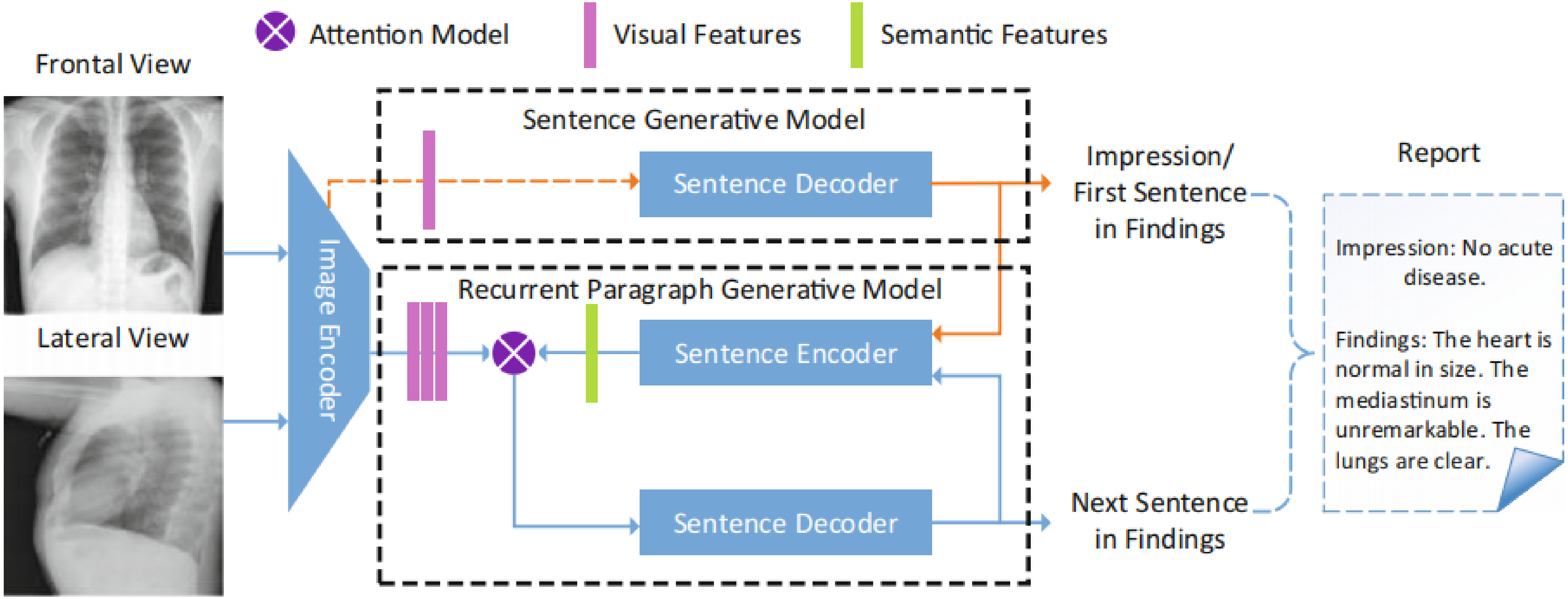
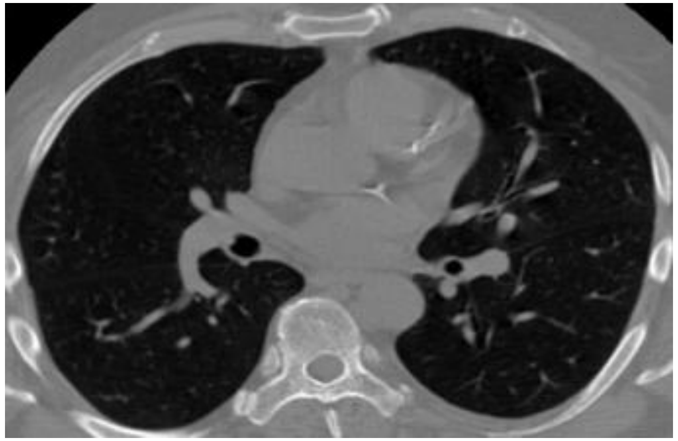


Fig. 2. The architecture of the proposed multimodal recurrent generation model with attention for radiology reports. Best viewed in color.

Strategy



CNN, DNN, RNN

MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

VA Medication History

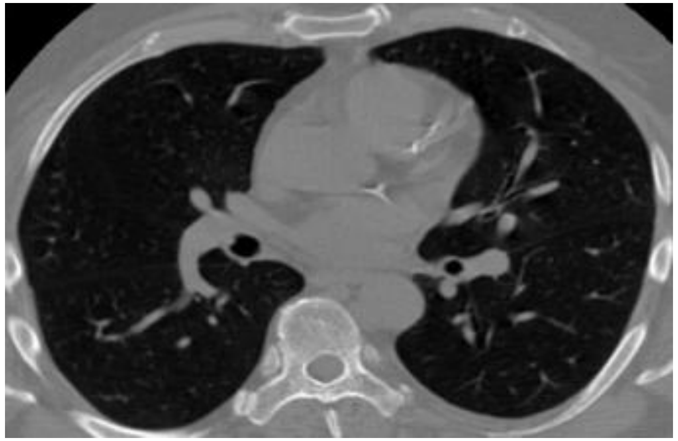
Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

Medication:	DOCUSATE NA 100MG CAP		
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY		
Status:	Active		
Refills Remaining:	10		
Last Filled On:	12 Apr 2012		
Initially Ordered On:	12 Apr 2012		
Quantity	**Days Supply**	**Pharmacy**	**Prescription Number**
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3

Strategy



CNN, DNN, RNN

MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

VA Medication History

Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

Medication:	DOCUSATE NA 100MG CAP		
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY		
Status:	Active		
Refills Remaining:	10		
Last Filled On:	12 Apr 2012		
Initially Ordered On:	12 Apr 2012		
Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3

MHVZZOED, TEST CONFIDENTIAL Page 11 of 27

VA Medication History

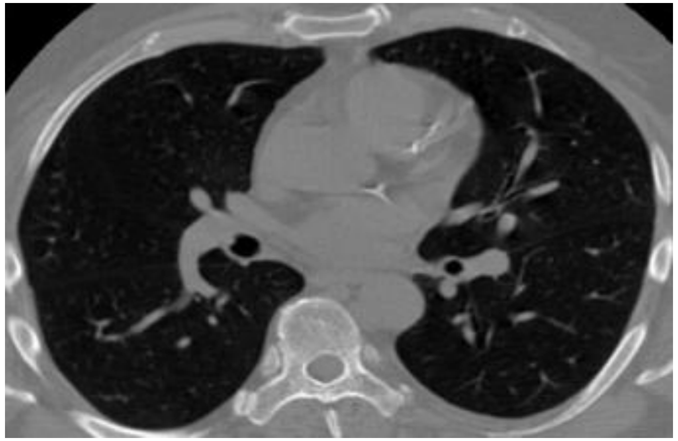
Source:	VA
Last Updated:	18 Jul 2012 @ 0002
Sorted By:	Last Filled On (Descending)

VA Medication History includes up to two years of medication history unless you select a different date range in your download request.

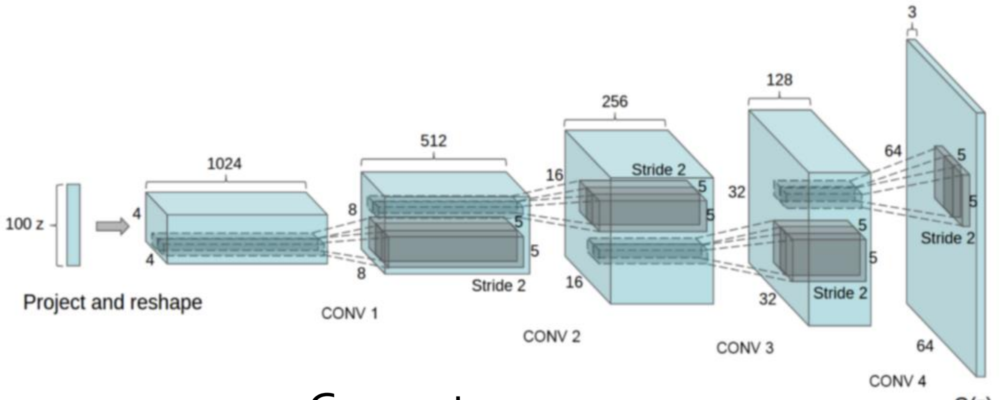
Medication:	DOCUSATE NA 100MG CAP		
Instructions:	TAKE ONE CAPSULE BY MOUTH EVERY DAY		
Status:	Active		
Refills Remaining:	10		
Last Filled On:	12 Apr 2012		
Initially Ordered On:	12 Apr 2012		
Quantity	Days Supply	Pharmacy	Prescription Number
30	30	PORTLAND PHARMACY	11305543A

Medication:	METOPROLOL TARTRATE 50MG TAB
Instructions:	TAKE ONE TABLET BY MOUTH EVERY 12 HOURS
Status:	Active
Refills Remaining:	3

CNN, DNN, RNN



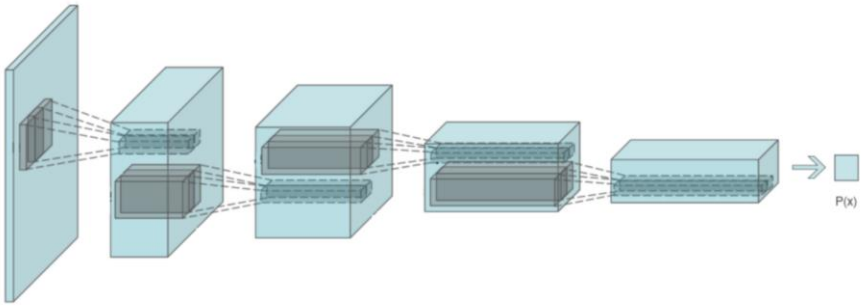
GAN



Generator

Fake

True



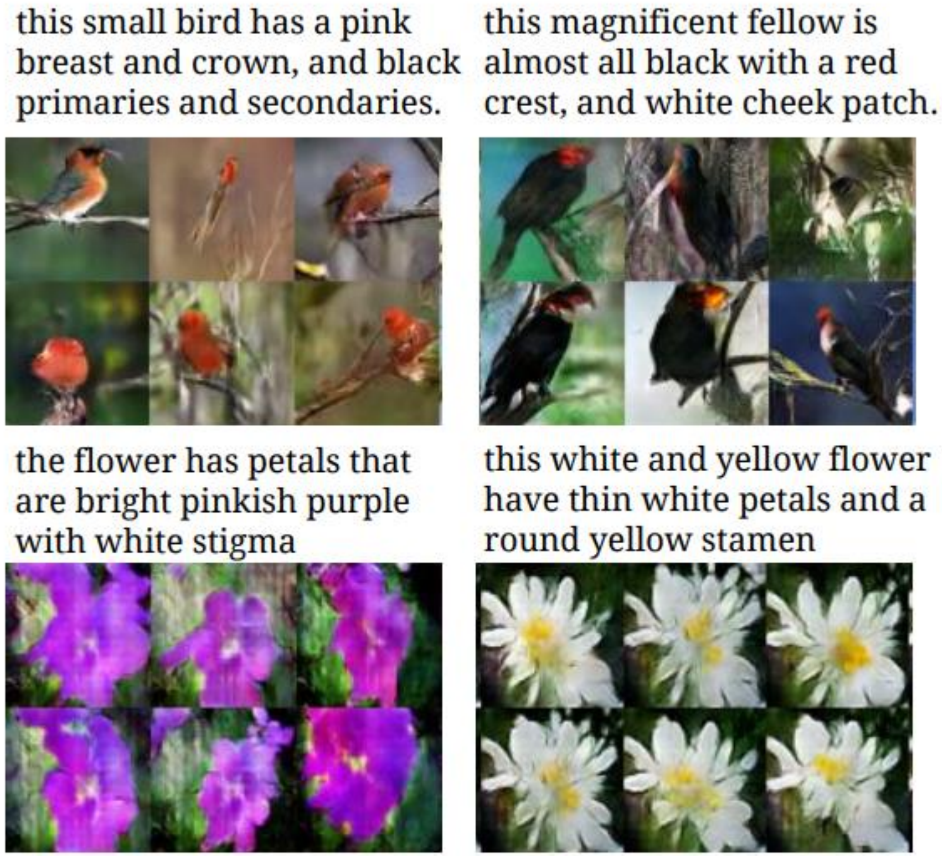
Discriminator

https://arxiv.org/abs/1511.06434

Text to Image

Generative Adversarial Text to Image Synthesis

Scott Reed, Zeynep Akata, Xinchen Yan, Lajanugen Logeswaran
Bernt Schiele, Honglak Lee
[1] University of Michigan, Ann Arbor, MI, USA (UMICH.EDU)



Text to Image

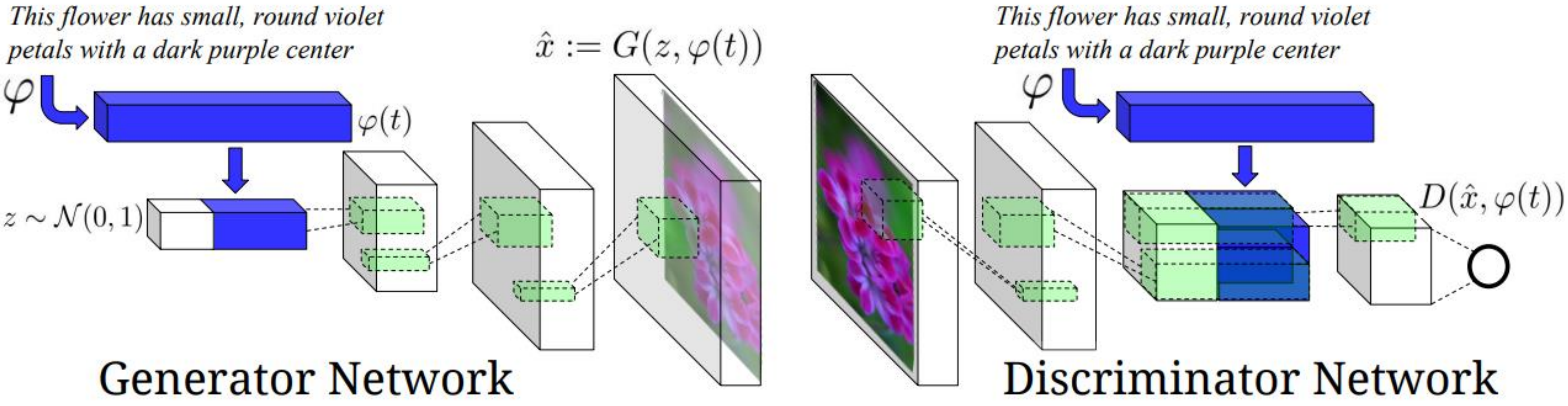
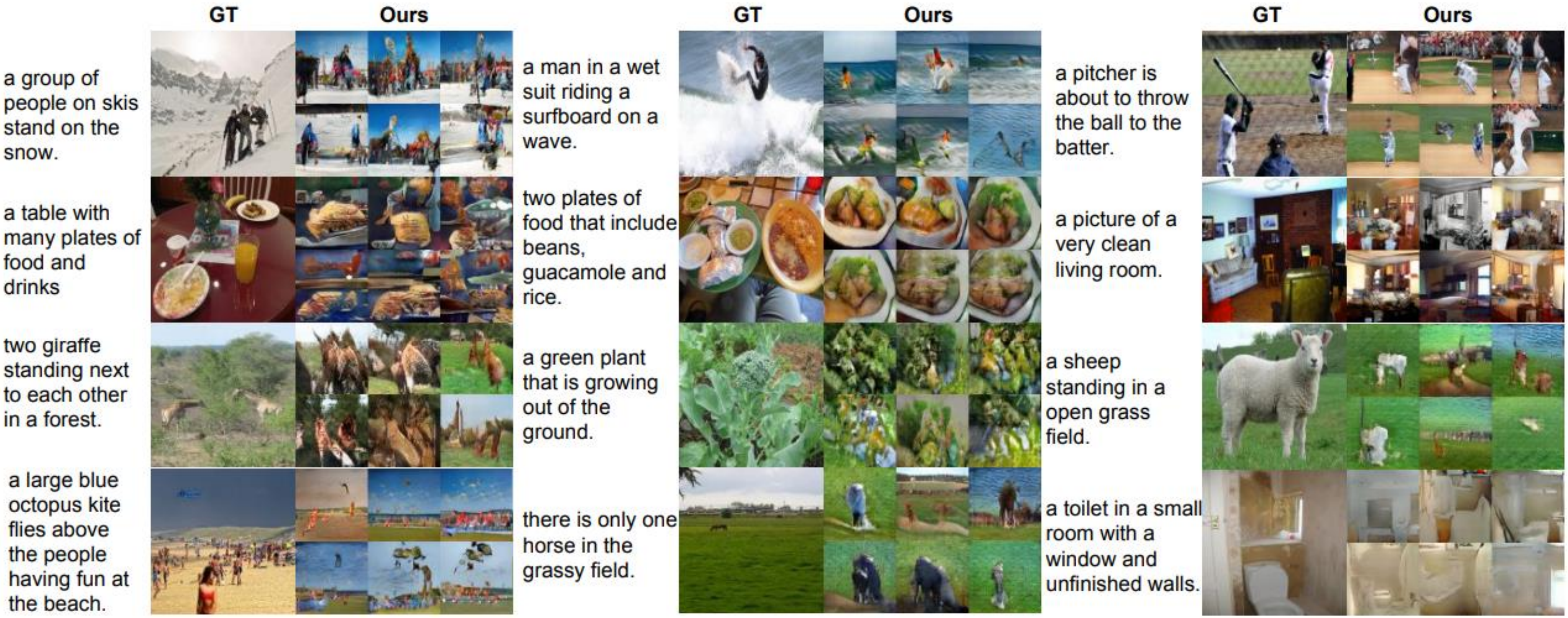


Figure 2. Our text-conditional convolutional GAN architecture. Text encoding $\varphi(t)$ is used by both generator and discriminator. It is projected to a lower-dimensions and depth concatenated with image feature maps for further stages of convolutional processing.

Text to Image



Text to Image

AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks

Tao Xu*[1], Pengchuan Zhang[2], Qiuyuan Huang[2],
Han Zhang[3], Zhe Gan[4], Xiaolei Huang[1], Xiaodong He[5]

[1]Lehigh University [2]Microsoft Research [3]Rutgers University [4]Duke University [5]JD AI Research
{tax313, xih206}@lehigh.edu, {penzhan, qihua, xiaohe}@microsoft.com
han.zhang@cs.rutgers.edu, zhe.gan@duke.edu, xiaodong.he@jd.com

Text to Image

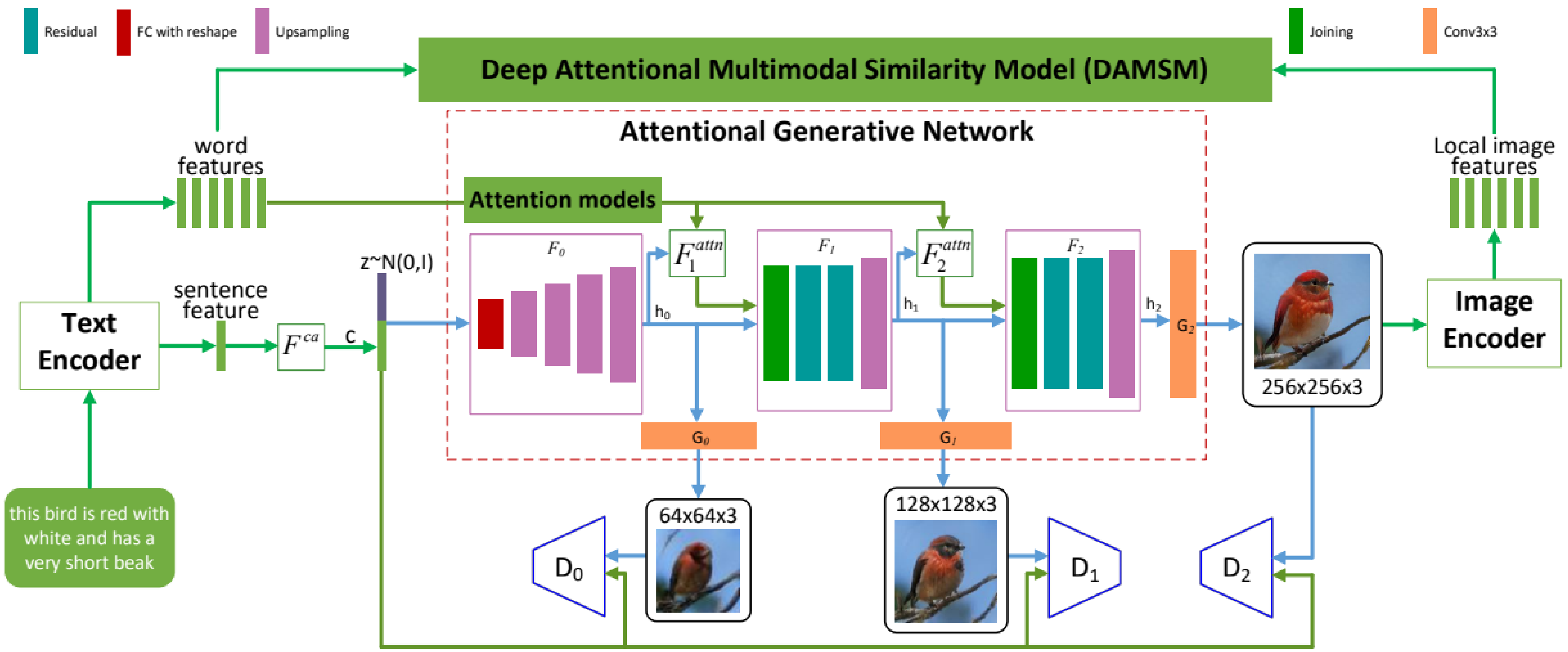


Figure 2. The architecture of the proposed AttnGAN. Each attention model automatically retrieves the conditions (*i.e.*, the most relevant word vectors) for generating different sub-regions of the image; the DAMSM provides the fine-grained image-text matching loss for the generative network.

Text to Image

Text to Image

https://t2i.cvalenzuelab.com/

Integration of Spatial Distribution in Imaging-Genetics

Vaishnavi Subramanian[1(✉)], Weizhao Tang[1], Benjamin Chidester[2], Jian Ma[2], and Minh N. Do[1]

[1] Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA
vs5@illinois.edu

[2] School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA

Abstract. To better understand diseases such as cancer, it is crucial for computational inference to quantify the spatial distribution of various cell types within a tumor. To this end, we used Ripley's K-statistic, which captures the spatial distribution patterns at different scales of both individual point sets and interactions between multiple point sets. We propose to improve the expressivity of histopathology image features by incorporating this descriptor to capture potential cellular interactions, especially interactions between lymphocytes and epithelial cells. We demonstrate the utility of the Ripley's K-statistic by analyzing digital slides from 710 TCGA breast invasive carcinoma (BRCA) patients. In particular, we consider its use in the context of imaging-genetics to understand correlations between gene expression and image features using canonical correlation analysis (CCA). Our analysis shows that including these spatial features leads to more significant associations between image features and gene expression.

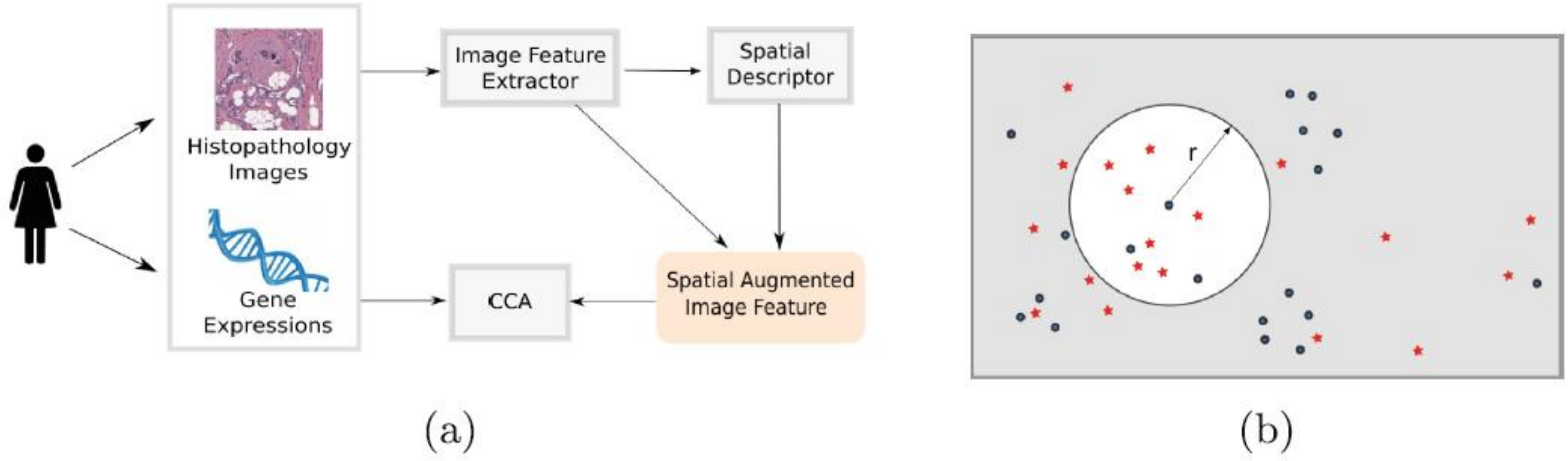
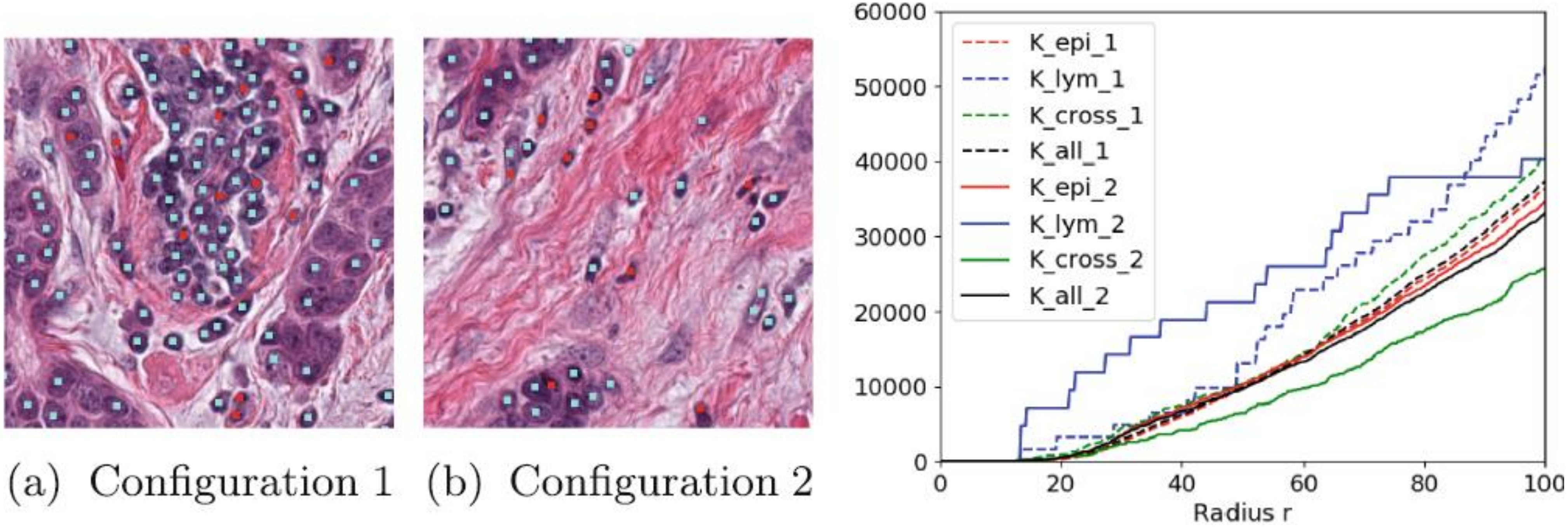
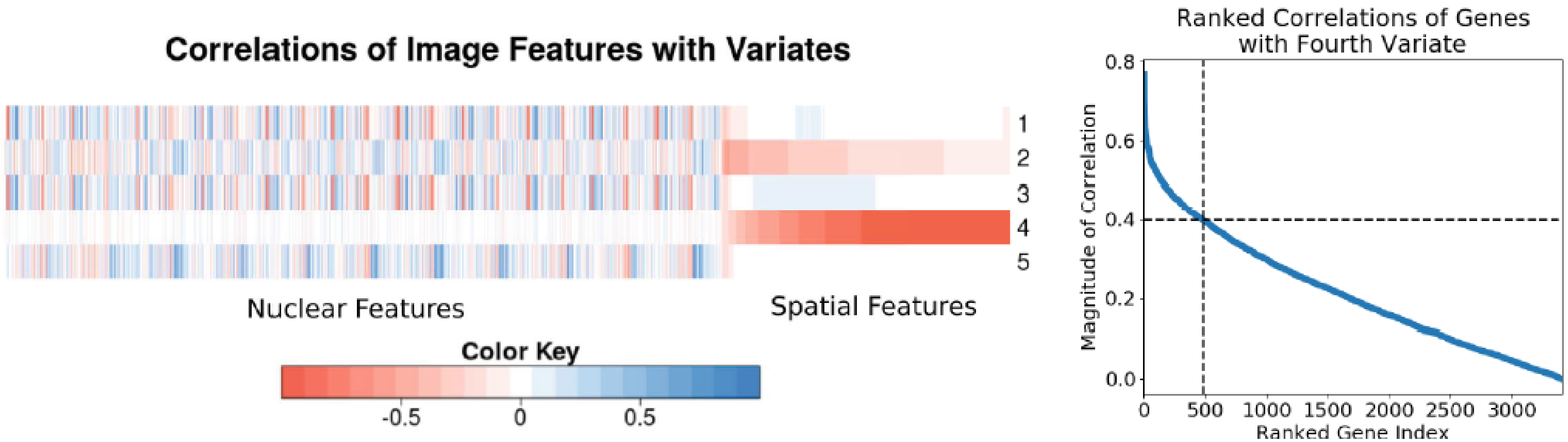


Fig. 1. (a) Our work-flow comprising CNN-based image feature extraction, spatial feature descriptor, and CCA between the features. (b) Pictorial representation of the K-function evaluated at radius r for the blue process, by counting blue points (self K-function) and red stars (cross K-function) within radius r.



(a) Configuration 1 (b) Configuration 2

(c) Corresponding K-functions



(a) Correlation of spatial-augmented image features with variates (K-function evaluated for $r \leq 500$).

(b) Ordered correlations of gene with spatial variate.

Fig. 3. SCCA for gene expression and image features, with spatial features added. (a) The fourth variate has a strong negative correlation with most spatial features. (b) The correlation of gene expression with this variate showed very high correlation with a few genes and then a linear decay in correlation. We chose the transition at the 480th gene, corresponding to a correlation threshold of 0.4, for use in pathway analysis.

TextRay: Mining Clinical Reports to Gain a Broad Understanding of Chest X-Rays

Jonathan Laserson[1(✉)], Christine Dan Lantsman[2], Michal Cohen-Sfady[1], Itamar Tamir[3], Eli Goz[1], Chen Brestel[1], Shir Bar[4], Maya Atar[5], and Eldad Elnekave[1]

[1] Zebra Medical Vision LTD, Shefayim, Israel
jonil@zebra-med.com
[2] Sheba Medical Center and Tel Aviv University, Ramat Gan, Israel
[3] Rabin Medical Center, Petah Tikva, Israel
[4] Technion, Israel Institute of Technology, Haifa, Israel
[5] Ben Gurion University, Beersheba, Israel

Abstract. The chest X-ray (CXR) is by far the most commonly performed radiological examination for screening and diagnosis of many cardiac and pulmonary diseases. There is an immense world-wide shortage of physicians capable of providing rapid and accurate interpretation of this study. A radiologist-driven analysis of over two million CXR reports generated an ontology including the 40 most prevalent pathologies on CXR. By manually tagging a relatively small set of sentences, we were able to construct a training set of 959k studies. A deep learning model was trained to predict the findings given the patient frontal and lateral scans. For 12 of the findings we compare the model performance against a team of radiologists and show that in most cases the radiologists agree on average more with the algorithm than with each other.

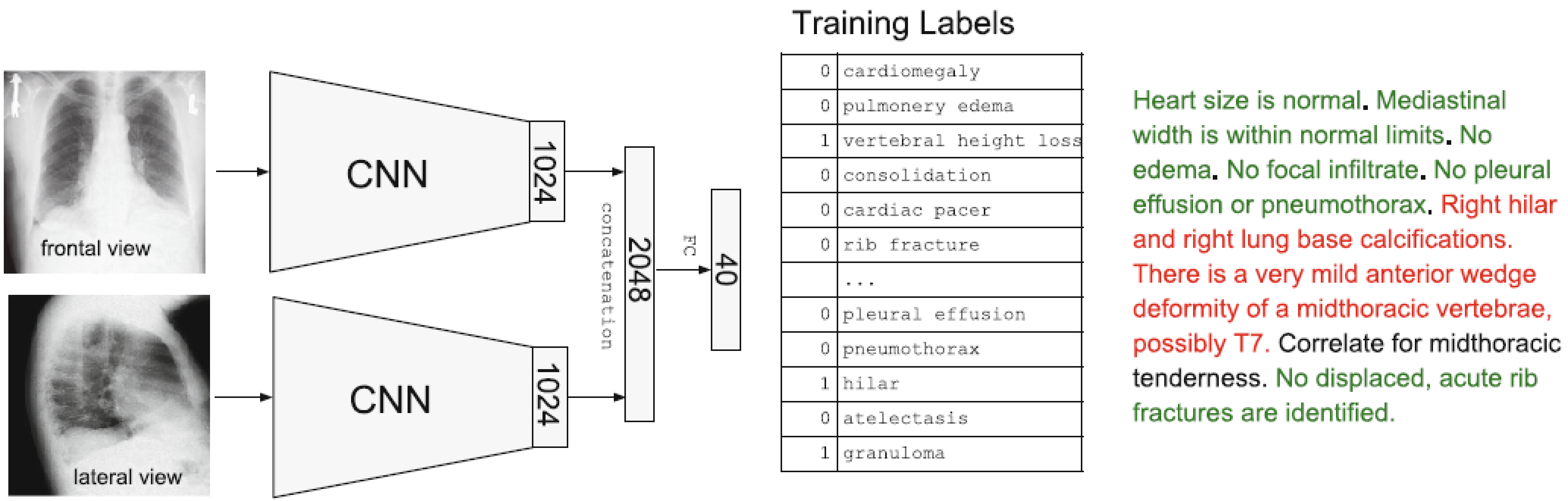


Fig. 1. TextRay Model Illustration. Frontal (PA) and lateral view images each go through a separate CNN. A fully-connected layer is applied on their concatenated feature vectors and emits the confidence for each finding. Training labels were extracted by analyzing the report sentences. Negative (green) and positive (red) sentences identified. Findings in positive sentences receive a positive training label. Negative or unmentioned findings receive a negative label.

Finding	Pool		Avg. agreement w/ rads			Δ (CI)
	pos	neg	Report	Avg. rad	Textray	Textray vs. rads
Pulmonary edema	128	482	0.613	0.639	0.730	+0.09 (0.07, 0.11)
Elevated diaphragm	202	77	0.731	0.675	0.754	+0.08 (0.05, 0.10)
Abnormal aorta	198	80	0.736	0.693	0.771	+0.08 (0.05, 0.11)
Hyperinflation	95	80	0.678	0.619	0.657	+0.04 (−0.02, 0.10)
Vertebral height loss	126	55	0.781	0.742	0.757	+0.02 (−0.02, 0.06)
Atelectasis	201	78	0.778	0.756	0.767	+0.01 (−0.03, 0.04)
Cardiomegaly	238	372	0.755	0.861	0.866	+0.01 (−0.02, 0.03)
Pleural effusion	207	73	0.905	0.893	0.896	+0.00 (−0.02, 0.03)
Consolidation	194	78	0.690	0.730	0.707	−0.02 (−0.07, 0.02)
Pneumothorax	111	124	0.830	0.855	0.823	−0.03 (−0.08, 0.01)
Rib fracture	183	76	0.683	0.799	0.745	−0.05 (−0.10, −0.01)
Hilar prominence	184	426	0.552	0.797	0.736	−0.06 (−0.09, −0.03)

Table 2 shows that TextRay is on par with human radiologists (within the 95% CI) on 10 out of 12 findings, with the exception of *rib fracture* and *hilar prominence*. On some findings (*elevated diaphragm*, *abnormal aorta*, and *pulmonary edema*), radiologists agree significantly more with our algorithm than

Ordinal Multi-modal Feature Selection for Survival Analysis of Early-Stage Renal Cancer

Wei Shao[1], Jun Cheng[2], Liang Sun[1], Zhi Han[4], Qianjin Feng[3], Daoqiang Zhang[1(✉)], and Kun Huang[4(✉)]

College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
dqzhang@nuaa.edu.cn
School of Biomedical Engineering, Shenzhen University, Shenzhen 518073, China
[3] School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China
[4] School of Medicine, Indiana University, Indianapolis, IN 46202, USA
kunhuang@iu.edu

Abstract. Existing studies have demonstrated that combining genomic data and histopathological images can better stratify cancer patients with distinct prognosis than using single biomarker, for different biomarkers may provide complementary information. However, these multi-modal data, most high-dimensional, may contain redundant features that will deteriorate the performance of the prognosis model, and therefore it has become a challenging problem to select the informative features for survival analysis from the redundant and heterogeneous feature groups. Existing feature selection methods assume that the survival information of one patient is independent to another, and thus miss the ordinal relationship among the survival time of different patients. To solve this issue, we make use of the important ordinal survival information among different patients and propose an ordinal sparse canonical correlation analysis (i.e., OSCCA) framework to simultaneously identify important image features and eigengenes for survival analysis. Specifically, we formulate our framework basing on sparse canonical correlation analysis model, which aims at finding the best linear projections so that the highest correlation between the selected image features and eigengenes can be achieved. In addition, we also add constrains to ensure that the ordinal survival information of different patients is preserved after projection. We evaluate the effectiveness of our method on an early-stage renal cell carcinoma dataset. Experimental results demonstrate that the selected features correlated strongly with survival, by which we can achieve better patient stratification than the comparing methods.

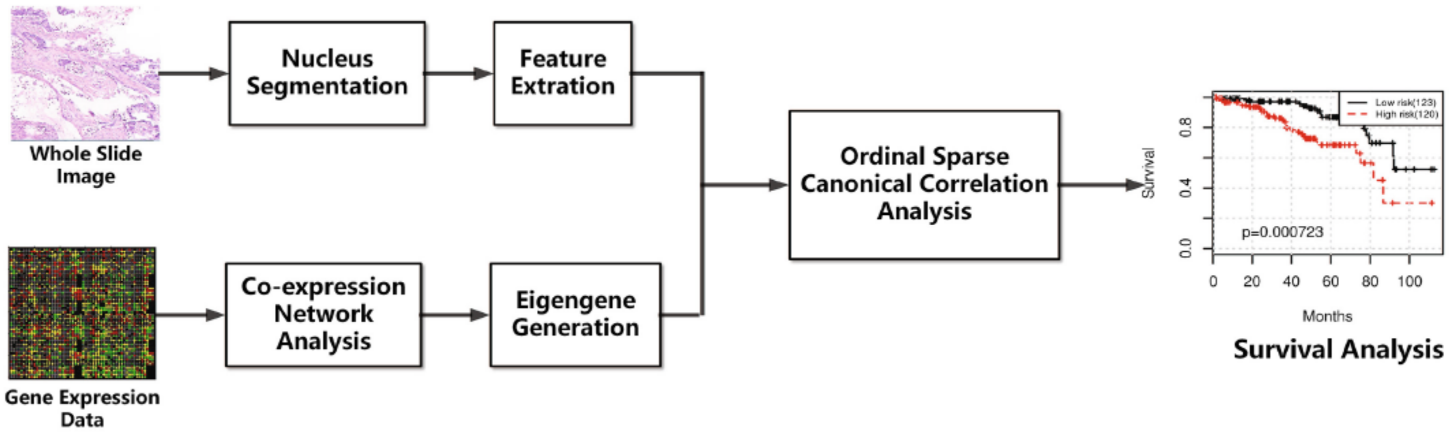


Fig. 1. The flowchart of the proposed method.

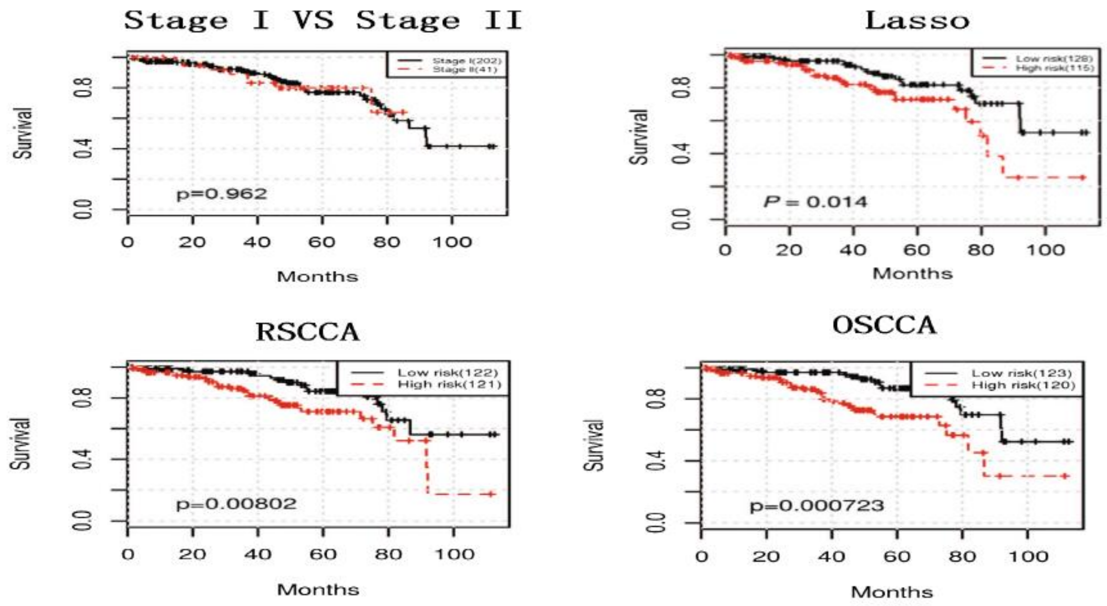


Fig. 2. Comparisons of the survival curves by applying different feature methods.

DeepEM: Deep 3D ConvNets with EM for Weakly Supervised Pulmonary Nodule Detection

Wentao Zhu[1(✉)], Yeeleng S. Vang[1], Yufang Huang[2], and Xiaohui Xie[1]

[1] University of California, Irvine, USA
{wentaoz1,ysvang,xhx}@uci.edu
[2] Lenovo Research, Beijing, China
huangyf8@lenovo.com

Abstract. Recently deep learning has been witnessing widespread adoption in various medical image applications. However, training complex deep neural nets requires large-scale datasets labeled with ground truth, which are often unavailable in many medical image domains. For instance, to train a deep neural net to detect pulmonary nodules in lung computed tomography (CT) images, current practice is to manually label nodule locations and sizes in many CT images to construct a sufficiently large training dataset, which is costly and difficult to scale. On the other hand, electronic medical records (EMR) contain plenty of partial information on the content of each medical image. In this work, we explore how to tap this vast, but currently unexplored data source to improve pulmonary nodule detection. We propose DeepEM, a novel deep 3D ConvNet framework augmented with expectation-maximization (EM), to mine weakly supervised labels in EMRs for pulmonary nodule detection. Experimental results show that DeepEM can lead to 1.5% and 3.9% average improvement in free-response receiver operating characteristic (FROC) scores on LUNA16 and Tianchi datasets, respectively, demonstrating the utility of incomplete information in EMRs for improving deep learning algorithms (https://github.com/uci-cbcl/DeepEM-for-Weakly-Supervised-Detection.git).

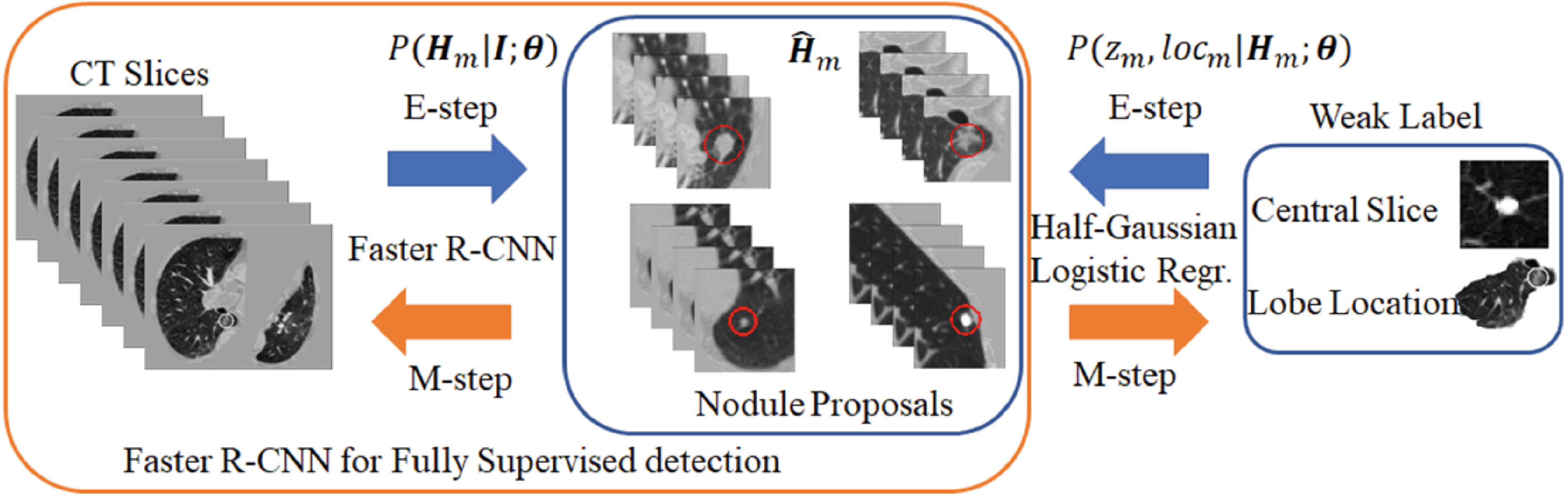


Fig. 1. Illustration of DeepEM framework. Faster R-CNN is employed for nodule proposal generation. Half-Gaussian model and logistic regression are employed for central slice and lobe location respectively. In the E-step, we utilize all the observations, CT slices, and weak label to infer the latent variable, nodule proposals, by maximum a posteriori (MAP) or sampling. In the M-step, we employ the estimated proposals to update parameters in the Faster R-CNN and logistic regression.

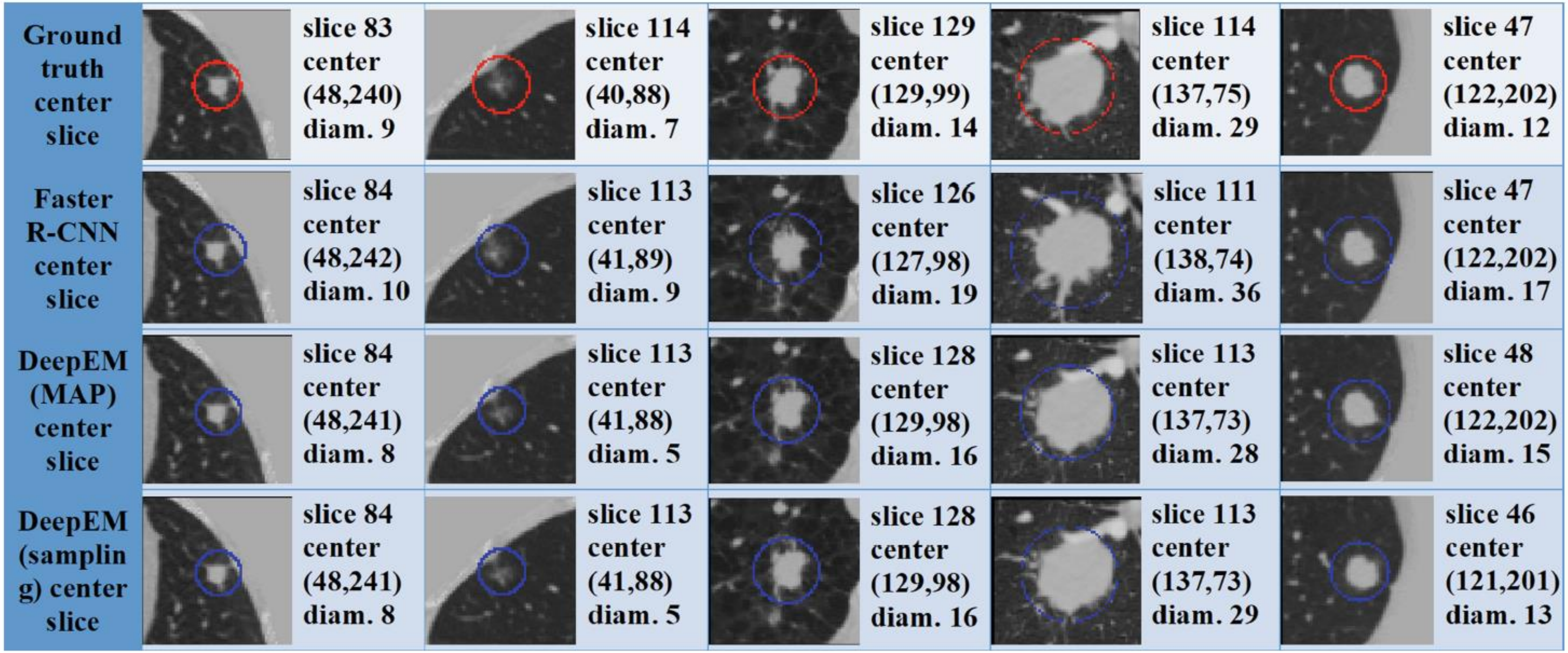


Fig. 3. Detection visual comparison among Faster R-CNN [14], DeepEM with MAP and DeepEM with Sampling on nodules randomly sampled from Tianchi. DeepEM provides more accurate detection (central slice, center and diameter) than Faster R-CNN.

Towards Automatic Report Generation in Spine Radiology Using Weakly Supervised Framework

Zhongyi Han[1,2], Benzheng Wei[1,2(✉)], Stephanie Leung[3,4], Jonathan Chung[3,4], and Shuo Li[3,4(✉)]

[1] College of Science and Technology, Shandong University of Traditional Chinese Medicine, Jinan, Shandong, China
wbz99@sina.com
[2] Computational Medicine Lab (CML), Shandong University of Traditional Chinese Medicine, Jinan, Shandong, China
[3] Department of Medical Imaging, Western Univeristy, London, ON, Canada
slishuo@gmail.com
[4] Digital Imaging Group (DIG), London, ON, Canada

Abstract. The objective of this work is to automatically generate unified reports of lumbar spinal MRIs in the field of radiology, i.e., given an MRI of a lumbar spine, directly generate a radiologist-level report to support clinical decision making. We show that this can be achieved via a weakly supervised framework that combines deep learning and symbolic program synthesis theory to overcome four inevitable tasks: semantic segmentation, radiological classification, positional labeling, and structural captioning. The weakly supervised framework using object level annotations without requiring radiologist-level report annotations to generate unified reports. Each generated report covers almost type lumbar structures comprised of six intervertebral discs, six neural foramina, and five lumbar vertebrae. The contents of each report contain the exact locations and pathological correlations of these lumbar structures as well as their normalities in terms of three type relevant spinal diseases: intervertebral disc degeneration, neural foraminal stenosis, and lumbar vertebrae deformities. This framework is applied to a large corpus of T1/T2-weighted sagittal MRIs of 253 subjects acquired from multiple vendors. Extensive experiments demonstrate that the framework is able to generate unified radiological reports, which reveals its effectiveness and potential as a clinical tool to relieve spinal radiologists from laborious workloads to a certain extent, such that contributes to relevant time savings and expedites the initiation of many specific therapies.

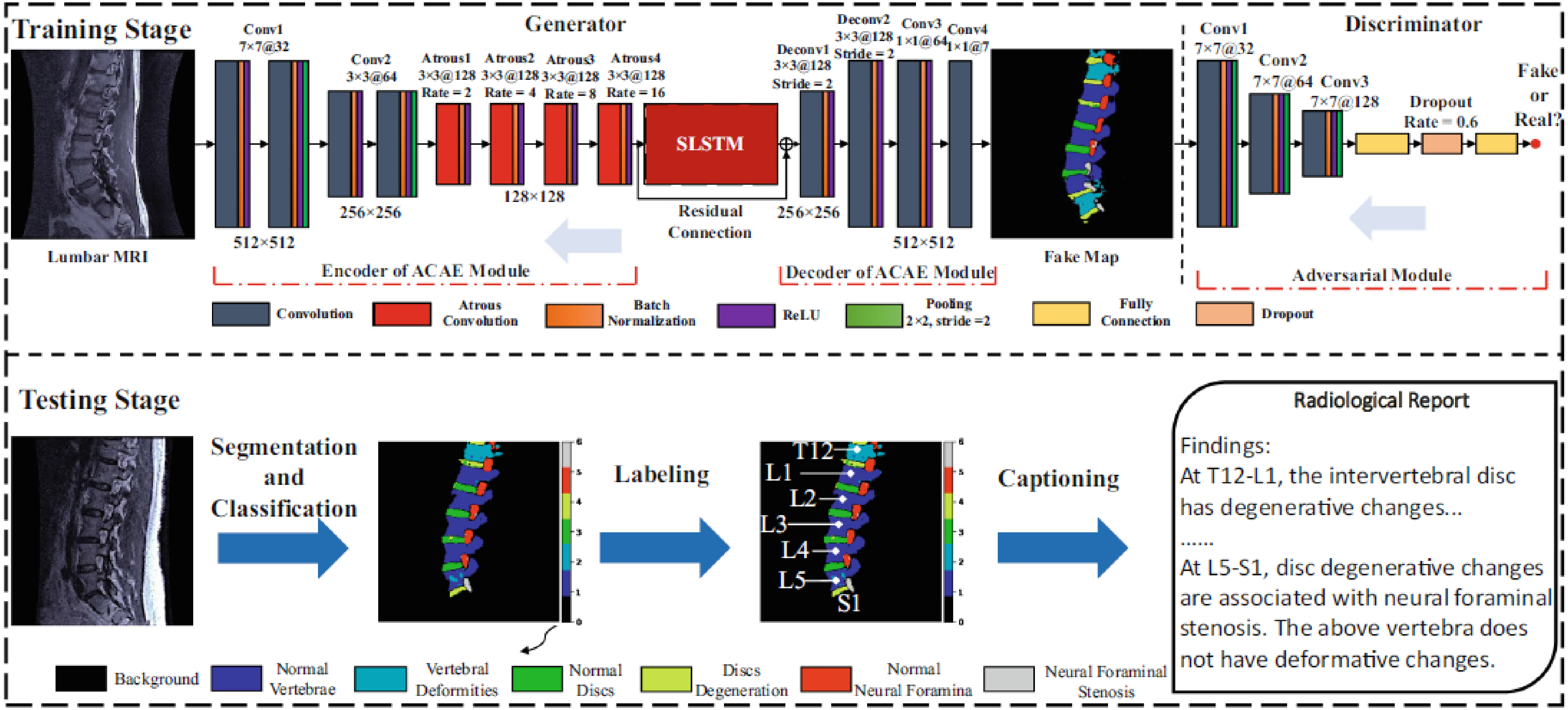
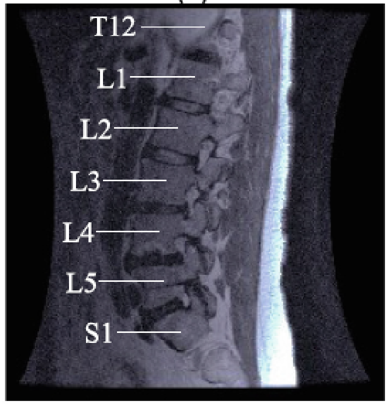


Fig. 1. The workflow of the proposed weakly supervised framework.

(a)



(a) Radiological Report

At T12-L1, the intervertebral disc has obvious degenerative changes. The neural foramen does not have stenosis.

At L1-L2, the above vertebra has deformative changes. The intervertebral disc does not have obvious degenerative changes. The neural foramen does not have obvious stenosis.

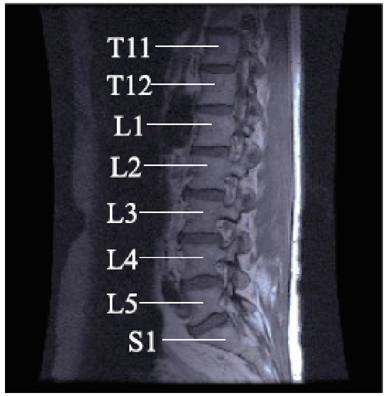
At L2-L3, the neural foramen has obvious stenosis. The intervertebral disc does not have obvious degenerative changes. The above vertebra does not have deformative changes.

At L3-L4, disc degenerative changes are associated with neural foraminal stenosis.

At L4-L5, disc degenerative changes are associated with neural foraminal stenosis.

At L5-S1, the intervertebral disc has obvious degenerative changes. The above vertebra also has deformative changes. They lead to the neural foraminal stenosis.

(b)



(b) Radiological Report

At T12-L1, the intervertebral disc has obvious degenerative changes. The neural foramen does not have stenosis.

At L1-L2, the above vertebra does not have deformative changes. The intervertebral disc does not have degenerative changes. The neural foramen also does not have stenosis.

At L2-L3, the above vertebra does not have deformative changes. The intervertebral disc does not have degenerative changes. The neural foramen also does not have stenosis.

At L3-L4, the neural foramen has obvious stenosis. The intervertebral disc does not have obvious degenerative changes. The above vertebra does not have deformative changes.

At L4-L5, disc degenerative changes are associated with neural foraminal stenosis.

At L5-S1, the intervertebral disc has obvious degenerative changes. The above vertebra also has deformative changes. They lead to neural foraminal stenosis to a certain extent.

Fig. 2. The generated radiological reports. The text in purple color represents that our framework is helpful for building comprehensive pathological analysis.

Accurate Weakly-Supervised Deep Lesion Segmentation Using Large-Scale Clinical Annotations: Slice-Propagated 3D Mask Generation from 2D RECIST

Jinzheng Cai[1,2(✉)], Youbao Tang[1], Le Lu[1], Adam P. Harrison[1], Ke Yan[1], Jing Xiao[3], Lin Yang[2], and Ronald M. Summers[1]

[1] National Institutes of Health Clinical Center, Bethesda, MD 20892, USA
{youbao.tang,le.lu,adam.harrison,ke.yan,rms}@nih.gov
[2] University of Florida, Gainesville, FL 32611, USA
jimmycai@ufl.edu, lin.yang@bme.ufl.edu
[3] Ping An Insurance (Group) Company of China, Ltd., Shenzhen 510852, People's Republic of China
xiaojing661@pingan.com.cn

Abstract. Volumetric lesion segmentation from computed tomography (CT) images is a powerful means to precisely assess multiple time-point lesion/tumor changes. However, because manual 3D segmentation is prohibitively time consuming, current practices rely on an imprecise surrogate called response evaluation criteria in solid tumors (RECIST). Despite their coarseness, RECIST markers are commonly found in current hospital picture and archiving systems (PACS), meaning they can provide a potentially powerful, yet extraordinarily challenging, source of weak supervision for full 3D segmentation. Toward this end, we introduce a convolutional neural network (CNN) based weakly supervised slice-propagated segmentation (WSSS) method to (1) generate the initial lesion segmentation on the axial RECIST-slice; (2) learn the data distribution on RECIST-slices; (3) extrapolate to segment the whole lesion slice by slice to finally obtain a volumetric segmentation. To validate the proposed method, we first test its performance on a fully annotated lymph node dataset, where WSSS performs comparably to its fully supervised counterparts. We then test on a comprehensive lesion dataset with 32,735 RECIST marks, where we report a mean Dice score of 92% on RECIST-marked slices and 76% on the entire 3D volumes.

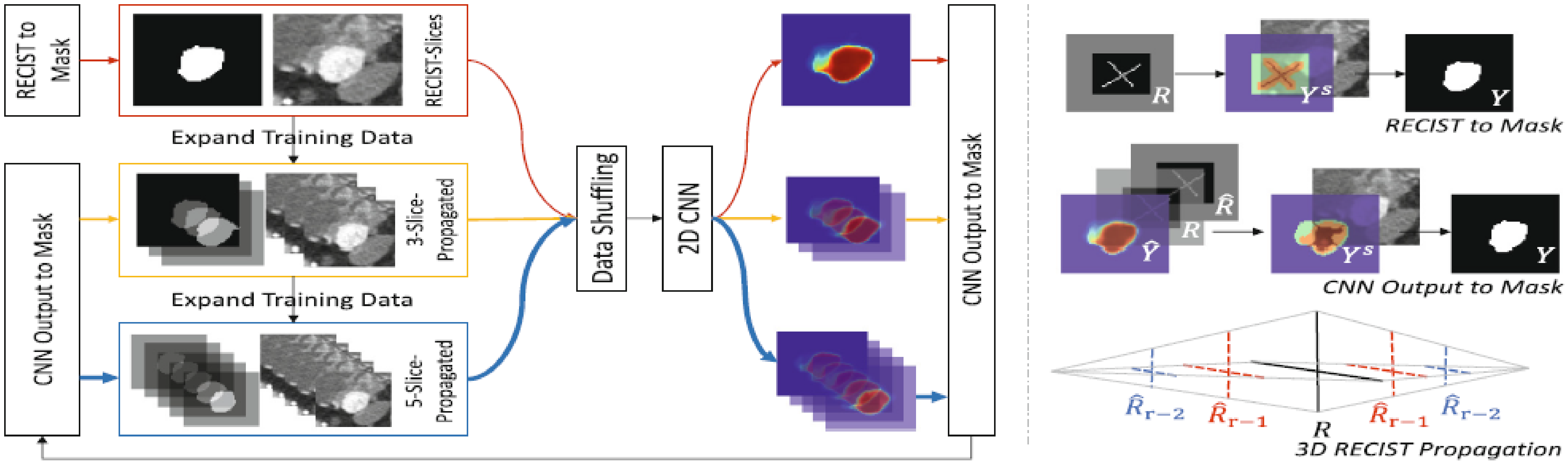


Fig. 1. Overview of the proposed method. **Right:** we use CNN outputs to gradually generate extra training data for lesion segmentation. Arrows colored in red, orange, and blue indicate slice-propagated training at its 1^{st}, 2^{nd}, and 3^{rd} steps, respectively. **Left:** regions colored with red, orange, green, and blue inside the initial segmentation mask Y present FG, PFG, PBG, and BG, respectively.