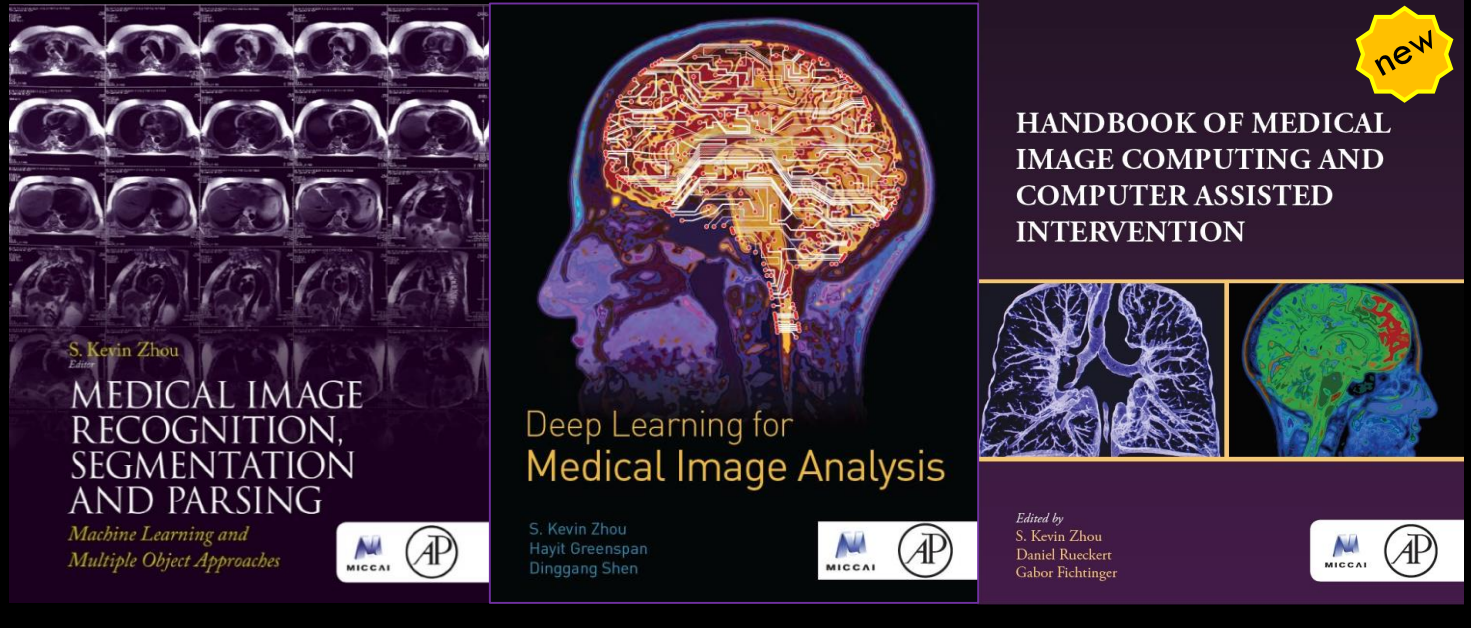
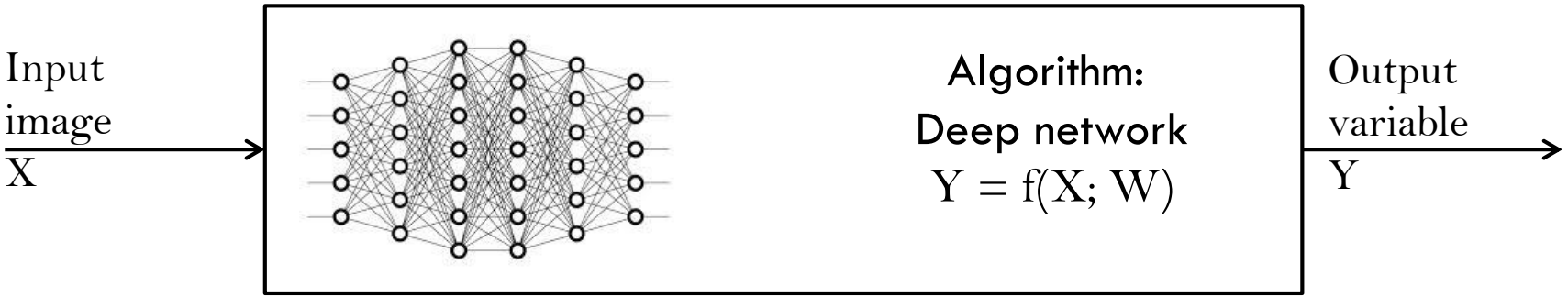
TOWARDS CREATING A 'KNOWLEDGE' GAP FOR DEEP LEARNING BASED MEDICAL IMAGE ANALYSIS



Dr. S. Kevin Zhou, Chinese Academy of Sciences

Deep learning



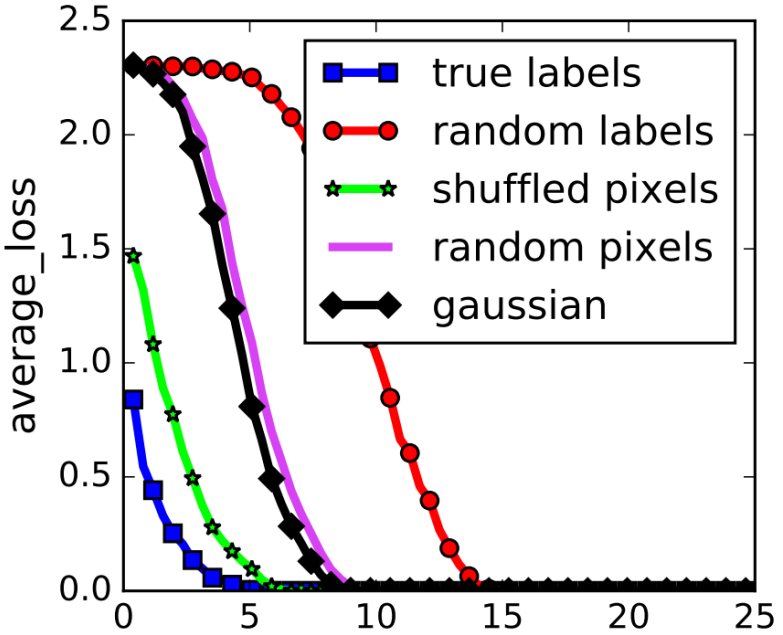
Learning:

$$\arg\min_W \Sigma_i \text{Loss}(Y_i, f(X_i; W)) + \text{Reg}(W)$$

Deep neural net = “super memorizer”

举’三’反一

- “state-of-the-art convolutional networks for image classification trained with stochastic gradient methods *easily fit a random labeling* of the training data.”



[Zhang et al. ICLR2017]

Deep neural nets = “super energy sucker” 以暴制人

- “AlphaGo consumed *~50,000x more energy* than Lee Sedol.”

AlphaGo v.s. in terms of watts human brain

Deep neural nets are overly parameterized
化简为繁

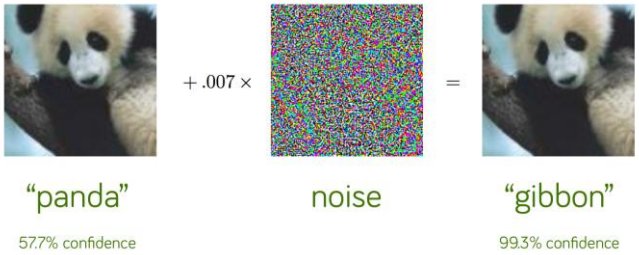
- It is possible to *'compress'* a deep network while maintaining similar accuracy

AlexNet

SqueezeNet
(50x less weights)
MobileNet, ShuffleNet

Adversarial learning & attacks
以假乱真

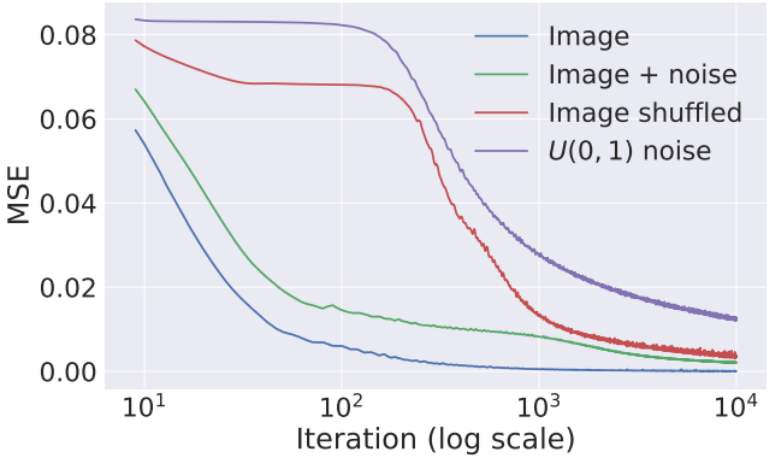
StyleGAN, CVPR'19



Explaining and Harnessing Adversarial Examples, arxiv 1412.6572

The learning process itself
先略后详

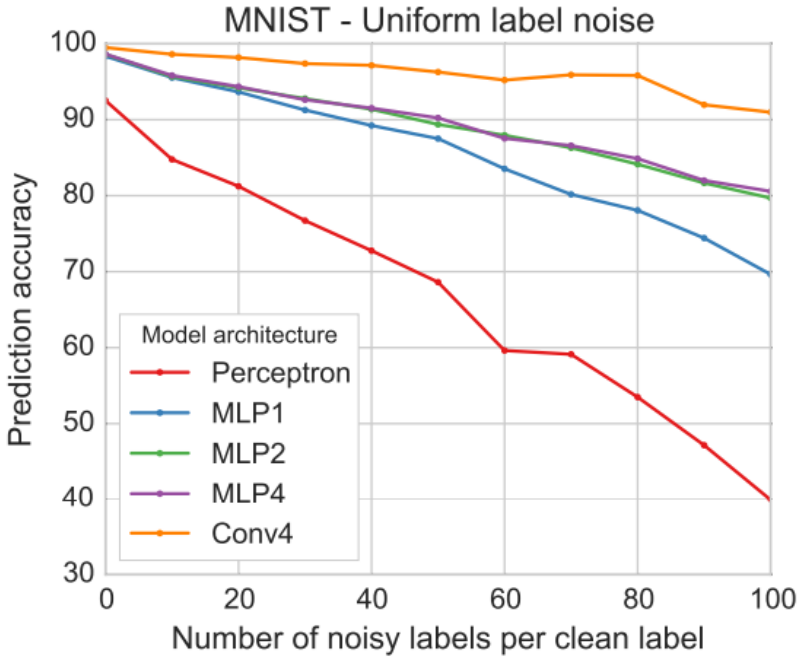
- Learning/fitting seems to proceed
  - from ‘easy’ to ‘difficult’ or
  - from ‘smooth’ to ‘noisy’
- Early stop



Deep image prior (arxiv 1711.10925)

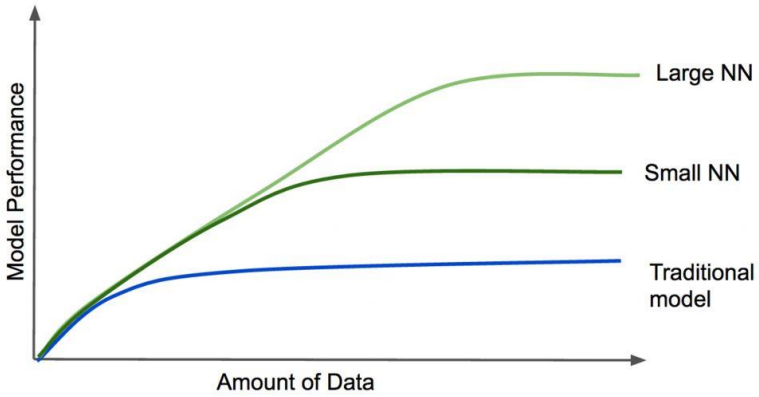
Robust to massive label noise
去芜存菁

- “Learning is robust to an essentially *arbitrary amount of label noise,* provided that the number of clean labels is sufficiently large”



arxiv 1705.10694

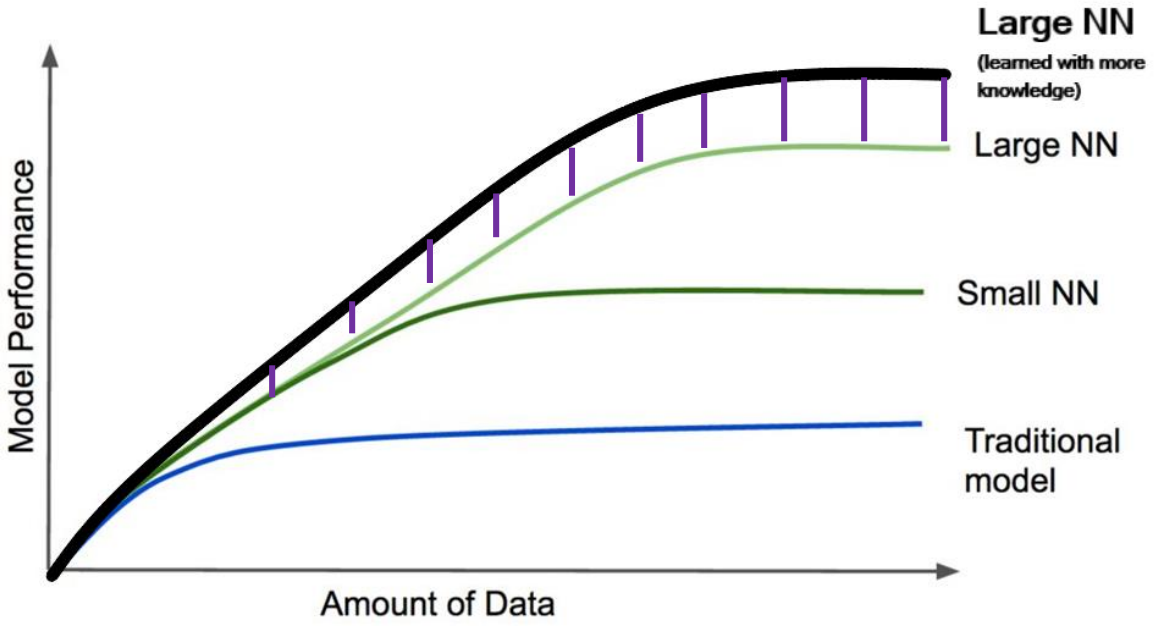
Performance vs amount of data



Recipe for performance improvement:

- Increasing data
- Increasing model capacity
- Repeat the above

Creating a 'knowledge gap'

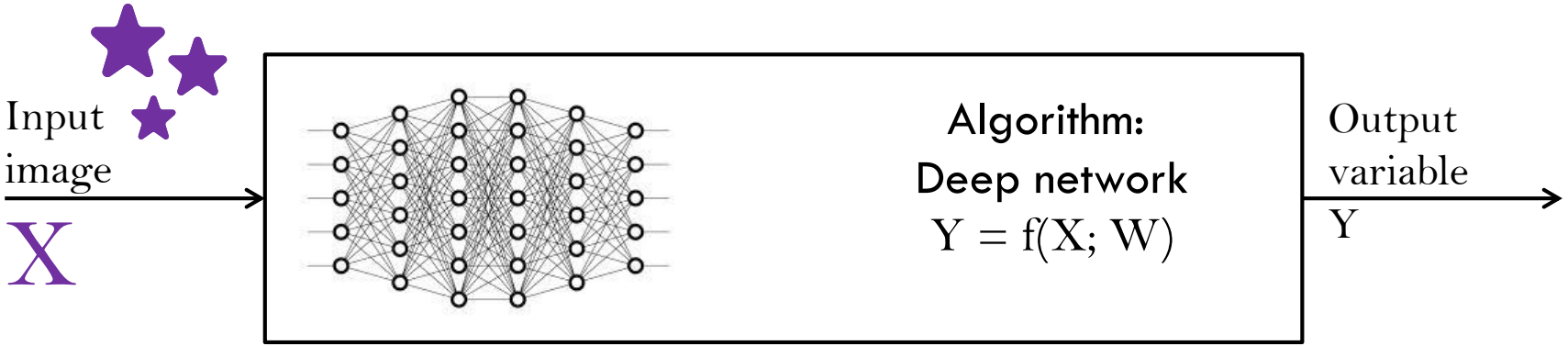


Deep learning with knowledge fusion

Knowledge fusion

- Input
- Output
- Algorithm

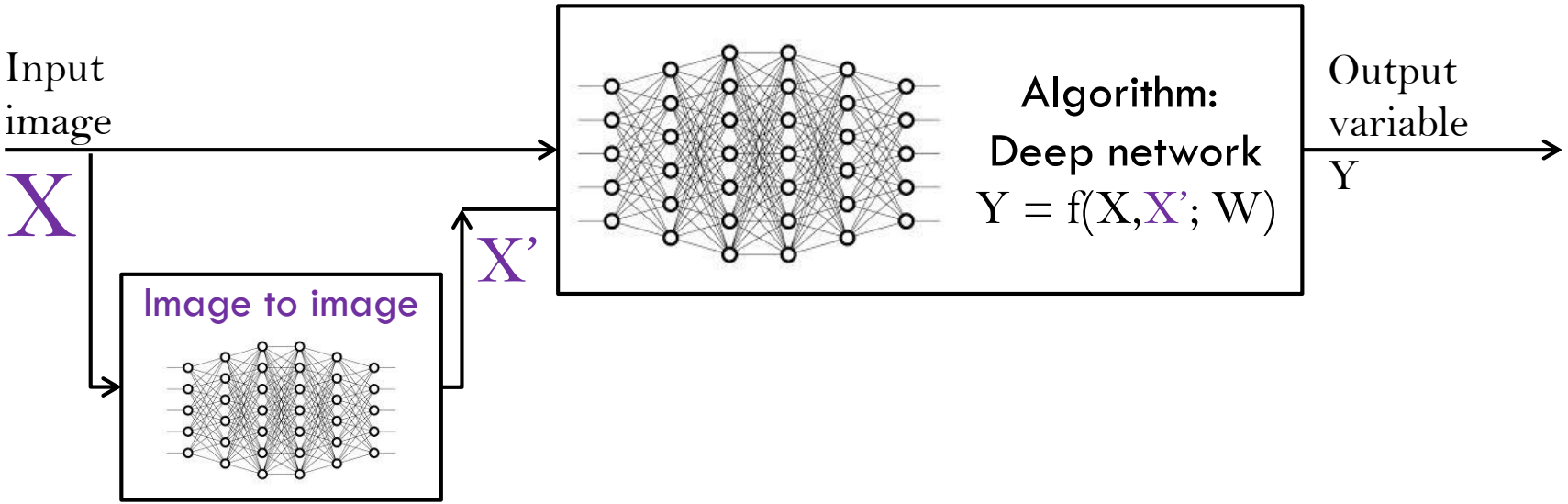
Knowledge in input



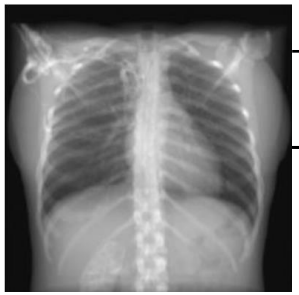
Knowledge in input

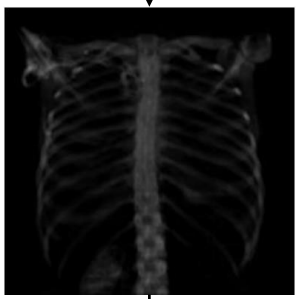
- Multi-modal inputs (RGBD, MR T1+T2, etc.)
- Synthesized inputs
- Other inputs

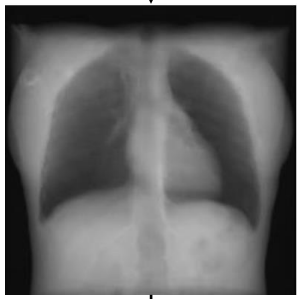
Synthesized inputs



Xray image decomposition and diagnosis







decomposition

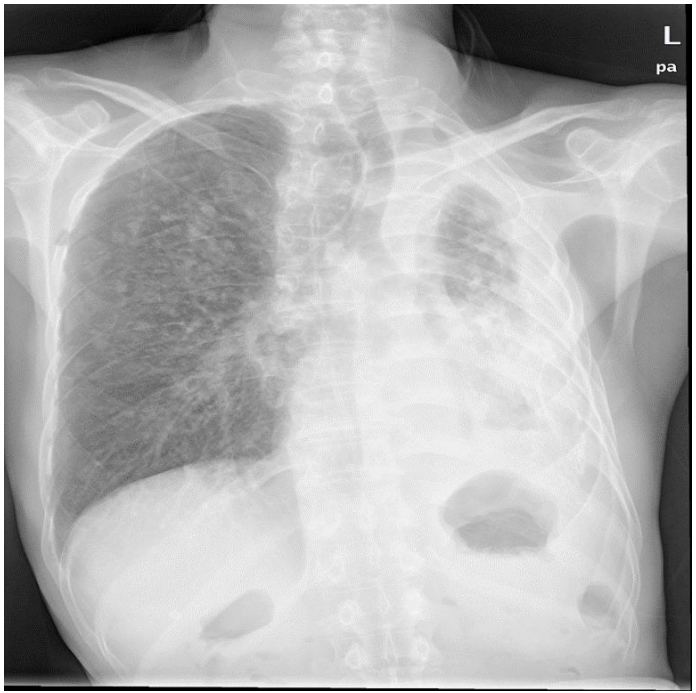
diagnosis

DNN

State-of-the-art accuracy in predicting **11 out of 14** common lung diseases based on Chest-xray14 dataset

Li et al., Encoding CT Anatomy Knowledge for Unpaired Chest X-ray Image Decomposition, MICCAI 2019. (patent pending)

Clinical evaluation



- Reading based on

  (i) the original & bone free images

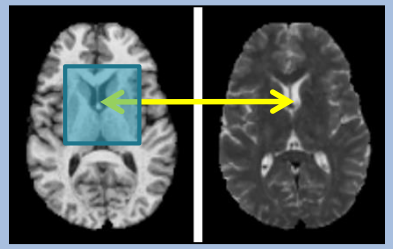
  (ii) only the original image

- Diagnosis accuracy ⬆
  + 8%
- Reading time ⬇
  - 27%

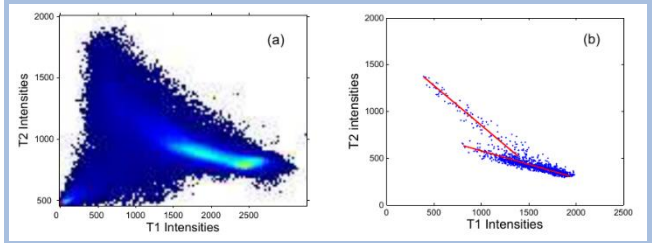
Joint work with Peking Union Medical College.

Supervised cross-domain image synthesis using location-sensitive deep network (LSDN) [MICCAI'2015]

Cross-domain image synthesis



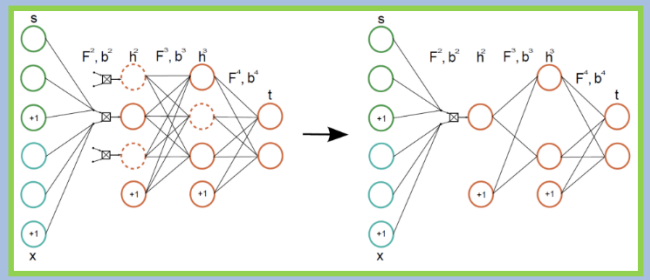
The importance of spatial info.



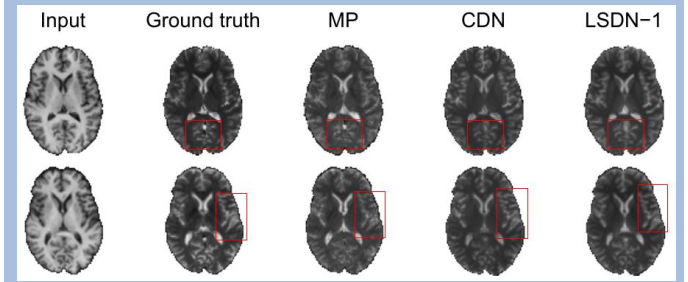
Whole image

Small region 10^3 voxels

Location-sensitive deep network (LSDN)

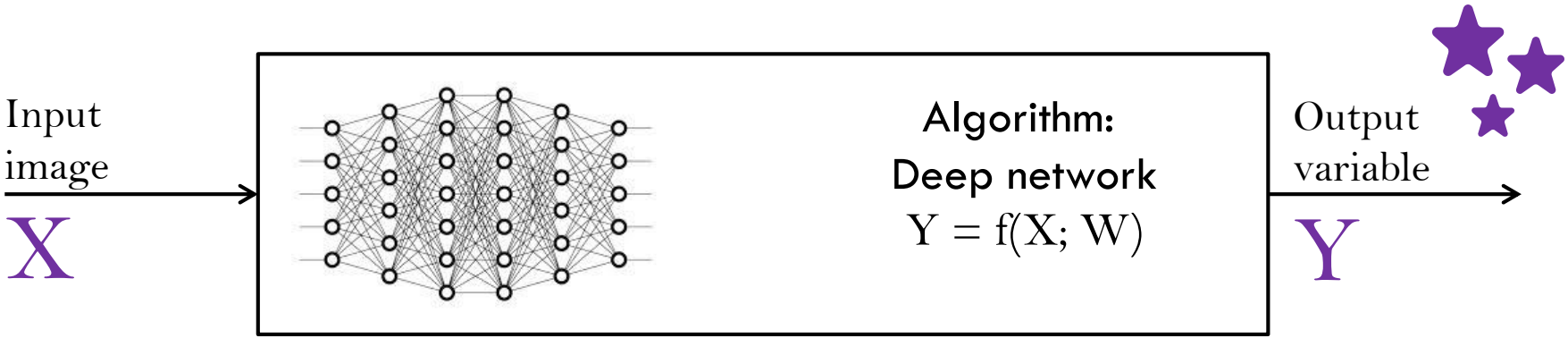


Accurate result



Nguyen, et al. Cross-Domain Synthesis of Medical Images Using Efficient Location-Sensitive Deep Network, MICCAI 2015.
Vemulapalli, et al. Unsupervised Cross-modal Synthesis of Subject-specific Scans, ICCV 2015.

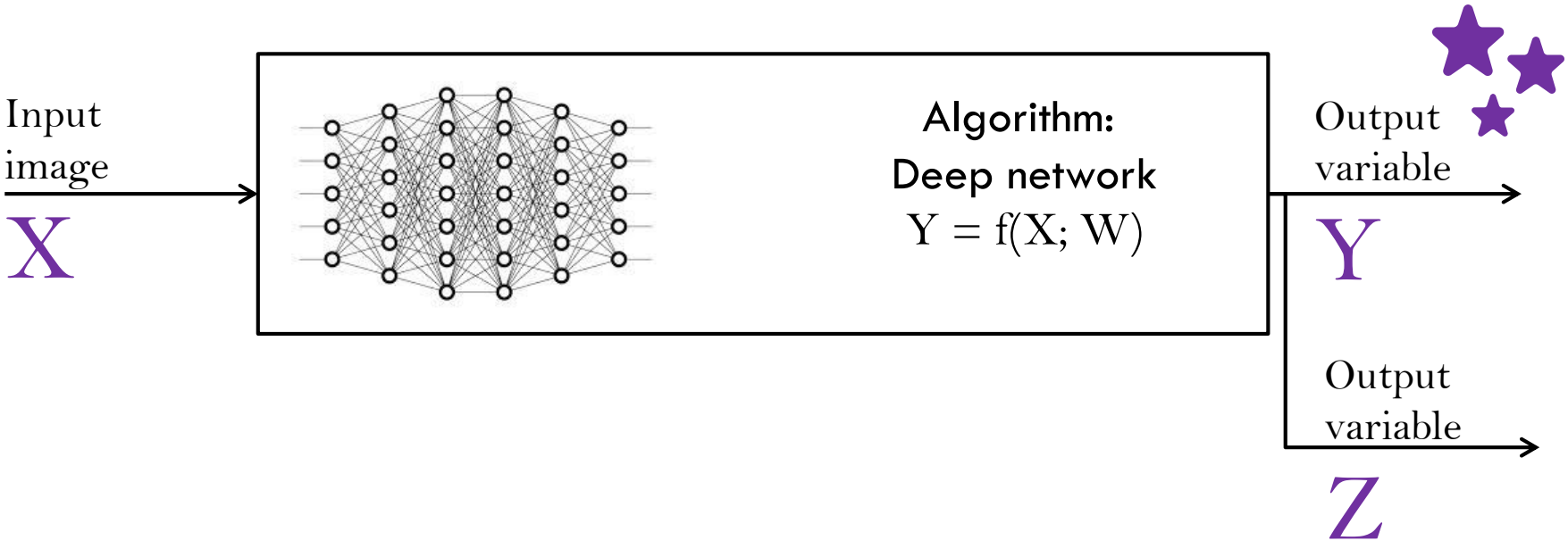
Knowledge in output



Knowledge in output

- Multitask learning
- New representation
- More priors

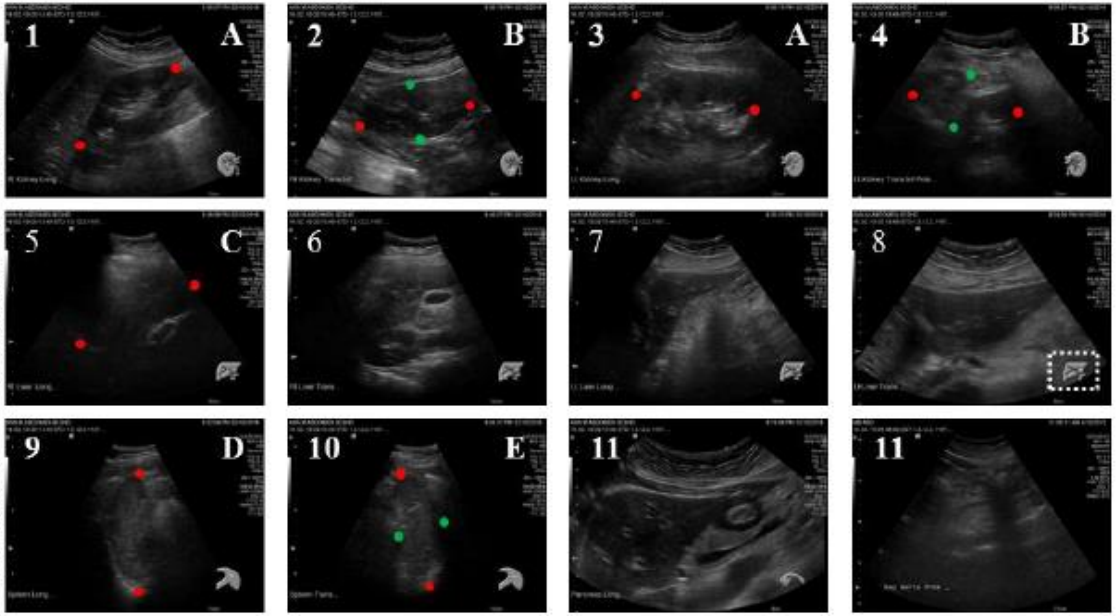
Multitask learning



View classification and landmark detection for abdominal ultrasound images

View Classification

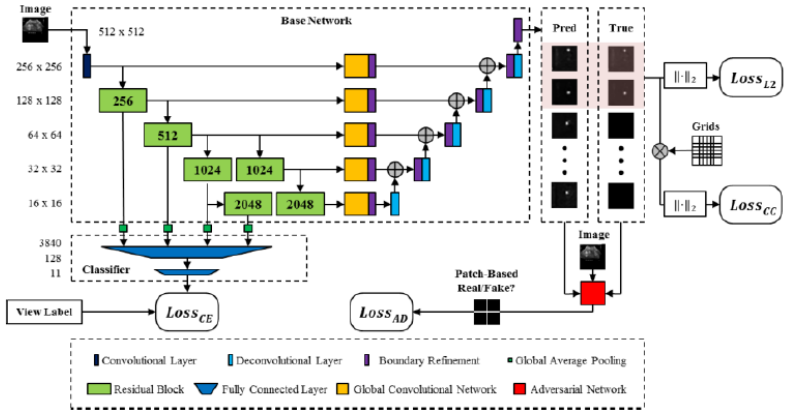
1. Kidney Right Long
2. Kidney Right Trans
3. Kidney Left Long
4. Kidney Left Trans
5. Liver Right Long
6. Liver Right Trans
7. Liver Left Long
8. Liver Left Trans
9. Spleen Long
10. Spleen Trans
11. Others



Landmark Detection

A. Kidney Long **B**. Kidney Trans **C**. Liver Long **D**. Spleen Long **E**. Spleen Trans

Simultaneous view classification and landmark detection for abdominal ultrasound images



View classification

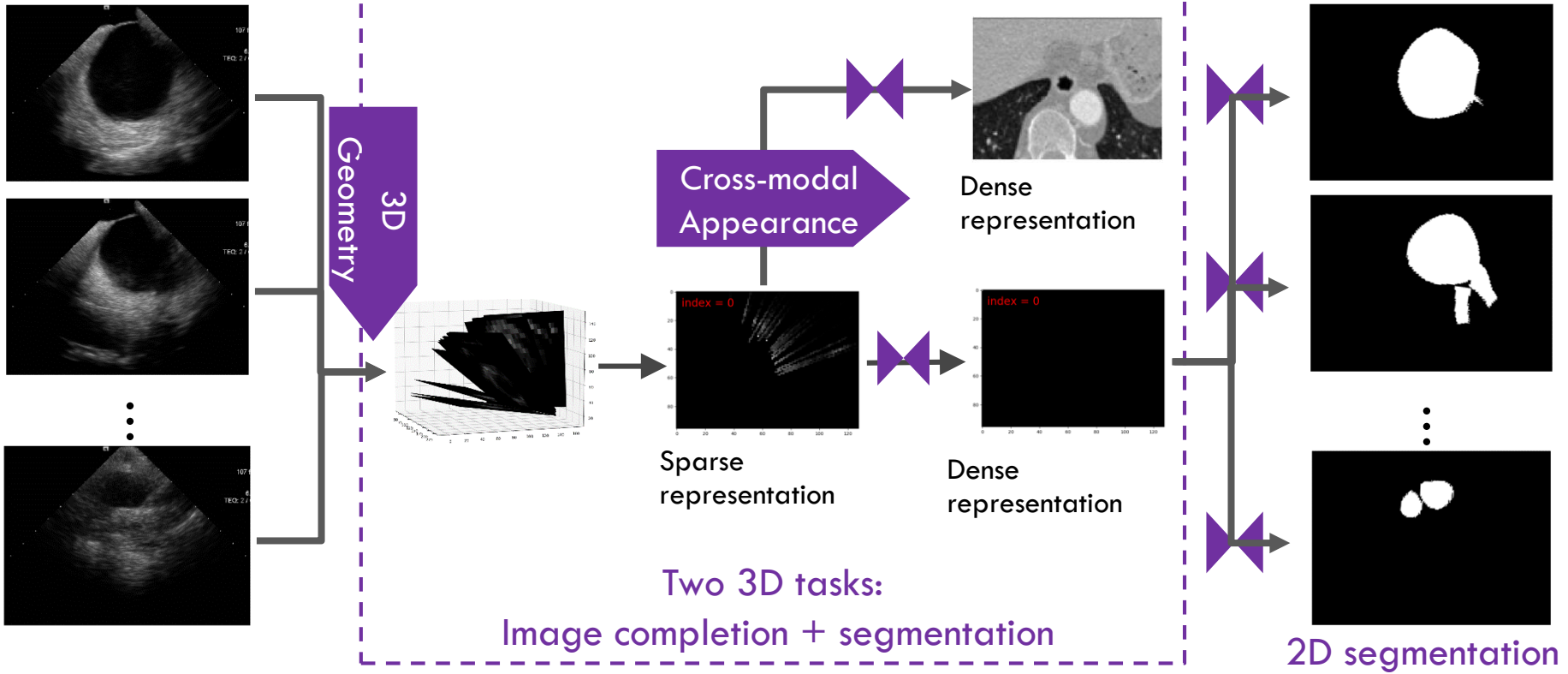
- MTL: 85.29%, STL: 81.22%, Human: 78.87%

Measurement

	KL_LA	KT_LA	KT_SA	LL_LA	SL_LA	ST_LA	ST_SA
Human	4.500	5.431	4.283	5.687	6.104	4.578	4.543
PBT [11]	11.036	9.147	8.393	11.083	7.289	9.359	12.308
SFCN	7.044	7.332	5.189	10.731	8.693	91.309	43.773
MGCN_R	**4.278**	**4.426**	**3.437**	**6.989**	**3.61**	**7.923**	**7.224**

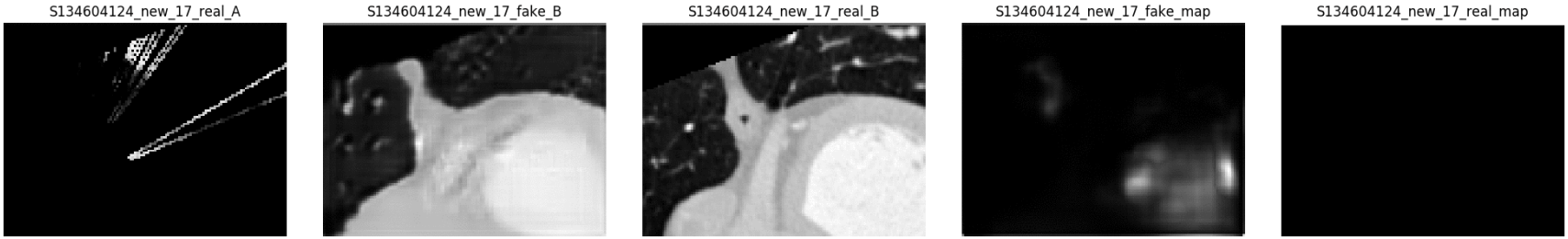
Xu et al., Less is More: Simultaneous View Classification and Landmark Detection for Abdominal Ultrasound Images, MICCAI 2018.

Intra-cardio echo (ICE) auto contouring



Results

index = 0

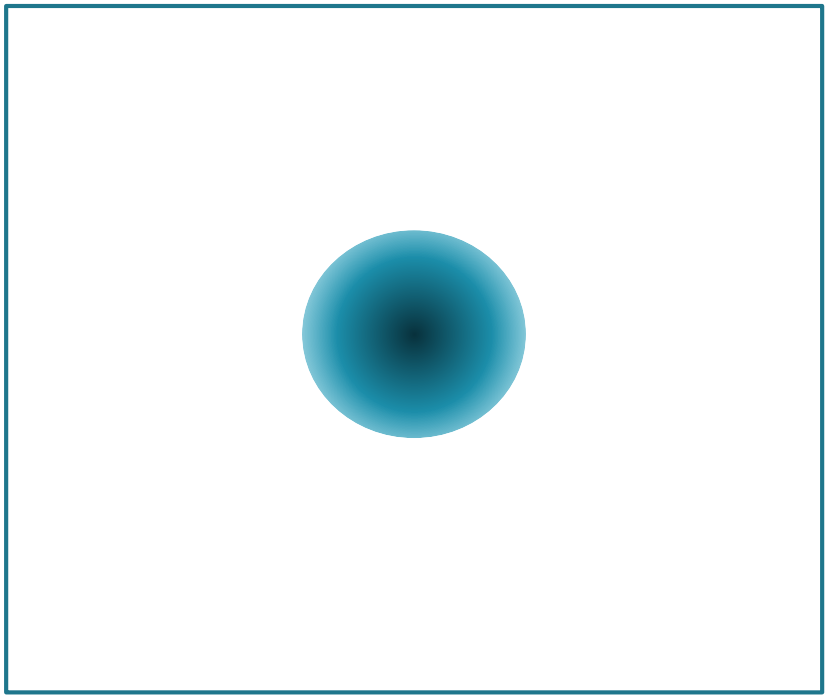


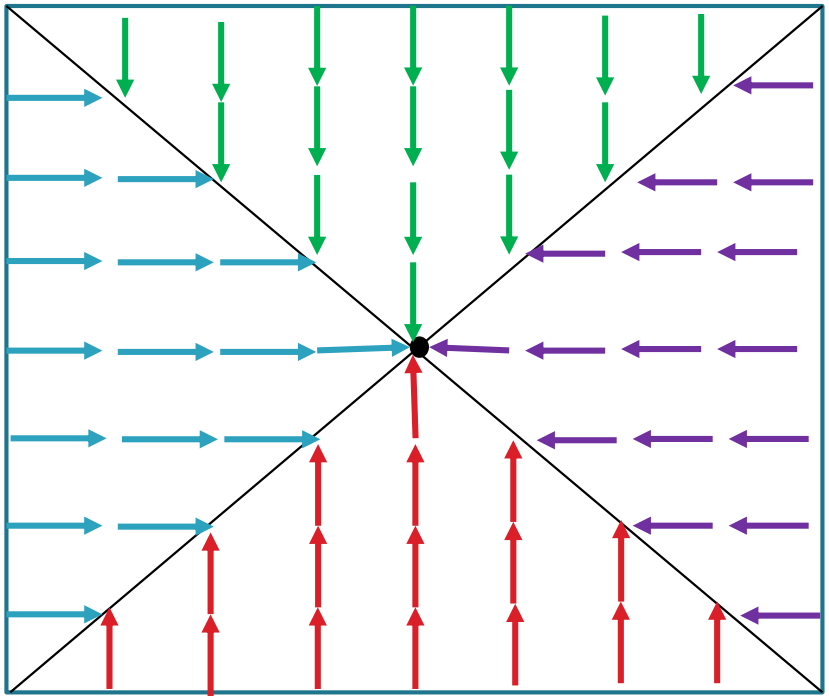
	LA	LAA	LIPV	LSPV	RIPV	RSPV	Total
2D only	0.926	0.443	0.553	0.483	0.549	0.242	0.872
3D only	0.907	0.363	0.546	0.418	0.603	0.403	0.853
2D + 3D	0.942	0.658	0.706	0.620	0.718	0.395	0.898

Liao et al. More knowledge is better: Cross-domain volume completion and 3D+2D segmentation for intracardiac echocardiography contouring, MICCAI 2018.

Novel representation for landmark

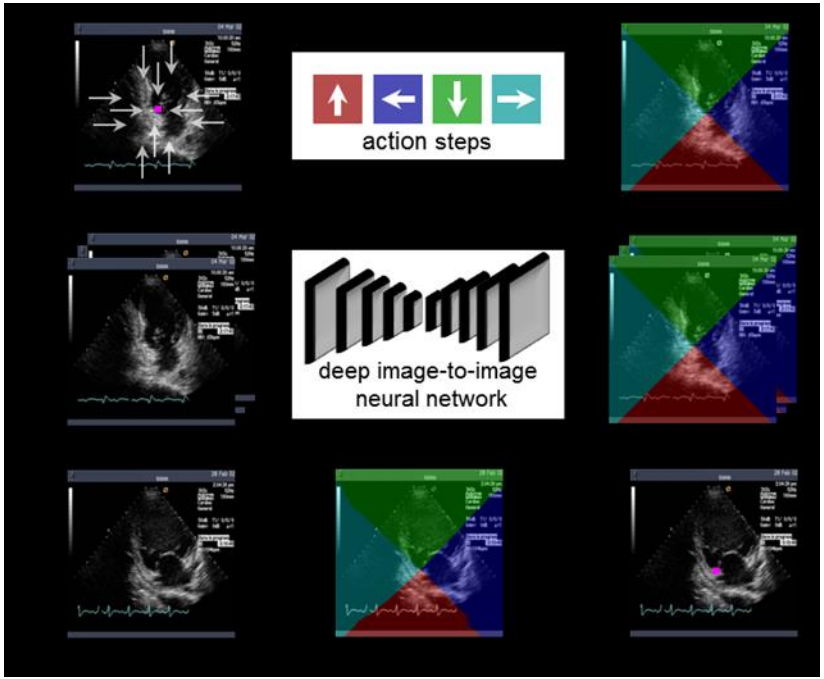
spatially local vs distributed





Xu et al., Supervised Action Classifier: Approaching Landmark Detection as Image Partitioning, MICCAI 2017.

Landmark detection using deep image-to-image network + supervised action map [MICCAI'2017]

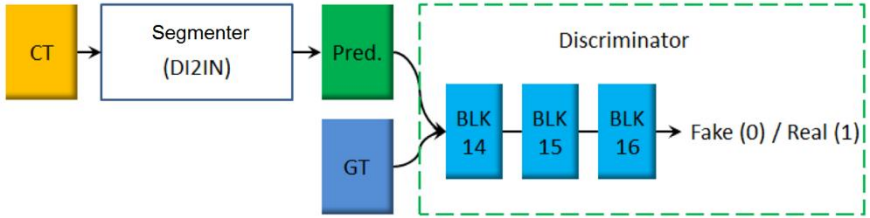


		PBT		DRL		I2I		SAC	
		lmk1	lmk2	lmk1	lmk2	lmk1	lmk2	lmk1	lmk2
CA	mean	10.45	13.85	7.69	10.02	6.73	9.02	**6.31**	**8.01**
	50%	5.74	8.11	5.43	7.63	5.00	6.40	**4.35**	**5.88**
	80%	11.11	16.18	9.33	13.73	8.54	11.40	**7.54**	**10.83**
OB	mean	59.23	130.66	29.99	32.45	30.07	21.97	**14.94**	**16.76**
	50%	35.31	139.49	11.69	13.17	5.39	6.08	**4.85**	**5.91**
	80%	109.84	193.64	43.98	45.76	13.34	15.54	**11.76**	**13.67**

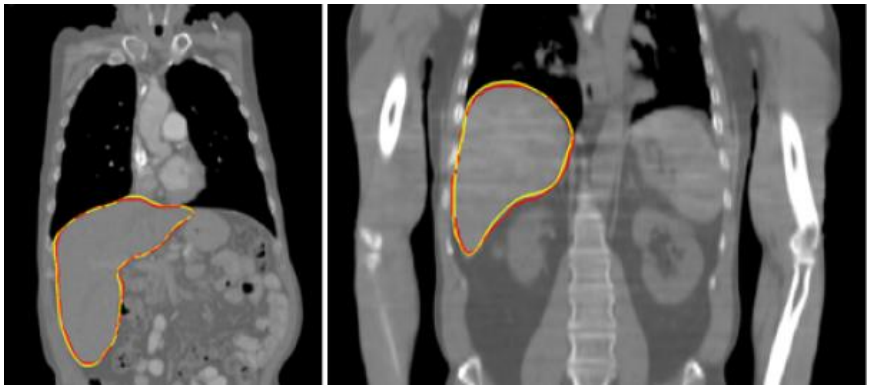
Xu et al., Supervised Action Classifier: Approaching Landmark Detection as Image Partitioning, MICCAI 2017.

Organ contouring with adversarial shape prior [MICCAI'2017]

- Using image2image network and adversarial shape prior
- Liver segmentation: 34% error reduction when using 1000 CT data sets

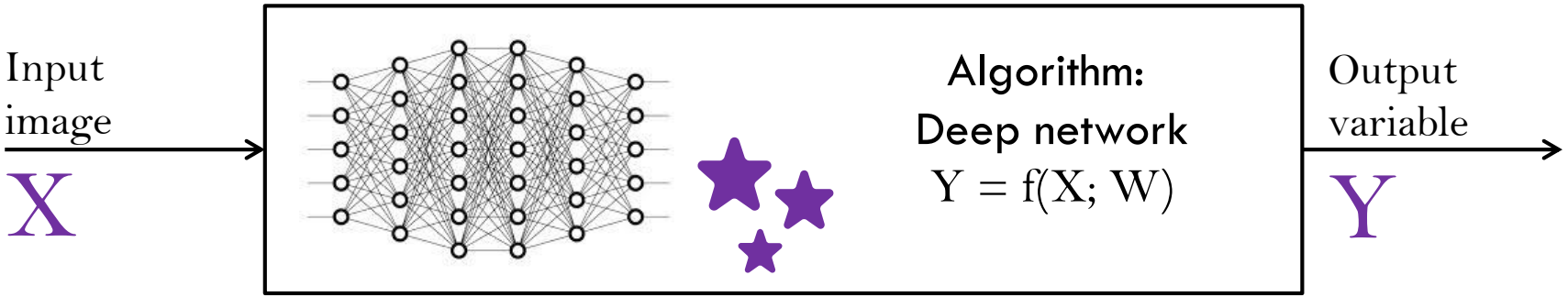


Method	ASD (mm)			
	Mean	Std	Max	Median
Ling *et al.* (400) [5]	2.95	5.07	37.45	2.01
DI2IN (400)	2.38	1.31	10.35	2.0
DI2IN-AN (400)	2.09	0.94	7.94	1.88
DI2IN (1000)	2.15	0.81	6.51	1.95
DI2IN-AN (1000)	**1.95**	0.75	**6.48**	**1.81**



Yang et al., Automatic Liver Segmentation Using an Adversarial Image-to-Image Network, MICCAI 2017

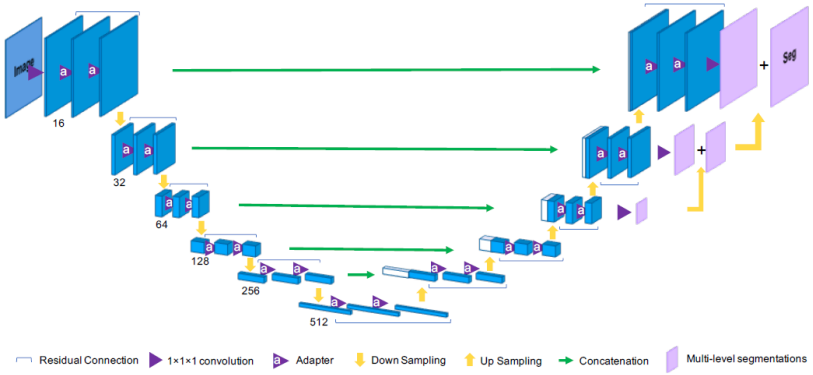
Knowledge in algorithm



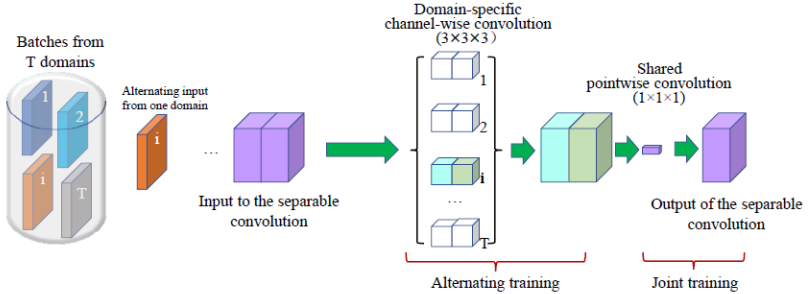
Knowledge in algorithm

- Network design
- Leveraging the imaging physics, geometry

U²-Net: universal u-net for multi-domain tasks



U²-Net

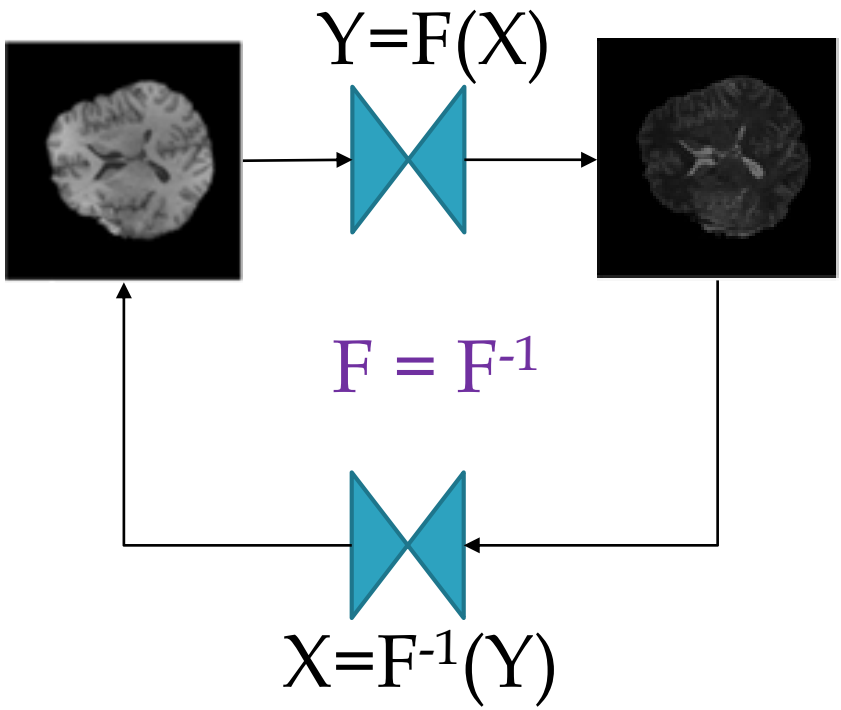


Adapter

- **One network with N adaptations** v.s. N independent networks
- Similar organ segmentation performance on 6 tasks but with 1% parameters
- Able to adapt to a new domain

* Huang et al., 3D U2-Net: A 3D Universal U-Net for Multi-Domain Medical Image Segmentation, MICCAI 2019. (patent pending)

Self-inverse network

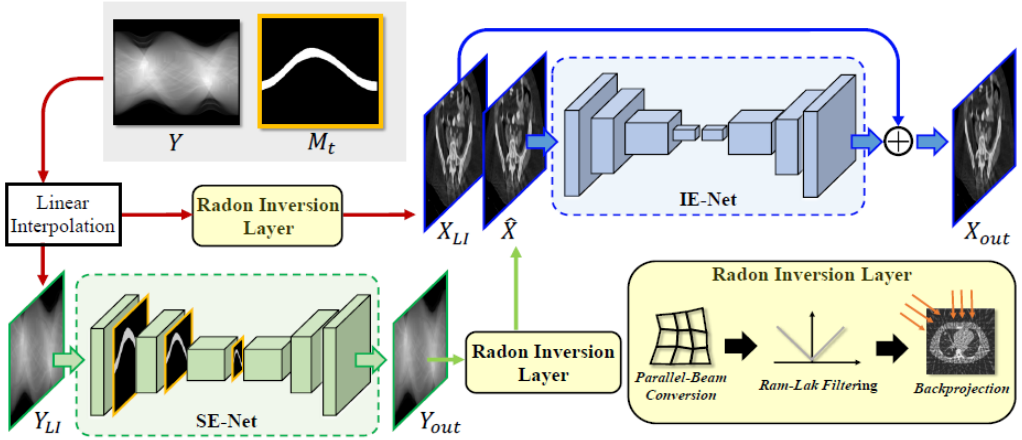


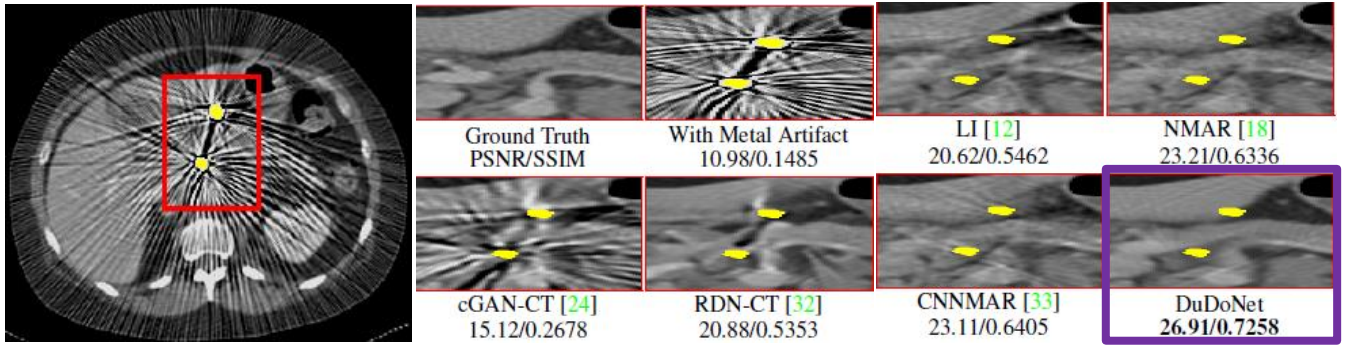
- Self-inverse
- Must be one2one

Direction	Method	(a)	L1	PSNR	SSIM
$T_1 \rightarrow T_2$	pix2pix		0.042	26.53	0.871
$T_1 \rightarrow T_2$	one2one		**0.039**	**29.23**	**0.875**
$T_2 \rightarrow T_1$	pix2pix		0.051	27.78	0.872
$T_2 \rightarrow T_1$	one2one		**0.048**	**30.99**	**0.876**

https://arxiv.org/abs/1909.04104
https://arxiv.org/abs/1909.04110

DuDoNet: Dual-domain network for CT metal artifact reduction





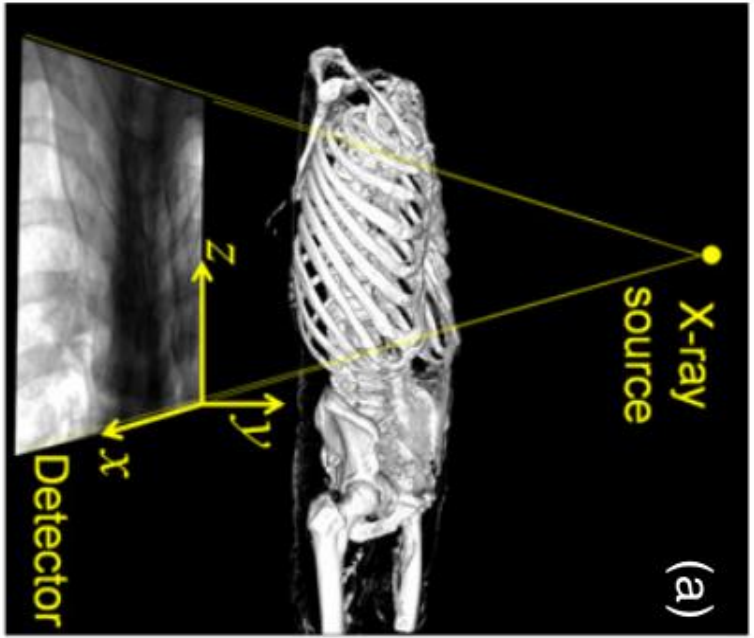
PSNR: 3dB better than state-of-the-art DL method.

Lin et al, DuDoNet: Dual Domain Network for CT Metal Artifact Reduction, CVPR2019. (patent pending)

Multiview 2d/3d rigid registration

Preoperative CT

Intraoperative X-Ray

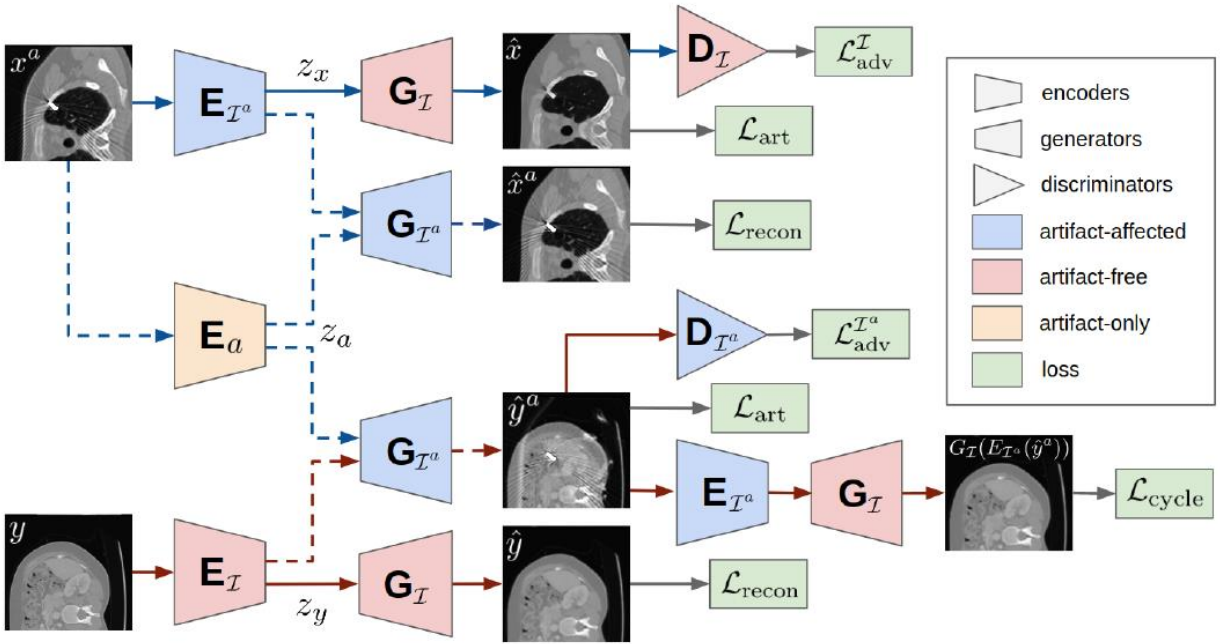


- POI tracking
- Multiview triangulation constraint

	mTRE (mm) 50th	mTRE (mm) 95th	GFR (>10mm)	Time (s)
Initial	20.4	29.7	92.9%	N/A
Opt.	0.62	57.8	40.0%	23.5s
DRL + opt.	1.06	24.6	15.6%	3.21s
Our + opt.	**0.55**	**5.67**	**2.7%**	**2.25s**

* Liao et al., Multiview 2D/3D Rigid Registration via a Point-Of-Interest Network for Tracking and Triangulation (POINT^2), CVPR2019. (patent pending)

Unsupervised artifact disentanglement network



Artifact Disentanglement Network

	PSNR(dB)	SSIM
ADN	**33.6**	**.924**
CycleGAN	30.8	.729
Deep Image Prior	26.4	.759
MUNIT	14.9	.750
DRIT	25.6	.797

Liao et al., Artifact Disentanglement Network for Unsupervised Metal Artifact Reduction, MICCAI 2019.

Why works?

思路	Idea	Examples
四两拨千金	Exploiting known information rather than brute force learning	ICE auto contouring, DuDoNet, disentanglement
升维思考	Making the pattern 'more' uniquely defined	more inputs/synthesized input
降维打击	Prior or regularization	multiview 2d/3d registration
梯度为王	Making problems more learnable	self-inverse learning, distributed landmark representation
量变产生质变	Allowing to see more examples	multitask learning, U^2Net

Acknowledgements

- Colleagues and students at MIRACLE (miracle.ict.ac.cn)

Acknowledgements

- Colleagues and students at MIRACLE
- Colleagues at Z²Sky
- Clinical collaborators at PUMC, JST, Fuwai, etc.
- Support from CAS, Alibaba, Tencent, etc.

智在天下

Z²Sky

Contact me if you are interested in …

zhoushaohua@ict.ac.cn

- Joining or visiting
- Collaborating with (clinical or R&D)
- Funding or investing in

奇迹

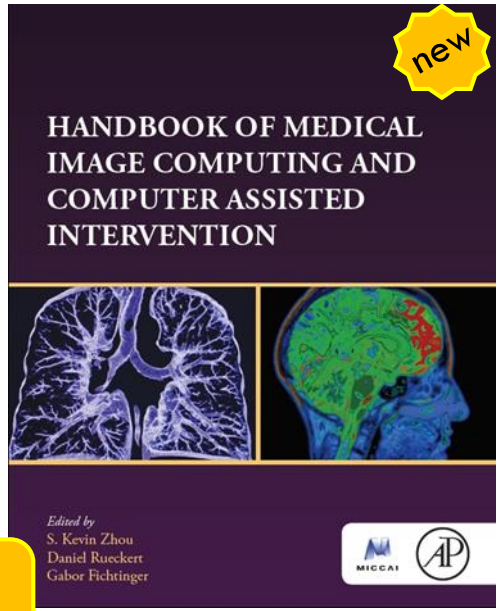
Medical Imaging, Robotics, Analytic Computing, Laboratory & Engineering

It is a MIRACLE !

智在天下

Z²Sky

Handbook of MICCAI



Pre-order
15% off

Editors: S. Kevin, Daniel Rueckert, Gabor Fichtinger

Hardcover ISBN: 9780128161760

Imprint: Academic Press

Published Date: 1st October 2019

Page Count: 1080

https://www.elsevier.com/books/handbook-of-medical-image-computing-and-computer-assisted-intervention/zhou/978-0-12-816176-0