西北工業大學
NORTHWESTERN POLYTECHNICAL UNIVERSITY

智能医学影像处理、分析与应用
（CCF-CV走进高校系列报告会 .第81期）

医学影像分析中的深度学习技术

夏勇 （教授）

空天地海一体化大数据应用技术国家工程实验室
西北工业大学.计算机学院

第七届医学图像计算青年研讨会（MICS 2020）

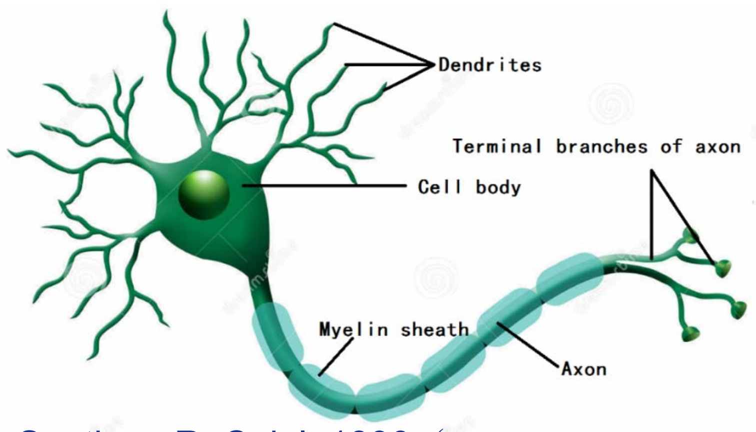
MICS在线学术讲座
微信公众号
QQ群：641894878

MICS
医学图像计算青年研讨会

Content

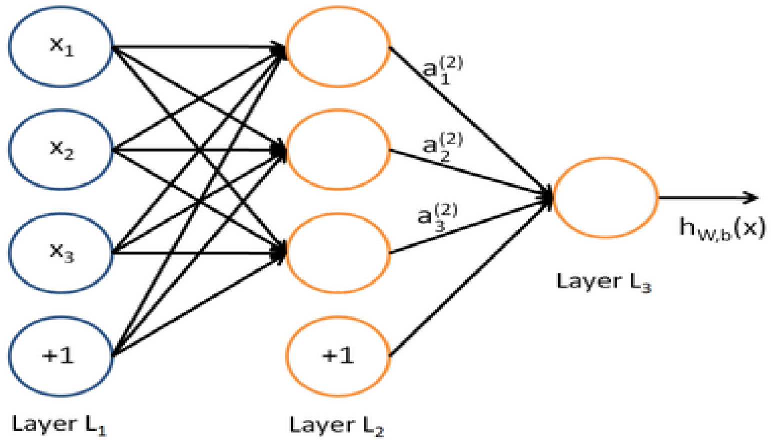
From Neural Network to Deep Learning

- Neuron Model



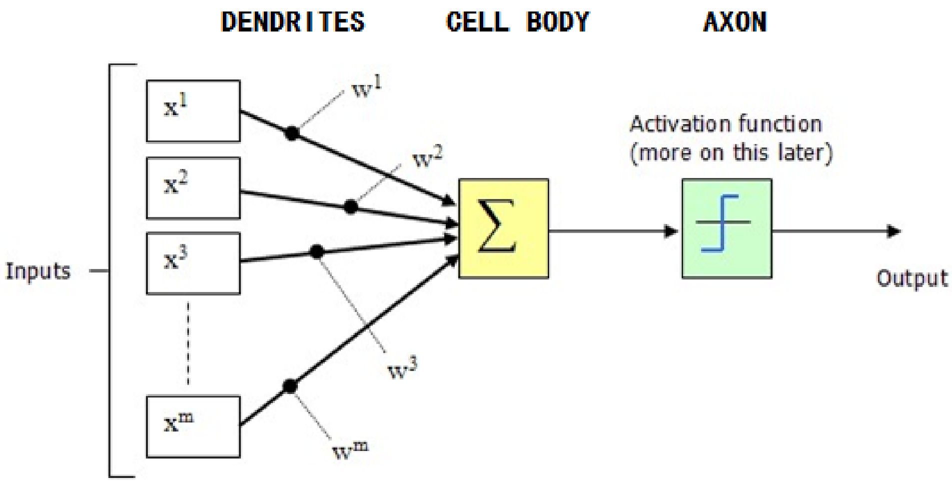
Santiago R. Cajal, 1890 's.

- Multi-Layer Perceptron (MLP)



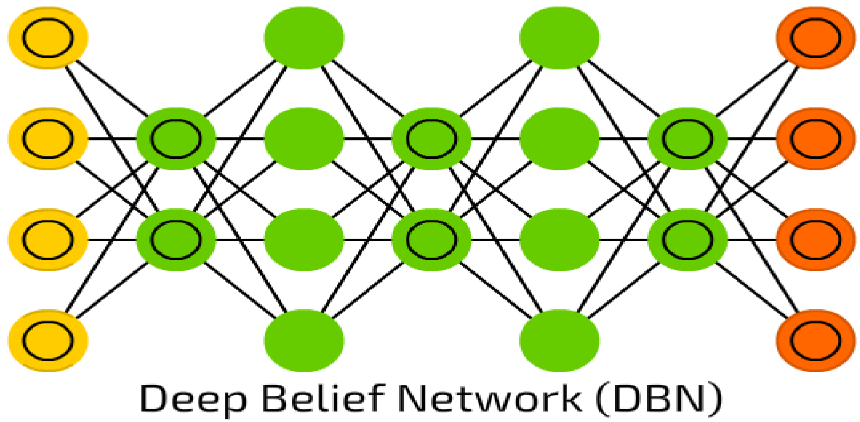
Rumelhart et al., Nature, 1986.

- Perceptron



Rosenblatt, Report 85-460-1, 1957.

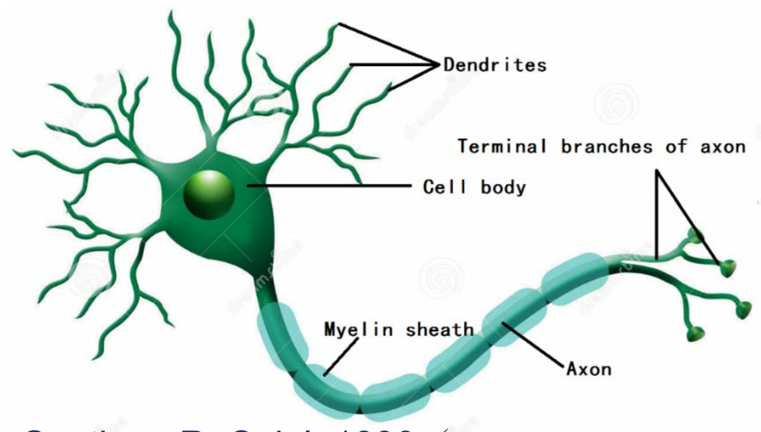
- Deep Learning



Hinton et al., Science, 2006.

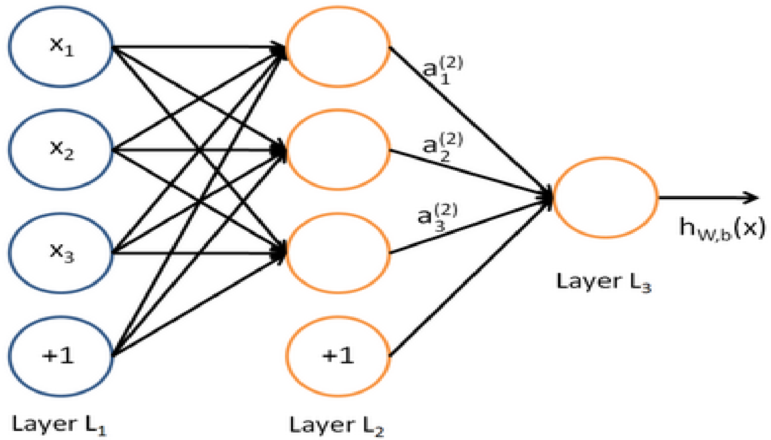
From Neural Network to Deep Learning

- Neuron Model



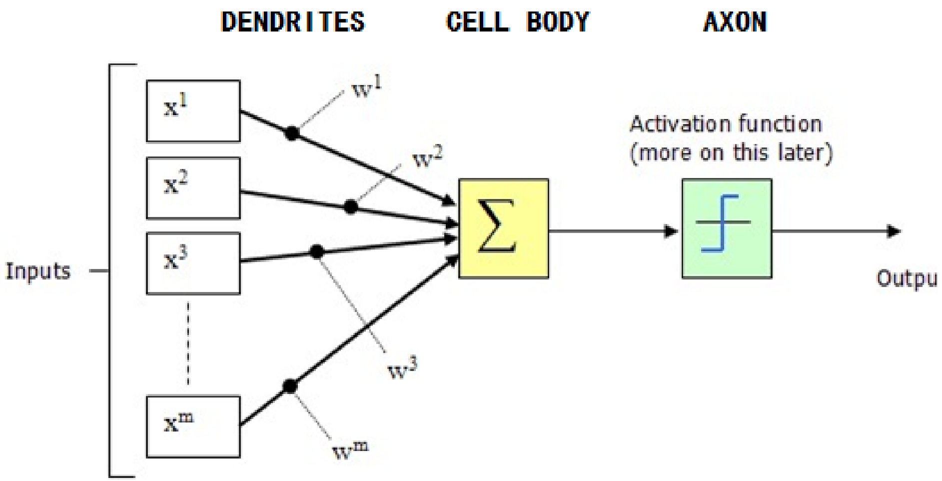
Santiago R. Cajal, 1890 ‘s.

- Multi-Layer Perceptron (MLP)



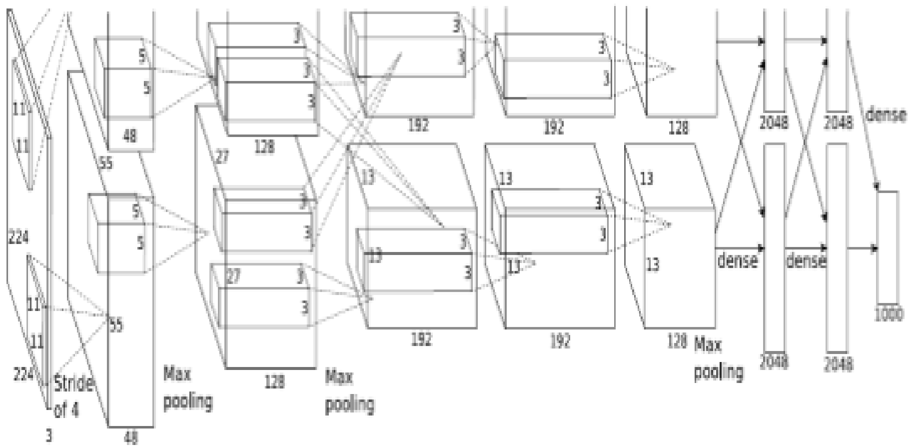
Rumelhart et al., Nature, 1986.

- Perceptron



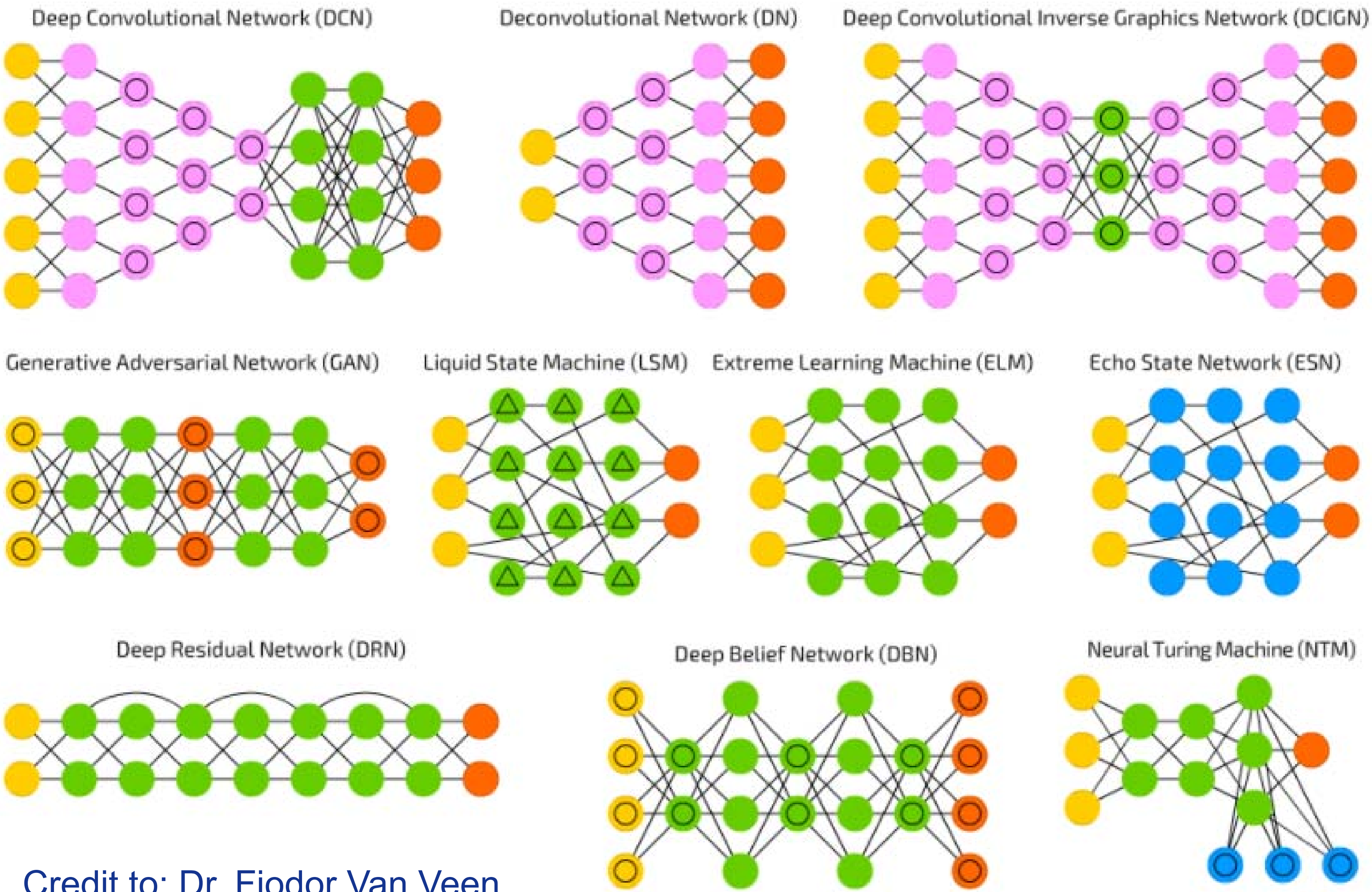
Rosenblatt, Report 85-460-1, 1957.

- Deep Learning



Krizhevsky et al., NIPS 2012.

Deep Learning Model Zoo



Credit to: Dr. Fjodor Van Veen

Deep Learning in Computer Vision

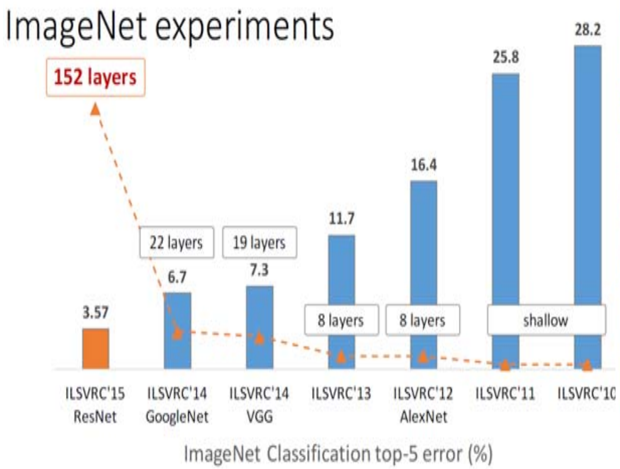


Image Classification

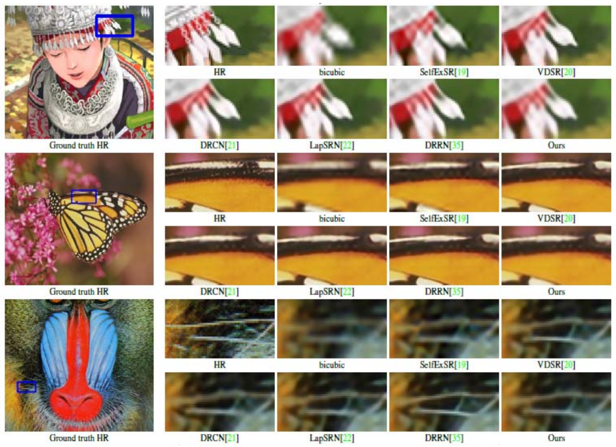
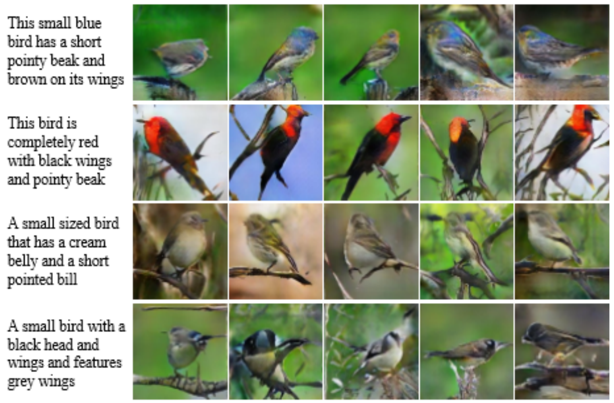


Image Superresolution

Image Style Transfer

Image-to-Image Generation



Text-to-Image Generation

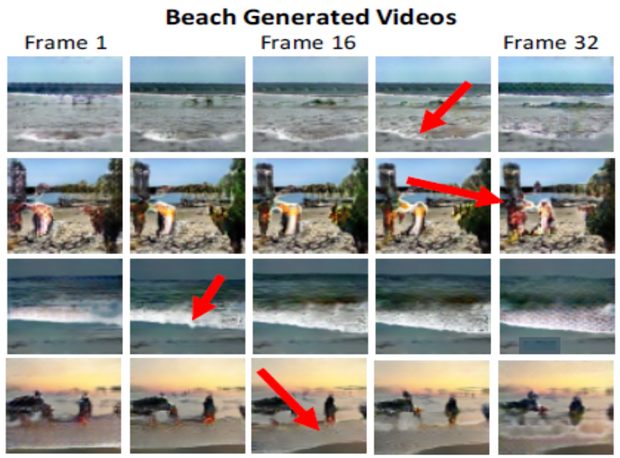
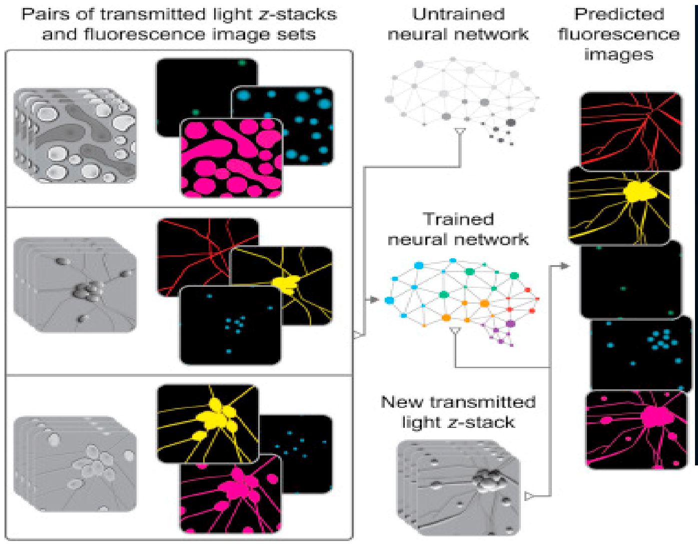
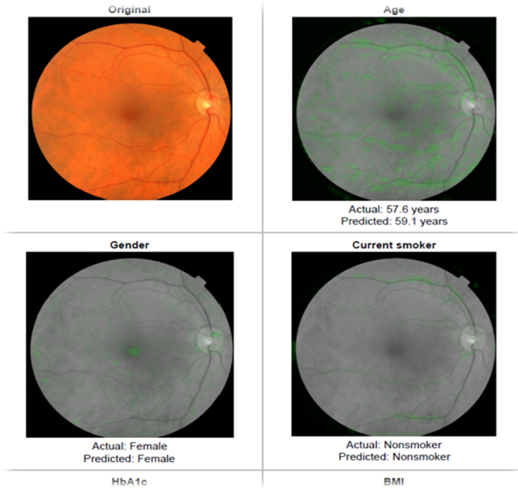


Image-to-Video Generation

Deep Learning in Medical Imaging Analysis

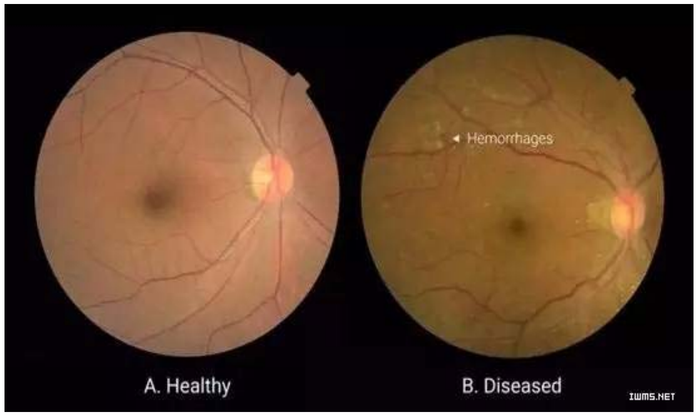


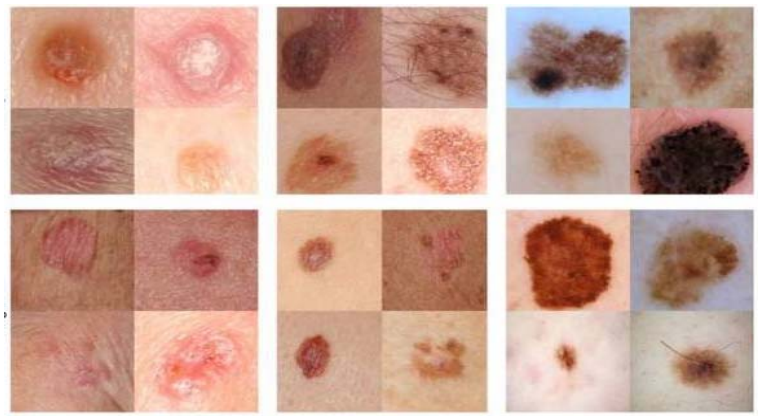


Google使用深度学习直接对细胞影像生成荧光标记 – 《Cell》

18年5月24日，Imagen公司用于检测成人手腕骨折的Osteo Detect软件获得FDA批准

谷歌通过视网膜图像预测心脑血管疾病风险，发表在《Nature》





Google利用深度学习开发的糖网诊断算法 –《JAMA》

18年4月11日,糖网诊断软件IDx-DR获得了FDA批准

斯坦福AI团队开发的皮肤癌辅助诊断方法 –《Nature》

Content

Lung Nodule Classification

- A “spot” on the lung, less than 3 cm in diameter and detected by anatomical imaging
  - Benign nodule: a benign lung tumour such as hamartomas
  - Malignant nodule: a primary cancerous lung tumor, lymphomas or a metastasis
- Benign-malignant lung nodule classification
  - Domain knowledge
  - Numerical representation
  - Incorporate the knowledge into deep learning

肺结节的良恶性鉴别

*轮廓——分叶？

*边缘——光滑锐利？毛刺？

*密度——均匀？空洞？钙化？脂肪？支气管充气征？空泡征？

*周围——血管集束征？胸膜凹陷？

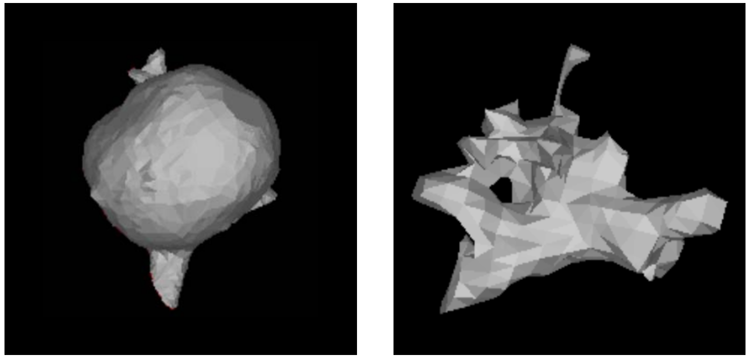
*强化程度——显著强化（≥25HU）？

*邻近——淋巴结增大？胸水？

Benign and Malignant Lung Nodules

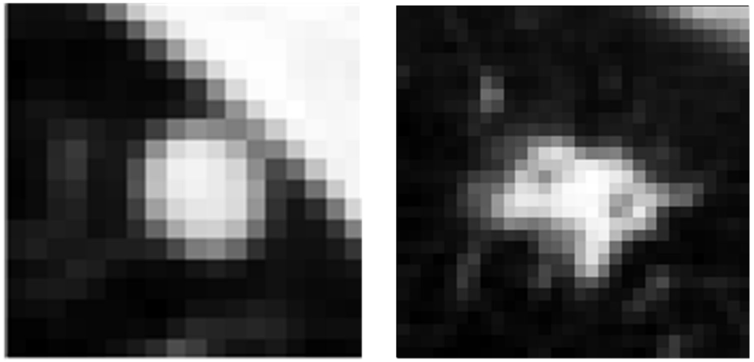
- Characteristics of malignant nodules
  - Heterogeneity in shape
  - Heterogeneity in voxel values (texture)

- Domain knowledge: There is a high correspondence between a nodule's malignancy and its heterogeneity



Benign Malignant

Heterogeneity in Shape (HS)



Benign Malignant

Heterogeneity in Voxel Values (HVV)

Fuse-TSD Algorithm

- GLCM-based Texture descriptor: Heterogeneity in voxel values,
- Fourier Shape descriptor: Heterogeneity in shape, and
- Deep model-learned features

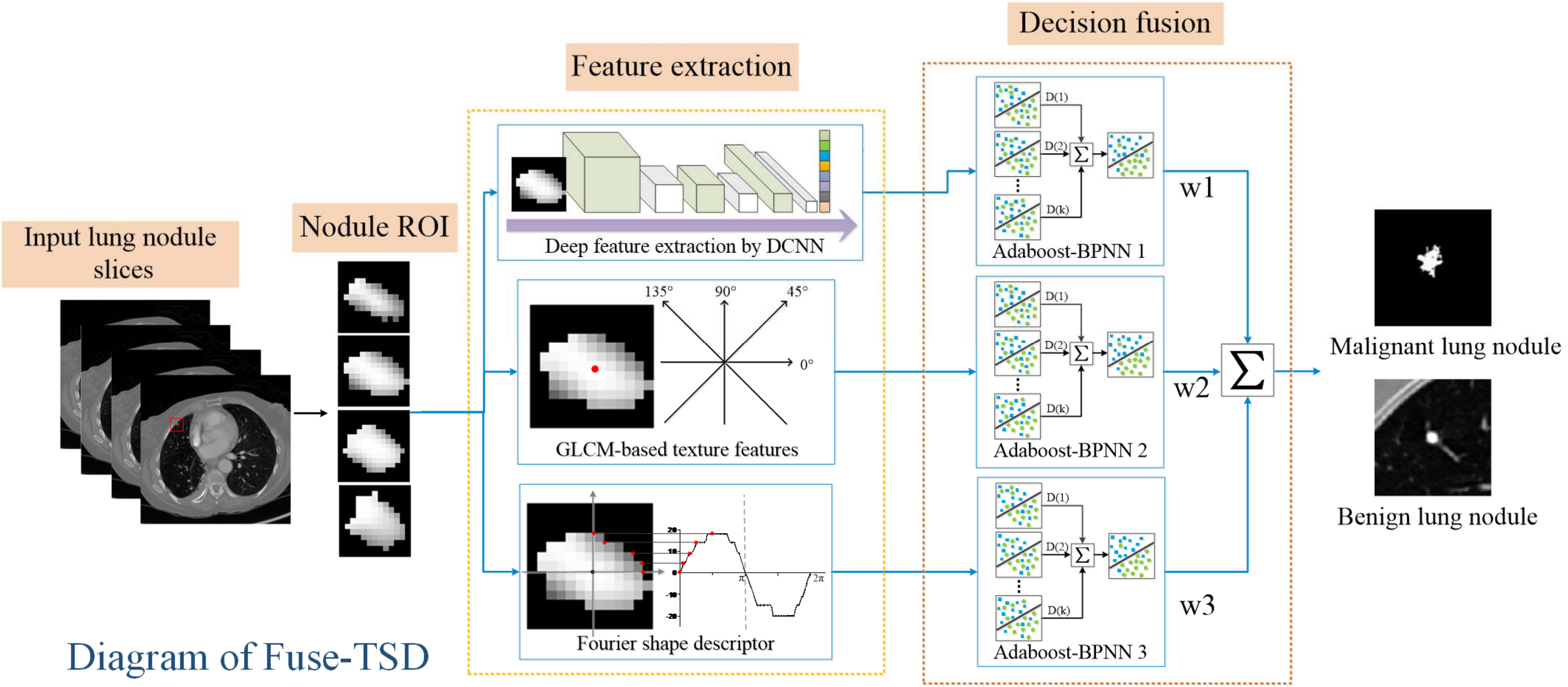


Diagram of Fuse-TSD

Yutong Xie, et al., *Information Fusion*, 2018.

Knowledge-Based Collaborative Deep Learning

- Using three ResNet-50 models to characterize a nodule's overall appearance (OA), heterogeneity of shape (HS) and heterogeneity of voxel values (HVV), respectively

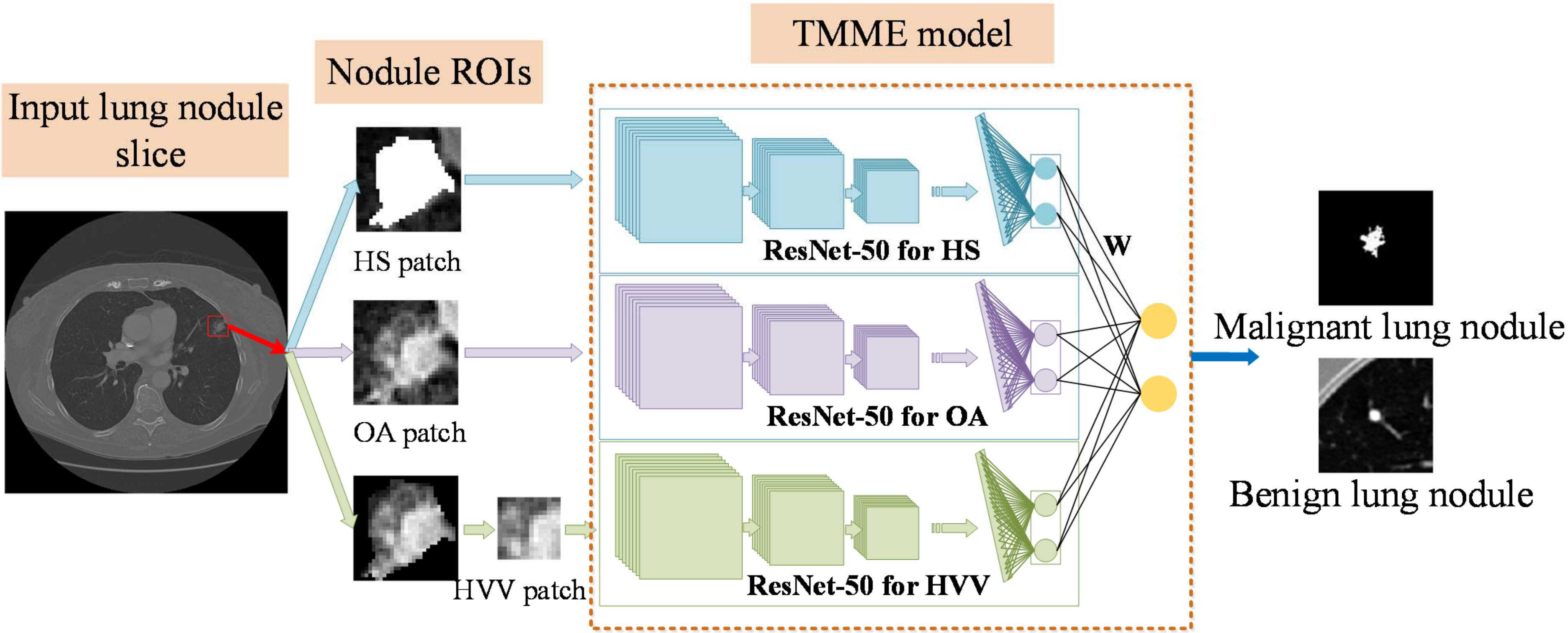


Diagram of the knowledge-based collaborative (KBC) deep learning algorithm for lung nodule classification

Yutong Xie, et al., *MICCAI 2017*; Yutong Xie, et al., *IEEE-TMI*, 2019.

Method

Semi-Supervised Adversarial Classification (SSAC) Model

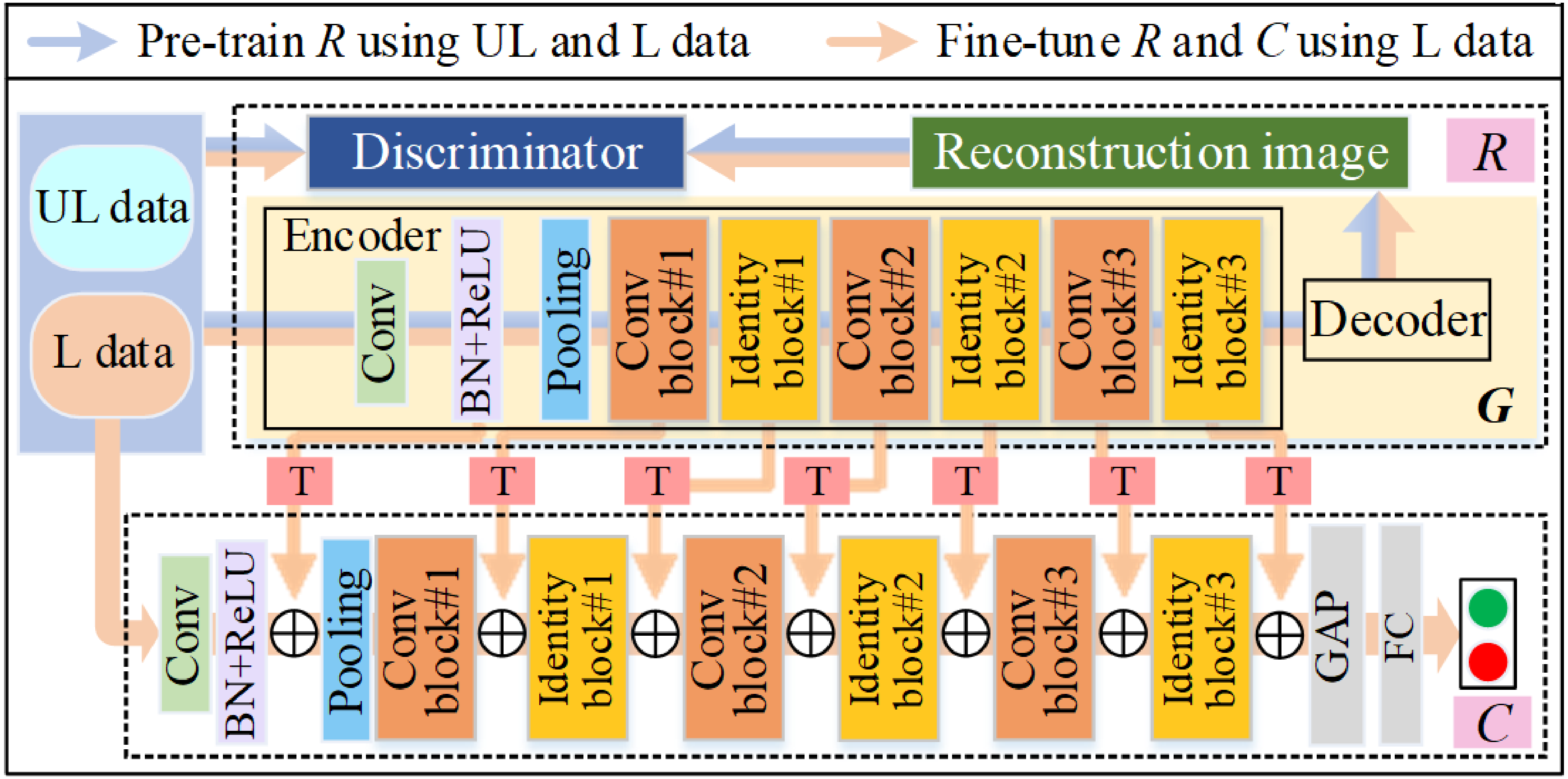


Diagram of SSAC Model

Yutong Xie, et al., *Medical Image Analysis*, 2019.

Method

Multi-View Knowledge-Based Collaborative SSAC Model

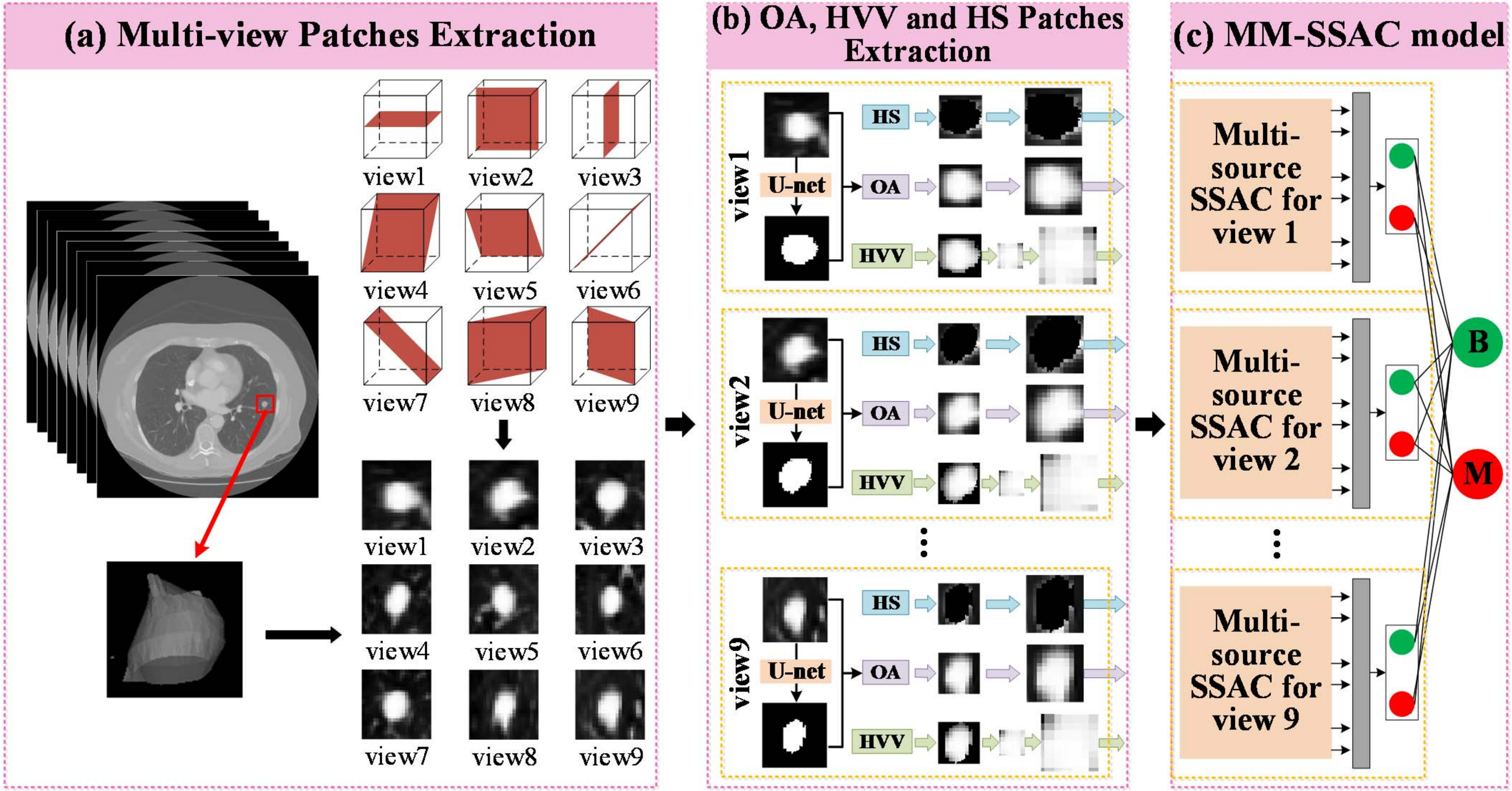


Diagram of proposed MK-SSAC model

Yutong Xie, et al., *Medical Image Analysis*, 2019.

Data and Results

- **LIDC-IDRC Dataset:** 1018 Chest CT studies

	Benign			Malignant	
Malignancy	1	2	3	4	5
Number	1301		637	644	

- **Tianchi Dataset:** 1839 unlabeled nodules

Performance of our MK-SSAC model and other models on the LIDC-IDRI dataset

Methods	Nodules	Results (%)			
	L/UL	Accuracy	Sensitivity	Specificity	AUC
MK-CatGAN	1945/1839	90.89±0.15	81.61±0.12	95.48±0.10	93.76±0.25
MK-AAE	1945/1839	91.13±0.15	82.92±0.20	95.19±0.12	94.00±0.21
MK-Ladder Network	1945/1839	92.11±0.27	83.07±0.19	96.59±0.10	95.36±0.20
3D GLCM+SVM, 2015	1945/0	85.38±0.10	70.20±0.15	92.80±0.20	88.19±0.16
MVF+SVM, 2016	1945/0	87.90±0.17	84.50±0.19	89.09±0.25	93.77±0.15
Fuse-TSD, 2018	1945/0	88.73±0.15	84.40±0.20	90.88±0.13	94.02±0.20
MV-KBC, 2018	1945/0	91.60±0.15	86.52±0.25	94.00±0.30	95.70±0.24
MK-SSAC	1945/1839	92.53±0.05	84.94±0.17	96.28±0.08	95.81±0.19

Yutong Xie, et al., *Medical Image Analysis*, 2019.

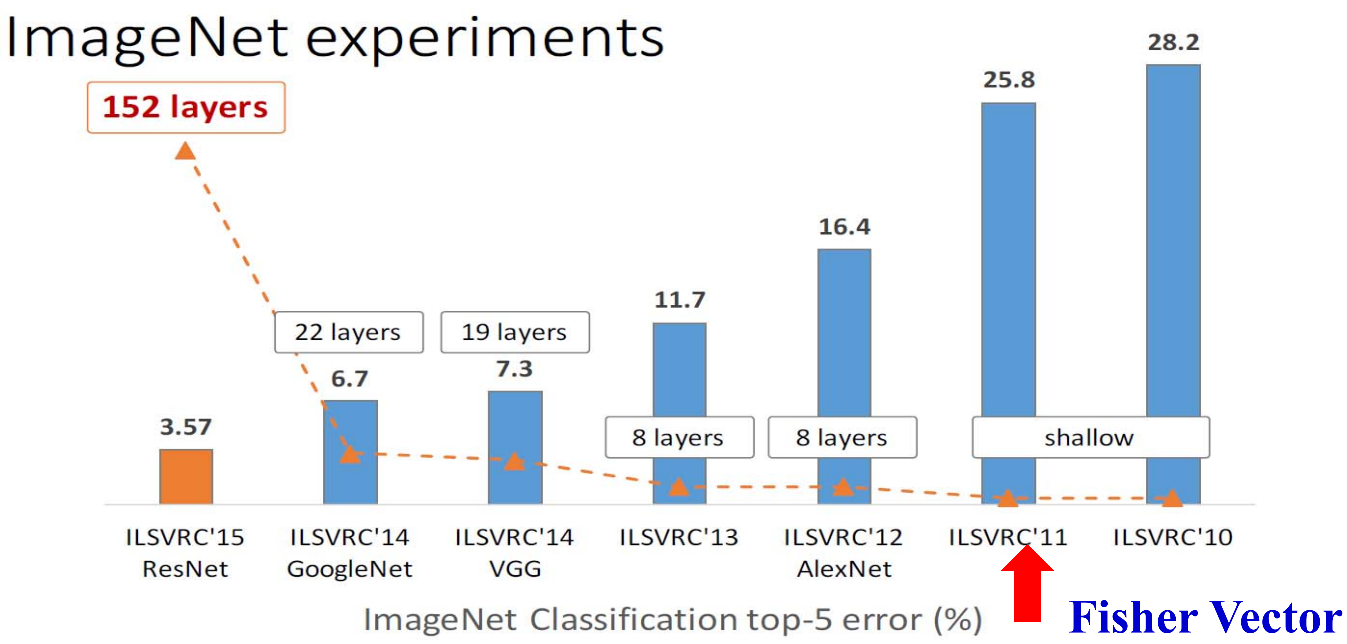
Content

CAN WE USE IMAGE-LEVEL LABELS TO FACILITATE IMAGE SEGMENTATION?

— Foreground Fisher Vector: A Traditional Approach to Image Classification

Before the Era of Deep Learning

- Fisher vector, deep convolutional neural networks (DCNNs), and the combination



K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition. In: CVPR 2016, pp. 770-778. IEEE Press, New York. (2016)

Theory of Fisher Vector

- The set of local descriptors of an image: $\boldsymbol{X} = \{x \in \mathcal{R}^D; x \sim p\}$
- **Fisher Score** is the gradient of the log-likelihood $P(\boldsymbol{X}|\lambda)$

$$G_\lambda^X = \nabla_\lambda P(\boldsymbol{X}|\lambda) = E_{x\sim p} \nabla_\lambda \log u_\lambda(x) = \nabla_\lambda \int_x p(x) \log u_\lambda(x)\, dx$$

  where u_λ is the probability density function of the generative process

- **Fisher Information Matrix**

$$F_\lambda = E_{x\sim u_\lambda}[\nabla_\lambda\, log\, u_\lambda(x) \nabla_\lambda\, log\, u_\lambda(x)^{\mathrm{T}}] = {L_\lambda}^{\mathrm{T}} L_\lambda$$

- **Fisher Kernel**

$$K_{FK}(X, Y) = {G_\lambda^X}^T F_\lambda^{-1} G_\lambda^Y = {\mathcal{G}_\lambda^X}^T \mathcal{G}_\lambda^Y$$

- The **Fisher Vector** of X is defined as $\mathcal{G}_\lambda^X = L_\lambda G_\lambda^X$.

Theory of Fisher Vector (cont.)

- Non-parametrically estimate p by $X = \{x_t \in R^D;\ t = 1, \ldots, T\}$,
- Parametrically estimate u_λ to be a GMM with components $\{\mathcal{N}(\mu_l, \sigma_l);\ l = 1, \ldots, L\}$ and $\lambda = \{\omega_l, \mu_l, \sigma_l;\ l = 1, \ldots, L\}$.

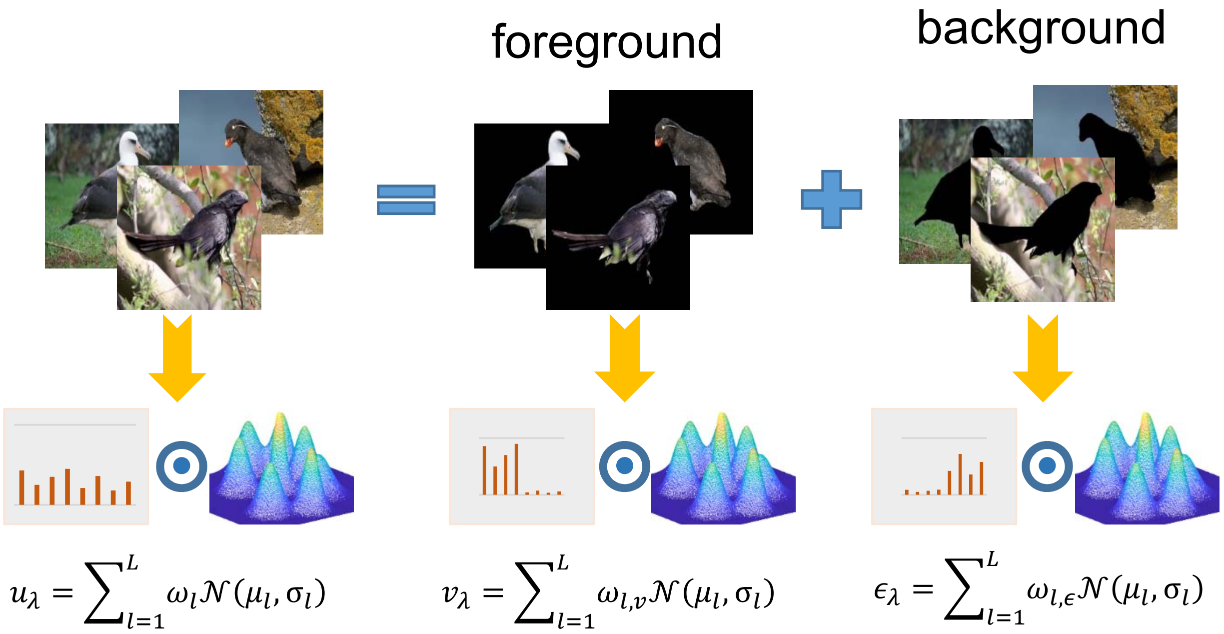
$$\begin{cases} \mathcal{G}_\lambda^X = \left(\mathcal{G}_{\mu,1}^{X\ \mathrm{T}}, \ldots, \mathcal{G}_{\mu,L}^{X\ \mathrm{T}}, \mathcal{G}_{\sigma,1}^{X\ \mathrm{T}}, \ldots, \mathcal{G}_{\sigma,L}^{X\ \mathrm{T}}\right)^{\mathrm{T}} \\ \mathcal{G}_{\mu,l}^X = \frac{1}{T\sqrt{\omega_l}} \sum_{t=1}^{T} \gamma_l(x_t) \left[\frac{x_t - \mu_l}{\sigma_l}\right] \\ \mathcal{G}_{\sigma,l}^X = \frac{1}{T\sqrt{2\omega_l}} \sum_{t=1}^{T} \gamma_l(x_t) \left[\frac{(x_t - \mu_l)^2}{\sigma_l{}^2} - 1\right], \end{cases}$$

$\gamma_l(x) = \frac{\omega_l u_l(x)}{\sum_{j=1}^{L} \omega_j u_j(x)}$ is the posterior probability of x assigned to the l-th Gaussian component u_l.

Foreground Fisher Vector

- Only foreground X_q is useful for classification

$$\boldsymbol{X} = X_q \cup X_r \leftrightarrow p(x) = wq(x) + (1-w)r(x)$$



- **Foreground FV**: $\check{\mathcal{G}}_\lambda^X = \check{L}_\lambda \check{G}_\lambda^X$, where $\check{G}_\lambda^X = E_{x\sim q} \nabla_\lambda \log v_\lambda(x)$.

Yongsheng Pan, et al., *IEEE-TIP*, 2019.

Foreground Fisher Vector

- $\mathcal{G}_\lambda^X \rightarrow \check{\mathcal{G}}_\lambda^X \leftrightarrow u_\lambda \rightarrow v_\lambda, p \rightarrow q$
- $\check{\gamma}_l(x) = \frac{1}{w}\gamma_l(x)$ if x is extracted from foreground, otherwise, $\check{\gamma}_l(x) = 0$.

$$\begin{cases} \check{\mathcal{G}}_\lambda^X = \left(\check{\mathcal{G}}_{\mu,1}^{X\ \mathrm{T}}, \ldots, \check{\mathcal{G}}_{\mu,L}^{X\ \mathrm{T}}, \check{\mathcal{G}}_{\sigma,1}^{X\ \mathrm{T}}, \ldots, \check{\mathcal{G}}_{\sigma,L}^{X\ \mathrm{T}}\right)^{\mathrm{T}} \\ \check{\mathcal{G}}_{\mu,l}^X = \frac{1}{T\sqrt{\omega_{l,v}}} \sum_{t=1}^{T} \check{\gamma}_l(x_t) \left[\frac{x_t - \mu_l}{\sigma_l}\right] = \frac{\sqrt{\omega_l}}{w\sqrt{\omega_{l,v}}} \mathcal{G}_{\mu,l}^X, \\ \check{\mathcal{G}}_{\sigma,l}^X = \frac{1}{T\sqrt{2\omega_{l,v}}} \sum_{t=1}^{T} \check{\gamma}_l(x_t) \left[\frac{(x_t - \mu_l)^2}{\sigma_l{}^2} - 1\right] = \frac{\sqrt{\omega_l}}{w\sqrt{\omega_{l,v}}} \mathcal{G}_{\sigma,l}^X, \end{cases} \xrightarrow{c_l = \frac{\sqrt{\omega_l}}{w\sqrt{\omega_{l,v}}}} \check{\mathcal{G}}_{*,l}^X = c_l \mathcal{G}_{*,l}^X$$

- Separating foreground and background equals solving $C = \{c_l, l = 1, \cdots, L\}$
- c_l indicates discriminative ability of $\left(\mathcal{G}_{\mu,l}^X, \mathcal{G}_{\sigma,l}^X\right)$ and the membership of u_l to foreground.
- $c_l = 0$ means $\mathcal{N}(\mu_l, \sigma_l)$ only models the background.

Yongsheng Pan, et al., *IEEE-TIP*, 2019.

Foreground Fisher Vector

- Calculate $\mathcal{G}_\lambda^X$ for each X in $\mathbb{X}_T$ and $\mathbb{X}_V$;
- Divide the training set into two subsets $\langle \mathbb{X}_T, Y_T \rangle$ and $\langle \mathbb{X}_V, Y_V \rangle$;
- Train linear classifier P by $\langle \mathbb{X}_T, Y_T \rangle$;
- Solve C by maximizing the prediction accuracy on $\langle \mathbb{X}_V, Y_V \rangle$.

$$C = \arg\max_{C} \sum_{X \in \mathbb{X}_V} \left| y_X - P\left(\frac{\mathcal{G}_\lambda^X \odot C}{\left\| \mathcal{G}_\lambda^X \odot C \right\|_2} \right) \right|.$$

- Then $\check{\mathcal{G}}_\lambda^X = \mathcal{G}_\lambda^X \odot C$.
- Membership grade of x to foreground

$$mg(x) = \sum_{l=1}^{L} c_l \gamma_l(x)$$

Yongsheng Pan, et al., *IEEE-TIP*, 2019.

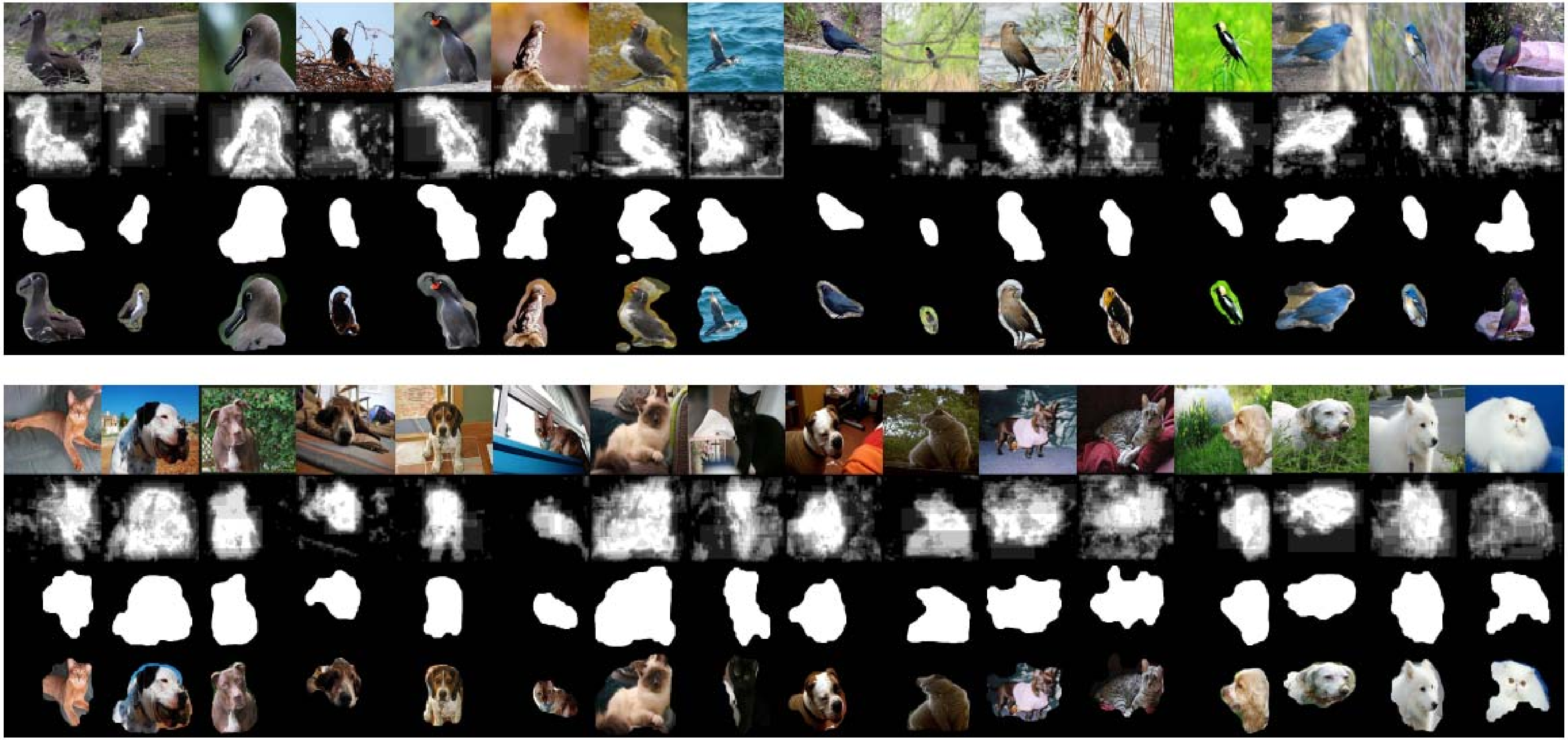
Data and Results

Results on Oxford Flower dataset (F102), Oxford Pet dataset (P37), Stanford Dog dataset (D120), Caltech-UCSD Bird dataset (B200), FGVC Aircraft dataset (A100), Stanford Car dataset (C196), Caltech-256 Object Category dataset (C256), PASCAL VOC 2007 dataset (V07), Fifteen Scenes dataset (S15), and MIT Indoor dataset (I67)

Method		F102	P37	D120	B200	C196	A100	C256	V07	I67	S15
AlexNet	FV	88.10	71.02	47.87	23.43	77.55	60.79	57.72	66.08	64.96	88.93
	IFV	92.72	83.61	60.89	47.28	84.29	74.88	68.62	76.58	72.72	91.57
	fFV	**95.34**	**89.66**	**70.01**	**73.31**	**87.10**	**81.26**	**74.66**	**78.99**	**74.12**	**91.80**
VGG-M	FV	89.87	71.99	50.20	22.11	80.88	61.09	61.26	70.16	69.04	89.88
	IFV	94.73	86.39	65.03	49.35	86.38	76.71	72.49	79.78	76.31	92.34
	fFV	**97.06**	**91.87**	**74.43**	**77.74**	**89.02**	**82.28**	**78.71**	**83.15**	**78.14**	**93.07**
VGG-VD16	FV	94.74	88.37	72.53	48.79	88.44	76.11	78.30	84.19	78.82	93.43
	IFV	97.31	92.89	79.03	70.42	90.78	83.34	83.95	89.23	82.57	95.10
	fFV	**98.47**	**94.85**	**83.63**	**84.95**	**92.27**	**86.63**	**87.46**	**91.37**	**84.15**	**95.15**
VGG-VD19	FV	94.79	89.76	69.77	49.81	88.87	77.37	79.39	85.42	79.95	93.51
	IFV	97.31	93.23	80.53	71.35	91.38	83.75	84.68	89.84	83.77	94.89
	fFV	**98.12**	**94.74**	**84.64**	**84.66**	**92.59**	**86.57**	**88.02**	**91.87**	**84.47**	**94.97**
ResNet50	FV	95.79	93.18	75.46	48.89	87.03	74.12	81.31	88.20	81.51	94.61
	IFV	97.33	94.81	83.95	70.26	89.34	81.11	86.50	91.28	84.79	95.66
	fFV	**98.53**	**95.44**	**87.28**	**86.51**	**90.96**	**85.29**	**89.56**	**93.00**	**85.35**	**95.70**
ResNet-101	FV	95.53	92.90	81.67	54.98	85.90	73.22	83.74	87.21	83.08	94.65
	IFV	97.44	95.33	87.53	74.59	88.93	80.42	88.72	92.29	86.20	95.93
	fFV	**98.78**	**96.06**	**89.27**	**86.46**	**90.92**	**85.34**	**90.89**	**93.87**	**86.26**	**96.11**
ResNet-152	FV	96.06	93.30	81.22	59.84	88.24	74.77	84.48	88.04	82.72	95.11
	IFV	97.68	95.49	87.60	75.65	90.09	80.49	89.00	92.39	85.61	95.78
	fFV	**98.65**	**96.12**	**89.82**	**87.90**	**91.39**	**84.98**	**91.16**	**94.22**	**86.65**	**96.05**

Yongsheng Pan, et al., *IEEE-TIP*, 2019.

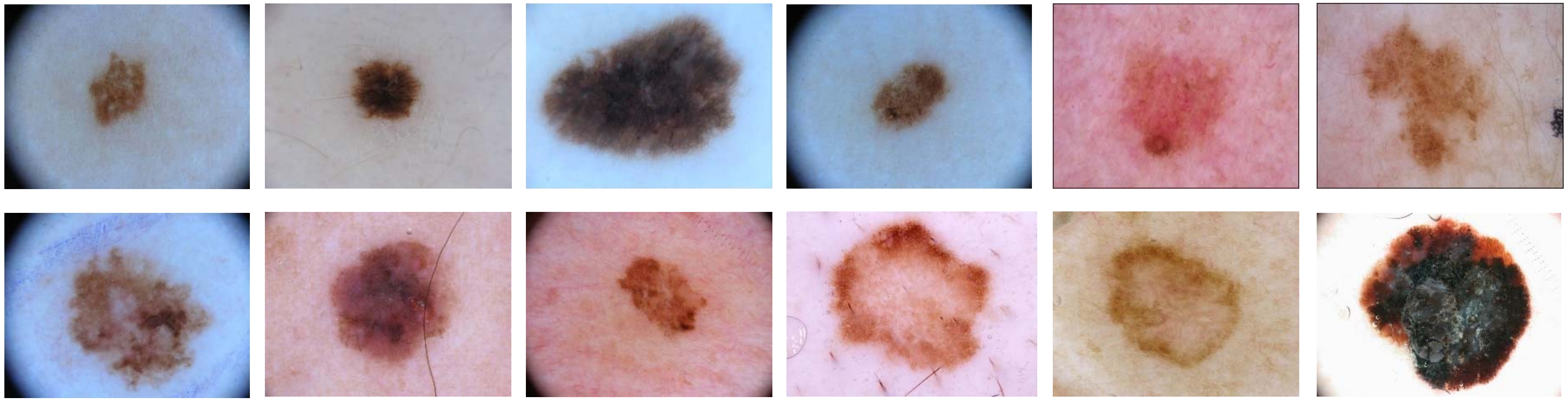
Data and Results



Ability of fgFV to separate the class-relevant foreground from class-irrelevant background. From top to bottom are the original images (1st and 5th rows, selected from B200 and P37), membership maps (2nd and 6th rows), binary membership maps (3rd and 7th rows), and the class-relevant content (4th and 8th rows).

Yongsheng Pan, et al., *IEEE-TIP*, 2019.

ISIC Skin Lesion Classification



- **ISIC Skin Lesion Classification Challenge 2016 Dataset**

  Training – 900 Test – 379

- **ISIC Skin Lesion Classification Challenge 2017 Dataset**

  Training – 2,000 Validation – 150 Test – 600

- **ISIC Skin Lesion Classification Challenge 2018 Dataset**

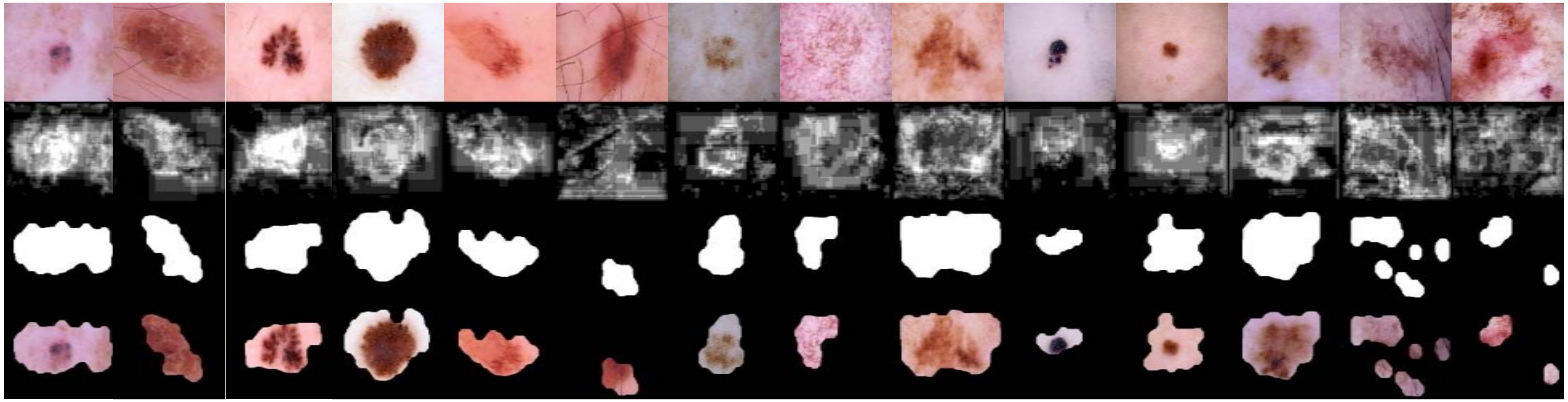
  Training – 10,015 Validation – 193 Test – 1,512

- **Task**: To classify the type of any skin lesion, benign or malignant?
- **Evaluation**: Average precision* (AveP), Area Under Curve (AUC), Sensitivity, Specificity

Data and Results

- ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection

Team	BMA	AUC	Accuracy	Sensitivity	Specificity	F1 score	PPV	NPV
DAISYLabs	0.854	0.982	0.970	0.836	0.982	0.843	0.860	0.971
test	0.835	0.964	0.964	0.732	0.987	0.794	0.883	0.959
ViDIR Group	0.821	0.971	0.949	0.820	0.968	0.767	0.734	0.957
Kevin	0.800	0.979	0.967	0.742	0.977	0.797	0.879	0.971
NWPU	0.792	0.952	0.928	0.702	0.980	0.680	0.697	0.924
yspan	0.785	0.949	0.947	0.746	0.975	0.733	0.726	0.948
Iyatomi Lab	0.780	0.871	0.943	0.780	0.963	0.738	0.711	0.953
Redha Ali	0.771	0.867	0.943	0.771	0.962	0.707	0.672	0.954
TUHH ISM	0.768	0.978	0.961	0.733	0.967	0.779	0.841	0.971
NWPU	0.754	0.949	0.947	0.732	0.973	0.725	0.728	0.948
Our Proposed FV-ftResNet	0.798	0.952	0.930	0.709	0.979	0.687	0.699	0.927

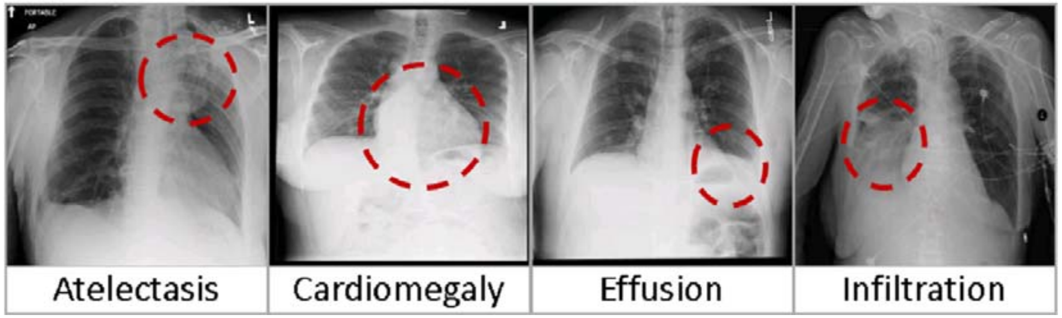


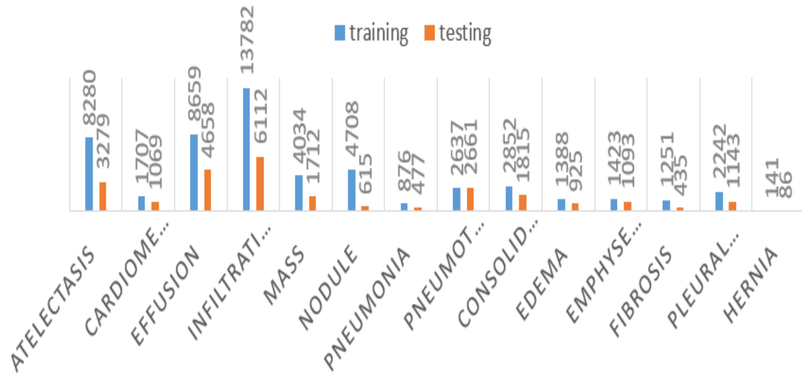
Class-relevant foreground regions detected during construction of fFV

CAN WE USE IMAGE-LEVEL LABELS TO FACILITATE IMAGE SEGMENTATION?

— Attention Residual Learning: A Deep Learning Approach to Image Classification

Deep Attention Learning





ChestX-ray14 Dataset: 14 Thoracic Diseases

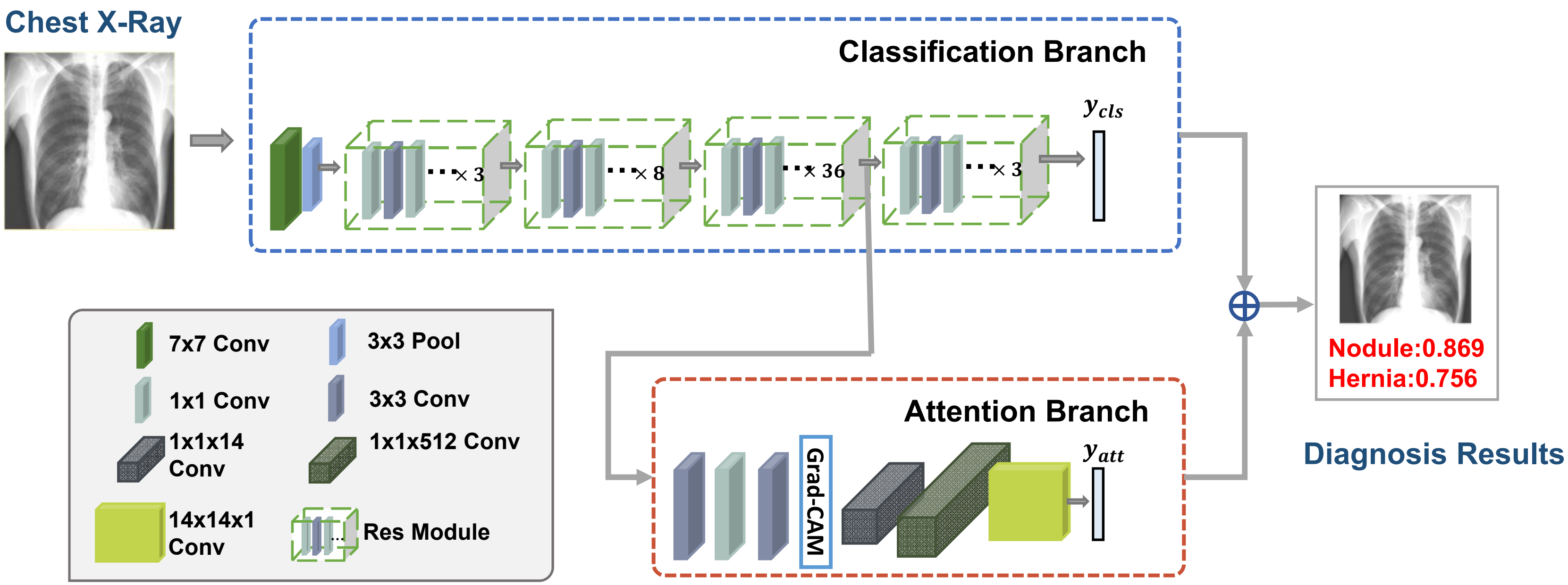
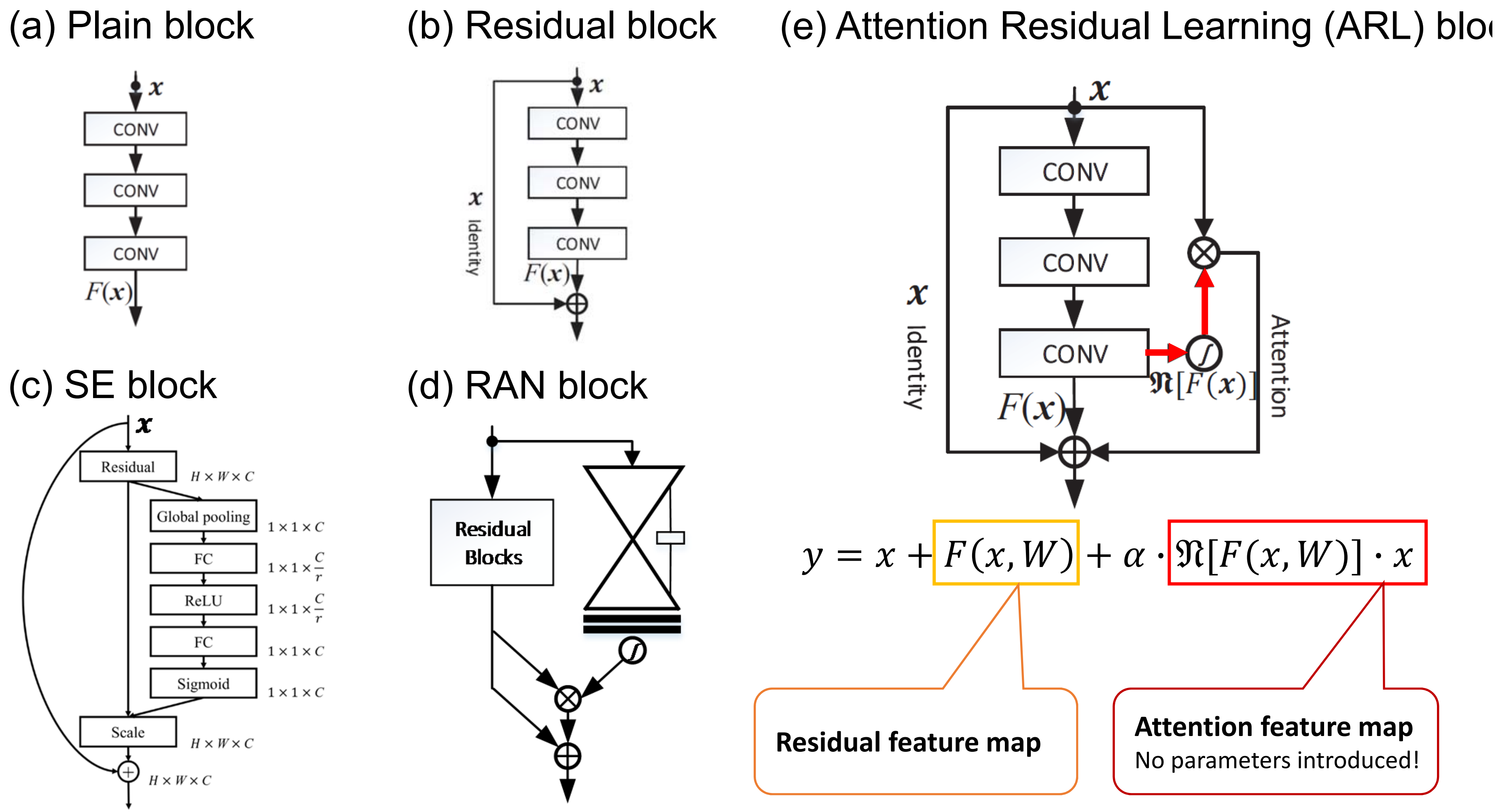


Diagram of the proposed ChestNet model

Hongyu Wang, et al., *IEEE-JBHI*, 2019.

Attention Residual Learning

- Different types of blocks used in deep learning models



Jinapeng Zhang, et al., *IEEE-TMI*, 2019.

ARL-CNN Model

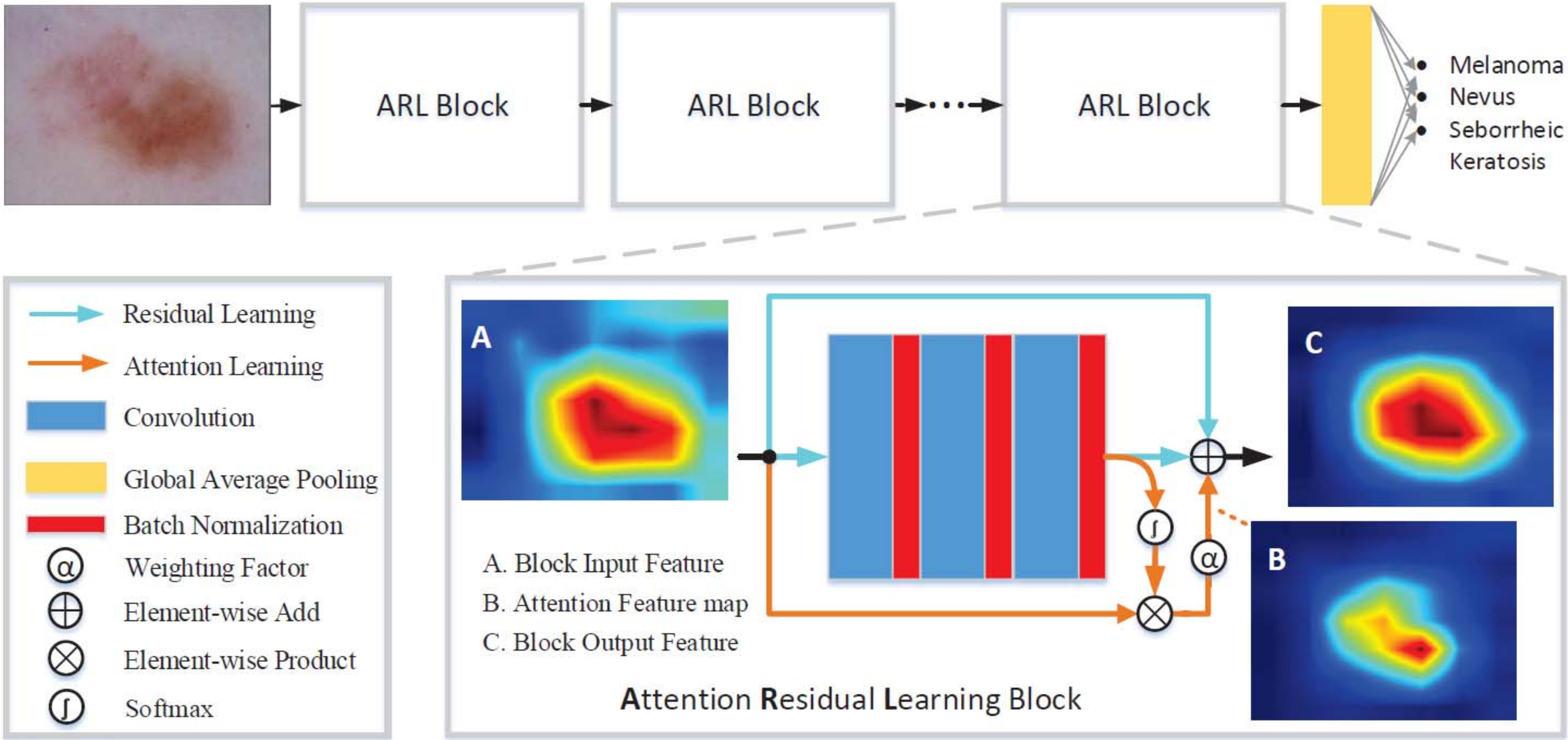


Diagram of the **A**ttention **R**esidual **L**earning based **C**onvolutional **N**eural **N**etwork (ARL-CNN)

Jinapeng Zhang, et al., *IEEE-TMI*, 2019.

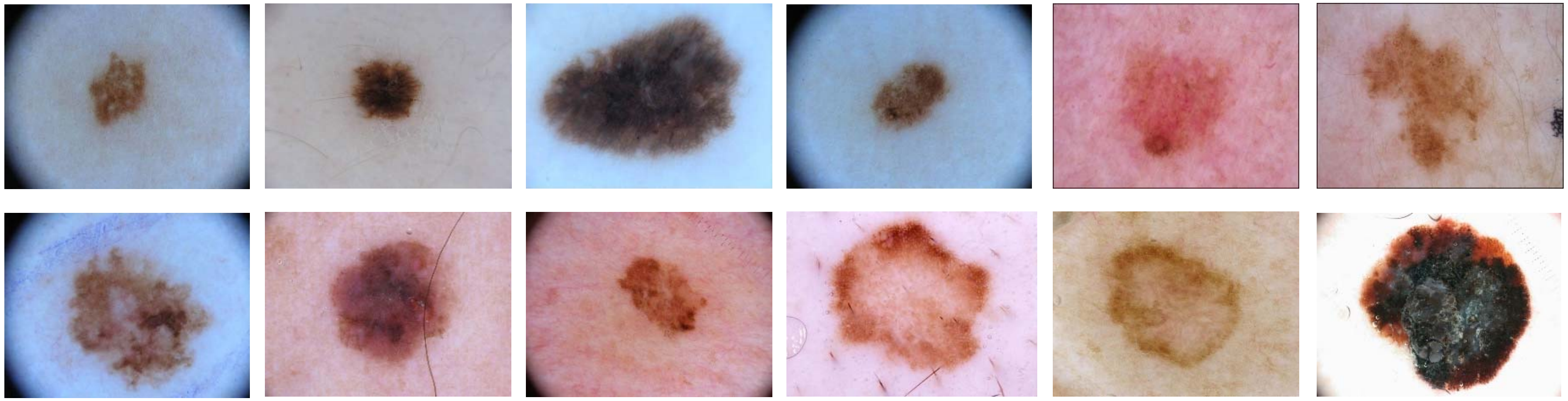
ARL-CNN Model

Architectures of ResNet14, ARL-CNN15, ResNet50 and ARL-CNN50

Layer name	Output size	ResNet14	ARL-CNN14	ResNet50	ARL-CNN50
Conv1	112x112x64	Conv, 7x7, stride 2	Conv, 7x7, stride 2	Conv, 7x7, stride 2	Conv, 7x7, stride 2
Conv2_x	56x56x256	3x3 max pool, stride 2 Residual block	3x3 max pool, stride 2 ARL	3x3 max pool,stride 2 [Residual block]x3	3x3 max pool,stride 2 [ARL]x3
Conv3_x	28x28x512	Residual block,stride 2	ARL,stride 2	Residual block, stride 2 [Residual block]x3	ARL, stride 2 [ARL]x3
Conv4_x	14x14x1024	Residual block,stride 2	ARL,stride 2	Residual block, stride 2 [Residual block]x5	ARL, stride 2 [ARL]x5
Conv5_x	7x7x2048	Residual block,stride 2	ARL,stride 2	Residual block, stride 2 [Residual block]x2	ARL, stride 2 [ARL]x2
GAP	1x1x2018	Global Average Pooling	Global Average Pooling	Global Average Pooling	Global Average Pooling
FC	2	Fully Connected	Fully Connected	Fully Connected	Fully Connected

Jinapeng Zhang, et al., *IEEE-TMI*, 2019.

ISIC Skin Lesion Classification



- **ISIC Skin Lesion Classification Challenge 2016 Dataset**

  Training – 900 Test – 379
- **ISIC Skin Lesion Classification Challenge 2017 Dataset**

  Training – 2,000 Validation – 150 Test – 600
- **ISIC Skin Lesion Classification Challenge 2018 Dataset**

  Training – 10,015 Validation – 193 Test – 1,512
- **Task**: To classify the type of any skin lesion, benign or malignant?
- **Evaluation**: Average precision* (AveP), Area Under Curve (AUC), Sensitivity, Specificity

Results

Comparison of the proposed ARL-CNN model with baseline models (ResNet-14 and ResNet-50), and state-of-the-art attention methods (Senet-14 and RAN-14) in melanoma classification and seborrheic keratosis classification in ISIC 2017 dataset.

Methods	Params ($\times 10^7$)	Melanoma Classification				Seborrheic Keratosis Classification			
		AUC	ACC	Sensitivity	Specificity	AUC	ACC	Sensitivity	Specificity
ResNet14 [23]	0.8	0.732	0.748	0.538	0.799	0.820	0.711	0.800	0.696
RAN14 [19]	1.2	0.767	0.762	**0.615**	0.797	0.852	0.758	**0.833**	0.745
SEnet14 [20]	1.1	0.758	0.757	0.598	0.795	0.847	0.727	0.811	0.712
ARL-CNN14	0.8	**0.777**	**0.778**	**0.615**	**0.818**	**0.875**	**0.763**	0.822	**0.753**
ResNet50*	2.3	0.857	0.838	0.632	0.888	0.948	0.842	0.867	0.837
ARL-CNN50*	2.3	**0.875**	**0.850**	**0.658**	**0.896**	**0.958**	**0.868**	**0.878**	**0.867**

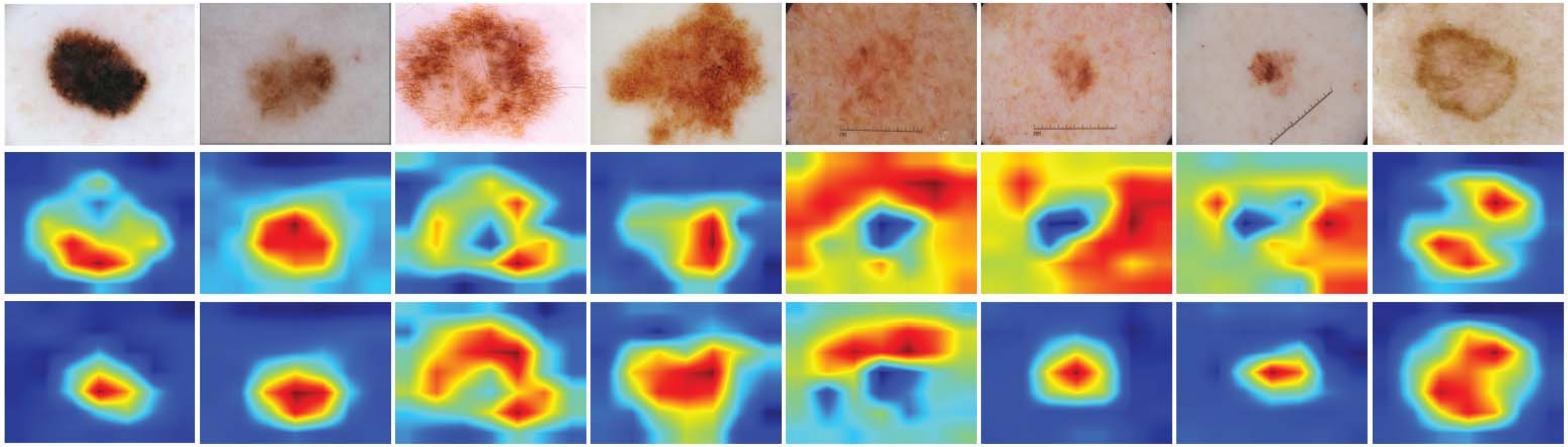
Comparison between the proposed ARL-CNN model and the top six ISIC 2017 challenge records. For each performance metric, **the highest and second highest** values were highlighted in "**red bold**" and "**black bold**", respectively.

Methods	Melanoma Classification				Seborrheic Keratosis Classification				Average AUC
	AUC	ACC	Sensitivity	Specificity	AUC	ACC	Sensitivity	Specificity	
Ours	**0.875**	0.850	**0.658**	0.896	**0.958**	0.868	**0.878**	0.867	**0.917**
#1 [31]	0.868	0.828	**0.735**	0.851	0.953	0.803	**0.978**	0.773	**0.911**
#2 [32]	0.856	0.823	0.103	**0.998**	**0.965**	0.875	0.178	**0.998**	0.910
#3 [33]	**0.874**	**0.872**	0.547	0.950	0.943	0.895	0.356	0.990	0.908
#4 [34]	0.870	**0.858**	0.427	0.963	0.921	**0.918**	0.589	0.976	0.896
#5 [35]	0.830	0.830	0.436	0.925	0.942	**0.917**	0.700	**0.995**	0.886
#6 [36]	0.836	0.845	0.350	**0.965**	0.935	0.913	0.556	0.976	0.886

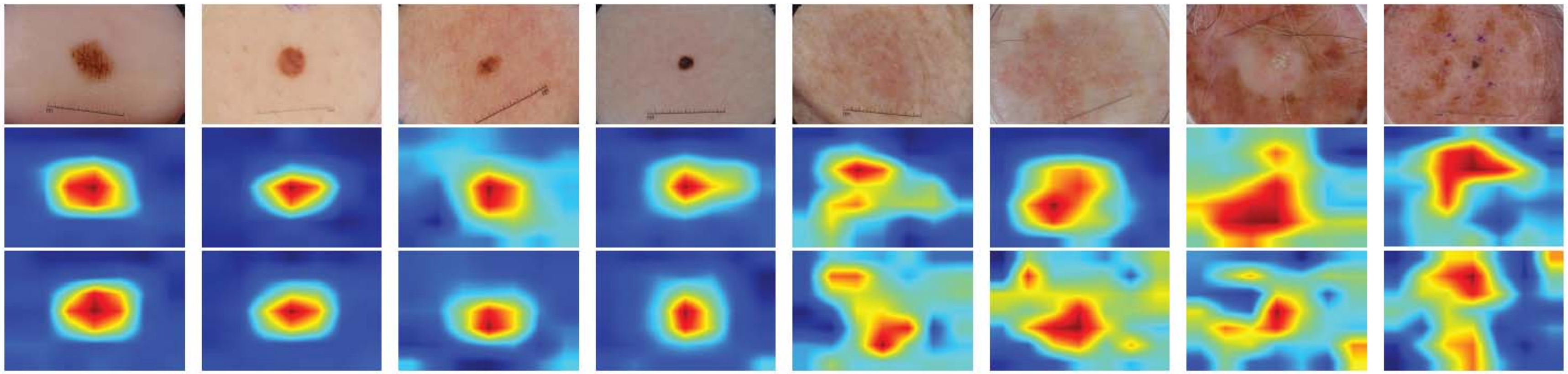
Jinapeng Zhang, et al., *IEEE-TMI*, 2019.

Results

Visualization of class activation mapping (CAM). Top row: dermoscopy images; Middle row: CAMs from ResNet50; Bottom row: CAMs from ARL-CNN50.



Visualization of attention regions learned in different tasks. Top row: dermoscopy images; Middle row: CAMs learned for melanoma classification; Bottom row: CAMs learned for seborrheic keratosis classification.



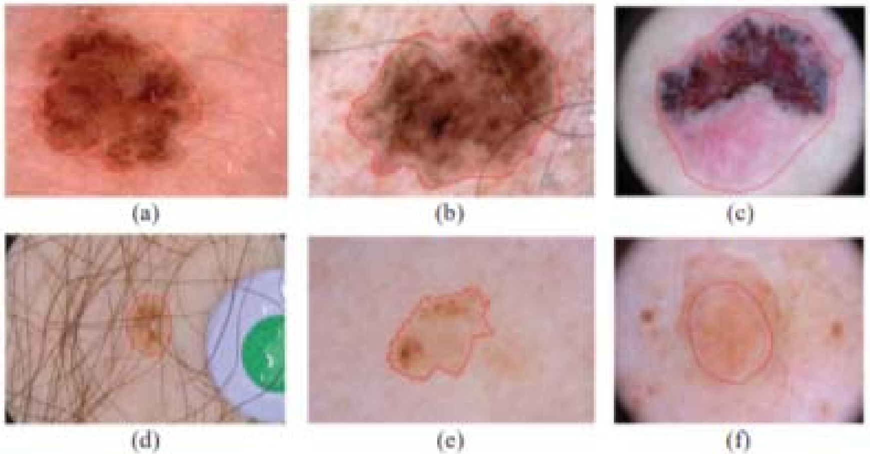
Jinapeng Zhang, et al., *IEEE-TMI*, 2019.

TOWARDS A UNIFIED IMAGE CLASSIFICATION-SEGMENTATION FRAMEWORK

— semi- and weakly-supervised Image segmentation

Medical Image Segmentation

- Segmentation of Skin Lesion



- Challenge of medical image segmentation

  → limited data with pixel-wise dense annotations
- Semi-supervised

  → labeled data + unlabeled data
- Weakly supervised：

  → weakly labeled data (e.g. bounding boxes / points / **image-level labels**)

Challenges on skin lesion segmentation. From these figures, we can see that skin lesion has variety of appearances, fuzzy object borders and some inevitable distractions. Red dash line represents the skin lesions contoured by dermatologists.

MB-DCNN Model: Skin Lesion Segmentation and Classification

➢ Leverage the **intrinsic correlation** existed in segmentation and classification tasks

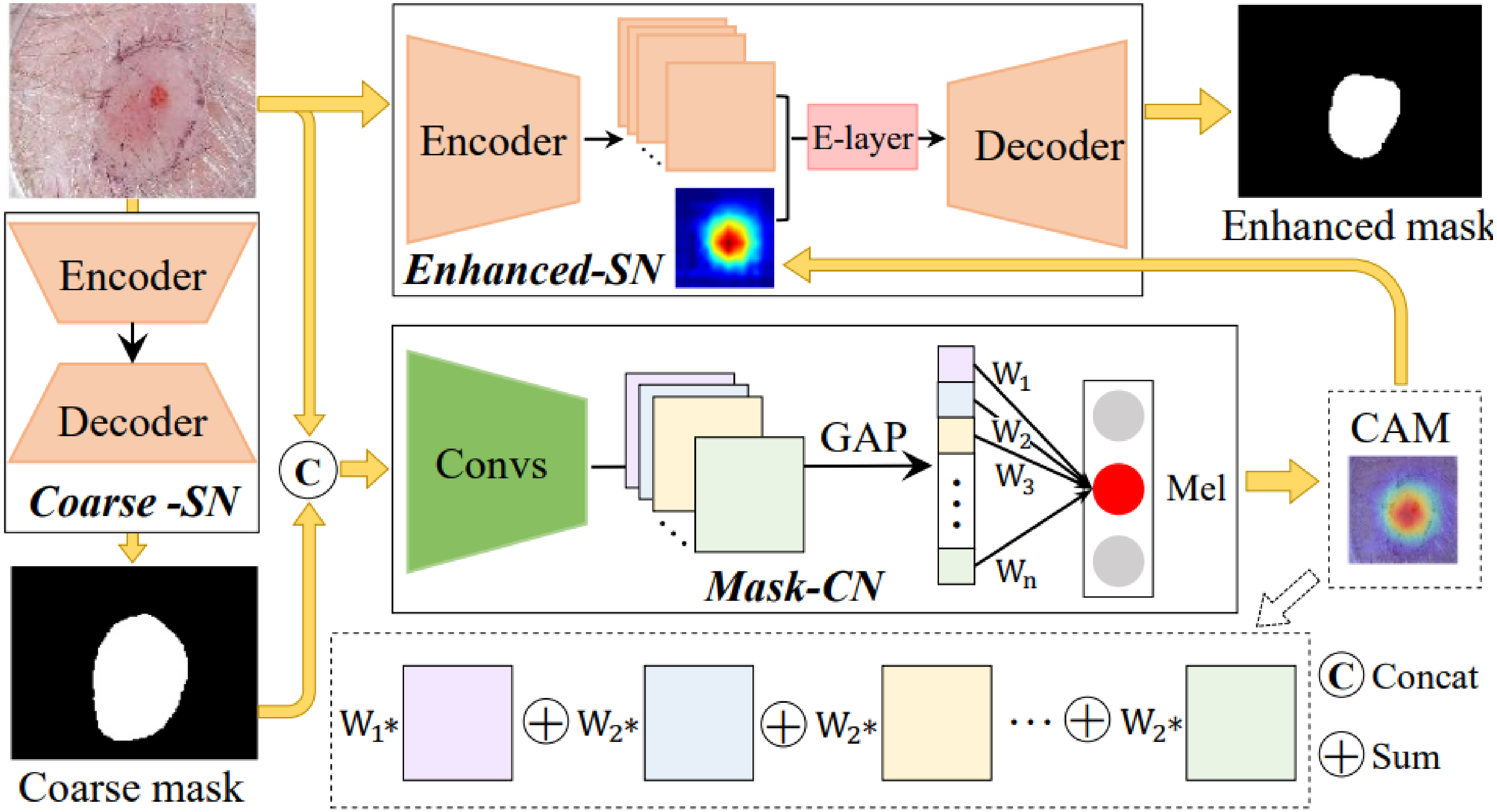


Diagram of mutual bootstrapping DCNNs (MB-DCNN) model

Yutong Xie, et al., arXiv:1903.03313.

MB-DCNN Model: Segmentation for Classification

- Transfer the masks produced by **coarse-SN to mask-CN** -> provide the prior lesion location
  (1) Training coarse-SN using pixel-level labels to roughly segment lesions

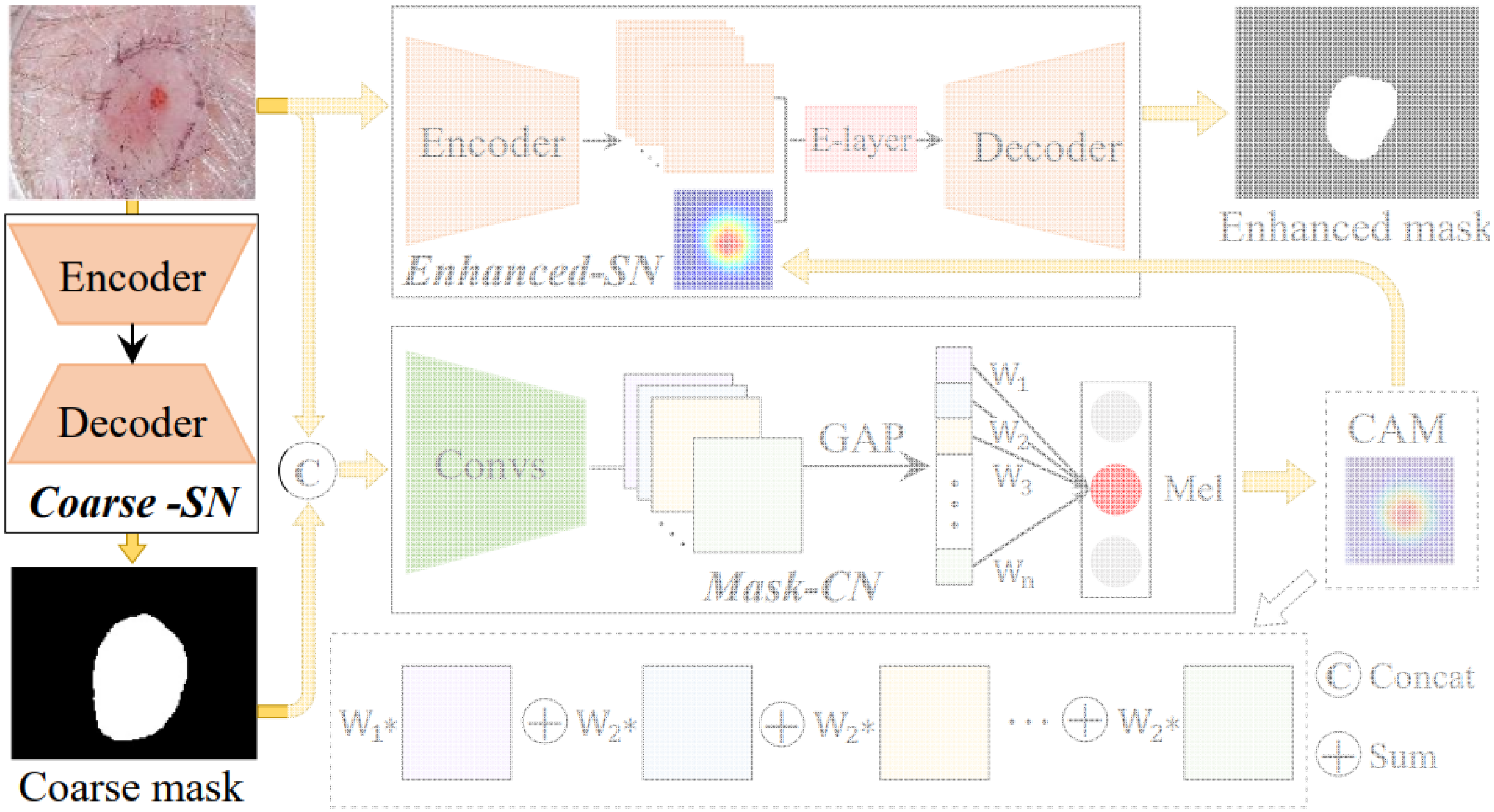


Diagram of mutual bootstrapping DCNNs (MB-DCNN) model

Yutong Xie, et al., arXiv:1903.03313.

MB-DCNN Model: Segmentation for Classification

- Transfer the masks produced by **coarse-SN to mask-CN** -> provide the prior lesion location
  (1) Training coarse-SN using pixel-level labels to roughly segment lesions
  (2) Employing coarse-SN to boost the localization ability of mask-CN

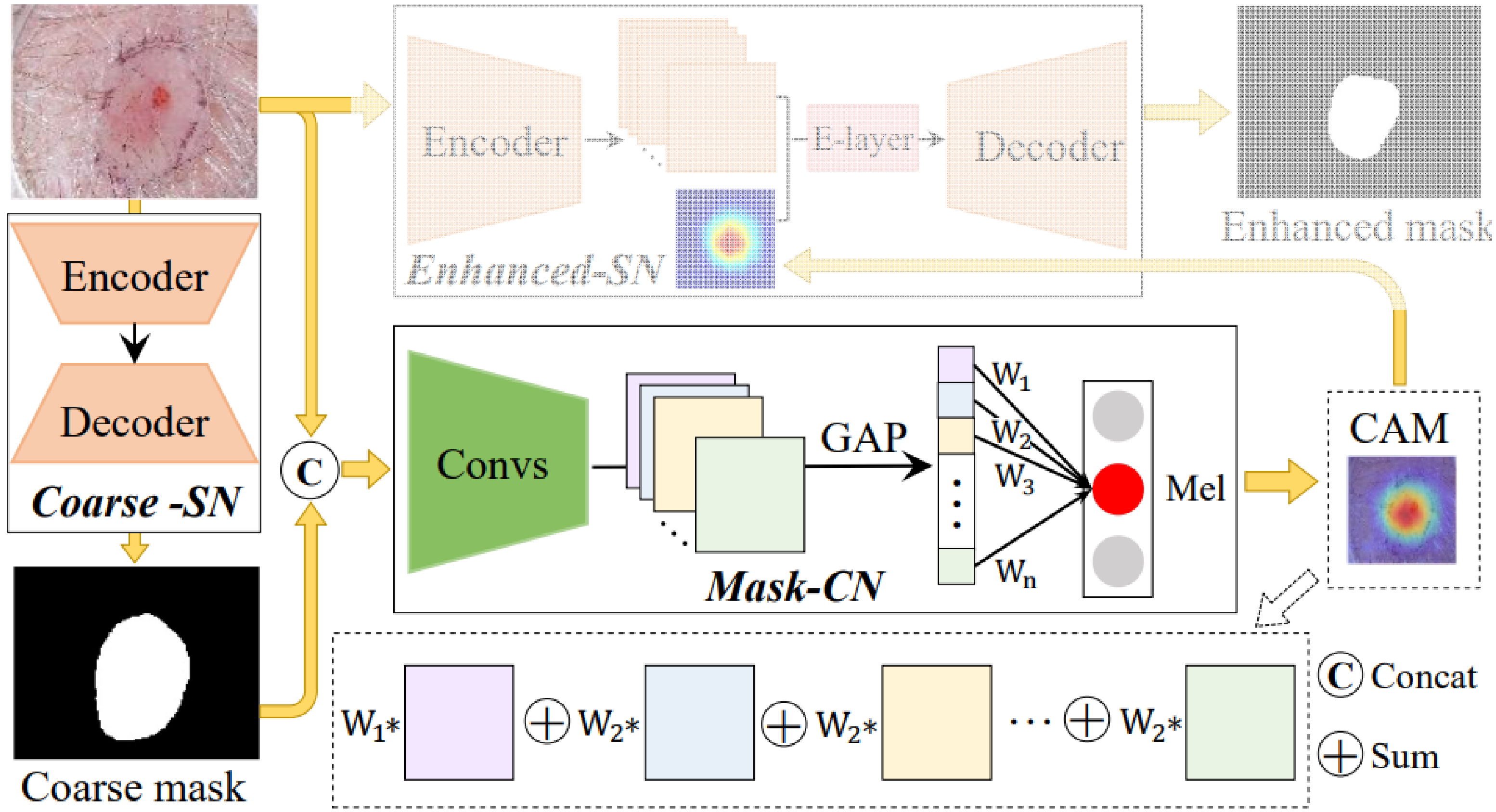


Diagram of mutual bootstrapping DCNNs (MB-DCNN) model

Yutong Xie, et al., arXiv:1903.03313.

MB-DCNN Model: Classification for Segmentation

- Leverage the **intrinsic correlation** existed in segmentation and classification tasks
- Transfer the masks produced by **coarse-SN to mask-CN** -> provide the prior lesion location
- Transfer the lesion location refined by **mask-CN to enhanced-SN** -> facilitate segmentation

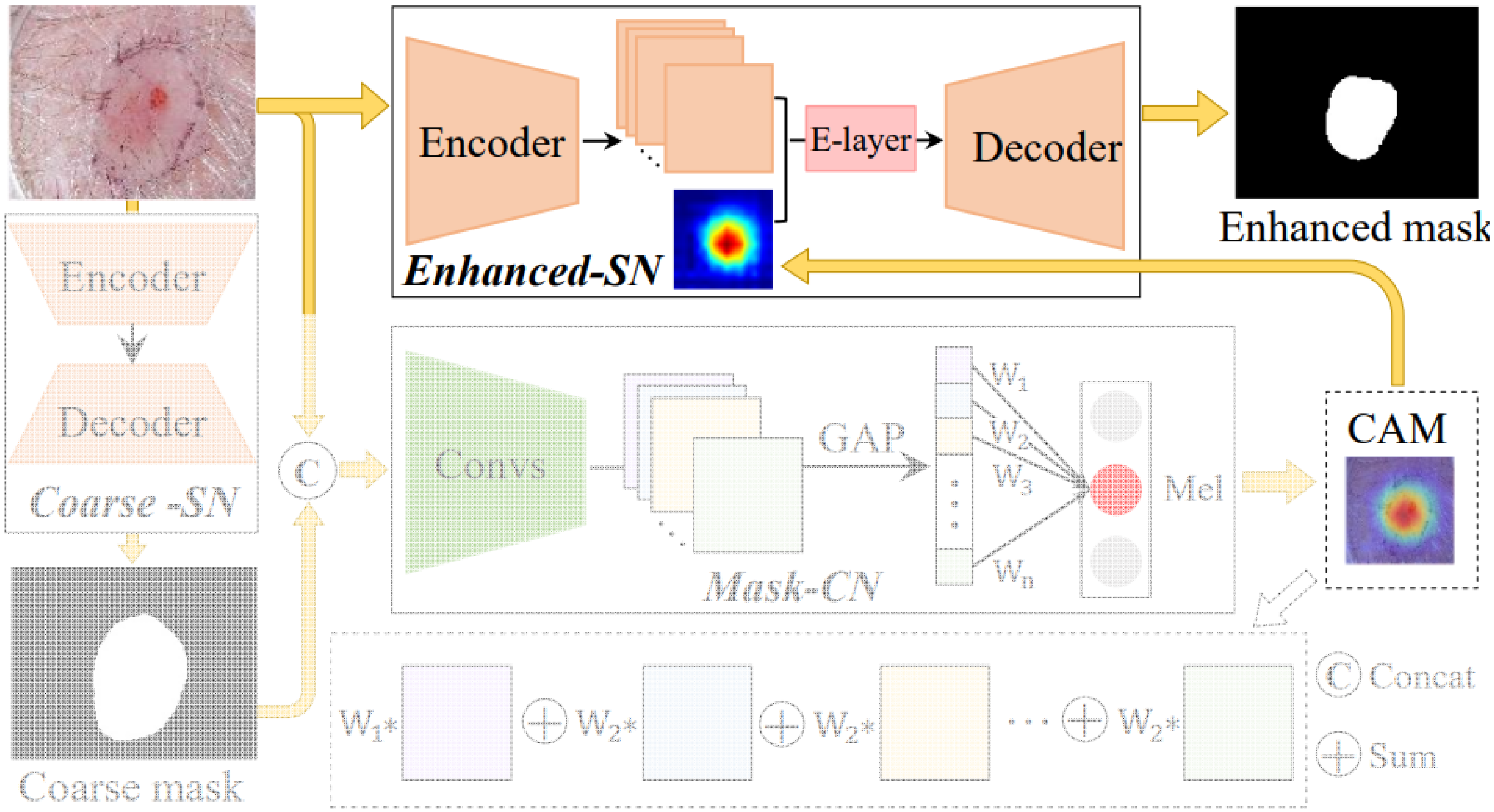


Diagram of mutual bootstrapping DCNNs (MB-DCNN) model

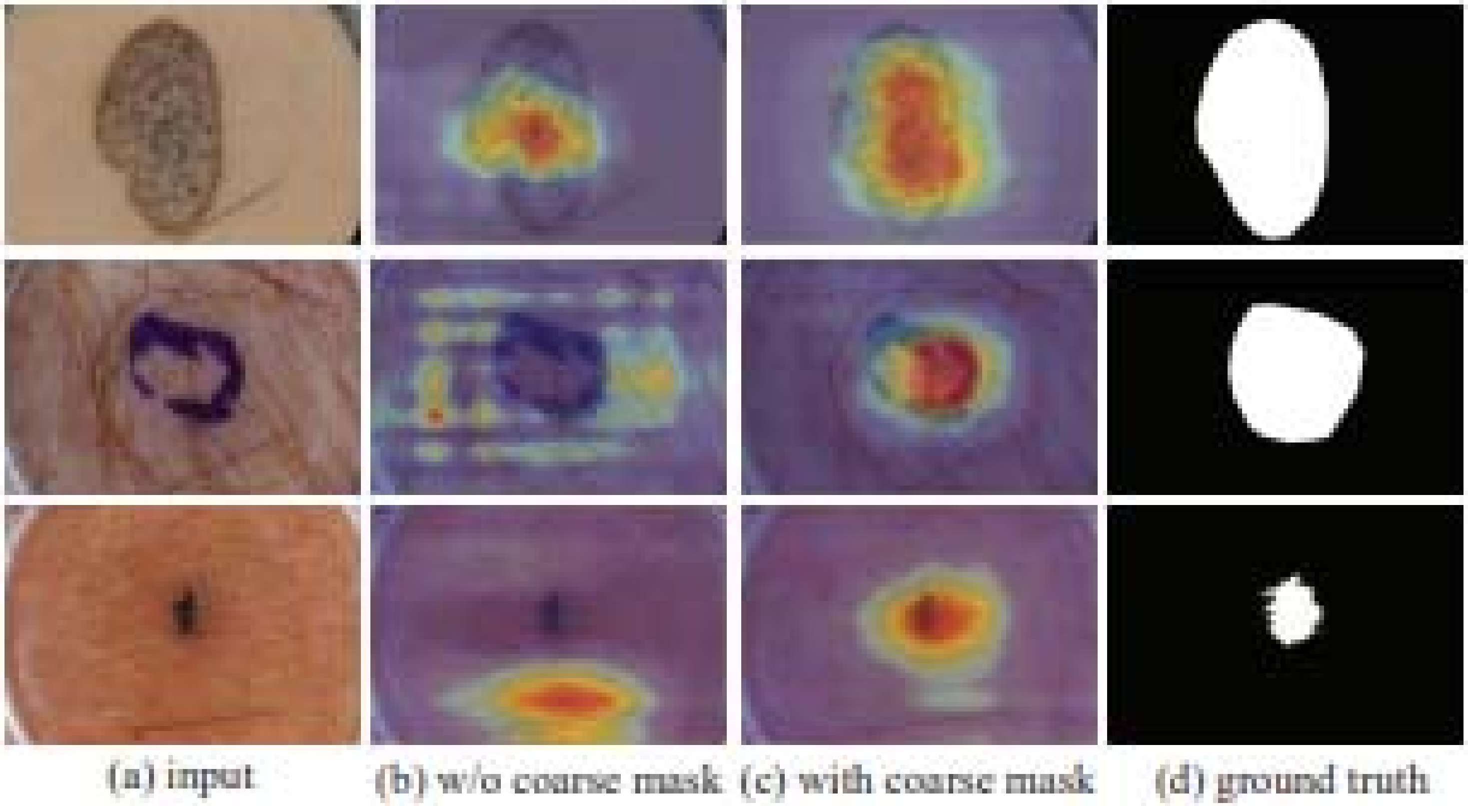
Yutong Xie, Jianpeng Zhang, **Yong Xia***, Chunhua Shen, "A Mutual Bootstrapping Model for Automated Skin Lesion Segmentation and Classification", arXiv:1903.03313.

MB-DCNN Model: Segmentation for Classification

- Transfer the masks produced by **coarse-SN to mask-CN** -> provide the prior lesion location

  (1) Training coarse-SN using pixel-level labels to roughly segment lesions

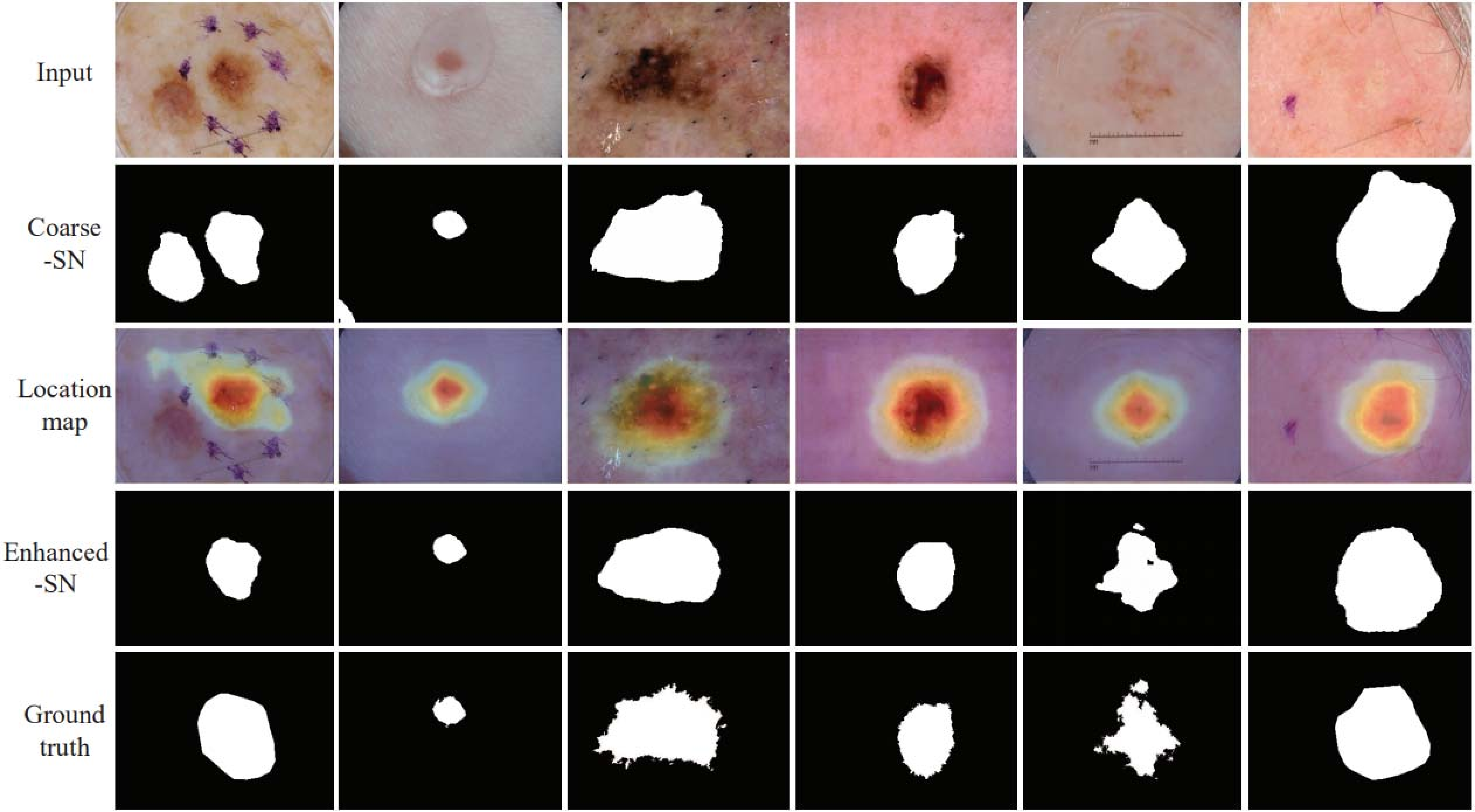
  (2) Employing coarse-SN to boost the localization ability of mask-CN



CAM obtained by mask-CN when using or not using the coarse mask

Yutong Xie, et al., arXiv:1903.03313.

MB-DCNN Model: Classification for Segmentation



Visualization of some segmentation results

Yutong Xie, et al., arXiv:1903.03313.

Segmentation Performance of MB-DCNN Model

Datasets	ISIC-2017				
Methods	CDNN, 2017 [31]	DDN, 2017 [17]	FCN+SSP, 2018 [21]	SLSDeep, 2018 [24]	**Our SWSDB**
JA	76.5	76.5	77.3	78.2	**80.4**
DI	84.9	86.6	85.7	87.8	**87.8**
AC	93.4	93.9	93.8	93.6	**94.7**
SE	82.5	82.5	85.5	81.6	**87.4**
SP	97.5	**98.4**	97.3	98.3	96.8

Datasets	PH2				
Methods	mFCNPI, 2017 [6]	RFCN, 2017 [30]	SLIC, 2018 [22]	Our SWSDB	**Fine-tuned SWSDB**
JA	84.0	-	-	86.7	**89.4**
DI	90.7	93.8	-	92.6	**94.2**
AC	94.2	-	90.4	95.8	**96.5**
SE	94.9	-	91	**97.9**	96.7
SP	94.0	-	89.7	**95.1**	94.6

Yutong Xie, et al., arXiv:1903.03313.

Classification Performance of MB-DCNN Model

- ISIC-2017 Challenge Dataset

  Training dataset：pixel-label: 2000 image-label: 1320

  Validation：150 Testing：600

- PH2 Dataset:

  Pixel-label：200

  Image-label：200

Methods	Melanoma Classification				Keratosis Classification				Average
	AC	SE	SP	AUC	AC	SE	SP	AUC	AUC (%)
Ours	87.8	72.7	91.5	90.3	93.0	84.4	94.5	97.3	**93.8**
ARL-CNN [29], 2019	85.0	65.8	89.6	87.5	86.8	87.8	86.7	95.8	91.7
SSAC [49], 2019	83.5	55.6	90.3	87.3	91.2	88.9	91.6	95.9	91.6
SDL [19], 2019	88.8	-	-	86.8	92.5	-	-	95.8	91.3
#1 [50]	82.8	73.5	85.1	86.8	80.3	97.8	77.3	95.3	91.1
#2 [51]	82.3	10.3	99.8	85.6	87.5	17.8	99.8	96.5	91.0
#3 [52]	87.2	54.7	95.0	87.4	89.5	35.6	99.0	94.3	90.8
#4 [53]	85.8	42.7	96.3	87.0	91.8	58.9	97.6	92.1	89.6
#5 [54]	83.0	43.6	92.5	83.0	91.7	70.0	99.5	94.2	88.6

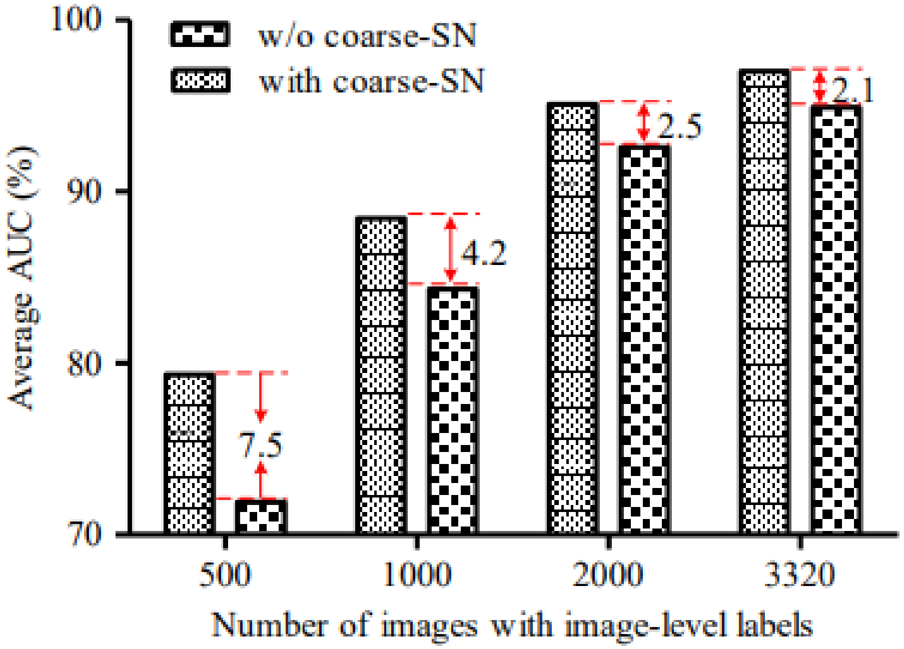
ISIC -2017

Methods	AC	SE	SP	AUC
Ours (Fine-tuned)	94.0	95.0	93.8	**97.7**
Ours	88.5	82.5	90.0	95.6
CICS [55], 2017	-	100.0	88.2	-
MFLF [56], 2015	-	98.0	90.0	-
CCS [57], 2015	-	92.5	76.3	84.3

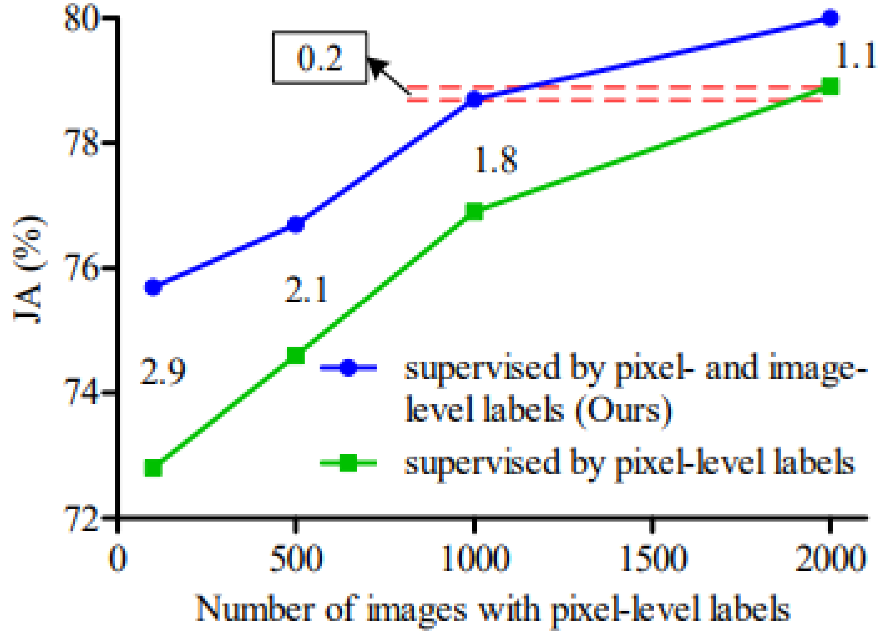
PH2

Yutong Xie, et al., arXiv:1903.03313.

Advantages of mutual bootstrapping



Average AUC values obtained on the ISIC-2017 validation set by our MB-DCNN model with or without the help of coarse-SN

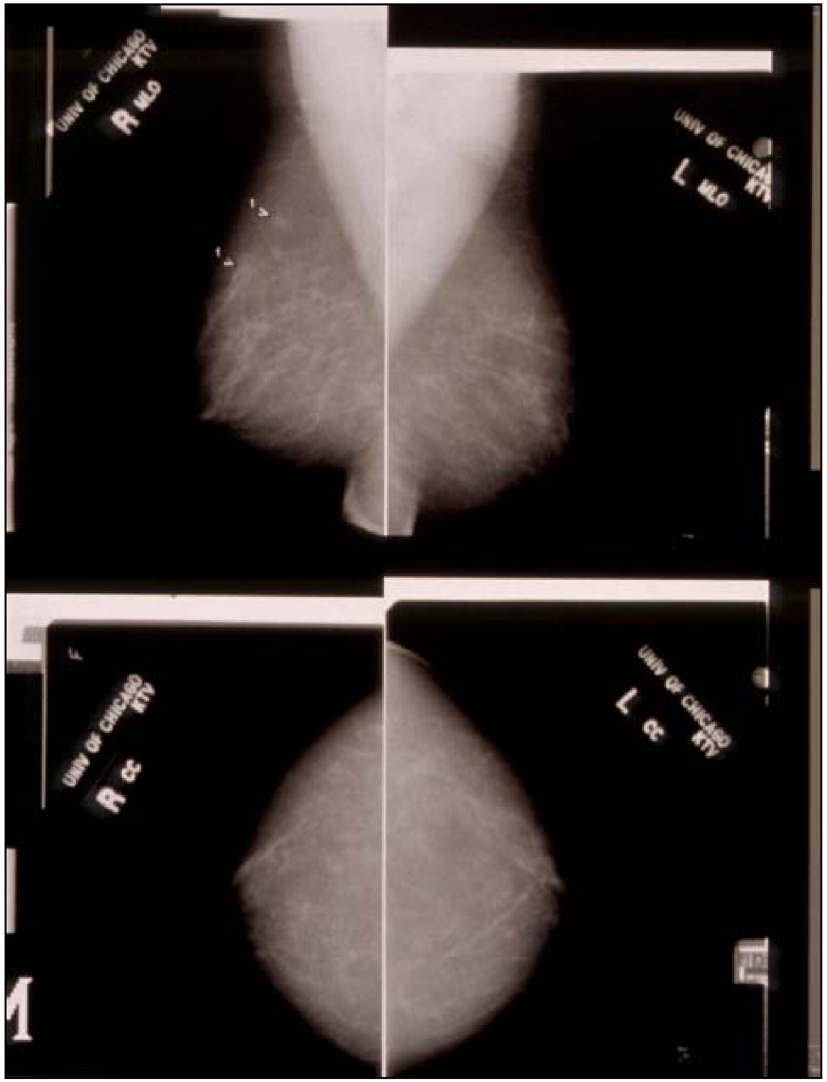


JA values of our MB-DCNN model and its full supervised counterpart on the ISIC-2017 validation set versus the number of training images with pixel-level labels.

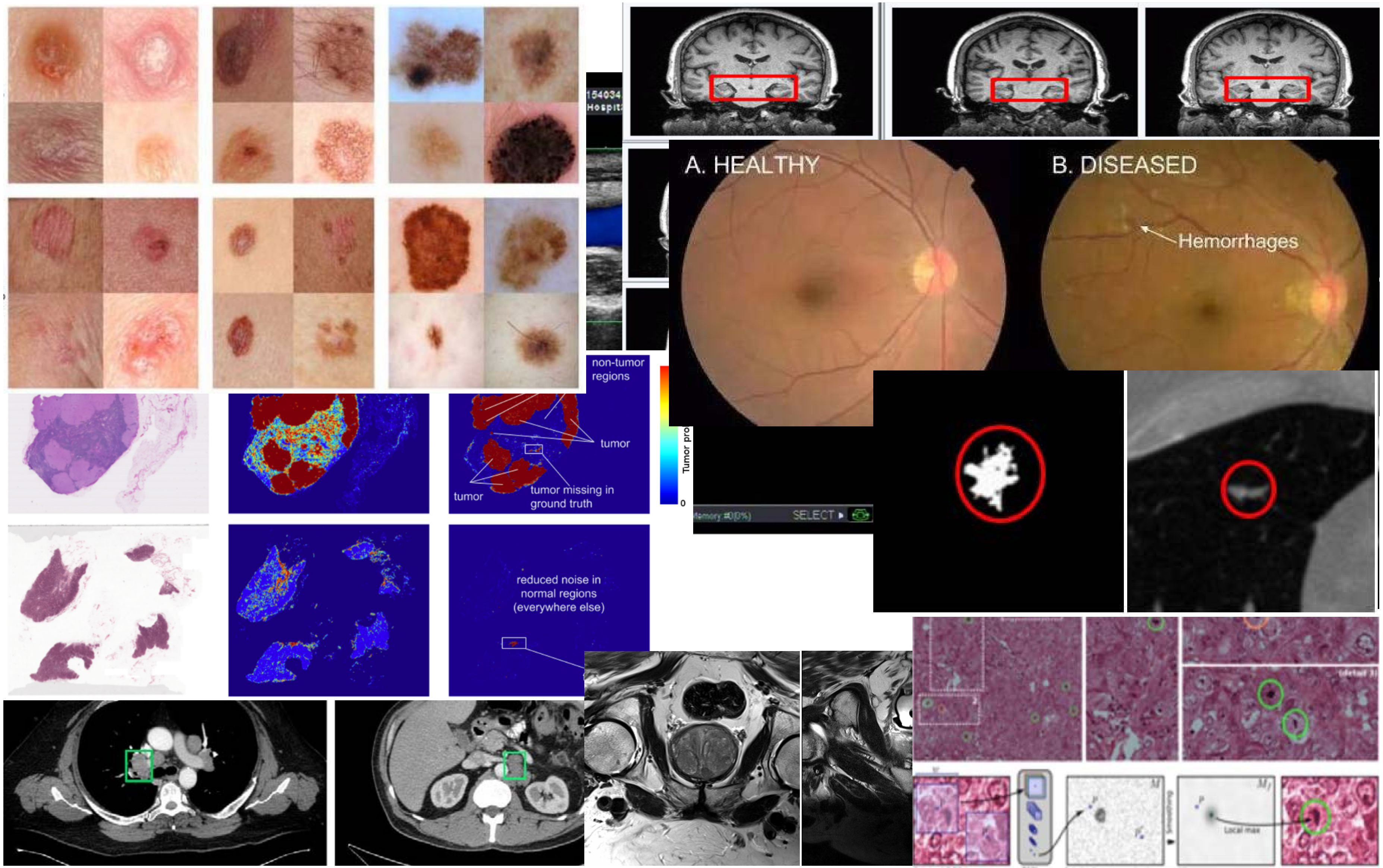
Yutong Xie, et al., arXiv:1903.03313.

Yesterday: First CAD System

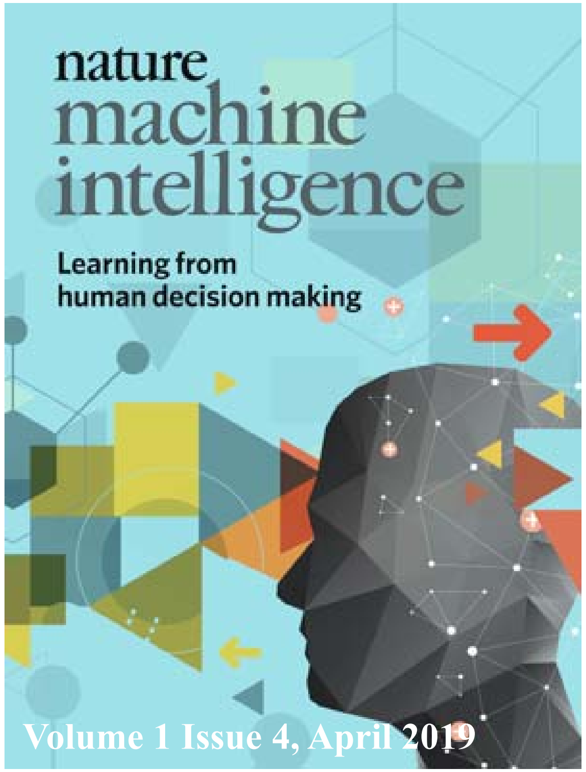
- The prototype CAD system for the detection of breast cancer on mammograms was developed in 1994 at the University of Chicago.
- The first CAD system for screening mammography was approved by the United States Food and Drug Administration in 1998.



Today：Breakthroughs in Medical Image Analysis



Challenges of CAD: Mistakes and Prejudices



4月《自然-机器智能》封面故事：人工智能和机器学习系统可能在许多任务上远超人类表现，但它们也***可能模仿或放大人类的错误和偏见***。心理学家数十年来对人类决策判断中的错误和偏见的发生及预防所开展的研究，在心理学和机器学习文本之间建立了联系，为发展和改善机器学习算法指明了方向。

Learning from human decision making

Artificial intelligence and machine learning systems may surpass human performance on a variety of tasks, but they may also mimic or amplify human errors or biases. This issue of *Nature Machine Intelligence* features a Perspective development and prevention of errors and biases in human judgment and decision making. The authors provide connections between the psychology and machine learning literatures, and offer guideposts for the development and improvement of machine learning algorithms.

See Alexander S. Rich and Todd M. Gureckis

Data Sampling Bias

- **Discrepancy between the distributions of training data and real data**

Computer-Aided Early Diagnosis of Dementia

Research – Most research focuses on Alzheimer‘s disease (AD) and mild cognitive impairment (MCI)

Clinics – There are many types of dementia, such as

Common Types

Alzheimer's disease (50 - 70%)

Vascular dementia (25%)

Lewy body dementia (155)

Frontotemporal dementia

Less Common Types

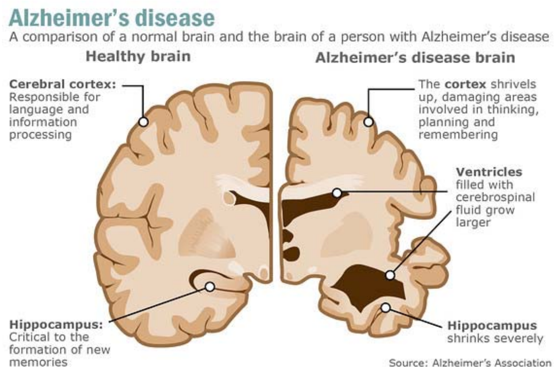
Normal pressure hydrocephalus

Parkinson's disease dementia

Syphilis

Creutzfeldt–Jakob disease

Others

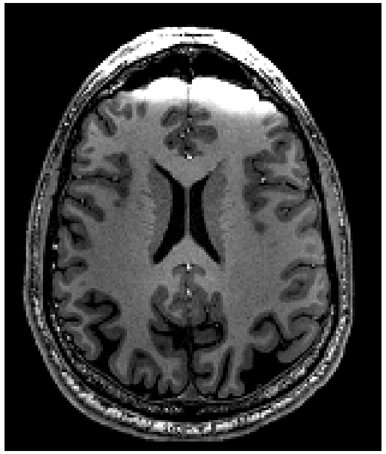


Dementia – A progressive neural degenerative disease with insidious onset

Dementia worldwide

Data Quality Bias

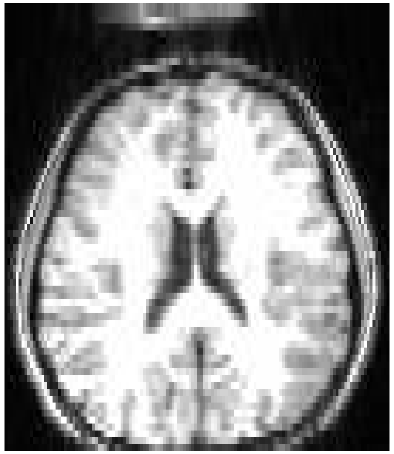
- **Variations of medical image quality comes from**
  - Different medical centers (operators)
  - Different scanners (types / vendors)
  - Different protocols
  - …



7T MR Image

3T MR Image

1.5T MR Image
IBSR V2.0

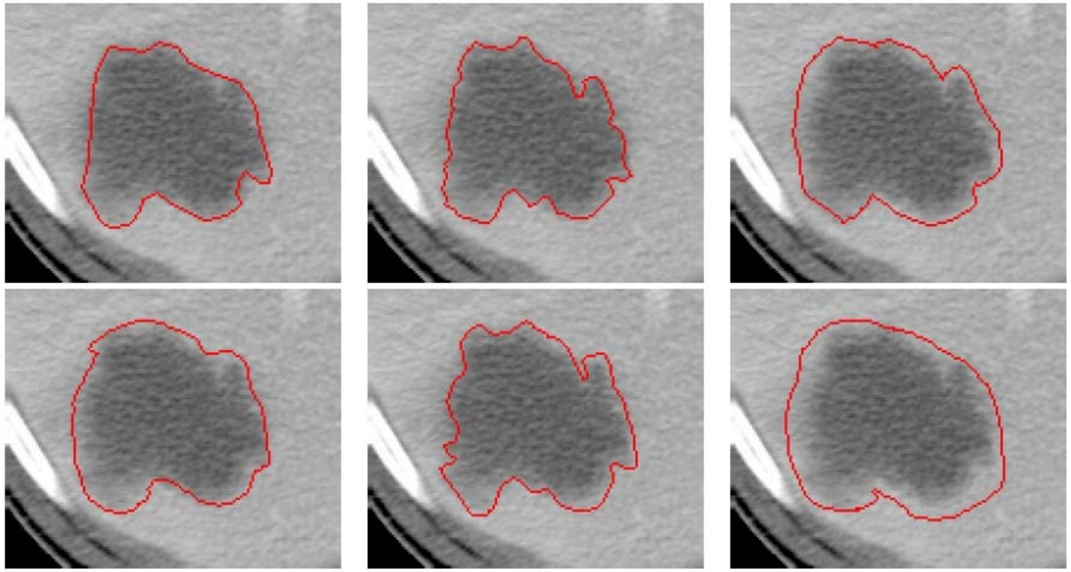


1.5T MR Image
IBSR Normal

Data Annotation Bias

- **Inter- and intra-observer variability of manual annotation of medical images**

Observer 1 Observer 2 Observer 3



Three trained observers (two radiologists and one radiotherapist) delineated this lesion twice with an interval of about one week (upper row and bottom row respectively). The area of delineated areas varies up to 10% per observer and the difference between observers amounts to more than 20%.

- Annotation of class labels is also difficult (e.g. the malignancy level of lung nodules)
- Consistent conclusions can be reached on easy cases – a cause of data sampling bias

Tomorrow: Challenges and Opportunities

Now this is not the end.
It is not even the beginning of the end.
But it is, perhaps, the end of the beginning.

-- Sir Winston Churchill

2017

英国未来制造业预测

2016

2015

新工业法国

2014

国家战略性新兴产业发展规划

2013

2012

德国工业4.0

新一代人工智能发展规划

"互联网+"行动

美国GE工业互联网

国家科技创新规划

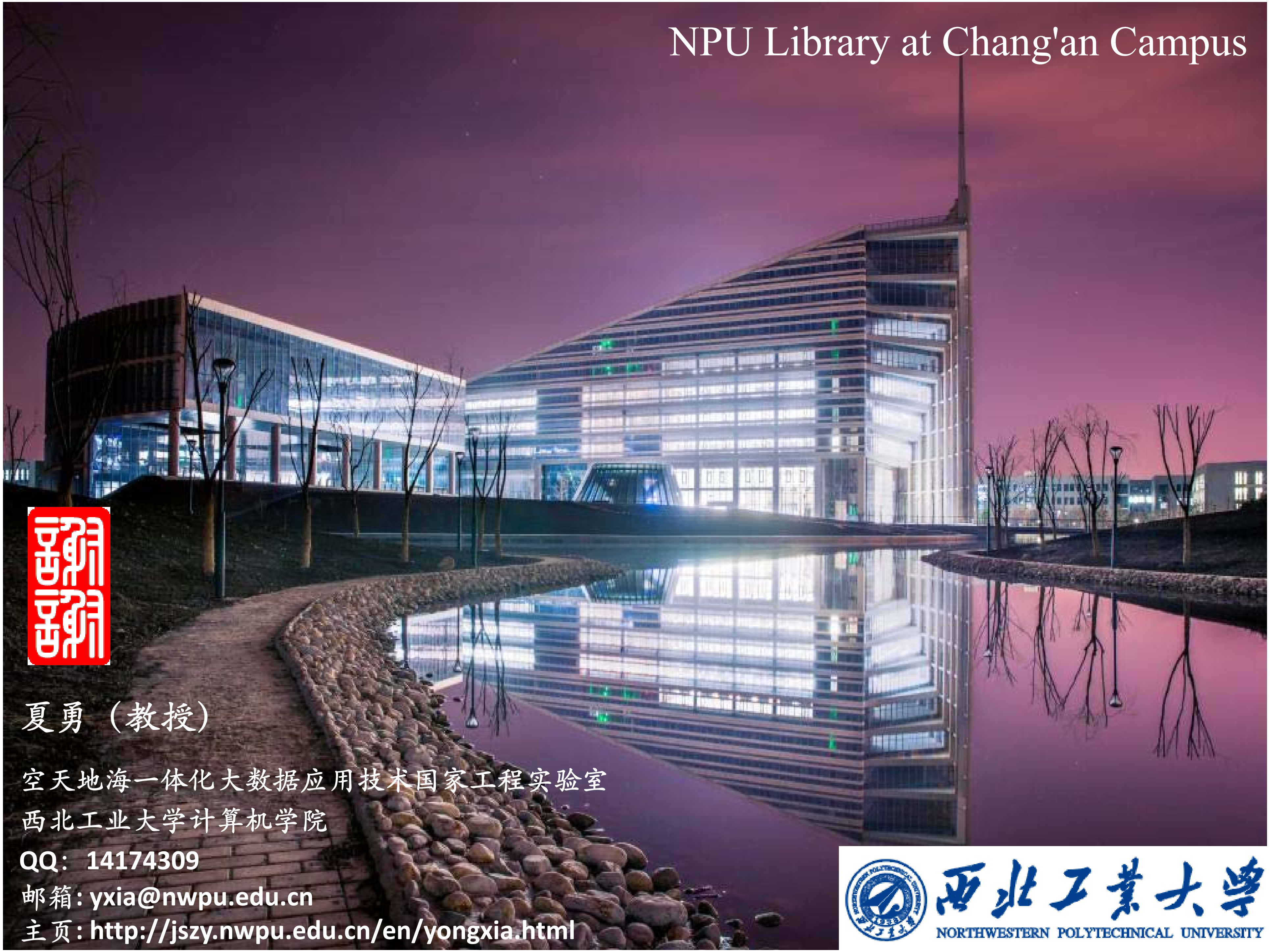
美国先进制造业国家战略

“健康中国2030”规划纲要

中国制造2025

----引自工信部王喜文《中国制造 2025 解读：从工业大国到工业强国》

The Advent of the 4th Industrial Revolution



谢谢

夏勇（教授）

空天地海一体化大数据应用技术国家工程实验室

西北工业大学计算机学院

QQ：14174309

邮箱: yxia@nwpu.edu.cn

主页: http://jszy.nwpu.edu.cn/en/yongxia.html

西北工業大學

NORTHWESTERN POLYTECHNICAL UNIVERSITY