MLA’18 -The 16th China Symposium on Machine Learning and Applications

从谱聚类到自注意力模型

—谈经典机器学习在深度学习时代的新形态

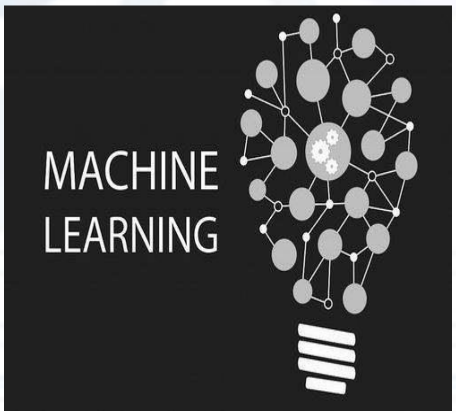
张兆翔

中国科学院自动化研究所

2018年11月3日，南京

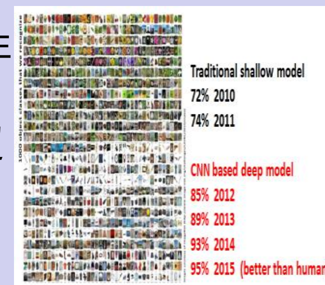
机器学习引领人工智能发展

人工智能近年来得到广泛关注，在**感知**、**交互**、**决策**等若干具体应用问题上甚至媲美人类性能，这得益于机器学习的发展与进步。



监督学习
集成学习
强化学习
主动学习
深度学习
.....

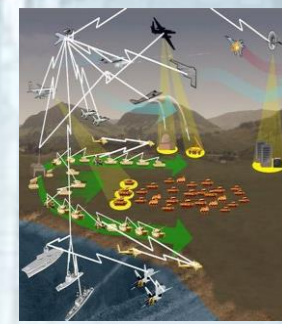
感知：2015年ImageNet的识别准确度已经超过人类。



交互：2018年Google智能语音助手既能听懂人说话，说的话又像人。

决策：2017年AlphaZero自学成为围棋顶尖高手。

工业制造



军事国防

AI+

家居

出行

生活服务

社会管理

机器学习的历史回顾

回顾机器学习数十年的历史，可以说是理论日益丰富，方法层出不穷，体系不断完善。

句法学习 Q学习 核PCA 谱聚类 迁移学习 贝叶斯程序学习 元学习

分类与回归树 主动学习

线性判别分析 Capsule

随机森林

AdaBoost AlexNet

决策树 对抗生成网络

PCA 贝叶斯网络 流形学习 RBM自编码器 DQN

可解释的深度学习

形式语言理论

支持向量机 递归神经网络 变分自编码器

EM算法 降噪AE

ResNet

贝叶斯分类器 子空间学习

卷积神经网络 DropOut

VGG Net

BP算法

DenseNet

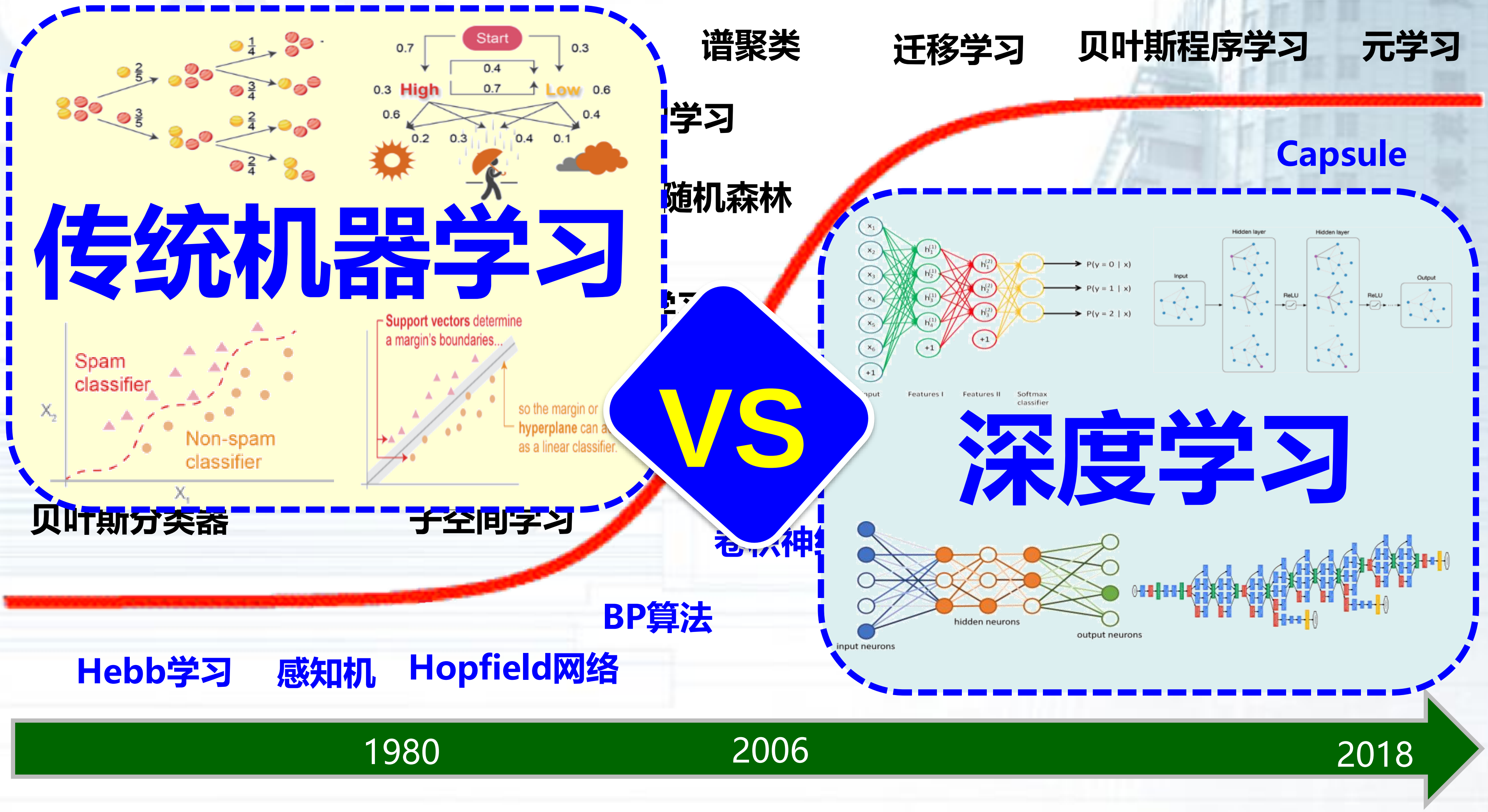
图卷积 R-CNN

Hebb学习 感知机 Hopfield网络

1980 2006 2018

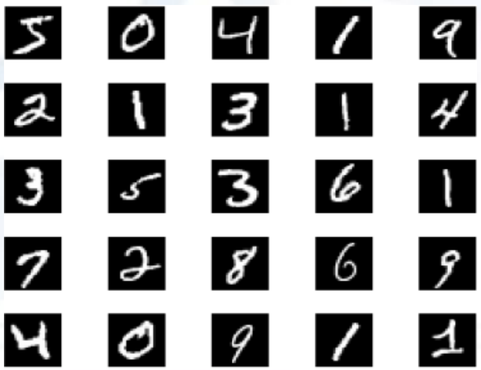
机器学习的历史回顾

回顾机器学习数十年的历史，可以说是理论日益丰富，方法层出不穷，体系不断完善。



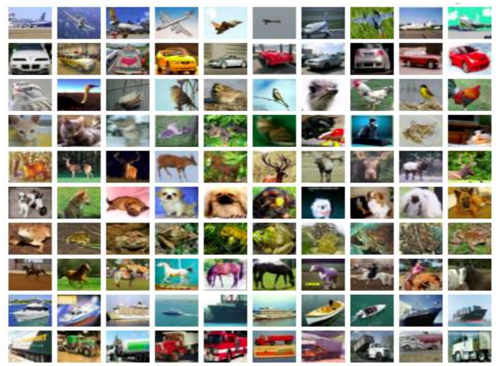
视觉感知与理解

视觉感知与理解一直以来与机器学习理论方法的引入密不可分。
以视觉物体识别为例：



MINST:10类，6万张图片

1998年



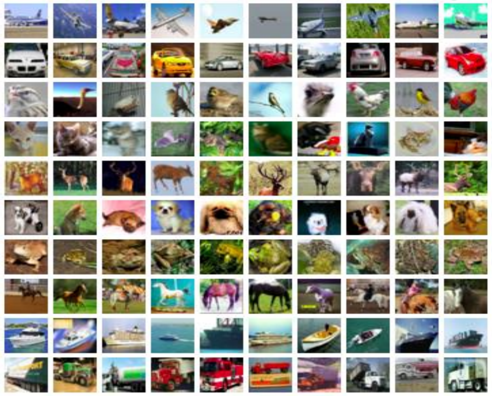
Tiny Images:7.5万类，7900万

2006年

SUN:397类，10万张

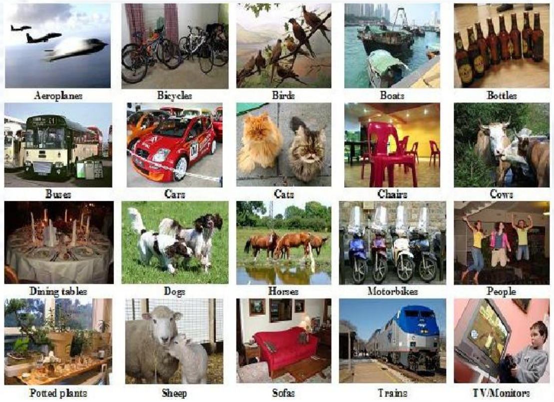
2010年

2004



Caltech101/256: 每类至少80张

2007年



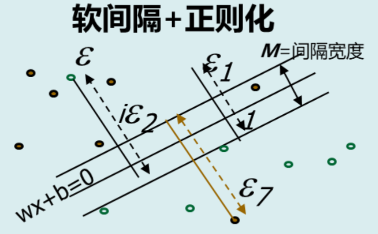
PASCALVOC2007：20类,9963张

2009年

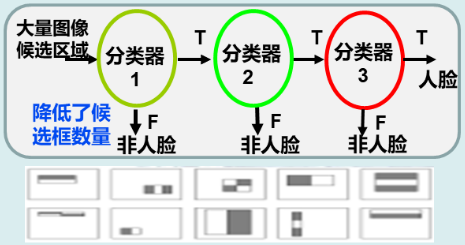
ImageNet：22k类，14M

视觉感知与理解

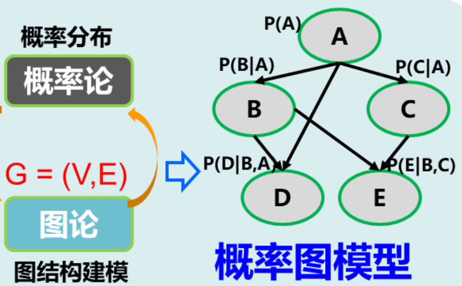
以视觉物体识别为例：



支持向量机(SVM)



Adaboost方法



概率图模型

小样本

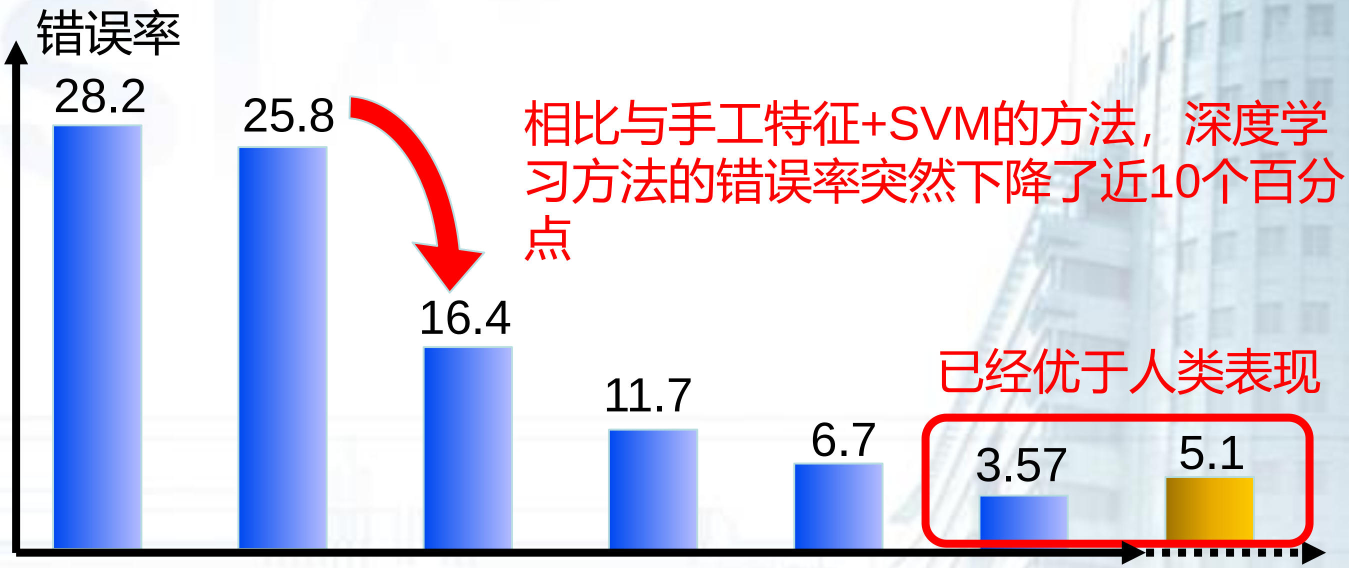
2012年

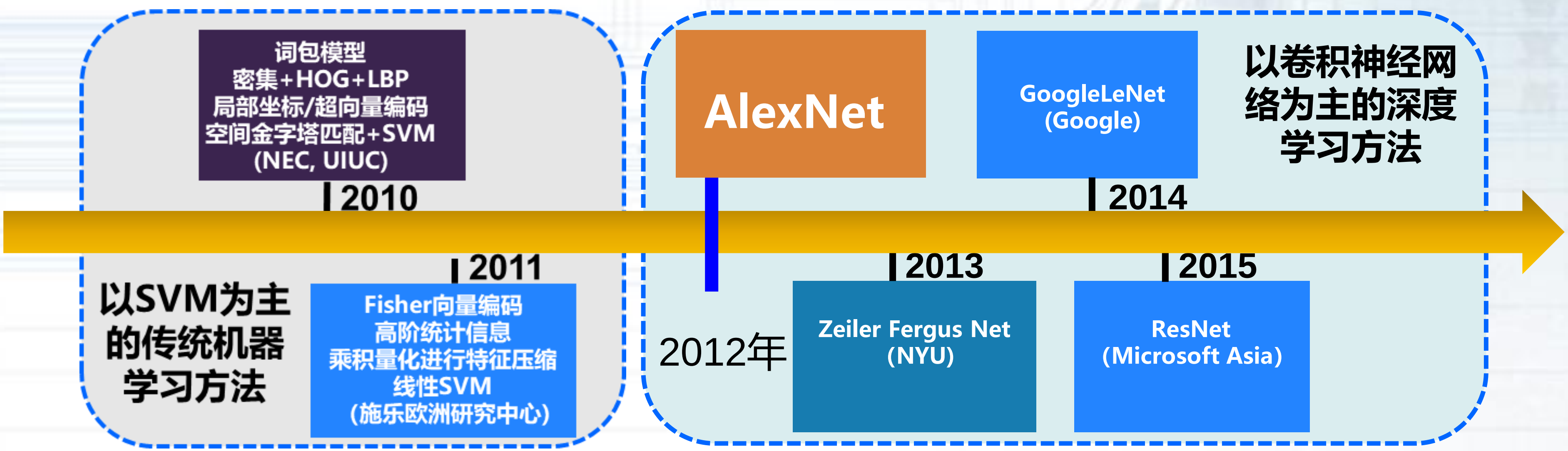
大数据

视觉感知与理解

以视觉物体识别为例：

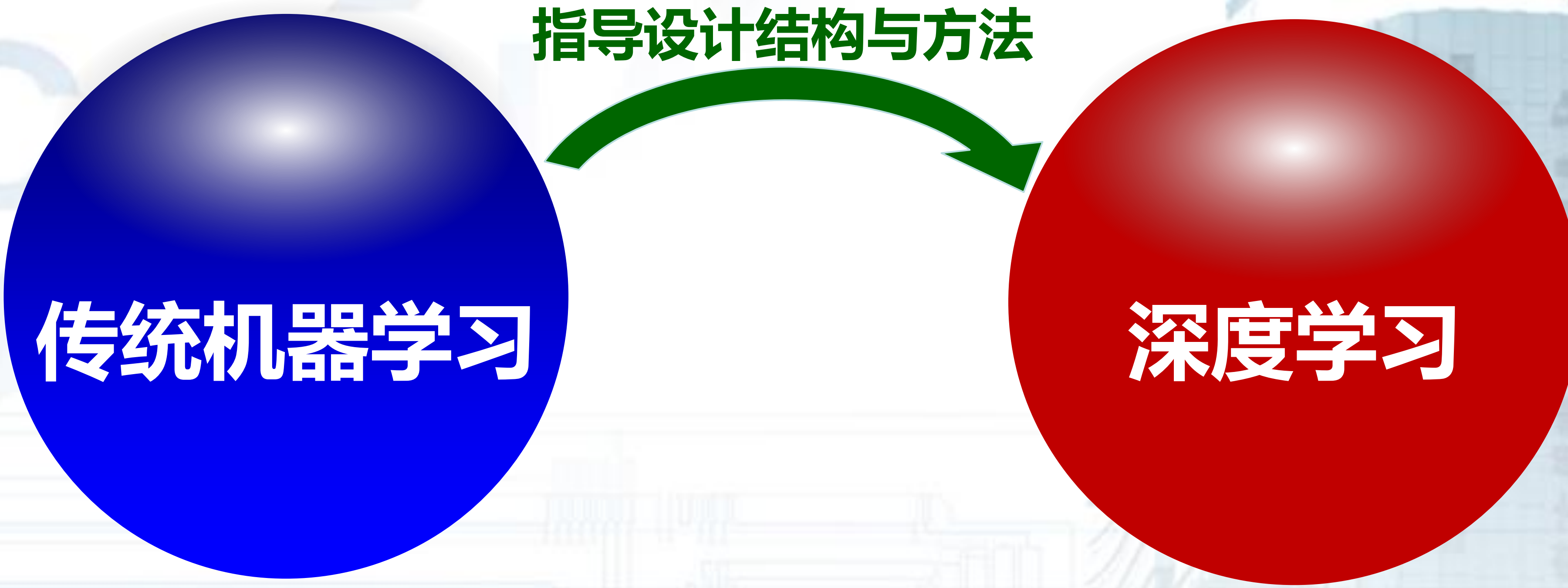
IMAGENET
Large
Scale
Visual
Recognition
Challenge
ILSVRC
2010~2017





ECCV2018文章关键字

传统机器学习与深度学习



传统机器学习：

理论性强

解释性好

不能充分利用大数据

深度学习：

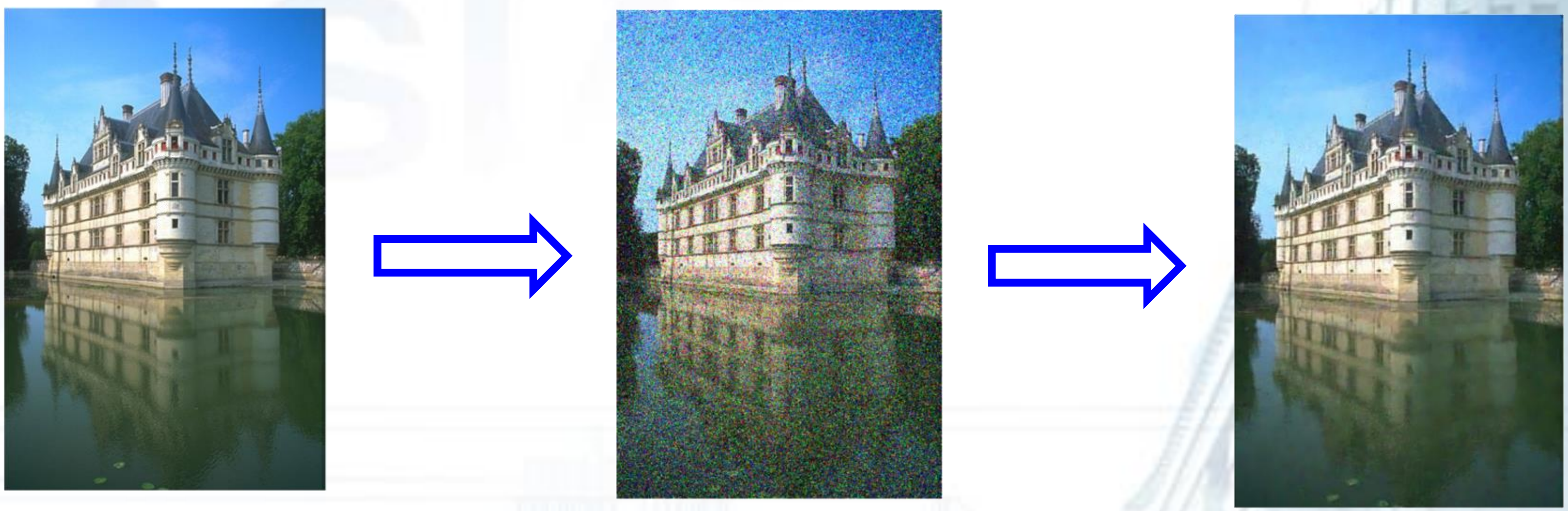
理论性弱

解释性差

能够充分利用大数据

结构：from Sparse Coding

以图像去噪为例：



(a) Original (b) Noisy (c) Denoised

SPARSE ASSUMPTIONS：

Natural image (patch) can be well represented as linear combination of **few** basis.

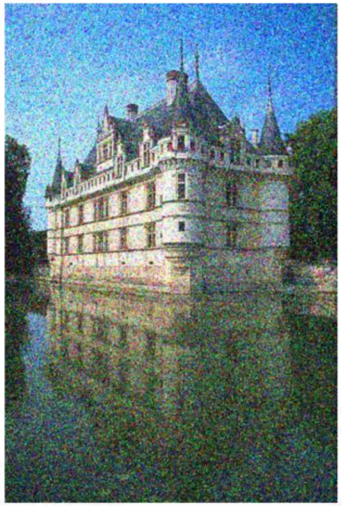
Noisy images generally do **NOT** follow the sparse assumption.

Mairal J, Elad M, Sapiro G. Sparse representation for color image restoration[J]. IEEE Transactions on image processing, 2008, 17(1): 53-69.

结构：from Sparse Coding

以图像去噪为例：

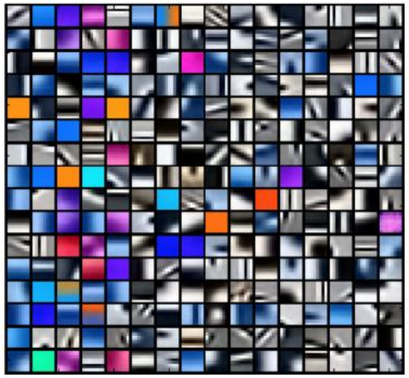
(a) Original



(b) Noisy

(c) Denoised

$$E_{W_d}(X, Z) = \frac{1}{2}||X - W_d Z||_2^2 + \alpha||Z||_1$$

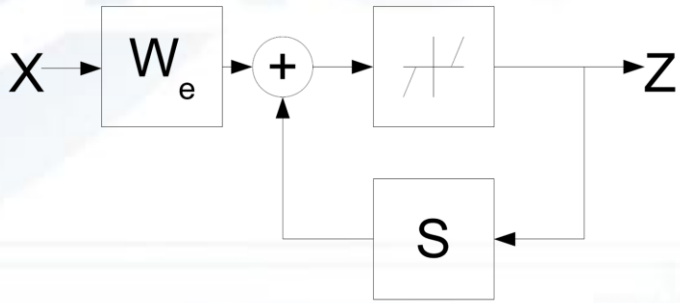


$$\begin{pmatrix} 0 \\ 0 \\ 0.3 \\ 0 \\ 0 \\ 0.8 \\ \dots \\ 0 \\ 0.5 \\ 0 \end{pmatrix}$$

结构：from Sparse Coding

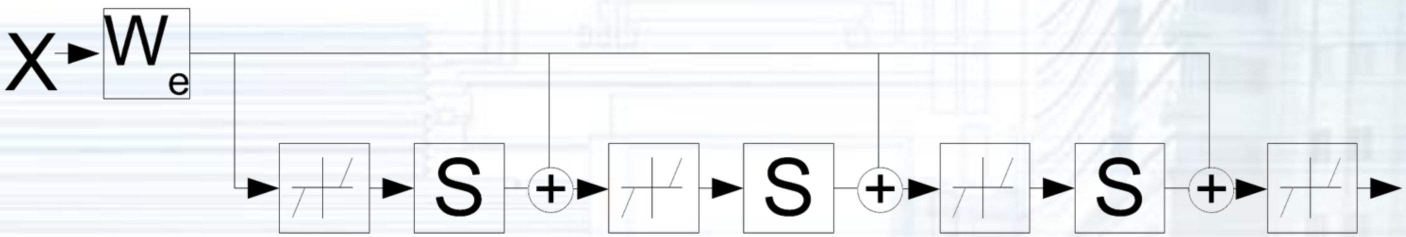
ISTA Algorithm：

$$Z(k+1) = h_\theta\left(W_e X + SZ(k)\right) \qquad Z(0) = 0$$



1. Slow testing
2. Difficult to transfer across Dataset
3. Inference time may fluctuate

Deep Learning Solution:

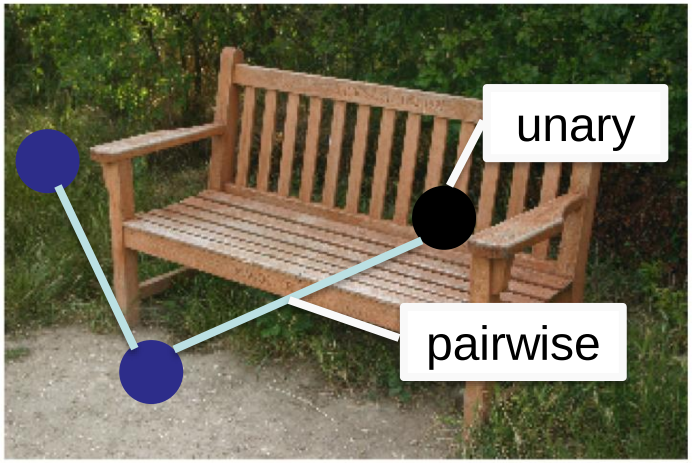


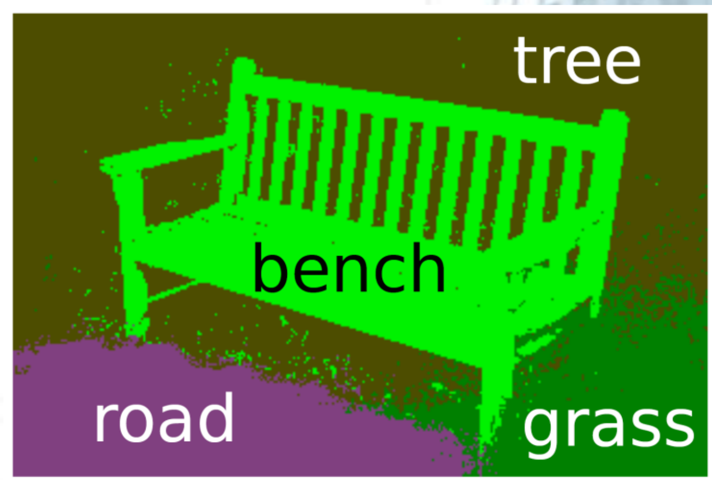
1. Fast inference
2. Easy to fine-tune across Dataset

K. Gregor and Y. LeCun. Learning Fast Approximations of Sparse Coding, ICML 2010.

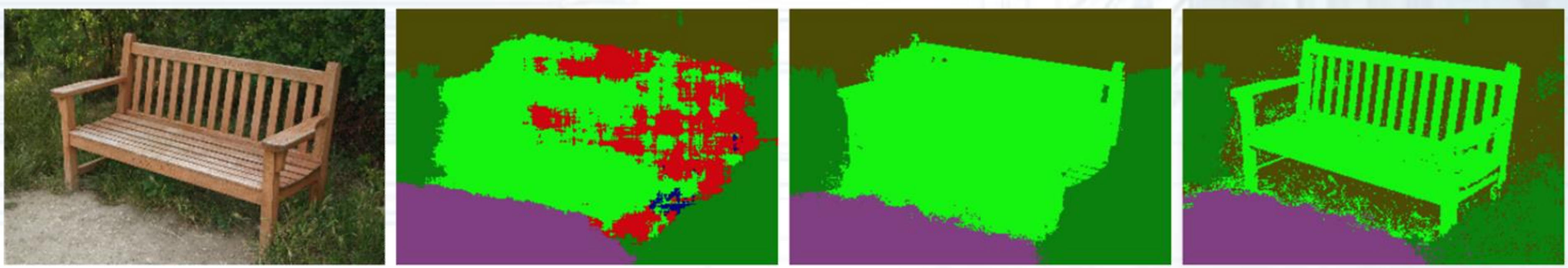
结构：from CRF

以图像分割为例：





$$E(\mathbf{x}) = \sum_{i} \psi_u(x_i) + \sum_{i<j} \psi_p(x_i, x_j)$$

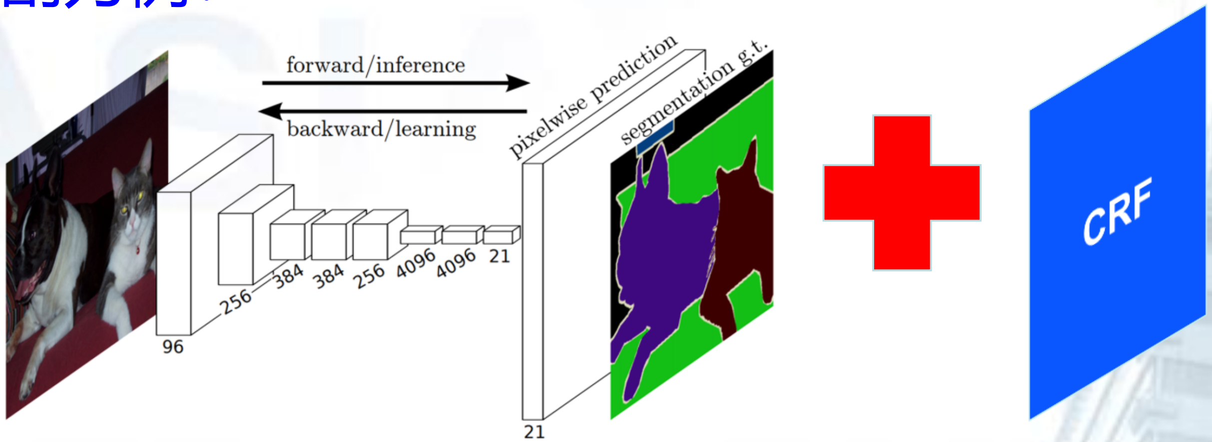


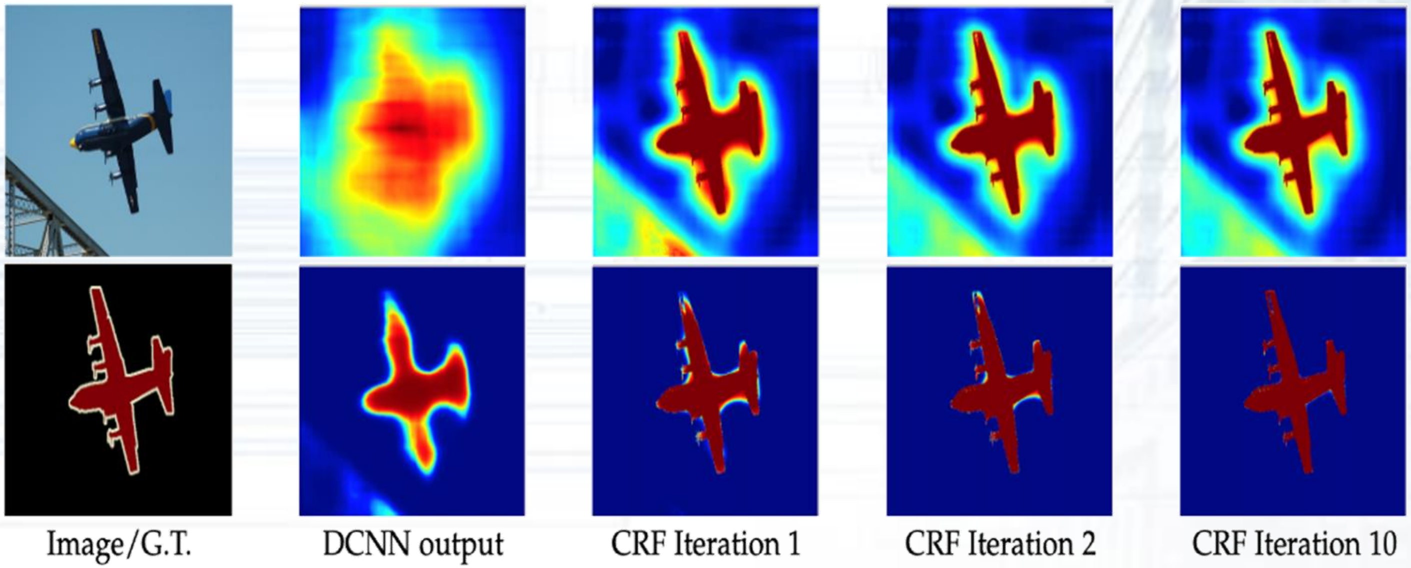
(a) Image (b) Unary classifiers (c) Robust P^n CRF (d) Fully connected CRF

Krähenbühl P, Koltun V. Efficient inference in fully connected crfs with gaussian edge potentials[C]//Advances in neural information processing systems. 2011: 109-117.

结构：from CRF

以图像分割为例：

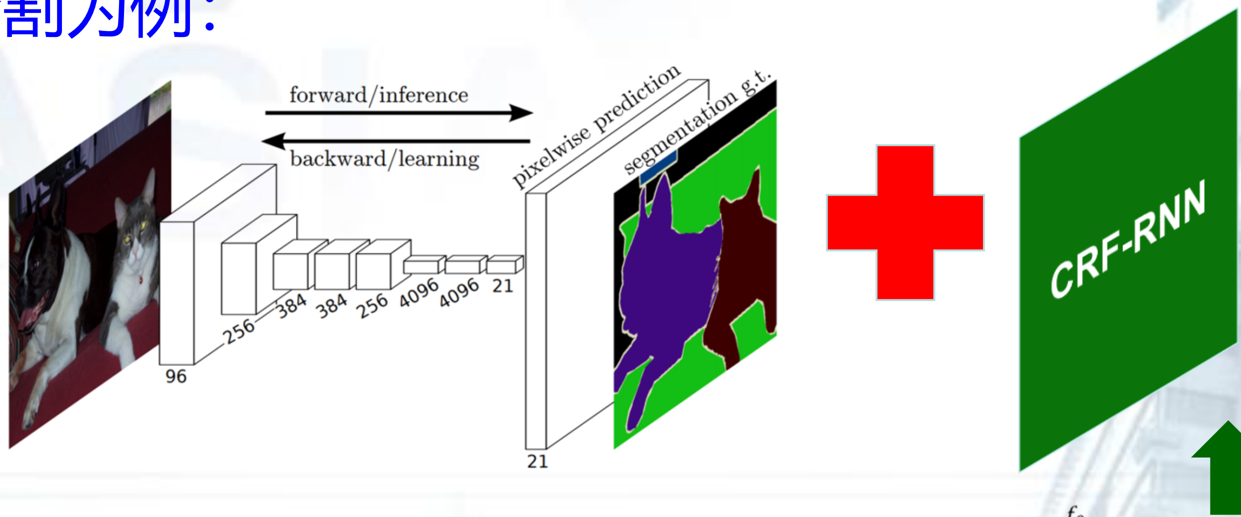


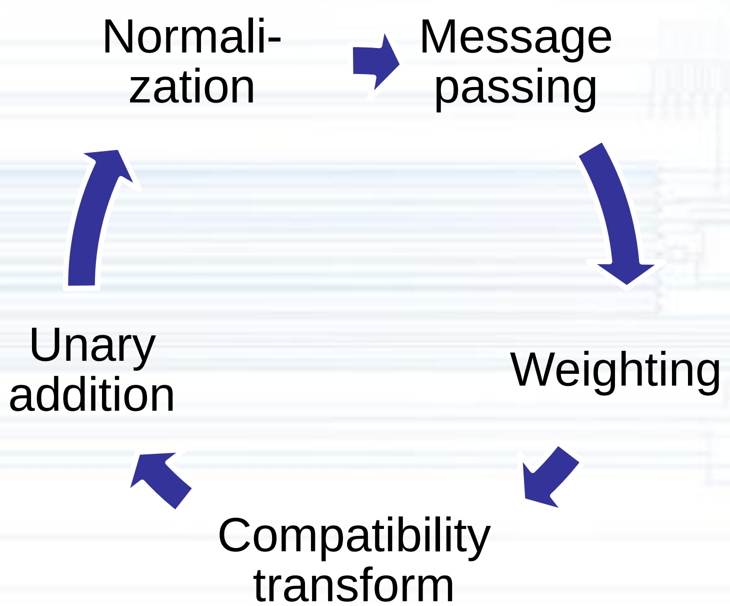


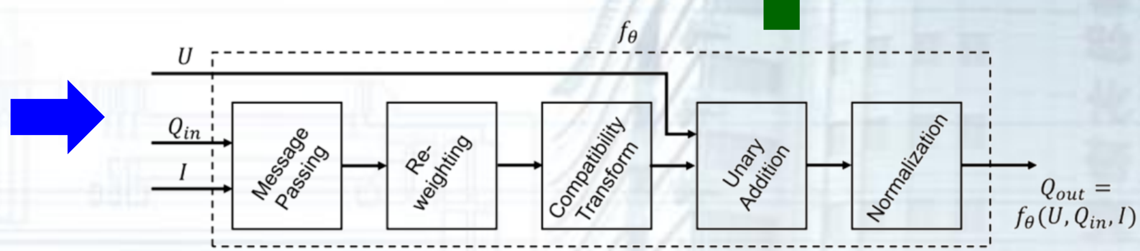
Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431-3440.
Chen L C, Papandreou G, Kokkinos I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs[J]. IEEE transactions on pattern analysis and machine intelligence, 2018, 40(4): 834-848.

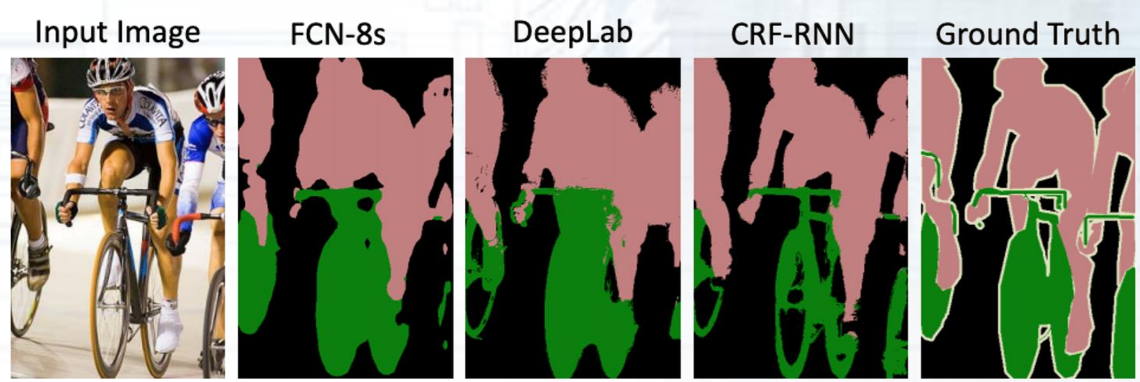
结构：from CRF

以图像分割为例：





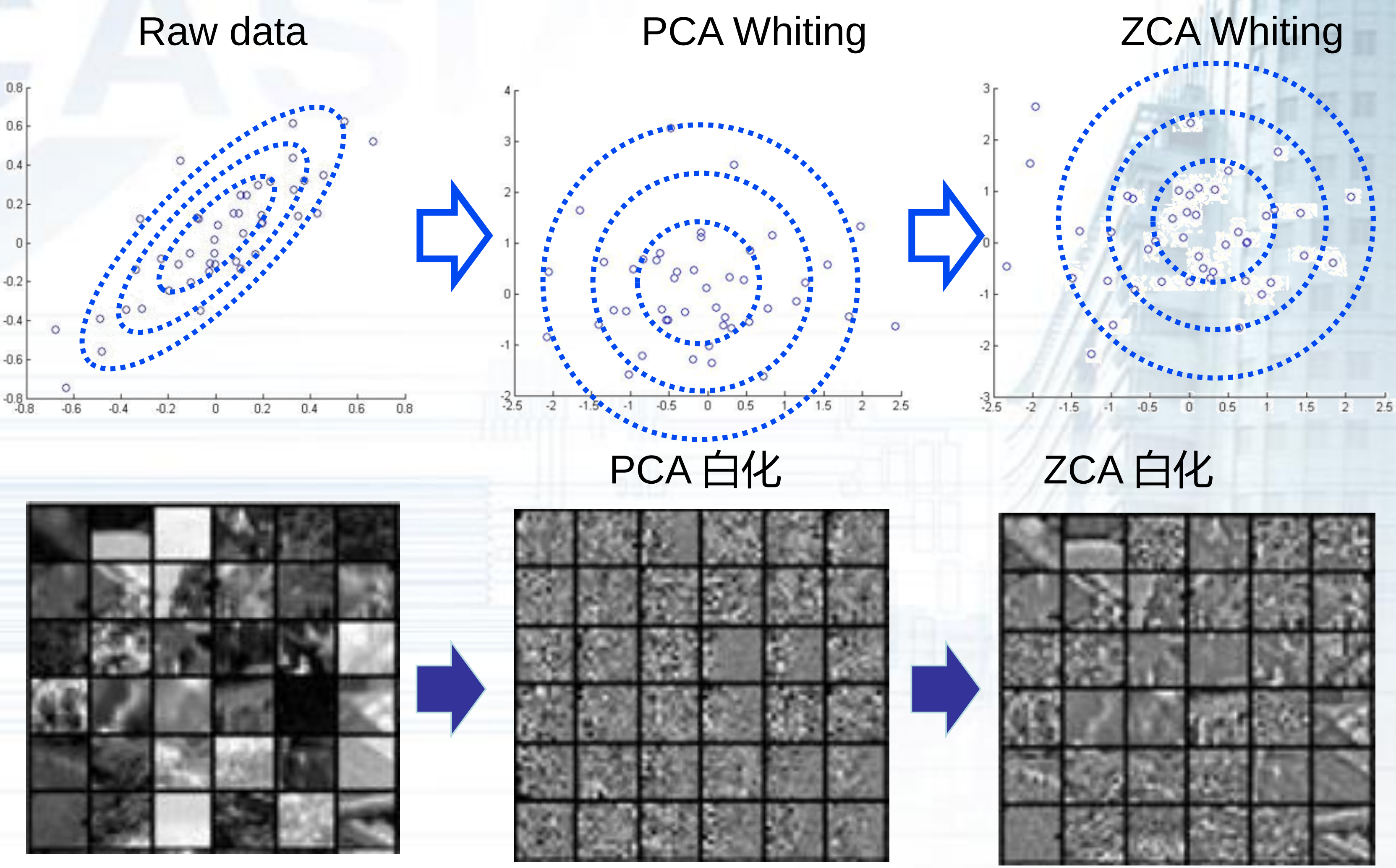




Zheng S, Jayasumana S, Romera-Paredes B, et al. Conditional random fields as recurrent neural networks[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1529-1537.

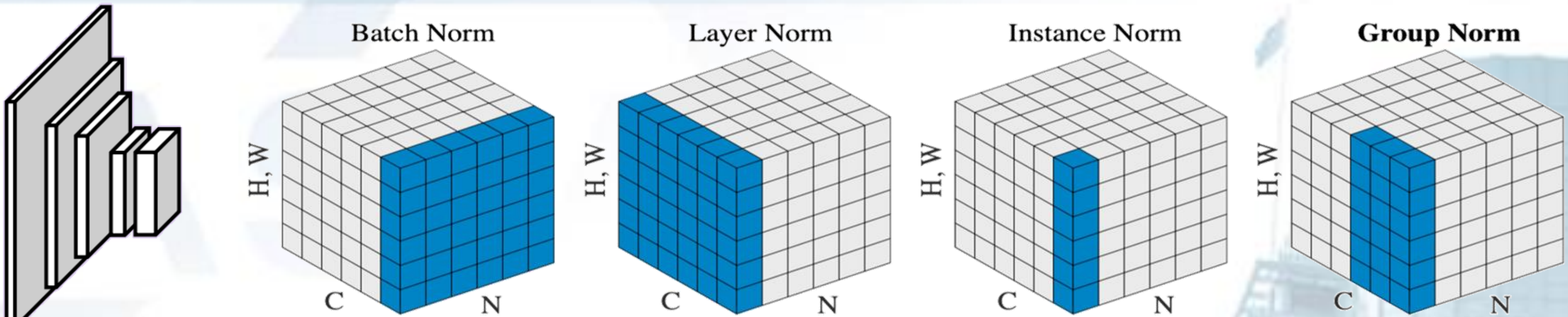
方法：whitening

去除维度之间的相关性与分布差异性，有利于加快收敛，防止过拟合。



Bell A J, Sejnowski T J. The “independent components” of natural scenes are edge filters[J]. Vision research, 1997, 37(23): 3327-3338.

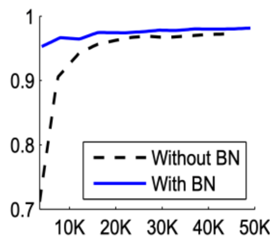
方法：whitening



Batch Normalization:

$$y_{nchw} = \frac{x_{nchw} - \mu_c}{\sqrt{\sigma_c^2 + \epsilon}}, \quad \mu_c = \frac{1}{NHW} \sum_{N,H,W} x_{nchw},$$

$$\sigma_c^2 = \frac{1}{NHW} \sum_{N,H,W} (x_{nchw} - \mu_c)^2$$



Accelerate the convergence of ConvNets

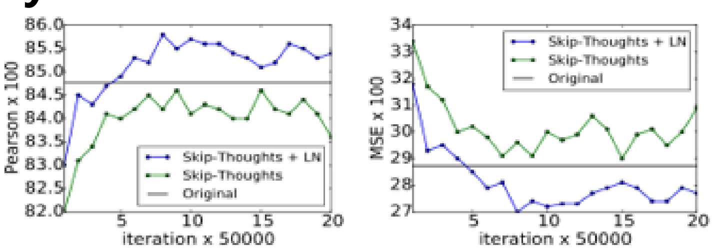
Instance Normalization:

$$y_{nchw} = \frac{x_{nchw} - \mu_{nc}}{\sqrt{\sigma_{nc}^2 + \epsilon}}, \quad \mu_{nc} = \frac{1}{HW} \sum_{H,W} x_{nchw},$$

$$\sigma_{nc}^2 = \frac{1}{HW} \sum_{H,W} (x_{nchw} - \mu_{nc})^2$$

Success in applications of image style transfer

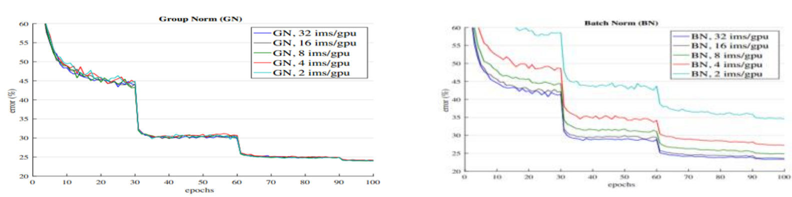
Layer Normalization:



(a) SICK(r) (b) SICK(MSE)

Improve the performance of ConvNets

Group Normalization:

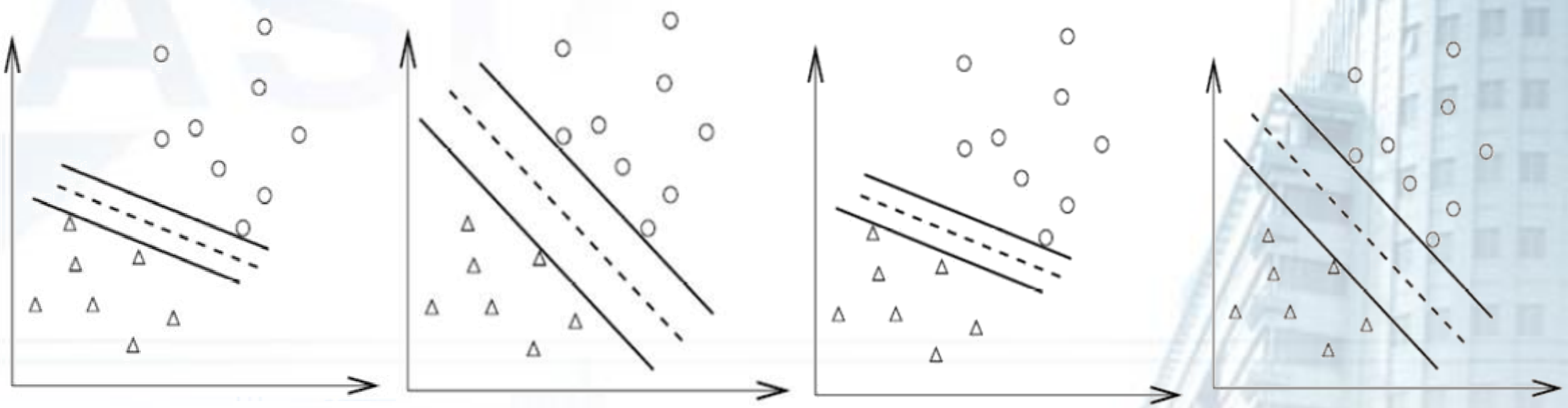


Independent to batch sizes, which can be utilized to handle large data

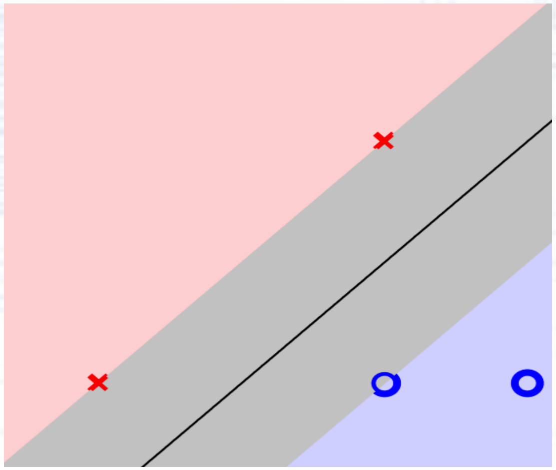
Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[J]. arXiv preprint arXiv:1502.03167, 2015.
Ba J L, Kiros J R, Hinton G E. Layer normalization[J]. arXiv preprint arXiv:1607.06450, 2016.
Huang X, Belongie S J. Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization[C]//ICCV. 2017: 1510-1519.
Wu Y, He K. Group normalization[J]. arXiv preprint arXiv:1803.08494, 2018.

方法：margin

Which is the best classifier:



SVM: find the best classifier through introducing **margin**.



$$\min_{\mathbf{w},\mathbf{b}} \frac{1}{2}\mathbf{w}^T\mathbf{w},$$

$$s.t. \quad y_i(\mathbf{w}^T x_i + b) \geq 1, i = 1 \ldots n$$

1. Robust to noise

2. Relieve overfit

Cortes C, Vapnik V. Support-vector networks[J]. Machine learning, 1995, 20(3): 273-297.

方法：margin

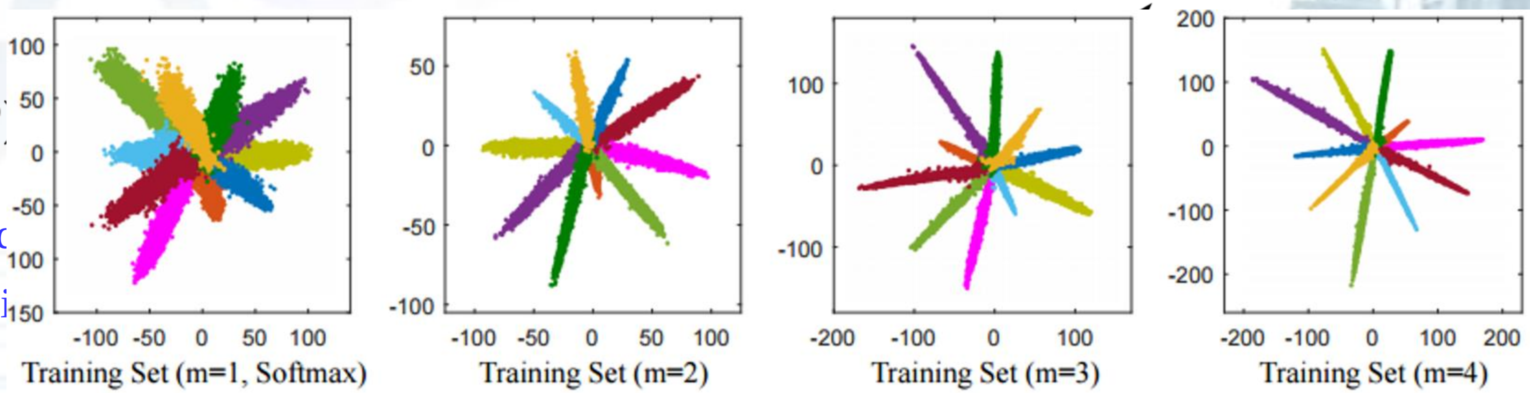
Margin被引入Deep Learning的Softmax层，从而提升特征的辨识力，在物体识别、行人重识别、人脸验证等问题上去的成功应用。

Large **margin** softmax

$\psi(\theta$

Introc

decisi



Angular **margin** softmax

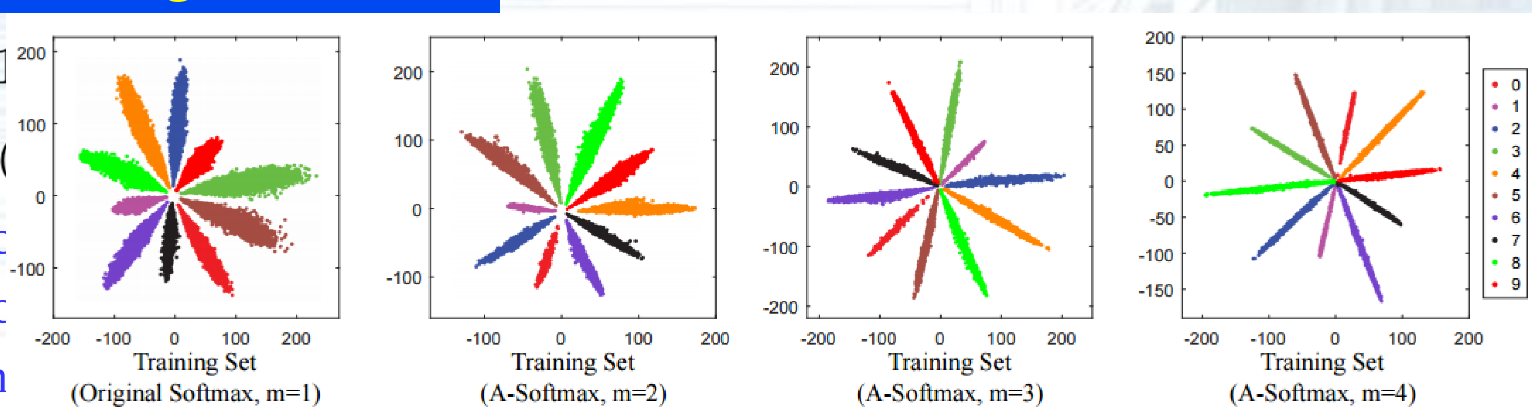
$\|W_i\|=1$

$\psi(\theta) = ($

Introduc

the dec

between

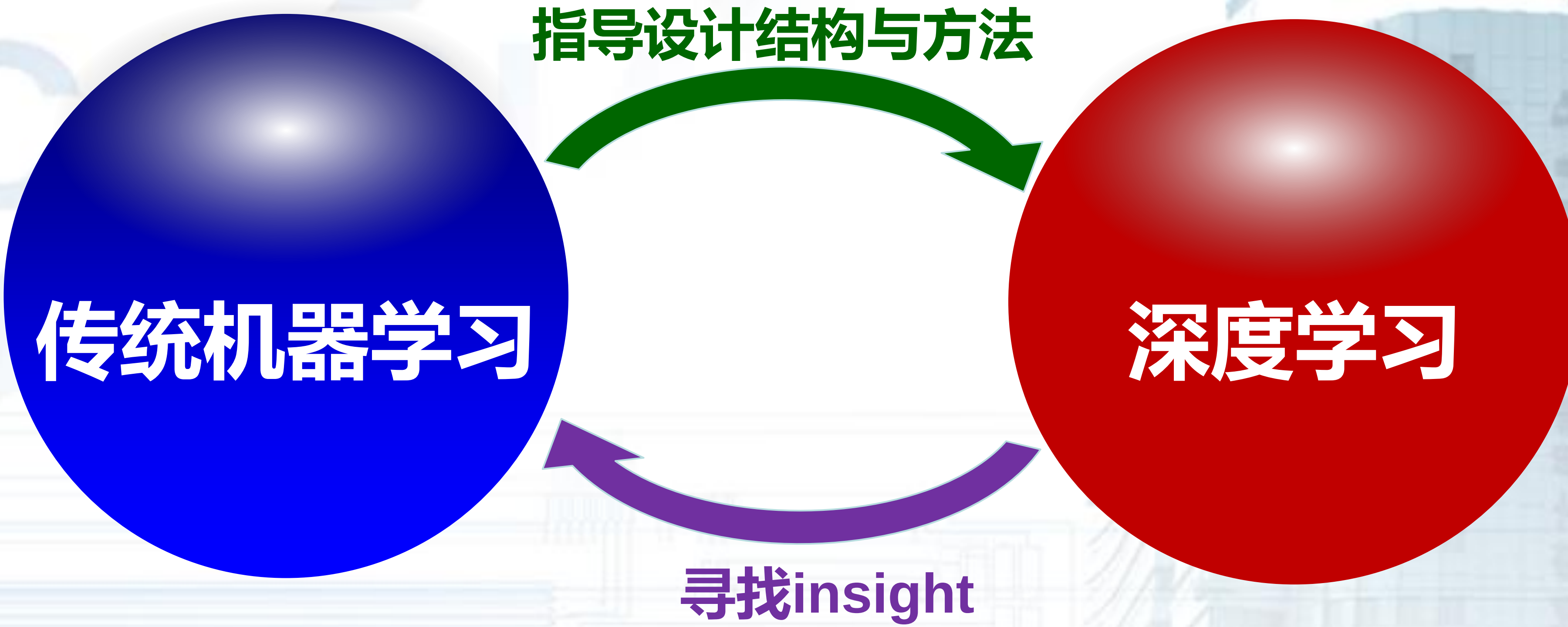


$\geq 2)$

Liu W, Wen Y, Yu Z, et al. Large-Margin Softmax Loss for Convolutional Neural Networks[C]//ICML. 2016: 507-516.

Deng J, Guo J, Zafeiriou S. Arcface: Additive angular margin loss for deep face recognition[J]. arXiv preprint arXiv:1801.07698, 2018.

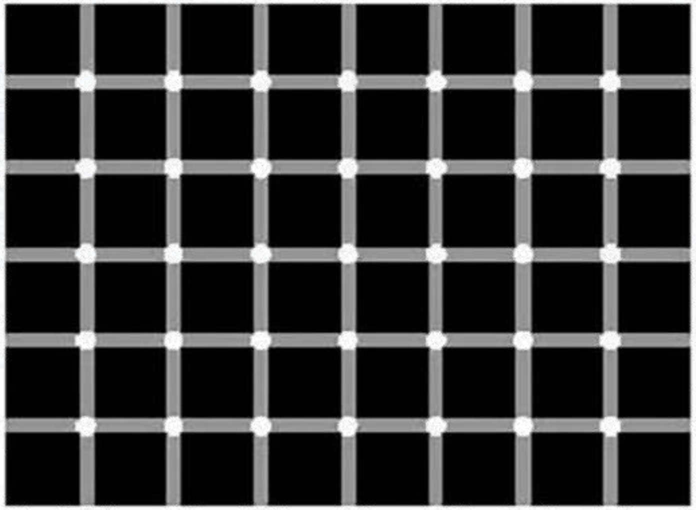
传统机器学习与深度学习



传统机器学习	深度学习
理论性强	理论性弱
解释性好	解释性差
不能充分利用大数据	能够充分利用大数据

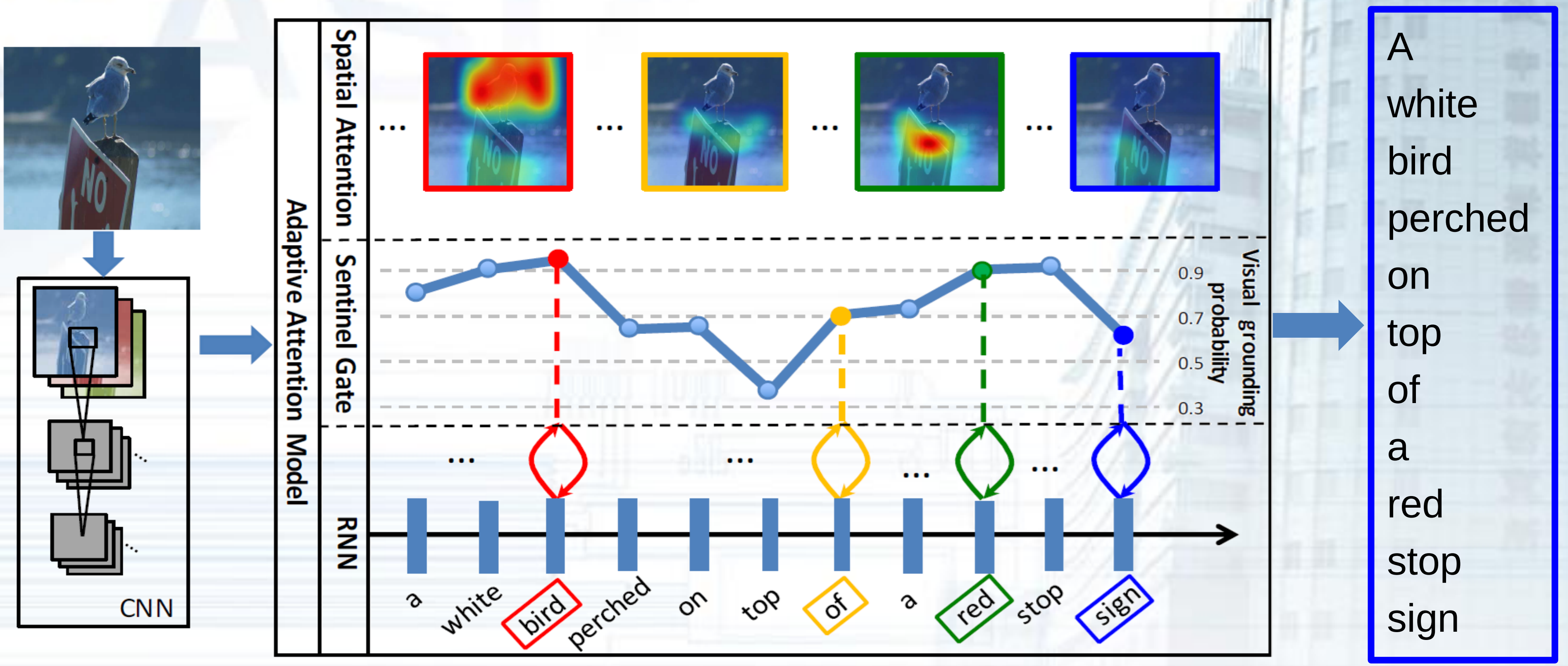
视觉注意

生物视觉注意机制：



视觉注意

深度学习中的注意模型，在多模态数据分析，如Visual Q&A、Video Caption和Image Caption等问题上有诸多成功应用。（一般结合Encoder-Decoder Model）



实现方式：**加权平均**　　　表现形式：**构建不同模态或者不同域数据之间的对齐。**

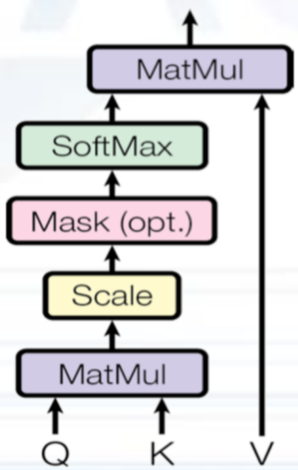
本质：**提取了用CNN或者LSTM无法刻画的长程关联信息**

Lu J, Xiong C, Parikh D, et al. Knowing when to look: Adaptive attention via a visual sentinel for image captioning[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017, 6: 2.

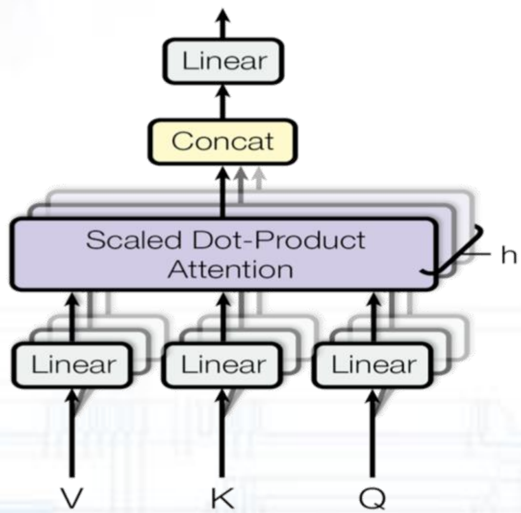
自注意力模型

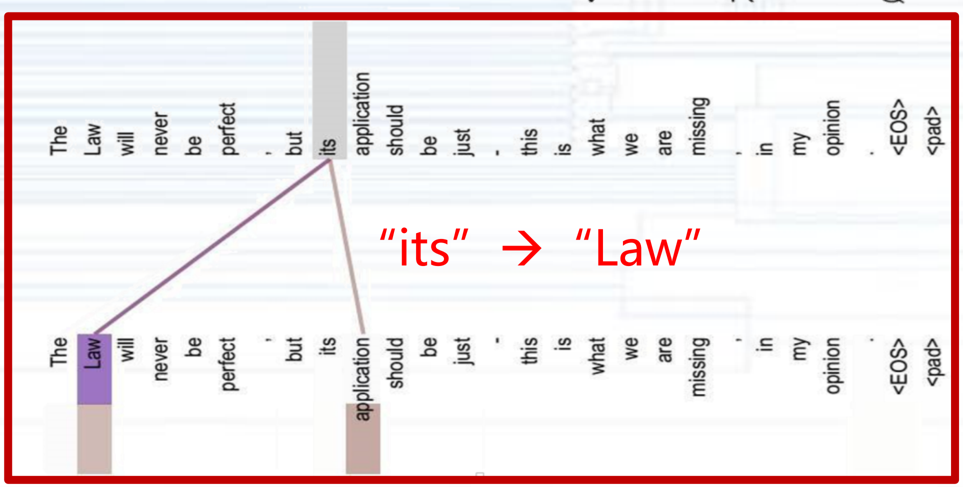
不仅仅是多模态间有长程关联问题，在单模态的时序间、像素间、通道间、样本间也同样存在长程连接，从而引伸为自注意力模型。

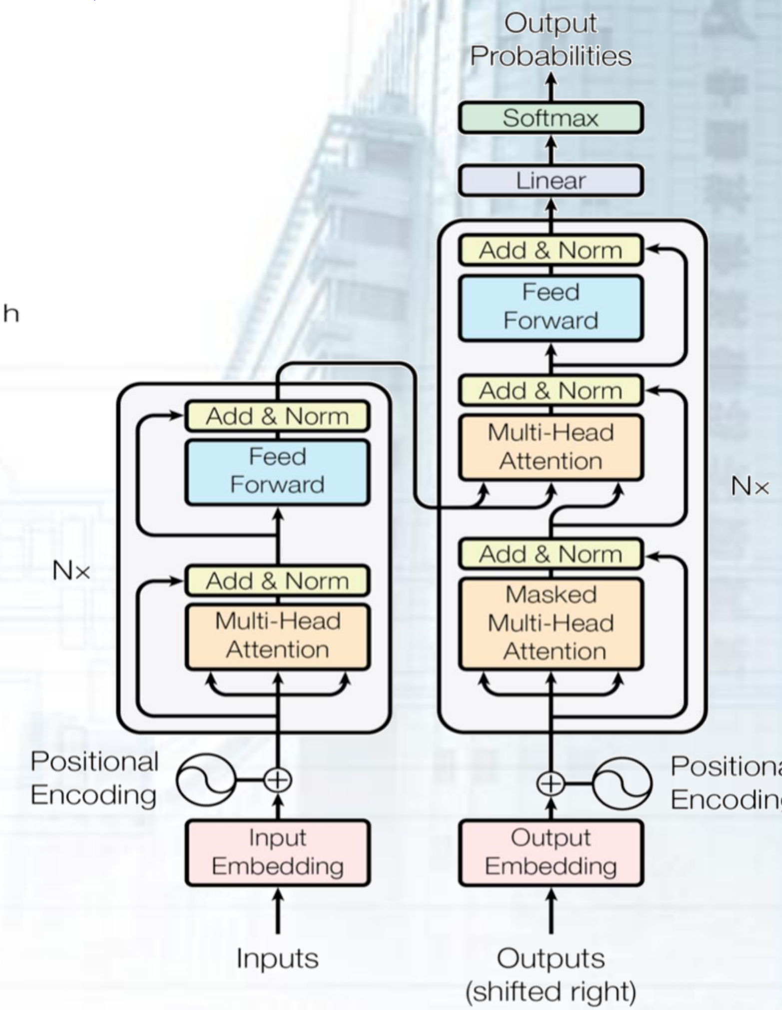
Scaled Dot-Product Attention



Multi-Head Attention



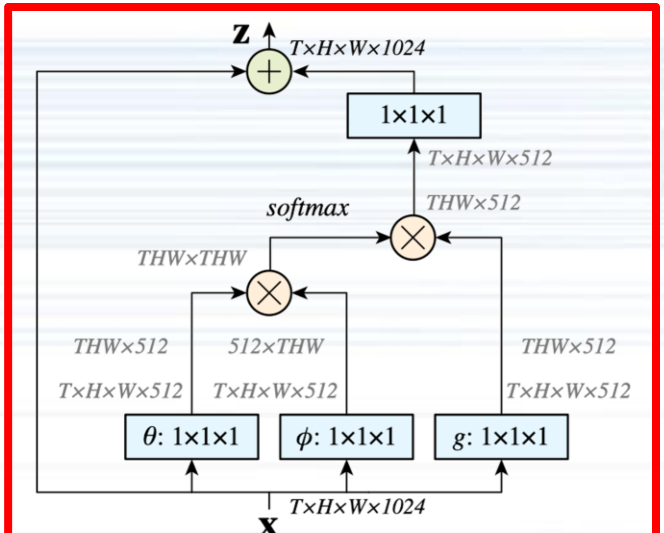




Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems. 2017: 5998-6008.

自注意力模型：视频分类

构建视频中存在时空中的像素间长程关联关系，克服CNN和LSTM等模型刻画长程关系的不足，进而提升视频分类的性能。

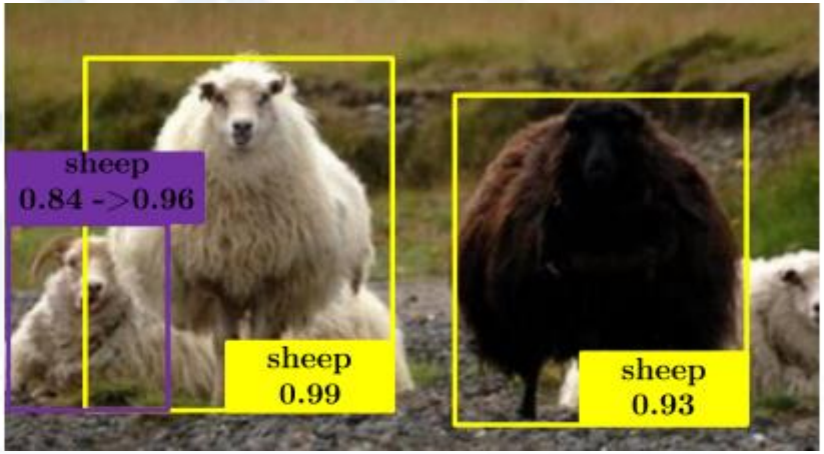


model	modality	*train/val*	*trainval/test*
2-Stream [43]	RGB + flow	18.6	-
2-Stream +LSTM [43]	RGB + flow	17.8	-
Asyn-TF [43]	RGB + flow	22.4	-
I3D [7]	RGB	32.9	34.4
I3D [ours]	RGB	35.5	37.2
NL I3D [ours]	RGB	**37.5**	**39.5**

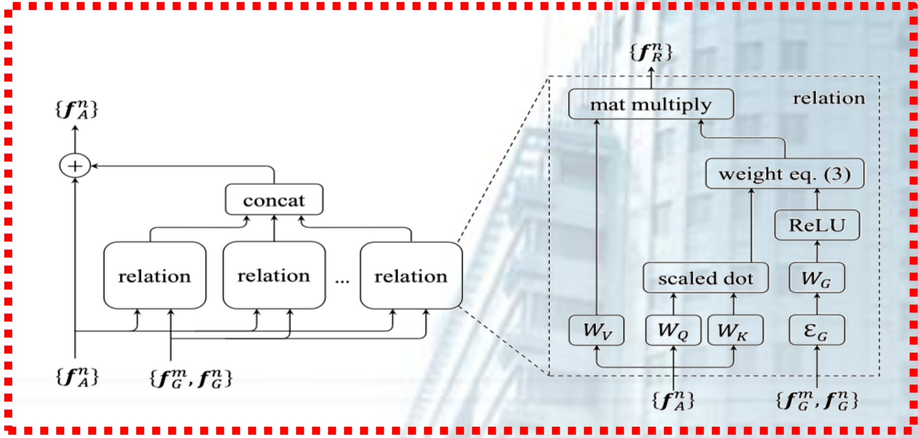
Non-local Neural Networks[J]. 2017.

自注意力模型：图像检测

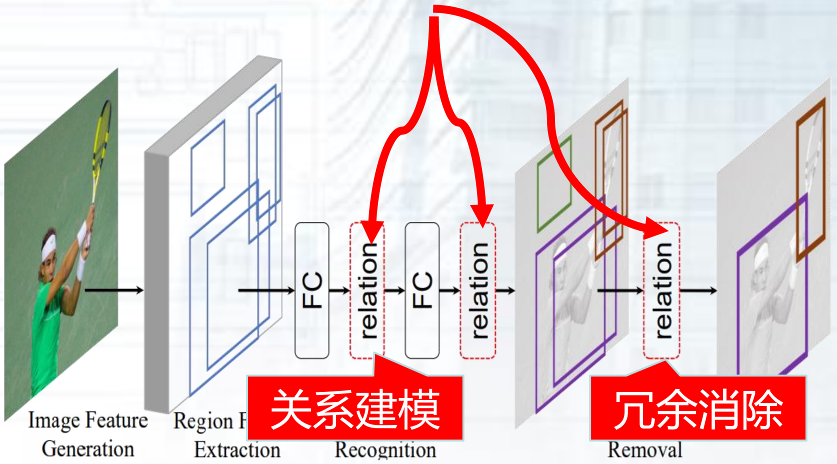
在自注意框架下，同时利用目标的语义特征和位置特征，实现检测任务中对目标关联关系建模，以及冗余目标关系建模，实现了端到端的检测网络。



关系建模有助于目标检测



Relation networks

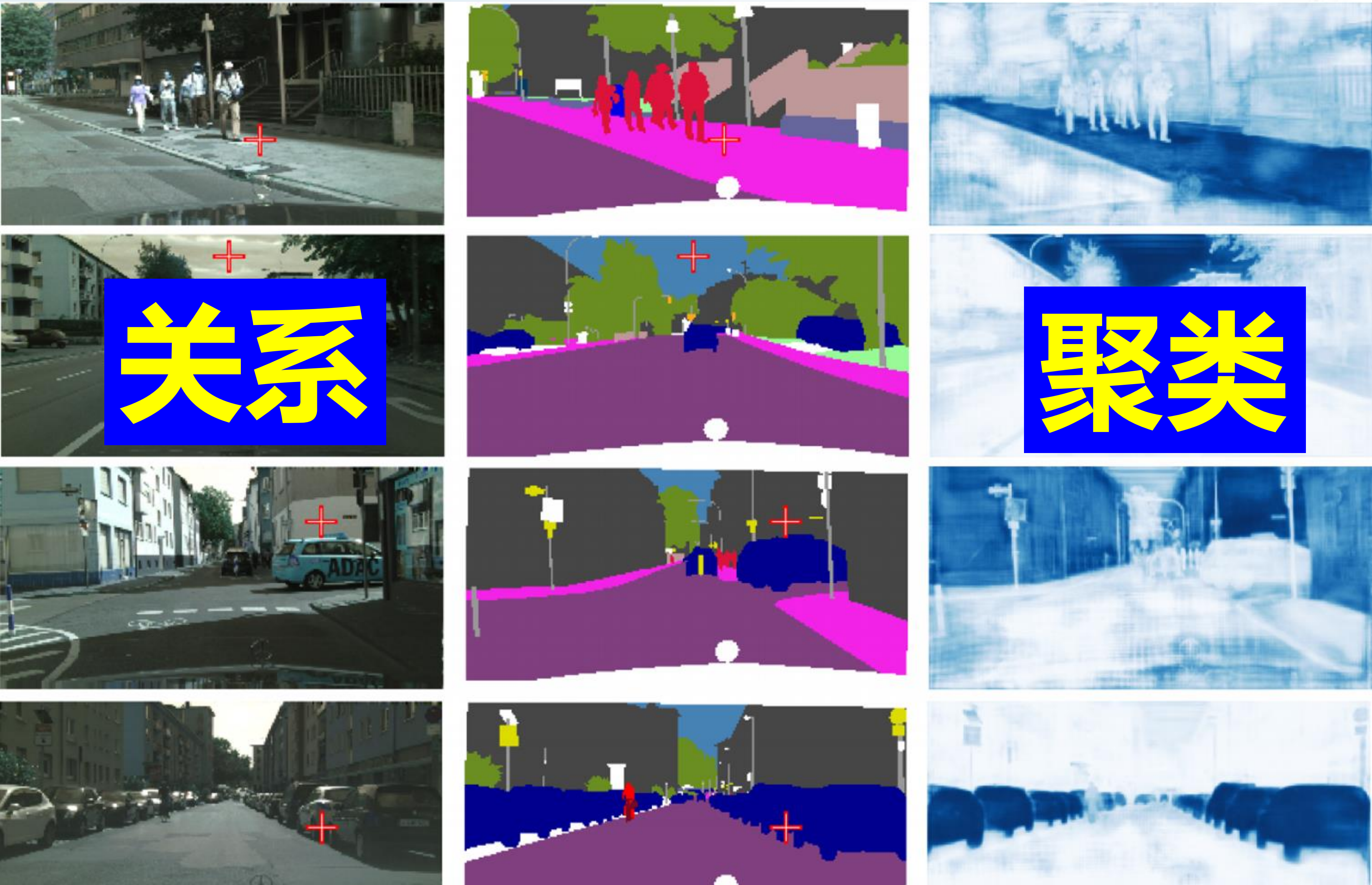


消除冗余检测结果　　建模人与手套的关系

建立端到端的目标检测网络

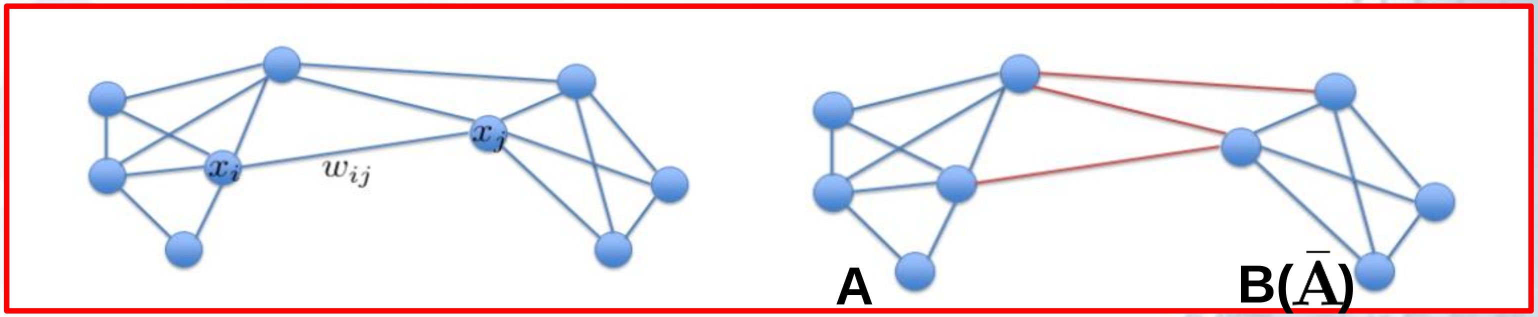
Relation Networks for Object Detection, CVPR 2018

自注意力模型：语义分割



基于关系的聚类：谱聚类

谱聚类对数据分布没有特别要求，能够在任意形状空间中进行样本聚类。



转化为图割问题（划分成 A, B 两个集合，使A, B 之间有最小关系度量。为了避免平凡解，一般会对关系度量进行归一化）

$$\mathrm{Ncut}(A,\bar{A}) = \frac{\mathrm{cut}(A,\bar{A})}{\mathrm{vol}(A)} + \frac{\mathrm{cut}(A,\bar{A})}{\mathrm{vol}(\bar{A})}$$

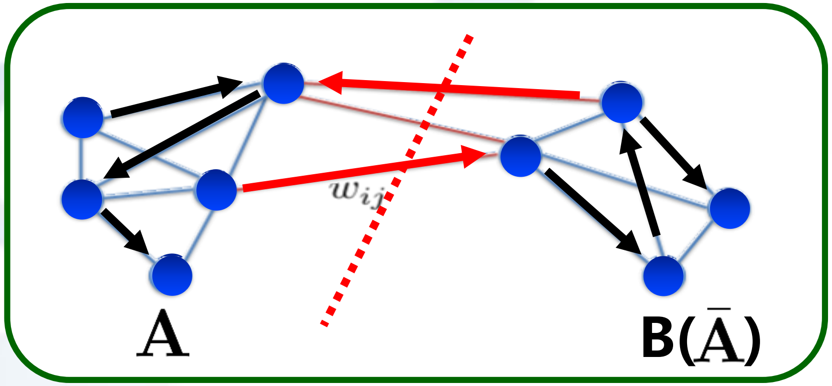
转化为随机游走问题（最小化随机游走时节点在不同clusters间转移的概率）

$$P(A|B) = P(X_t \in A | X_{t+1} \in B)$$

$$\mathrm{Ncut}(A,\bar{A}) = P(A|\bar{A}) + P(\bar{A}|A)$$

Meila and Shi (2001)

谱聚类：随机游走视角



1. 给定样本点集合:

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n]$$

2. 根据样本点之间的关系，构建相似性矩阵W。

$$\mathbf{w}_{i,j} = exp(\mathbf{x}_i^T \mathbf{x}_j / \sigma)$$

3. 利用对角矩阵D，将相似性矩阵W归一化为概率转移矩阵。

$$\mathbf{T} = \mathbf{D}^{-1}\mathbf{W} \quad \mathbf{d}_i = \sum_{j=1}^{n} w_{i,j}$$

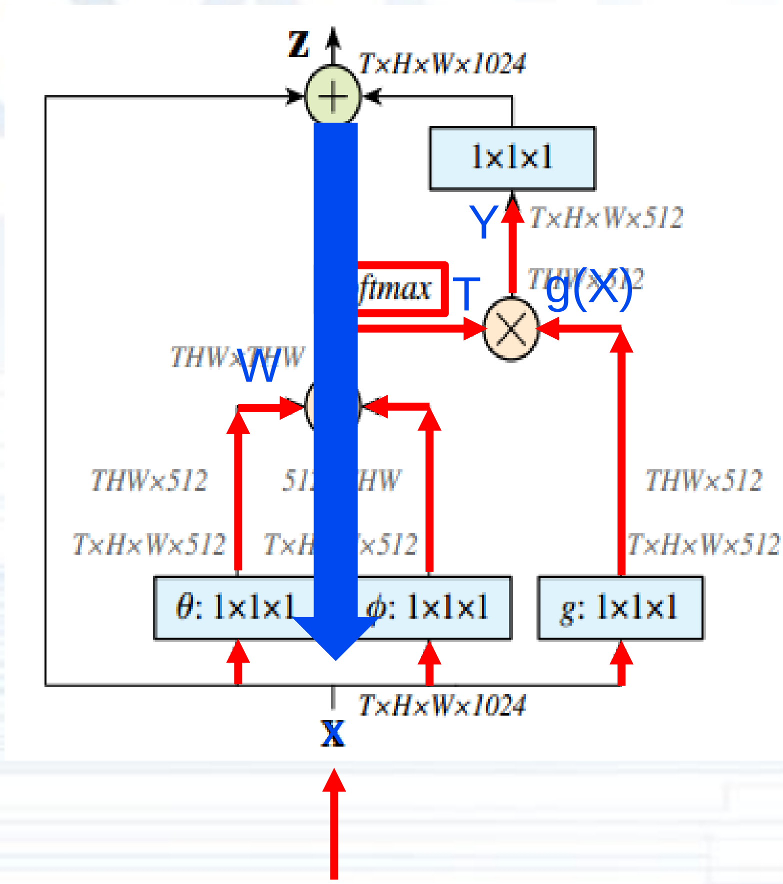
4. 利用概率转移矩阵计算一步随机游走后的图。

$$\mathbf{X}' = \mathbf{T}\mathbf{X}$$

5. 反复迭代2-4，优化目标使得不同类样本之间游走概率最小。

$$\arg\min_{\mathbf{A}} P(\mathbf{A}|\bar{\mathbf{A}}; \mathbf{T})$$

自注意力模型：训练视角



1. 给定样本集合

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$$

2. 构建关系矩阵 W：

$$\mathbf{W} = \phi(\mathbf{X})^T \varphi(\mathbf{X})$$

3. 利用softmax函数，将关系矩阵W加以归一化，形成自注意矩阵T。

$$\mathbf{T} = softmax(\mathbf{W}) \quad \mathbf{T}_{i,j} = \frac{exp(\mathbf{w}_{i,j}/\sigma)}{\sum_{j=1}^{n} exp(\mathbf{w}_{i,j}/\sigma)}$$

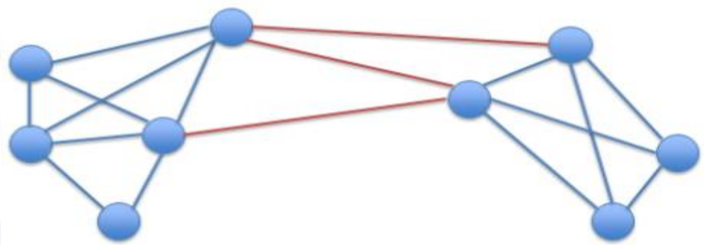
4. 利用自注意矩阵T构建新的自注意特征。

$$\mathbf{Y} = \mathbf{T}g(\mathbf{X})$$

5. 通过BP算法，反复优化。

$$\arg\min_{\mathbf{T}} L(\mathbf{A}, \bar{\mathbf{A}}; \mathbf{T})$$

自注意力模型→深度谱聚类模型



A　　　　B($\bar{\mathbf{A}}$)

样本点集合: $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$

相似性矩阵: $\mathbf{W}$,其中 $\mathbf{w}_{i,j} = exp(\boxed{\mathbf{x}_i^T \mathbf{x}_j}/\sigma)$

对角矩阵: $\mathbf{D}$ ， 其中 $\mathbf{d}_i = \sum_{j=1}^{n} w_{i,j}$

随机游走矩阵: $\mathbf{T} = \mathbf{D}^{-1}\mathbf{W}$

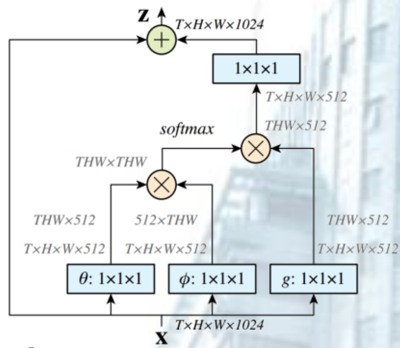
$$\mathbf{T}_{i,j} = \frac{\mathbf{w}_{i,j}}{\sum_{j=1}^{n} \mathbf{w}_{i,j}}$$

变换后特征: $\mathbf{X}' = \mathbf{T}\mathbf{X}$

优化目标:

$$\arg\min_{\mathbf{A}} P(\mathbf{A}|\bar{\mathbf{A}}; \boxed{\mathbf{T}})$$

T固定，A可学



样本点集合: $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$

关系矩阵: $\mathbf{W} = \phi(\mathbf{X})^T \varphi(\mathbf{X})$ 其中

$\phi(\bullet), \varphi(\bullet)$是通过卷积的特征降维映射

自注意矩阵: $\mathbf{T} = softmax(\mathbf{W})$

$$\mathbf{T}_{i,j} = \frac{exp(\mathbf{w}_{i,j}/\sigma)}{\sum_{j=1}^{n} exp(\mathbf{w}_{i,j}/\sigma)}$$

自注意特征: $\mathbf{Y} = \mathbf{T}g(\mathbf{X})$

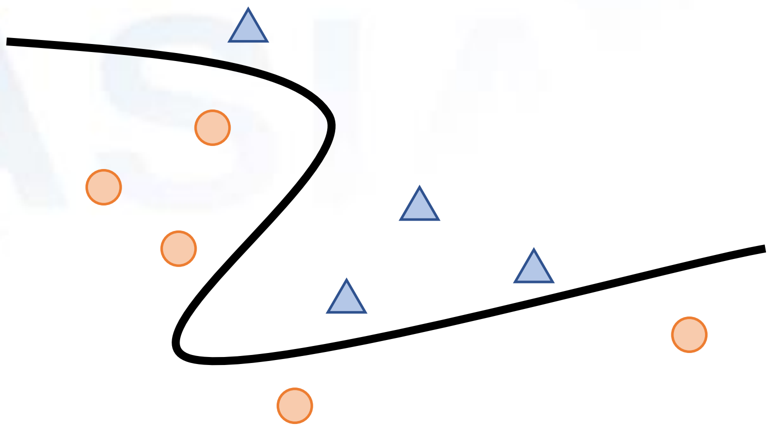
优化目标:

$$\arg\min L(\mathbf{A}, \bar{\mathbf{A}}; \mathbf{T})$$

$\boxed{\mathbf{T}}$ → T可学，A固定

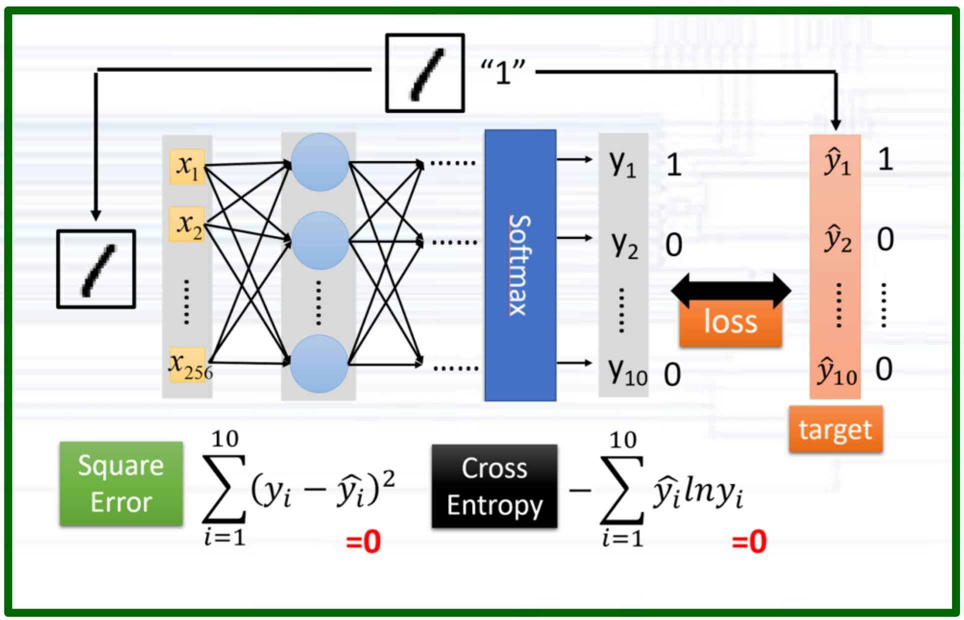
Insight：深度谱聚类模型

划分问题：



- 类间尽可能分开
- 类内尽可能紧致

深度学习的主流分类模型：Softmax+Cross Entropy



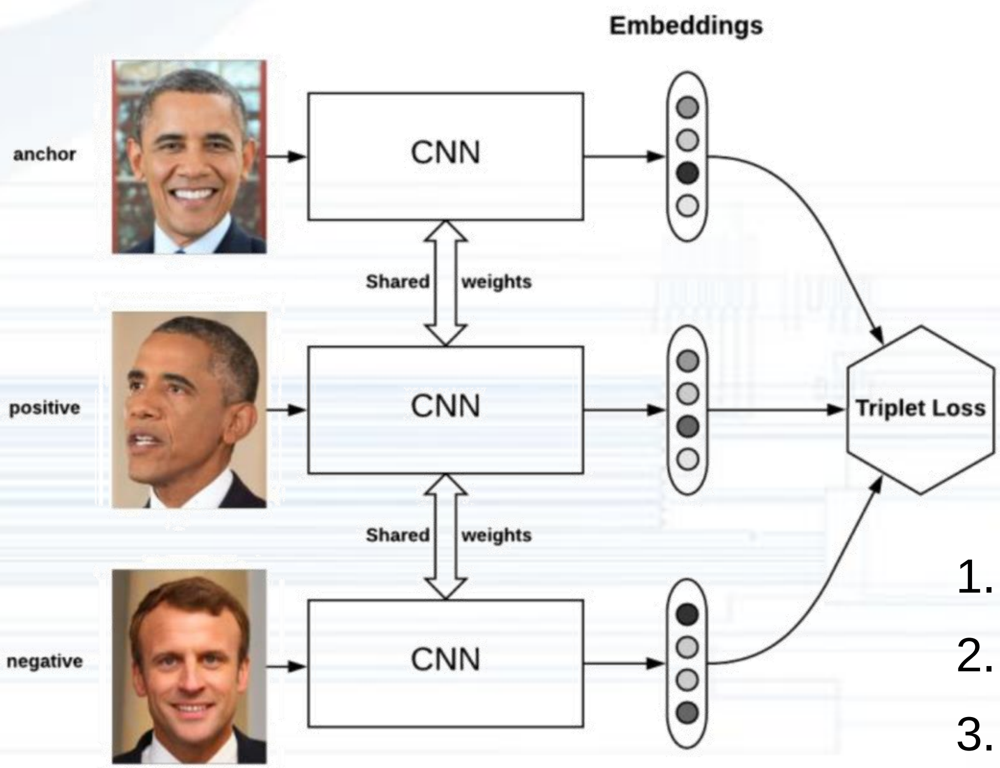
- 类间差异性很直接考虑
- 类内紧致性几乎不考虑

Image from ImageNet, augmentation from NVIDIA DALI

Insight：深度谱聚类模型

在较分类问题更为复杂的如检测、分割等问题中，仅仅通过数据增广的手段考虑类内紧致度不能胜任。

Triplet Loss：



$$\|f(x_i^a)-f(x_i^p)\|_2^2+\alpha<\|f(x_i^a)-f(x_i^n)\|_2^2$$

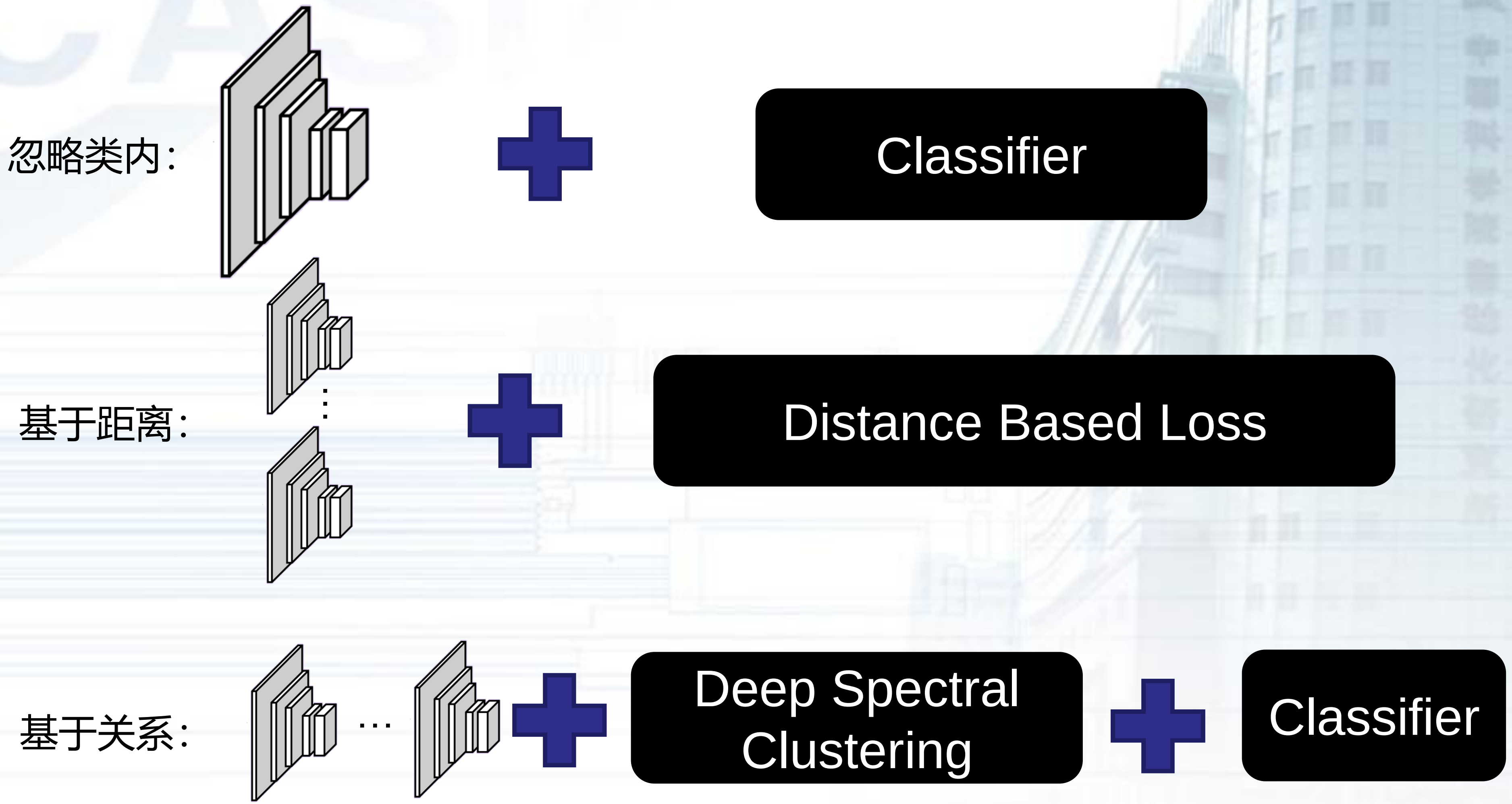
$$\forall\left(f(x_i^a), f(x_i^p), f(x_i^n)\right) \in \mathcal{T}$$

$$\sum_i^N\left[\|f(x_i^a)-f(x_i^p)\|_2^2-\|f(x_i^a)-f(x_i^n)\|_2^2+\alpha\right]_+$$

1. Triplelet数目多，需要采样，损失信息；
2. Triplelet正负样本不均衡，影响性能；
3. 距离或者度量受限定，难以定义。

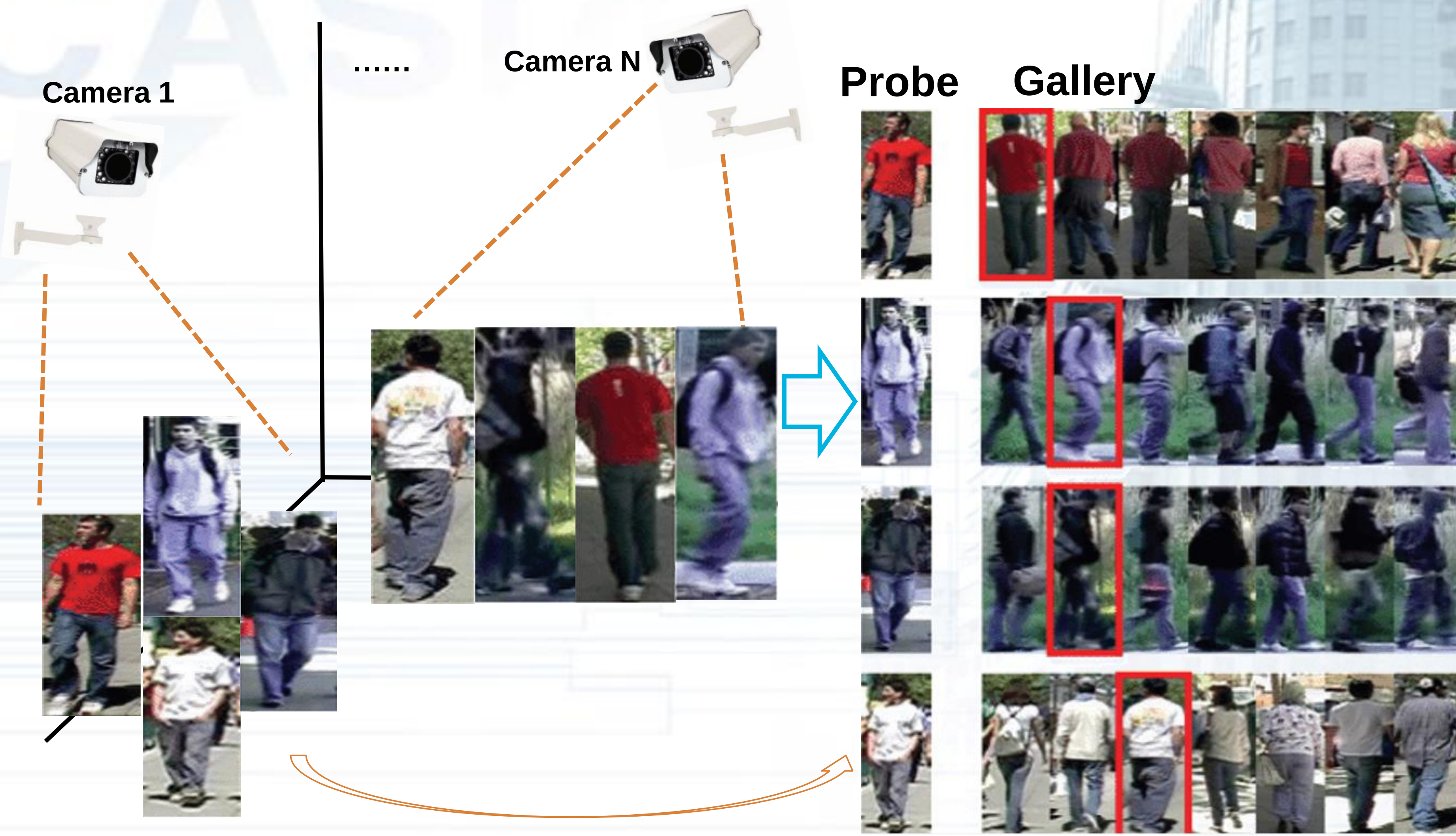
Insight：深度谱聚类模型

基于关系的深度谱聚类模型能够通过有效的变换，通过构建元素关系，学习Affinity矩阵，降低了类之间随机游走概率，增加了类内紧致度。



Our Work: Person Re-identification

Person re-identification is defined as given a query image, rank all the gallery images according to their similarity to the query image.



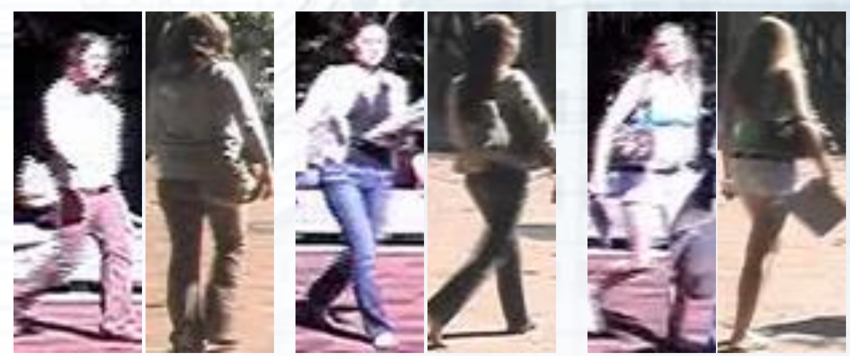
Our Work: Person Re-identification

Person Re-identification is challenging due to background cluster, occlusion, view/pose/illumination variation and so on.

Non-overlapping cameral views

View/Pose changes

Partial occlusion

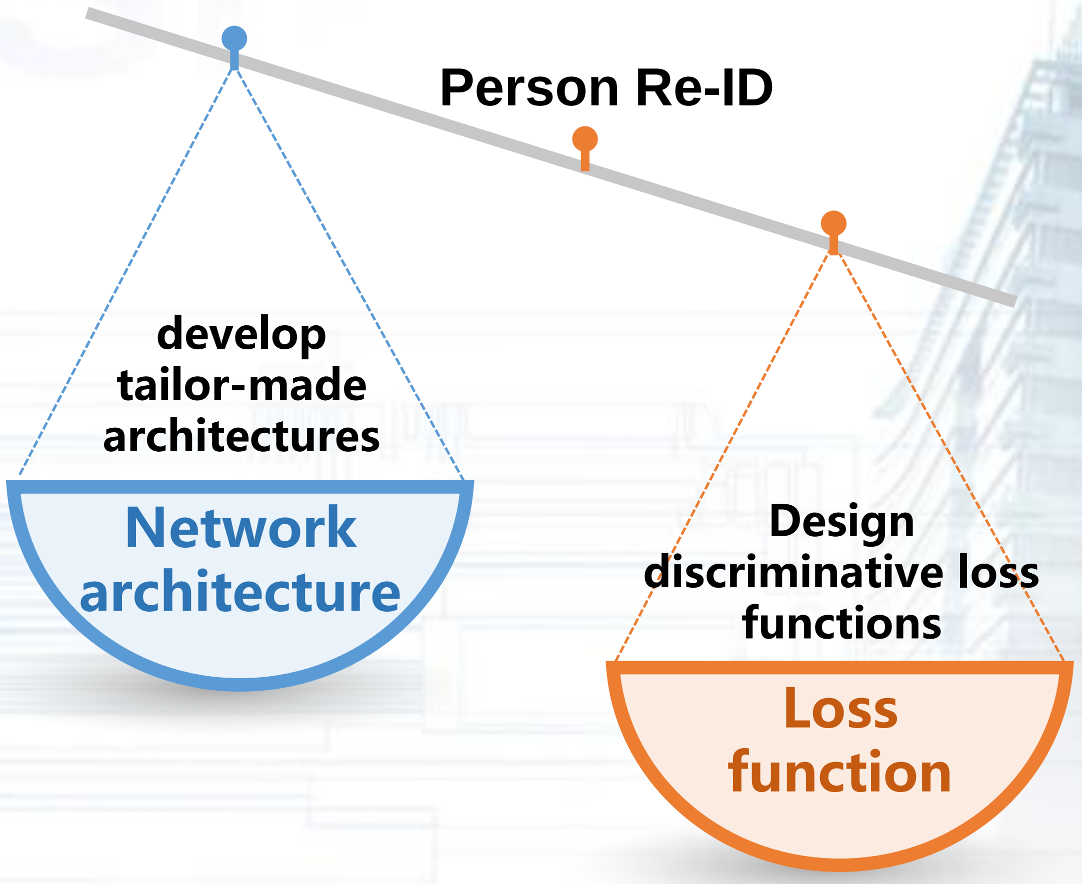


Illumination variation

…

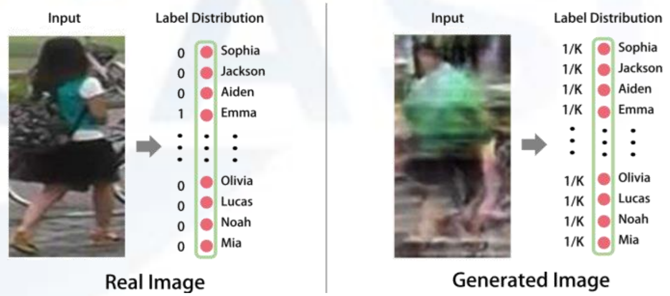
Our Work: Person Re-identification

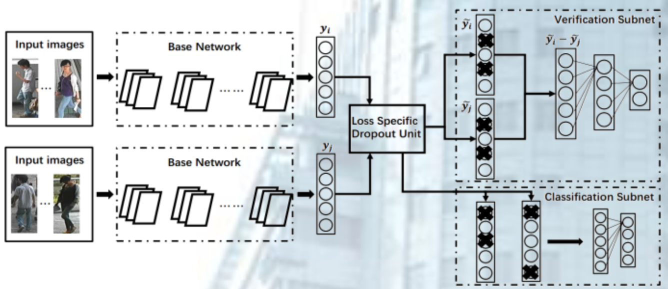
Related Works:

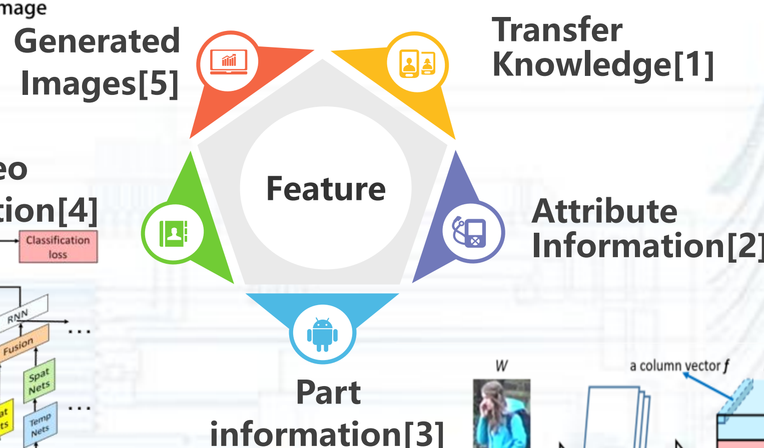


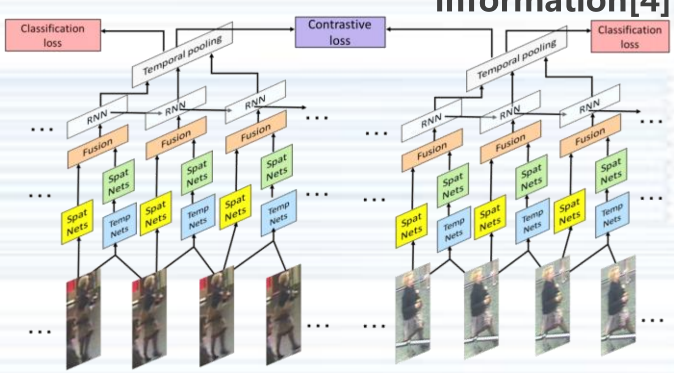
Our Work: Person Re-identification

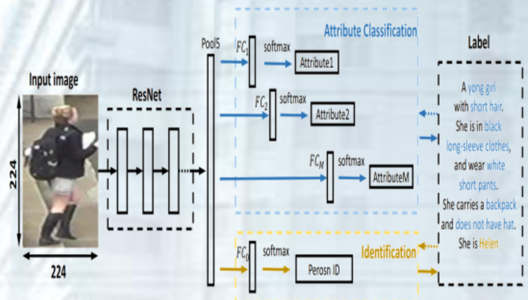
Related works of Network architecture: trying to embed more prior knowledge for learning better feature representation.

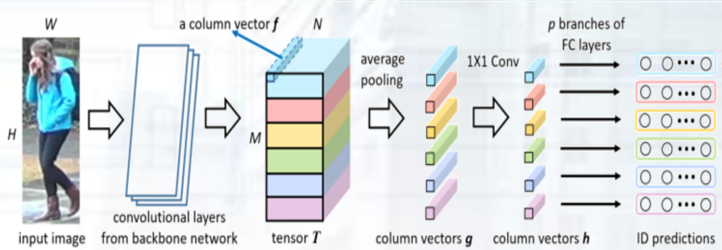








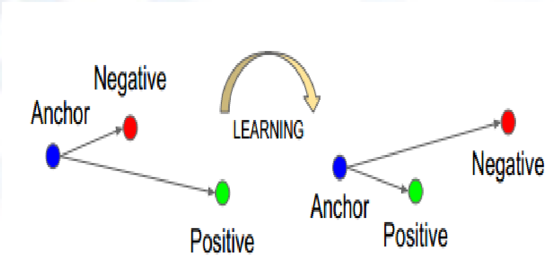


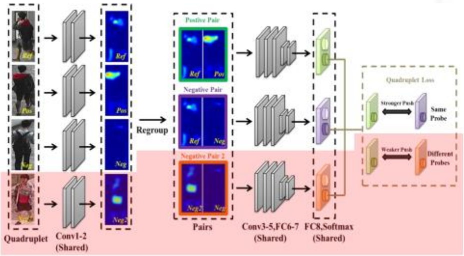


[1]. Mengyue Geng, Yaowei Wang, Tao Xiang, Yonghong Tian. Deep transfer learning for person reidentification. arXiv 2016.
[2]. Lin Y, et al. Improving person re-identification by attribute and identity learning. arXiv preprint arXiv:1703.07220, 2017.
[3]. Y. Sun,et al, Beyond Part Models: Person Retrieval with Refined Part Pooling, arXiv 2017.
[4]. Liu H, et al. Video based person re-identification with accumulative motion context. arXiv preprint arXiv:1701.00193,2017.
[5]. Zheng Z,et al. Unlabeled samples generated by gan improve the person re-identification baseline in vitro. ICCV2017.

Our Work: Person Re-identification

Related works of Loss Function: trying to define metrics to make the intra-class distance be less than the inter-class distance.





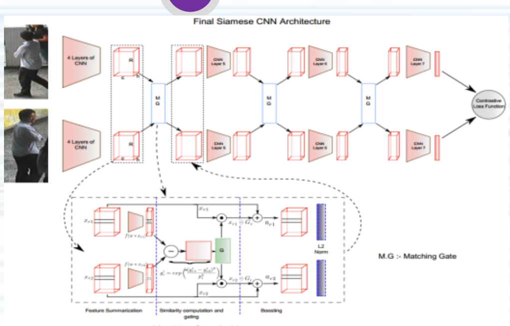
ECCV2016 — Contrastive loss

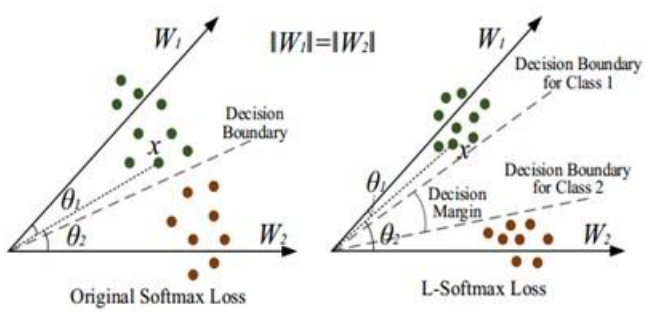
CVPR2015 — Triplet Loss

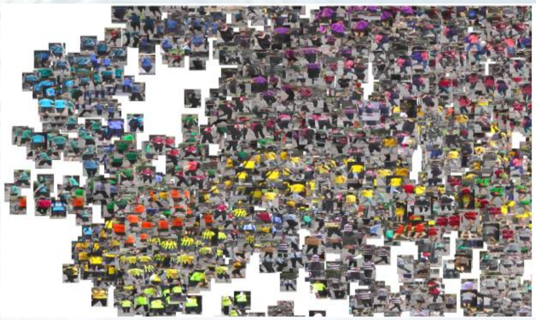
ICML2016 — Large margin softmax loss

CVPR2017 — Quadruplet Loss

AAAI2018 — DarkRank Loss







[1]. R. Varior,et al. Gated siamese convolutional neural network architecture for human re-identification. ECCV2016

[2]. F. Schroff,et al, FaceNet: A Unified Embedding for Face Recognition and Clustering, CVPR2015

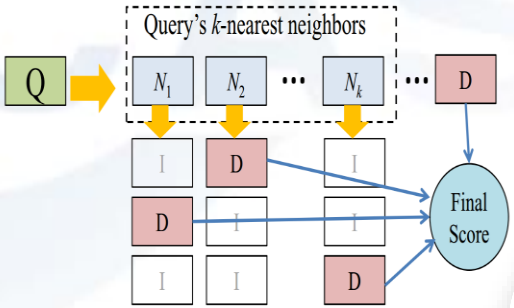
[3]. W. Liu, et al, Large-Margin Softmax Loss for Convolutional Neural Networks, ICML2016.

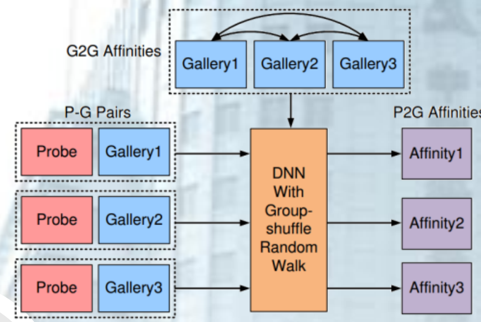
[4]. W. Chen,et al. Beyond triplet loss: a deep quadruplet network for person re-identification. CVPR2017.

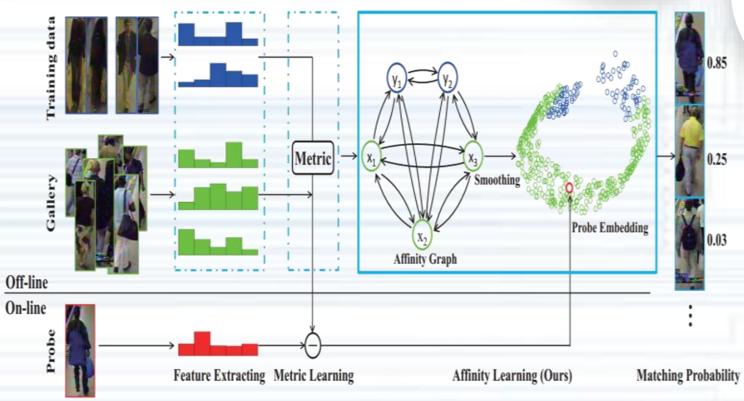
[5]. Z. Zhang, et al. DarkRank: Accelerating Deep Metric Learning via Cross Sample Similarities Transfer. AAAI2018.

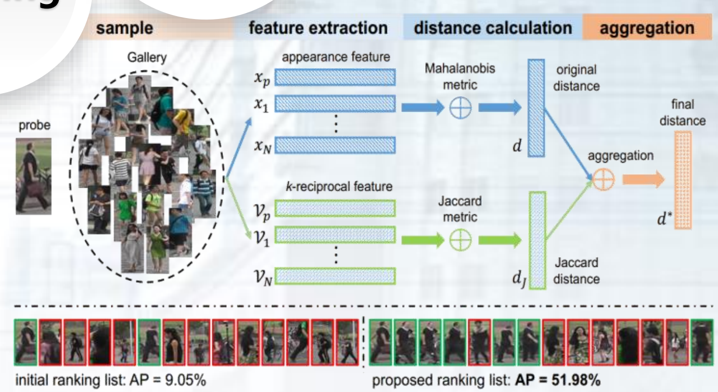
Our Work: Person Re-identification

Related Works of Post-processing: re-ranking based on relations.





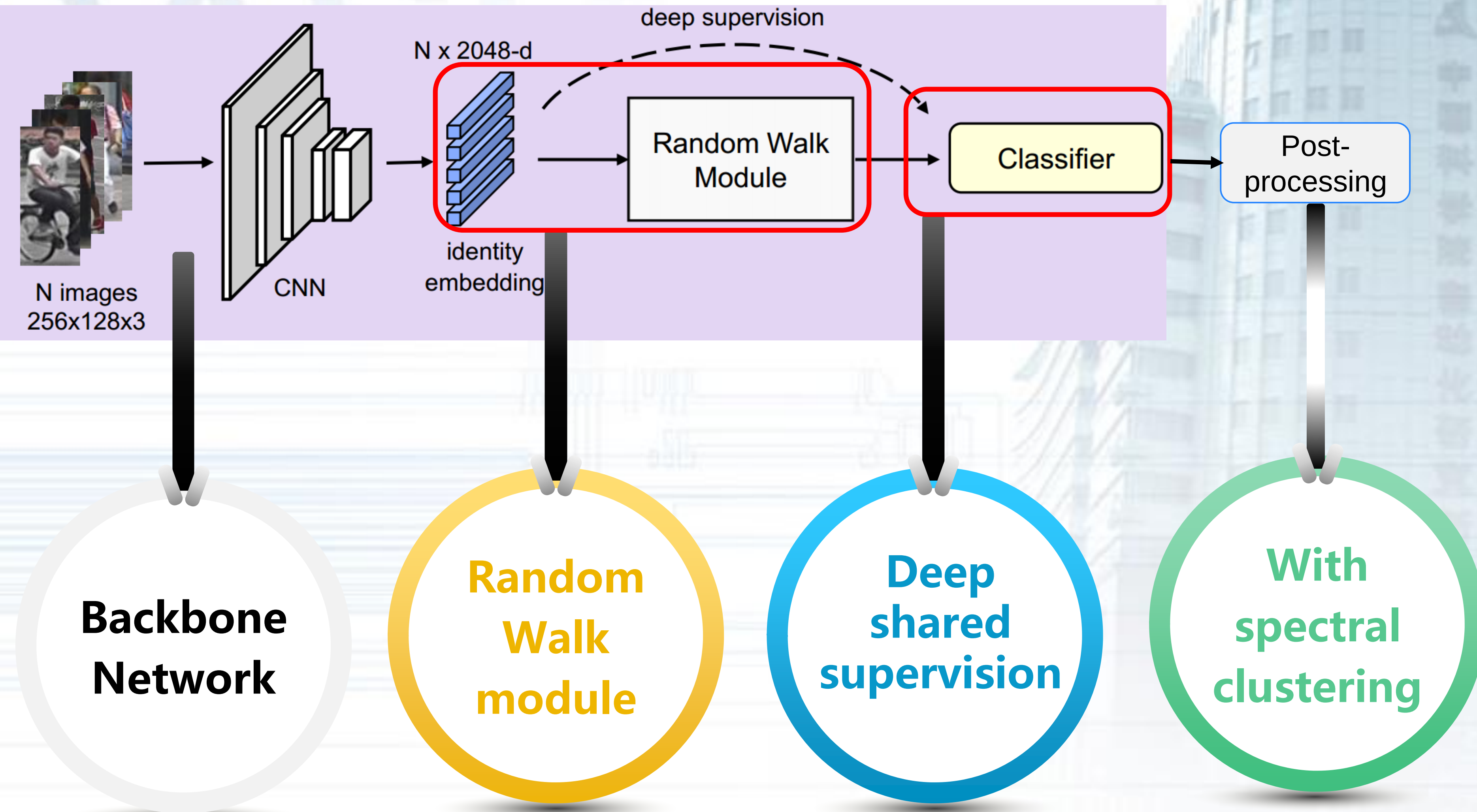




1. X. Shen, et al. Object retrieval and localization with spatially-constrained similarity measure and K-NN re-ranking, CVPR, 2012.
2. S. Bai, et al. Scalable person re-identification on supervised smoothed manifold, CVPR 2017.
3. Zhong Z, et al. Re-ranking person re-identification with k-reciprocal encoding, CVPR2017.
4. Y. Shen, et al. Deep Group-shuffling Random Walk for Person Re-identification, CVPR 2018

Our Work: Person Re-identification

Our framework:



https://arxiv.org/abs/1811.11405

Our Work: Person Re-identification

We conduct our experiments in the largest public datasets.

CUHK03
(CVPR2014)

Bboxes:28192 Identities:1467
Cameras:2 Detector:hand, DPM
Scene:indoor

DukeMTMC
(ICCV2017)

Bboxes:32,668 Identities:1501
Cameras:6 Detector:DPM
Scene:outdoor

Market-1501
(ICCV2015)

Bboxes:36,411 Identities:1812
Cameras:8 Detector:hand
Scene:outdoor

MSMT17
(CVPR2018)

Bboxes:126,441 Identities:4101
Cameras:15 Detector: Faster RCNN
Scene: indoor, outdoor

[1]. W. Li, R. Zhao, T. Xiao, and X. Wang. Deepreid: Deep filter pairing neural network for person re-identification. CVPR 2014
[2]. Z. Zheng, et al. Unlabeled samples generated by gan improve the person re-identification baseline in vitro. ICCV, 2017
[3]. L. Zheng,et al. Scalable person re-identification: A benchmark. ICCV, 2015.
[4]. L. Wei, et al, Person Transfer GAN to Bridge Domain Gap for Person Re-Identification, CVPR2018

Our Work: Person Re-identification

Experiments achieved the state-of-the-art.

Methods	Reference	MSMT17		
		mAP	R-1	R-5
GoogleNet [38]	CVPR15	23.0	47.6	65.0
PDC [34]	ICCV17	29.7	58.0	73.6
GLAD [45]	ACMMM17	34.0	61.4	76.8
Proposed		**47.3**	**73.6**	**86.0**

13.3% 12.2% 9.8%

https://arxiv.org/abs/1811.11405

Our Work: Person Re-identification

Experiments achieved the state-of-the-art.

Methods	Reference	DukeMTMC		
		mAP	R-1	R-5
PSE [25]	CVPR18	62.0	79.8	89.7
HA-CNN [17]	CVPR18	63.8	80.5	-
MLFN [3]	CVPR18	62.8	81.0	-
DuATM [32]	CVPR18	64.6	81.8	90.2
PCB+RPP [37]	ECCV18	69.2	83.3	-
Part-aligned [35]	ECCV18	69.3	84.4	92.2
Mancs [40]	ECCV18	71.8	84.9	-
Proposed		**73.2**	**86.9**	**93.9**

1.4% 2.0% 1.7%

https://arxiv.org/abs/1811.11405

Our Work：Person Re-identification

Experiments achieved the state-of-the-art.

Methods	Reference	Market-1501		
		mAP	R-1	R-5
PSE [25]	CVPR18	69.0	87.7	93.1
DPFL [6]	ICCV17	73.1	88.9	-
GLAD [45]	ACMMM17	73.9	89.9	-
MLFN [3]	CVPR18	74.3	90.0	-
HA-CNN [17]	CVPR18	75.7	91.2	-
DuATM [32]	CVPR18	76.6	91.4	97.1
Part-aligned [35]	ECCV18	79.6	91.7	96.9
PCB [37]	ECCV18	77.4	92.3	97.2
Mancs [40]	ECCV18	82.3	93.1	-
Proposed		**82.4**	**93.2**	**97.4**

https://arxiv.org/abs/1811.11405

Our Work：Person Re-identification

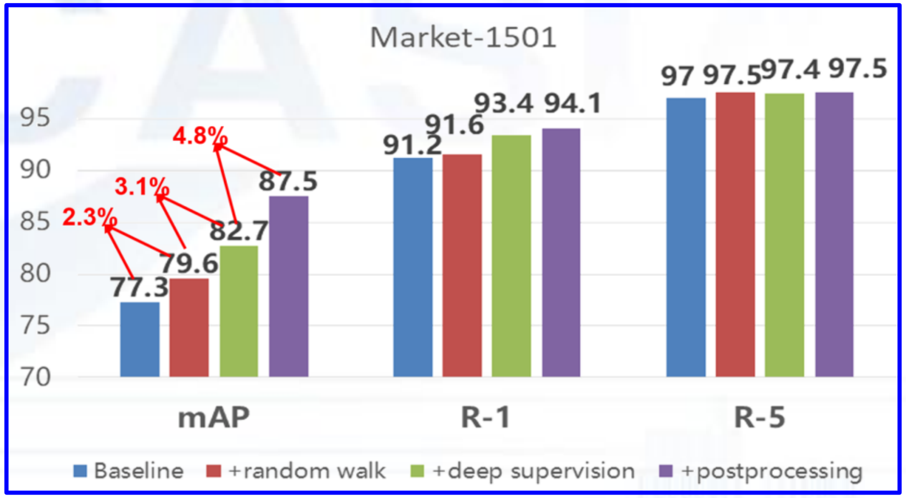
Experiments achieved the state-of-the-art.

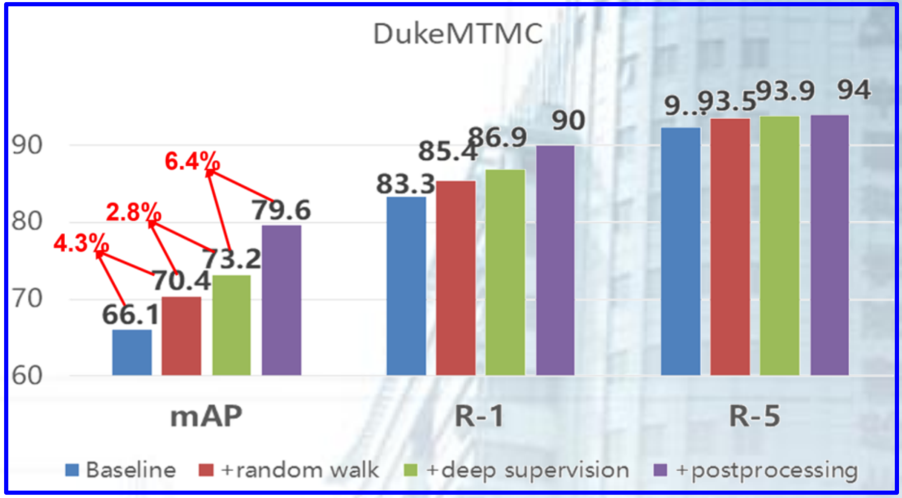
Methods	Reference	CUHK03		
		mAP	R-1	R-5
SVDNet [36]	ICCV17	37.8	40.9	-
DPFL [6]	CVPR18	40.5	43.0	-
HA-CNN [17]	CVPR18	41.0	44.4	-
MLFN [3]	CVPR18	49.2	54.7	-
DaRe [42]	CVPR18	61.6	66.1	-
Proposed		**62.4**	**68.2**	**84.4**

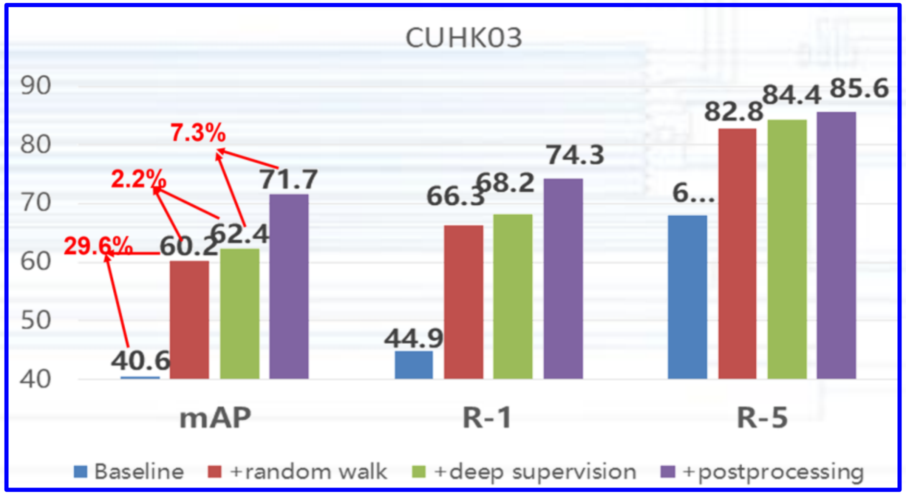
https://arxiv.org/abs/1811.11405

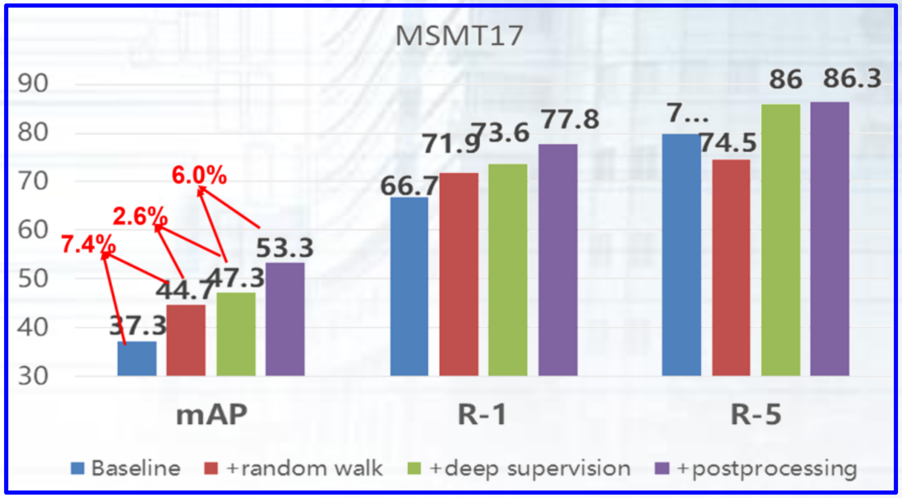
Our Work: Person Re-identification

Experiments achieved the state-of-the-art.









https://arxiv.org/abs/1811.11405

Our Work: Person Re-identification

Illustration of the comparison:

(a) baseline

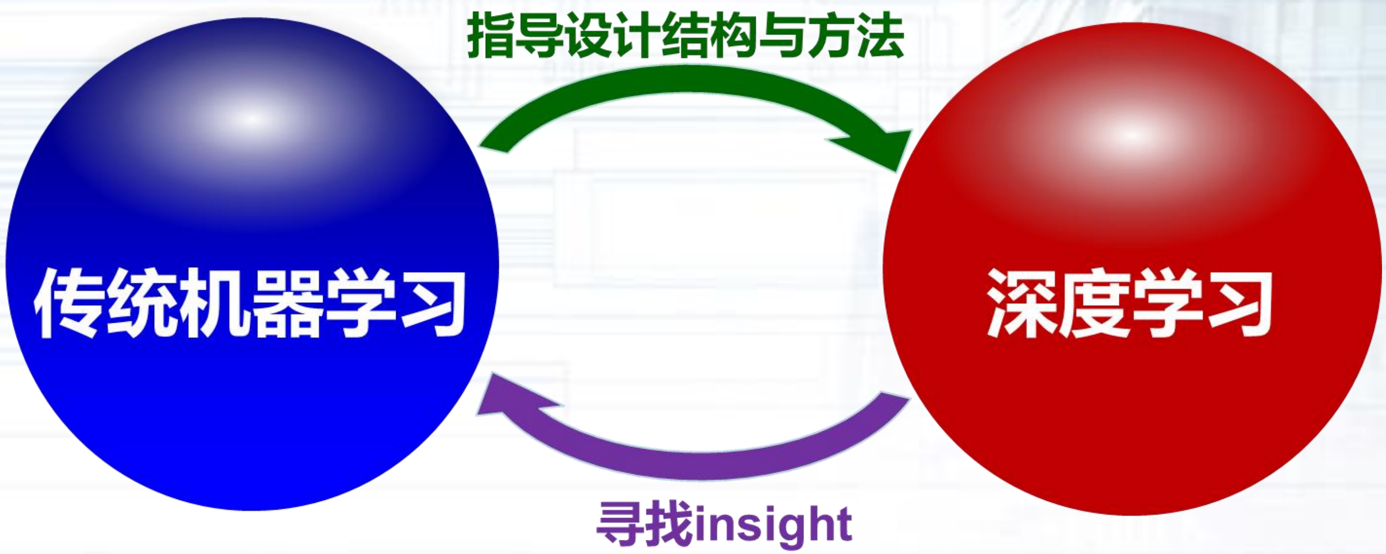
(b) + random walk
+ deep supervision

(c) +Post-processing

https://arxiv.org/abs/1811.11405

Take-home message

- discussed the relations between deep learning and the traditional machine learning.
- showed that the so-called self-attention is deep spectral clustering and analyzed its new insight.
- applied deep spectral clustering to various tasks based on the new insight with the state-of-the-art performances achieved.



谢谢，

请批评指正！