

人工智能时代 后深度学习的挑战与思考

焦李成

中国人工智能学会副理事长、西安电子科技大学教授

智能感知与计算国际联合研究中心主任

智能感知与计算国际合作联合实验室主任

IEEE GRSS 西安分会主席, IEEE TGRS 副主编



51CTO WOT

WOTD

World Of Tech
2017年12月1-2日

全球软件开发技术峰会

[深圳站]

报名咨询：010-68478816

议题提交：wot@51cto.com

市场合作：yangxh@51cto.com

商务合作：songjc@51cto.com

媒体合作：yankk@51cto.com

在线咨询（微信）：18401576051

团 · 购 · 享 · 受 · 更 · 多 · 优 · 惠

5折 优惠（截止8月31日）
现在报名，立省1400元 / 张



一、引言

01

深度学习时代的机遇

02

神经网络七十年回顾与展望

03

深度学习当前研究热点

04

深度学习模型的演化及思考

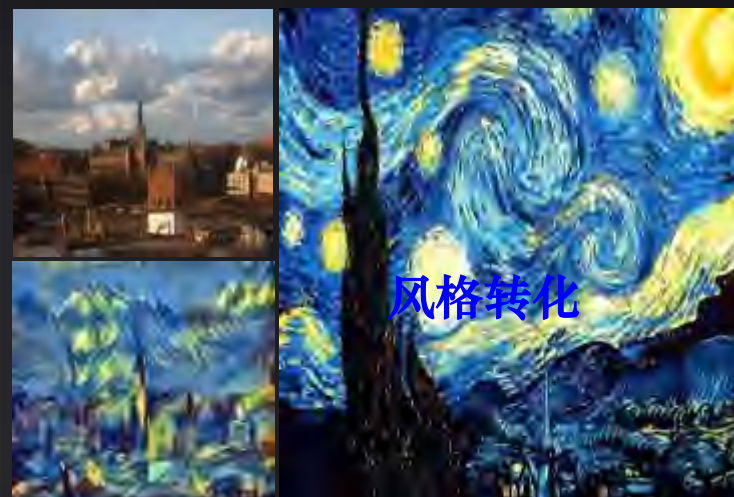
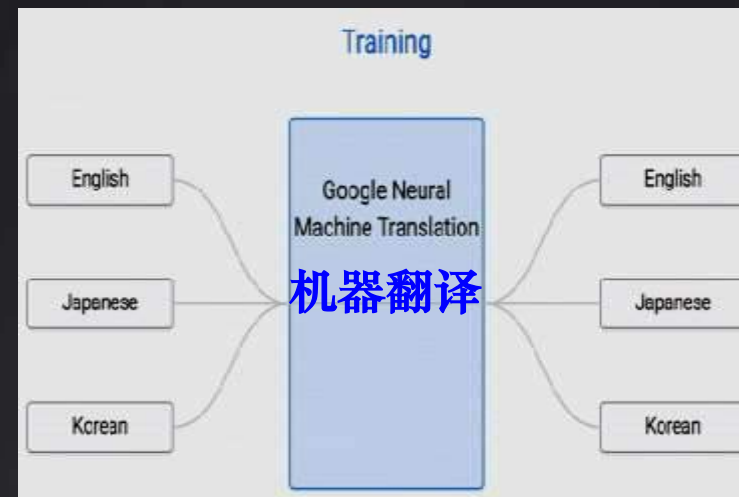
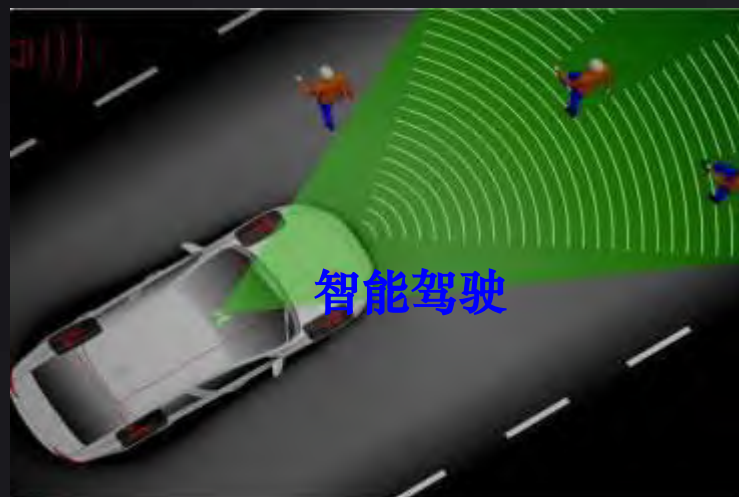


深度学习智源大数据

1.1 深度学习时代的机遇

面向泛在（如移动计算）、高风险（如精准医疗）、高可靠性（如智能交通）等应用场景，突破深度学习理论基础薄弱、模型结构单一、资源消耗过高、数据依赖性强的瓶颈。研究下一代深度学习理论基础；非神经网络、资源节约型深度学习模型、方法及高效优化技术；适于小样本/无监督学习范式、强化/迁移/博弈学习的深度学习方法与技术。

应用驱动包括 自然图像处理，自然语音处理，遥感影像理解与解译



1.2 神经网络七十年回顾与展望



Donald Hebb
Hebb 学习规则 1949年

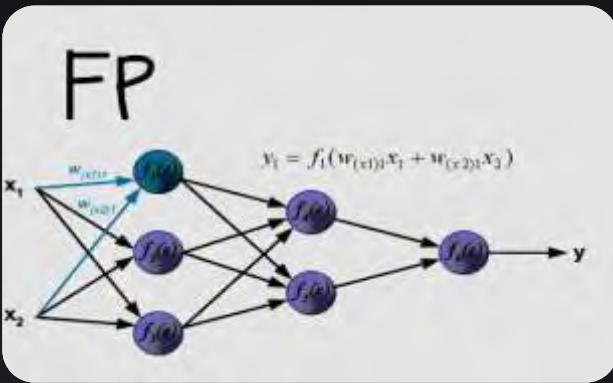
W.S.MeCulloch & W.Pitts
MP模型 1943年



F. Rosenblatt
感知器 1957年
第一次研究高潮



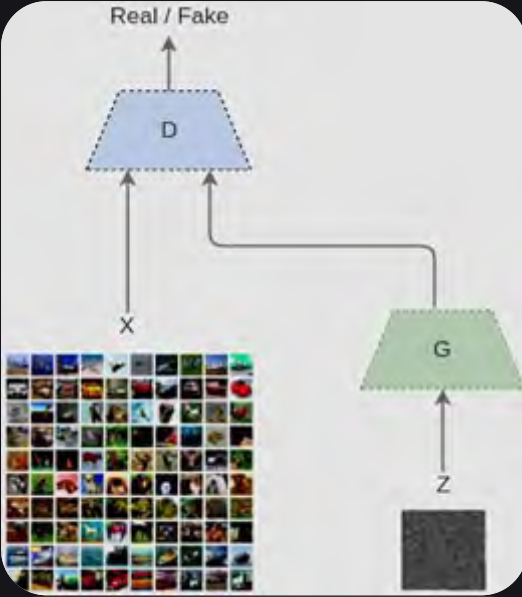
Hopfield
Hopfield模型 1982年
第二次研究高潮



Rumelhart
BP网络 1986年

51CTO WOTI
深层模型

DBN, CNN, AE, VAE, RNN, GAN
2006年至今
第三次研究高潮



1.3 深度学习当前研究热点

深度学习是推动人工智能发展的核心技术之一，目前，关于它的研究主要集中在以下方向：

- 深度强化、迁移学习（感知能力与决策能力的有力结合，AlphaGo，无人超市值守）；
- 深度对偶学习（对偶框架下的无监督学习）
- 深度递归/循环神经网络（自然语音处理，如机器翻译）；
- 并行化深度学习（如基于TensorFlow，Caffe基于算法设计下的多通路、多分辨深度网络搭建与平台设计）；
- 深度生成网络（深度置信网络，生成式对抗网络，变分自编码器，深度贝叶斯网络）；
- 深度融合网络（深度森林、深度ADMM，深度极限学习机等）；
- 基于卷积神经网络的目标检测、识别与追踪（RCNN，SPP-Net，Fast-RCNN，Faster-RCNN，R-FCN，YOLO，SSD等等）

1.4 深度学习模型的演化及思考

第一代深度学习模型及其基础模块

深度前馈神经网络— FNN , MLP

深度卷积神经网络— CNN

深度循环/递归神经网络— RNN , $LSTM$

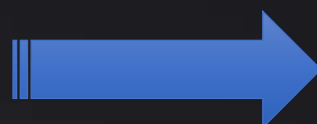
稀疏深度神经网络— $S-HMAX$

深度堆栈网络—自编码网络 AE

深度置信网络—受限玻尔兹曼机 RBM

深度贝叶斯网络—贝叶斯网络

特点：输入与输出单网络描述，以及不同策略（稀疏性、正则化及DropOut等）下深度网络性能的提升，有效克服数据量少、网络非凸性以及层级优化收敛性弱等缺陷。



后深度学习时代网络及代表模型

深度融合网络—深度森林、深度极限学习机

深度生成网络—生成式对抗网络 GAN

深度强化学习—深度 Q 网络

深度迁移学习—对偶机制

深度复域网络—深度复卷积神经网络

深度二值/三值神经网络

深度脉冲神经网络

特点：从单网络到多网络间的协同优化；感知、决策与认知的有效结合；数据规模及硬件加速平台的发展。以及通过数域的延拓与计算的简化，提升网络的时效。

二、人工智能与智造

- 01 深度学习应用于无人驾驶
- 02 深度学习应用于无人超市值守
- 03 深度学习应用于机器翻译
- 04 深度学习时代: *Data is King , Model is Queen , & Compute is Wing*



中国智造·工业大数据是智能制造的基础

人工智能 简单来说分为两大类，一类是模拟人脑工作机制、行为方式，是仿脑类脑的技术。另一类是快速的认知，因为人的大脑对大数据的认知本身没有那么快，但获取数据的速度极强，可以从数据中分析出人类认知问题特定的方式方法，这部分就是数据智能，也叫人工智能。

中国科学院院士·徐宗本

2.1 深度学习应用于无人驾驶



目前，无人驾驶要具备**环境感知**，**高精度地图导航**和**驾驶决策**

三个部分的能力。首先，**环境感知**包括车道线、车辆、行人、交通标志等的自动检测，使用深度学习可以训练出一个识别率非常高的分类器从而能够使环境感知部分以高精度完成，为驾驶决策模块提供正确的环境信息与认知，保证无人驾驶正常的完成。**核心技术：卷积神经网络，如YOLO，SSD模型。**

另外，在**高精度地图的创建与导航**方面也会用到深度学习，同样是因为深度学习有较高的识别率，基于视觉的高精度地图，就是通过识别某个特别的建筑标志，来定位车辆的位置，所以这也用了深度学习高识别率的特性。**核心技术：CNN，GAN。**

最后，**驾驶决策部分**，通过试错与自我修正的方法完成驾驶策略的决策。**核心技术：强化学习，如Q-Learning。**



2.2 深度学习应用于无人超市值守



现有无人超市

Amazon go

Take go

缤果盒子

淘咖啡

便利蜂

.....

目前，无人超市面向“人与商品”之间发生的关系（动作），通过传感器采集数据（人与商品间动作关系集），进一步使用深度学习技术建立人与商品之间的动作判别模型，提高系统反作弊/识别能力的同时，结合超市进出口支付系统，完成购物。

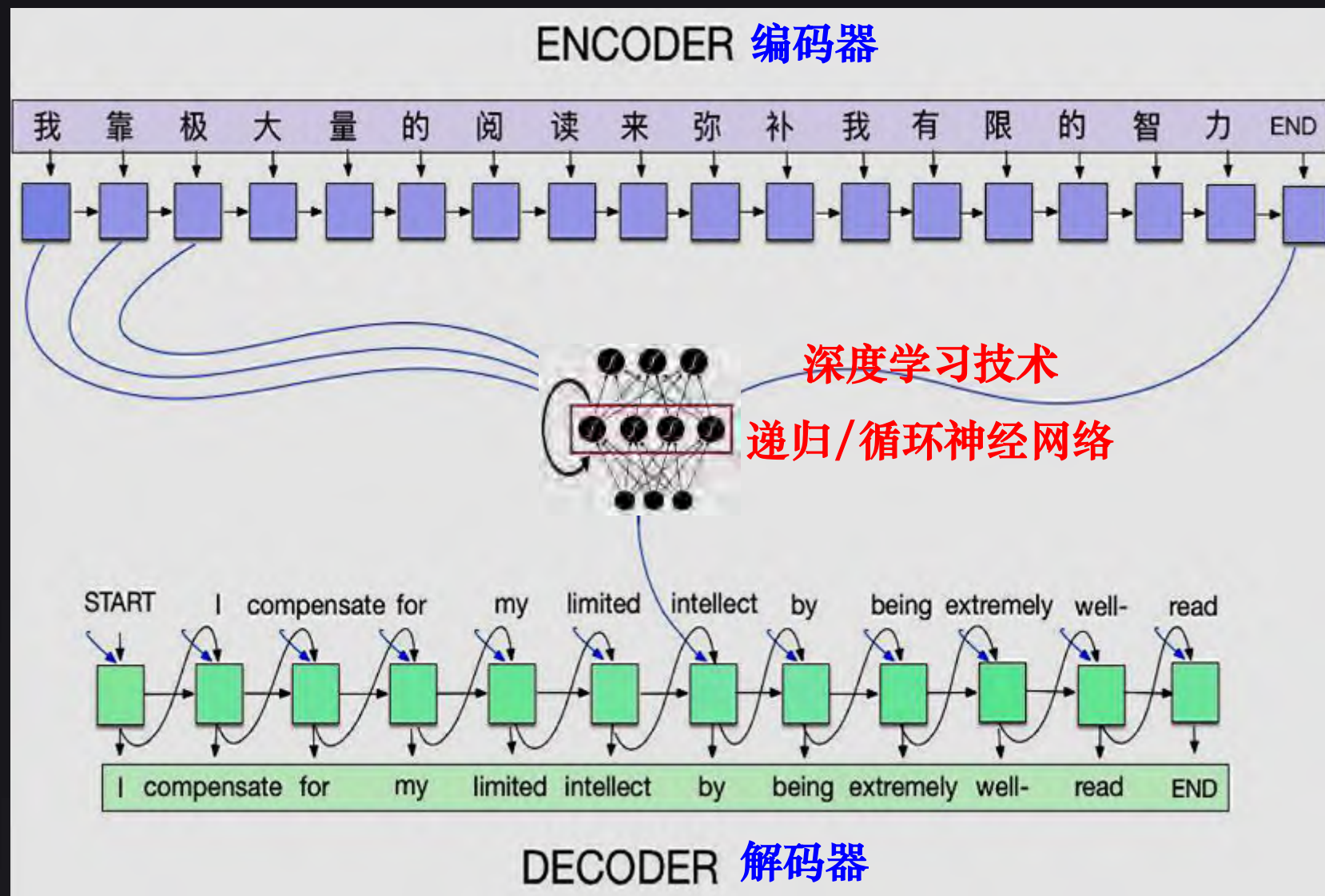
数据获取结构——商品位置变更，人行进路径的追踪，人与商品间动作的有效确定，商品识别；

模型设计原则——动作判别模型识别有效动作（拿或不拿商品），区分作弊动作（如商品放入书包等）；

核心模块构建——进出口扫码支付，行进路径追踪，人与商品间动作有效识别，数据采集与传输；

核心技术驱动——基于卷积神经网络的识别系统-例如网络模型SSD，YOLO...，因为无区域搜索技术的识别对应的fps高；

2.3 深度学习应用于机器翻译



机器翻译中深度学习核心技术

卷积神经网络; seq2seq;

递归/循环神经网络 *LSTM*, *RNN*;

生成式对抗网络、对偶机制——数据扩充、修复

技术: 编码和解码器均采用*RNN*网络

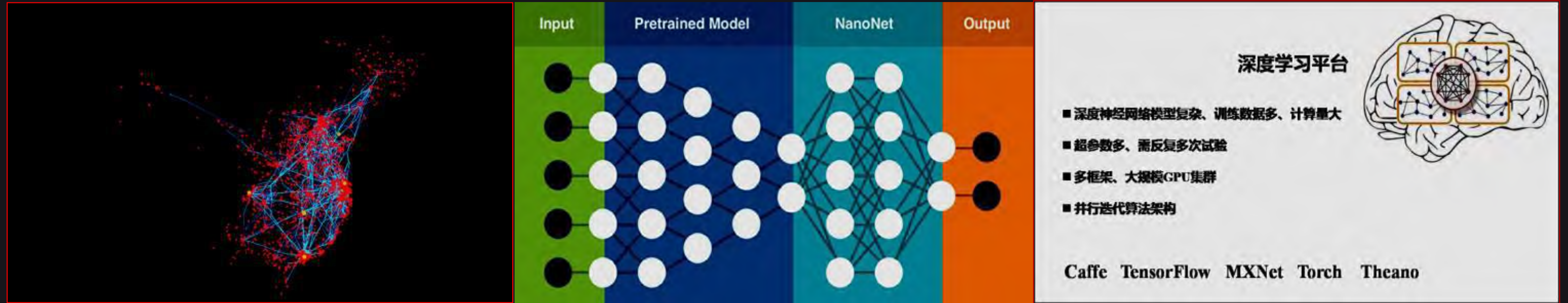
编码阶段: *RNN*把源语言句子映射成一个固定长度的输出隐向量;

解码阶段: 仍使用*RNN*把固定的长度的隐向量解码成目标语言句子。

- 1、编、解码均采用单层的*LSTM*;
- 2、比*LSTM*结构更为简单的*GRU*单元。

2.4 深度学习时代：*Data is King, Model is Queen, & Compute is Wing*

51CTO WOTI



数据获取与清洗技术

异常数据筛检—奇异性检测

数据去量纲化—归一化处理

数据预处理—降噪、修复、增强

模型构建与网络选择

图像处理—卷积神经网络

自然语音处理—递归神经网络

数据扩充与生成—深度生成模型

软计算与硬件平台

学习算法—批量随机梯度下降

硬件平台—Caffe等平台搭建

平构计算—权值共享、迁移学习

中国科学院院士 张钹：利用深度学习在解决实际问题时通常不需要专业知识，即深度网络能自动提取特征，不需人工参与，但高质量的数据是基础。



三、深度学习理论与应用实践

- ◆ 01 基于深度卷积神经网络的并行化与Caffe平台的实现
- ◆ 02 对偶范式下的稀疏深度神经网络及其应用
- ◆ 03 基于CPU/GPU并行计算下的遥感影像大数据深度学习

大数据+大模型+大计算=深度学习

深度学习 第一，它可以解决一些不能完全表述的问题，也就是说“知其然，不知其所以然”的问题；第二，它能针对不确定性的问题，不断的变化，不断的再学习。深度学习还有一个完全改变传统神经网络的作用，它不仅仅做函数映射，更重要提取那些多层次重复的模式，这是我们讲的真正公共的特征。

中国科学院院士·张钹

3.1 基于深度卷积神经网络的并行化 与Caffe平台的实现

深度卷积网络模型的改进与加速

如何学习滤波器组?

- 稀疏编码
- 卷积稀疏编码 (反卷积网络)

如何使得特征具有变换不变特性?

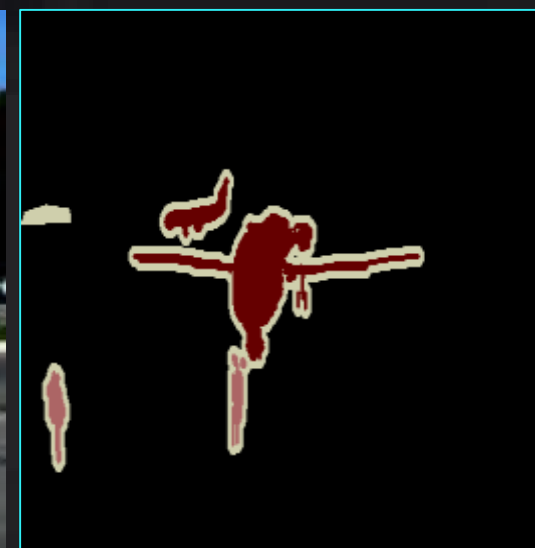
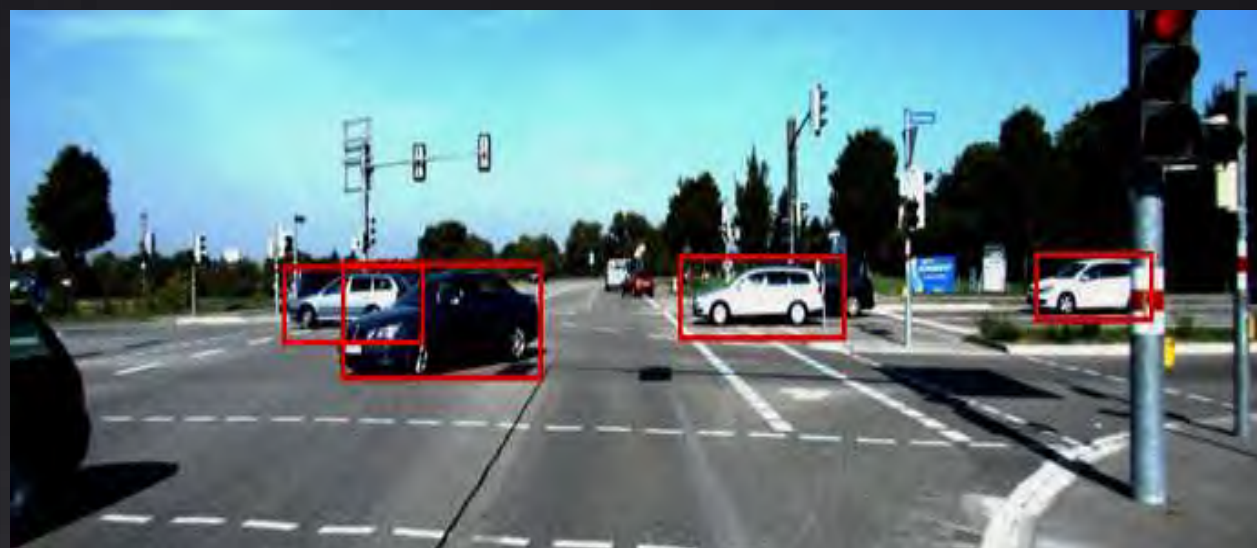
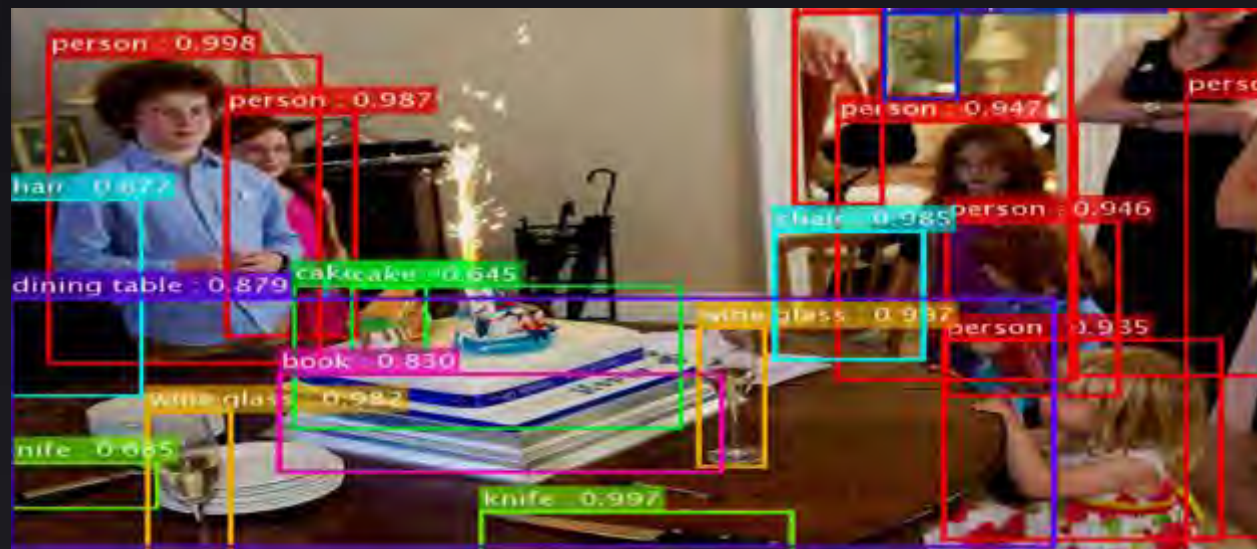
- 池化方式 (最大、平均、正则池化等)

如何加速深度卷积网络的学习?

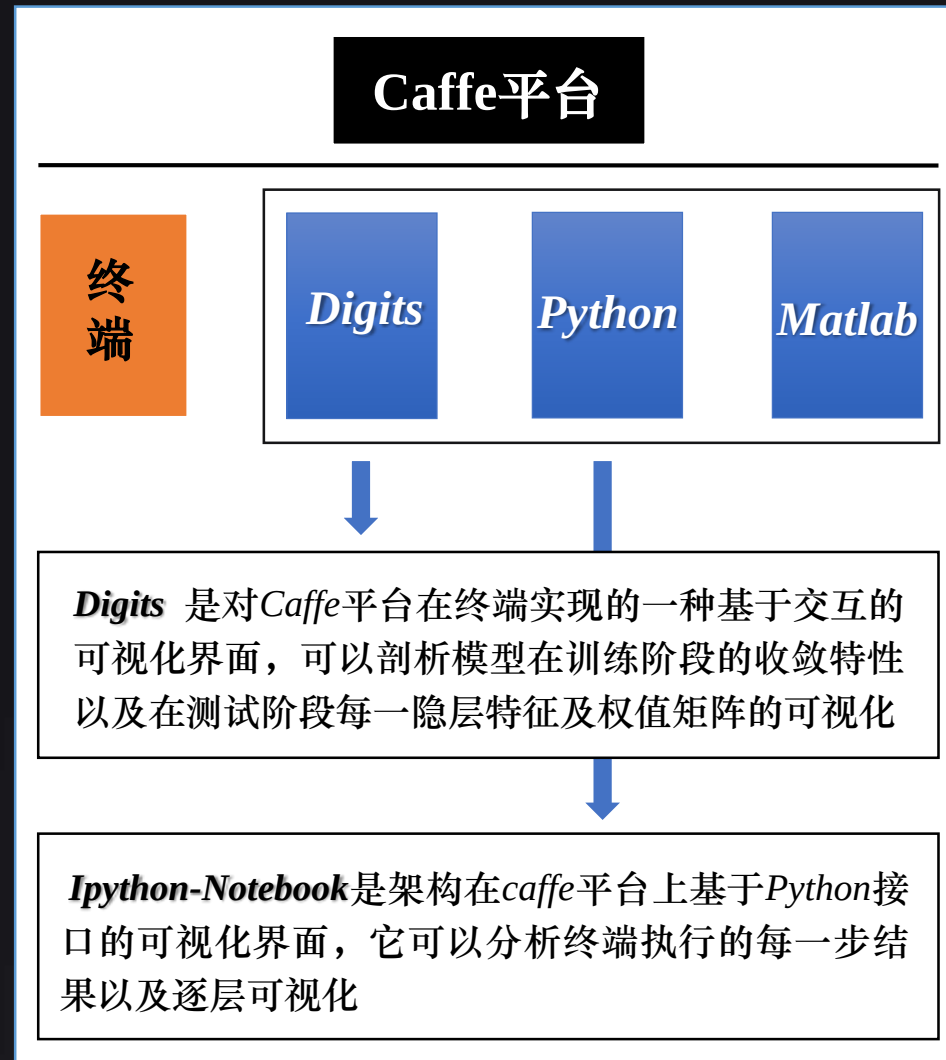
- 模型各个模块的合理架构
- 各种技巧的使用
- 多GPU并行加速

相比于浅层学习, 深度学习较少的依赖数据先验知识!

—— Yoshua Bengio



Caffe: Convolutional architecture of fast features embedding



Caffe Demos

The Caffe neural network library makes implementing state-of-the-art computer vision systems easy.

Classification

Click for a Quick Example

Classification	Score
vehicle	0.59614
conveyance	0.57206
craft	0.49646
vessel	0.47302
wheeled vehicle	0.37942

CNN took 0.061 seconds.

Provide an image URL

Or upload an image: 未选择文件。

© BVLC 2014

目前平台可实现功能

分类

分割

回归

目标检测、识别

医学影像病灶检测

数据生成、修复

.....

平台底层模块

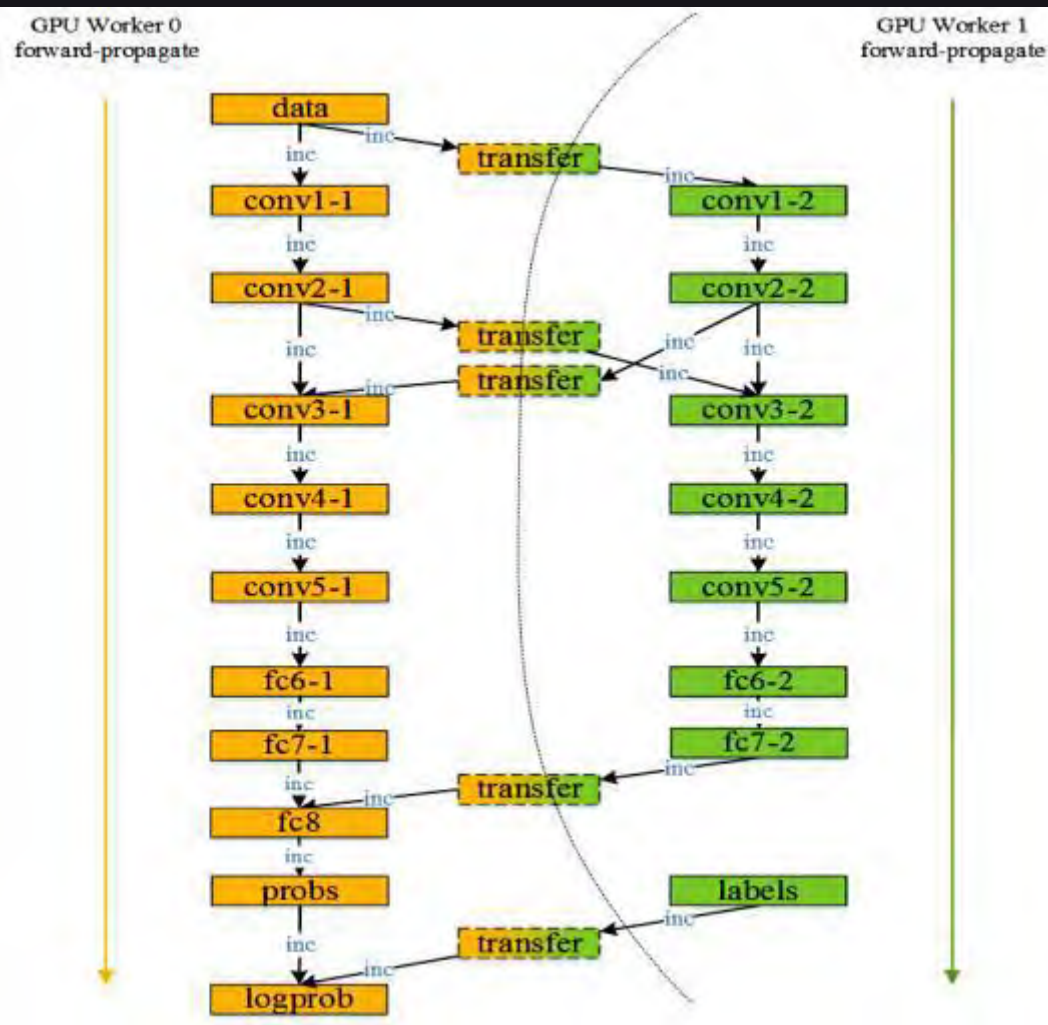
数据导入与生成

模型构建与传输

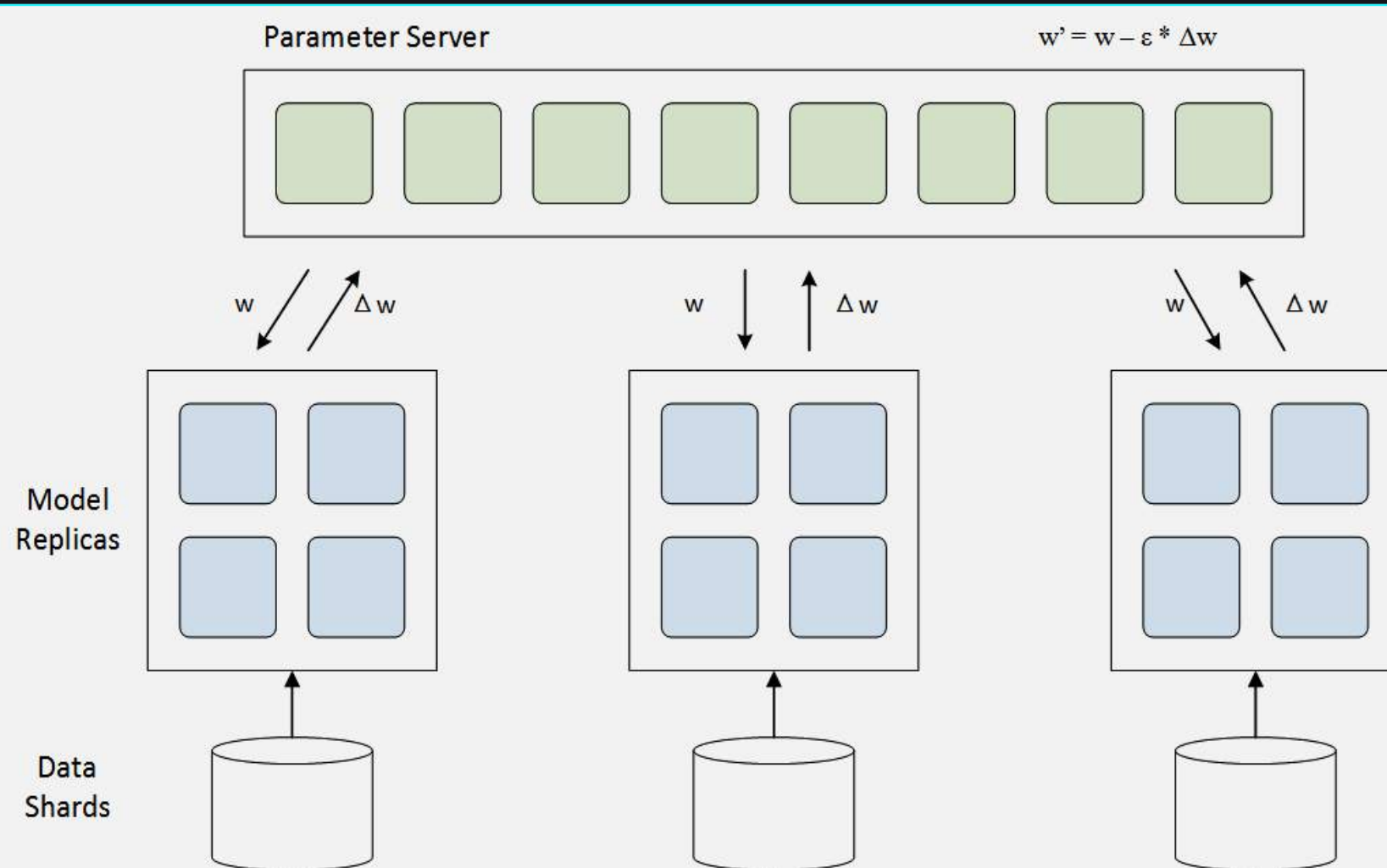
参数传输协议

网络参数更新与反馈

模型并行 适当拆分模型到不同的计算单元上，利用任务的可并行性达到整个模型在计算过程中并行化效果



数据并行 对输入数据集合进行分区，以期多个并行处理单元能够同时对不同的子数据集合进行操作

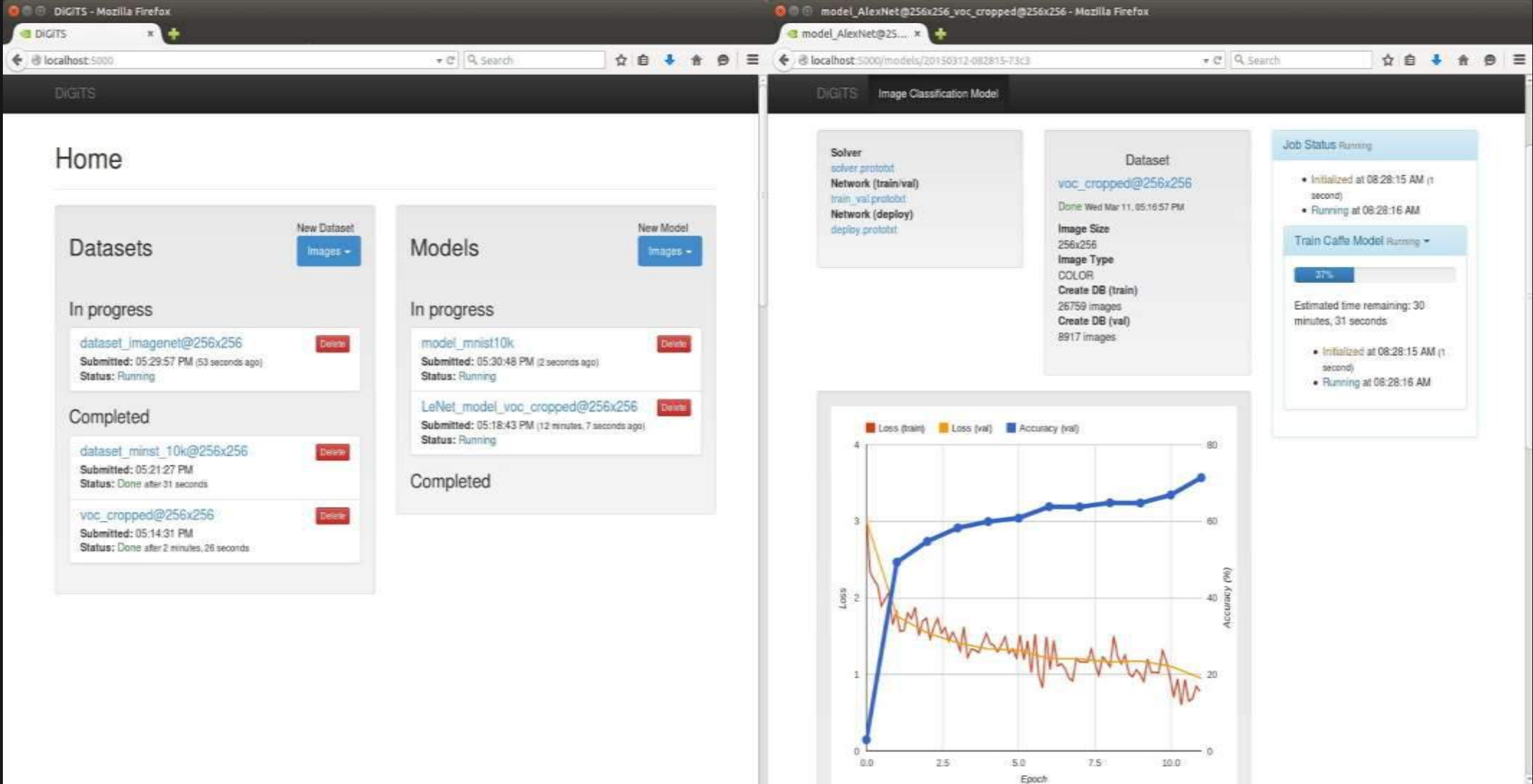


注意：在数据并行的实现中，影响性能的瓶颈在于多GPU间的参数交换；由于参数交换是硬件设计的瓶颈问题

深度检测网络——训练模型



基于Caffe平台的Digits可视化界面



Digits模型	CLI argument	Classes
DetectNet	CoCo -airplane	airplanes
DetectNet	CoCo-bottle	bottles
DetectNet	CoCo-chair	chairs
DetectNet	CoCo-dog	dogs
Ped-100	Pednet	Pedestrians
Multiped-500	multiped	Pedestrians luggage
FaceNet	face	faces

51CTO WOT!



3.2 对偶范式下的稀疏深度神经网络及其应用

研究动机——对偶学习机制对深度学习的意义

对偶学习：核心是利用任务的对称属性使其获得更有效的反馈/正则化，特别是在数据量少的情况下，从而引导、加强学习过程，内蕴合作机制。

人工智能中任务的对称属性，常见的有：

机器翻译：有英文翻译为中文，也有中文翻译为英文

图像理解：图像描述（图像到文字）与图像生成（文字到图像）

对话任务：问答系统与问题生成

语音处理：语音识别（语音到文本）与语音合成（文本到语音）

.....



深度网络的性能：引导（反馈，正则化，先验认知） + 小数据集 等价于 大(海量)数据集？

研究动机——稀疏性对于深度神经网络的有效性

稀疏性：深度神经网络与稀疏性之间的关系深刻而且本质，从机器学习中的特征工程到深度学习中的特征学习，无论是以显性还是隐性的嵌入方式，稀疏性都在模型中扮演着重要的角色。

常见的稀疏性有：

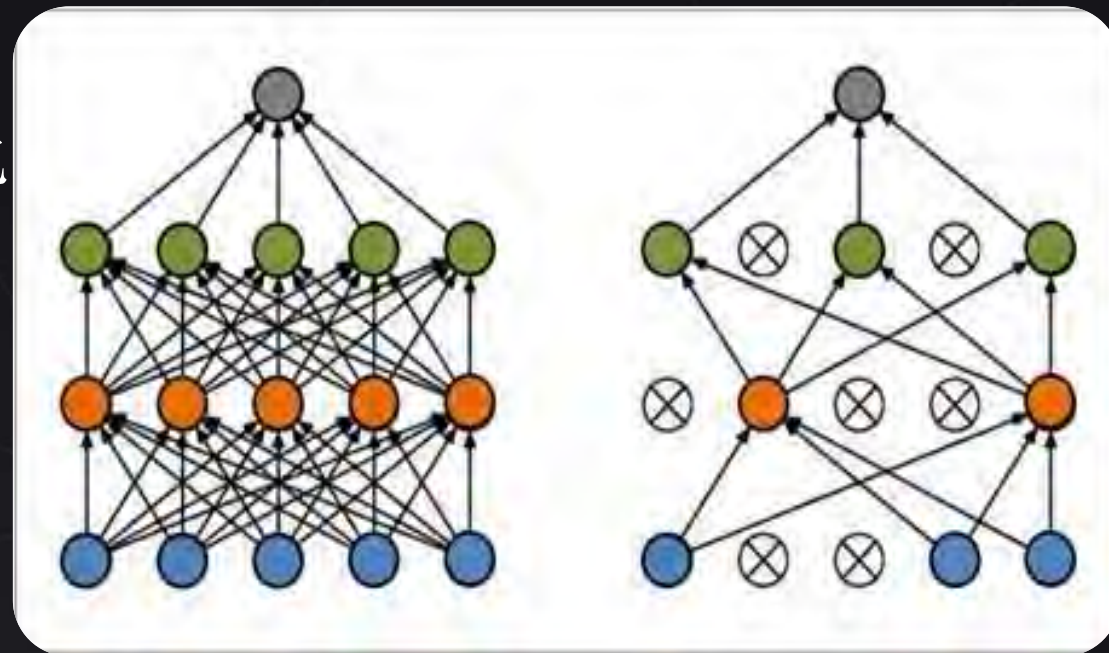
数据的稀疏性：数据在某种变换下对应的表示系数具有非零元素较少的状况

稀疏正则：使用 $L1$ 、 $L0$ 等范数或伪范数可以诱导出稀疏特性。

稀疏链接：训练技巧DropOut等。

稀疏激活响应：修正线性单元ReLU。

稀疏分类器：从Softmax到Sparsemax。



稀疏性能否提升深度网络的性能？参数稀疏化+小数据等价于大(海量)数据？

研究动机——生成性与判别性对深度学习框架的冲击

判别性：不考虑数据的分布，由训练数据直接学习预测函数；另外对数据进行特征表示，以期简化学习任务，进而提高学习效率及准确率。**生成性：**学习数据的概率密度函数，来获取同类、不同风格数据的扩张，并且根据分布特性引入“特定”先验正则约束。

深度学习框架：

端到端的设计模式：层级特征学习/表示+分类/回归器设计

数据量的规模：如深度网络($DFNN$, $DCNN$, DBN)—参数量约等于训练数据量

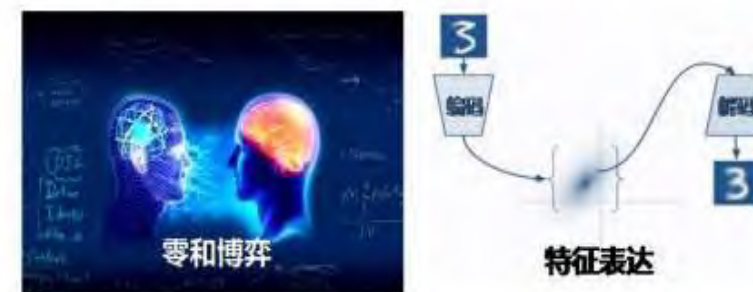
生成性引入深度网络中的冲击：数据扩张

典型网络框架：生成式对抗网络GAN，变分自编码器VAE，深度贝叶斯网络

深度生成网络——生成样本、数据缺失填补

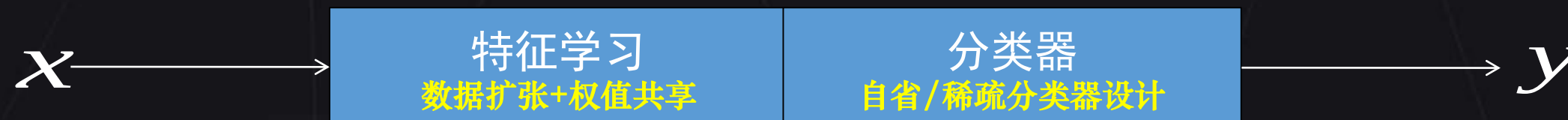
■ 生成式对抗网络——数据空间

■ 变分自编码网络——特征空间



少量真实数据集（取决于生成风格各异数据的质量） + 生成的伪数据集 等价于 大(海量)数据

实践应用——分类任务--逐阶段参数初始化、数据扩张



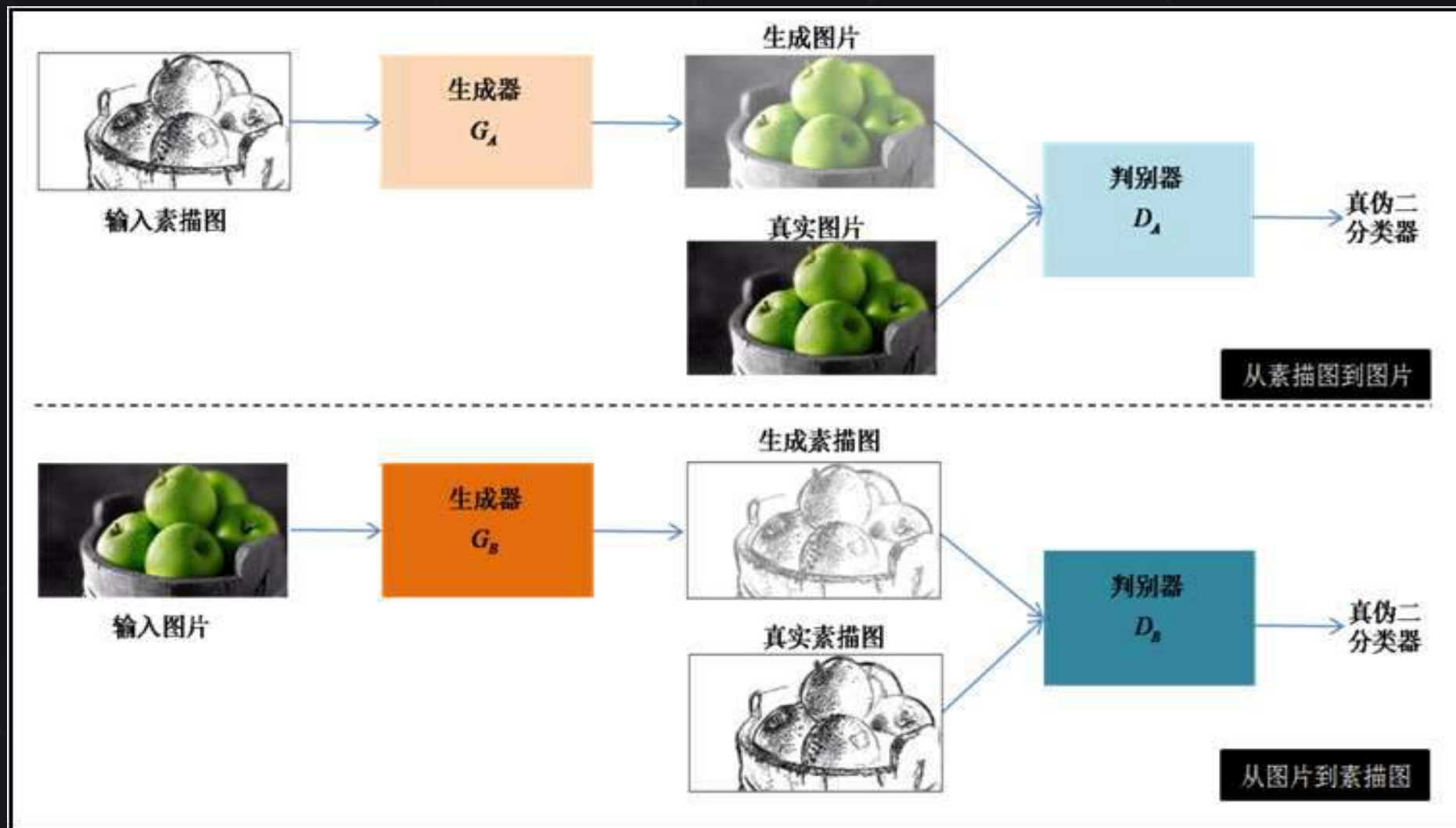
- 逐阶段学习+精调机制。
- 深度网络层级规模不限于数据的拓扑特性与数量。
- 生成式判别模型——开放式。
- 网络泛化性能自适应训练集与测试集的相似程度。
- *Reclassification—Step*，样本自省更新。

数据扩张：生成器，无监督学习（生成器的参数）+有监督精调（每一类精调更新）——每一类扩张

权值共享：判别器（除去真伪二分类器，即特征学习部分的参数）

分 类 器：自省支撑矩阵机 [Introspective SMM Classifier Learning](#)； 稀疏分类器设计 [Sparsemax](#)

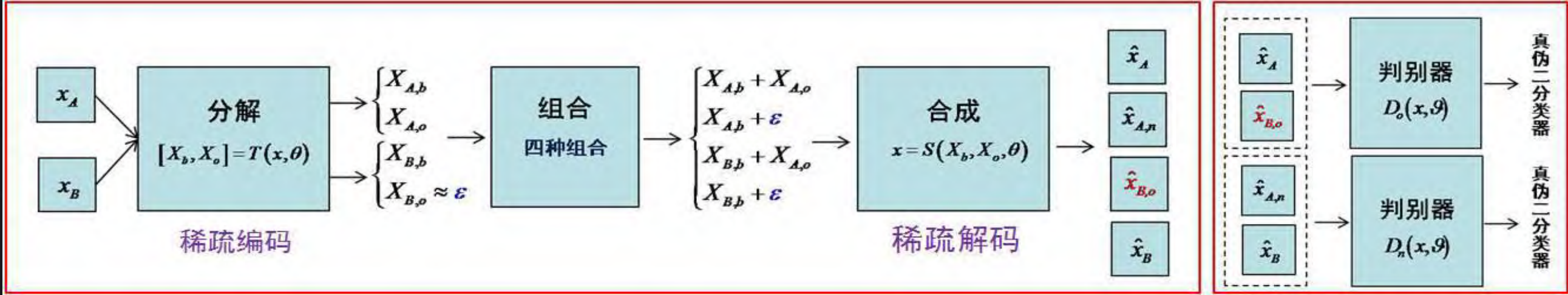
实践应用——图片风格迁移——对称数据对



从素描图到彩色图片间的风格转化，利用对偶范式下的稀疏深度神经网络。充分利用未标注的数据，提高对偶任务中的两个“翻译”模型的性能。



目标变形:
针对非对称数据对，即没有逻辑或拓扑结构上的输入与输出间关系，期望将一幅图中的目标特性移植到另一幅图中，前提是另一幅图中没有目标特性。例如，对于人，有笑脸、眼镜、帽子、发型等都可视为目标特性。



3.3 基于CPU/GPU并行计算下的遥感 影像大数据深度学习

遥感影像大数据类脑解译系统

影像大数据类脑解译系统是由30个惠普Z840工作站构成，总计算速度可达近600万亿次的高性能计算集群。该系统能够快速地对大规模影像数据（遥感影像、视频、图像等）的类脑解译处理，具有计算快、稳定性强、吞吐量高、任务并行等优势。系统主要配置参数如下：

处理器：60颗10核Intel Xeon E5-2630v4/10核心；

内存：2TB 2400MHz DDR4 ECC；

硬盘：15TB SSD，120TB SATA3 6Gb/S；

显卡：60块NVIDIA GTX 1080；



分布式深度学习系统

分布式深度学习系统是由20台AMAX工作站构成，总计算速度可达近800万亿次的高性能分布式计算集群。该系统通过数据并行或模型并行的形式将模型训练任务下发到各个节点，实现基于大规模数据的深层复杂模型训练。系统主要配置参数如下：

处理器：40颗Intel Broadwell E5-2650V4/12核心；

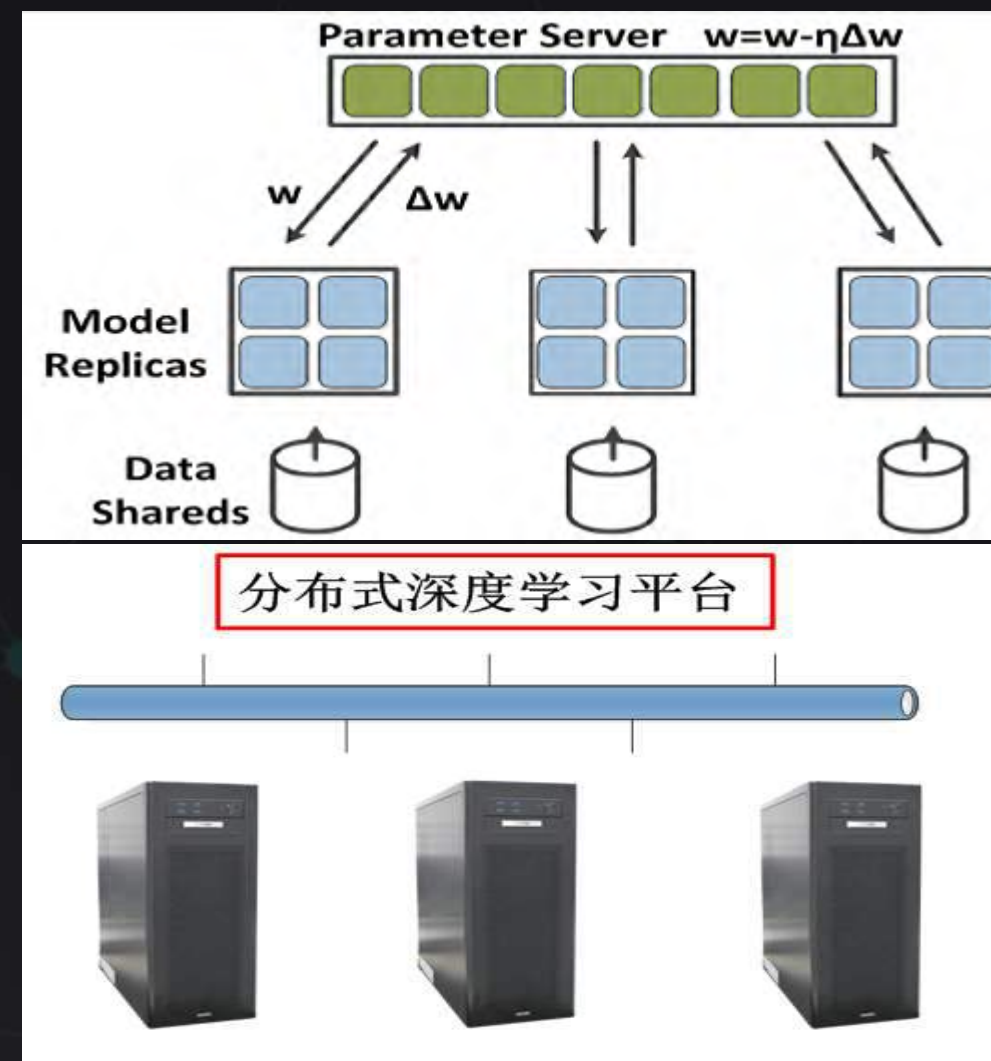
内存：5TB 2400MHz DDR4 ECC；

硬盘：10TB SSD, 160TB SATA3 6Gb/S；

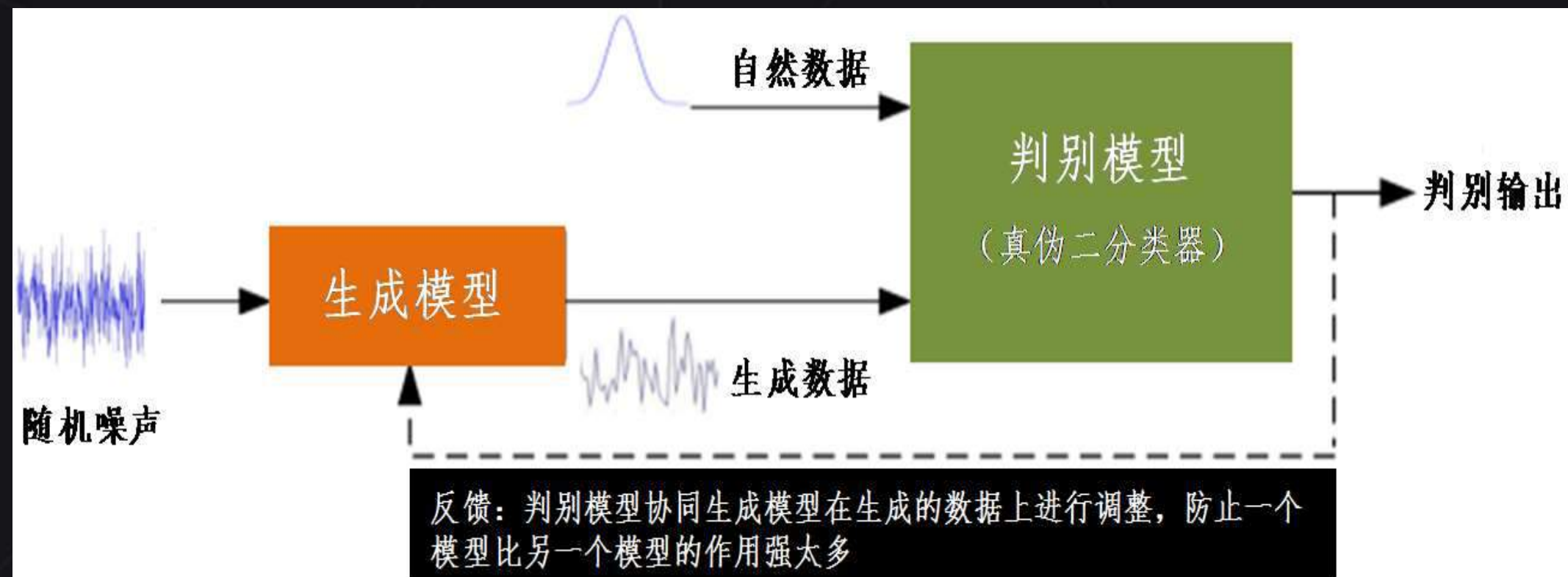
显卡：80块Nvidia GTX 1080；

分布式深度学习平台：

Tensorflow、PaddlePaddle、Caffe、MxNet

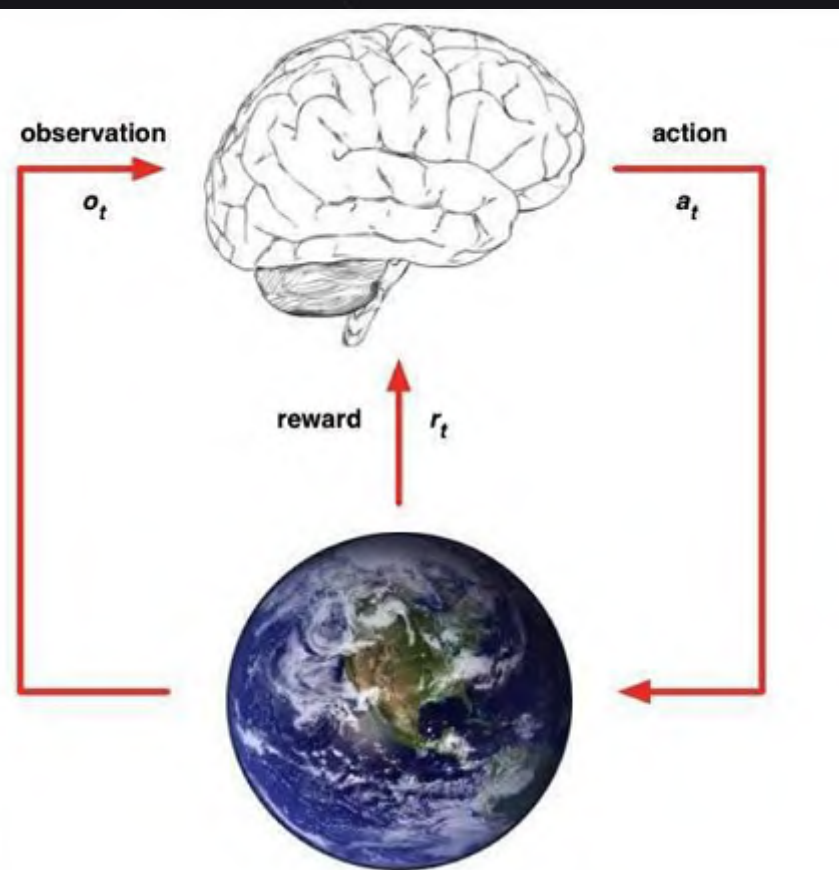
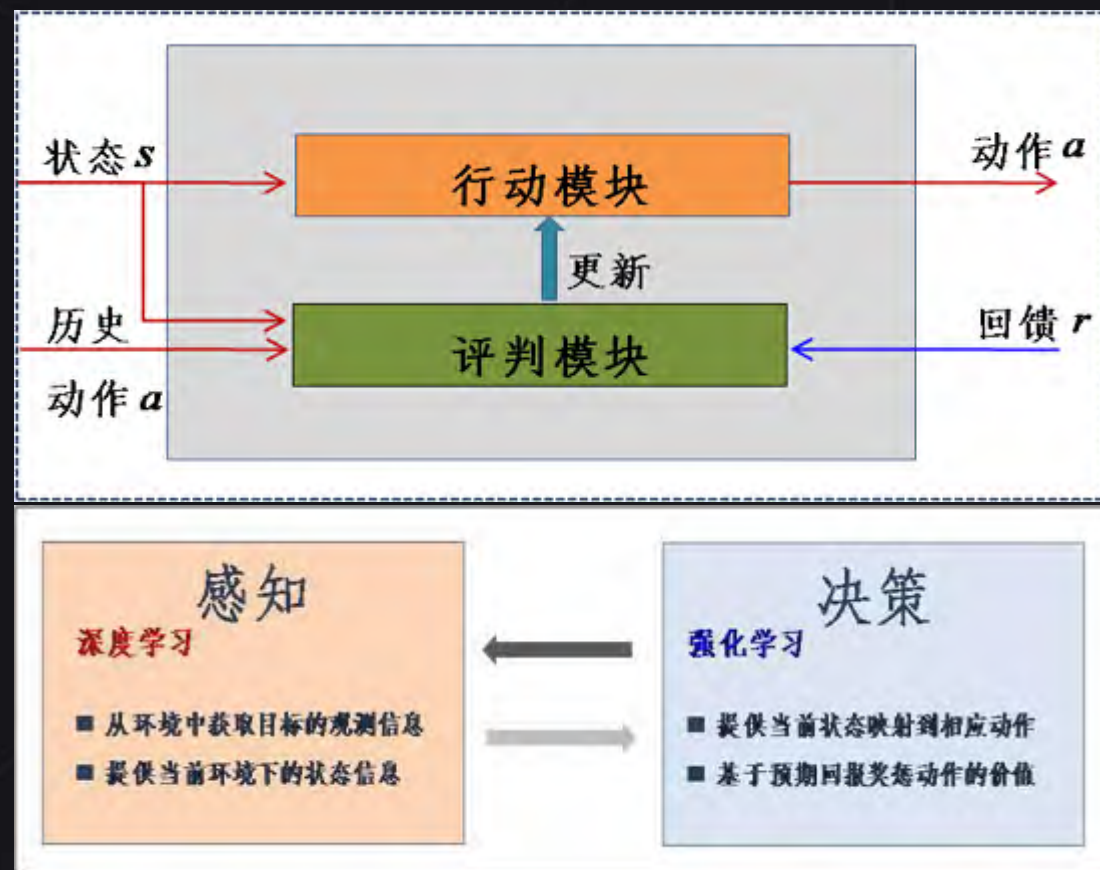


后深度学习网络模型——生成式对抗网络



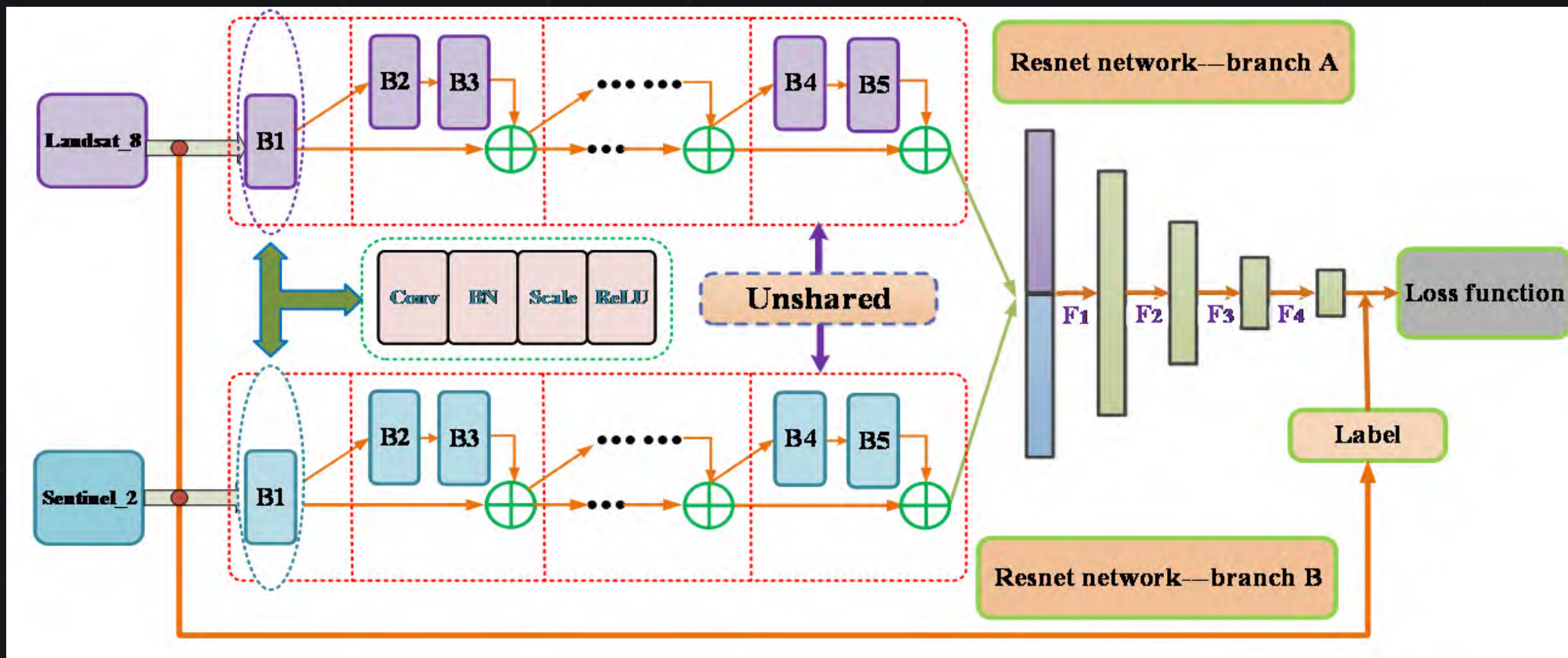
对抗网络结构包括生成网络和判别网络，其中的生成网络用于捕获数据的分布，判别网络来估计一个概率，即样本来自于训练样本而不是生成网络所产生的样本的概率。优化时，训练生成网络以最大的概率去使得判别网络做出错误的决定；而判别网络希望不被生成网络所产生的数据所欺骗。常见网络模型：*DCGAN*，*W-GAN*，*LS-GAN*，*EB-GAN*，.....

深度强化学习（感知 + 决策）



强化学习 是受到生物能够有效适应环境的启发，以试错的机制与环境进行交互，通过最大化累积奖赏的方式来学习最优策略。

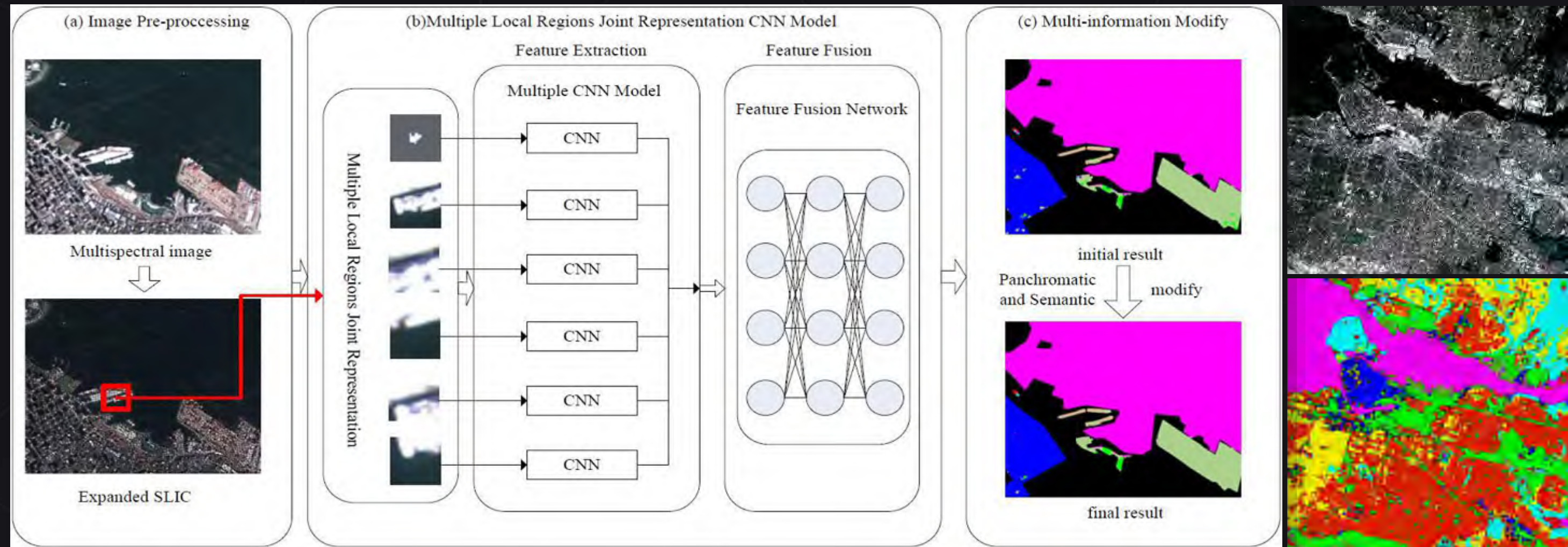
Two-branch ResNet Network for Multisource Remote Sensing Imagery Classification



双通道深度残差网



Supervoxel-based Multiple Local CNN for Panchromatic and Multispectral Image Classification





四、后深度学习的挑战与思考

- 01 深度学习的局限性及其未来
- 02 智能制造的大数据机遇与挑战
- 03 最新后深度学习的模型进展
- 04 深度学习未来之路

4.1 深度学习的局限性及其未来

深度学习的局限性： 某种准则下，深层网络不见得比浅层网络性态好

1. 数据：深度网络应用于小数据集，使得生成模型、对偶机制等逐渐融入深度学习理念中…
2. 模型：深度网络模型预测或表征能力强大，但却可解释性差，其模型被视为“黑箱” …
3. 优化目标函数：训练深度网络在计算和调试方面成本代价太高，并且对于许多日常预测问题并没有任何意义…
4. 求解：初始化策略与梯度更新相结合，使得网络模块化贡献可区分性越来越模糊，并且依赖众多平台使得求解比重减小…

深度学习的未来： 商业应用将进一步推动深度学习技术日臻成熟

1. 应用：唇语识别，作诗，智能驾驶技术，无人超市值守，智能盲杖，语音识别，自动新闻标题生成，图像修复，推荐系统，…
2. 理论框架：卷积神经网络，生成式对抗网络—深度生成模型，递归/循环神经网络，复杂深度神经网络系统，深度强化/迁移学习…
3. 硬件平台：分布式深度学习系统提升计算能力…

4.2 智能制造的大数据机遇与挑战

明确目标是前提：每个地区、每个政府、每个企业，要解决的问题不一样，必须要真正解决问题，大数据才有用。

拥有数据是基础：大数据产业就是以现代技术设施为基础，以数据为生产要素，以数据的价值挖掘为创新活动的产业，叫大数据产业。因而没有数据谈不上大数据产业，“中国制造2025”制造大数据也非常重要。

计算平台是支撑：没有一定的计算架构和计算平台，计算不了。它是支撑作用，但做企业的人不必过分强化，也不必过分低估。

分析技术是核心：模型是核心，分析模型的稳定性、收敛性等，以及模型的合理架构及背后的技术经验。

产生效益是根本：数据是基础，平台是支撑，技术是核心，赚钱是王道。

大数据是首先解决数据采集问题，就是互联互通问题；定制化服务是中心，基本模式要转变，懂数据会分析是关键。

4.3 后深度学习模型的研究进展

4.3.1 形变卷积神经网络——目标检测、识别与追踪

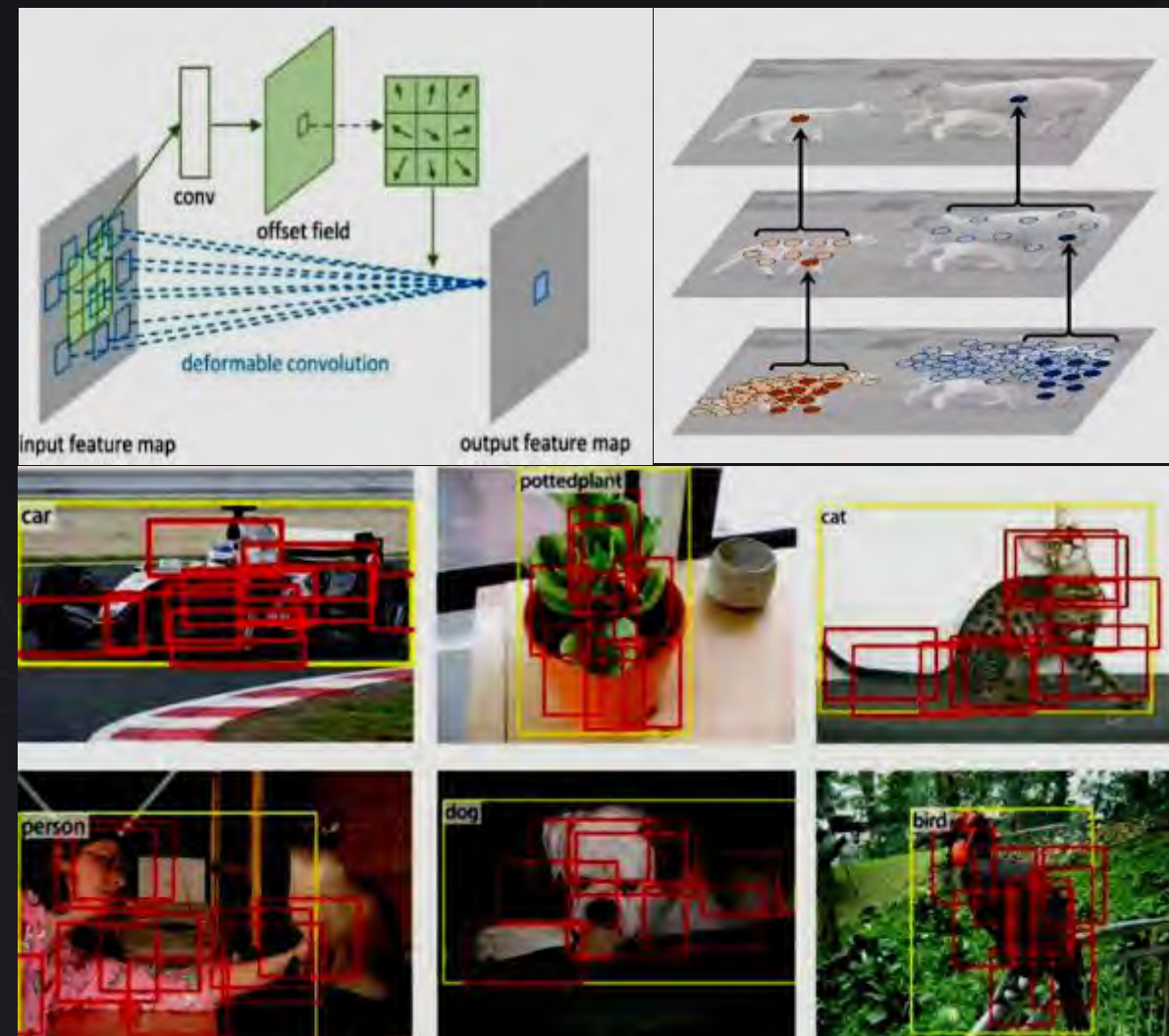
生物机理： 固定(局部)感受野到自适应(局部)感受野。

数学分析： 卷积操作到形变卷积操作；池化操作到形变池化操作。

物理解释： 加入形变因子（待优化参数），随着层级的增加，可保持父节点（上一层特征映射图中的某一点）与子节点（下一层特征映射图中若干点，子节点的个数取决于滤波器的大小）的相似性。

应用场景： 图中显著区域提取、场景候选目标提取等。

参考论文： *Deformable Convolutional Networks*



基于区域的目标检测模型： R-CNN, SPPNet, Fast-RCNN, Faster-RCNN, R-FCN, Mask-RCNN... **非区域的目标检测模型：** SSD, YOLO, YOLO2, ...

4.3 后深度学习模型的研究进展

4.3.2 深度复域神经网络——深度复卷积神经网络

动机：目前，深度学习的绝大多数构建模块、技术和架构都基于实数值的运算和表征。但是，近来在循环神经网络和其它更古老的基础理论上的分析表明复数可以有更加丰富的表征能力，也可以促进对噪声鲁棒的记忆检索机制。

创新：复卷积流，即复卷积、复池化，复非线性，复归一化操作

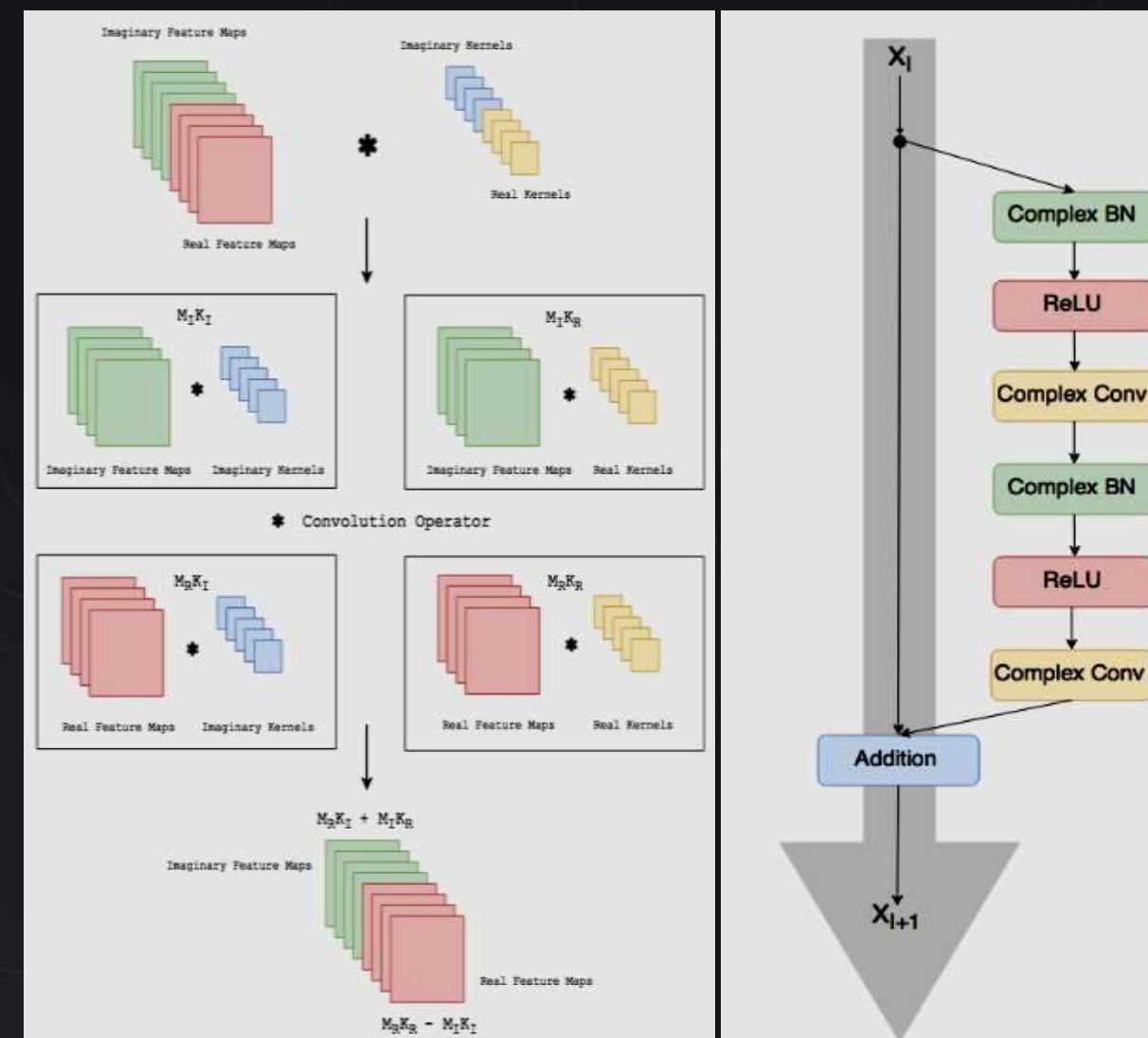
贡献：对复数批量归一化给出了数学描述；复数权值初始化，...

应用场景：遥感影像大数据相关处理任务，如飞机、舰船等的检测。

参考论文： *Deep complex network*

Deep complex network = Deep real network + Deep imaginary network ?

源代码： https://github.com/ChihebTrabelsi/deep_complex_networks



4.3 后深度学习模型的研究进展

4.3.3 深度生成模型——生成式对抗网络

动机：深度学习应用于小样本数据集？如何做数据扩充？并且使得扩充后的样本质量较高？

创新：生成式对抗网络（生成器与判别器），零和博弈，无监督学习范式。

贡献：将判别性模型框架与生成性模型框架融合，给出了深度学习的新范式

应用场景：数据扩充、图像修复，风格转化，图像超分辨，语义分割，目标检测…

参考论文： *Generative Adversarial Nets*

方法整理： W-GAN, DC-GAN, EB-GAN, f-GAN, ……



图像风格转化

附更多生成式对抗网络文章、代码、应用下载地址：<https://github.com/zhangqianhui/AdversarialNetsPapers>

4.3 后深度学习模型的研究进展

4.3.4 从残差网络和稠密网络到双通道网络

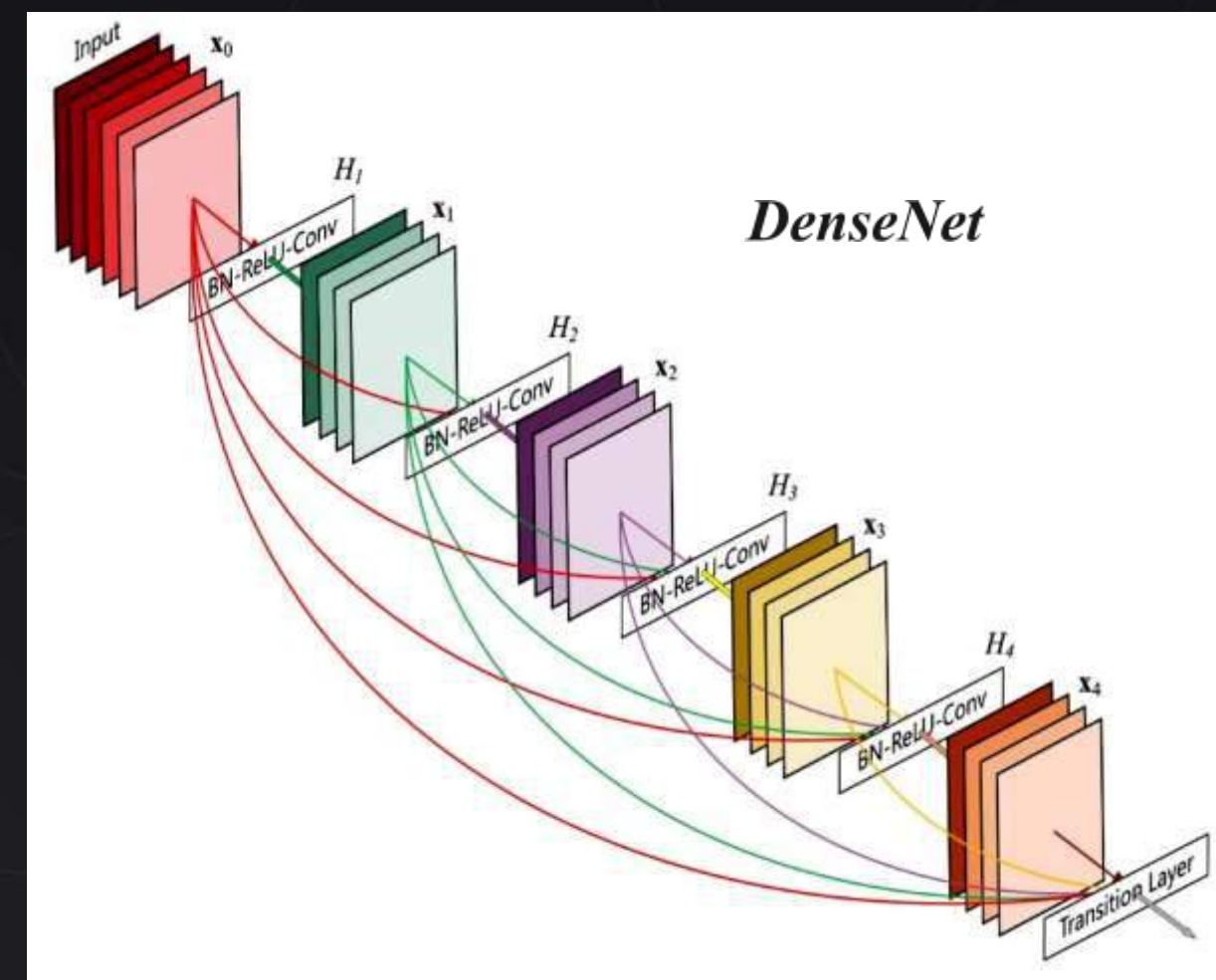
动机：构建高性能、低资源占用的网络结构一直是深度学习里一个重要的研究方向。一个有效的网络拓扑结构，不仅能全面提高图像识别任务中准确性，更能通过直接替换底层网络的方式受益图像生成、检测、分割、美化等等诸多应用。

关系：残差网络是稠密网络的一种特例，双通道网络结合了残差网络与稠密网络的各自的优势。稠密网络可以通过共享参数的方式，退化为残差网络。

认知：残差网络可有效地降低特征的冗余度，并重复利用已有特征，但缺点是难以利用高层信息再发掘底层特征；

三篇参考论文： *ResNet*, *DenseNet*, *Dual Path Networks*

源代码 <https://github.com/liuzhuang13/DenseNet>



4.3 后深度学习模型的研究进展

4.3.5 深度适配网络

背景及贡献：解决迁移学习和机器学习中经典的 $domain\ adaptation$ 问题，只不过是深度网络为载体来进行适配迁移。

创新：多层适配和多核 MMD 。

评价：DAN作为深度迁移学习的代表性方法，充分利用了深度网络的可迁移特性，然后又把统计学习中的 $MK-MMD$ 距离引入，取得了很好的效果。

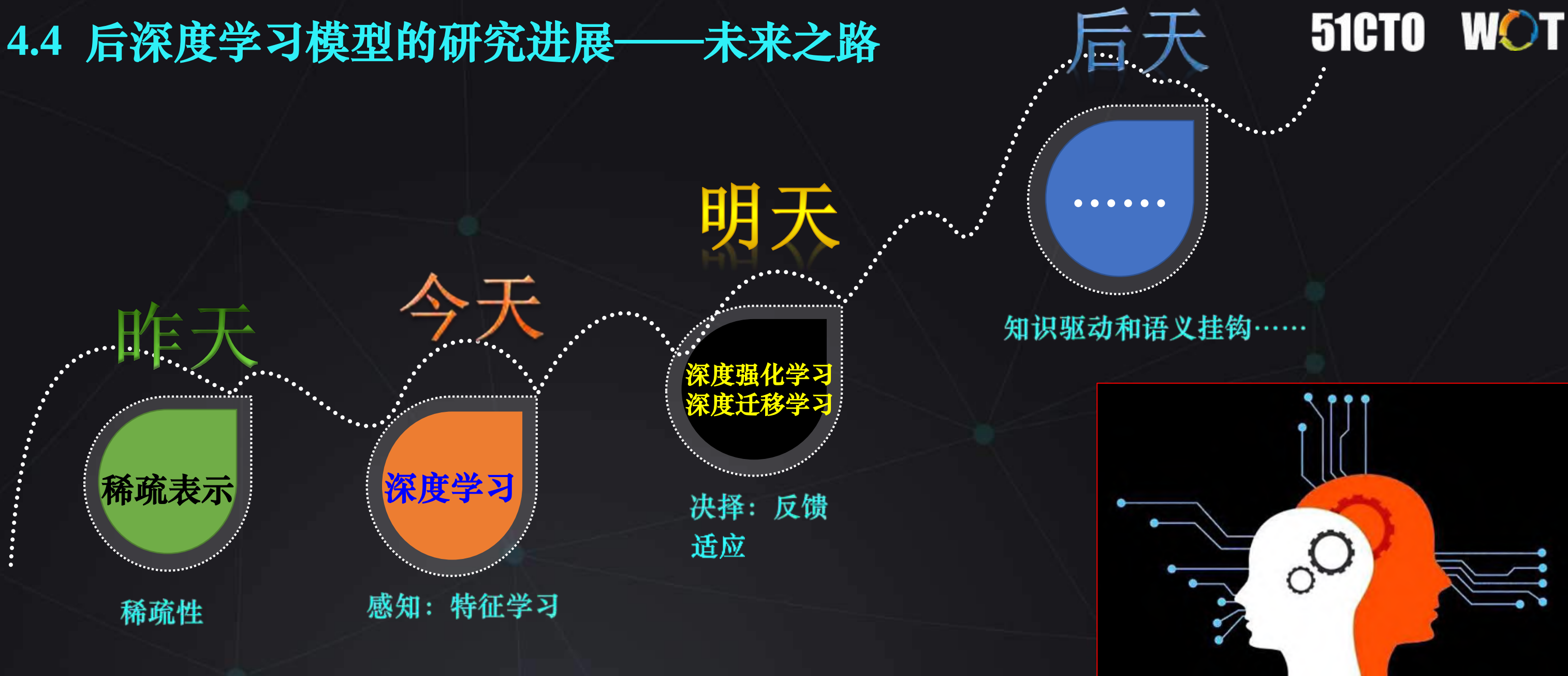
应用场景：作为深度迁移学习领域的代表性网络，可应用于非平衡数据集、小规模数据集等应用任务。

参考论文：[Deep Adaptation Networks](#)



4.4 后深度学习模型的研究进展——未来之路

51CTO WOTI



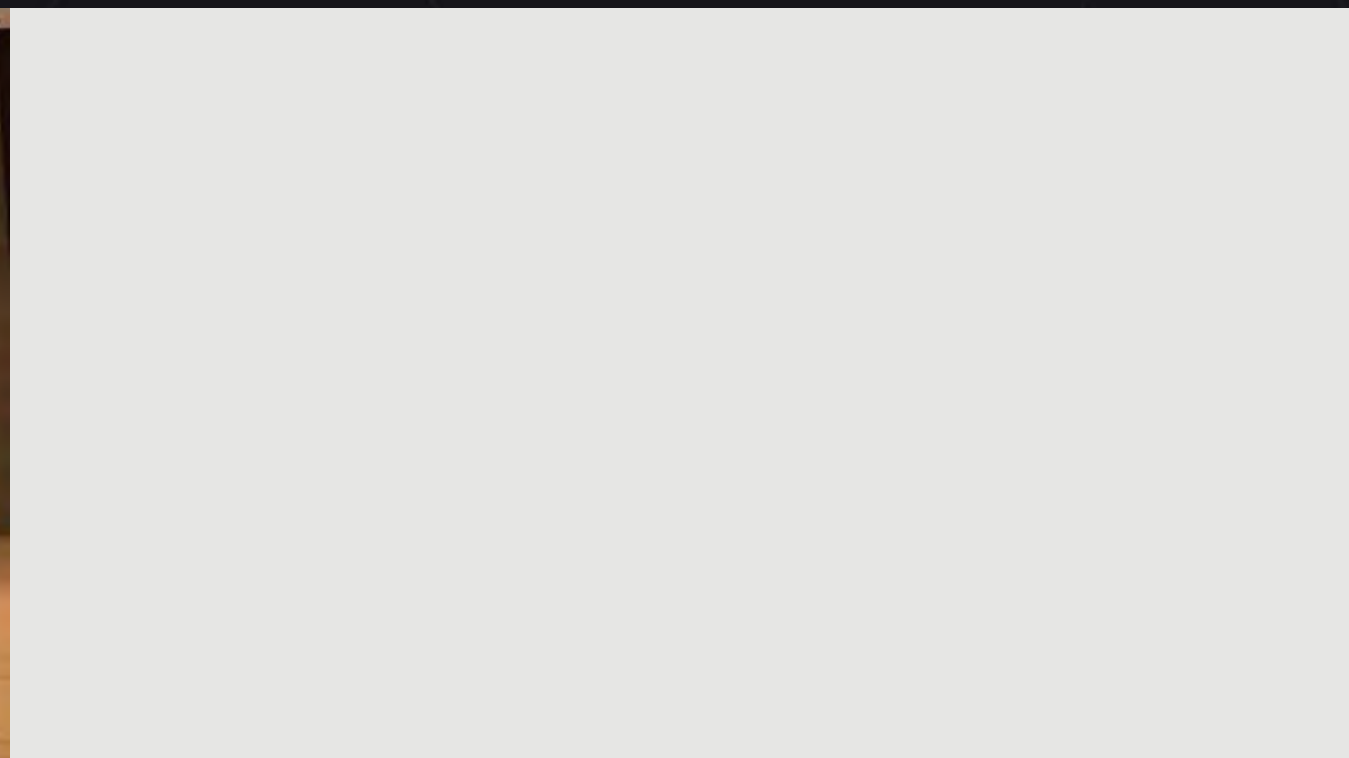
实验室介绍

51CTO WOT!



智能感知与图像理解教育部重点实验室 · 智能盲杖

51CTO WOTI



介绍： 本项目由盲杖与服务器端两部分组成。盲杖基于深度学习技术，通过摄像头采集盲人前方的道路图像，实时地规划出一条安全的“虚拟盲道”，通过声音及振动反馈道路状况，提高盲人的出行安全。为拓宽盲人朋友圈，我们创新性地引入“盲杖社交”的理念。服务器完成配套的处理及通信。系统按功能划分，分为“虚拟盲道”以及“盲杖社交”两部分。

真诚期待与您合作 · 智缘人工智能深度学习时代

51CTO WOTI

我们实验室今后的发展目标是面向国际学术前沿和国家重大需求、国际协同创新，在已有优势和积累的基础上，开拓新的理论方法的研究和应用，重点围绕大数据的智能感知与类脑计算中的关键问题及其在网络、视频、遥感大数据领域中的应用，不断创新，取得基础性和系统化的示范性研究成果。在未来几年内把实验室建设成为以国际学术前沿和国家战略需求为导向的、具有国际综合创新能力的科研与高层次人才培养及社会服务的国家基地。

扫描二维码，了解更
多实验室信息



51CTO WOT!

