

深度学习

主讲：王亚星、刘夏雷、郭春乐
南开大学计算机学院

致谢：本课件主要内容来自浙江大学吴飞教授、
南开大学程明明教授

提纲

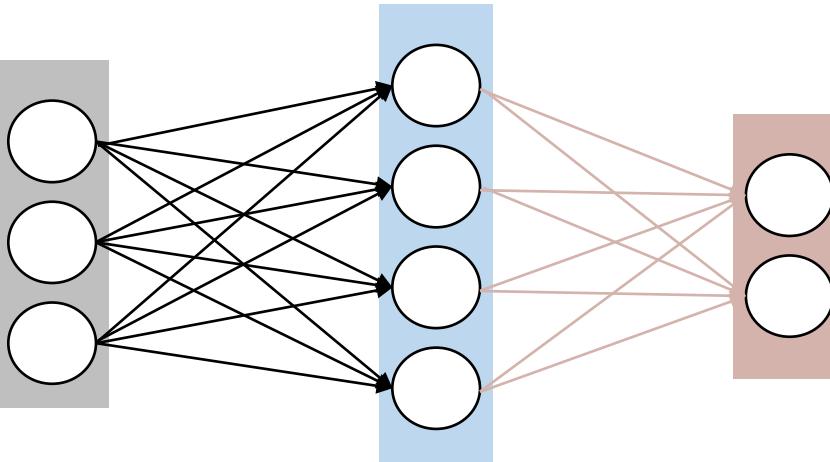
- 深度学习历史发展
- 前馈神经网络
- 卷积神经网络
- 循环神经网络（了解）
- 深度生成学习（了解）
- 深度学习应用

课程信息

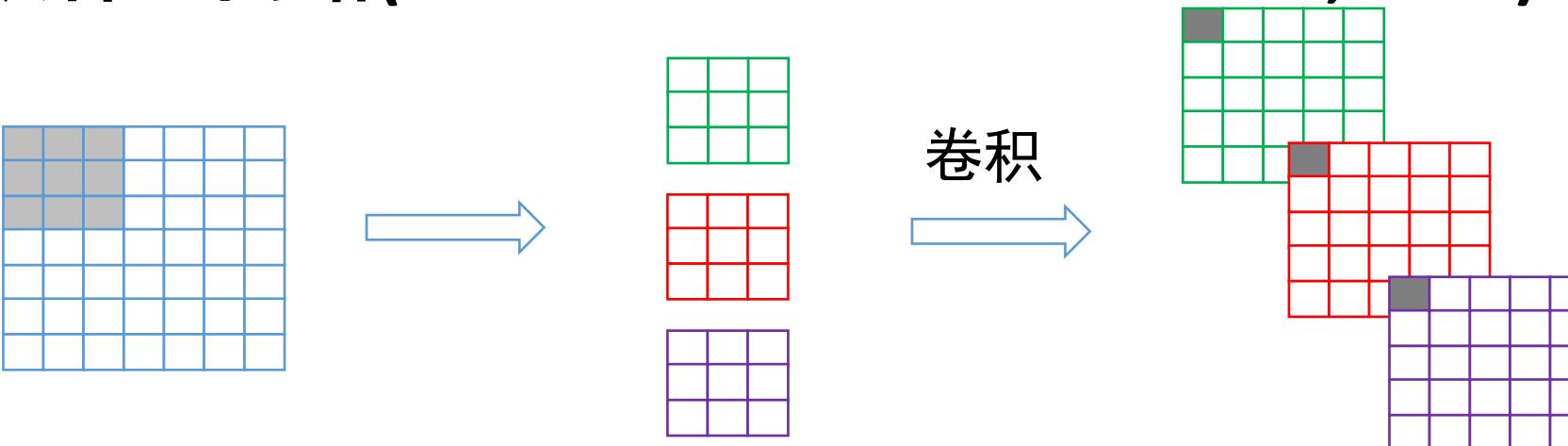
- 地点：B402，实验楼A304
- 时间：1-17周讲授，3-17周实训
- 教材：《人工智能导论：模型与方法》，高等教育出版社
- 作业：课上随堂测试考核(10%)、**研讨内容(10%)**、实验内容考核(40%)和期末考试(40%)
- 期末考试：30道选择题（每小题2分）4道简答题（每小题5分）
2道解答题（每小题10分）

从前馈神经网络到卷积神经网络

- 三层前馈神经网络

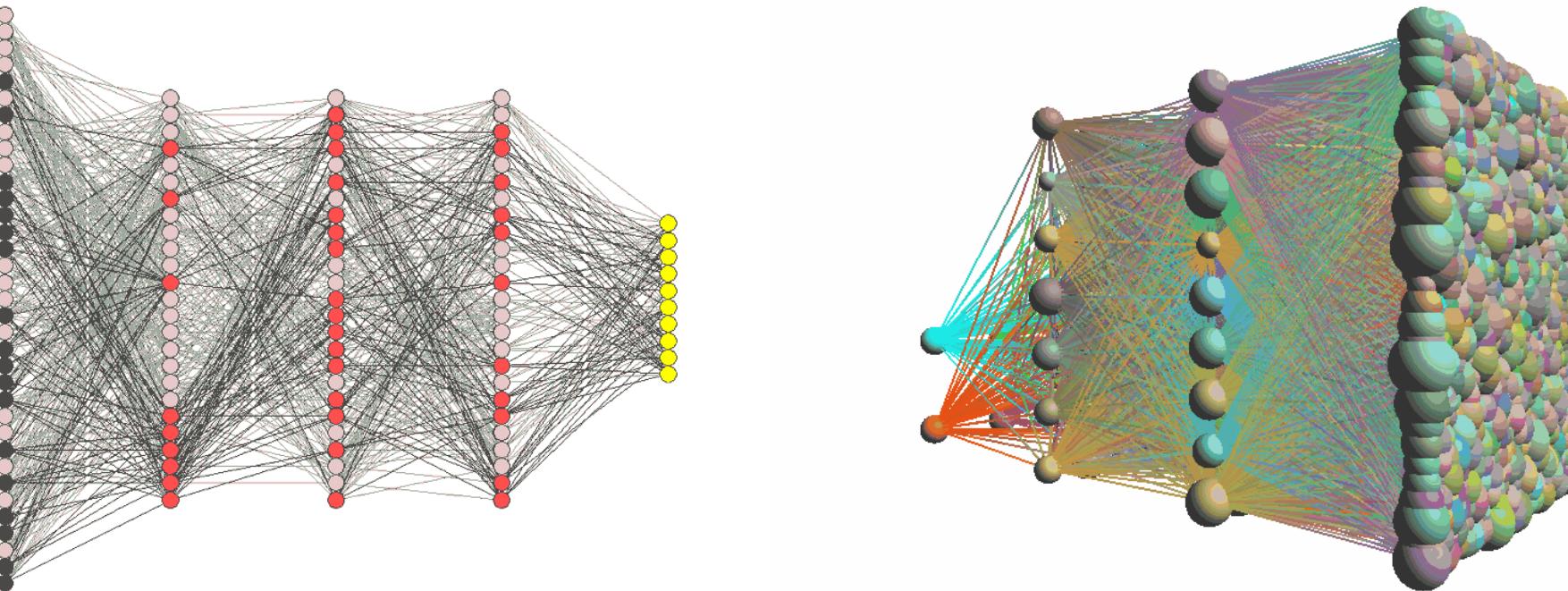


- 卷积神经网络(convolution neural network, CNN)



卷积神经网络 (CNN)

- 前馈神经网络中，输入层的输入直接与第一个隐藏层中所有神经元相互连接。



卷积神经网络 (CNN)

- 前馈神经网络中，输入层的输入直接与第一个隐藏层中所有神经元相互连接。
 - 如果输入数据是一幅 $1000*1000$ 的图像，即1,000,000维的向量，考虑输入数据与第一个隐藏层（与输入维度一致）中所有神经元均相连，则输入层到第一个隐藏层之间有 10^{12} 个参数。

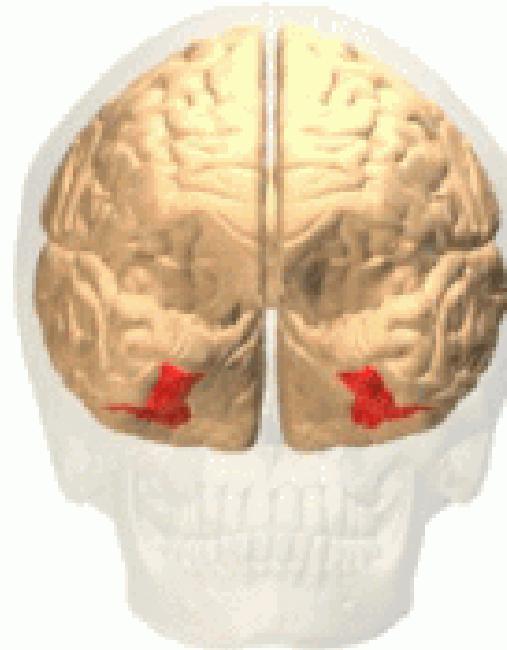
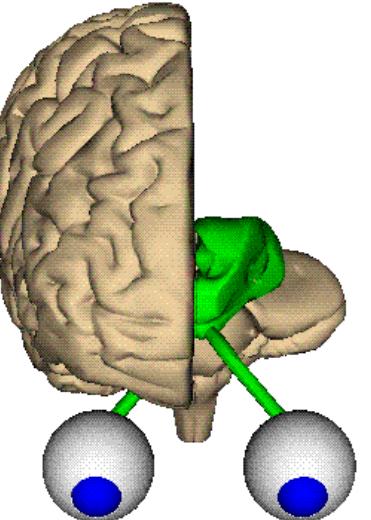
卷积神经网络 (CNN)

- 前馈神经网络中，输入层的输入直接与第一个隐藏层中所有神经元相互连接。
 - 如果输入数据是一幅 $1000*1000$ 的图像，即1,000,000维的向量，考虑输入数据与第一个隐藏层（与输入维度一致）中所有神经元均相连，则输入层到第一个隐藏层之间有 10^{12} 个参数。
 - 模型参数数量如此巨大不仅会占用大量计算机内存，同时也使神经网络模型变得难以训练收敛。因此，对于图像这样的数据，不能直接将所构成的像素点向量与前馈神经网络神经元相连。

卷积神经网络 (CNN)

- 人脑“视觉系统的信息处理”这一机制

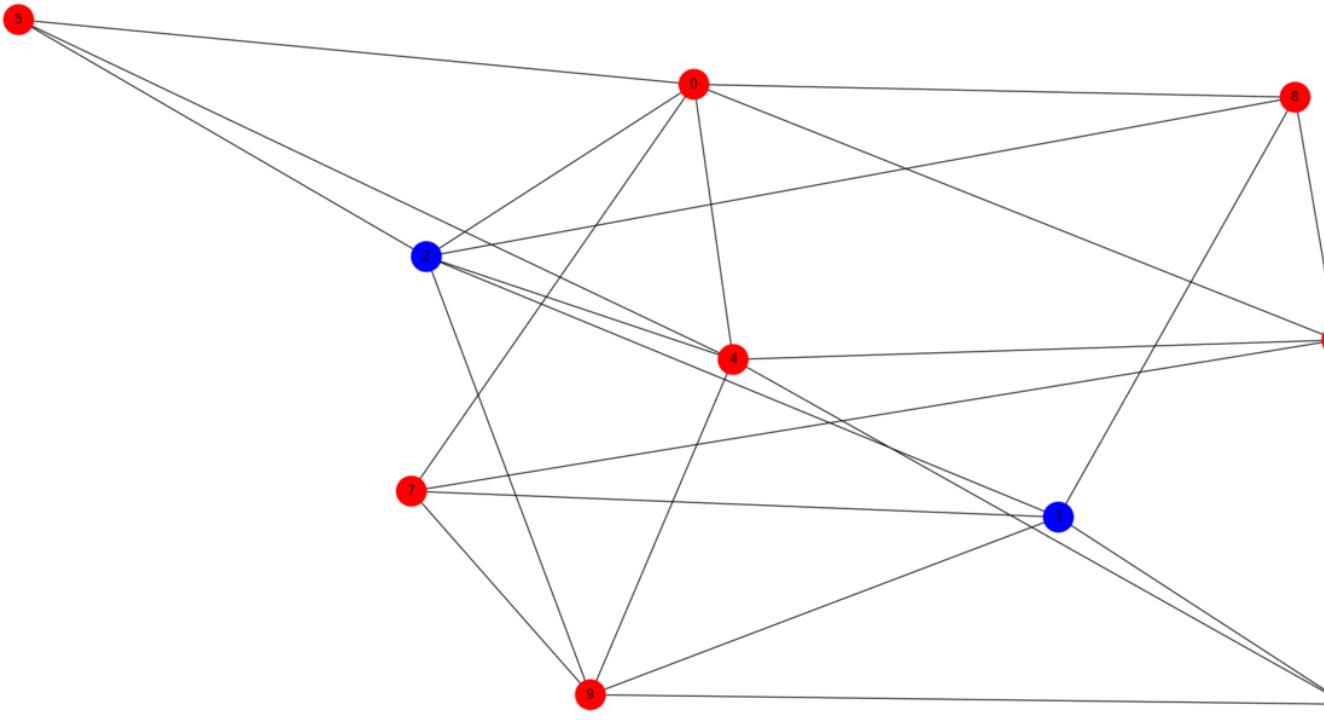
- 即可视皮层对外界信息是分级感受的
- 1959年，David Hubel和Torsten Wiesel发现



卷积神经网络 (CNN)

- 卷积神经网络的前身：满足平移不变性的网络Neocognitron

- 1980年Kunihiko Fukushima，将神经科学所发现的结构进行了计算机模拟，提出级联方式（cascade，即逐层滤波）



卷积神经网络 (CNN)

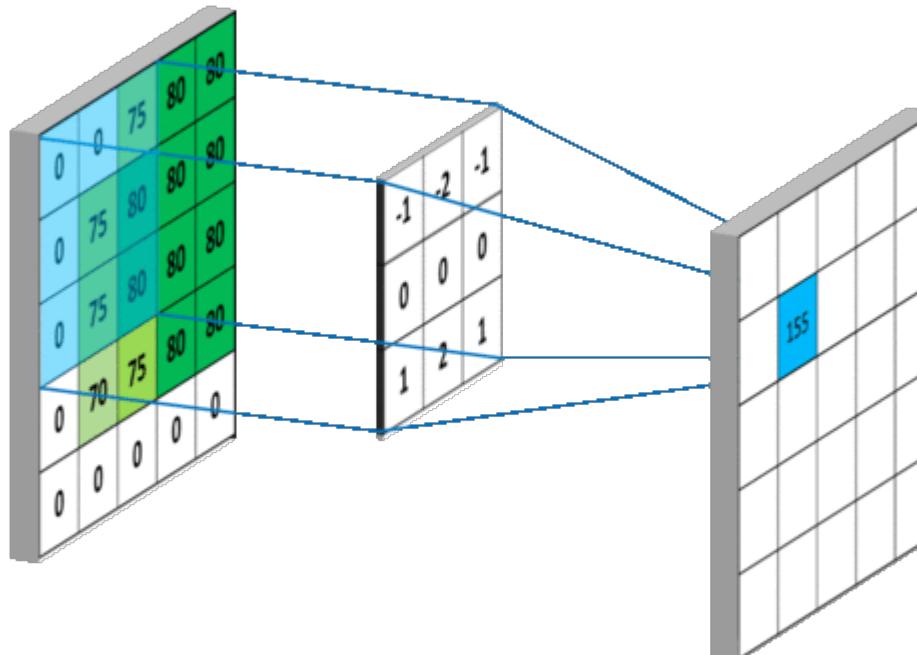
- 图像中像素点具有很强的空间依赖性
 - 卷积就是针对像素的空间依赖性来对图像进行处理的一种技术。

11, 11
12, ,11



卷积神经网络 (CNN)

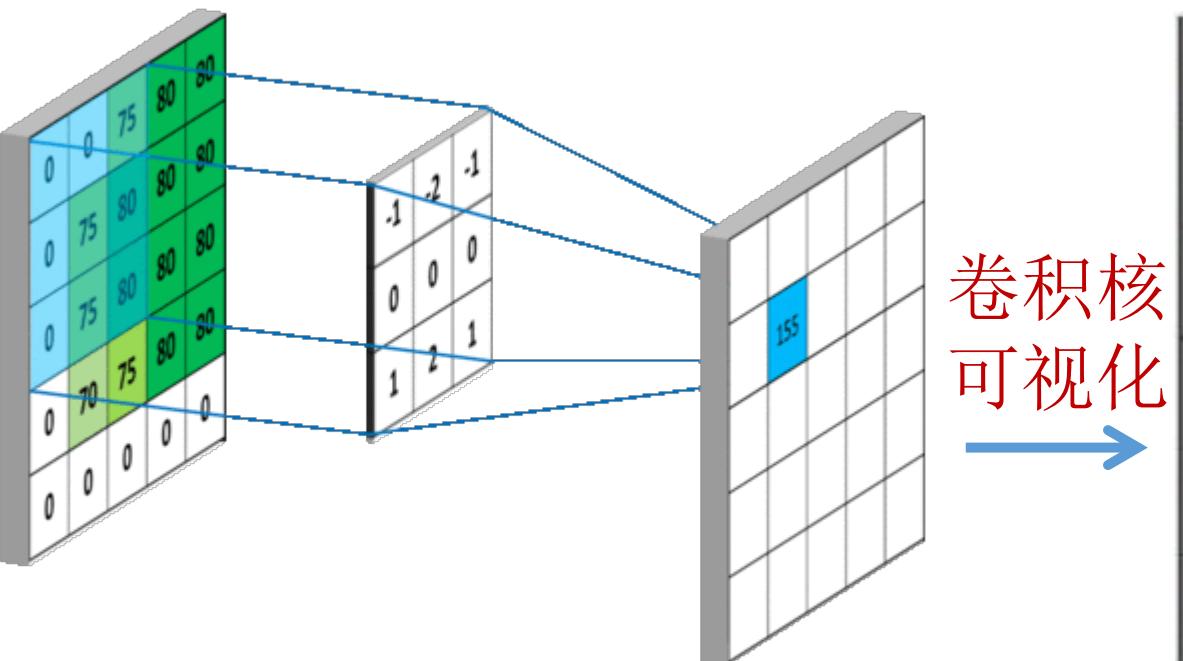
- 在图像卷积计算中，需要定义一个卷积核（kernel）
 - 卷积核是一个二维矩阵，矩阵中数值为对图像中与卷积核同样大小的子块像素点进行卷积计算时所采用的权重。



卷积核

卷积神经网络 (CNN)

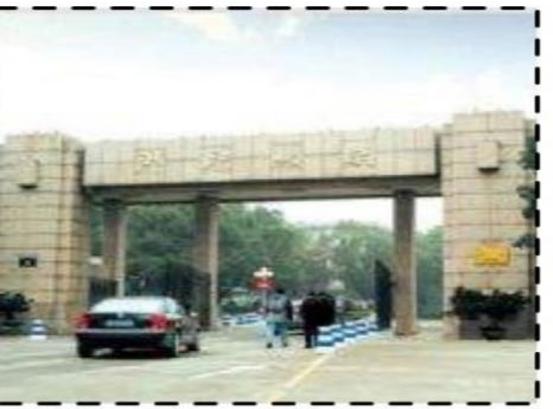
- 卷积核中的权重系数 w_i 是通过数据驱动机制学习得到
 - 用来捕获某像素点及其邻域像素点所构成的特有空间模式。
 - 一旦从数据中学习得到权重系数，这些权重系数就刻画了图像中像素点构成的空间分布不同模式。



卷积神经网络：卷积操作

- 图像经过特定卷积矩阵滤波后，所得到的卷积结果可认为是保留了像素点所构成的特定空间分布模式

卷积神经网络：卷积操作示意



转变为灰度图



$$(77 \times 1 + 75 \times 2 + 76 \times 1 + 74 \times 2 + 65 \times 4 + 77 \times 2 + 75 \times 1 + 75 \times 2 + 74 \times 1) / (1 + 2 + 1 + 2 + 4 + 2 + 1 + 2 + 1) = 68.125 = 68$$

77	75	76
74	65	77
75	75	74

灰度图中的一个 3×3 子块

\times

1	2	1
2	4	2
1	2	1

3×3 高斯滤波器

得到

68

$$(77 \times 1 + 75 \times 2 + 76 \times 1 + 74 \times 1 + 65 \times 4 + 77 \times 2 + 75 \times 1 + 75 \times 2 + 74 \times 1) / (1 + 2 + 1 + 2 + 4 + 2 + 1 + 2 + 1) = 68.125 = 68$$



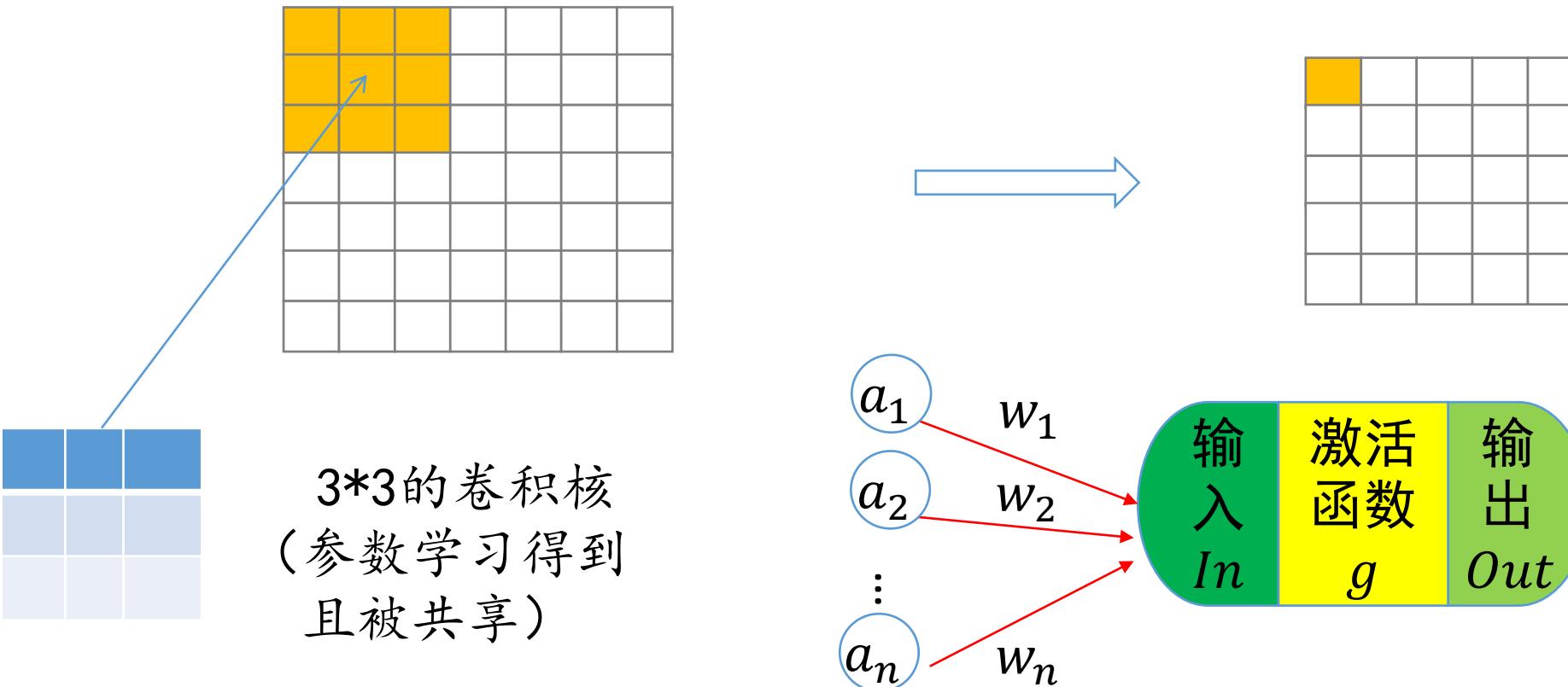
使用了 5×5 高斯模版卷积后的结果图



使用了 3×3 高斯模版卷积后的结果图

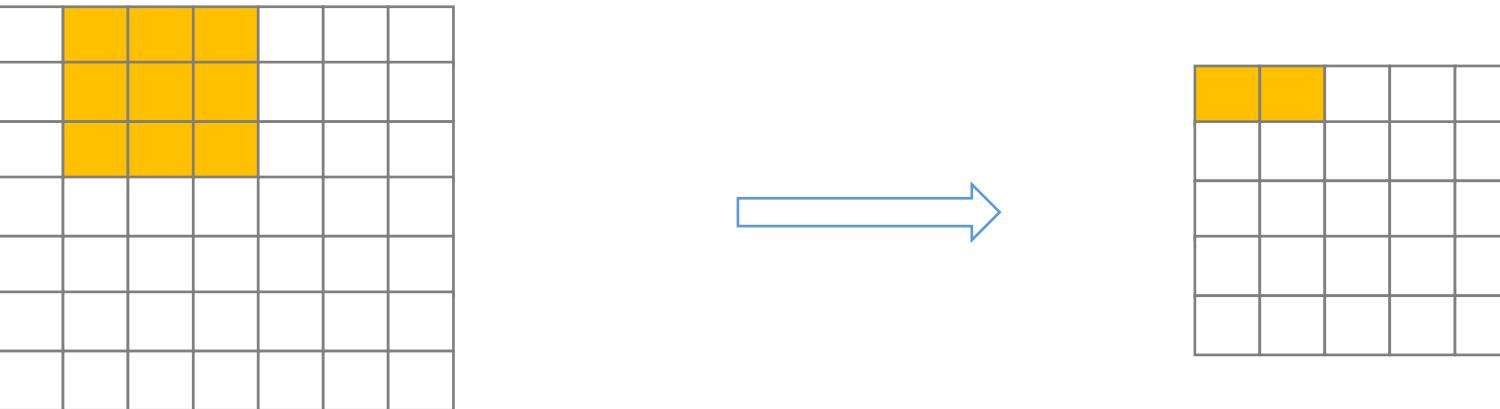
卷积神经网络：卷积操作

7*7大小的图像，通过3*3大小卷积矩阵，以1的步长进行卷积操作，可得到5*5大小的卷积结果



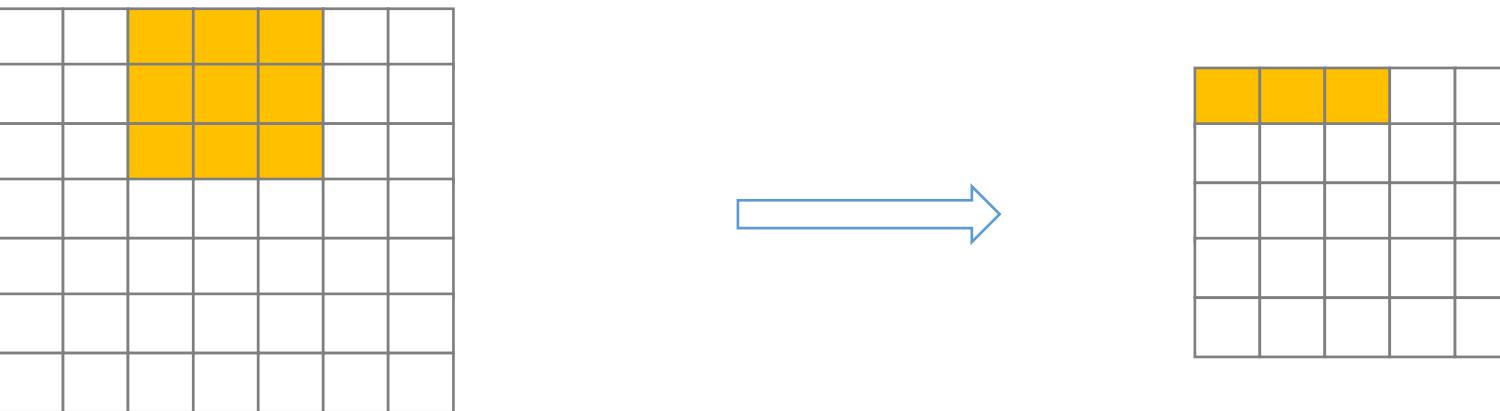
卷积神经网络：卷积操作

7*7大小的图像，通过3*3大小卷积矩阵以1的步长
进行卷积操作，可得到5*5大小的卷积结果



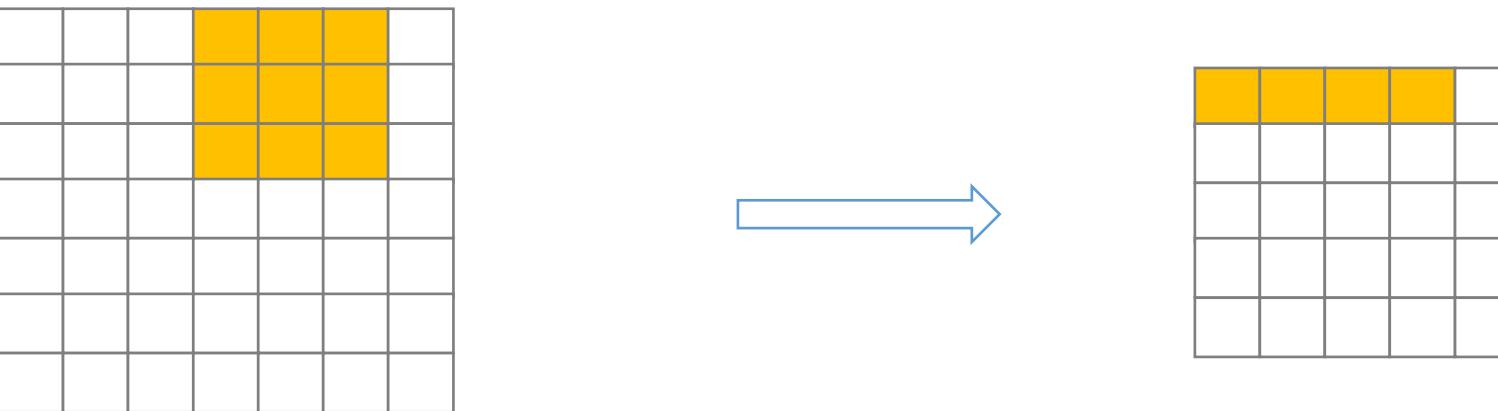
卷积神经网络：卷积操作

7*7大小的图像，通过3*3大小卷积矩阵以1的步长
进行卷积操作，可得到5*5大小的卷积结果



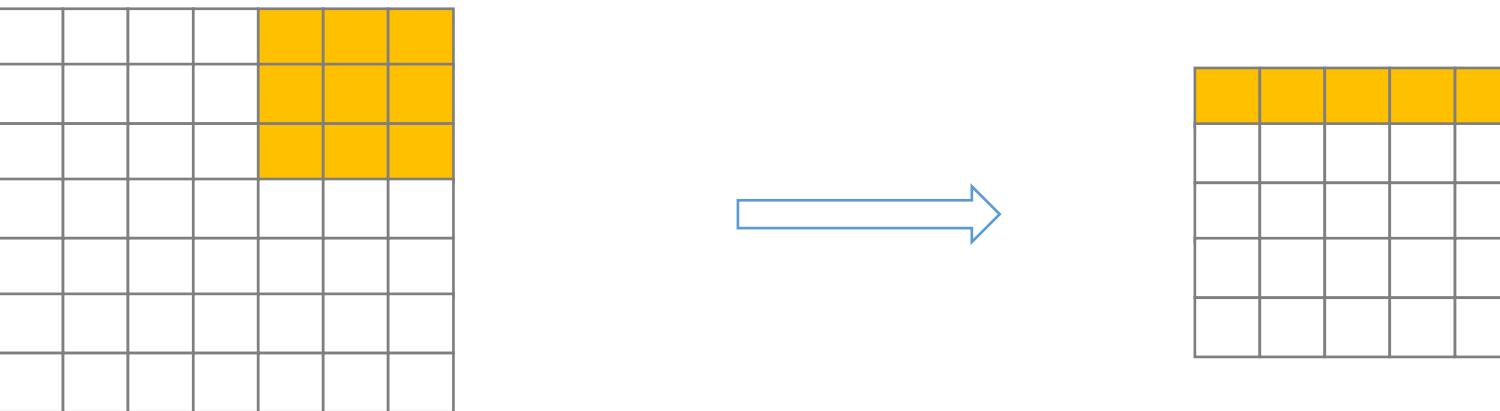
卷积神经网络：卷积操作

7*7大小的图像，通过3*3大小卷积矩阵以1的步长
进行卷积操作，可得到5*5大小的卷积结果



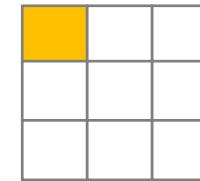
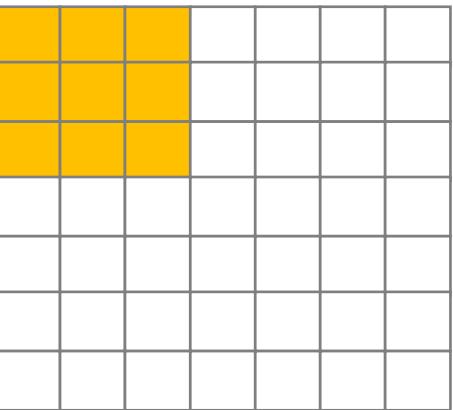
卷积神经网络：卷积操作

7*7大小的图像，通过3*3大小卷积矩阵以1的步长
进行卷积操作，可得到5*5大小的卷积结果



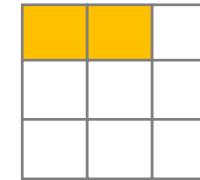
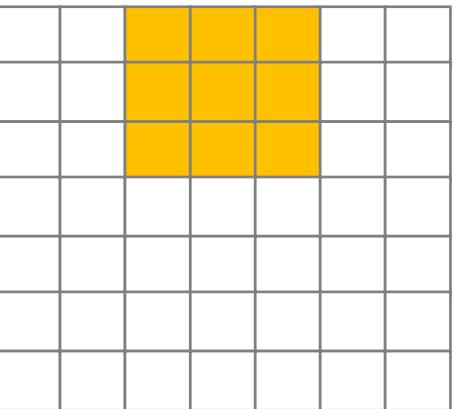
卷积神经网络：卷积操作

如果步长改为2



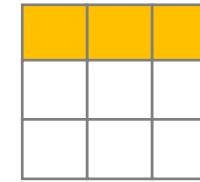
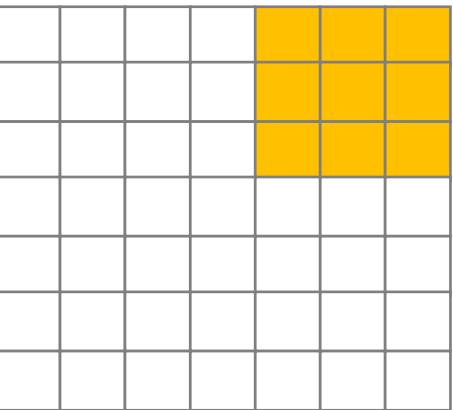
卷积神经网络：卷积操作

如果步长改为2



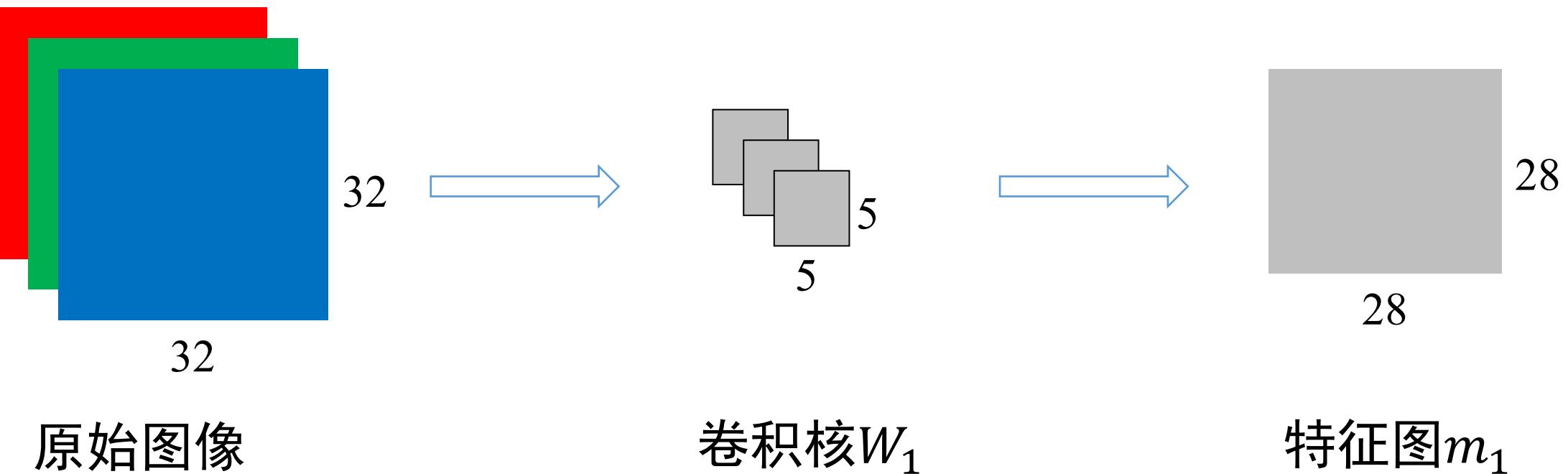
卷积神经网络：卷积操作

如果步长改为2



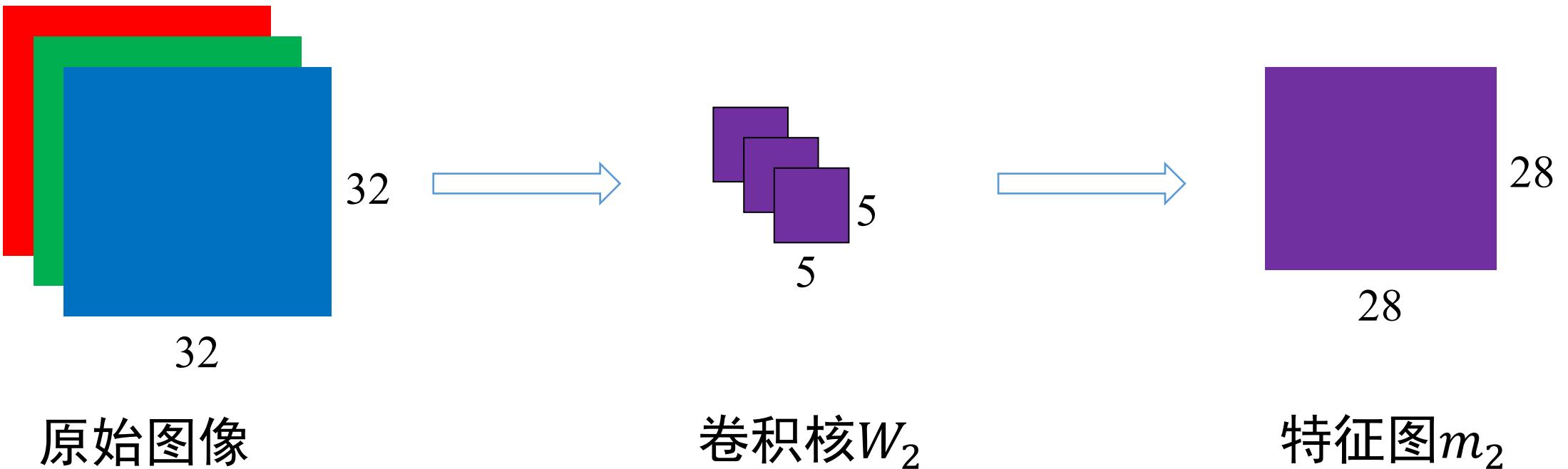
卷积神经网络：卷积操作

有一张 $32*32*3$ (RGB)的图像，使用 $5*5*3$ 的卷积核 W_1 ，步长为1对其进行卷积操作。卷积核 W_1 在原始图像上从左到右、从上到下进行计算，改变 $5*5$ 子块区域中的中心像素点值，得到 $28*28$ 的特征图 m_1



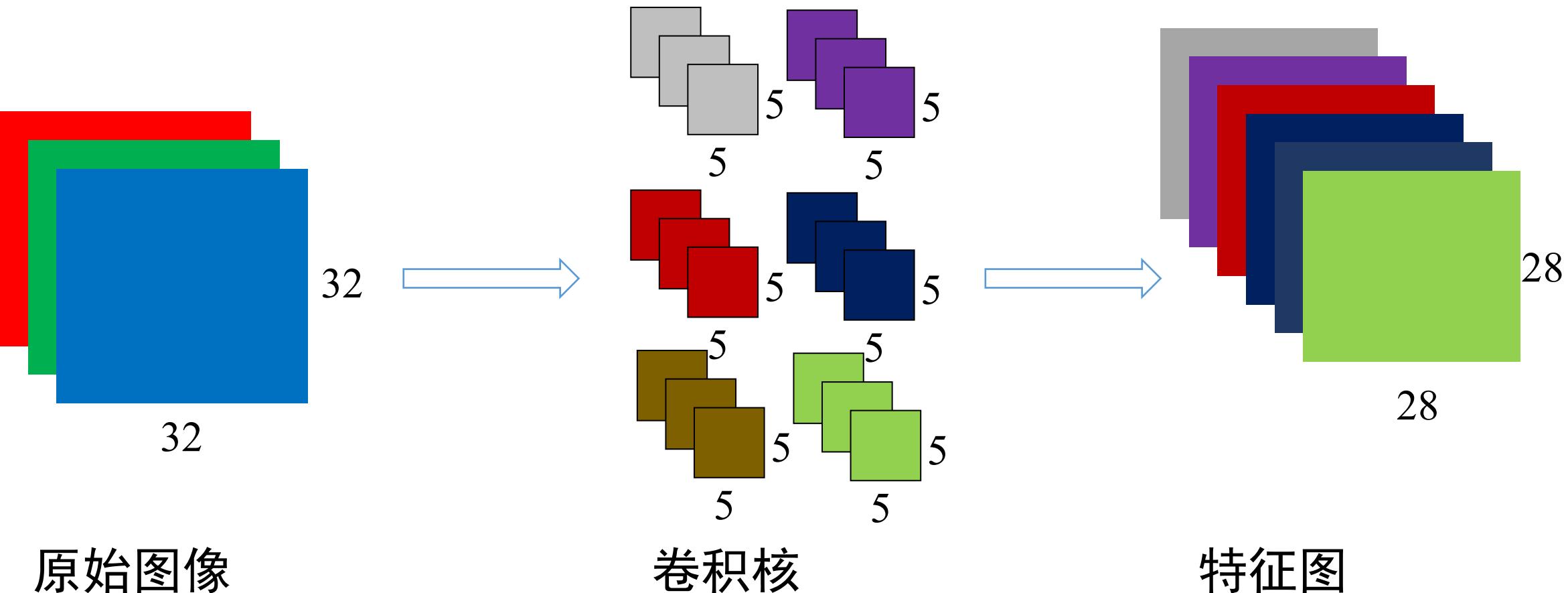
卷积神经网络：卷积操作

使用另一个 $5*5*3$ 的卷积核 W_2 与原始图像做卷积操作，得到特征图 m_2



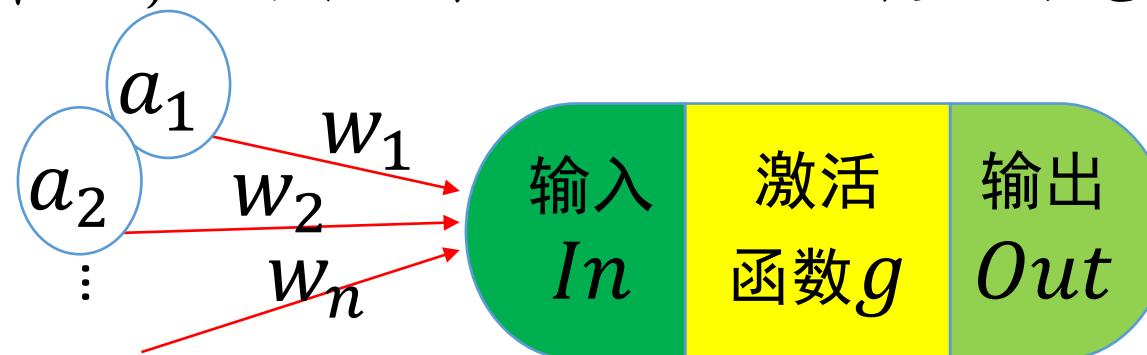
卷积神经网络：卷积操作

- 使用6个 $5 \times 5 \times 3$ 的卷积核与原始图像做卷积操作，则得到6个 28×28 的特征图
- 注意：6个 $5 \times 5 \times 3$ 的卷积核均是数据驱动学习得到，其刻画了不同的视觉模式



卷积神经网络：卷积+激活函数(非线性映射)

在对原始图像做卷积操作后，可使用ReLU激活函数对卷积后结果进行处理



卷积神经网络：卷积操作

- 神经科学家发现，人的视觉神经细胞对不同的视觉模式具有特征选择性（Feature Selectivity）

卷积神经网络：池化(pooling)操作

- 对输入的特征图进行下采样，以获得最主要信息

卷积神经网络：最大池化操作

在输入特征图中每一个区域寻找最大值。

1	5	4	3
2	6	4	0
3	1	7	9
2	8	6	5

对2*2大小区域，按照步长为2进行最大池化操作



6	4
8	9

卷积神经网络：平均池化操作

对输入特征图的每一个区域的值求平均值（取整）

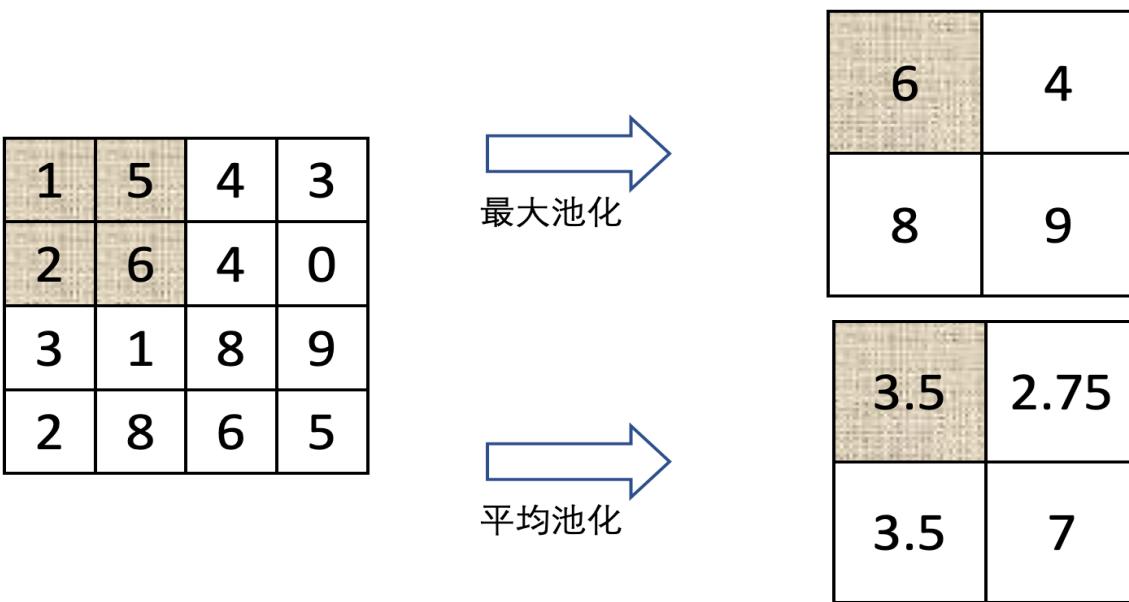
1	5	4	3
2	6	4	0
3	1	7	9
2	8	6	5

对2*2大小区域，按照步长为2进行平均池化操作



3	2
3	6

卷积神经网络：池化操作



由于图像中存在较多冗余，在图像处理中，可用某一区域子块的统计信息（如最大值或均值等）来刻画该区域中所有像素点呈现的空间分布模式，以替代区域子块中所有像素点取值，这就是卷积神经网络中池化(pooling)操作。池化操作对卷积结果特征图进行约减，实现了下采样，同时保留了特征图中主要信息。

卷积神经网络：池化操作

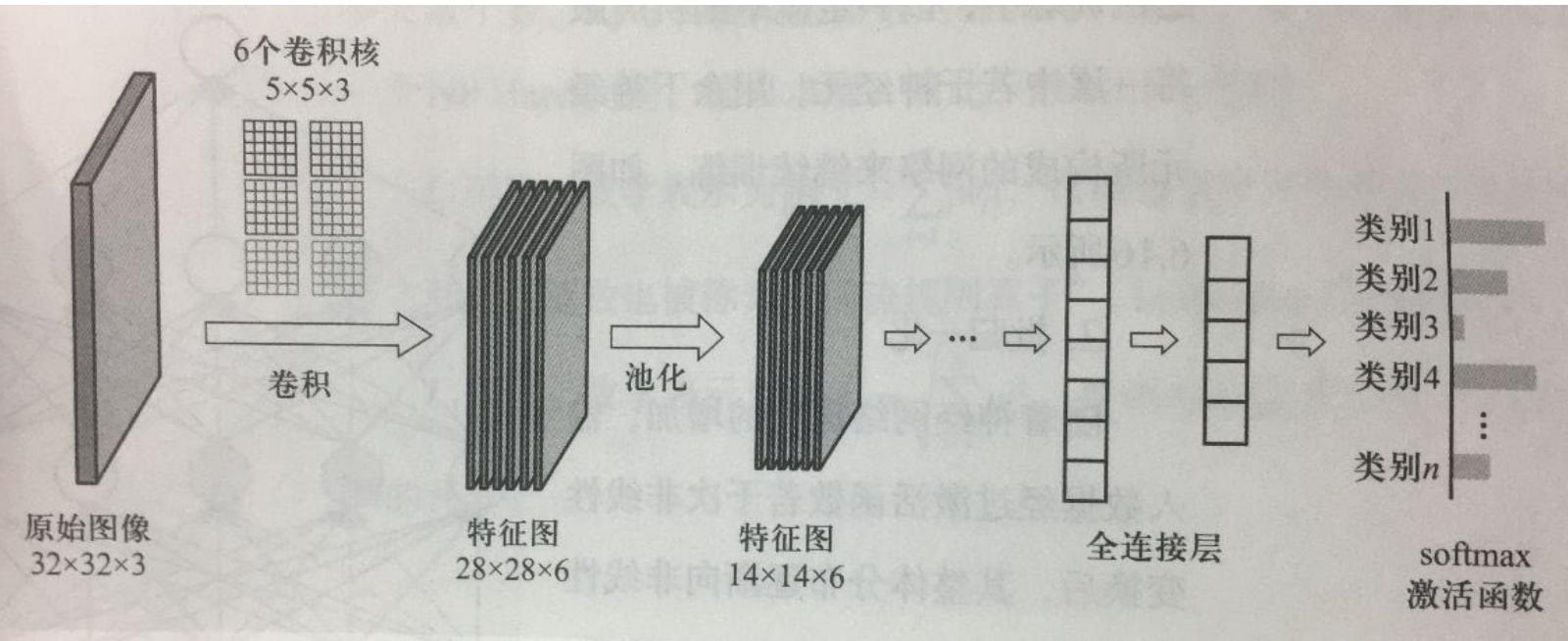


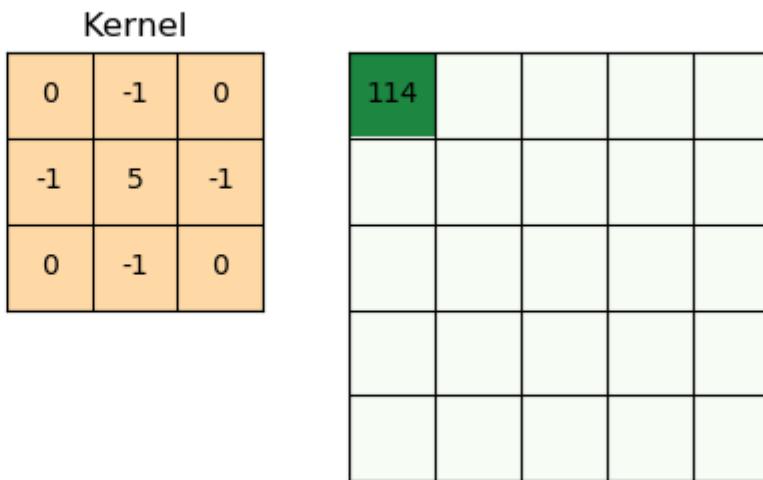
图6.15 基于卷积神经网络的图像分类示意图

对于输入的海量标注数据，通过多次迭代训练，卷积神经网络在若干次卷积操作、接着对卷积所得结果进行激活函数操作和池化操作下，最后通过全连接层来学习得到输入数据的特征表达，即**分布式向量表达(distributed vector representation)**。

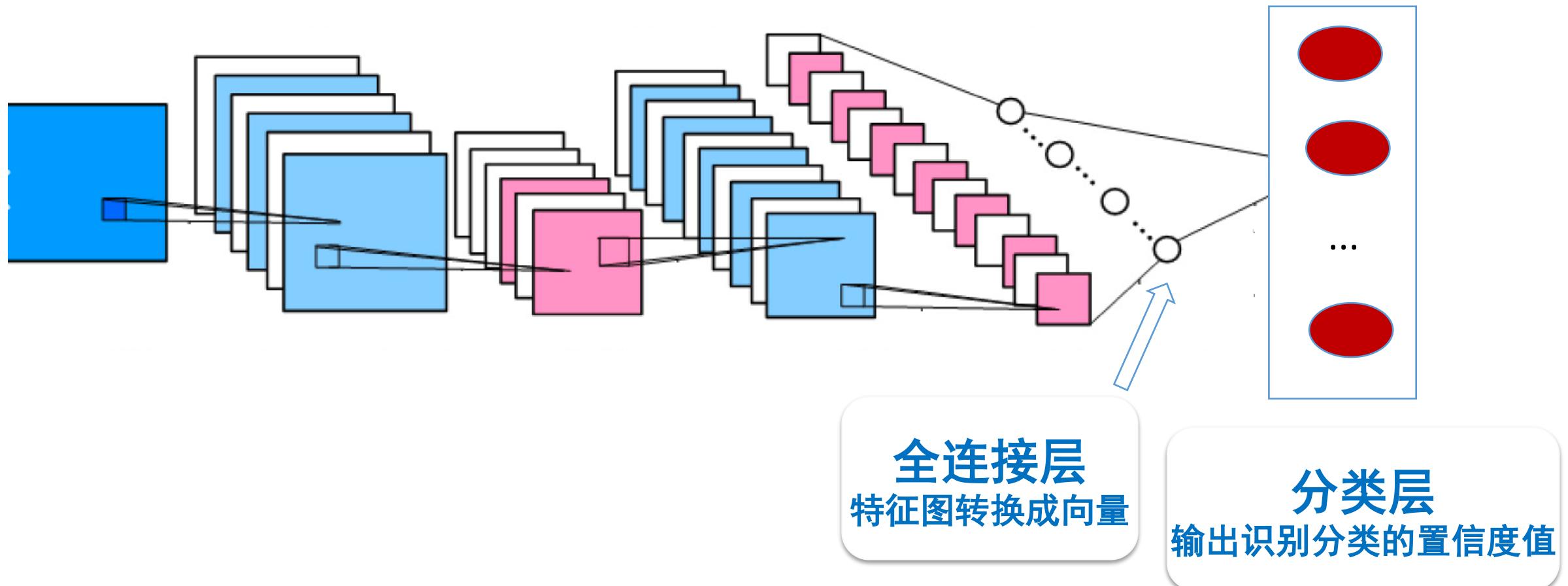
卷积神经网络：padding操作

对特征进行填充，一般用0

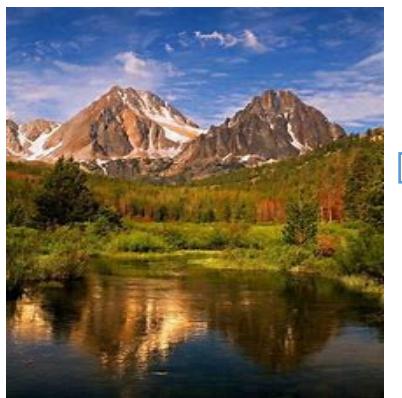
0	0	0	0	0	0	0	0
0	60	113	56	139	85	0	0
0	73	121	54	84	128	0	0
0	131	99	70	129	127	0	0
0	80	57	115	69	134	0	0
0	104	126	123	95	130	0	0
0	0	0	0	0	0	0	0



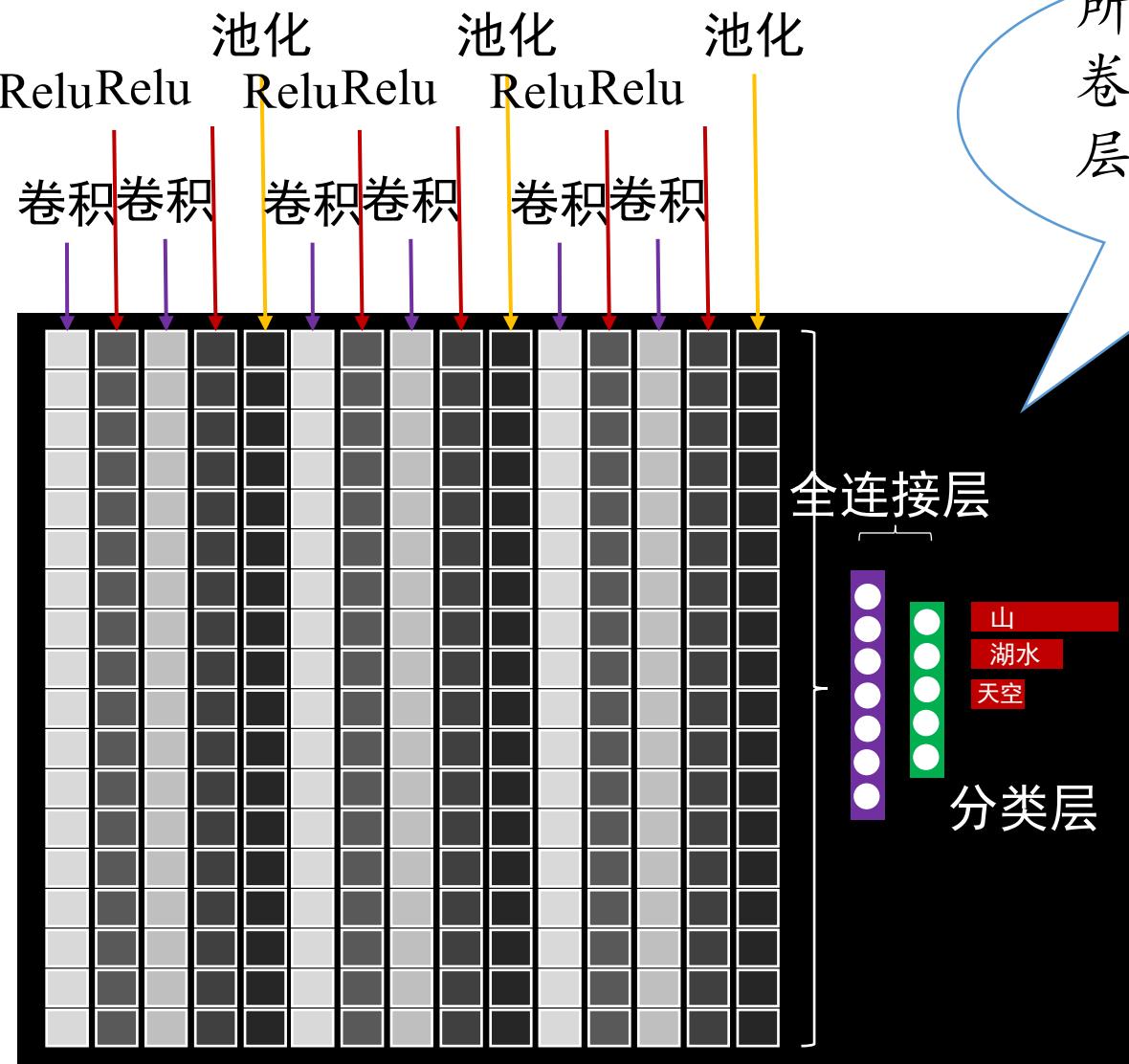
卷积神经网络：全连接层与分类层



卷积神经网络基本架构

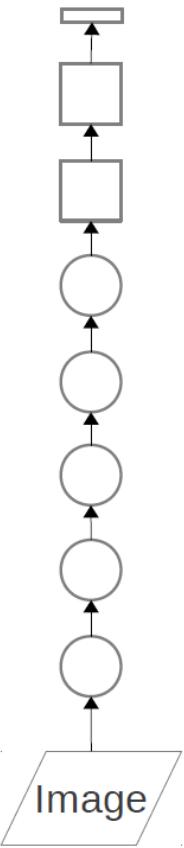


输入
→



所需学习参数：
卷积核、全连接
层权重、激活函
数参数

卷积神经网络的参数



- Trained with stochastic gradient descent on two NVIDIA GPUs for about a week
- 650,000 neurons
- 60,000,000 parameters
- 630,000,000 connections
- **Final feature layer:** 4096-dimensional

体现了数据、模型和算力的结合



Convolutional layer: convolves its input with a bank of 3D filters, then applies point-wise non-linearity



Fully-connected layer: applies linear filters to its input, then applies point-wise non-linearity

Alex Krizhevsky, et.al., ImageNet Classification with Deep Convolutional Neural Networks

神经网络正则化

为了缓解神经网络在训练过程中出现的过拟合问题，需要采取一些正则化技术来提升神经网络的泛化能力 (generalization)

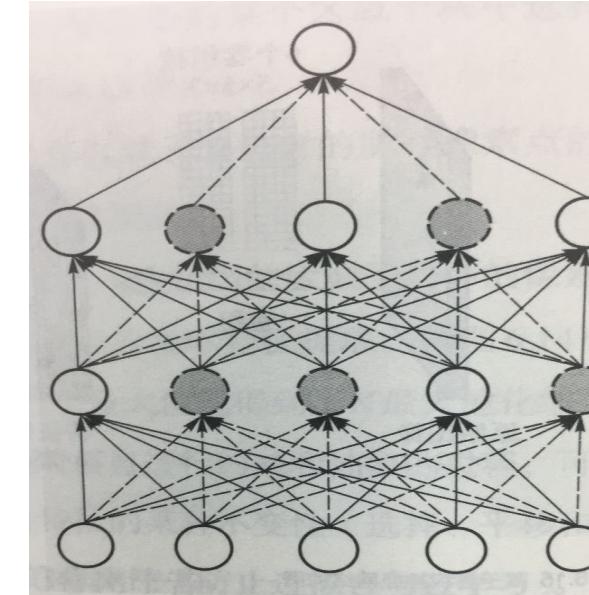
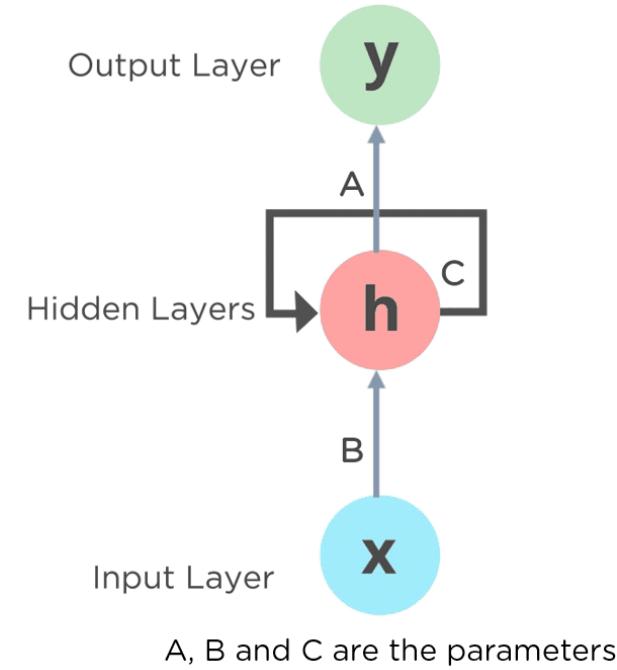


图6.16 使用Dropout的神经网络模型

提纲

- 深度学习历史发展
- 前馈神经网络
- 卷积神经网络
- 循环神经网络
- 深度生成学习
- 深度学习应用

循环神经网络 (RNN)



循环神经网络 (RNN)

- 循环神经网络 (Recurrent Neural Network, RNN) 是一类处理序列数据 (如文本句子、视频帧等) 时所采用的网络结构。先前所介绍的前馈神经网络或卷积神经网络所需要处理的输入数据一次性给定，难以处理存在前后依赖关系的数据。

循环神经网络 (RNN)

- 循环神经网络在处理数据过程中构成了一个循环体
 - 对于序列数据，在 t 时刻循环神经网络单元会读取当前输入数据 x_t 和前一时刻输入数据 x_{t-1} 所对应的隐式编码结果 h_{t-1} ，一起生成 t 时刻的隐式编码结果 h_t 。接着将 h_t 后传。

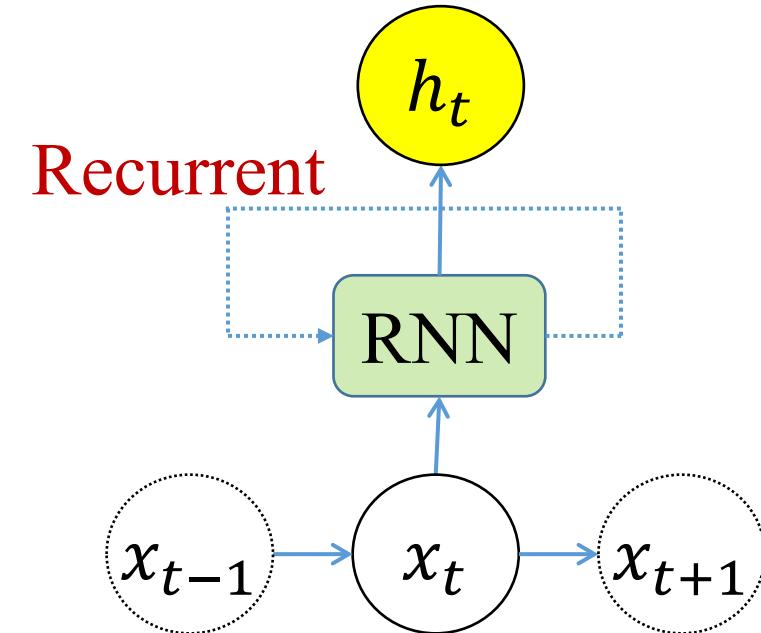
循环神经网络 (RNN)

- 在处理数据过程中构成了一个循环体

- 时刻 t 所得到的隐式编码 h_t 是由上一时刻隐式编码 h_{t-1} 和当前输入 x_t 共同参与生成的。

隐式编码 h_{t-1} 已经“记忆”了 t 时刻之前的时序信息。或者说前序时刻信息影响了后续时刻信息的处理。

- 与前馈神经网络和卷积神经网络在外理时



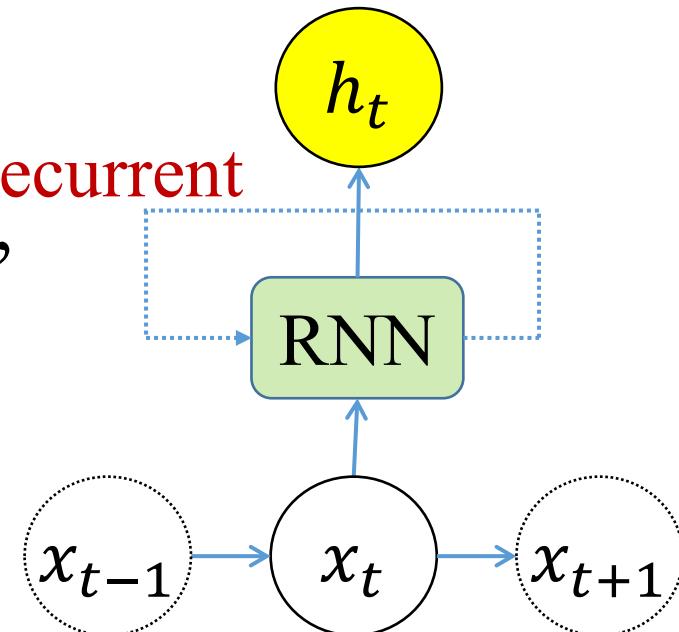
循环神经网络 (RNN)

- 在处理数据过程中构成了一个循环体

- 对当前输入数据 x_t , RNN会结合 $t - 1$ 时刻的隐式编码 h_{t-1} , 产生当前隐式编码 h_t :

$$h_t = \Phi(U \times x_t + W \times h_{t-1})$$

- 这里 $\Phi(\cdot)$ 是激活函数, 可为Sigmoid或者Tanh, 使模型能够忘掉无关的信息, 同时更新记忆内容。U与W为模型参数。
- 当前时刻的隐式编码输出 h_t 不仅仅与当前输入数据 x_t 相关, 与网络已有的“记忆” h_{t-1} 也有着密不可分的联系。



循环神经网络 (RNN)

- 为了更加直观展示在循环神经网络中前序时刻信息被“记住”来影响当前时刻信息编码，这里对上述公式进行如下变化：

$$h_t = \Phi(U \times x_t + W \times h_{t-1}) = \Phi(U \times x_t + W \times \Phi(U \times x_{t-1} + W \times h_{t-2}))$$

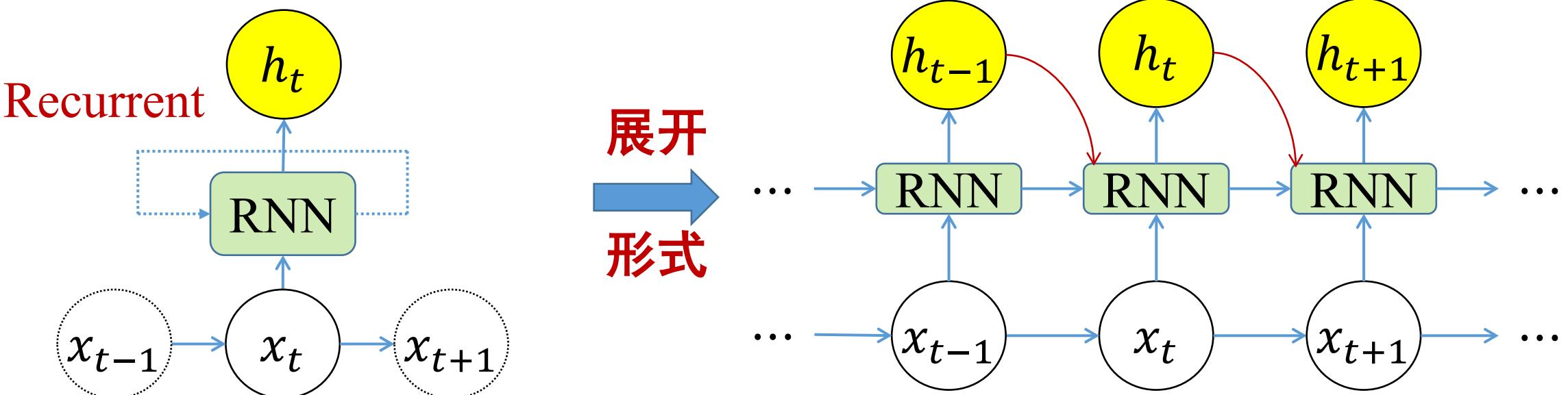
$$= \Phi\left(U \times \underbrace{x_t}_{t \text{时刻输入}} + W \times \Phi\left(U \times \underbrace{x_{t-1}}_{t-1 \text{时刻输入}} + W \times \Phi(U \times \underbrace{x_{t-2}}_{t-2 \text{时刻输入}} + \dots)\right)\right)$$

- 当前时刻编码中，的确均包含了历史信息，这也说明了循环神经网络能够记忆的原因，使得一些需要记忆历史过往信息的任务能够被循环神经网络有效处理。

循环神经网络 (RNN)

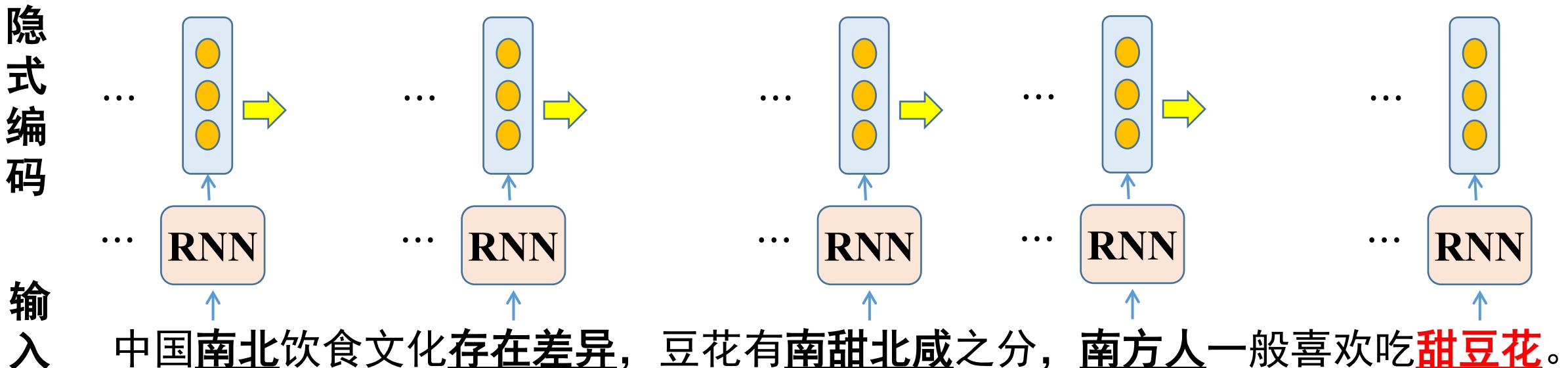
• 将RNN展开，可以得到和前馈神经网络相似的网络结构。

- 这个网络结构可利用反向传播算法和梯度下降算法来训练参数，称为“沿时间反向传播算法(backpropagation through time, BPTT)”。
- 由于RNN每个时刻都有一个输出，所以在计算循环神经网络的损失时，通常需要将所有时刻上的损失进行累加。



循环神经网络 (RNN)

- 单词之间存在前后依赖，这种依赖会被有效利用起来，得到各自单词的隐式编码以及句子的向量表达。



提纲

- 深度学习历史发展
- 前馈神经网络
- 卷积神经网络
- 循环神经网络
- 深度生成学习
- 深度学习应用

深度生成学习模型 (deep generative learning model)

在本章之前的介绍中，神经网络模型从数据中提取出高层语义在数据中所蕴含的“模式”，并利用这些模式实现对数据的分类和检测等，这种模型通常称为判别模型，判别模型不关心数据如何生成，它只关心数据蕴含哪些模式以及如何将数据进行分类。与之相对的模型类型被称为生成模型(generative model)，生成模型需要学习目标数据的分布规律，以合成属于目标数据空间的新数据。

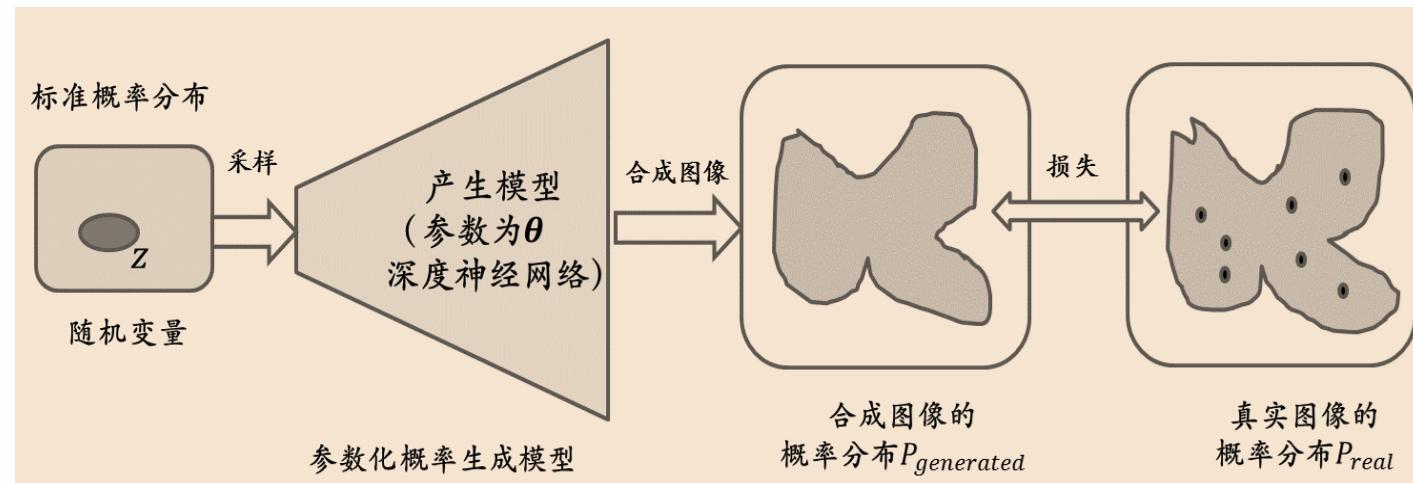
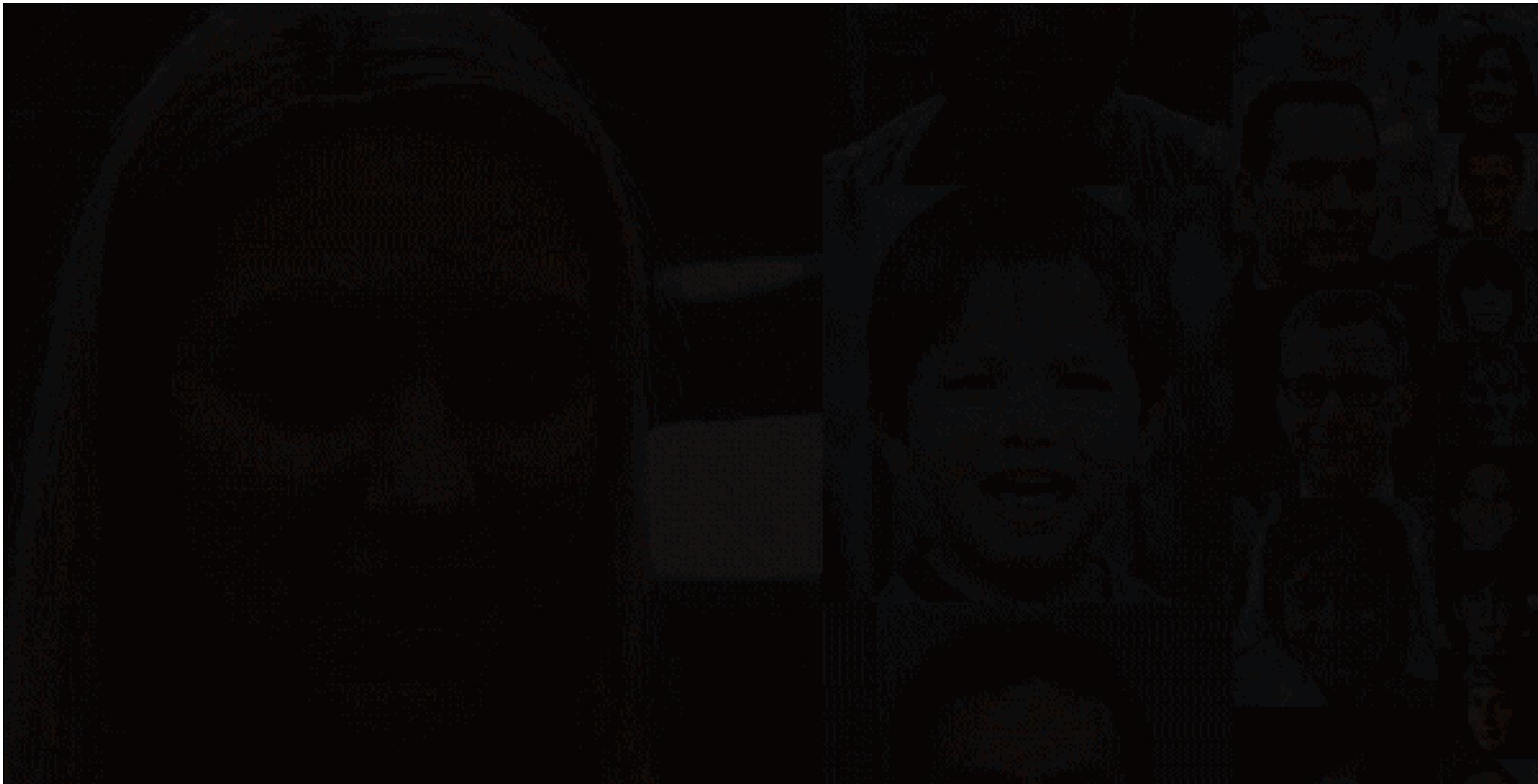


图6.23 生成模型的示意图

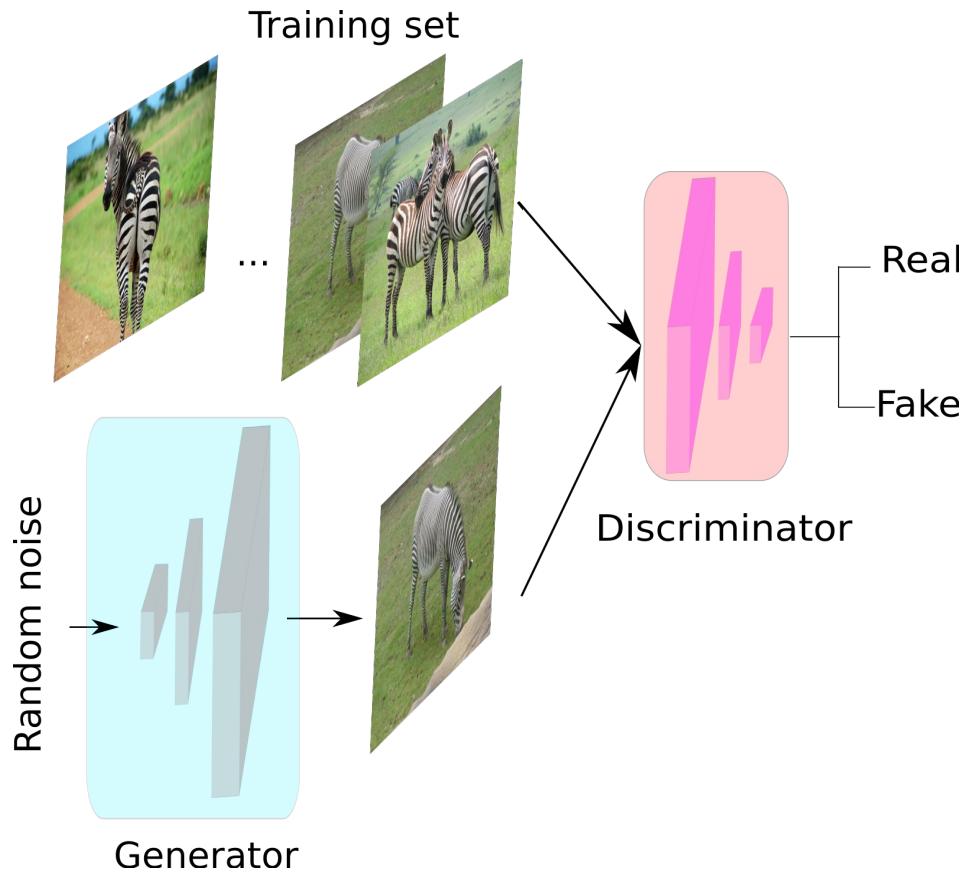
变分自编码器(variational auto-encoder, VAE)、自回归模型(Autoregressive models)与生成对抗网络 (generative adversarial network, GAN) 等.....

生成对抗网络 (deep generative learning model)



图片来自StyleGAN

生成对抗网络 (deep generative learning model)



生成对抗网络由一个生成器 (generator, 简称G) 和一个判别器 (discriminator, 简称D) 组成。

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

生成对抗网络 (deep generative learning model)

算法6.5.1 训练生成对抗网络

输入：神经网络 D 、 G ，噪声分布 $p_z(z)$ ，真实数据分布 $p_{data}(x)$

输出：神经网络参数 θ_d 、 θ_g

算法步骤：

每轮训练循环执行：

训练 k 轮判别器：

从噪声分布 $p_z(z)$ 中采样 m 个样本 $\{z^{(1)}, z^{(2)}, z^{(3)} \dots z^{(m)}\}$

从真实数据分布 $p_{data}(x)$ 中采样 m 个样本 $\{x^{(1)}, x^{(2)}, x^{(3)} \dots x^{(m)}\}$

沿梯度上升方向更新判别器参数：

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m [\log D(x^{(i)}) + \log (1 - D(G(z^{(i)})))]$$

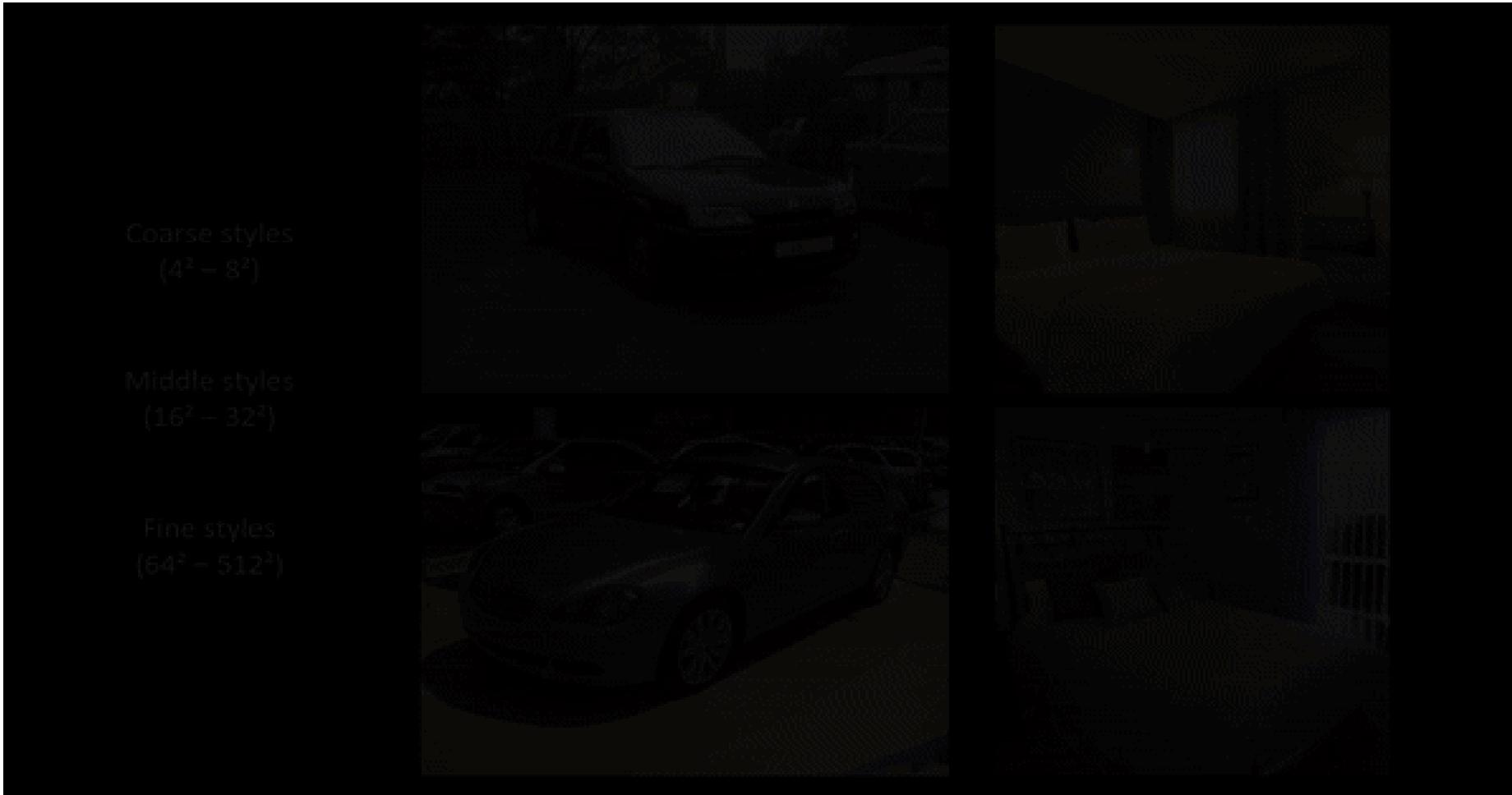
从噪声分布 $p_z(z)$ 中采样 m 个样本 $\{z^{(1)}, z^{(2)}, z^{(3)} \dots z^{(m)}\}$

沿梯度下降方向更新判别器参数：

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^m \log (1 - D(G(z^{(i)})))$$

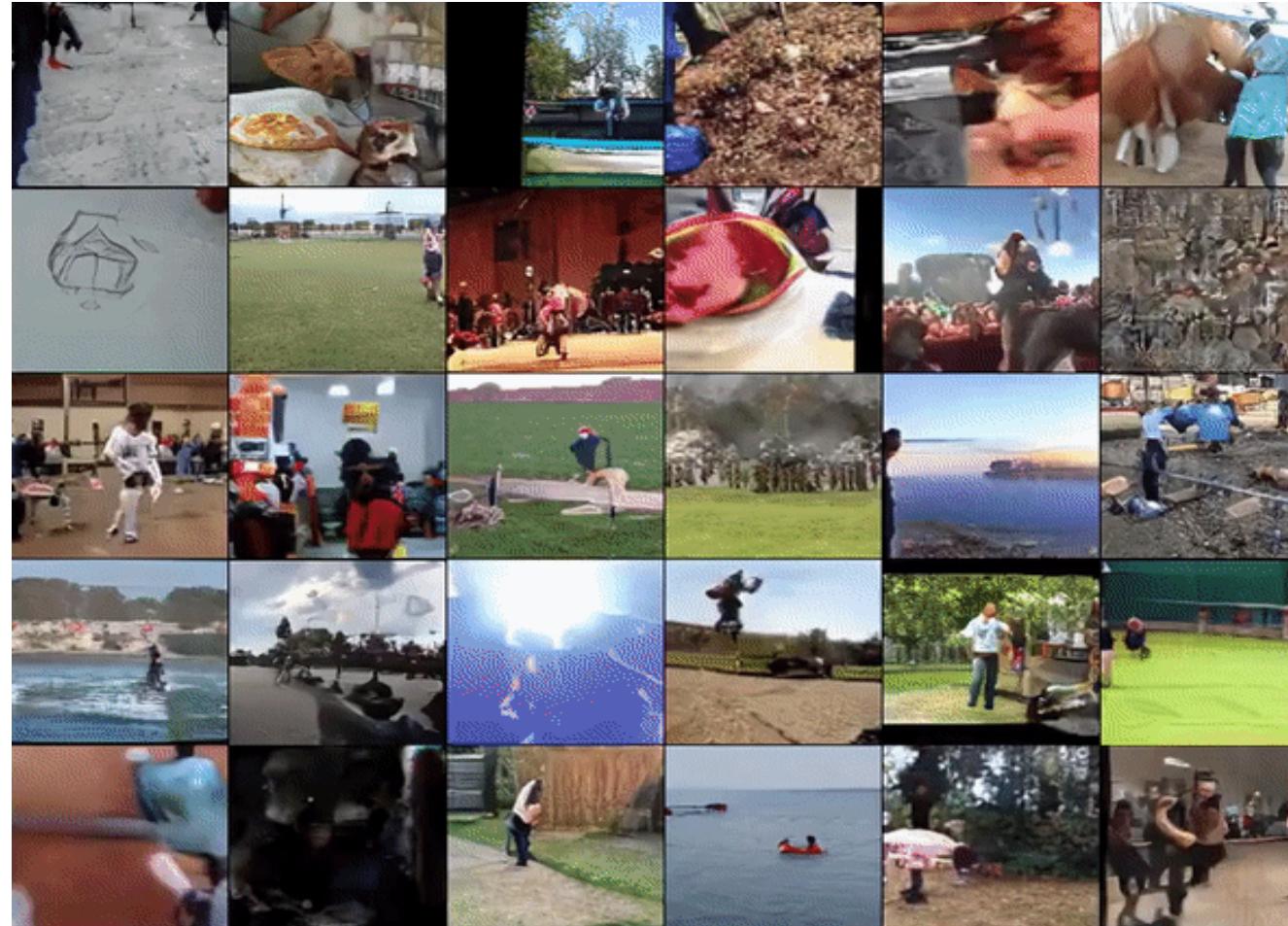
在训练初期，因为生成数据质量较低，判别器 D 可以轻易将其区分，此时 $\log(1 - D(G(z)))$ 无法为生成器 G 提供显著的梯度，因此可以使用 $-\log D(G(z))$ 来代替损失函数。这么做并不会影响其能够收敛至最优解的性质，且可以在训练初期提供足够的梯度。

条件生成对抗网络



条件生成对抗网络

Unconditional GANs: vedio.



Video are from DVDGAN

条件生成对抗网络

Unconditional GANs: Chinese Font.

字種成東字推
符利用亞型斷的
到字條網言進新
字符件絡字行方
一生對體自動法

千萬孤獨
山徑舟釣
鳥蓑寒江
飛笠雪
絕滅翁

Font are from zi2zi

提纲

- 深度学习历史发展
- 前馈神经网络
- 卷积神经网络
- 循环神经网络
- 深度生成学习
- 深度学习应用

研讨内容一

- 选题：图像超分、图像生成、图像分类、目标检测
- 组队：3人
- 提交报告：问题描述、每篇文章核心内容、相互联系等
- 文档工具：基于latex编写
(https://nankai.feishu.cn/file/SovubR3AToVdZoxXSFkc7jUCnrT?from=from_copylink)
- 截止日期：2024年5月31日23: 59
- 组队信息递交时间：2024年4月30日23: 59

研讨内容二

- 自选选题：任何
- 提交报告：问题描述、每篇文章核心内容、相互联系等
- 文档工具：基于latex编写
- 截止日期：2024年5月31日23: 59

图像分类和目标定位

Classification



Car

Localization

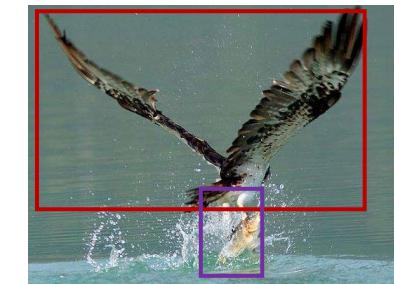


(x, y, w, h)

...

Single Object

Object Detection



Eagle, Fish

...

Multi Objects

图像分类

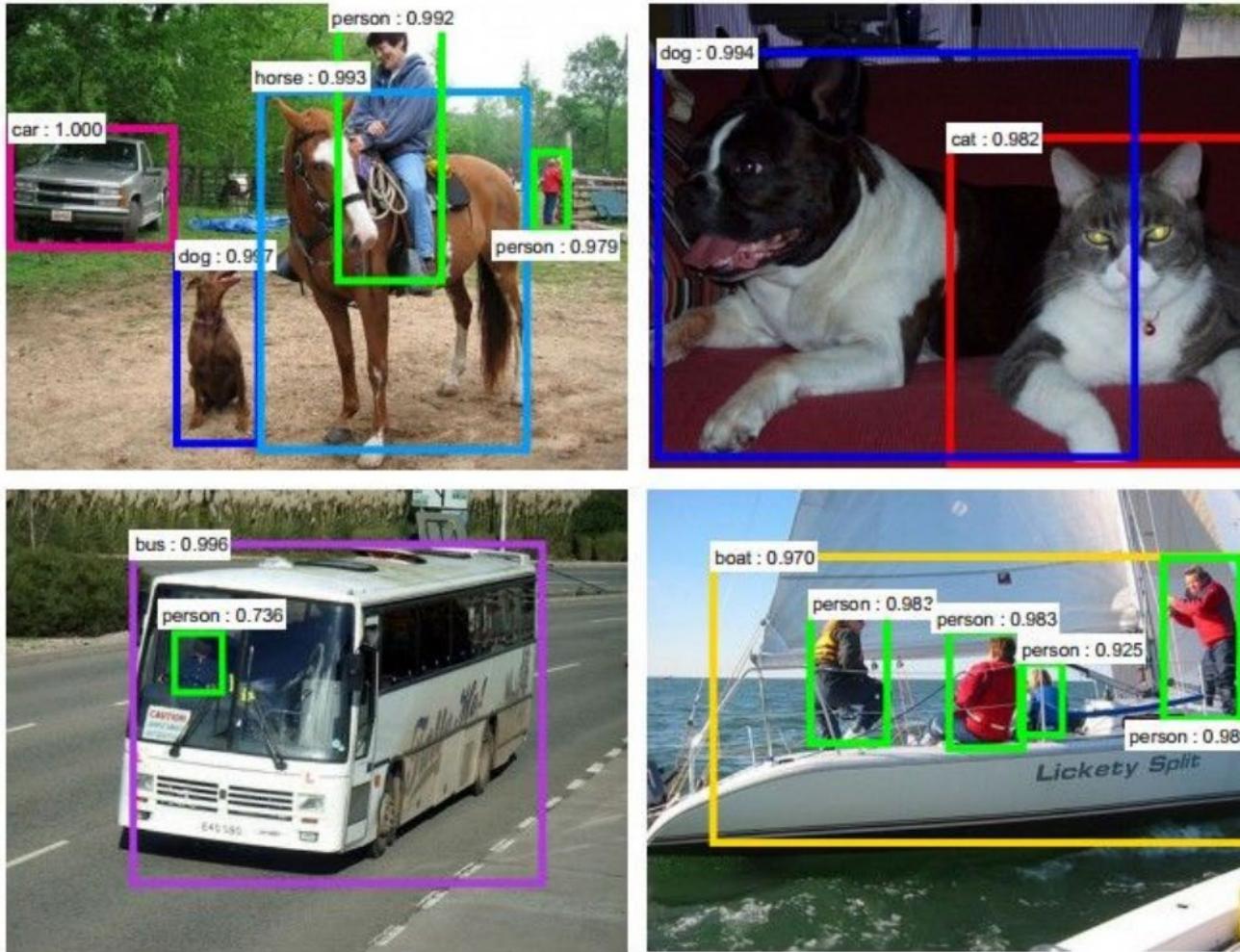


Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems* 25 (2012).

图像分类

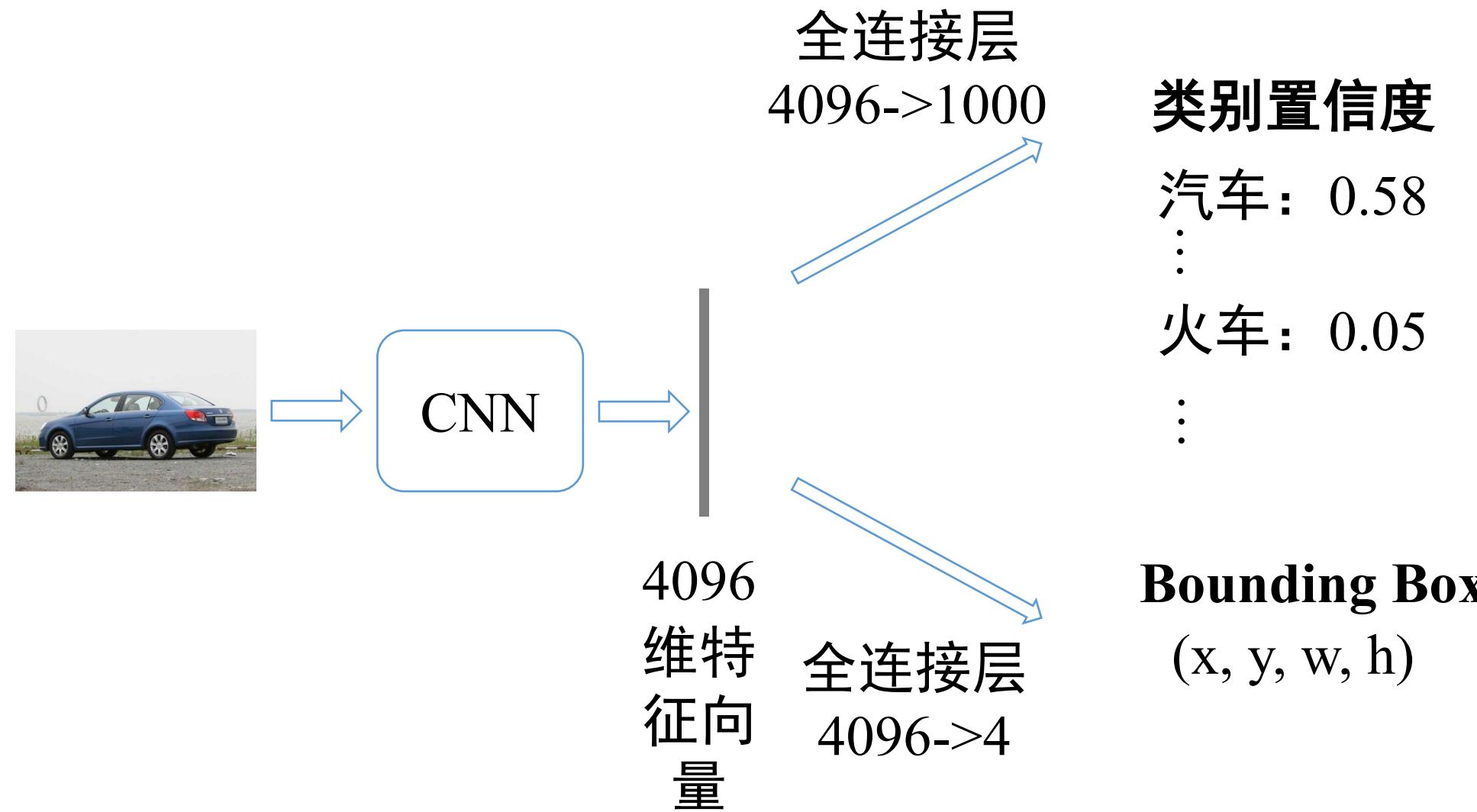
- [ResNet] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." *CVPR* 2016.
- [Swin Transformer] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, Baining Guo. "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows." *ICCV* 2021
- [MAGE] Tianhong Li, Huiwen Chang, Shlok Kumar Mishra, Han Zhang, Dina Katabi, Dilip Krishnan, MAGE: MAsked Generative Encoder to Unify Representation Learning and Image Synthesis, *CVPR* 2023

目标检测

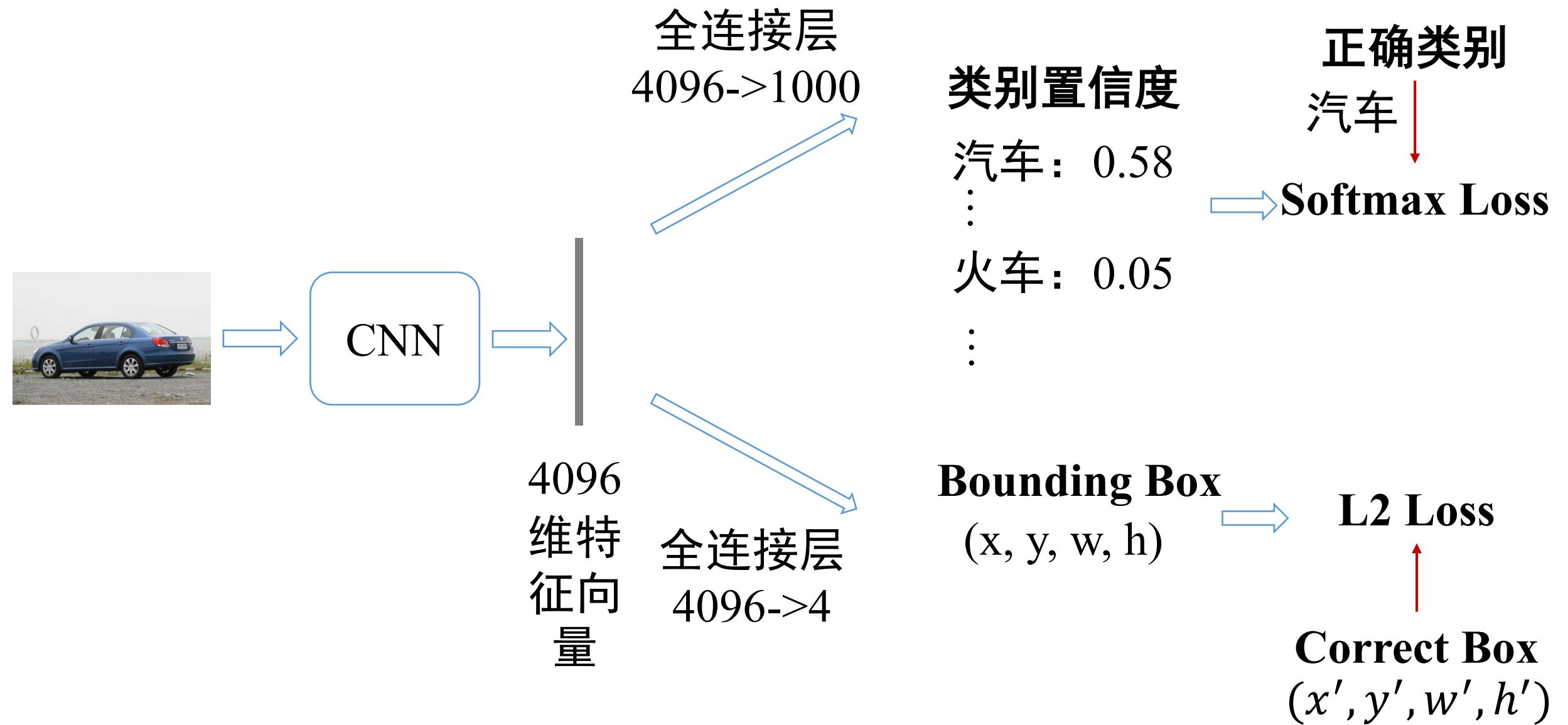


Girshick, Ross. "Fast r-cnn." *Proceedings of the IEEE international conference on computer vision*. 2015.

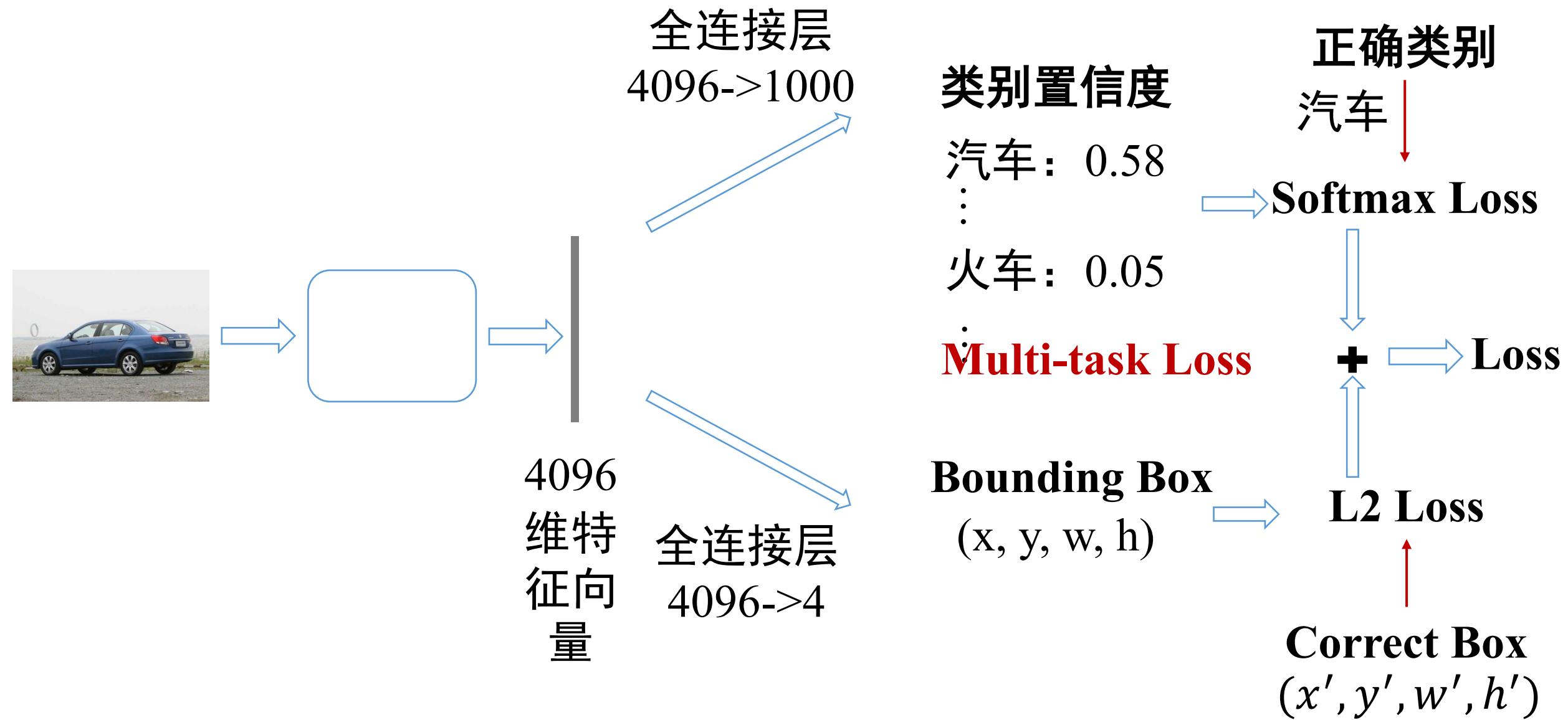
图像分类和目标定位



图像分类和目标定位



图像分类和目标定位



目标检测

- **[EfficientDet]** Mingxing Tan, Ruoming Pang, Quoc V. Le.
EfficientDet: Scalable and Efficient Object Detection. CVPR2020.
- **[DETR]** Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, Sergey Zagoruyko, End-to-End Object Detection with Transformers, ECCV2020 .
- **[DetCLIPv2]** Lewei Yao, Jianhua Han, Xiaodan Liang, Dan Xu, Wei Zhang, Zhenguo Li, Hang Xu. "DetCLIPv2: Scalable Open-Vocabulary Object Detection Pre-training via Word-Region Alignment." cvpr2023.

图像超分辨率重建

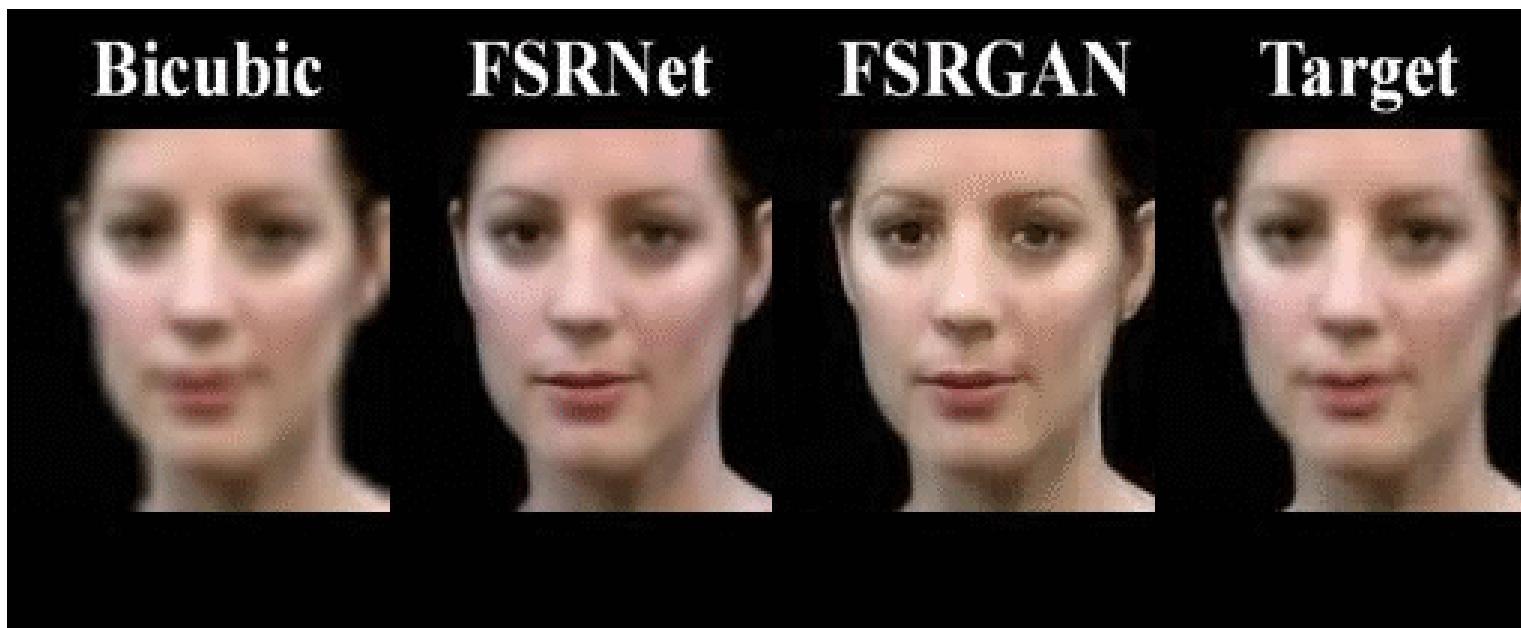
图像分辨率是一组用于评估图像中蕴含细节信息丰富程度的性能参数，包括时间分辨率、空间分辨率及色阶分辨率等，体现了成像系统实际所能反映物体细节信息的能力。相较于低分辨率图像，高分辨率图像通常包含更大的像素密度、更丰富的纹理细节及更高的可信赖度。但在实际上情况中，受采集设备与环境、网络传输介质与带宽、图像退化模型本身等诸多因素的约束，我们通常并不能直接得到具有边缘锐化、无成块模糊的理想高分辨率图像。



部分图片和文字来自网络：https://blog.51cto.com/u_15404184/4323223

图像超分辨率重建

图像的超分辨率重建技术指的是将给定的**低分辨率图像**通过特定的算法**恢复成相应的高分辨率图像**。具体来说，图像超分辨率重建技术指的是利用数字图像处理、计算机视觉等领域的相关知识，借由特定的算法和处理流程，从给定的低分辨率图像中重建出高分辨率图像的过程。



图像超分辨率重建

基于深度学习的超分辨率重构算法：

1. Dong C, Loy C C, He K, et al. Learning a deep convolutional network for image super-resolution[C]//European conference on computer vision. Springer, Cham, 2014: 184-199.
2. Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 1646-1654.
3. Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4681-4690.

图像生成

生成模型是一种训练模型进行无监督学习的模型，给模型一组数据，希望从数据中学习到信息后的模型能够生成一组和训练集尽可能相近的数据。图片生成模型就是更具体的指向说给模型一组图片作为训练集让模型进行学习，希望模型生成一组和训练集图片尽可能相近的图片。



利用DCGAN生成动漫人脸实例，分别为0, 100, 200次迭代结果。

部分图片和文字来自网络：https://blog.csdn.net/like_red/article/details/79192374

https://blog.csdn.net/qq_34739497/article/details/79902356

图像生成

图像生成经典算法：

1. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[J]. NeurIPS 2014.
2. Yang Song, Diederik P. Kingma, How to Train Your Energy-Based Models. arXiv 2021.
3. Tim Salimans, Andrej Karpathy, Xi Chen, Diederik P. Kingma, PixelCNN++: Improving the PixelCNN with Discretized Logistic Mixture Likelihood and Other Modifications, arxiv 2021.

图像编辑

利用生成模型，对用户提供的图像做各种个性化操作。



input: "A man with a beard wearing glasses and a beanie in blue shirt"



"A man with a beard wearing **glasses** and a beanie in blue shirt"



"A man with **a beard** wearing glasses and a beanie in blue shirt"



Input: "Yoshua Bengio with beard"



"Yoshua Bengio with **beard**"



Input: "Yoshua Bengio with glasses"



"Yoshua Bengio with **glasses**"

图像编辑

图像编辑经典算法：

1. Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, Björn Ommer. High-Resolution Image Synthesis with Latent Diffusion Models, CVPR 2022.
2. Lvmin Zhang, Anyi Rao, Maneesh Agrawala. Adding Conditional Control to Text-to-Image Diffusion Models. ICCV2023.
3. Senmao Li, Joost van de Weijer, Taihang Hu, Fahad Shahbaz Khan, Qibin Hou, Yaxing Wang, Jian Yang. Get What You Want, Not What You Don't: Image Content Suppression for Text-to-Image Diffusion Models. ICLR2024.

图像去噪（低光场景）

原始图像文件（RAW图像文件）包含从数字相机、扫描器或电影胶片扫描仪的图像传感器所处理数据。由于传感器排列不能重叠，因此通常是单通道数据。此外相较于RGB图像的8Bit存储格式，RAW数据可以存储更多位也意味着更丰富的细节，在低光场景下有着极大的优势。相机中内置的图像信号处理流水线（Image Signal Processing Pipeline, ISP Pipeline）可以将RAW图像转化为常见的RGB图像。由于传感器记录的数据与接收到的光子数量呈正相关关系，因此RAW数据的像素值与曝光时间成正比。

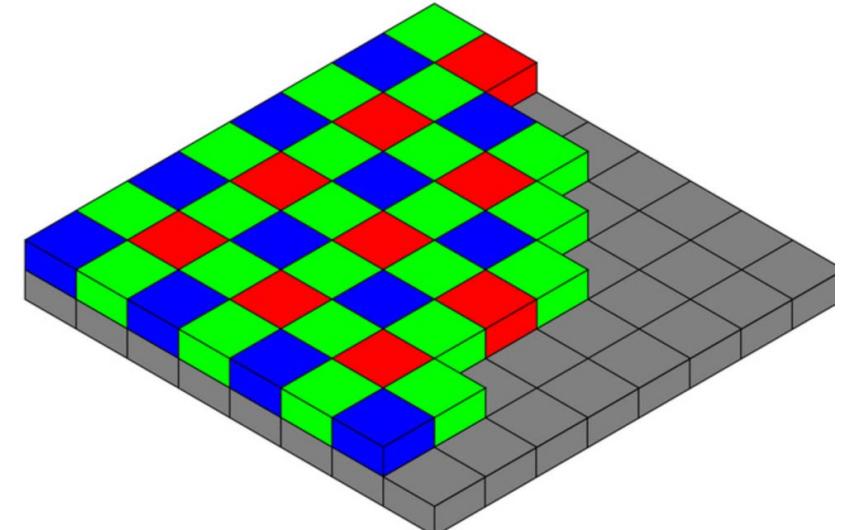


图1: 排列方式为BGGR的传感器

高质量视频修复（低光场景）

低光场景下的成像往往遇到光照条件不足、信噪比低的情况。基于RGB的方法由于ISP的影响，其噪声模型被破坏，因此最近的研究主要聚焦于光照条件的调节，如图2所示。对于RAW数据来讲，由于其噪声分布更为原始，因此通常可以处理更加极端的低光环境，如图3所示。

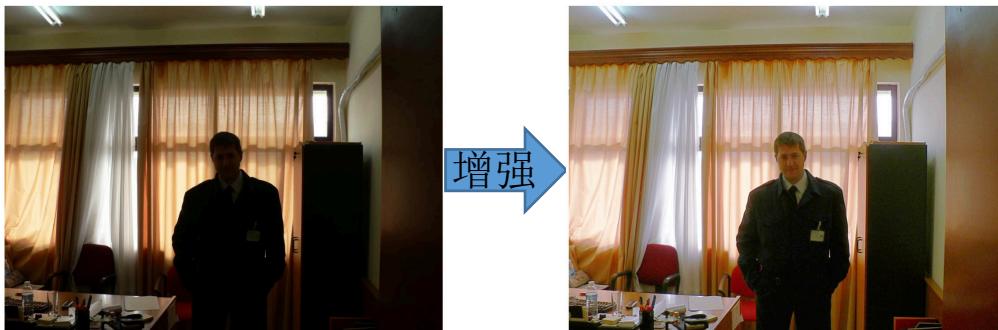


图2: 基于RGB图像的低光增强算法

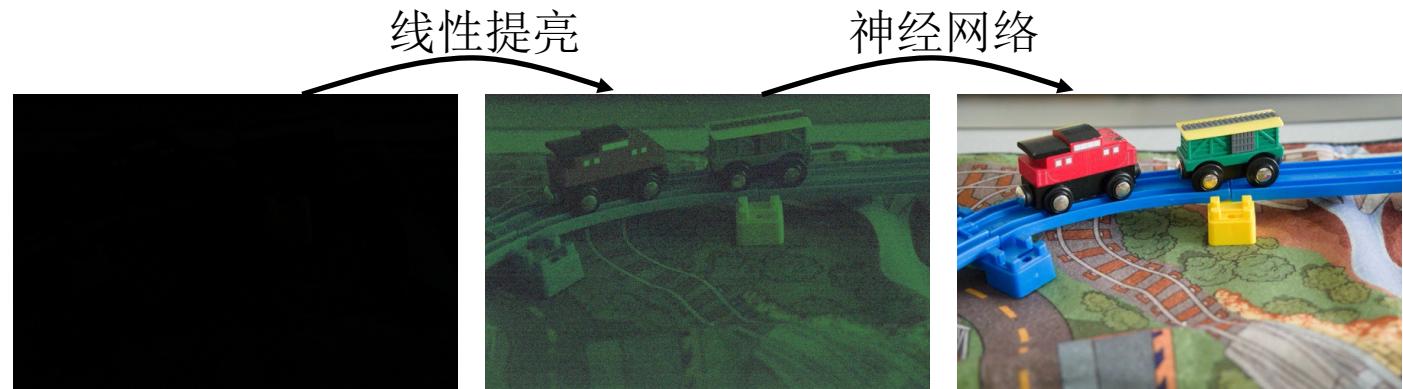


图3: 基于RAW图像的低光增强算法

高质量视频修复

1. Li Z, Lu C Z, Qin J, et al. **Towards an end-to-end framework for flow-guided video inpainting**[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 17562-17571.
2. Li Z, Zhu Z L, Han L H, et al. **Amt: All-pairs multi-field transforms for efficient frame interpolation**[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 9801-9810.
3. Zhou S, Yang P, Wang J, et al. **Upscale-A-Video: Temporal-Consistent Diffusion Model for Real-World Video Super-Resolution**[J]. arXiv preprint arXiv:2312.06640, 2023.

图像去噪（低光增强）

基于深度学习的低光增强算法：

1. Guo C, Li C, Guo J, et al. **Zero-reference deep curve estimation for low-light image enhancement**[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 1780-1789.
2. Jin X, Han L H, Li Z, et al. **Dnf: Decouple and feedback network for seeing in the dark**[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 18135-18144.
3. Jin X, Xiao J W, Han L H, et al. **Lighting every darkness in two pairs: A calibration-free pipeline for raw denoising**[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 13275-13284.

连续学习

- 什么是连续学习？
 - 机器学习模型像人类大脑一样连续更新
- 挑战：灾难性遗忘 (Catastrophic forgetting)
- Continual learning / Lifelong learning / Incremental learning



连续学习 (continual learning)

国家重大项目牵引

1.2 面向开放环境的自适应感知（2018年）

The screenshot shows the official website of the Chinese government (www.gov.cn). The main navigation bar includes links for国务院 (State Council), 总理 (Premier), 新闻 (News), 政策 (Policies), 互动 (Interaction), 服务 (Services), 数据 (Data), 国情 (National Conditions), and 国家政务服务平台 (National Government Service Platform). The page title is "中华人民共和国中央人民政府" (The State Council of the People's Republic of China). The main content area displays the "科技部关于发布科技创新2030—“新一代人工智能”重大项目2018年度项目申报指南的通知" (Notice of the Ministry of Science and Technology on the Release of the Application Guide for the Key Project of the New Generation Artificial Intelligence in the 2030 Innovation Plan for 2018). It provides detailed information about the project, including the title, issuing agency, document number, source, classification, date, and content.

1.2 连续学习理论和方法（2021年）

科技创新 2030—“新一代人工智能”重大 项目 2021 年度项目申报指南

为落实《新一代人工智能发展规划》，启动实施科技创新 2030—“新一代人工智能”重大项目。根据重大项目实施方案的部署，科技部组织编制了 2021 年度项目申报指南。

考核指标：研究建立针对知识和数据相融合的连续学习理论和模型；围绕不少于 3 个具有代表性任务场景，构建基准测试数据集；实现零样本和小样本学习，学习方法鲁棒可靠和可验证；实现连续学习过程中知识稳定增扩；开源学习数据、模型和软件框架等，开源成果下载量和注册使用等方面具有一定影响力。

连续学习 (continual learning)

基于深度学习的连续学习算法：

1. Rebiffé, Sylvestre-Alvise, et al. "**icarl: Incremental classifier and representation learning.**" Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. 2017.
2. Dhar, Prithviraj, et al. "**Learning without memorizing.**" Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019.
3. Zhai, Jiang-Tian, et al. "**Masked autoencoders are efficient class incremental learners.**" Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.

扫描填写组队信息



2024人工智能导论课程报告
分组-王亚星

每组由组长填写，每组只填一次，三人一组。

*1. 选择课程报告题目

- 图像超分
- 图像生成
- 图像分类
- 目标检测
- 图像编辑
- 低光增强
- 连续学习
- 高质量视频修复

*2. 组长

*3. 组员

*4. 组员

提交

问卷星 提供技术支持

谢谢！