

《人工智能引论》第一次讲解

闫博凯 11.4

- 有关课程
- 问题讲解

有关课程

- 课程特点
 - 提纲挈领，范围广泛
 - 目的是增加同学们对AI各领域的了解
 - 深度一定程度上取决于自己
- 如何学好这门课
 - 基础知识点很重要
 - 认真对待实验和作业
 - 重视细节，更要理解原理
- 辅助资料
 - 有关领域**经典**论文
 - B站等课程资源如：李宏毅深度学习(<https://www.bilibili.com/video/BV1JE411g7XF?p=3>),斯坦福大学CS224n(<https://www.bilibili.com/video/BV18Y411p79k?p=4>),CS231n(<https://www.bilibili.com/video/BV1i14y1w7GJ?p=1>)
 - 以我为主，为我所用

问题讲解

极小极大方法, $\alpha - \beta$ 剪枝

- MAX电脑, 修改 α 值
- MIN玩家, 修改 β 值
- 递归调用

第二次讲解

闫博凯 12.9

- 常见激活函数
- 自然语言处理模型
- 计算机视觉任务

常见激活函数

- Sigmoid函数

$$y = \frac{1}{1 + e^{-x}}$$

- 将输出映射到0-1内的连续区间
- 梯度饱和

- Tanh函数

$$y = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

- 梯度饱和
- 多用于序列问题

- ReLU函数

$$y = \begin{cases} x, & x \geq 0, \\ 0, & otherwise. \end{cases}$$

- 解决了梯度饱和问题
- 应用最为广泛

- LeakyReLU

$$y = \begin{cases} x, & x \geq 0, \\ \alpha x, & otherwise. \end{cases}$$

- ELU

-

$$y = \begin{cases} x, & x \geq 0, \\ \alpha(e^x - 1), & otherwise. \end{cases}$$

- Swish

$$y = \frac{x}{1 + e^{-\beta x}}$$

自然语言处理模型

- RNN

$$\mathbf{h}_t = \sigma(\mathbf{U}\mathbf{h}_{t-1} + \mathbf{W}\mathbf{x}_t)$$

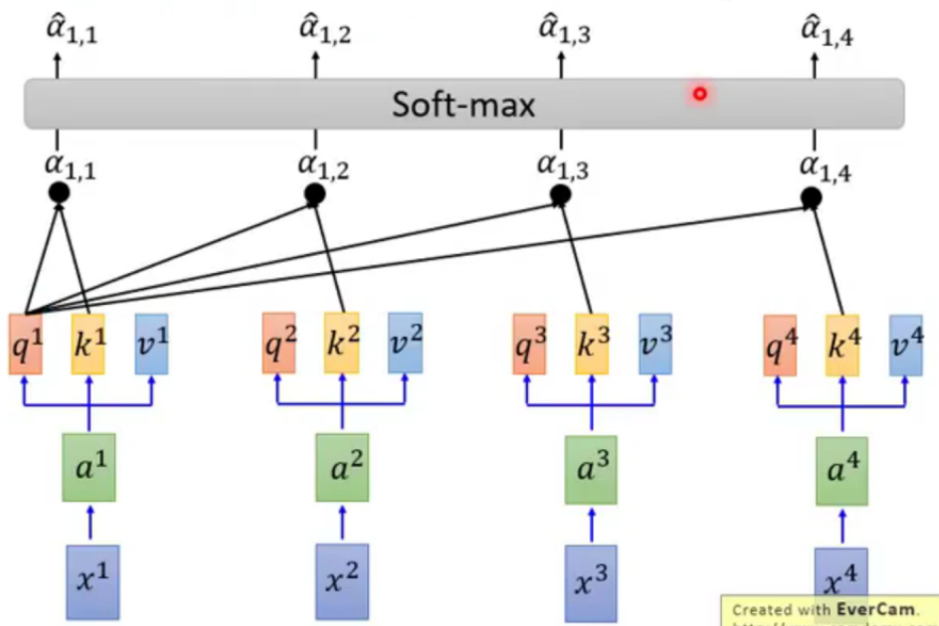
- LSTM

$$\begin{aligned}f_t &= \sigma(W_{fh}h_{t-1} + W_{fx}x_t + b_f), \\i_t &= \sigma(W_{ih}h_{t-1} + W_{ix}x_t + b_i), \\\tilde{c}_t &= \tanh(W_{\tilde{c}h}h_{t-1} + W_{\tilde{c}x}x_t + b_{\tilde{c}}), \\c_t &= f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t, \\o_t &= \sigma(W_{oh}h_{t-1} + W_{ox}x_t + b_o), \\h_t &= o_t \cdot \tanh(c_t).\end{aligned}$$

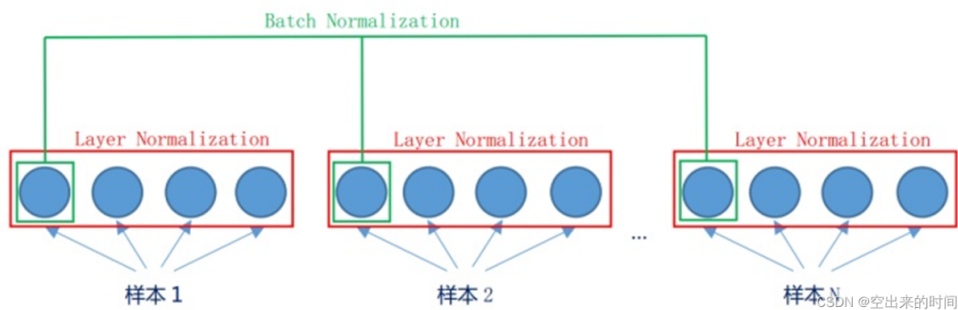
- Transformer
 - 自注意力机制

Self-attention

$$\hat{\alpha}_{1,i} = \exp(\alpha_{1,i}) / \sum_j \exp(\alpha_{1,j})$$

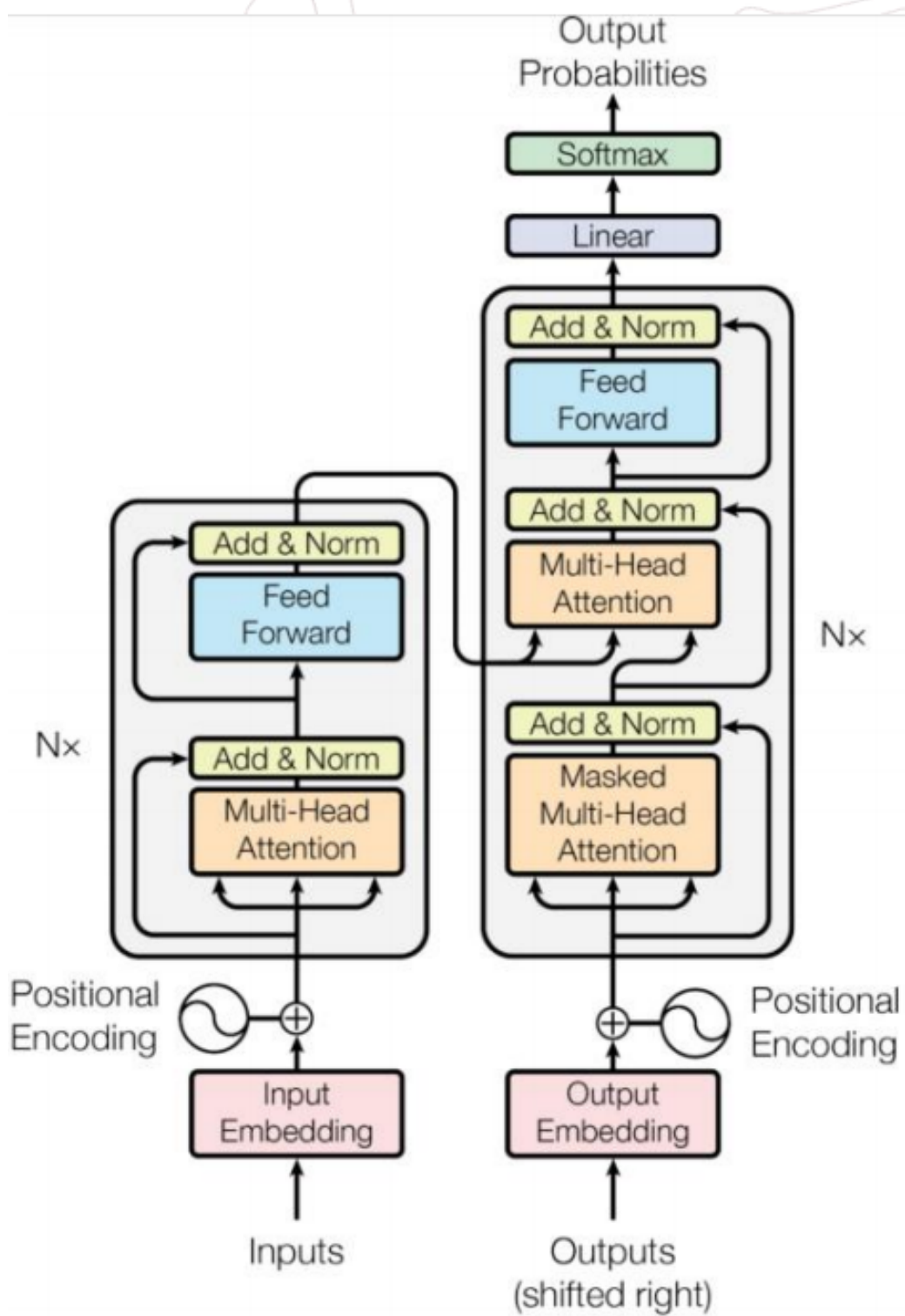


- 层归一化(Layernorm)



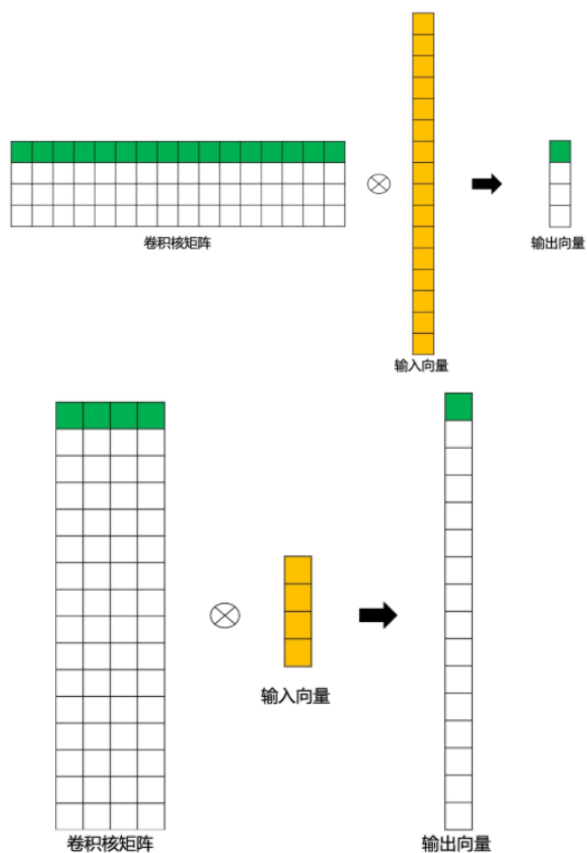
- 残差连接
- Encoder+Decoder
- Positional Encoding

$$p_t = \begin{bmatrix} \sin(w_1 \cdot t) \\ \cos(w_1 \cdot t) \\ \sin(w_2 \cdot t) \\ \cos(w_2 \cdot t) \\ \vdots \\ \sin(w_{d/2} \cdot t) \\ \cos(w_{d/2} \cdot t) \end{bmatrix}, w_k = \frac{1}{10000^{\frac{2k}{d}}}$$

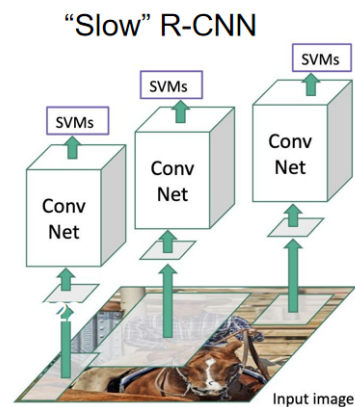
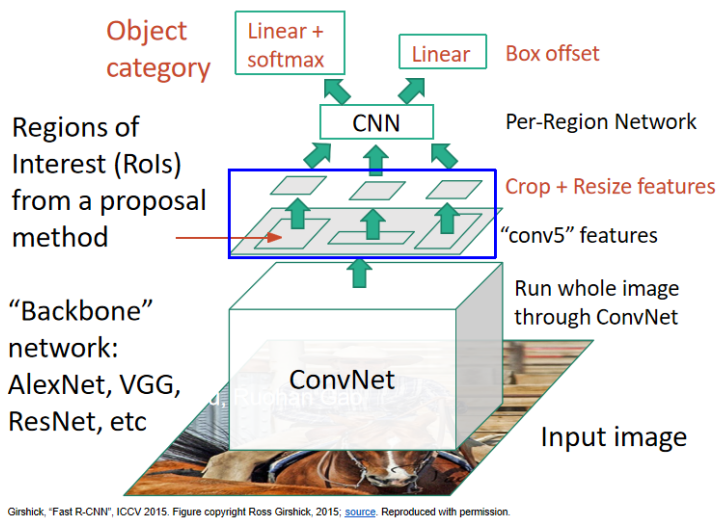


计算机视觉任务

- 图像分类
- 语义分割
 - 对每一个像素点进行分类
 - 转置卷积



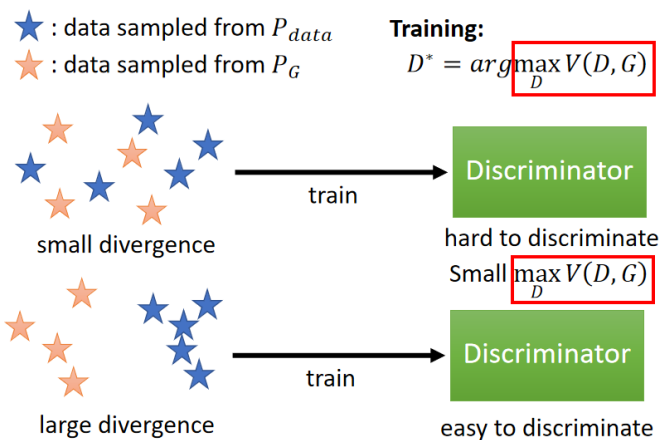
- 目标检测
 - R-CNN
 - Fast-RCNN、Faster-RCNN
 - YOLO模型



- 实例分割
 - Mask-RCNN
- 生成模型
 - GAN: 生成对抗网络

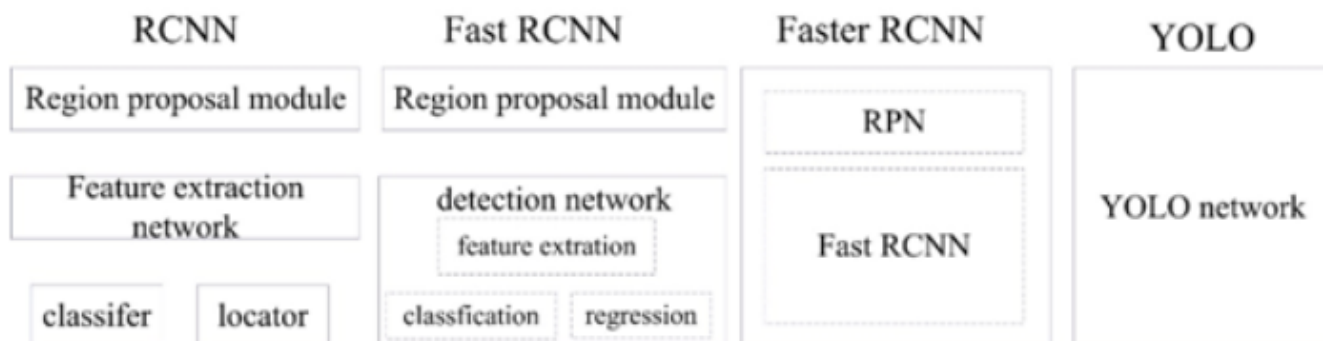
Discriminator

$$G^* = \arg\min_G \text{Div}(P_G, P_{data})$$



146

总结



第三次讲解

闫博凯 12.23

- 傅里叶变换
- 语音信号处理

傅里叶变换

- 离散时间傅里叶变换(DTFT)

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}$$

- z变换

$$X(z) = \sum_{n=-\infty}^{+\infty} x[n]z^{-n}$$

- 离散傅里叶变换(DFT)

$$X_N[k] = \sum_{n=0}^{N-1} x[n]e^{-j\frac{2\pi k}{N}n}$$

- 快速傅里叶变换(FFT)

$$\begin{aligned} X_N[k] &= \sum_{n=0}^{N-1} x[n]W_N^{kn} \\ &= \sum_{r=0}^{N/2-1} x[2r]W_N^{2rk} + \sum_{r=0}^{N/2-1} x[2r+1]W_N^{(2r+1)k} \\ &= \sum_{r=0}^{N/2-1} x[2r]W_{N/2}^{rk} + W_N^k \sum_{r=0}^{N/2-1} x[2r+1]W_{N/2}^{rk} \text{(旋转因子可约性)} \\ &= \sum_{r=0}^{N/2-1} g[r]W_{N/2}^{rk} + W_N^k \sum_{r=0}^{N/2-1} h[r]W_{N/2}^{rk} \end{aligned}$$

语音信号处理

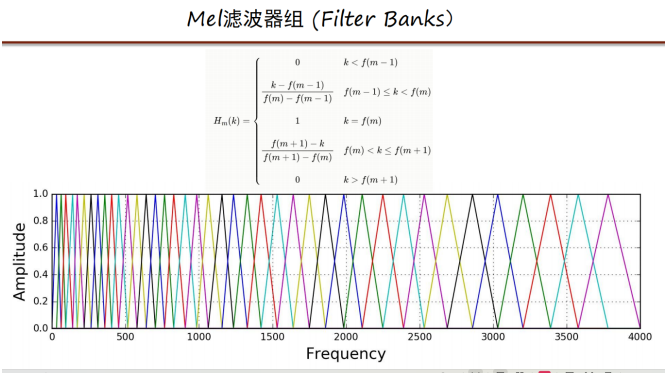
- 预加重

$$y[n] = x[n] - ax[n - 1]$$

- 分帧加窗
- 短时傅里叶变换(STFT)

$$X_N[m, k] = \sum_{n=-\infty}^{\infty} x[n + mT]w[n]e^{-j\frac{2\pi k}{N}n}$$

- Mel滤波
 - 更关注低频的部分



- 离散余弦变换(DCT)

$$X_{DCT}[t, k'] = \sum_{k=0}^{M-1} P_{\text{Mel}}[t, k] \cos \frac{\pi k'(2k + 1)}{2M}, k' = 0, \dots, L - 1$$

划重点

- 第一章
图灵测试、中文房间、各种流派、恐怖谷、神经元、各种发展史
- 第二、三章
搜索算法、解空间的画法、启发式搜索、A、A*算法、评价函数
- 第四章
知识如何表示、常见知识图谱(了解)，重点:Entity、Relation、Precision、Recall、F1
- 第五章
回归、分类问题、激活函数事例、loss优化、梯度下降、batch、epoch概念、过拟合问题、Kmeans方法
- 第六章
强化学习的基本概念和方法、马尔可夫过程、计算 γ 值等、价值函数和动作价值函数、贝尔曼方程、动态规划、蒙特卡洛搜索(着重掌握)、时序差分、Q学习(了解)
- 第七章
交叉熵、LSA、本吉奥的神经语言模型
- 第八章
RNN、LSTM、conditional generation、attention的计算
- 第九章
self-attention、q、k、v,multi-head attention、transformer、BERT、GPT(了解)
- 第十章
滤波技术、PCA算法
- 第十一章
梯度下降、反向传播原理、卷积过程、参数意义
- 第十二章
视觉任务及目标、转置卷积、GAN的概念和过程
- 第十三章
听觉的基本概念、分贝怎么计算、Mel滤波、Mel与频率的关系、MFCC理解
- 第十四章
token的选择、downsampling、Tacotron、fastspeech、length、regulator等