

计算能力与算法
注意力机制

- 动机与由来 \Rightarrow 减少信息汲取；利用有限人脑处理能力 \Rightarrow 帮助系统忽略无关紧要信息
- 注意力发生在突触水平上（改变突触的“权重”）
 - 提高对输入信号的敏感性，并锐化这些信号的精确度
 - 选择性地提高吸引注意力信息的传输，降低噪音水平或减弱干扰注意力信息的传输
- 注意力是一个用来分配有限信息处理能力的选择机制
 - 抑制某个区域的Input/Output以阻止干扰信息的进入
 - 加强某个区域的Signal以促进该信息的处理
- 自下而上（基于显著性的注意力）：见山是山，见水是水。即大脑内呈现的信息是通过视觉通路传递的外界刺激原始的物理特征，包括颜色、强度、方向等等。这是外部环境信息驱动的注意力。
 - 眼睛接收到像素信息，根据具体的像素信息成像。这属于Bottom-up过程即人眼根据某个像素及其周围的像素信息形成图像区域并进一步抽象为某一个物体形状。
- 自上而下（聚焦式注意力）：指的是大脑的（高级）联合皮层，包括前额叶皮层和后顶叶皮层根据当前任务的目标和以往的知识对视觉通路中的信息进行的调控。这是大脑内部信息驱动的注意力。
 - 眼睛扫描整个场景元素，寻找感兴趣的影像区域。这属于Top-down过程，即获得第一步物体形状之后，人眼会快速扫描整个视野场景，然后快速定位到感兴趣的地区。

定义：将类模型从认知注意力的技术

旨在通过适当方法将有限计算资源用在处理更重要信息，提高效率

目前通过引入多头交互机制，增强模型对感兴趣区域的感知力，从而提升Attention关注小而重要目标

简介：引入 Q (query), K (key), V (value) 三元组

目标 对象 距离度

① 点积矩阵 $A(Q, K, V) = \sum_{i=1}^n \text{similarity}(Q_i, K_i) \cdot V_i$

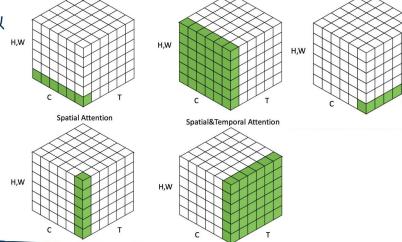
一强注意力 $D_{ij} \approx 1$ 二道分布 $D_{ij} \approx 0$ 不关注 $D_{ij} \approx 0$ 关注 $D_{ij} \approx 1$ 有吸引出注意力，可通过通过归一化

一软注意力 D_{ij} 连续分布 $D_{ij} \in [0, 1]$ 强度高低由 D_{ij} 表示 关注区域通过可微，可利用梯度逆向传播学习注意力权重

点积注意力

根据作用对象不同可以

- 空间注意力
- 通道注意力
- 混合注意力
- 时间注意力



d. 空间注意力

输入图像用 x 表示，特征提取网络用 ϕ ，则 $\phi(x)$ 代表卷积操作提取特征图。

注意力模型的locations随着时间一直变化，每一时刻的定位组成 $\{l_1, l_2, l_3, \dots, l_t\}$ ，其中 $l_i \in \pi(\{\phi(x)\})$ ， π 代表一个选择标准（选择峰值响应区域）。

E. 时间注意力

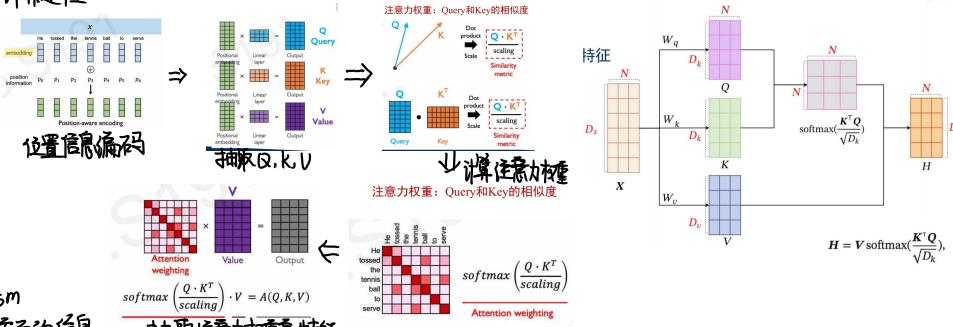
对于时间域的注意力机制，应用最早的当属NLP领域。斯坦福大学的一篇Effective Approaches to Attention-based Neural Machine Translation提出了一种软时间注意力机制。

自注意力

缩放点积注意力模块

(Q, K, V) 三元组捕捉长距离依赖建模式： $\text{dot}(\text{dot}(k, Q), V)$

Scaled Dot-Product Attention 计算过程



擅长捕捉数据或特征的内部相关性

多头自注意力 \Rightarrow Multi head selfattention mechanism

特征模型分为多个头形成若干子空间，多模型关注不同方面的信息，抽取注意力权重高特征

典型应用：① 文本翻译

② 图像分类 Image $\xrightarrow{\text{slice}}$ small unit $\xrightarrow{\text{transformer}}$ Encode $\xrightarrow{\text{NN}}$ train

③ 目标检测 DETR

④ 语义分割 SETR

⑤ 点云分类与切割 PCTRC Point cloud ...

⑥ 行为识别 MAD (多标签动作识别)

Transformer 回归偏置 \Rightarrow 学习算法对学习的局限的一些假设

CNN：假设特征具有局部性 \Rightarrow 将相邻特征融合更易得解

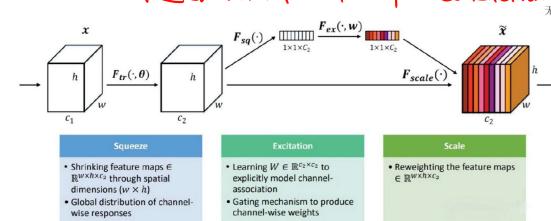
RNN 假设每一时刻计算依赖于历史计算

Transformer 避免了 CNN 局部性回归偏置问题

a. 空间注意力 \Rightarrow 自注意力关键位置

b. 通道注意力 \Rightarrow SENet，学习不同通道特征重要程度

提出 Block 将通道注意力分成三个部分：Squeeze, Excitation, Scale



c. 空间通道注意力

空间通道注意力是一种混合注意力，顾名思义即将空间注意力和通道注意力结合。

卷积块注意力模块 (CBAM) 是一个直观的混合注意力机制模型，它在通道域和空间域同时加入了注意力机制。空间注意力引导网络关注“where”；通道注意力引导网络关注“what”。

