# 1.思路

[关于注意力机制的学习\_注意力机制放在深层还是浅层-CSDN博客](https://blog.csdn.net/hands_up_down/article/details/121973082)

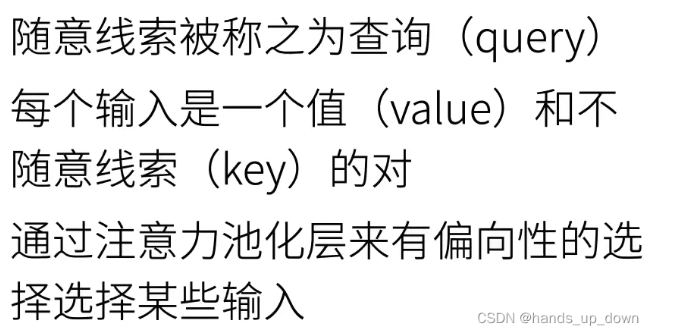
人类根据随意线索和不随意线索选择注意点

在我的理解里面，这里的随意是指跟随主观意识，他后面又直接提到像一般的CNN网络都属于是根据不随意线索来选择注意点，即CNN网络是根据不随意线索来对网络进行的训练，我们对网络要看什么是不会做限定的，网络自己决定先看什么再看什么

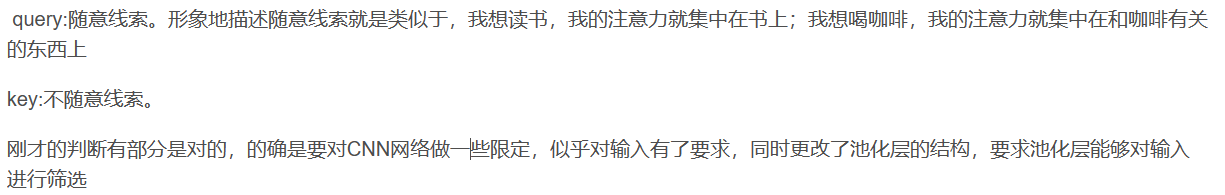
那我又可以想，这里的主观意识究竟是谁的，是人的，还是网络自己的。网络本身也是由人类来搭建的，网络所拥有的所谓的主观意识是否也仅仅是创造它的人的主观意识的反映呢？？

从ZFnet中做的可视化来看，CNN网络一般是看一些浅层的信息，线条、边缘、形状，之后会出现纹理、颜色，最后才是高度具象、与要识别的类别目标高度类似的信息

那我马上就可以产生一个猜想，attention mechanism是不是要对卷积网络做一些限定，以使他根据随意线索来进行训练



注意力池化层？



对于detr系列的重新回顾

对于注意力机制、特征金字塔、卷积块、池化、sppf、损失函数的新的东西

# 模型

1. ECM加哪里

想的应该是放到c3分支部分，然后有个问题，直接加上去，然后1\*1，还是怎么替换

1. 对于upsample的创新
2. 对于decoder部分是否可以改进
3. 对于rep部分
4. 对于sppf部分

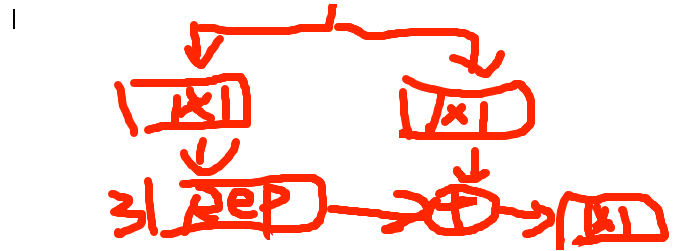
先是那个LPE失败了，后面可以试一下看它在什么种类好

然后热力图上一下，分析每个层上热力图

还有就是将每个模块的网络画出。

Repconv

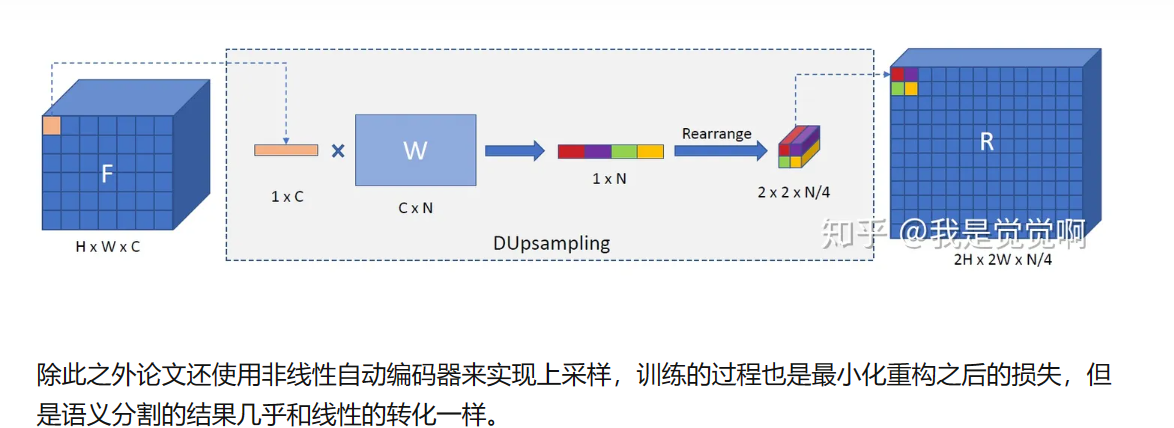
Repconv3



38. ultralytics/cfg/models/rt-detr/rtdetr-Conv3XCC3.yaml

    使用[Swift Parameter-free Attention Network](https://github.com/hongyuanyu/SPAN/tree/main)中的Conv3XC改进RepC3.

dupsampling



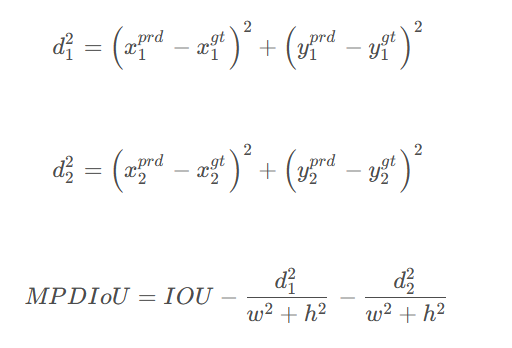
还有对这个可变形卷积以及偏移有点感觉，不知道是否有点想的偏

就是可变形卷积可以针对特征在尺寸变化的时候根据一个参考（可以是原，也可以是前一层特征图作为参考）

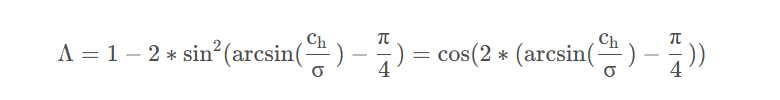
拼接后的特征在进行卷积操作之前可能蕴含了更多的语义信息，卷积操作可以帮助模型更好地利用这些信息来生成偏移量。dcn会根据offset

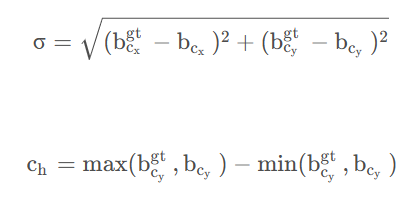
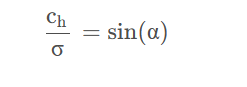
我是否可以这样：，一个使用FSM，一个使用1\*1，然后让两个再和upsample后的特征进行拼接，然后卷积作为offset。

而且里面的dcn看是换成v3还是其他改进。

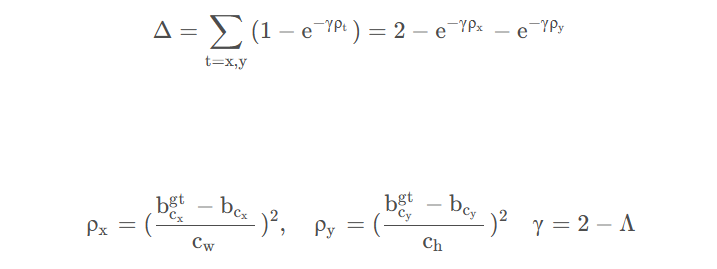


1. 角度





2.距离



34. ultralytics/cfg/models/rt-detr/rtdetr-AIFI-DAttention.yaml

    使用[Vision Transformer with Deformable Attention CVPR2022](https://github.com/LeapLabTHU/DAT)中的DAttention改进AIFI.

这个aifi改变不一定有好的，但是要尝试

关键我感觉还是在concat前后

46. ultralytics/cfg/models/rt-detr/rtdetr-slimneck-ASF.yaml

    使用[SlimNeck](https://github.com/AlanLi1997/slim-neck-by-gsconv)中的VoVGSCSP\VoVGSCSPC和GSConv和[ASF-YOLO](https://github.com/mkang315/ASF-YOLO)中的Attentional Scale Sequence Fusion改进rtdetr中的CCFM.

