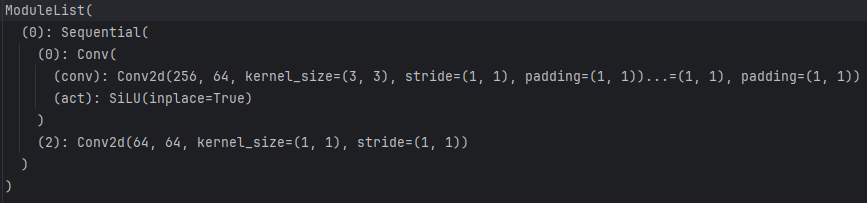
## 检测头分析(WorldDetect)

### 结构分析

self.cv2：



self.cv3：三个不同的推理分支

self.cv3[0]:

Conv(256, c3, 3)

Conv(c3, c3, 3)

nn.Conv2d(c3, embed, 1)

self.cv3[1]:

Conv(x, c3, 3)

Conv(c3, c3, 3)

nn.Conv2d(c3, embed, 1)

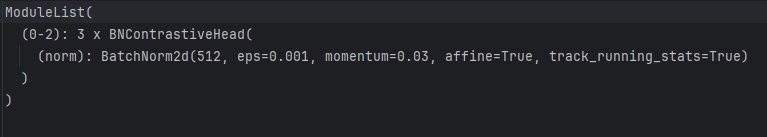
self.cv3[2]:

Conv(x, c3, 3)

Conv(c3, c3, 3)

nn.Conv2d(c3, embed, 1)

self.cv4：三个完全一致的BNContrastiveHead：



self.nl=3

### 前向算法分析(WorldDetect.forward)

输入张量：

* x：
  + 输入尺寸：List[Tensor1, Tensor2, Tensor3]
    - Tensor1：(batch, 256, 80, 80)
      * 数据源来自15层C2fAttn的输出
    - Tensor2：(batch, 512, 40, 40)
      * 数据源来自18层C2fAttn的输出
    - Tensor3：(batch, 512, 20, 20)
      * 数据源来自21层C2fAttn的输出
* text：
  + 输入尺寸：(batch, n, 512)，n为输入的目标文本数量

#### 推理分支[1]