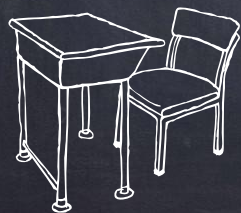


# CNN FOR CV

## 答疑课-0825

谭师兄



# CONTENTS

---

**01** | 神经网络基础梳理&细节

**02** | 关于Pytorch的问题

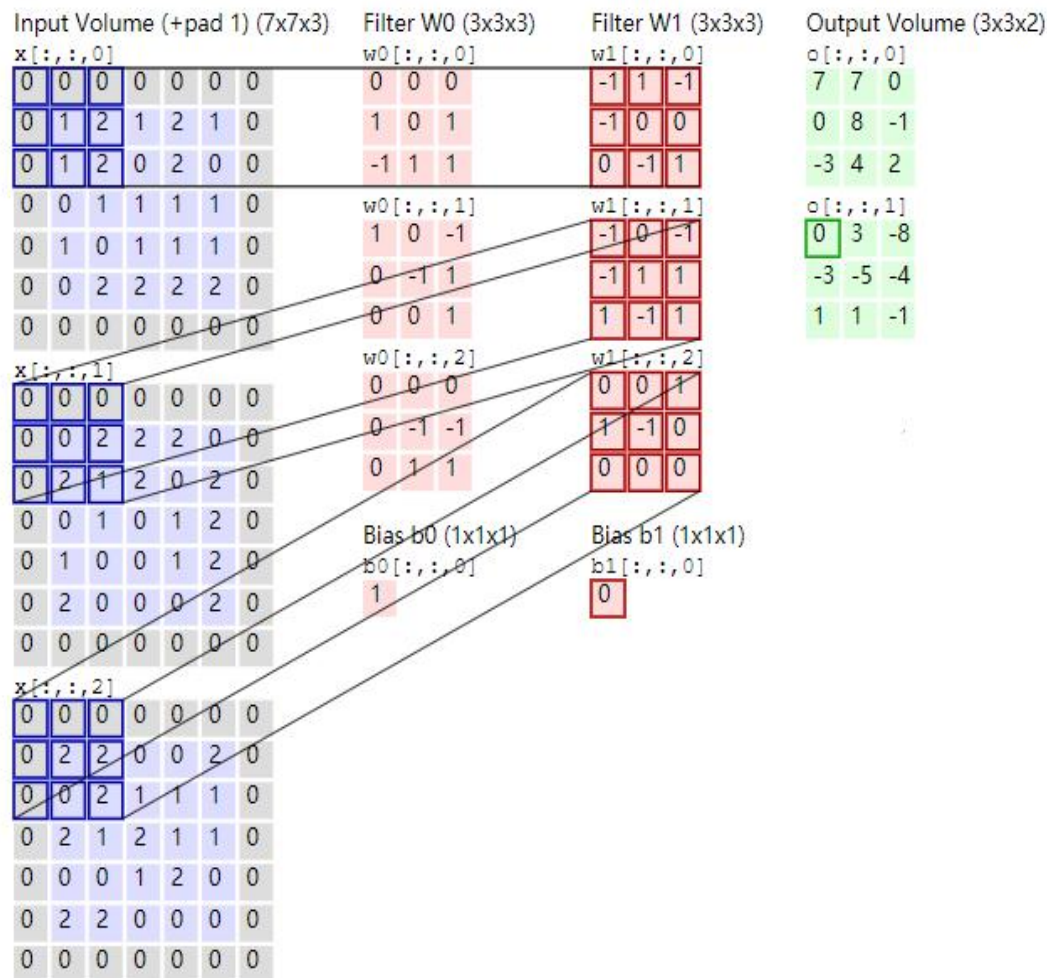
**03** | 关于项目的问题



训练出的model是个啥？

什么是learn？

训练和预测模型是一回事吗？

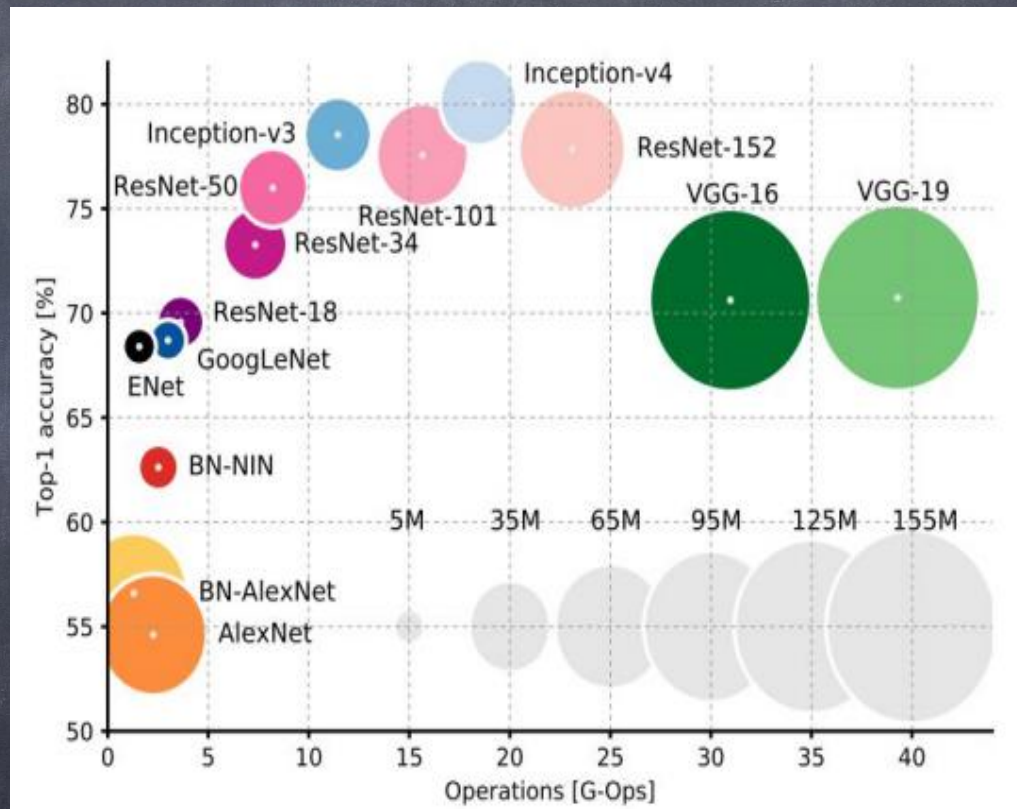
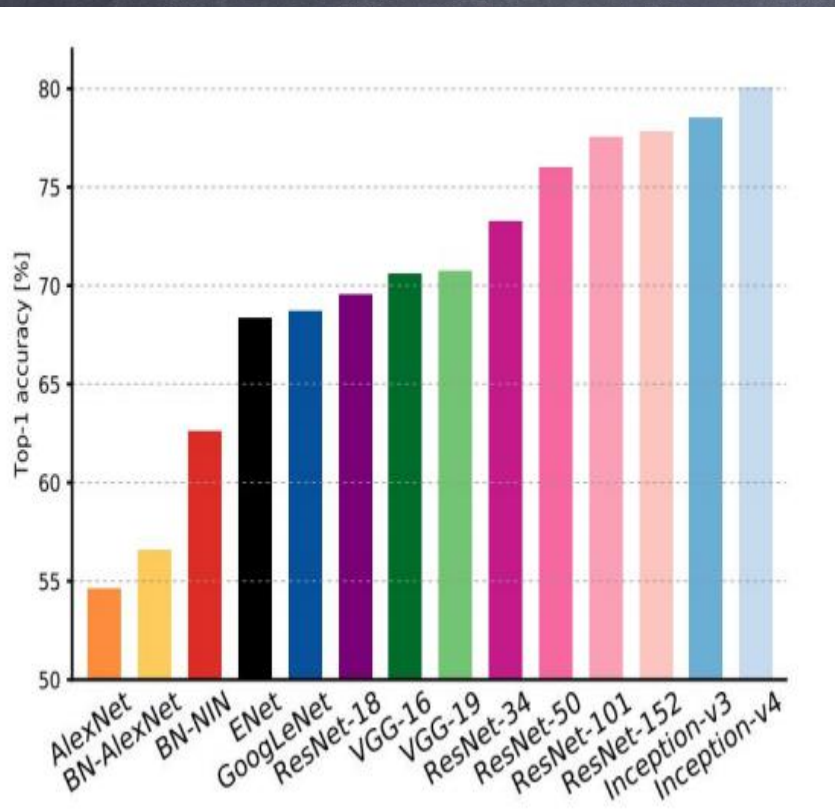


BP 在 FC、Conv 和 Pooling layer 的推导？

<https://www.cnblogs.com/pinard/p/6422831.html>

<https://www.cnblogs.com/pinard/p/6494810.html>

# 如何选择Backbone?





## 如何选择Backbone?

VGG--highest memory, most operations

GoogLeNet--most efficient

AlexNet--smaller compute, still memory heavy, lower accuracy

RestNet--moderate efficiency depending on model, highest accuracy

*(Current Best default)*

## VGG16, Resnet等, 固定输入尺寸, 除了resize还有什么方法?

resize

crop

提前pooling (会很耗内存, 只能用很小的batchsize)

## Nesterov 推导处有问题

### III. Optimization

#### D. Momentum Based

##### D2. Nesterov

把③代入④

原式:

$$\begin{cases} v_t = \gamma v_{t-1} - \alpha \nabla \bar{J}(\theta_t + \gamma v_{t-1}) & \textcircled{1} \\ \theta = \theta + v_t & \textcircled{2} \end{cases}$$

变形:

$$\begin{cases} v_t = \gamma v_{t-1} - \alpha \nabla \bar{J}(\theta_{t+1}) & \textcircled{3} \\ \theta_t = \theta_{t-1} + \gamma v_t - \alpha \nabla \bar{J}(\theta_{t+1}) & \textcircled{4} \end{cases}$$

$$\Rightarrow \theta_t = \theta_{t-1} + \gamma^2 v_{t-1} - (1+\gamma) \alpha \nabla \bar{J}(\theta_{t+1})$$

$$\Rightarrow \phi_{t-1} = \theta_{t-1} + \gamma v_{t-1} \quad \textcircled{5}$$

$$\Rightarrow v_t = \gamma v_{t-1} - \alpha \nabla \bar{J}(\phi_{t-1})$$

$$= \theta_t + \gamma v_t$$

$$= \theta_{t-1} + (1+\gamma) v_t$$

由⑤

$$= \phi_{t-1} - \gamma v_{t-1} + (1+\gamma) [\gamma v_{t-1} - \alpha \nabla \bar{J}(\phi_{t-1})]$$

$$= \phi_{t-1} + \gamma^2 v_{t-1} - (1+\gamma) \alpha \nabla \bar{J}(\phi_{t-1})$$

学员ID 99197



若初始模型在训练过程中表现不好，该如何对模型进行调整？

1. 数据角度（增强数据集）
2. 模型角度（在数据充足的情况下，增加模型复杂度）
3. 调参优化（初始化方法、优化算法、学习率、正则、loss...）

输入数据的transform 和 dataloader 的细节

**torchvision.transforms**

**torch.utils.data.DataLoader**

<https://pytorch.org/docs/stable/torchvision/transforms.html>

<https://pytorch.org/docs/stable/data.html?highlight=dataloader#torch.utils.data.DataLoader>



在load模型时，optimizer的参数需不需要load？可以load的数据，除了model，model.params，还有些啥？

<https://zhuanlan.zhihu.com/p/38056115>

Pytorch 如何训练/冻结特定的层？

<https://www.cnblogs.com/jiangkejie/p/11199847.html>

训练中loss出现Nan，怎么分析？

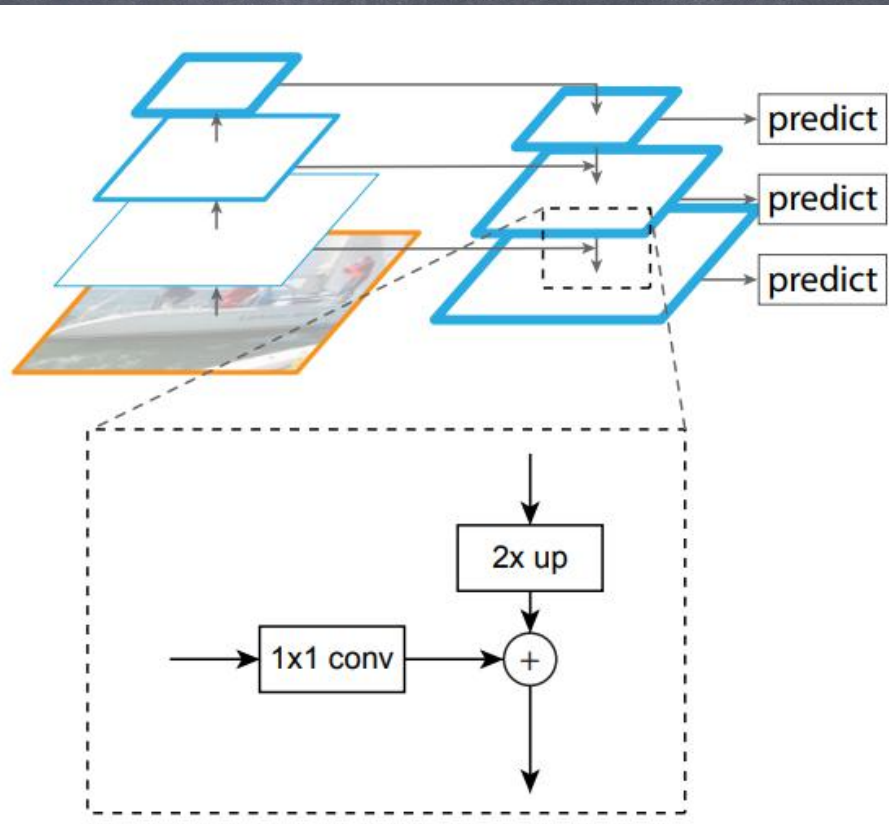
如果数据没问题的话看看是不是learning rate的问题 一般先从数据排查，看看数据是否存在nan。再就是loss 公式本身，有没有 /0 或者log(0)的情况。还有就是看lr是不是过大 和 是否梯度爆炸。

## FPN (Feature Pyramid Network. ) 上采样怎么理解?

上层语义信息丰富，更抽象；

下层图像定位细节信息丰富，  
有利于小目标的检测。

upsampled by a factor of 2 using  
Nearest-neighbor interpolation





像给的代码案例中， $w = \text{rect}[2] - \text{rect}[0] + 1$   $h = \text{rect}[3] - \text{rect}[1] + 1$ ，这个+1的细节是否需注意？或者是处于怎样的考虑？

0-9 9-0+1 = 10 个数

Batch Normalization 的  
最后一个公式怎么实现平移？

**Input:** Values of  $x$  over a mini-batch:  $\mathcal{B} = \{x_1 \dots x_m\}$ ;  
Parameters to be learned:  $\gamma, \beta$

**Output:**  $\{y_i = \text{BN}_{\gamma, \beta}(x_i)\}$

$$\mu_{\mathcal{B}} \leftarrow \frac{1}{m} \sum_{i=1}^m x_i \quad // \text{ mini-batch mean}$$

$$\sigma_{\mathcal{B}}^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_{\mathcal{B}})^2 \quad // \text{ mini-batch variance}$$

$$\hat{x}_i \leftarrow \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \epsilon}} \quad // \text{ normalize}$$

$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv \text{BN}_{\gamma, \beta}(x_i) \quad // \text{ scale and shift}$$

`optimizer.zero_grad()` 为什么梯度需要初始化为0?

因为求出每个minibatch的梯度，`backward()`会累积，如果不清零在下个minibatch会继续累加

`running_loss += loss.item() * inputs.size(0)` loss结果为什么要乘上  
`input.size(0)`?

`inputs.size(0)` would give the total images in each batch.

SVD能不能重构图像?

<https://blog.csdn.net/ouening/article/details/81814020>



VGG、resnet模型，图像数据输入前，要有减去均值；  
densenet是缩放到（0，1），这些是否必要？

可以在反向传播中加快网络中每一层权重参数的收敛

<https://blog.csdn.net/hai008007/article/details/79718251>

[https://blog.csdn.net/elaine\\_bao/article/details/50889856](https://blog.csdn.net/elaine_bao/article/details/50889856)

Adagrad优化方法，其速度更新是用梯度的平方加上过往的梯度平方，  
在这里为什么用梯度平方？它有什么意义？RMSProp和Adam的更新速度，  
都是对梯度平方进行加权平均，这个加权平均是为什么？

<https://www.cnblogs.com/wuliytTaotao/p/9479958.html>

[https://zh.d2l.ai/chapter\\_optimization/adam.html](https://zh.d2l.ai/chapter_optimization/adam.html)

怎么回归heatmap? 能否讲解一下, 然后贴些资料, 网上好难找

<https://zhuanlan.zhihu.com/p/42968117>

<https://github.com/raymon-tian/hourglass-facekeypoints-detection>



THANK YOU

