

# 法律声明

本课件包括演示文稿、示例、代码、题库、视频和声音等内容，深度之眼和讲师拥有完全知识产权；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或者机构不得盗版、复制、仿造其中的创意和内容，我们保留一切通过法律手段追究违反者的权利。

## 课程详情请咨询

- 微信公众号：深度之眼
- 客服微信号：deepshare0920



公众号



微信

关注公众号深度之眼，后台回复论文，获取60篇AI必读经典前沿论文





deepshare.net

深度之眼

# BN、LN、IN and GN

导师：余老师

---

关注公众号深度之眼，后台回复论文，获取60篇AI必读经典前沿论文



# 目录

1/ 为什么要Normalization?

2/ 常见的Normalizaton——BN、LN、IN and GN

3/ Normalization 小结



# Normalization in DL

## Normalization in DL

### Why Normalization

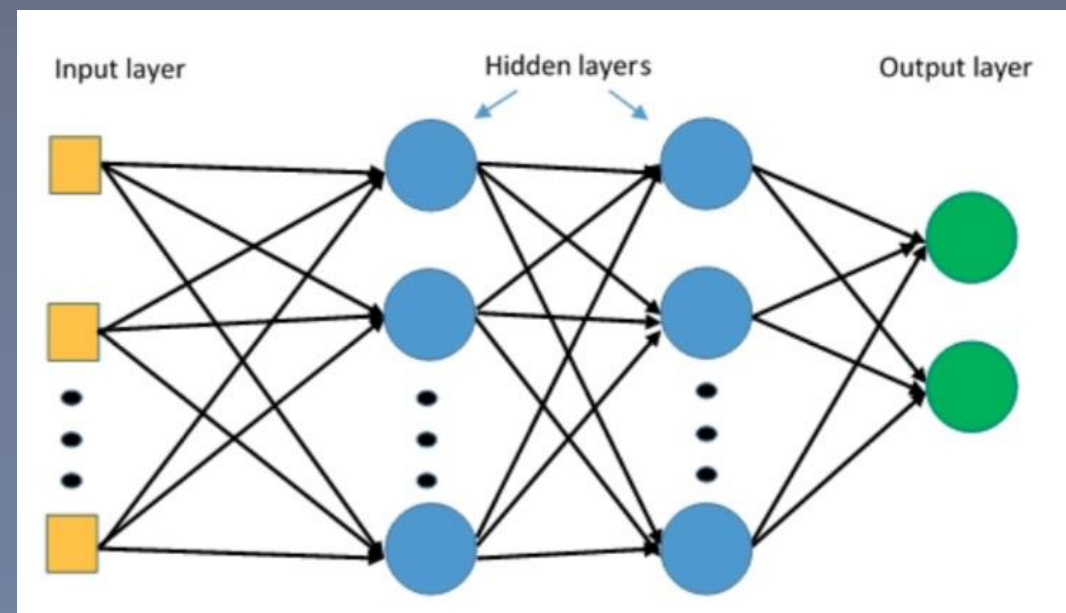
**Internal Covariate Shift (ICS): 数据尺度/分布异常, 导致训练困难**

$$H_{11} = \sum_{i=0}^n X_i * W_{1i}$$

$$\begin{aligned} D(H_{11}) &= \sum_{i=0}^n D(X_i) * D(W_{1i}) \\ &= n * (1 * 1) \\ &= n \end{aligned}$$

$$\text{std}(H_{11}) = \sqrt{D(H_{11})} = \sqrt{n}$$

$$D(H_1) = n * D(X) * D(W)$$



$X$   $W_1$   $H_1$   $W_2$   $H_2$   $W_3$   $out$

关注公众号深度之眼, 后台回复论文, 获取60篇AI必读经典前沿论文

# Normalization in DL

Normalization in DL

## 常见的Normalization

1. Batch Normalization (BN)
2. Layer Normalization (LN)
3. Instance Normalization (IN)
4. Group Normalization (GN)

同

异

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

$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv \mathbf{N}_{\gamma, \beta}(x_i)$$

均值和方差求取方式

# Normalization in DL

## Normalization in DL

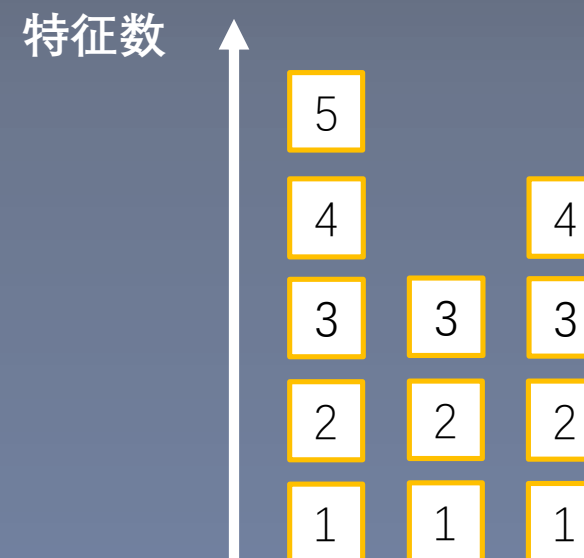
### 1. Layer Normalization

起因：BN不适用于变长的网络，如RNN

思路：逐层计算均值和方差

注意事项：

1. 不再有running\_mean和running\_var
2. gamma和beta为逐元素的



# Normalization in DL

## Normalization in DL

### nn.LayerNorm

主要参数：

- **normalized\_shape**: 该层特征形状
- **eps**: 分母修正项
- **elementwise\_affine**: 是否需要affine transform

```
nn.LayerNorm(  
    normalized_shape,  
    eps=1e-05,  
    elementwise_affine=True)
```

# Normalization in DL

## Normalization in DL

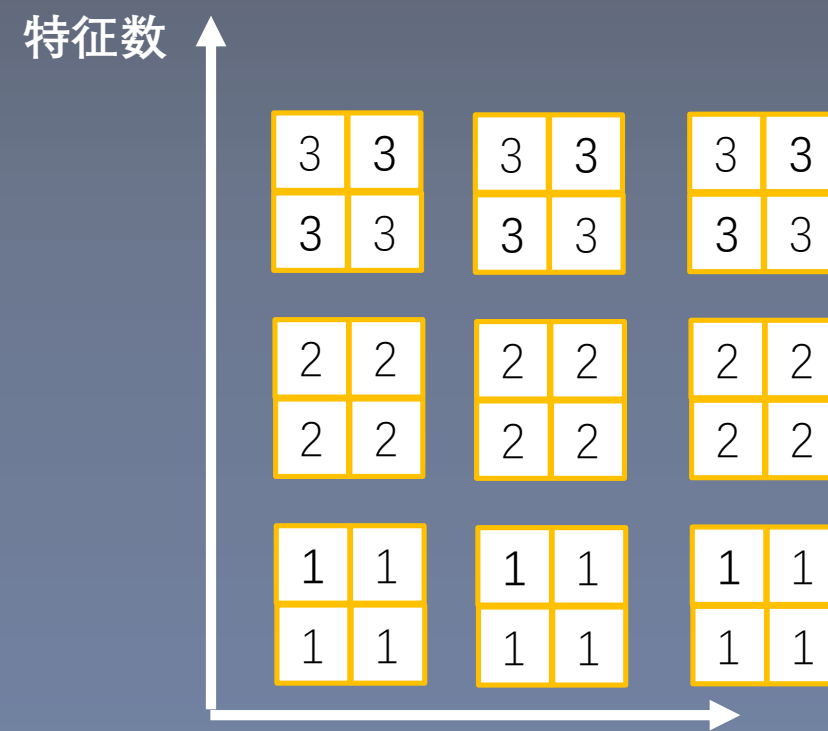
### 2. Instance Normalization

起因：BN在图像生成（Image Generation）中不适用

思路：逐Instance (channel) 计算均值和方差



Figure 1: Artistic style transfer example of Gatys et al. (2016) method.



《Instance Normalization: The Missing Ingredient for Fast Stylization》

《Image Style Transfer Using Convolutional Neural Networks》

关注公众号深度之眼，后台回复论文，获取60篇AI必读经典前沿论文



# Normalization in DL



## Normalization in DL

### nn.InstanceNorm

主要参数:

- **num\_features**: 一个样本特征数量 (最重要)
- **eps**: 分母修正项
- **momentum**: 指数加权平均估计当前mean/var
- **affine**: 是否需要affine transform
- **track\_running\_stats**: 是训练状态, 还是测试状态

```
nn.InstanceNorm2d(  
    num_features,  
    eps=1e-05,  
    momentum=0.1,  
    affine=False,  
    track_running_stats=False)
```

# Normalization in DL

## Normalization in DL

### 3. Group Normalization

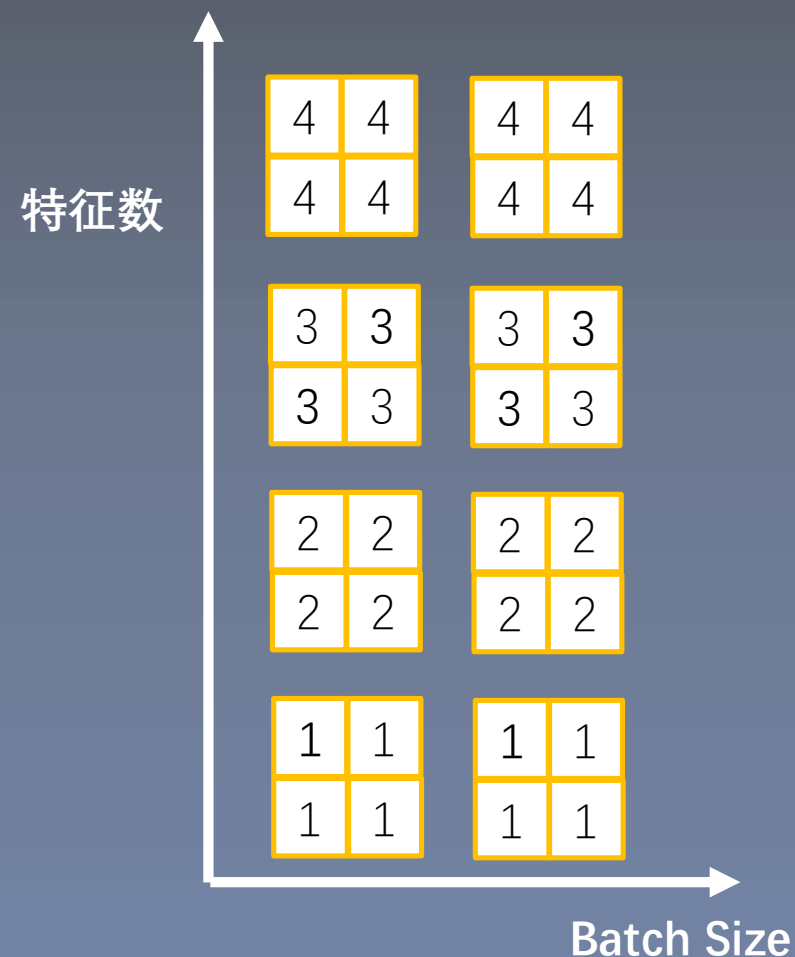
起因：小batch样本中，BN估计的值不准

思路：数据不够，通道来凑

注意事项：

1. 不再有running\_mean和running\_var
2. gamma和beta为逐通道（channel）的

应用场景：大模型（小batch size）任务





# Normalization in DL

## Normalization in DL

---

### nn.GroupNorm

主要参数：

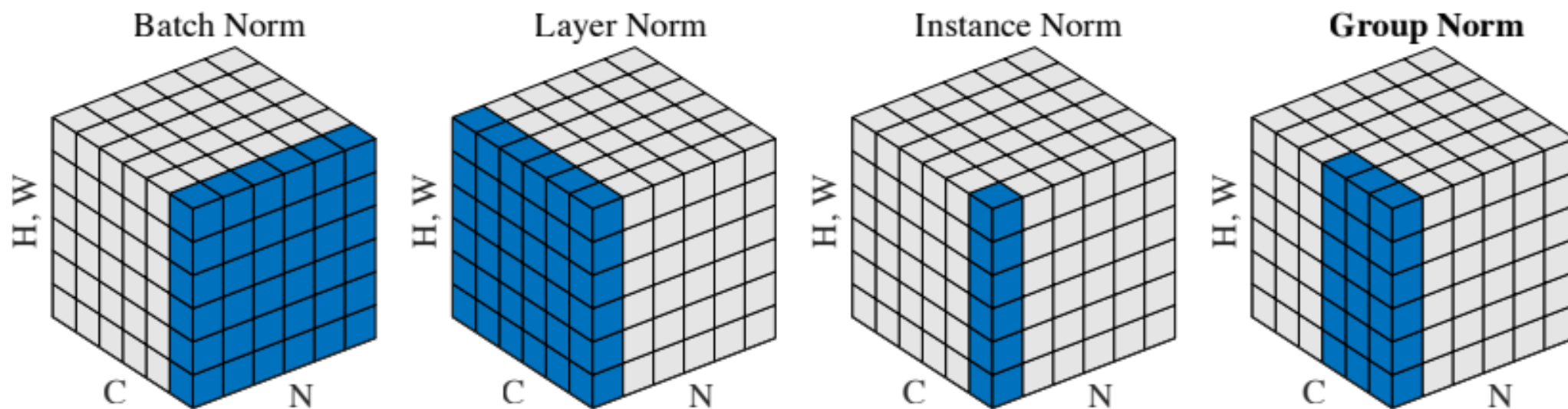
- **num\_groups**：分组数
- **num\_channels**：通道数（特征数）
- **eps**：分母修正项
- **affine**：是否需要affine transform

```
nn.GroupNorm(  
    num_groups,  
    num_channels,  
    eps=1e-05,  
    affine=True)
```

# Normalization in DL

Normalization in DL

小结：BN、LN、IN和GN都是为了克服**Internal Covariate Shift (ICS)**

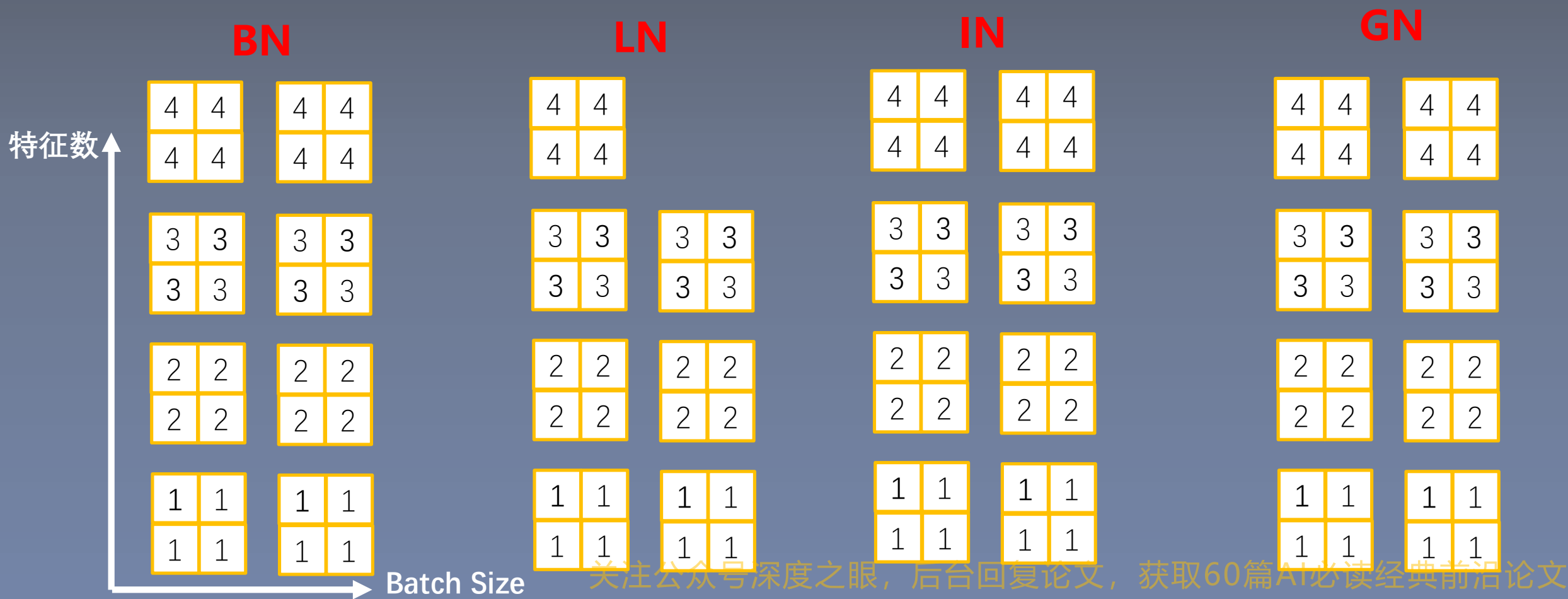




# Normalization in DL

## Normalization in DL

小结：BN、LN、IN和GN都是为了克服**Internal Covariate Shift (ICS)**



# —— 结 语 ——

在这次课程中，学习了深度学习中的Normalization

在下次课程中，我们将会学习

**模型序列化与加载**



关注公众号深度之眼，后台回复论文，获取60篇AI必读经典前沿论文





**deepshare.net**

深度之眼

联系我们:

电话: 18001992849

邮箱: [service@deepshare.net](mailto:service@deepshare.net)

QQ: 2677693114



公众号



客服微信

关注公众号深度之眼，后台回复论文，获取60篇AI必读经典前沿论文