

4-3 卷积神经网络 II

王中雷

厦门大学王亚南经济研究院和经济学院, 2025

内容摘要

1. LeNet5

2. AlexNet

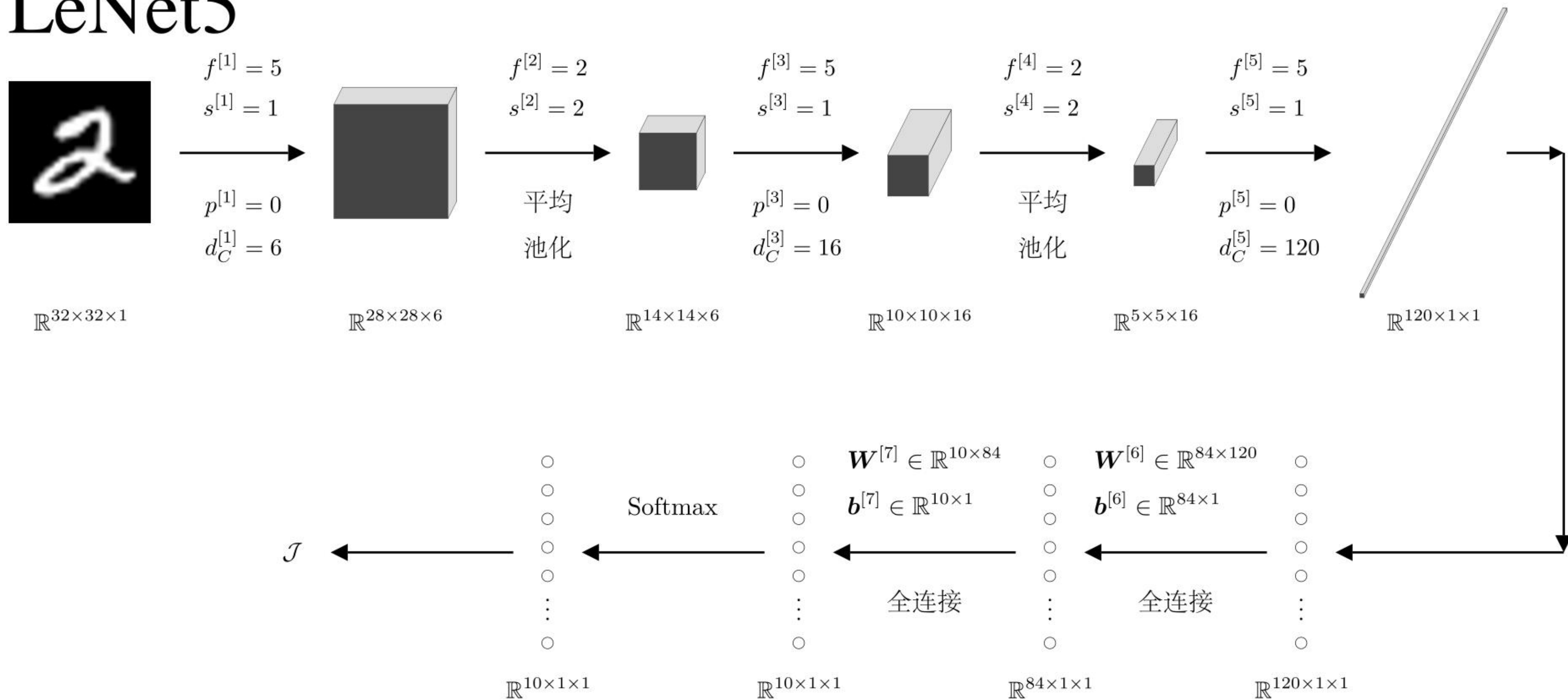
3. VGG16

LeNet5

1. LeNet5 于 1998 年被 LeCun 提出

- 被用于识别 32×32 灰度手写数字图片 (10 类分类问题)
- $\tanh$ 是隐藏层的激活函数
- 最早的几个卷积神经网络模型之一
- 在计算机视觉领域具有里程碑式的重要意义
- 名称中的“5”表示模型有 5 (或者 7) 层

LeNet5

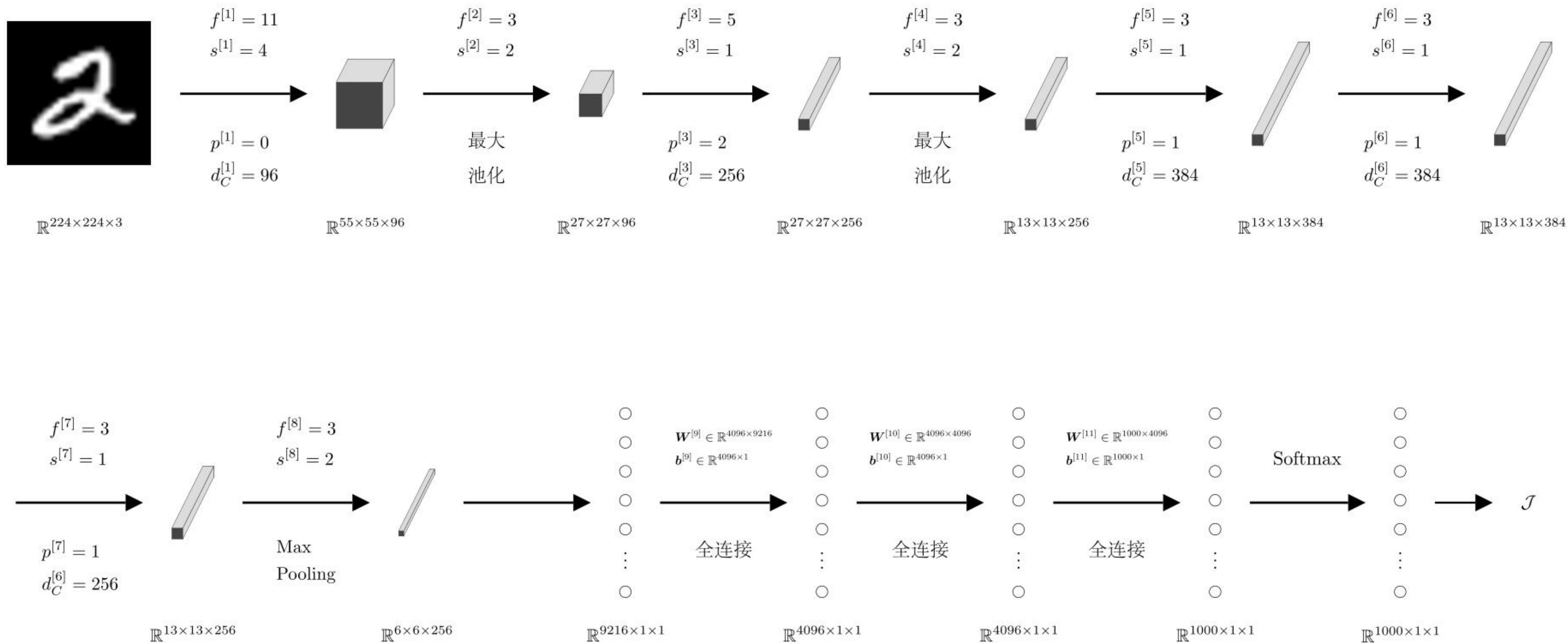


AlexNet

1. 计算机视觉领域的里程碑式的重要模型

- 在 ImageNet Large Scale Visual Recognition Challenge 2012 比赛中获得第一名
- 用于识别分辨率为 $224 \times 224 \times 3$ 彩色图片的类别（1000 类分类问题）
- 模型性能显著超越其他方法
- 极大促进了深度学习在计算机视觉领域的发展

AlexNet



VGG16

1. 计算机视觉领域的经典模型

- 赢得 ImageNet Large Scale Visual Recognition Challenge 2014 比赛第一名
- 用于识别分辨率为 $224 \times 224 \times 3$ 彩色图片的类别（1000 类分类问题）

2. 符号

- CONV[num]: 进行“num”次卷积运算，对应卷积核尺寸为 3×3 、步幅为1 以及填充为1
- POOL: 进行最大池化，对应的（池化）核尺寸为 2×2 、步幅为2

3. 注意

- 卷积运算**不改变**“输入图片”的尺寸
- 最大池化将“输入图片”的尺寸**缩小一半**

VGG16

