

《基于全卷积神经网络的图像分类》

Presented by meijun

答辩人：姚美君

指导老师：喻玲娟

01 选题背景

02 CNN和FCN原理

03 研究方法

04 总结分析

目录 CONTENTS

The background of the slide features a stylized, layered mountain range. The mountains are rendered in various shades of gray, with the foreground mountains being a dark charcoal gray and the background mountains becoming progressively lighter, creating a sense of atmospheric perspective. The peaks are soft and rounded, and the overall composition is minimalist and modern.

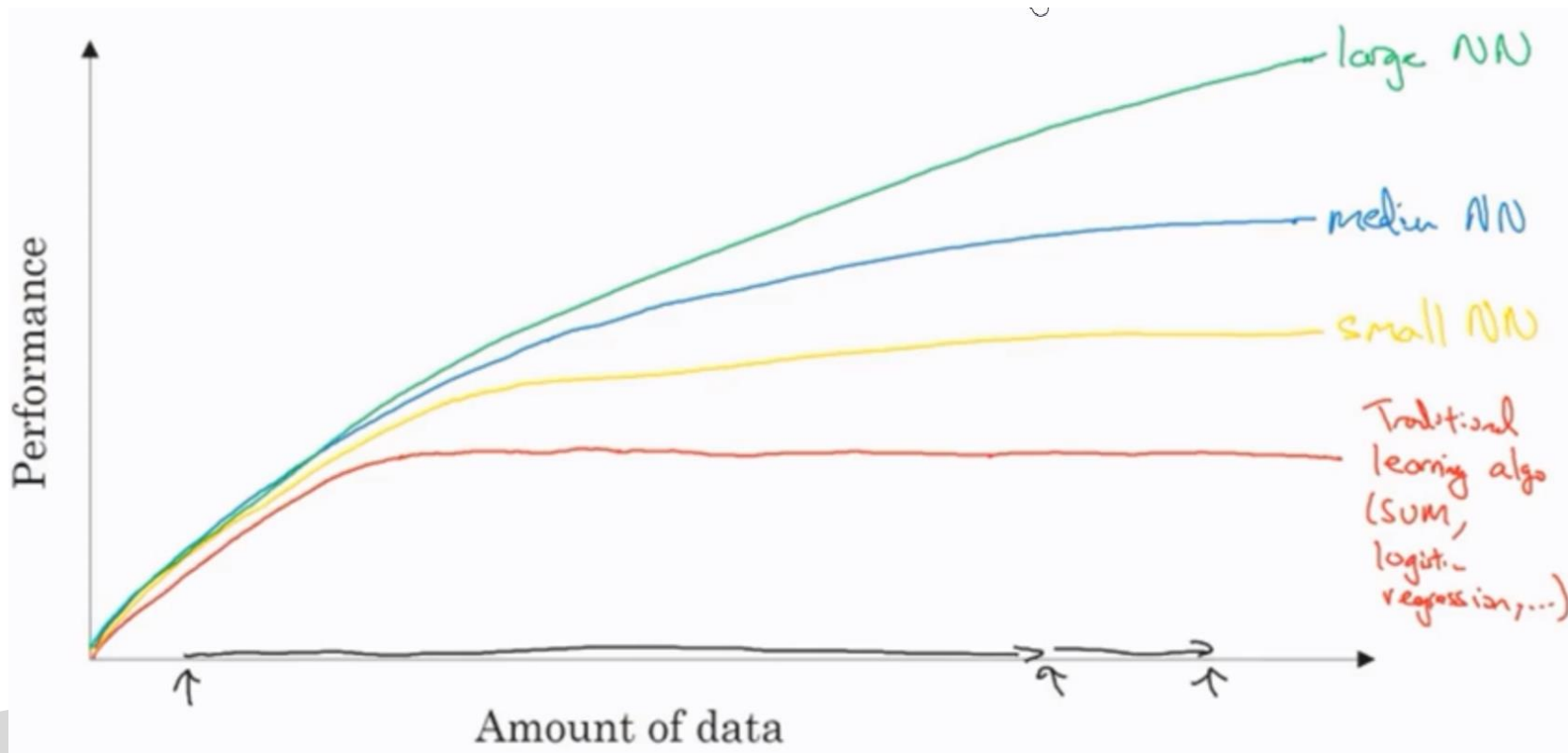
01

选题背景

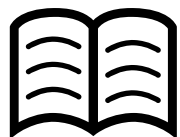
01

选题背景

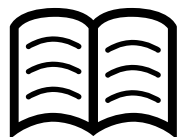
深度学习和神经网络的理念已经有几十年了，现在才流行起来



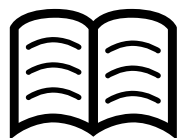
01 选题背景



目前，图像分类最常用的网络模型是卷积神经网络CNN。它由交替的卷积层和池化层，其次是少量的全连接层。大多数的优化方案也是基于此基础架构。



研究发现，用相同步长的卷积层替代池化层之后，精确度并无损失。以找到更简单的适用于图像分类的神经网络架构为目标，研究仅有卷积层组成的全卷积神经网络FCN。

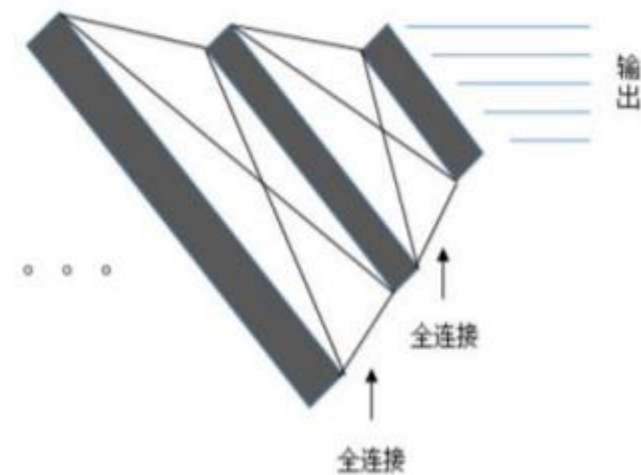
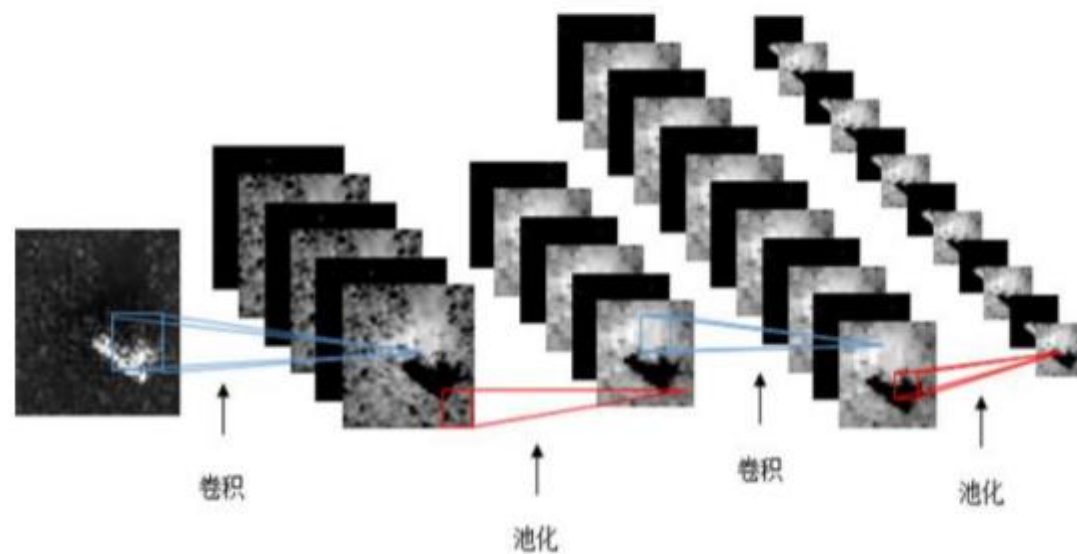
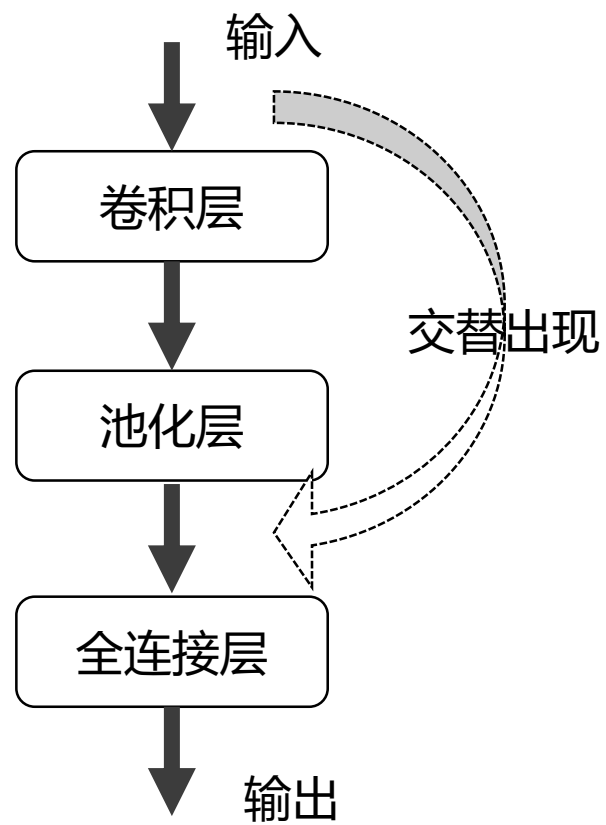


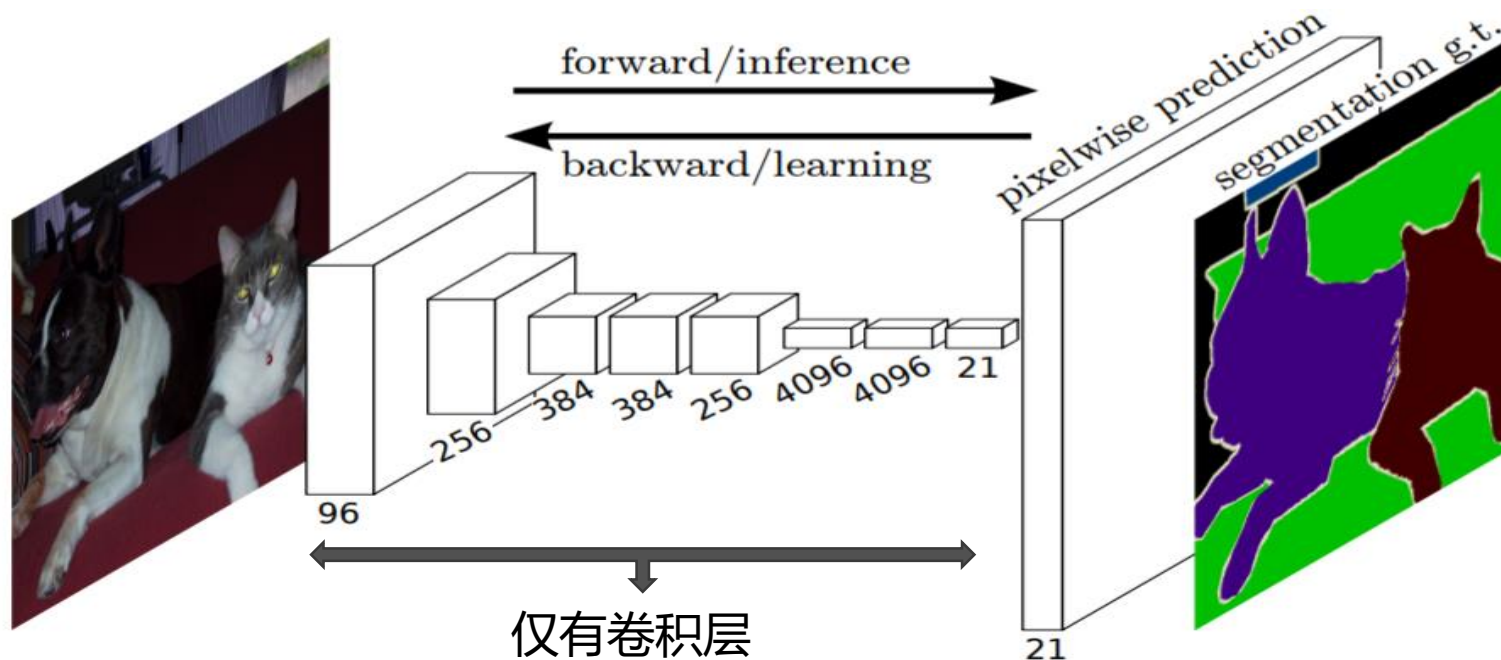
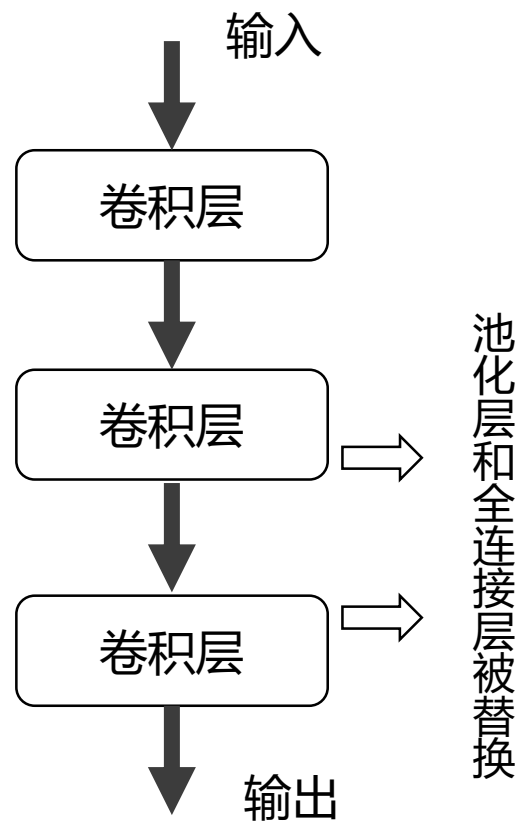
目标，重新评估CNN中各个基本组件的必要性，从两者在不同大小的数据集上的表现，在时间效率和精确度两个方面，研究和分析基于FCN的图像分类。

A stylized, layered mountain range graphic in shades of gray and black, spanning the top half of the image. The mountains are composed of several overlapping, wavy horizontal bands of varying heights and shades, creating a sense of depth and texture. The topmost layer is a light gray, while the bottom layer is a dark gray, with intermediate layers in various shades of gray.

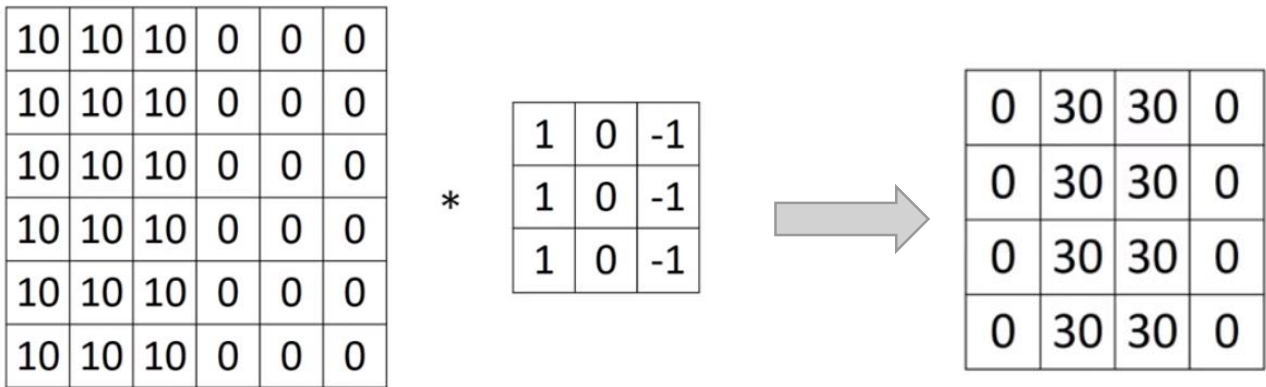
02

CNN和FCN原理





边缘检测:



卷积层



Conv1:filter size 5x5x3, stride 1



Conv2:filter size 5x5x3, stride 1



Conv3:filter size 5x5x20, stride 1



最大池化:

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

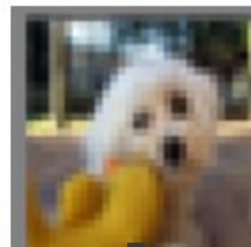
max pool with 2x2 filters
and stride 2

6	8
3	4

池化层

f: 过滤器大小
p: 补零
s: 步长

conv: $f=5*5$ $p=2$ $s=1$



Size: 32x32



Size: 32x32

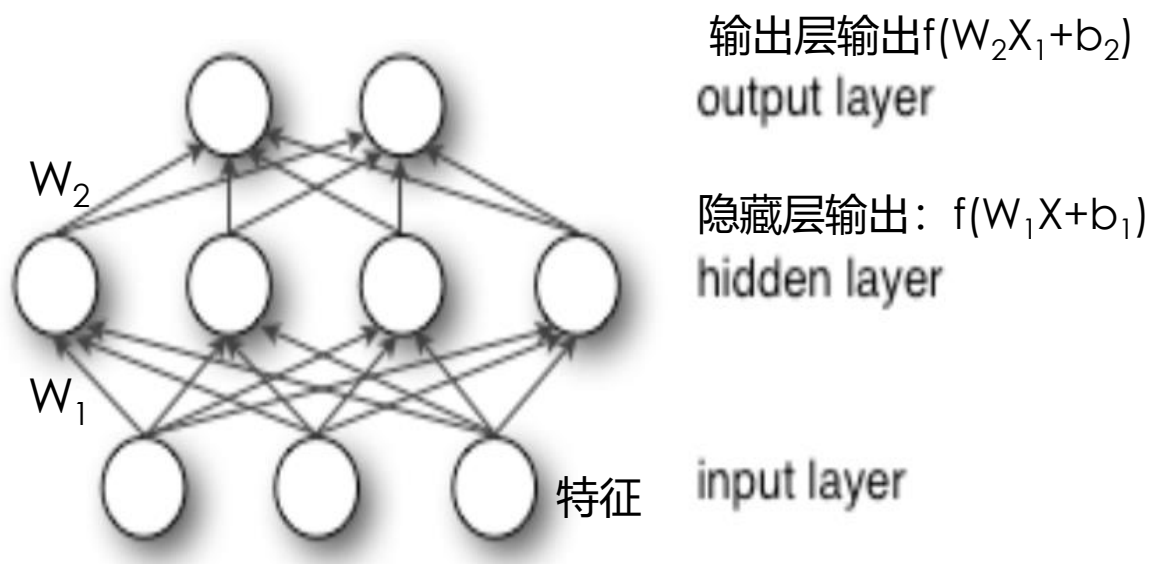
pool: $f=2*2$ $s=2$

池化



Size: 16x16

最简单的全连接层：仅一层隐藏层

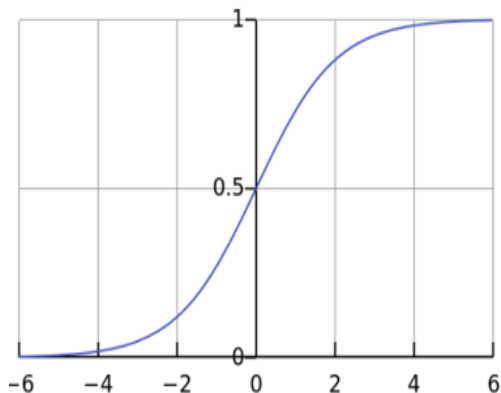


全连接层表达式: $f(x) = G(W_2 * f(W_1X + b_1) + b_2)$

f 为激活函数: sigmoid、tanh、relu; 若为多分类任务, G 通常为softmax

- ①初始化 W 和 b
- ②输入特征, 通过如全连接层表达式的运算过程, 得到预测值 y
- ③通过梯度下降(迭代过程)寻找使损失函数 J 最小的 W 和 b

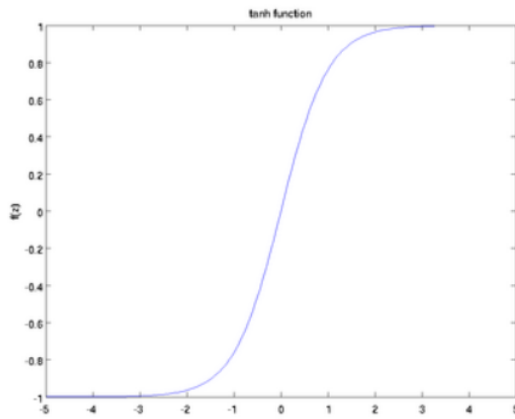
Sigmoid



$$x \in R$$

$$0 < y < 1$$

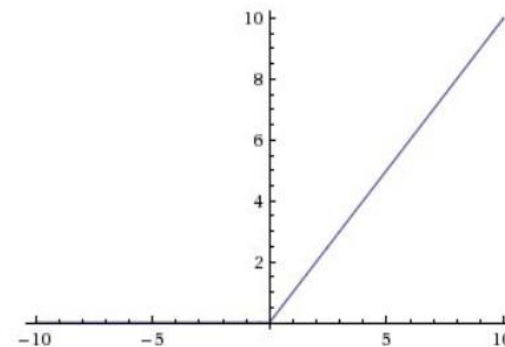
tanh



$$x \in R$$

$$-1 < y < 1$$

relu



$$x \in R$$

$$y = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

02 CNN和FCN原理

优化器(optimizer)

SGD(默认小批量梯度下降)

RMSProp

Adagrad

Adam

Adamax

.....

损失函数(loss function) 凸函数：全局最优解

均方误差 (MSE) :

$$\frac{1}{n} \sum_{i=1}^n (y_{pred} - y_{true})^2$$

平均离差:

$$\frac{1}{n} \sum_{i=1}^n |y_{pred} - y_{true}|$$

交叉熵:

$$-\frac{1}{n} \sum_{i=1}^n y_{true} \log(y_{pred}) + (1 - y_{true}) \log(1 - y_{pred})$$

.....

02 CNN和FCN原理

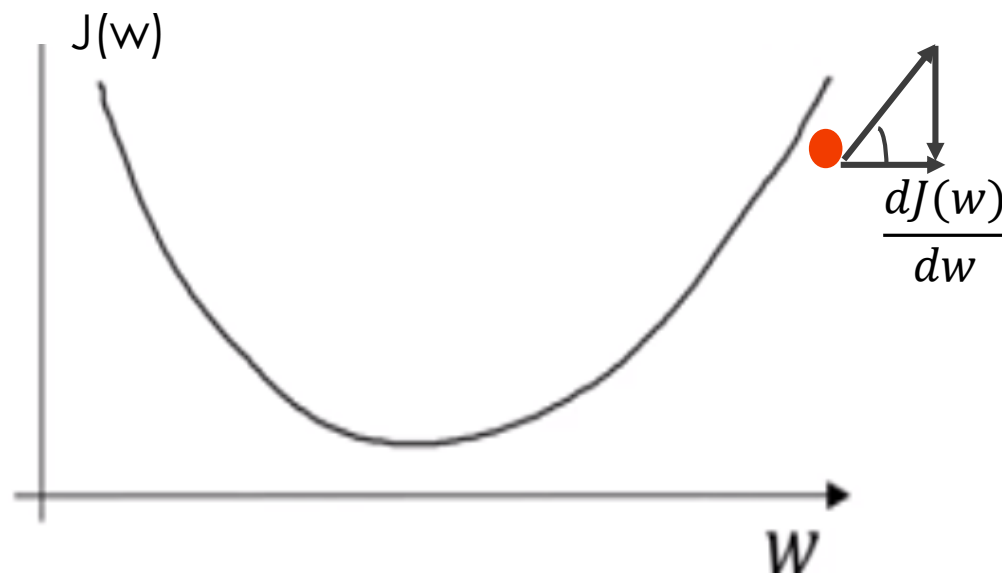
梯度下降(Gradient descent):最快下降的方向

损失函数(loss function) 均方误差 (MSE)

$$\text{Repeat } \left\{ w := w - \alpha \frac{dJ(w,b)}{dw} \right\}$$

$$\text{Repeat } \left\{ b := b - \alpha \frac{dJ(w,b)}{db} \right\}$$

α 学习率: 决定梯度下降法中的步长;



02 CNN和FCN原理

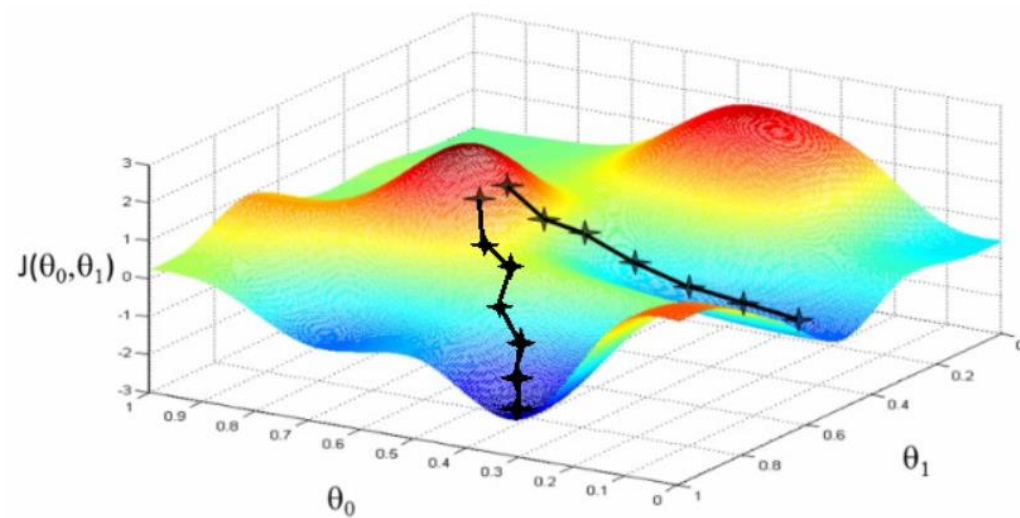
梯度下降(Gradient descent):最快下降的方向

损失函数(loss function) 均方误差 (MSE)

$$\text{Repeat } \left\{ \theta_0 := \theta_0 - \alpha \frac{\sigma J(\theta_0, \theta_1)}{\sigma \theta_0} \right\}$$

$$\text{Repeat } \left\{ \theta_1 := \theta_1 - \alpha \frac{\sigma J(\theta_0, \theta_1)}{\sigma \theta_1} \right\}$$

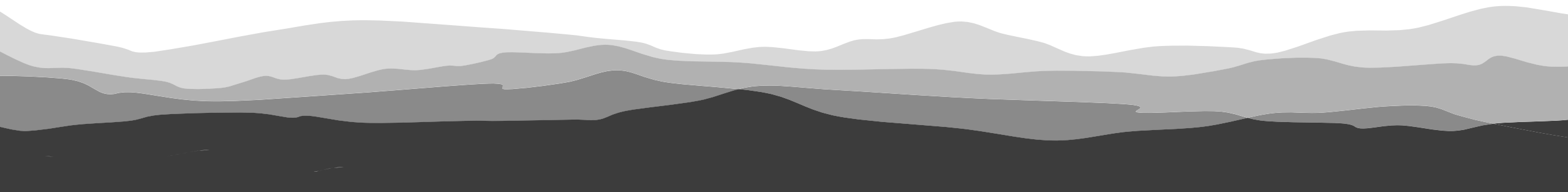
α 学习率: 决定梯度下降法中的步长;



Dropout: 训练过程中丢弃参数（防止过拟合）

Flatten: 将多维阵列展平成一维向量（全连接层前）

Softmax: 多分类任务的全连接层的输出层常使用的激活函数



A stylized, layered mountain range graphic in shades of gray and black, spanning the top half of the slide. The mountains are composed of several overlapping, wavy horizontal bands of varying heights and shades, creating a sense of depth and texture. The topmost layer is a light gray, while the bottom layer is a dark gray, with intermediate layers in various shades of gray.

03

研究方法

03 研究方法

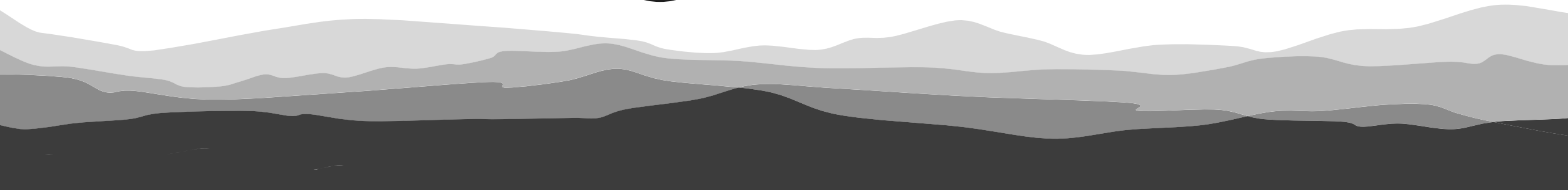
编程语言: python

深度学习框架: keras

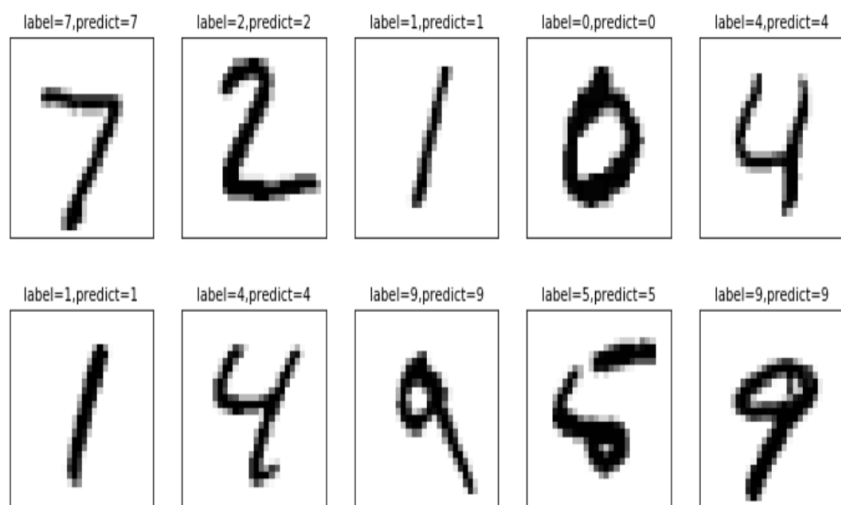


编译器: jupyter notebook

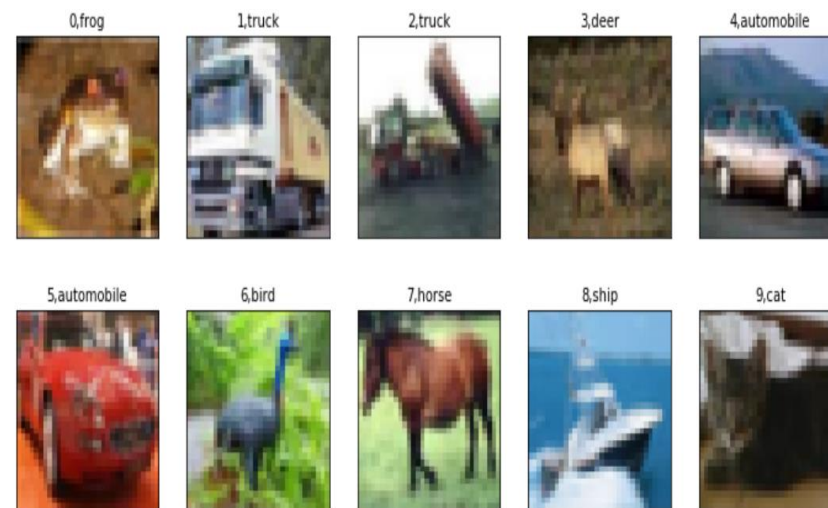
算法: CNN、FCN



Mnist: 手写数字图像集



Cifar-10: 飞机、鸟、猫...



mnist和cifar-10数据集对比

	Mnist 数据集	Cifar数据集
Size	11.5MB	175.5MB
颜色格式	灰度图像	RGB图像
数据内容	手写数字图像	飞机、车等十类图像
总数据样本数	60000	60000
训练集样本数	50000	50000
测试集样本数	10000	10000

Mnist数据集：CNN和FCN实验模型

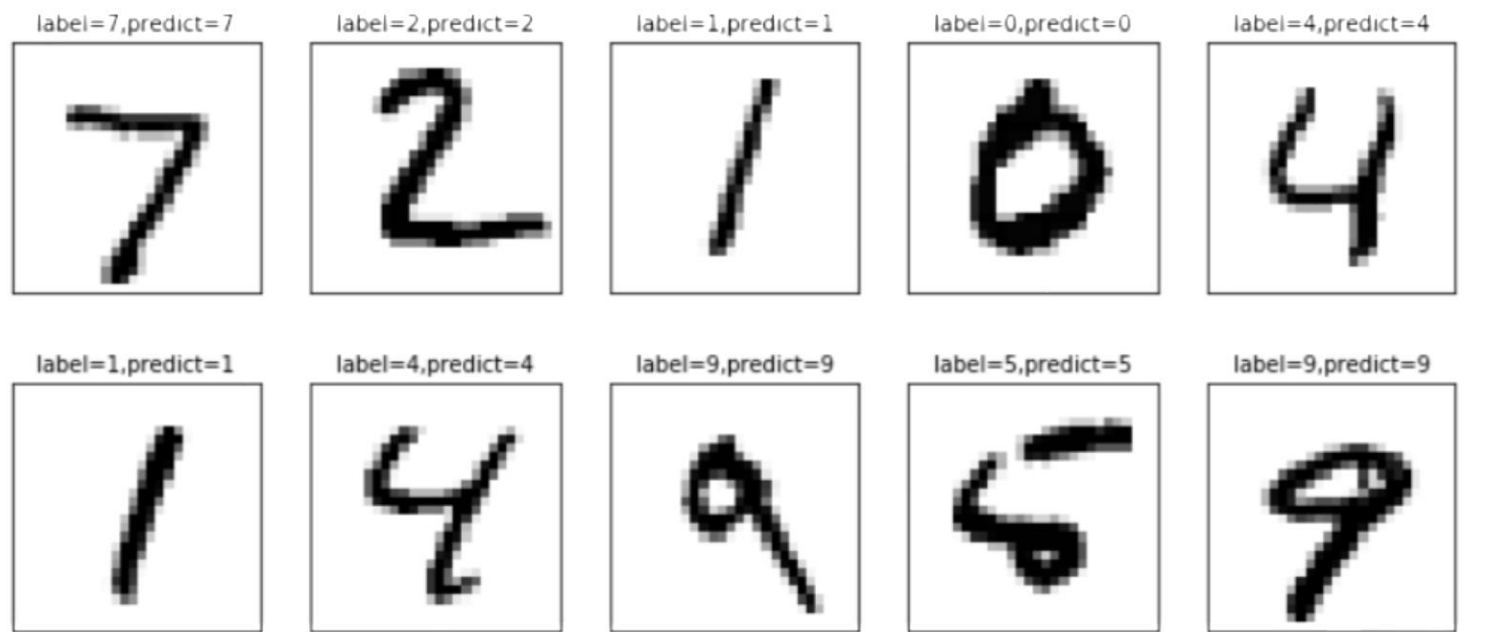


CNN

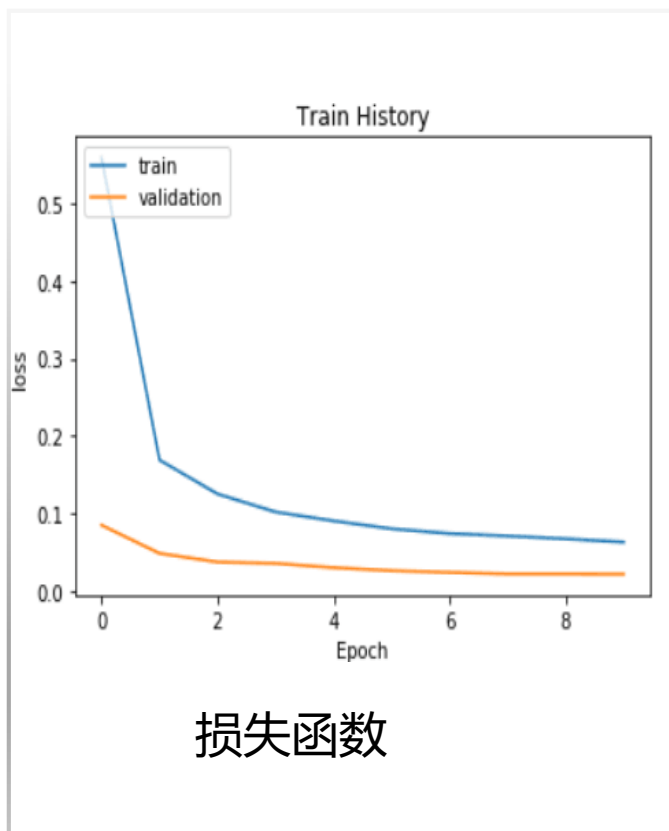
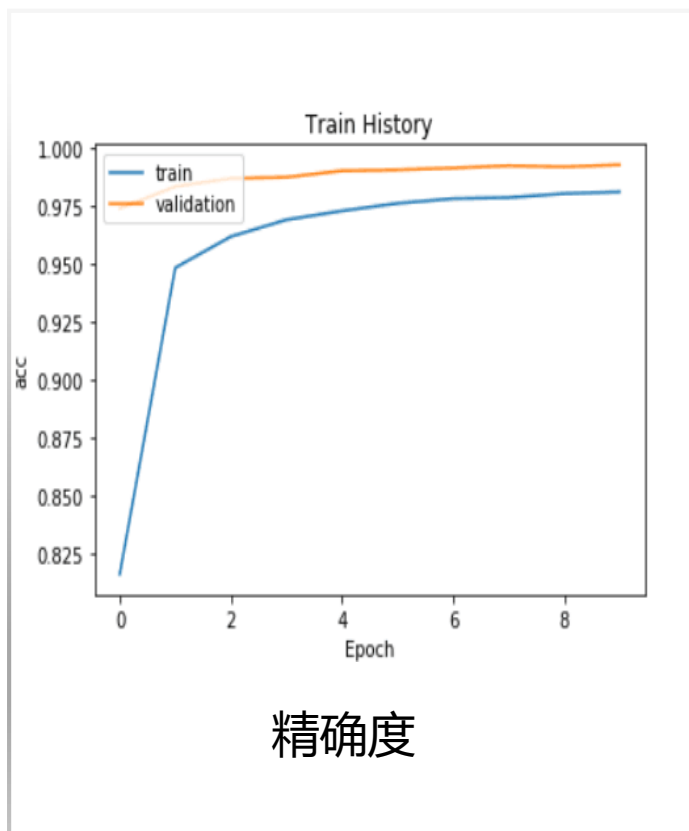


FCN

Mnist 预测结果展示



Mnist / FCN 精确度和混淆矩阵



predict	0	1	2	3	4	5	6	7	8	9
label										
0	977	0	0	0	0	0	1	1	1	0
1	0	1131	0	0	0	0	1	1	2	0
2	1	0	1024	0	2	0	0	3	2	0
3	0	0	0	999	0	4	0	1	4	2
4	0	0	0	0	975	0	1	0	2	4
5	1	0	0	3	0	884	1	0	1	2
6	3	2	0	0	2	1	942	0	8	0
7	0	3	3	0	2	0	0	1015	0	5
8	2	0	0	0	0	0	0	0	970	2
9	0	0	0	0	5	1	0	1	7	995

混淆矩阵

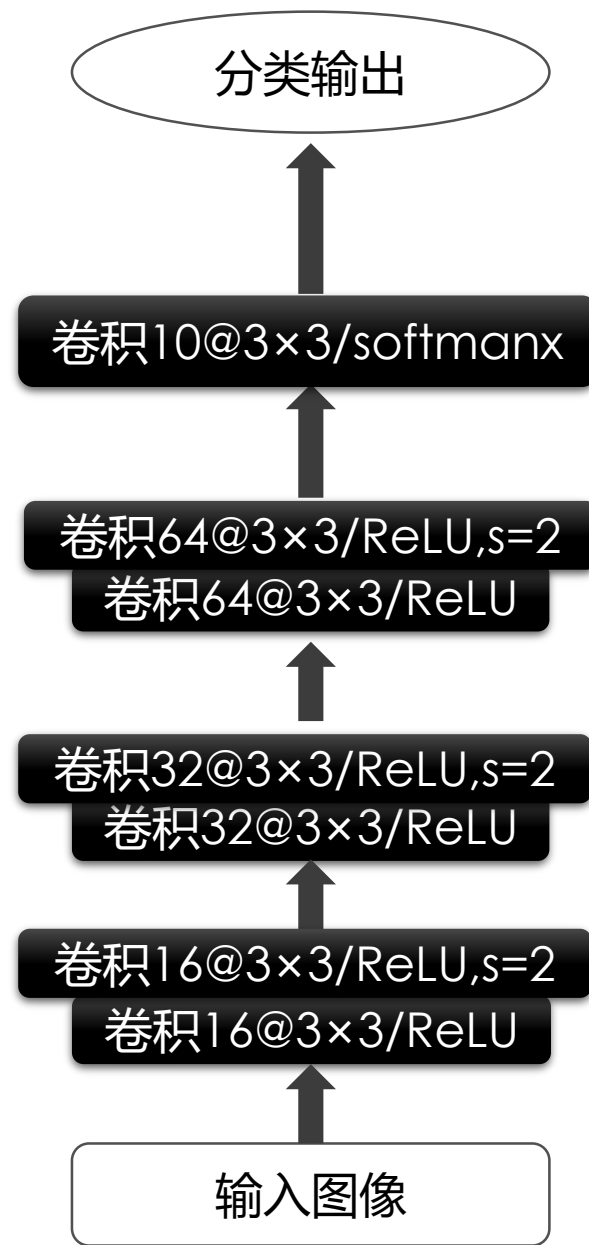
Mnist数据集：CNN和FCN实验结果对比

	FCN	CNN
时间	410s	490s
参数params	74362	98442
损失函数loss	0.0317	0.0632
训练集精确度acc	0.9901	0.9807
测试集损失函数val_loss	0.0230	0.0229
测试集精确度val_acc	0.9925	0.9922

Cifar-10数据集：CNN和FCN实验模型



CNN

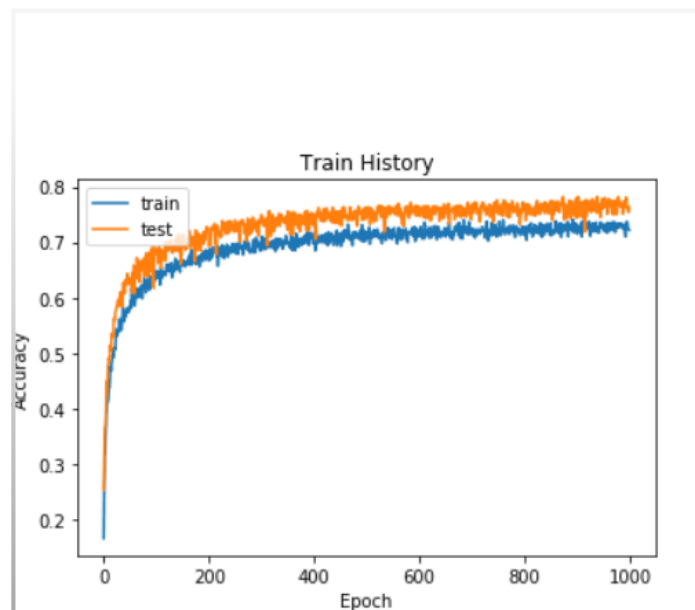


FCN

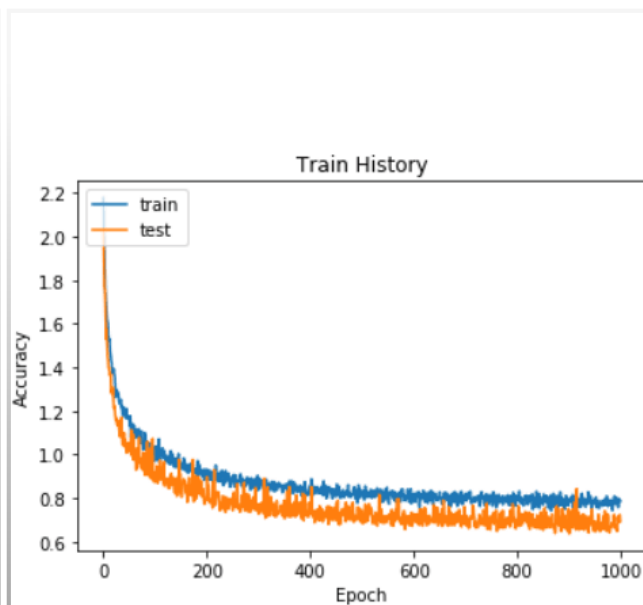
Cifar-10 预测结果展示



Cifar-10 / FCN精确度和混淆矩阵



精确度



损失函数

predict	0	1	2	3	4	5	6	7	8	9
label										
0	659	226	30	0	0	6	0	3	13	63
1	8	948	0	0	0	0	0	0	1	43
2	152	322	145	0	0	207	16	22	21	115
3	76	284	20	2	1	425	4	22	39	127
4	86	308	24	3	11	295	25	61	29	158
5	37	214	17	2	0	603	2	26	15	84
6	30	443	13	3	2	254	75	15	39	126
7	68	220	9	0	2	201	2	290	21	187
8	330	271	11	0	0	3	0	1	250	134
9	51	306	0	0	0	13	0	3	6	621

混淆矩阵

Cifar-10数据集：CNN和FCN实验结果对比

	FCN	CNN
时间	8000s	10000s
参数params	79184	553514
损失函数loss	0.7853	0.8370
训练集精确度acc	0.7219	0.7100
测试集损失函数val_loss	0.6935	0.7375
测试集精确度val_acc	0.7632	0.7512

A stylized, layered mountain range graphic in shades of gray and black, spanning the top half of the slide. The mountains are composed of several overlapping, wavy horizontal bands of varying heights and shades, creating a sense of depth and texture. The topmost layer is a light gray, while the bottom layer is a dark gray, with intermediate layers in various shades of gray.

04

总结分析

精确度

Mnist 数据集，FCN和CNN的精确度相差不大，FCN略高一点点。

Cifar数据集，FCN相比于CNN，模型精确度表现更优异。

时间效率

Mnist 数据集，FCN和CNN中的参数数量，没有很大的区别，在时间效率上并无较大的区别。

Cifar 数据集，CNN中的参数数量相比于FCN有了一个数量级的增加，FCN的时间上的效率远远高于CNN。

卷积层参数是局部连接，全连接层中是全部连接，卷积可以使训练参数更少，从而时间效率更高。

- (1) 如果数据集比较小：CNN和FCN都能表现的不错。
- (2) 若数据集较大，选择FCN能让你的模型在精确度是时间效率上都更有优势。

CNN和FCN算法

应用场景分析

感谢聆听！

《基于全卷积神经网络的图像分类》

Presented by meijun

答辩人：姚美君

指导老师：喻玲娟

