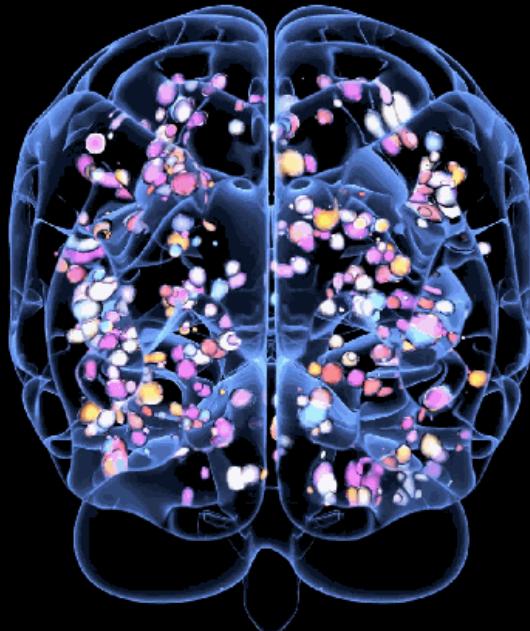


2024年
秋季学期



人工
神经
网络
作业

◎ 说明： 成作业可以使用你所熟悉的编程语言和平台，比如 C, C++、 MATLAB、 Python等。作业链接。

01 深度网络

一、作业内容

1、作业要求

- 练习搭建深度学习网络基本网络（CNN）实现数据分类与参数回归；
- 掌握深度学习训练中基本方法：
 - 训练集、测试机、验证集合的划分；
 - 训练过程中的超参调试：学习速率、样本批次数量、归一化处理
 - 对于样本增强、正则参数调整、Dropout选取等
- 熟悉几大类深度网络特性以及应用；

任意选择后面给出的六种数据集合中的两个数据集合，按照数据集合要求，完成数据分类或者参数回归。

2、实验平台

搭建深度神经网络完成作业，可以选择自己熟悉的实验环境完成，包括但不限于：

- 百度的 AI Studio 2021 ANN课程，发布算力卡；
- MATLAB
- Keras, Pytorch
- 华为，阿里，Google AI云平台等

(1) AI Studio

请大家根据网络学堂公布的AI Studio 邀请码进入课堂，可以使用到百度提供的算力卡以及获得作业中的AI Studio工程数据集合。



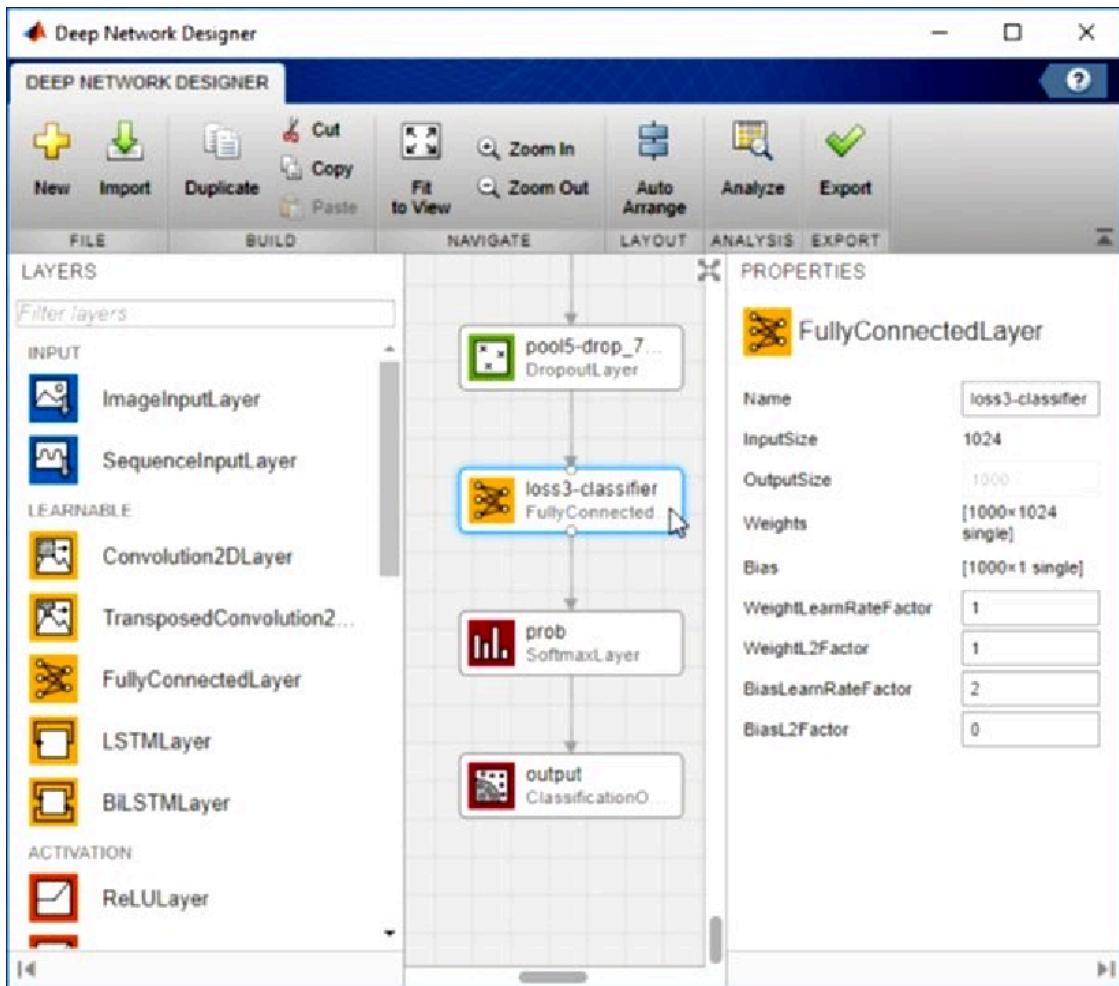
▲ 图1.2.1 AI Studio 学习界面

大家进入AI Studio之后可以根据其中相关文档自行联系搭建基本深度学习网络，来完成作业要求。

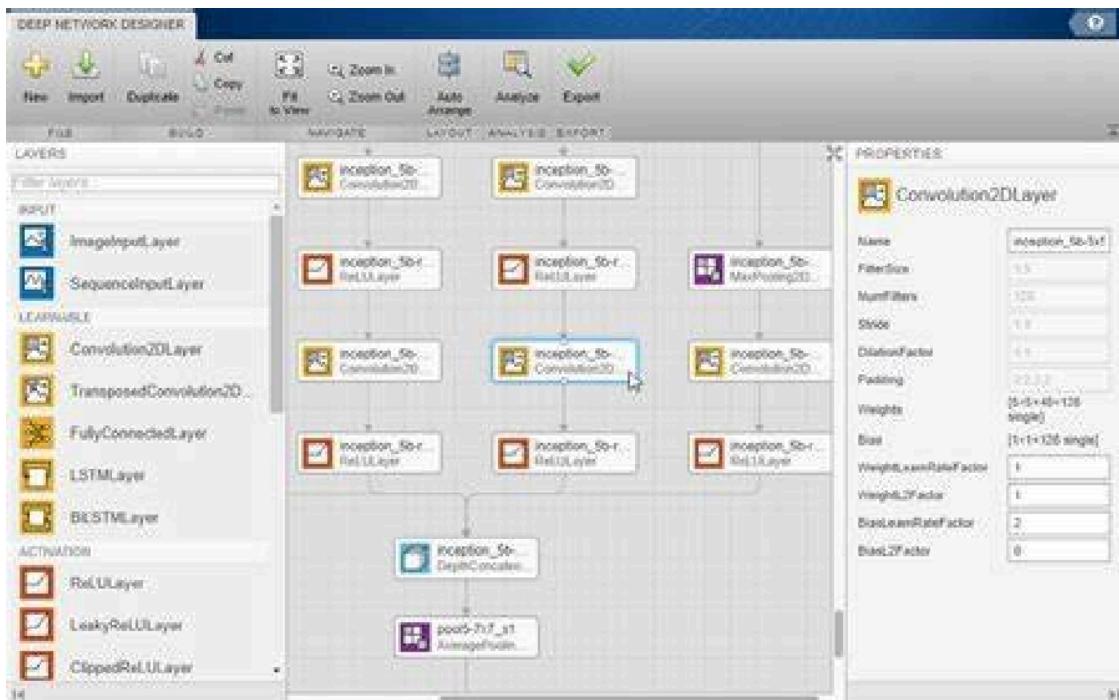
(2) MATLAB

对于没有PYTHON基础的同学可以使用MATLAB 2019b完成作业。

在学校内可以免费安装2019b之后版本的MATLAB软件，详细的安装说明可以参见 **MATLAB深度学习2019** 中的相关介绍。



▲ 图1.2.3 MATLAB 中深度学习网络设计界面



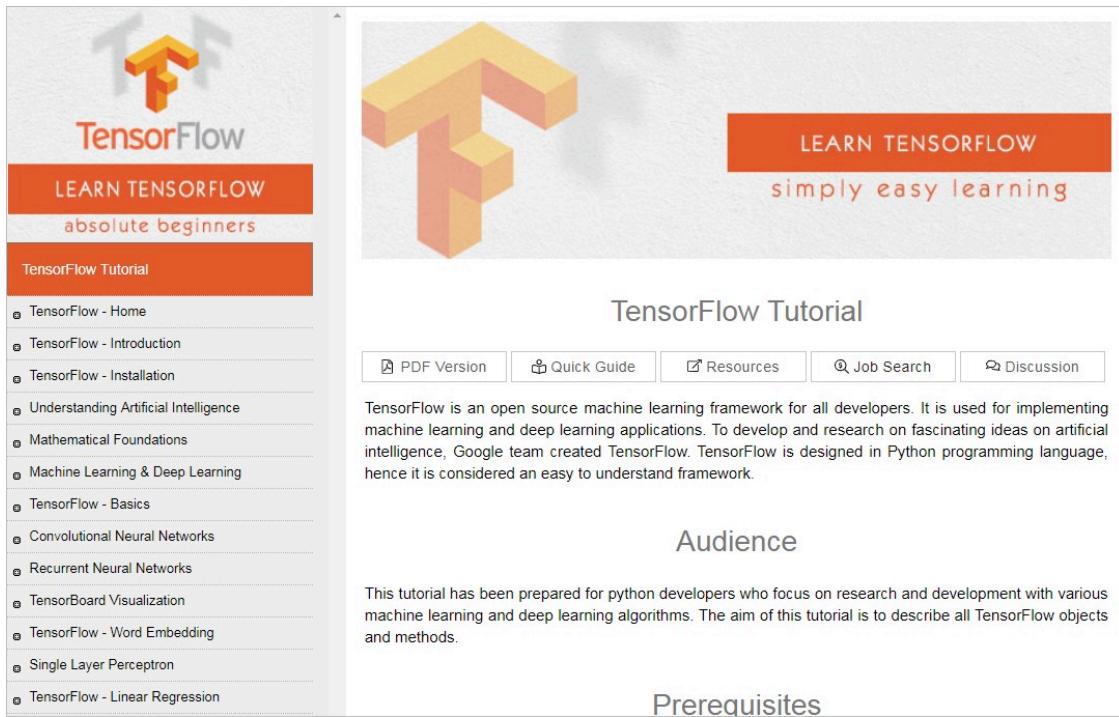
▲ 图1.2.4 在MATLAB中搭建深度学习网络

(3) Python

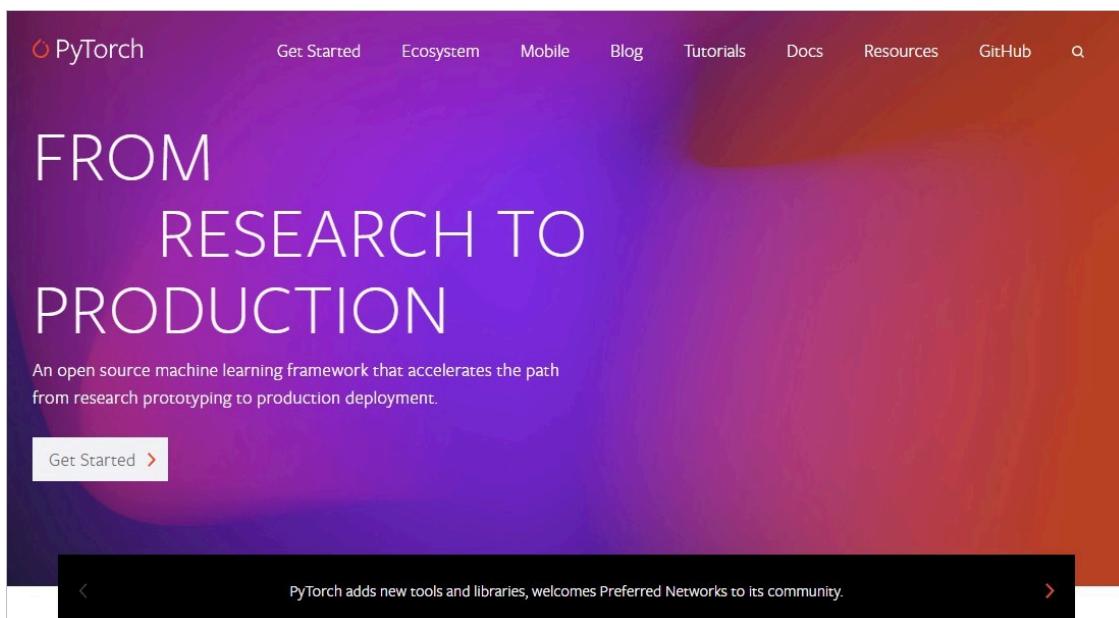
具有python基础的同学可以参考《Python核心编程（第二版）》，TensorFlow、Pytorch和Keras三个深度学习框架。

对应的文档连接分别为：

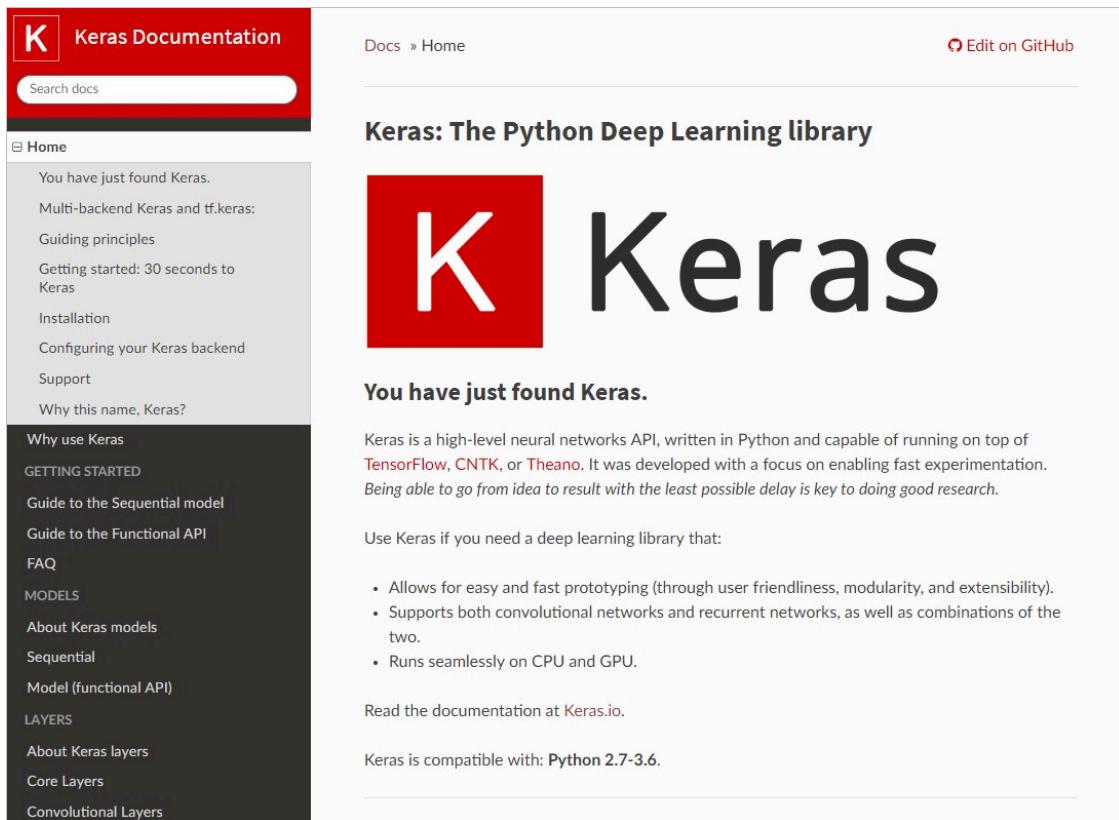
- https://www.tensorflow.org/api_docs/
- <https://pytorch.org/docs/stable/index.html>
- <https://keras.io/>



▲ 图1.2.5 TENSORFLOW开发深度学习网络



▲ 图1.2.6 pyTorch开发深度学习网络



▲ 图1.2.7 Keras 开发深度学习网络

3、数据库以及下载

作业中所需要的数据库，可以在[课程文件](#)中下载第五次作业数据文件。也可以在AI Studio中找到对应的作业文件。

实验数据总共有六组实验数据：

- 水果与动物数据：图片目录压缩文件；
- MNIST数据； MATLAB数据格式
- CIFAR10数据集合； MATLAB数据格式
- 转动数字数据集合； MATLAB数据格式
- 危险品数据基集合； MATLAB数据格式
- AI Studio工程数据集合； AI Studio中数据文件

前五种数据集合可以在网络学堂的“[课程文件](#)”中下载，其中MNIST，CIFAR10,转动数字，危险品数据是MATLAB数据格式。如果使用其它深度学习平台，可以从MATLAB读取并转存成所需要的格式。

下面是网络学堂中课程文件下载的第五次作业数据集合解压缩之后的文件目录，其中包括有五个数据集合。

```
|---CIFAR-10  
|---FORBIDDEN  
|---FruitAnimal  
|---MNIST  
|---ROTATEDIGIT
```

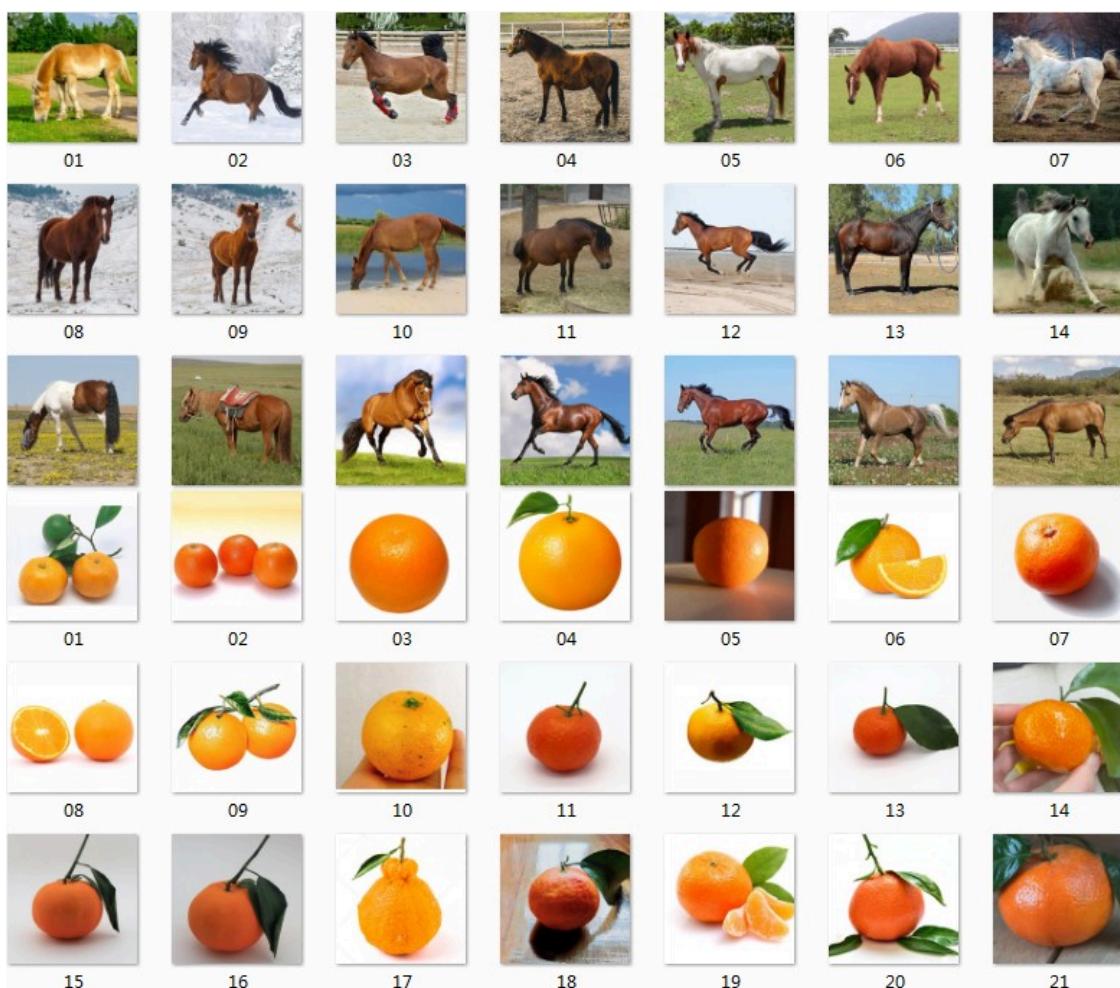
二、数据库

1、水果与动物

(1) 数据集合简介

请设计一个CNN，完成对于给定水果（五类）和动物（五类）总共十类物体的识别。

数据库还是第一次作业中，请大家设计BP网络识别所使用的动物，水果数据集合。只是在第一次作业中，要求仅仅是将Apriltag，数字，动物，水果这四种物品分出所属的大类。但本次作业中只使用其中的动物、水果两大类，供十小类的物体，要求识别它们的小类别。



▲ 图2.1.1 水果与动物数据库的样本

数据库图片可以由网络学堂数据文件下载。也可以在AI Studio的公开数据集合中下载：**五种水果与五种动物数据集合**：

<https://aistudio.baidu.com/aistudio/datasetdetail/120001>



▲ 图2.1.2 AI Studio 数据集合

下载文件之后，加压缩会得到两个目录：**水果、动物**。它们各自包括五个子目录，其中分别存放着大约100张左右的水果、动物图片，用于训练和测试。

```
|---FruitAnimal
|   |---动物
|   |   |---动物
|   |   |   |---牛
|   |   |   |---狗
|   |   |   |---猪
|   |   |   |---猫
|   |   |   |---马
|   |---水果
|       |---水果
|           |---榴莲
|           |---橙子
|           |---苹果
|           |---葡萄
|           |---香蕉
```

(2) 数据处理预处理

在训练网络之前，大家可以对数据库进行预处理和增强；

II.图片预处理

- **尺寸归一化**：数据库中的图片尺寸已经归一在 283×283 ，可以自行选择是否做进一步的调整图片尺寸；

III.数据库增强

为了提高训练之后的网络泛化能力，可以选择适当的图片处理方法，对于训练数据集合进行增强：

- 图片旋转

- 图片亮度调整
- 图片色度调整
- 图片尺度变化
- 图片增噪声

(3) 作业要求

1. 构建深度学习网络，对于水果动物数据集合完成小类别识别任务；
2. 讨论在不同的网络结构，参数数量对于网络识别精度的影响；
3. 讨论与对比网络训练超参、数据库增强对于网络识别精度的影响；
4. 寻找一个兼顾网络模型大小与识别性能的网络，即模型参数少，识别性能下降不多的网络。

(4) 参考答案

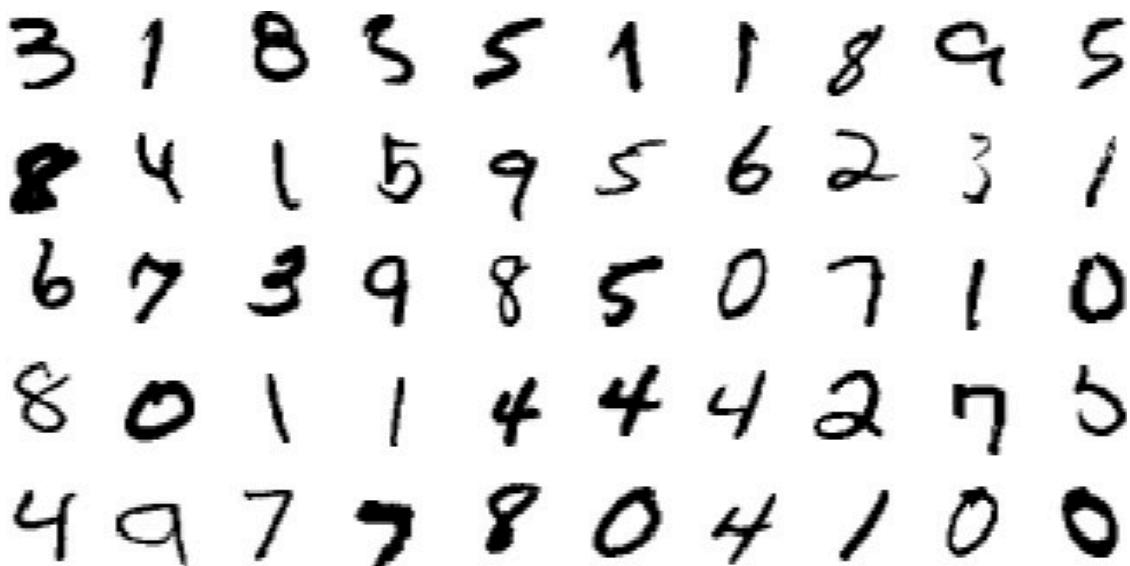
对于第一题的参考答案可以参见如下两篇博文：

- [2021年人工神经网络第四次作业-第一题：LeNet对于水果与动物进行分类](#)
- [对于FAMNIST中的十种动物和水果进行识别测试](#)

2、MNIST

(1) 数据库简介

MNIST数据集合是手写体数字识别图片，图片是尺寸为 28×28 的灰度图片。



▲ 图3.1.1 mnist手写体数字图片示例

MNIST数据库官方下载：[THE MNIST DATABASE of handwritten digits](#)。

在网络学堂中给出了MATLAB格式的MNIST数据库。使用MATLAB命令
load('mnist.mat')可以将附件数据库：

MNIST下的数据文件调入MATLAB。其中包含有四个变量：

(1) **test_images**: 10000×784 int32

- 测试集合10000张手写体数字图片，图片大小为28×28，
- 灰度图片，取值范围(0 ~ 255).

(2) **test_labels**: 1×10000 int32

- 测试集合数字图片的标签：取值0 ~ 9

(3) **train_images**: 60000×784 int32

- 训练集合60000张数字手写体图片，格式与测试集合相同。
-

(4) **train_labels**: 1×60000 int32

- 训练集合数字图片的标签：取值0 ~ 9

该数据集合在各大深度学习框架内基本上都有相应的连接可以直接下载，大家也可以直接使用。

(2) 作业要求

① 构造两类深度学习网络对MNIST数据集合完成数字分类任务。一类是由全连接层构成的深度学习网络；另一类是带有卷积层的神经网络。

② 讨论两类神经网络在不同的网络参数下的训练收敛情况、测试精度的变化情况。

③ 探索寻找到一个能够兼顾网络模型大小与识别性能相对优化的网络模型，即模型参数少，识别性能下架不多。

(3) 参考答案

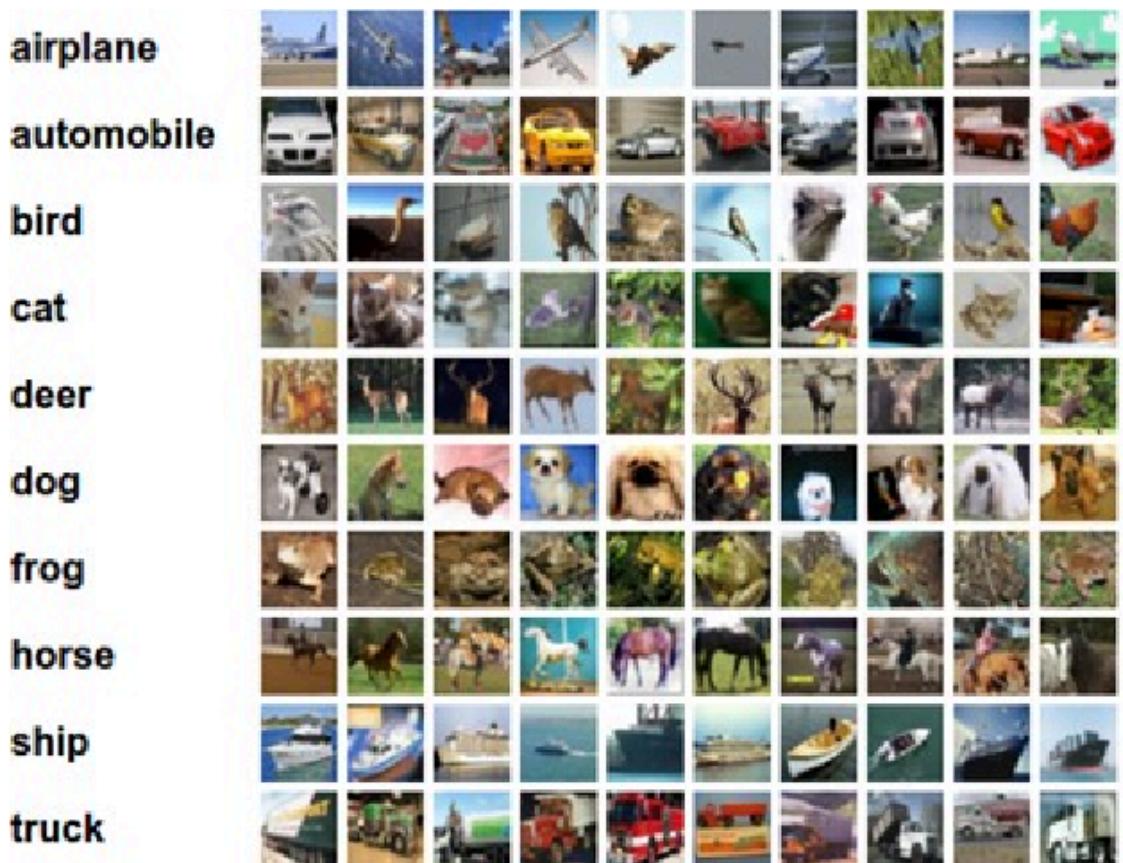
如果在PaddlePaddle中完成作业，可以看一下实验指导：

- [2021年人工神经网络第四次作业 - 第二题MNIST手写体识别](#)

3、CIFAR-10

(1) 数据库简介

对CIFAR-10数据集的分类是机器学习中一个公开的基准测试问题，其任务是对一组大小为32x32的RGB图像进行分类，这些图像涵盖了10个类别：飞机，汽车，鸟，猫，鹿，狗，青蛙，马，船以及卡车。



▲ 图4.1.1 CIFAR-10数据库

CIFAR-10数据集合可以从如下网站下载：<http://www.cs.toronto.edu/~kriz/cifar.html>

在CIFAR-10数据目录中包括有以下六个数据文件：

- (1) data_batch_1.mat
- (2) data_batch_2.mat
- (3) data_batch_3.mat
- (4) data_batch_4.mat
- (5) data_batch_5.mat
- (6) test_batch.mat
- (7) batches.meta.mat:label_names,各类别的名称

以上数据集合在调入MATLAB之后，都会包括有：

- data:10000×3072 uint8:图片信息
- labels:10000×1 uint8

在MATLAB中可以使用如下命令显示图片的内容：

```
imshow(permute(reshape(data(1,:),[32,32,3]),[2,1,3]))
```

数据命令中这里的参数'1'是对应第一个图片，可以将其修改成其它1~10000之间的数字来显示其它的图片。

想了解更多信息请参考CIFAR-10 page，以及Alex Krizhevsky写的技术报告

(2) 作业要求

① 建立一个用于识别图像的相对较小的卷积神经网络，完成CIFAR-10数据集合的识别问题。

- 建立一个规范的网络结构，训练并进行评估；
- 为建立更大规模更加复杂的模型提供一个范例

选择CIFAR-10是因为它的复杂程度足以用来检验TensorFlow中的大部分功能，并可将其扩展为更大的模型。与此同时由于模型较小所以训练速度很快，比较适合用来测试新的想法，检验新的技术。

② 大家可以参见网络上的“CIFAR-10 教程”来讨论卷积神经网络中如下各部分参数对于卷积神经网络训练的影响“

- 相关核心数学对象，如卷积、修正线性激活、最大池化以及局部响应归一化；
- 训练过程中一些网络行为的可视化，这些行为包括输入图像、损失情况、网络行为的分布情况以及梯度；
- 算法学习参数的移动平均值的计算函数，以及在评估阶段使用这些平均值提高预测性能；
- 实现了一种机制，使得学习率随着时间的推移而递减；
- 如果有条件，可以讨论网络在多个GPU上并行训练，多个GPU之间实现参数共享和更新变量值。

③ 讨论数据增强对于网络性能的改善。

- 讨论DROP-OUT机制对于网络训练和泛华能力的影响。

(3) 参考答案

如果在PaddlePaddle中完成作业，可以参考：

- 2021年人工神经网络第四次作业 - 第三题Cifar10

- 在Paddle中利用AlexNet测试CIFAR10数据集合

4、旋转数字

(1) 数据库简介

ROTATEDIGIT数据集合是采集到机械电能表显示的数字。这类数字中包含有数字滚轮在不同角度下数字的图片。这些数字不同角度除了对应了数字本身，还代表了整数之间的小数分量。



▲ 图5.1.1 机械电表上的旋转数字

在ROTATEDIGIT目录下包括有如下六个数据文件：

(1) black_1_1:



(4) red_10:



(2) black_10:



(5) red_20:



(3) black_11:



(6) red_21:



▲ 图5.1.2 六个数据文件

将数据库使用MATLAB命令LOAD调入之后，会产生以下数据变量：

1. `digitSize`: 表示数字图像的尺寸
2. `digit`: $1000 \times 2128 \times 3$, 数字图片。前面是数字图片的个数, 中间是宽×高, 后面是RGB三维。
3. `label`: 数字标签: 0 ~ 9
4. `fraction`: 表示数字整数之后的小数。即反映了数字旋转的程度。fraction部分, 有的数据集合存在, 有的不存在。

(2) 作业要求

- ① 选择合适的神经网络, 完成对数据集合的字符识别。
- ② 讨论使用某一类数字集合训练之后的网络, 在另外数字集合上的泛化性能。

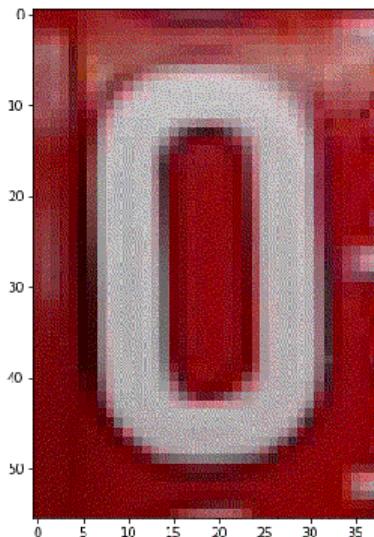
③ 讨论网络正则化的方法对于网络泛化能力的影响：

1. 对于网络权系数范数约束；
2. 对输入样本增加噪声；
3. 对输入样板进行变形增强；
4. 使用Drop-out技术训练网络

④ 对于有Fraction标签的数据集合，构造一个具有输出带有小数点数字输出的回归网络。并讨论它与前面分类网络在网络结构，训练方法等方面的特点。

(3) 参考答案

实验指导可以参见：[2021年人工神经网络第四次作业-第四题：旋转的数字](#)。

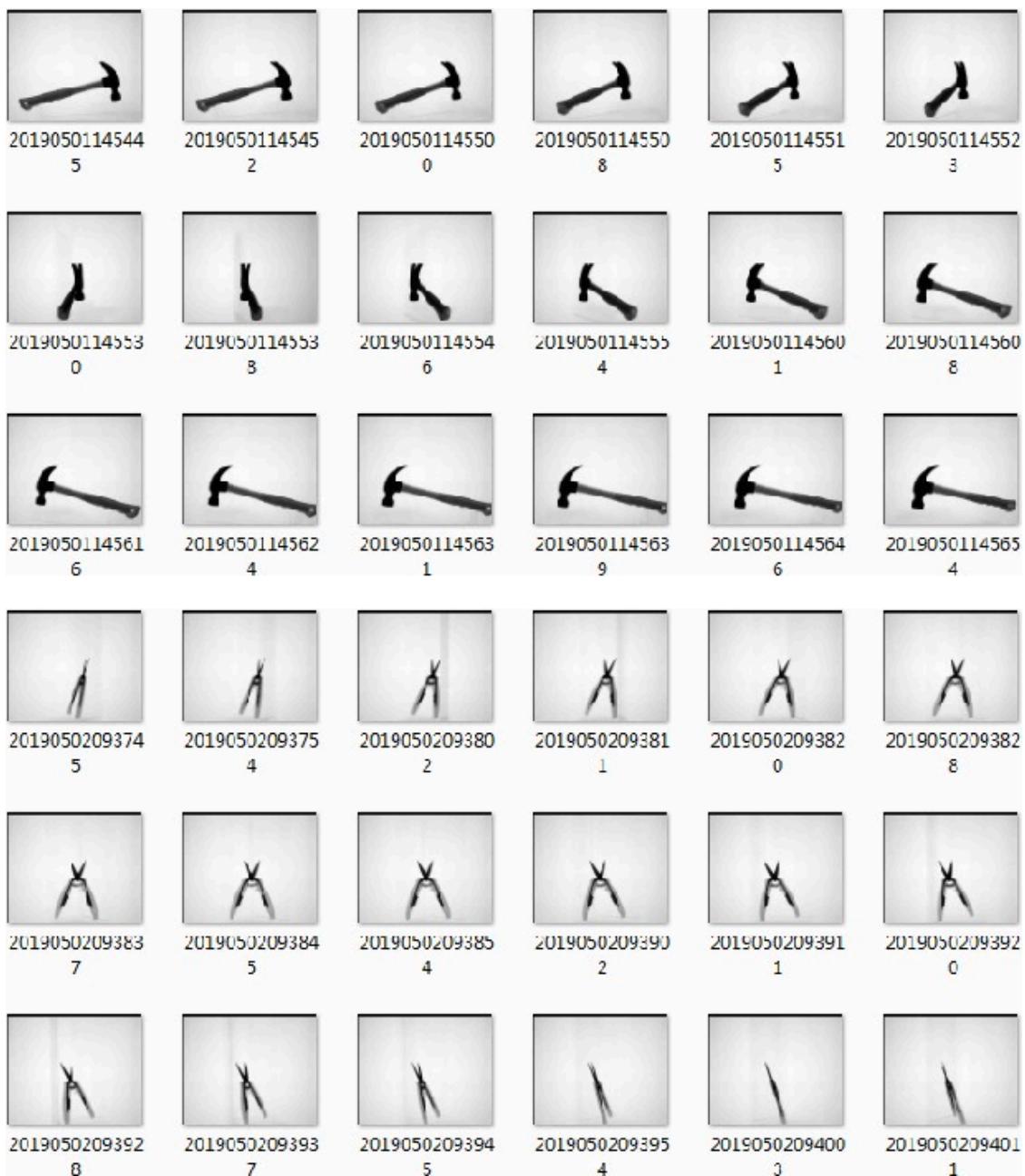


▲ 图5.2.1 第四次作业时间指导

5、危险品识别

(1) 数据库简介

FORBIDDEN数据集合是拍摄的一些物品的X-射线图片。同一个物品拍摄有不同的角度：



▲ 图6.1.1 危险品数据库

在目录FORBIDDEN中存在三个数据文件：

1. XRAYDATA-96-80.MAT: 尺寸为96×80灰度图像
2. XRAYDATA-120-100.MAT: 尺寸为120×100灰度图像
3. XRAYDATA-240-200.MAT: 尺寸为240×200灰度图像

三个文件的内容都是一样的，只是图片的尺寸有区别。

数据文件使用MATLAB命令load调入内存后，会产生三个变量：

- images: 7401×48000 uint8: 物品图片
- labels: 物品种类标签
- labeldir: 物品拍摄方位
-

同一个物品在拍摄的过程中，大都存在三个不同的方位，每个方位具有18个左右不同角度。

在图片集合中，同一物品在同一方位下，不同的角度拍摄的图片在数据集合中是顺序排列的。可以根据这个信息获得图频拍设的角度。

在数据集合中，标号小于400的样本为单个物品的数据，标号大于等于400的样本为多个物品重叠后的样品。大家实验中，可以只使用标号小于400的样品。

(2) 作业要求

① 构造合适的神经网络，对于数据库中物品X射线图片进行识别分类。可以根据物品不同的大小，在训练网络之前对于图片进行预处理。

② 讨论实验结果中，造成识别错误的因素包括哪些。

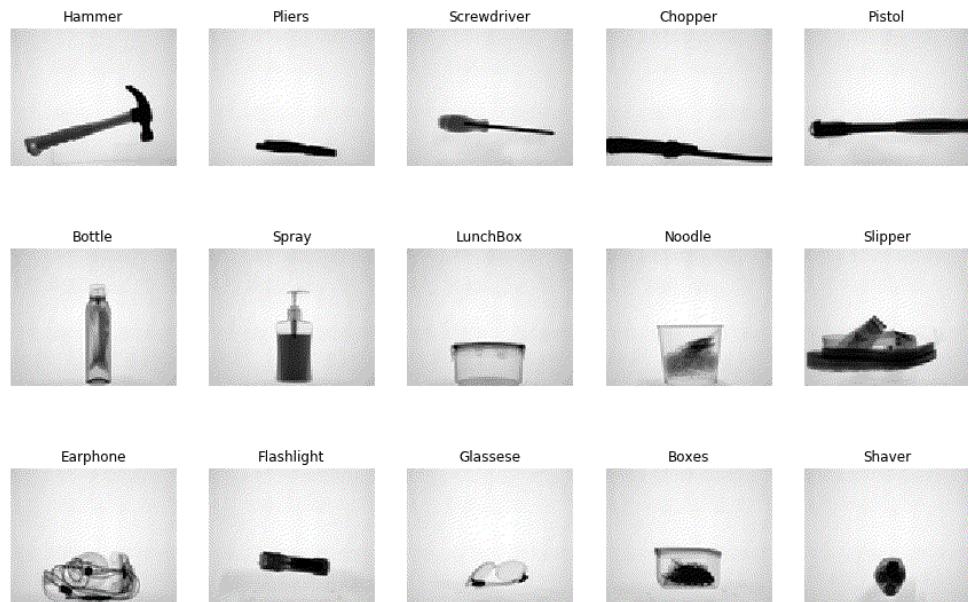
③ 探讨如何利用图频的序列信息完成物品的识别的方法。根据自己的时间和经验，选做是否可以通过LSTM网络来利用图频的不同角度序列信息提高识别的精度。

④ 可以根据自己的时间和经验，选做一些目标定位的神经网络：比如YOLO3, SSD等来进行实验。

(3) 参考答案

作业实验指导说明可以参见：**2021年人工神经网络第四次作业-第五题：危险品识别**

◦



▲ 图6.2.1 第五次作业中用于识别的数据库

(4) AI Studio数据库

(5) 数据库简介

在AI Studio中存在很多的公开数据集合。可以自行选择其中的数据集合用于识别、检测和参数回归。

The screenshot shows the AI Studio interface. At the top, there's a navigation bar with '飞桨 AI Studio' logo, '项目', '数据集' (highlighted in blue), '课程', '比赛', '认证', '更多', '论坛', '访问飞桨官网', and language switch '中 | En'. Below the navigation is a search bar with placeholder '交通标志' and a magnifying glass icon. To the right of the search bar is a '创建数据集' button. The main area is titled '公开数据集' (Public Dataset) with tabs for '我的数据集' and '我喜欢的'. A dropdown menu shows '全部标签' (All Labels) and '综合排序' (Comprehensive Sorting). A search input field contains '交通标志'. Below this, a list of datasets is shown:

- 交通标志-new 2020-05-04 (by skylee, 42 views, 179 likes, 0 comments)
- 交通标志 2021-08-08 (by qween994, 10 views, 94 likes, 0 comments)

To the right, a sidebar titled '数据集贡献达人' (Dataset Contribution Leader) lists top contributors:

名次	用户名	贡献数	引用次数
1	小度AIStudio	52	93.7K
2	PaddleTalent	56	6.2K
3	代码生成器	73	3.4K
4	深渊上的坑	56	4.2K
5	笨笨	17	53.8K
6	哦吼	17	9.7K

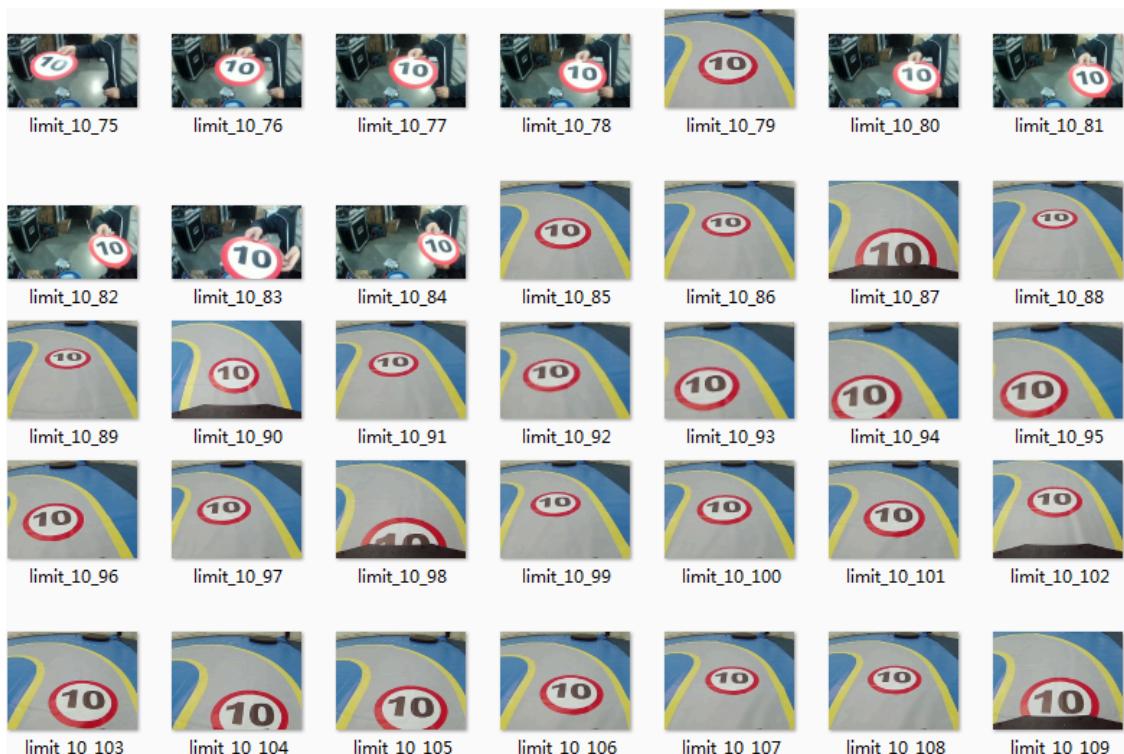
▲ 图7.1.1 AI Studio中的公开数据集合

下面给出几种典型的数据集合：

II. 交通标志

下载地址：<https://aistudio.baidu.com/aistudio/datasetdetail/49552>

该数据集合给出了5018个从驾驶员视角的路面交通标志。标识的种类有图片的名称给出。



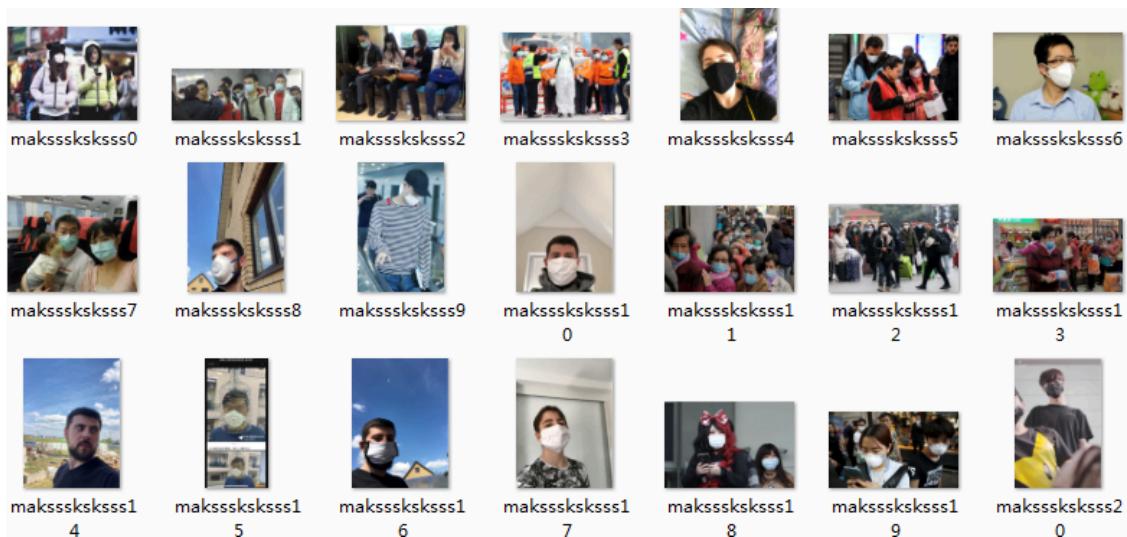
▲ 图7.1.2 驾驶员视角对应的道路上的交通标示

III. 目标检测

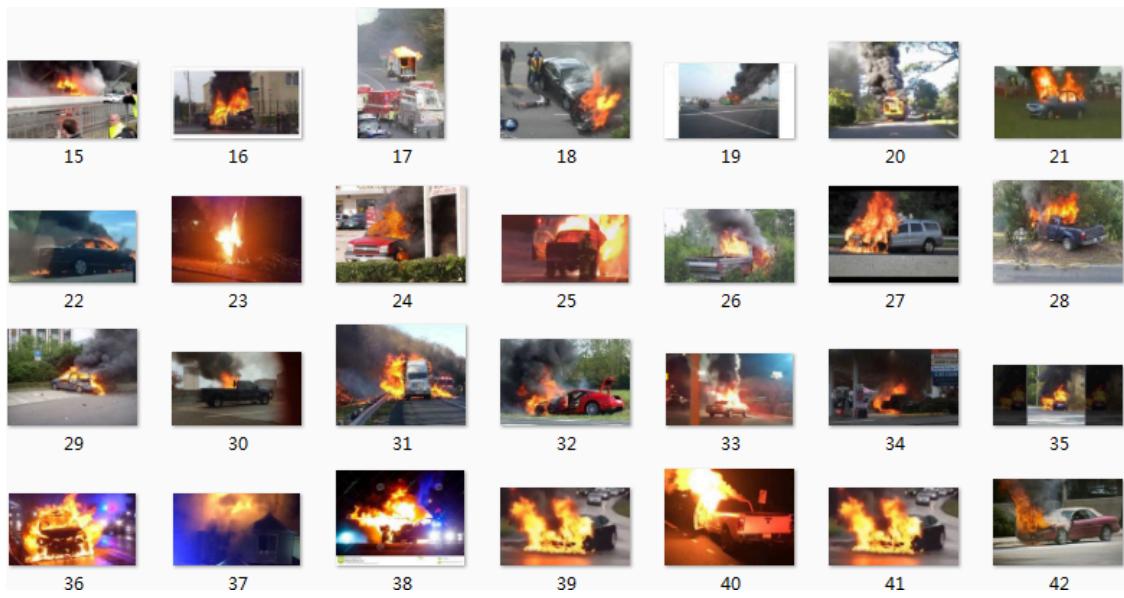
目标检测数据集合集 : <https://aistudio.baidu.com/aistudio/datasetdetail/103743>

该集合中给出了五类目标检测的集合，在数据集中分别包含了检测的图片以及对应的XML格式的标示结果。

```
|─barricade
|  ├Annotations
|  └JPEGImages
├facemask
|  ├Annotations
|  └JPEGImages
├fire
|  ├Annotations
|  └JPEGImages
├MidAutumn
|  ├Annotations
|  └JPEGImages
└roadsign_voc
  ├Annotations
  └JPEGImages
```



▲ 图7.1.3 人脸检测



▲ 图7.1.4 着火现场火源检测

IV. 水果数据集合

5种水果数据集：<https://aistudio.baidu.com/aistudio/datasetdetail/12236>

解压缩之后的水果所在的目录：

```

├── cantaloupe
├── carrot
├── cherry
├── cucumber
└── watermelon

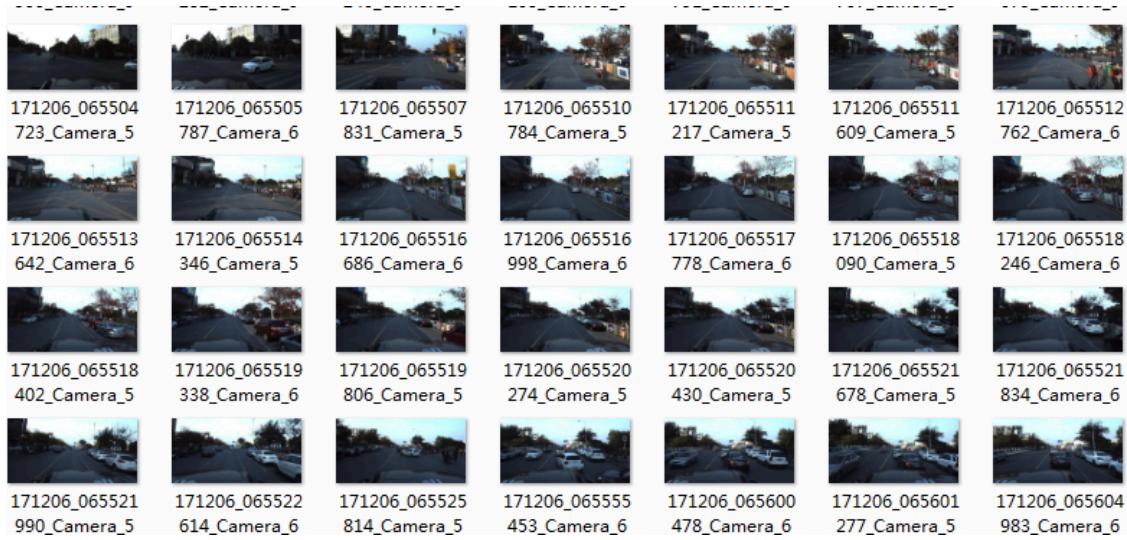
```



▲ 图7.1.5 蔬菜数据集合

V. 车道监测

无人车车道检测数据集合 : <https://aistudio.baidu.com/aistudio/datasetdetail/3625>



▲ 图7.1.6 车道线检测数据集合

VI. 交通灯

红绿灯及交通标志 : <https://aistudio.baidu.com/aistudio/datasetdetail/117209>

该数据集合中包含有红绿灯以及交通标志的检测图片和标注。



▲ 图7.1.7 交通灯目标检测

VII. 作业要求

参照前面五个数据集合的要求，设计深度神经网络完成数据集合中分类和定位要求。

■ 相关文献链接：

- [2023年秋季学期人工神经网络第五次作业](#)
- [MATLAB深度学习 2019](#)
- [《Python核心编程（第二版）》，TensorFlow、Pytorch和Keras三个深度学习框架](#)
- [五种水果与五种动物数据集合](#)
- [2021年人工神经网络第四次作业-第一题：LeNet对于水果与动物进行分类](#)
- [对于FAMNIST中的十种动物和水果进行识别测试-CSDN博客](#)
- [THE MNIST DATABASE of handwritten digits](#)
- [2021年人工神经网络第四次作业 - 第二题MNIST手写体识别](#)
- [2021年人工神经网络第四次作业 - 第三题Cifar10](#)
- [在Paddle中利用AlexNet测试CIFAR10数据集合](#)
- [2021年人工神经网络第四次作业-第四题：旋转的数字](#)
- [2021年人工神经网络第四次作业-第五题：危险品识别](#)
- [目标检测数据集合集](#)
- [5种水果数据集](#)
- [无人车车道检测数据集合](#)
- [红绿灯及交通标志](#)

● 相关图表链接：

- [图1.2.1 AI Studio 学习界面](#)
- [图1.2.3 MATLAB 中深度学习网络设计界面](#)
- [图1.2.4 在MATLAB中搭建深度学习网络](#)
- [图1.2.5 TENSORFLOW开发深度学习网络](#)
- [图1.2.6 pyTorch开发深度学习网络](#)
- [图1.2.7 Keras 开发深度学习网络](#)
- [图2.1.1 水果与动物数据库的样本](#)
- [图2.1.2 AI Studio 数据集合](#)
- [图3.1.1 mnist手写出谁出图片](#)
- [图4.1.1 CIFAR-10数据库](#)

- 图5.1.1 机械电表上的旋转数字
- 图5.1.2 六个数据文件
- 图5.2.1 第四次作业时间指导
- 图6.1.1 危险品数据库
- 图6.2.1 第五次作业中用于识别的数据库
- 图7.1.1 AI Studio中的公开数据集合
- 图7.1.2 驾驶员视角对应的道路上的交通标示
- 图7.1.3 人脸检测
- 图7.1.4 着火现场火源检测
- 图7.1.5 蔬菜数据集合
- 图7.1.6 车道线检测数据集合
- 图7.1.7 交通灯目标检测