



DeepL

订阅DeepL Pro以翻译大型文件。

欲了解更多信息，请访问www.DeepL.com/pro。

机器学习HW3--图像分类

ML TAs

ntu-ml-2021-spring-ta@googlegroups.com

目标

1. 用卷积神经网络解决图像分类问题。
2. 用数据增强的方式提高性能。
3. 了解如何利用未标记的数据以及它的好处。

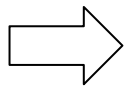
任务--食品分类



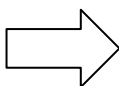
任务--食品分类

- 这些图像收集自food-11数据集，所有这些图像都被分类为11类。
- 这里的数据集稍作修改。
- 训练集。280 * 11张标记的图像 + 6786张未标记的图像
- 验证集。30 * 11个标记的图像
- 测试集。3347张图片
- 不要利用原始数据集或标签。
 - 这就是作弊。

任务--食品分类



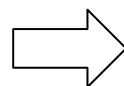
有线电
视新闻



逻辑数
(未规范化)

-0.25
+3.02
+0.56
-3.90
-0.01
+0.25
-0.47
+5.00
-1.32
+1.14
-0.28

软体



概率(归一化
)

0.004
0.114
0.009
0.001
0.005
0.007
0.003
0.831
0.001
0.017
0.004

要求

Kaggle链接:

<https://www.kaggle.com/c/ml2021spring-hw3>

- 这个家庭作业分三个层次。
 - 顺利
 - 中型
 - 困难
- 你可以通过运行示例代码轻松完成简单的级别。
- 对于其余的，我们建议你从相同的代码开始。
 - 我们已经为你准备了一些**TODO**块。
- 不要在其他数据集上预训练你的模型。

- 如果你使用一些著名的模型架构（如**ResNet**），请确保不加载预训练的权重作为初始化。

要求 - 简单

- 使用标记的图像和提供的代码建立一个卷积的**Nueral**网络。
- 简单公共基线：44.862（准确率，%）。

要求 - 中等

- 通过不同的模型架构或数据增量，使用标记的图像来提高性能。
- 公共正常基线：**52.807**（准确率，%）。
- 你可以通过在示例代码中增加几行来实现基线。

```
# It is important to do data augmentation in training.
# However, not every augmentation is useful.
# Please think about what kind of augmentation is helpful for food recognition.
train_tfm = transforms.Compose([
    # Resize the image into a fixed shape (height = width = 128)
    transforms.Resize((128, 128)),
    # You may add some transforms here.
    # ToTensor() should be the last one of the transforms.
    transforms.ToTensor(),
])
```

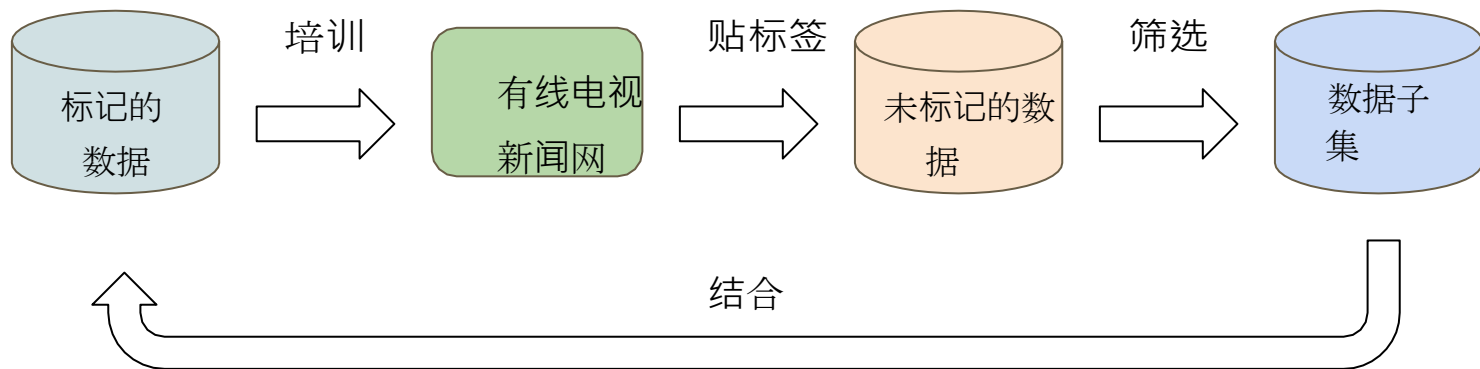
要求--困难

- 用更多的未标记的图像来提高性能。
- 公共强基线：82.138（准确率，%）。
- 自己做吧（通过完成示例代码中的TODO块）。
- 这里允许使用未标记的测试数据。
- 提示：半监督学习、自监督学习

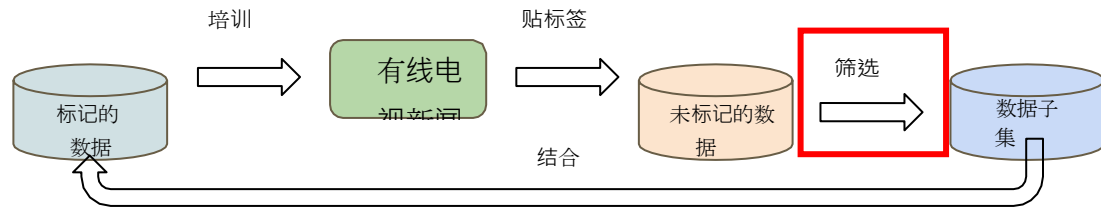
```
def get_pseudo_labels(dataset, model, threshold=0.65):  
    # This functions generates pseudo-labels of a dataset using given model.  
    # It returns an instance of DatasetFolder containing images whose prediction confidences exceed a given threshold.  
    # You are NOT allowed to use any models trained on external data for pseudo-labeling.
```

半监督学习

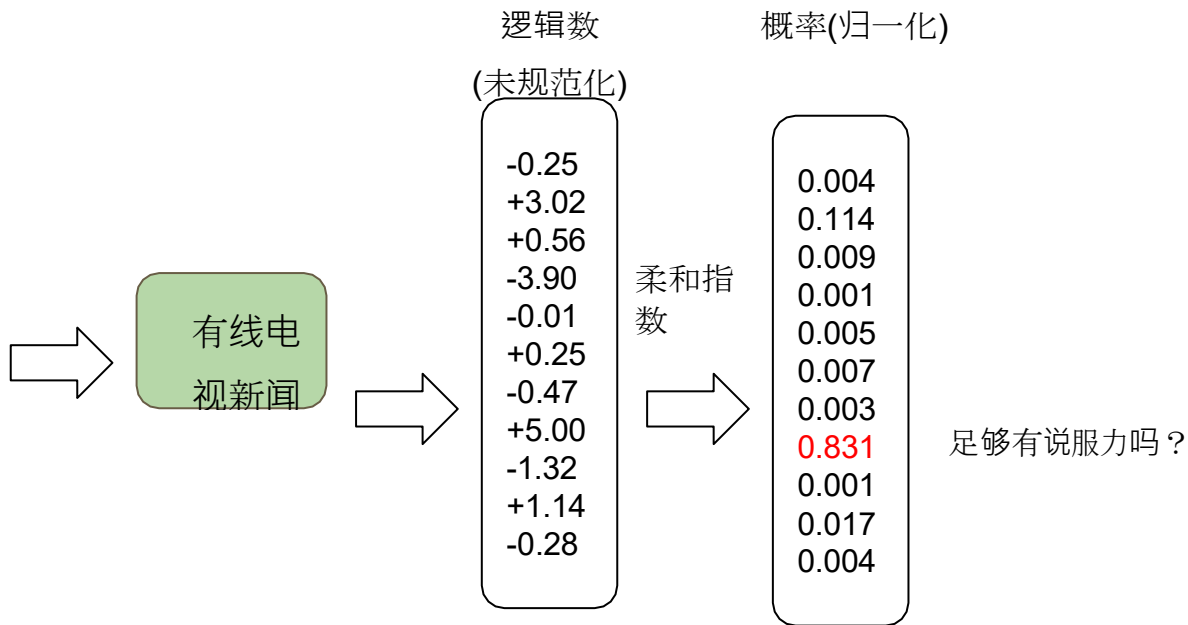
- 半监督学习有许多变种。
- 例如，为未标记的数据生成伪标签，并用它们进行训练。

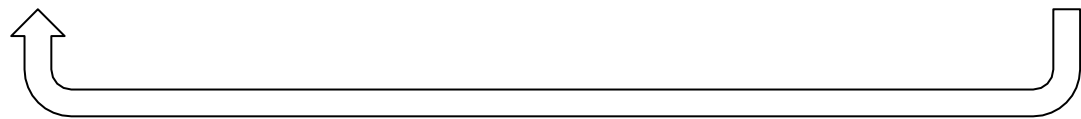


伪标签



伪标签=7





是

Kaggle提交格式

- 预测结果应以**csv**格式提交。
- 第一行是 **"id, label"**
- 其余的行是 "{id}, {prediction}"（例如，0005, 8）。
- 总共应该有（3347+1）行。

身份证	标签
0001	0
0002	9
0003	4
0004	5

评级政策

- 公共简单基线。 +1pt
- 公共媒介基线。 +1pt
- 公共强势基线。 +1pt
- 私人简单基线。 +1pt
- 私营中型基线。 +1pt
- 私人强大的基线。 +1pt

- 提交你的代码。

+4pt

代码提交

- 通过NTU COOL提交你的代码。

<student_id>_hw3.zip

- 哆啦**A**梦
 - 指定你的代码的来源。
 - 组织你的代码并使其易于阅读（不是必须的）。
- 请勿
 - 提交一个空文件或垃圾文件。
 - 提交数据集或模型。
 - 将你的代码压缩成其他格式，如.rar或.7z，并简单地将其重命名为.zip。

- 如果我们发现你作弊或你的代码有问题，你将受到惩罚。
 - 第一次的课程最终得分*0.9，否则就不能通过课程。

最后期限

- Kaggle截止日期： 2021/04/16 23:59:59
- 代码提交： 2021/04/18 23:59:59
- 逾期提交的文件不被接受。

如果你有任何问题...

- NTU COOL（推荐）。
 - <https://cool.ntu.edu.tw/courses/4793>
- 电子邮件
 - ntu-ml-2021spring-ta@googlegroups.com
 - 标题必须以**[hw3]**开头。
- TA小时
 - 星期五 14:00 - 18:00

有用的资源

- 半监督学习

- [https://speech.ee.ntu.edu.tw/~tlkagk/courses/ML_2016/Lecture/semi%20 \(v3\) .pdf](https://speech.ee.ntu.edu.tw/~tlkagk/courses/ML_2016/Lecture/semi%20(v3).pdf)
- https://www.youtube.com/watch?v=fX_guE7JNnY&ab_channel=Hung-yiLee
- MixMatch: <https://arxiv.org/abs/1905.02249>
- 吵闹的学生 : <https://arxiv.org/abs/1911.04252>

- 蟒蛇

- <https://pytorch.org/>

- 炬视

- <http://pytorch.org/vision/stable/index.html>