

# lesson1

## 1. 基本介绍

### 1.1 参加 ML/DL 比赛的目的

中心点：如何在简历中体现出自己的能力从而在千份挑一的简历中脱颖而出？

比赛的排名不重要！！

重要的是一定要思路清晰的表达出自己的解决方案，兼顾创新性和性能

## 1.1 参加ML/DL比赛的目的

### 1.1 purpose of participating a ML/DL competition

#### 1. 为什么会有这么多的ML/DL比赛？

##### 1) 公司:

通过奖金方式寻求业务痛点的解决方案...

提升自己的影响力，通过比赛谈业务..

寻求人才

##### 2) 学术界:

模型突破推动学科影响力...

#### 2. 比赛平台介绍

1) kaggle <https://www.kaggle.com/>

2) tianchi <https://tianchi.aliyun.com/>

3) CodaLab <https://codalab.lisn.upsaclay.fr/>

4) DataFountain <https://www.datafountain.cn/>

- 1st Place - \$12,000
- 2nd Place - \$8,000
- 3rd Place - \$5,000

TOP8队伍均可获奖，最高可获得10万元现金奖励，总奖金20万！

一等奖 10 万 (1 名)  
二等奖 4 万 (1 名)  
三等奖 2 万 (1 名)  
创新奖 8 千 (5 名)

为进一步激励参赛者，还特别设置了更多竞赛福利。

- 优秀人才推荐：高校精英-有机会获得博世中国校招及实习绿卡，直通部门面试；社会大牛-有机会加入博世中国人才库定期获得职位推送，与博世HR或技术专家一对一交流
- 参加颁奖典礼：获奖选手受邀参加在2022WAC大会开幕式上举行的颁奖典礼
- 颁发荣誉证书：获奖选手可获得博世人工智能大会组委会颁发的证书
- 获得超多奖品：比赛过程中有机会获得博世咖啡机、面包机等在内的超多奖品

kaggle

阿里云 | TIANCHI 天池

CodaLab  
Accelerating reproducible computational research.

DataFountain



# 1.1 参加ML/DL比赛的目的

## 1.1 purpose of participating a ML/DL competiton

### 3. 从哪里知晓比赛

公司 & 学术界.

<https://aideadlin.es/?sub=ML,CV,CG,NLP,RO,SP,DM>

小享的推送

### 4. 通过比赛我能学到什么？

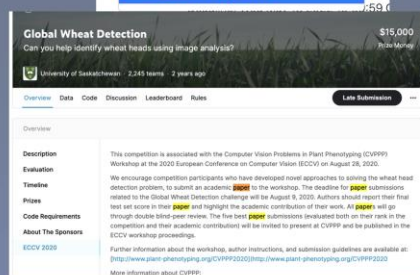
1) 公司：

比赛包含的技能、面试  
业务方向代码的储备。

2) 学术界：

workshop paper

4、有机器学习/深度学习相关项目落地经验，或在Kaggle等比赛有优异成绩者，优先。



# 1.1 参加ML/DL比赛的目的

## 1.1 purpose of participating a ML/DL competiton

### 5. Kaggle Progression System

介绍：<https://www.kaggle.com/progression>

**Novice**  
You've joined the community.  
☒ Register!

**Contributor**  
You've completed your profile, engaged with the community, and fully explored Kaggle's platform.

- ☒ Run 1 notebook or script
- ☒ Make 1 competition submission
- ☐ Make 1 comment
- ☐ Give 1 upvote

**Expert**  
You've completed a significant body of work on Kaggle in one or more categories of expertise. Once you've reached the expert tier for a category, you will be entered into the site wide Kaggle Ranking for that category.

Competitions	Datasets	Notebooks	Discussions
<input type="checkbox"/> 2 bronze medals	<input type="checkbox"/> 3 bronze medals	<input type="checkbox"/> 5 bronze medals	<input type="checkbox"/> 50 bronze medals

**Master**  
You've demonstrated excellence in one or more categories of expertise on Kaggle to reach this prestigious tier. Masters in the Competitions category are eligible for exclusive Master-Only competitions.

Competitions	Datasets	Notebooks	Discussions
<input type="checkbox"/> 1 gold medal <input type="checkbox"/> 2 silver medals	<input type="checkbox"/> 1 gold medal <input type="checkbox"/> 4 silver medals	<input type="checkbox"/> 10 silver medals	<input type="checkbox"/> 50 silver medals 200 medals in total

**Grandmaster**  
You've consistently demonstrated outstanding performance in one or more categories of expertise on Kaggle to reach this pinnacle tier. You're the best of the best.

Competitions	Datasets	Notebooks	Discussions
<input type="checkbox"/> 5 gold medals Solo gold medal	<input type="checkbox"/> 5 gold medals <input type="checkbox"/> 5 silver medals	<input type="checkbox"/> 15 gold medals	<input type="checkbox"/> 50 gold medals 500 medals in total

## 1.2 入门路径

中心点：如何在发现一个好的 AI 问题 并解决并且 publish 一篇论文？

## 2.1 入门路径

### 1.1 RoadMap

#### 1. 代码能力

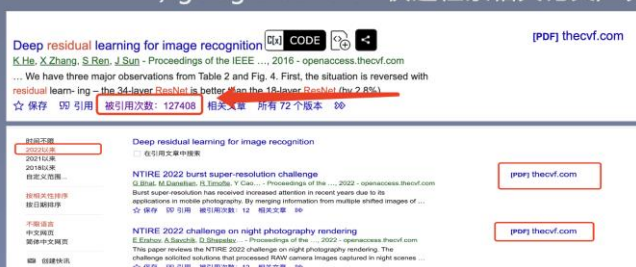
- 1) ML/DL框架: Pytorch...
- 2) 各种python包 : scikit-learn, pandas...

PyTorch



#### 2. 论文检索学习能力

- 1) google scholar 快速检索相关论文/ 以及后续改进工作



Dear Authors,

ICTAI 2022 uses a double-blind review process. To facilitate this, please ensure that your names and affiliations do not appear anywhere in the submitted paper.

Best Regards!

ICTAI 2022 Organizing Committee

### Important Dates – Extended

Paper Submission: ~~July 31, 2022~~ **August 8, 2022**  
 Paper Notification: ~~August 31, 2022~~ **September 5, 2022**  
 Camera Ready: ~~September 20, 2022~~ **September 26, 2022**

最近的会议链接 : <https://ictai.computer.org/2022>

7:34



Q 搜索宝贝



20

宝贝

评价

详情

推荐

2

舒适鼠标

经久耐用 多买多送



直播讲解

## 【不同配置供你挑选】

配置	处理器	显卡选择	内存+硬盘	价格
配置一	E5-2680 C2*2 (双路16核32线程)	GTX960M 4G	32G+360G SSD固态	¥2290
配置二	E5-2670 V2*2 (双路20核40线程)	GTX960 4G /GTX850M 8G	48G+360G SSD固态	¥2690
配置三	E5-2680 V2*2 (双路20核40线程)	GTX960 4G /GTX850M 8G	64G+500G NVME M.2 固态	¥3190
配置四	E5-2666 V3*2 (双路20核40线程)	GTX1060 5G	64G+500G NVME M.2 固态	¥3490
配置五	E5-2673 V3*2 (双路24核48线程)	GTX1060 5G	64G+500G NVME M.2 固态	¥3690
配置六	E5-2680 V3*2 (双路24核48线程)	GTX1060 6G	64G+500G NVME M.2 固态	¥4090
配置七	E5-2680 V4*2 (双路28核56线程)	GTX1060 6G	64G+500G NVME M.2 固态	进店逛逛
配置八	E5-2697 V3*2 (双路28核56线程)	GTX1070 8G	64G+500G NVME M.2 固态	¥4790
配置九	E5-2696 V3 *2 (双路36核72线程)	GTX1070 8G	64G+500G NVME M.2 固态	¥5690
配置十	E5-2682 V4 *2 (双路32核64线程)	GTX1070 8G	64G+500G NVME M.2 固态	¥5890
配置十一	E5-2673 V4 *2 (双路40核80线程)	GTX1070TI 8G	96G+1TB NVME M.2 固态	¥6990
配置十二	E5-2696 V4 *2 (双路44核88线程)	RTX2060 12G	128G+1TB NVME M.2 固态	¥8590
配置十三	E5-2696 V4 *2 (双路44核88线程)	RTX3060 12G	128G+1TB NVME M.2 固态	¥9990



## 2. Kaggle 平台使用

中心点 : Kaggle & 天池的基本介绍

### 2.1 Overview

Overview

[Overview](#) [Data](#) [Code](#) [Discussion](#) [Leaderboard](#) [Rules](#) [Team](#) [My Submissions](#) [Submit Predictions](#) ...

**Description** This competition is evaluated on the mean [Dice coefficient](#). The Dice coefficient can be used to compare the pixel-wise agreement between a predicted segmentation and its corresponding ground truth. The formula is given by:

$$\frac{2 * |X \cap Y|}{|X| + |Y|}$$

where X is the predicted set of pixels and Y is the ground truth. The Dice coefficient is defined to be 1 when both X and Y are empty. The leaderboard score is the mean of the Dice coefficients for each image in the test set.

**Supervised ML Evaluation** **评价指标**

**Timeline** **比赛时间、奖金**

**Prizes**

**Judges Prize**

**Code Requirements** **Submission File** **kernel赛限制比赛时间**

**Organizers & Sponsors**

In order to reduce the submission file size, our metric uses run-length encoding on the pixel values. Instead of submitting an exhaustive list of indices for your segmentation, you will submit pairs of values that contain a start position and a run length. E.g. '1 3' implies starting at pixel 1 and running a total of 3 pixels (1,2,3).

Note that, at the time of encoding, the mask should be binary, meaning the masks for all objects in an image are joined into a single large mask. A value of 0 should indicate pixels that are not masked, and a value of 1 will indicate pixels that are masked.

The competition format requires a space delimited list of pairs. For example, '1 3 10 5' implies pixels 1,2,3,10,11,12,13,14 are to be included in the mask. The metric checks that the pairs are sorted, positive, and the decoded pixel values are not duplicated. The pixels are numbered from top to bottom, then left to right: 1 is pixel (1,1), 2 is pixel (2,1), etc.

The file should contain a header and have the following format:

```
id, rle
2,1 3
5,3 4
6,4 5 66 77
etc.
```

### 2.2 Data

Research Code Competition

**HuBMAP + HPA - Hacking the Human Body** **\$60,000**  
Segment multi-organ functional tissue units **Prize Money**

**HuBMAP + HPA** · 616 teams · 2 months to go (a month to go until merger deadline)

[Overview](#) [Data](#) [Code](#) [Discussion](#) [Leaderboard](#) [Rules](#) [Team](#) [My Submissions](#) [Submit Predictions](#) ...

**Data Description**

## Data Explorer

9.39 GB

- test\_images
  - 10078.tiff
- train\_annotations
- train\_images
- sample\_submission.csv
- test.csv
- train.csv

### < sample\_submission.csv (19 B)



Detail Compact Column

2 of 2 columns

id	rle
1 total values	1 unique value
10078	12 34

## ● 比赛数据描述

### Files

[train/test].csv Metadata for the train/test set. Only the first few rows of the test set are available for download.

- id** - The image ID.
- organ** - The organ that the biopsy sample was taken from.
- data\_source** - Whether the image was provided by HuBMAP or HPA.
- img\_height** - The height of the image in pixels.
- img\_width** - The width of the image in pixels.
- pixel\_size** - The height/width of a single pixel from this image in micrometers. All HPA images have a pixel size of 0.4  $\mu\text{m}$ . For HuBMAP imagery the pixel size is 0.5  $\mu\text{m}$  for kidney, 0.2290  $\mu\text{m}$  for large intestine, 0.7562  $\mu\text{m}$  for lung, 0.4945  $\mu\text{m}$  for spleen, and 6.263  $\mu\text{m}$  for prostate.
- tissue\_thickness** - The thickness of the biopsy sample in micrometers. All HPA images have a thickness of 4  $\mu\text{m}$ . The HuBMAP samples have tissue slice thicknesses 10  $\mu\text{m}$  for kidney, 8  $\mu\text{m}$  for large intestine, 4  $\mu\text{m}$  for spleen, 5  $\mu\text{m}$  for lung, and 5  $\mu\text{m}$  for prostate.
- rle** - The target column. A run length encoded copy of the annotations. Provided for the training set only.
- age** - The patient's age in years. Provided for the training set only.
- sex** - The sex of the patient. Provided for the training set only.

## ● 比赛如何提交

注意：每场 kaggle 比赛稍有不同

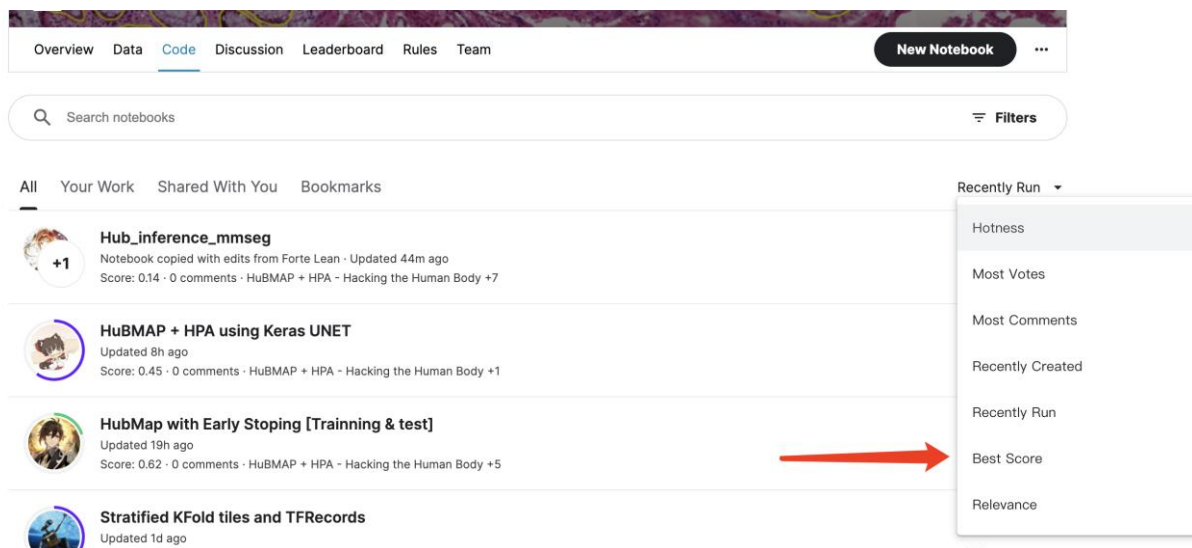
sample\_submission.csv

- id** - The image ID.
- rle** - A run length encoded mask of the FTUs in the image.

[train/test]\_images/ The images. Expect roughly 550 images in the hidden test set. All HPA images are 3000  $\times$  3000 pixels with a tissue area within the image around 2500  $\times$  2500 pixels. The HuBMAP images range in size from 4500 $\times$ 4500 down to 160 $\times$ 160 pixels. HPA samples were stained with antibodies visualized with 3,3'-diaminobenzidine (DAB) and counterstained with hematoxylin. HuBMAP images were prepared using Periodic acid-Schiff (PAS)/hematoxylin and eosin (H&E) stains. All images used have at least one FTU. All tissue data used in this competition is from healthy donors that pathologists identified as pathologically unremarkable tissue.

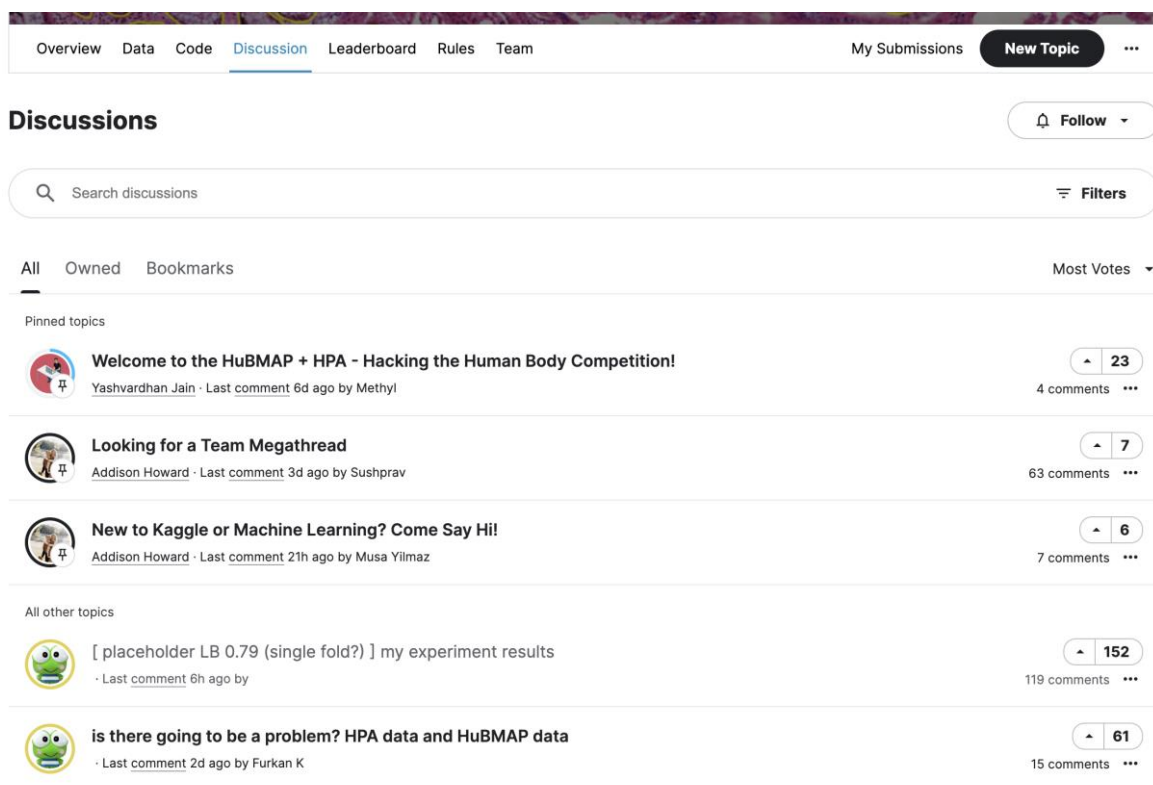
train\_annotations/ The annotations provided in the format of points that define the boundaries of the polygon masks of the FTUs.

## 2.3 Code




The screenshot shows the Kaggle Code page. At the top, there are tabs for Overview, Data, Code (selected), Discussion, Leaderboard, Rules, and Team. A 'New Notebook' button is on the right. Below the tabs is a search bar labeled 'Search notebooks' and a 'Filters' button. Under the search bar, there are tabs for All, Your Work, Shared With You, and Bookmarks. A list of notebooks is displayed, including 'Hub\_inference\_mmseg', 'HuBMAP + HPA using Keras UNET', 'HubMap with Early Stopping [Training & test]', and 'Stratified KFold tiles and TFRecords'. A red arrow points to a dropdown menu on the right side of the list, which contains the following options: Recently Run, Hotness, Most Votes, Most Comments, Recently Created, Recently Run, Best Score, and Relevance.

## 2.4 Discussion



The screenshot shows the Kaggle Discussion page. At the top, there are tabs for Overview, Data, Code, Discussion (selected), Leaderboard, Rules, and Team. A 'My Submissions' button and a 'New Topic' button are on the right. Below the tabs is a search bar labeled 'Search discussions' and a 'Filters' button. Under the search bar, there are tabs for All, Owned, and Bookmarks. A list of discussion topics is displayed, including 'Welcome to the HuBMAP + HPA - Hacking the Human Body Competition!', 'Looking for a Team Megathread', 'New to Kaggle or Machine Learning? Come Say Hi!', '[ placeholder LB 0.79 (single fold?) ] my experiment results', and 'is there going to be a problem? HPA data and HuBMAP data'. Each topic has a profile picture, a title, a description, a timestamp, and a comment count.

## 2.5 LeaderBoard

 HuBMAP + HPA · 616 teams · 2 months to go (a month to go until merger deadline)

OverviewDataCodeDiscussionLeaderboardRulesTeam

My SubmissionsSubmit Predictions...




### Leaderboard

[Raw Data](#)[Refresh](#)

PublicPrivate

This leaderboard is calculated with approximately 55% of the test data. The final results will be based on the other 45%, so the final standings may be different.

☒ Prize Contender

#	Team	Members	Score	Entries	Last	Code	Join
1	[placeholder] ★ ☆ ★ HP ☆ ☆ ★ ☆Z by		0.80	126	27m		
2	nw zzzz	 	0.80	79	20h		

OverviewDataCodeDiscussionLeaderboardRulesTeam

My SubmissionsSubmit Predictions...

### Leaderboard

[Refresh](#)

PublicPrivate



**The Private Leaderboard isn't available yet.**

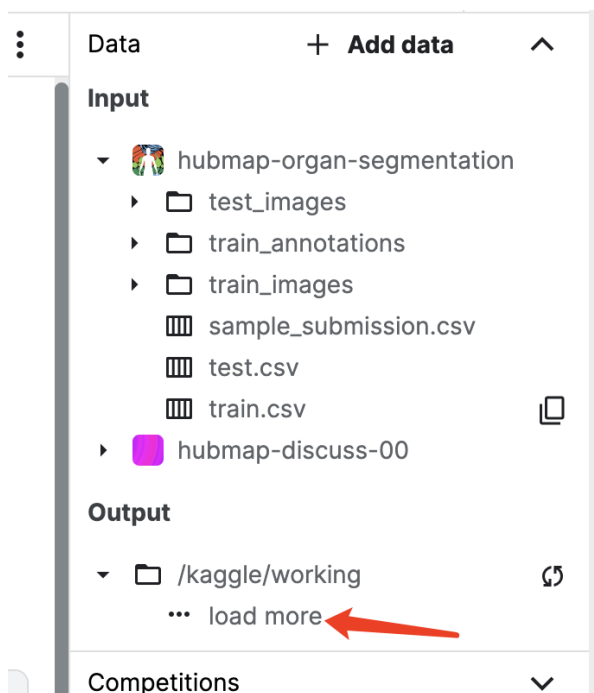
The final ranks and medals will display here after the competition closes.

## 3. Baseline

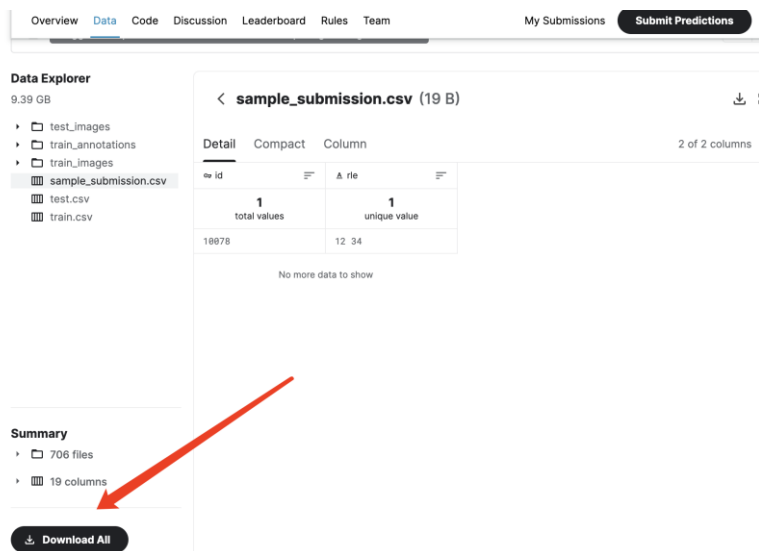
中心点：本次比赛的任务目的. & Baseline 基本结构

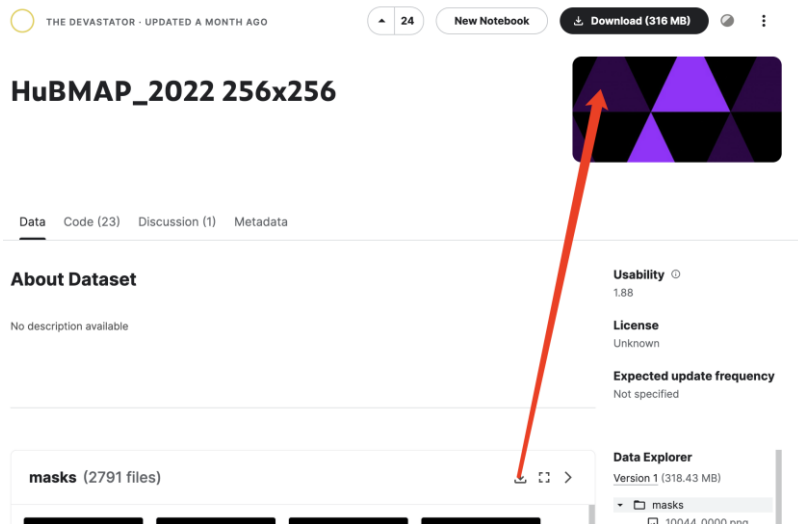


## 3.1 基本结构



### 本地构造的目录结构





<https://www.kaggle.com/datasets/thedevastator/hubmap-2022-256x256>

```
(base) → input ls
ckpt-frank      hubmap-2022-256x256.zip  hubmap-organ-segmentation.zip
hubmap-2022-256x256  hubmap-organ-segmentation  tmp
(base) → input
```

```
/home/wangtsh/kaggle/huBMAP/working
(base) → working git:(main) x ls
apex  logs    note      submission.csv      train_resnext101.py  utils
data  models  scripts  submit_ensembl_tta.py  train_resnext50.py
(base) → working git:(main) x
```

## 3.2 训练

# step0: 连接服务器

```
~ ssh frank@10.19.124.245 -p 2022
```

# step1: 查看 cuda 版本

```
~ nvcc -V
```

```
nvcc: NVIDIA (R) Cuda compiler driver
```

```
Copyright (c) 2005-2019 NVIDIA Corporation
```

```
Built on Wed_Oct_23_19:24:38_PDT_2019
```

```
Cuda compilation tools, release 10.2, V10.2.89
```

```
~ ll /usr/local
```

```
lrwxrwxrwx 1 root root 9 5月 14 15:14 cuda -> cuda-10.2
```

```
# step2: 创建 anaconda 环境 & pytorch
# https://pytorch.org/get-started/previous-versions/
conda create --name python3.8_torch1.8_cuda10.2 python=3.8
conda install pytorch==1.8.1 torchvision==0.9.1 torchaudio==0.8.1
          cudatoolkit=10.2 -c pytorch

# step5: 安装 sklearn
# https://scikit-learn.org/stable/auto_examples/release_highlights/plot_release_highlights_1_1_0.html?highlight=install
pip install --upgrade scikit-learn # 安装最新版

# step7: 安装 opencv
pip install opencv-python
```

缺什么包就安装什么包

**(python3.9\_torch1.11\_cuda11.3) → working git:(main) ✕ sh scripts/run\_resnext50.sh**

训练好的模型保存在：

```
(python3.9_torch1.11_cuda11.3) → working git:(main) ✕ ll ../input/ckpt-frank/resnext50
总用量 689M
-rw-rw-r-- 1 wanglsh wanglsh 138M 8月  3 11:26 fold0.pth
-rw-rw-r-- 1 wanglsh wanglsh 138M 8月  3 10:37 fold1.pth
-rw-rw-r-- 1 wanglsh wanglsh 138M 8月  3 10:37 fold2.pth
-rw-rw-r-- 1 wanglsh wanglsh 138M 8月  3 10:37 fold3.pth
-rw-rw-r-- 1 wanglsh wanglsh 138M 8月  3 10:37 fold4.pth
```

### 3.3 提交

线下验证是否正常运行：

直接 python **submit\_ensembl\_tta.py** .

快速提交方法：找到一个类似提交 notebook: 如：

<https://www.kaggle.com/code/thedevastator/inference-fastai-baseline>

THE DEVASTATOR · 1MO AGO · 4,732 VIEWS

76 Edit My Copy 365

## [Inference] - FastAI Baseline

Python · [Private Datasource], HuBMAP + HPA - Hacking the Human Body, [Training] - FastAI Baseline

Notebook Data Logs Comments (11)



Competition Notebook

HuBMAP + HPA - Hacking the Human Bo...

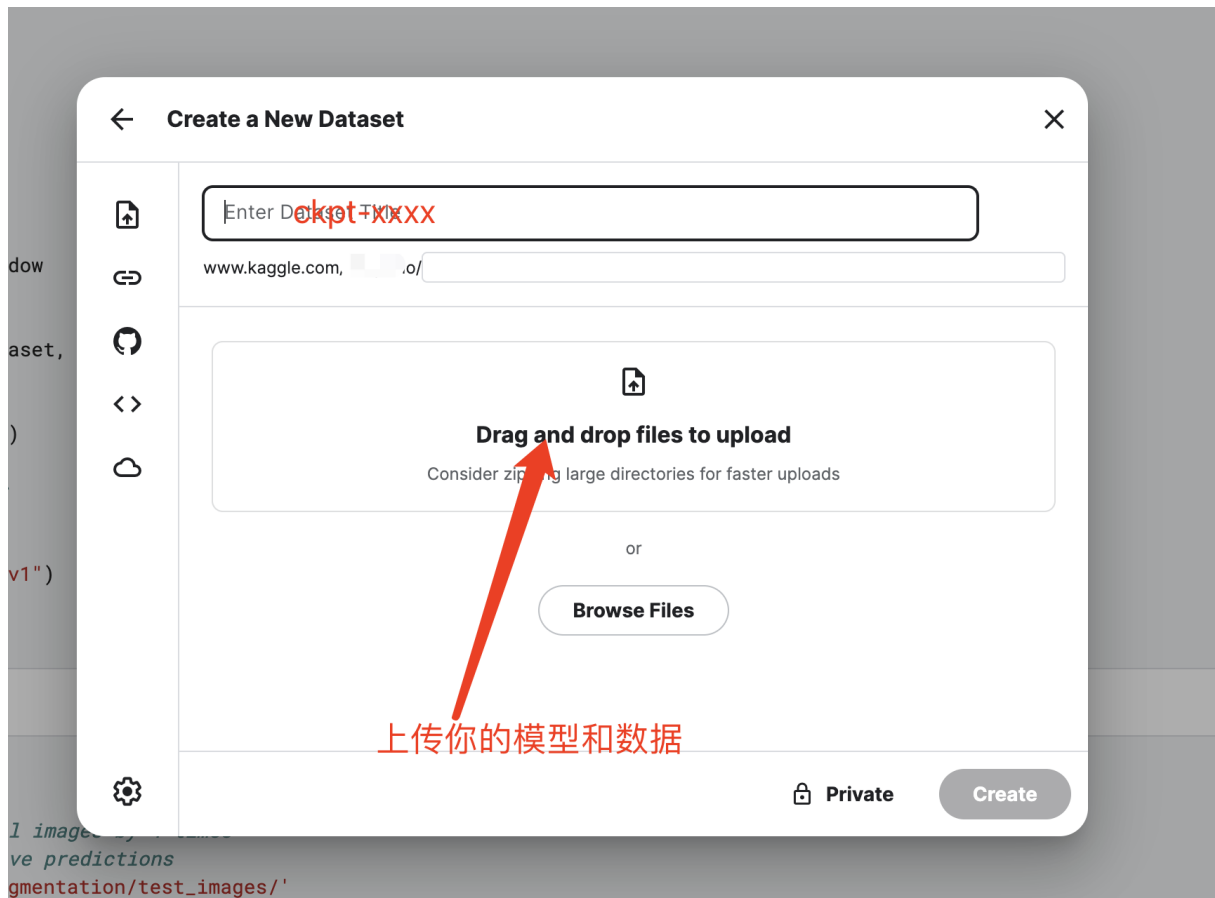
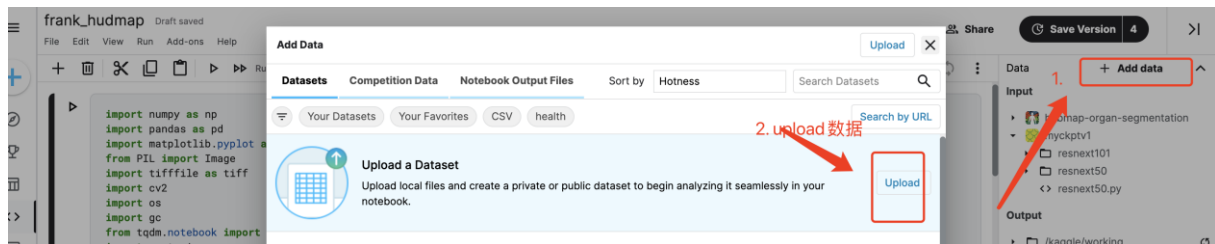
Run

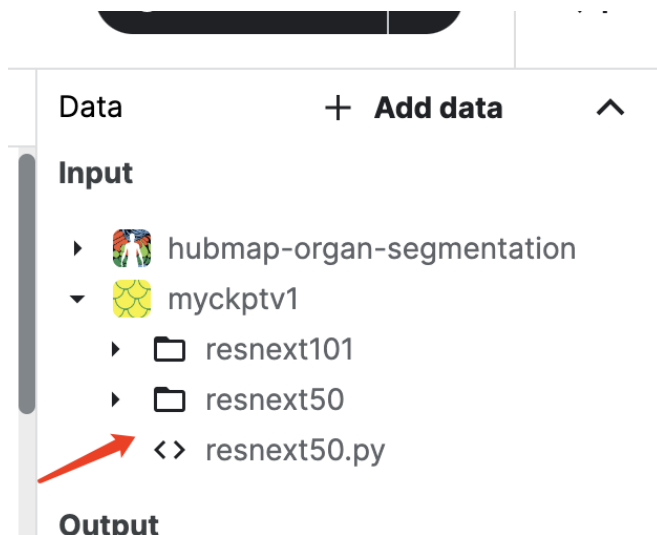
56.0s - GPU

Public Score

0.56

Version 21 of 21





右边在 input 目录下就有你的各种文件。

此时，粘贴 codebase 里面的 **submit\_ensembl\_tta.py** 文件



下面全部粘贴即可，但是要注意模型的路径



```
[ ]: bs = 64
sz = 256 # the size of tiles
reduce = 4 # reduce the original images by 4 times
TH = 0.4 # threshold for positive predictions
DATA = '../input/hubmap-organ-segmentation/test_images/'
MODELS_rsxt50 = [f'../input/myckptv1/resnext50/fold{i}.pth' for i in range(5)]
MODELS_rsxt101 = [f'../input/myckptv1/resnext101/fold{i}.pth' for i in range(5)]

df_sample = pd.read_csv('../input/hubmap-organ-segmentation/sample_submission.csv')
device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')

# functions to convert encoding to mask and mask to encoding
def enc2mask(encs, shape):
    img = np.zeros(shape[0]*shape[1], dtype=np.uint8)
    for m, enc in enumerate(encs):
        if isinstance(enc, np.float) and np.isnan(enc): continue
        s = enc.split()
        for i in range(len(s)//2):
            start = int(s[2*i]) - 1
            length = int(s[2*i+1])
            img[start:start+length] = 1 + m
    return img.reshape(shape).T

def mask2enc(mask, n=1):
    pixels = mask.T.flatten()
    encs = []
    for i in range(1, n+1):
        p = (pixels == i).astype(np.int8)
        if p.sum() == 0: encs.append(np.nan)
        else:
```

Save Version 4

Code ▾ Draft Session off (run a cell to start)

flips as TTA  
[[[-1], [-2], [-2, -1]]  
flips:  
torch.flip(x, f)  
model in self.models:  
p = model(xf)  
p = torch.flip(p, f)  
py += torch.sigmoid(p).detach()  
+len(f  
lf.mod  
ple(py  
ute(0,  
len(p  
ge(bat  
[i], y  
1  
et)  
import

Save Version

Version Name (optional):

Version 5 9 / 50

✓ Save & Run All (Commit)  
Run a fresh copy of your notebook and save the output

Advanced Settings Cancel Save

Data + Add data

Input

- hubmap-organ-segmentation
- myckptv1
  - resnext101
  - resnext50
  - resnext50.py

Output

- /kaggle/working

Competitions

Settings

Schedule a notebook run

Code Help

FIND CODE HELP

Find Code Help

Search for examples of how to do this

## Submit to competition



HuBMAP + HPA - Hacking the Human Body

You have 3 submissions remaining today. Submissions reset daily.

### Submission File:

SELECT NOTEBOOK\*

frank\_hudmap

NOTEBOOK VERSION\*

1200

OUTPUT FILE\*

submission.csv

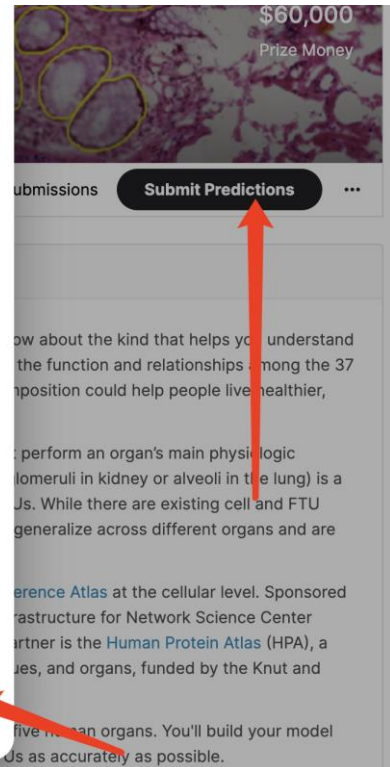
SUBMISSION DESCRIPTION

Notebook frank\_hudmap | 1200

In this competition, we will privately re-run your selected Notebook Version with a hidden test set substituted into the competition dataset. We then extract your chosen Output File from the re-run and use that to determine your score.

Cancel

Submit



点击 submit 即可，等待成绩即可。

## 3.2 trick 总结

见 UW lesson 总结 - 当成 wiki 来用

## 4. To do list

中心点：基本的 Training - 修改模型的能力

4.1 替换 backbone-resnext50/101 为 efficientNet

参考：Ucode HuBMAP <https://www.kaggle.com/code/yuliknormanowen/ucode-hubmap-efficientnet-submission>

## Ensemble EfficientNet Results and Submitting The Result

### Training

Two ways: train on local machine or on Kaggle

#### Local Machine

Use dataset [HuBMAP \(EfficientNet\) Train\\_Local](#)

1. Download dataset, run: `bash download.sh`
2. Training for fold 0, run: `python3 train.py -f 0` Ensemble all 4 folds

#### Kaggle

Use notebook [HuBMAP \(EfficientNet\) Train](#)

### Reference

- [HuBMAP fast.ai starter \(EfficientNet\)](#) by wangkui
- [Converting to 256×256](#) by The Devastator
- [\[Training\] - FastAI Baseline](#) by The Devastator
- [\[Inference\] - FastAI Baseline](#) by The Devastator


## 4.2 替换 backbone-resnext50/101 为 swintransformer

参考 : [\[ placeholder LB 0.79 \(single fold?\) \] my experiment results](#)

<https://www.kaggle.com/competitions/hubmap-organ-segmentation/discussion/332941>

<https://www.kaggle.com/code/hengck23/lb-0-75-variable-size-swin-transformer-v1-and-v2/data>

### Data

**HuBMAP + HPA - Hacking the Human Body**  
Segment multi-organ functional tissue units  
Last Updated: a month ago

**About this Competition**

The goal of this competition is to identify the locations of each functional tissue unit (FTU) in biopsy slides from several different organs. The underlying data includes imagery from different sources prepared with different protocols at a variety of resolutions, reflecting typical challenges for working with medical data.

This competition uses data from two different consortia, the [Human Protein Atlas \(HPA\)](#) and [Human BioMolecular Atlas Program \(HuBMAP\)](#). The training dataset consists of data from public HPA data, the public test set is a combination of private HPA data and HuBMAP data, and the private test set contains only HuBMAP data. Adapting models to function properly when presented with data that was prepared using a different protocol will be one of the core challenges of this competition. While this is expected to make the problem more difficult, developing models that generalize is a key goal of this endeavor.

This competition uses a hidden test. When your submitted notebook is scored the actual test data (including a full length sample submission) will be made available to your notebook.

**Input (9.4 GB)**

- Data Sources
  - HuBMAP + HPA - Hacki...
  - hubmap-discuss-00
    - train\_log
    - upernet-swin-v1v2

## 5. 下节课预告

数据讲解

## 6.答疑

1

1.机器配置

colab pro 可以完成本次比赛吗，需要什么方法能加速模型训练，调参如何调呢

2.检测跟分割用的 backbone 一样吗？分割跟农业方向可以怎么结合？需不需要别的技术？

3.目前的高分方案，是否都是基于 transformer？efficientnet 还有一战之力吗

4.colab 国内需要翻墙才能用吗

5.enc2mask(encs, shape)这个函数能讲解一下吗

2. 病理图片很大，怎么分 patch 来做了，一直不清楚，目前实际的工作中不分 patch 可能损失信息更多，比如缩小图片大小等，所以老师讲讲 patch 这个思路吧

3. 问题 1.目前分割 sota 前三是什么，大概讲下优缺点或者创新点 2. 打分割比赛常用 tricks

4. 有没有深度学习也可以借鉴和使用的特征提取方式，因为之前有类似的比赛纯靠提特征 +lgbm 也前三。

5. mask2former 是否适合本场比赛，为什么，咱们的代码可以修改为使用这个模型吗

6. 医学图像第一次做，问下常用的图像预处理都有哪些，需要做怎么评估和判断(从效果和理论上都参考)。另外对于不同的图像数据和不同的下层特征提取网络或者任务，选取预处理的方法有什么不一样的地方？

7. 之前胃肠道的比赛看到有人用 MONAI 分数也不错，老师可以讲一下吗

8. kaggle 的平台如何查看报名成功，如何组队，后期如何提交任务老师可以讲一下吗

9. 讲一下 main729 中导入的包在程序中哪个地方用了

10. 模型融合的代码可以详细讲解下吗，比如 resnet50 和 efficiennetb5 融合

11. github 开源的语义分割模型众多，每个作者的项目结构都不一样，请问如何有效整合不同的模型，来方便自己使用呢？

12. 可以讲一下如何修改 smp 那个包里封装的 unet 模型嘛？比如如何引入 ASPP、FPN 等。

13. 可以讲一下这个课程大体流程吗？比如什么时候以怎样的形式组队等

14. 请老师讲一下如何用 HPA pixel\_size=0.4 的图像去造 HubMAP 的数据？pixel\_size 的变化仅仅通过 resize 大小就可以实现吗？

15. SegFormer 训练太长了，如何加速收敛呢？