

Why 1.0 on Kaggle?

Kaggleで1.0になった理由は？

**These slides are at Connpass. So you can
check them out.**

スライドはConnpassにアップされています。是非みてください。

**My username on
Connpass is
"globophobe".**

**Connpassのユーザ名は
「globophobe」です。**



A little while ago, I finished Fast.ai lesson 3.

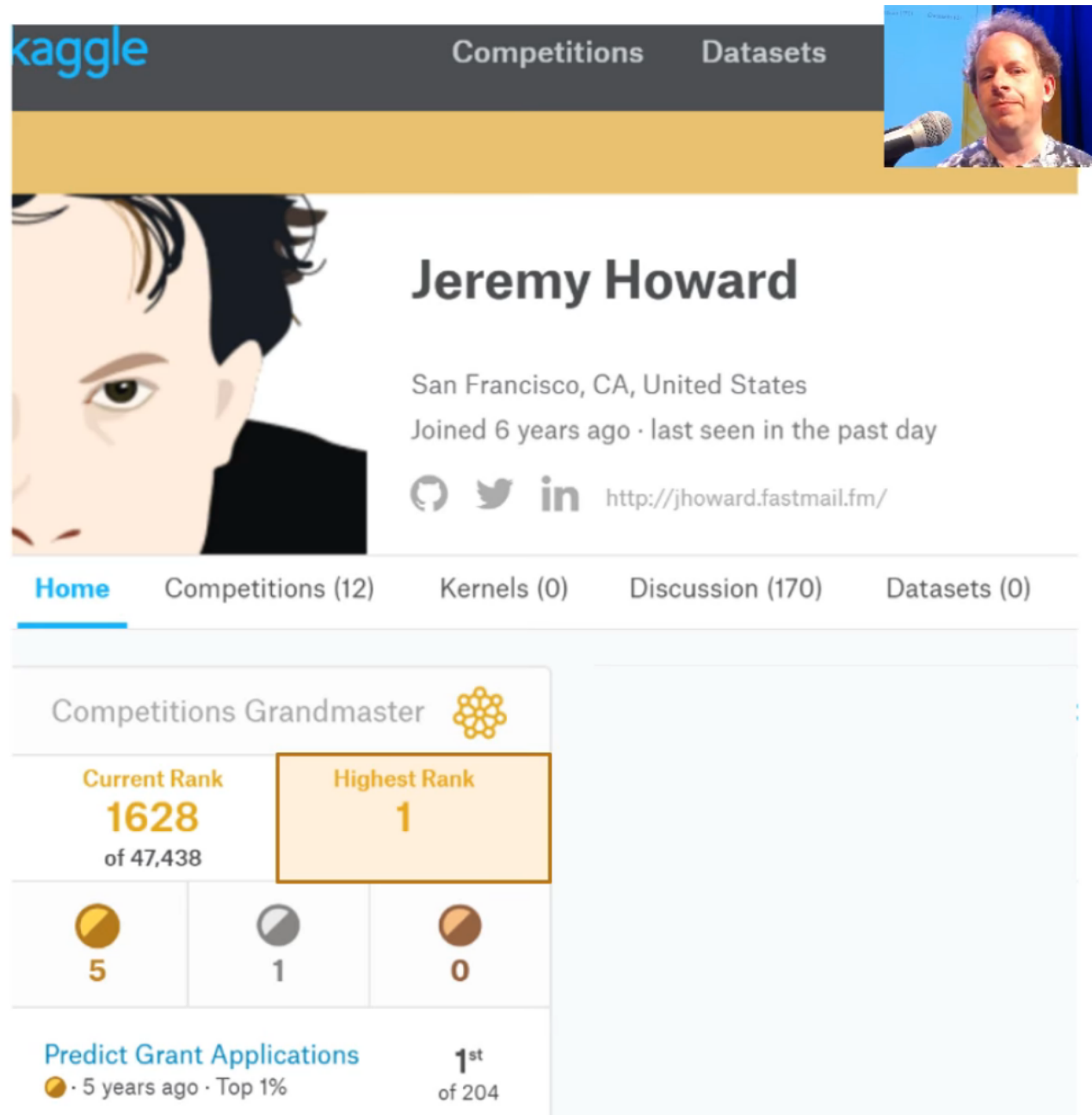
少し前、Fast.aiレッスン3を終えました。

Fast.ai is a free deep learning curriculum.

無料機械学習カリキュラムです。

The teacher was Kaggle #1, and its president.

先生はKaggleの1位、そして
Kaggleの組織の会長でした。



The image shows a screenshot of the Kaggle website profile for Jeremy Howard. The profile includes a stylized illustration of his face, his name, location (San Francisco, CA), and social media links. It also displays his Kaggle statistics, including his current and highest ranks, and a recent competition win.


Kaggle Competitions Datasets

Jeremy Howard




San Francisco, CA, United States
Joined 6 years ago · last seen in the past day

[GitHub](#) [Twitter](#) [LinkedIn](#) <http://jhoward.fastmail.fm/>

Home Competitions (12) Kernels (0) Discussion (170) Datasets (0)

Competitions Grandmaster 

Current Rank	Highest Rank
1628 of 47,438	1

5	1	0
		

Predict Grant Applications 1st
● · 5 years ago · Top 1% of 204

Lesson 1 and 2 briefly explain how to create a CNN with ResNet for transfer learning.

レッスン1と2は、簡単に転送学習用のResNetを使用してCNNを作成方法を説明します。

**Lesson 3 explains in more detail how to use the
Fast.ai library.**

**レッスン3では、Fast.aiライブラリの使用方法についてさらに詳しく
説明します。**

The teacher starts with a Kaggle contest.

先生はKaggleコンテストから始まります。



Planet: Understanding the Amazon from Space

地球：宇宙からはアマゾンを理解する

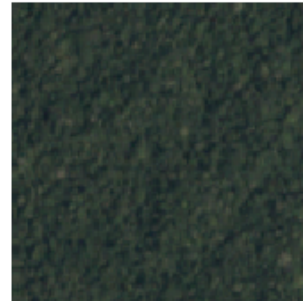
Data is multilabel satellite images.

データはマルチラベル衛星画像です。

primary; clear



primary; clear



primary; clear; agriculture; road



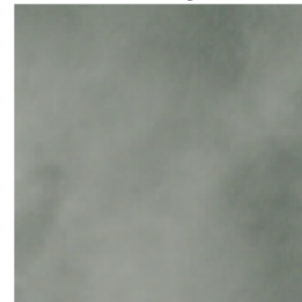
primary; clear; agriculture; road



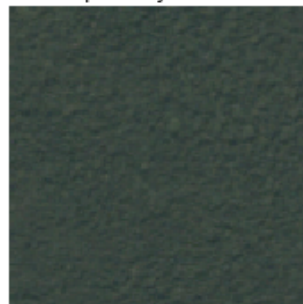
primary; partly_cloudy



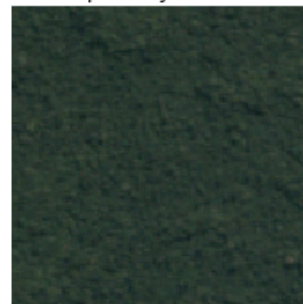
cloudy



primary; clear



primary; clear



primary; clear; water; agriculture; cultivation



**How to use the Fast.ai data block API to create
a CNN is explained.**

**Fast.aiデータブロックAPIを使用してCNNを作成する方法について説
明されています。**

Where is the data? データはどこ?

```
src = (ImageFileList.from_folder(path)
      .label_from_csv(
          'train_v2.csv',
          sep=' ',
          folder='train-jpg', suffix='.jpg'
      )
      .random_split_by_pct())
```

How to augment? データ増強方法？

```
tfms = get_transforms(  
    flip_vert=True,  
    max_lighting=0.1,  
    max_zoom=1.05,  
    max_warp=0.  
)
```

Create Fast.ai DataBunch instance.

Fast.ai DataBunchのインスタンスを作成します。

```
data = (  
    src.datasets()  
    .transform(tfms, size=128)  
    .databunch().normalize(imagenet_stats)  
)
```

**DataBunch is train, validation, and optionally
test PyTorch DataLoaders.**

DataBunchはPyTorch DataLoaderの束です。トレーニング、検証、そしてオプションでテストデータです。

After 5 epochs, about top 50 on Kaggle. トップ 50位ぐらいでした。

```
learn.fit_one_cycle(5, slice(0.01))
```

Total time: 04:17

epoch	train_loss	valid_loss	accuracy_t	fbeta
1	0.115247	0.103319	0.950703	0.910291
...
5	0.091275	0.085441	0.958006	0.926234

**Interesting point, Planet data is 256x256, but
he resized to 128x128.**

**興味深い点は、データは256 x 256ですけど、128 x 128にサイズ
が変更されました。**

Then, he made a new dataset at 256x256, and continued training with his pretrained model.

**そして、256 x 256の新しいデータセットを作って、そのデータで
128 x 128訓練されたモデルの訓練を続けました。**

In the end, about top 25 on Kaggle.

最終に、先生はトップ25位ぐらいでした。

Total time: 18:23

epoch	train_loss	valid_loss	accuracy_t	fbeta
1	0.083591	0.082895	0.968310	0.928210
...
5	0.074927	0.080691	0.968819	0.931414

**I wanted to practice, so I thought I would try a
Kaggle contest for the first time.**

**練習したかったので、Kaggleコンテストを初めてしようと思いまし
た。**



Aerial Cactus Identification


空中サボテンの同定


However, the Fast.ai fam was already there.

しかし、Fast.aiの生徒はすでに集まっていた。

From the forum, "Why are people using Fast.ai getting 1.0 score?"

フォーラムで、Fast.aiを使用している人はなぜ1.0スコアを得ているのですか。




Kurian Benoy
45th place

Why is everyone who is using fast.ai getting 1.0 score?

posted in [Aerial Cactus Identification](#) 2 months ago



I am asking this question because other libraries like keras, kekaskas. .. and all couldn't get a perfect score of 1.

**One comment, "Because of the transforms,
including warping."**

1つのコメントは「ワープを含むデータ増強方法のデフォルトは良いから。」

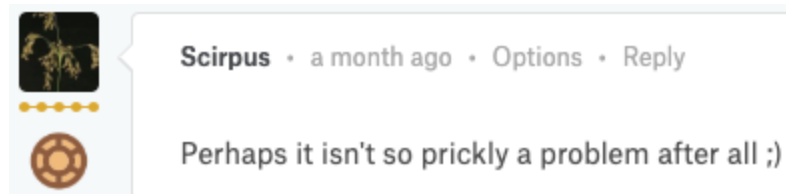


JohnBeuving • (175th in this Competition) • 2 months ago • Options • Reply

Their transform functions are really elaborate (includes eg. warping) and with good defaults.







Another comment, "Perhaps it's not such a prickly problem."

**もう一つのコメントは「おそらく空中サボテンの同定の問題はあまり
難しくない。」**



In any case, I added my name to the leaderboard.

リーダーボードに自分の名前を追加しました。

93	Keewon Shin		1.0000	10	1d
94	gmgm		1.0000	1	18h
95	globophobe		1.0000	9	2h
Your Best Entry ↑					
Your submission scored 1.0000, which is an improvement of your previous score of 0.9999. Great job!					
 Tweet this!					
96	stormchasingg		0.9999	2	3mo
97	a.tanaka		0.9999	1	3mo

Unexpectedly, the 1.0 scores had the same data augmentation parameters.

意外と、1.0のスコアのデータ増強のパラメータは同じでした。

1.0 public kernel augmentation:

1.0の公開カーネルはこれをデータ増強のため：

```
transformations = get_transforms(  
    do_flip=True, flip_vert=True, max_rotate=10.0,  
    max_zoom=1.1, max_lighting=0.2, max_warp=0.2,  
    p_affine=0.75, p_lighting=0.75  
)
```

Except for one parameter, those are all default parameters.

1つを除いて、すべてのパラメータはFast.aiのデフォルトのパラメータです。

This is all you need.

これだけで十分です。

```
transformations = get_transforms(flip_vert=True)
```

flip_vert defaults to False, because not all images are location invariant.

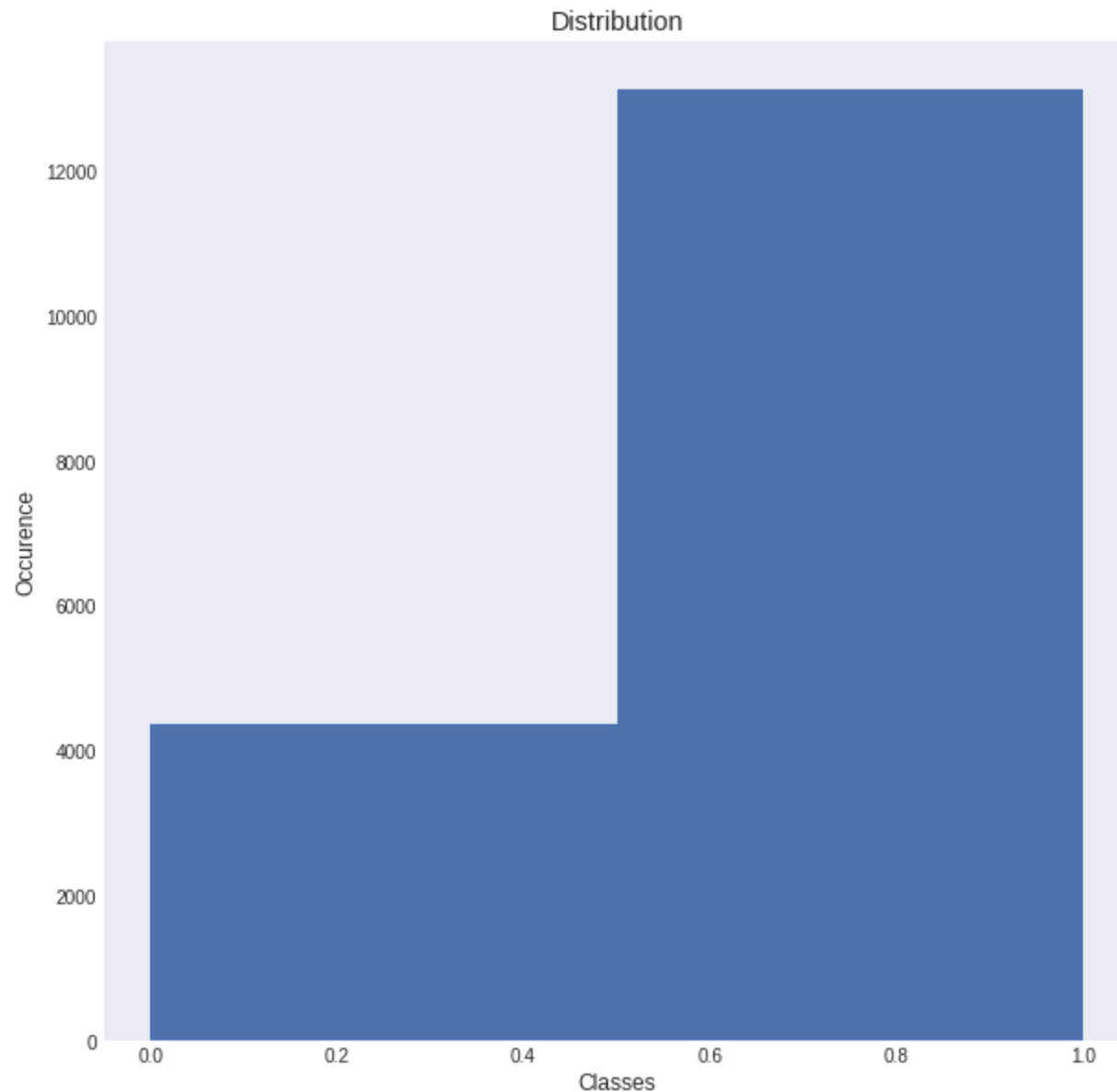
すべての画像は逆さになるのは可能ではないので、flip_vertのデフォルトはFalseです。

That was enough for a 0.9999 score.

それは0.9999のスコアのためは十分でした。

For 1.0, correct imbalanced classes.

1.0の解決策、不均衡なクラスを修正する事でした。



I'm interested to see what happens on the private leaderboard when the contest ends.

Kaggleコンテスト終わってから、プライベートリーダーボードはどうなるかなを楽しみに。

I think it was a good first experience with Kaggle.

最初のKaggleの経験は良かったと思います。

That's all. Thanks for listening.

以上です。ご清聴ありがとうございます。