Google Landmark Challenge

2018/7/19

松井勇佑
国立情報学研究所

Research Prediction Competition

Google Landmark Retrieval Challenge

Given an image, can you find all of the same landmarks in a dataset?

$2,500
Prize Money

Google · 218 teams · 8 days ago

Overview

Image retrieval is a fundamental problem in computer vision: given a query image, can you find similar images in a large database? This is especially important for query images containing landmarks, which accounts for a large portion of what people like to photograph.

In this competition, Kagglers are given query images and, for each query, are expected to retrieve all database images containing the same landmarks (if any).

The new dataset is the largest worldwide dataset for image retrieval research, comprising more than a million images of 15K unique landmarks. We hope that this release will accelerate progress in this important research problem.

This challenge is organized in conjunction with the Landmark Recognition Challenge (https://www.kaggle.com/c/landmark-recognition-challenge). In particular, note that the test set for both challenges is the same, to encourage participants to compete in both. We also encourage participants to use the training data from the recognition challenge to train models which could be useful for the retrieval challenge. Note, however, that there are no landmarks in common between the training/index

NII佐藤真一研チームで参加

- Retrieval task: **7th** place (in 209 teams)
- Recognition task: **7th** place (in 483 teams)

どちらも日本からのチームのうち一位

Aimin Su (University Grenoble Alpes, France)
Steven Ly (University of Southern California, USA)
インターン

Fan Yang (University of Tokyo, NII)
Ryota Hinami (University of Tokyo, NII)
博士学生

Qier Meng (NII)
Sang Phan (NII)
Yusuke Matsui (NII)
Zheng Wang (NII)
ポスドク

Shin’ichi Satoh (NII, University of Tokyo)
教授

Task

$N_q = 100K$ $\quad$ $N_g = 1M$

$K = 100$

N_q

ab4fc	r4hfd, iuk2m, …

- ➢ Report this
- ➢ Evaluated by mAP
- ➢ Groundtruth: same landmark

Task

$K = 100$

N_q

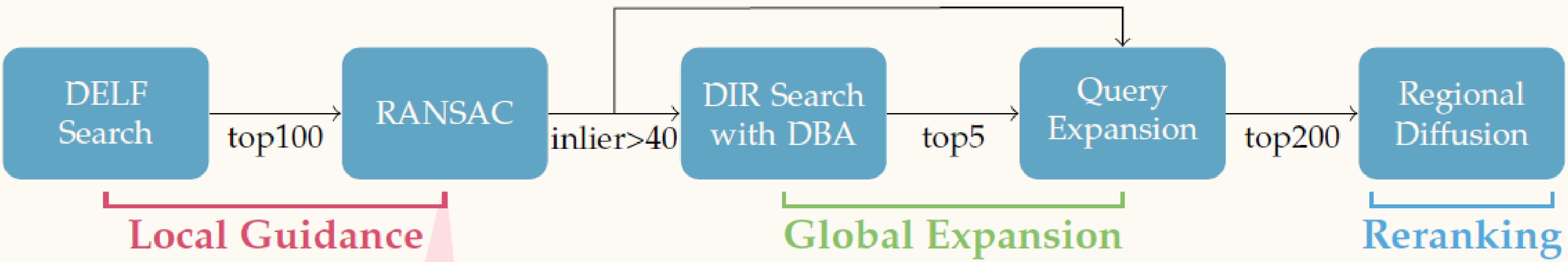
ab4fc	r4hfd, iuk2m, …

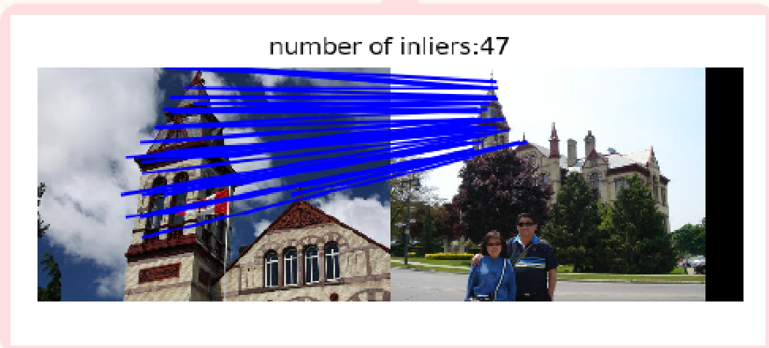
- Report this
- Evaluated by mAP
- Groundtruth: same landmark

$N_q = 100K$ $\quad N_g = 1M$

Query True False

Landmark Retrieval via Local Guidance and Global Expansion





Best result

	$mAP_{@100}$
public[3rd]	**0.600**
private[7th]	**0.561**

Main features:

- **Spatial verification** takes highly confident candidates;
- **Global expansion** allows better generalization;
- **Regional diffusion** reranks results on manifolds.

Overview | Data | Kernels | Discussion | Leaderboard | Rules | Team | My Submissions | Late Submission

Refresh

This competition has completed. This leaderboard reflects the final standings.

In the money | Gold | Silver | Bronze

7位

#	△pub	Team Name	Kernel	Team Members	Score	Entries	Last
1	▲1	CVSSP & Visual Atoms			0.627	176	8d
2	▼1	Layer 6 AI			0.608	200	8d
	3	SevenSpace			0.598	138	9d
		Naver Labs Europe			0.586	138	8d
		VPP			0.583	169	9d
6	▼	VRG Prague			0.576	197	9d
7	▼4	NII Satoh Lab		+6	0.561	185	8d
8	—	ETRI_VB			0.559	181	9d
9	—	VisionIR			0.542	32	8d
10	—	Callidior			0.522	59	10d
11	—	aixili2014			0.495	16	9d
12	▲1	Kohei			0.491	82	14d
13	▼1	MCC			0.491	61	10d

Overview Data Kernels Discussion Leaderboard Rules Team My Submissions Late Submission

Refresh

This competition has completed. This leaderboard reflects the final standings.

In the money Gold Silver Bronze

7位

チェコ工科大学
学画像検索の歴史を作ってきた人たち

#	△pub	Team Name	Kernel	Team Members	Score	Entries	Last
1	▲1	CVSSP & Visual Atoms			0.627	176	8d
2	▼1	Layer 6 AI			0.608	200	8d
	3	SevenSpace					
		Naver Labs Eur					
		VPP					
6	▼	VRG Prague			0.576	197	9d
7	▼4	NII Satoh Lab		+6	0.561	185	8d
8	—	ETRI_VB			0.559	181	9d
9	—	VisionIR			0.542	32	8d
10	—	Callidior			0.522	59	10d
11	—	aixili2014			0.495	16	9d
12	▲1	Kohei			0.491	82	14d
13	▼1	MCC			0.491	61	10d

Overview Data Kernels Discussion Leaderboard Rules Team My Submissions Late Submission

Refresh

This competition has completed. This leaderboard reflects the final standings.

16歳（！）

In the money Gold Silver Bronze

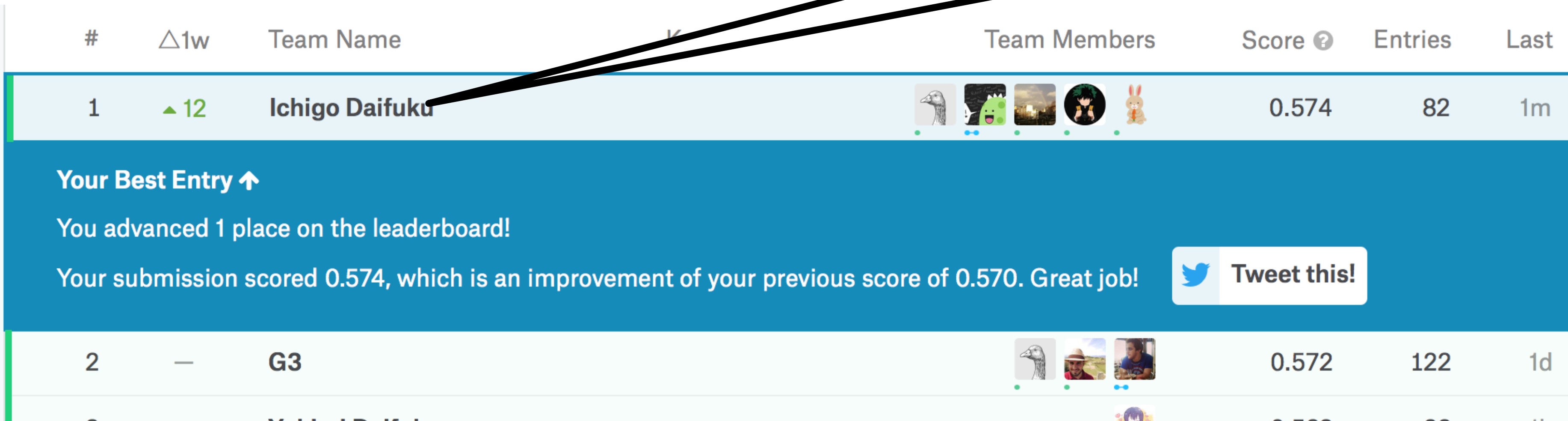
#	△pub	Team Name	Kernel	Team Members	Score	Entries	Last
1	▲1	CVSSP & Visual Atoms			0.627	176	8d
2	▼1	Layer 6 AI			0.608	200	8d
	3	SevenSpace					
		Naver Labs Eur					
		VPP					
6	▼	VRG Prague			0.576	197	9d
7	▼4	NII Satoh Lab		+6	0.561	185	8d
8	—	ETRI_VB			0.559	181	9d
9	—	VisionIR			0.542	32	8d
10	—	Callidior			0.522	59	10d
11	—	aixili2014			0.495	16	9d
12	▲1	Kohei			0.491	82	14d
13	▼1	MCC			0.491	61	10d

7位

チェコ工科大学
学画像検索の歴史を作ってきた人たち

余談

一時期１位だった・・



#	△1w	Team Name	Team Members	Score	Entries	Last
1	▲12	Ichigo Daifuku		0.574	82	1m

Your Best Entry

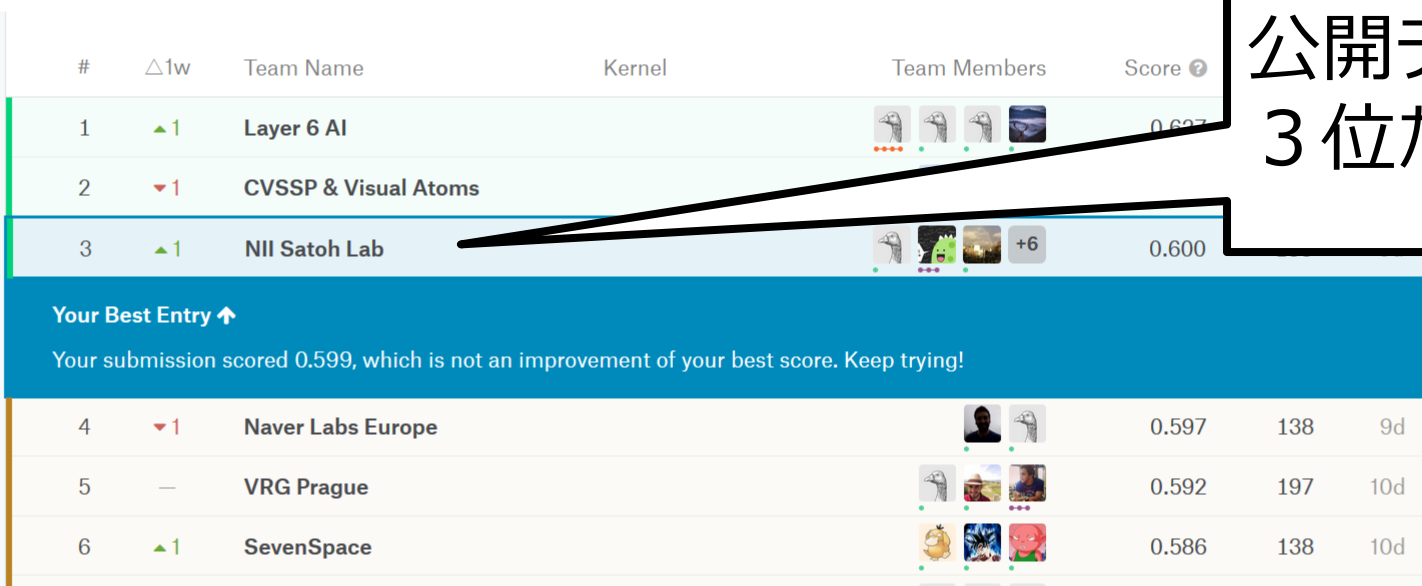
You advanced 1 place on the leaderboard!

Your submission scored 0.574, which is an improvement of your previous score of 0.570. Great job!

Tweet this!

#	△1w	Team Name	Team Members	Score	Entries	Last
2	—	G3		0.572	122	1d

公開データでは
３位だった・・



#	△1w	Team Name	Kernel	Team Members	Score	Entries	Last
1	▲1	Layer 6 AI					
2	▼1	CVSSP & Visual Atoms					
3	▲1	NII Satoh Lab		+6	0.600		

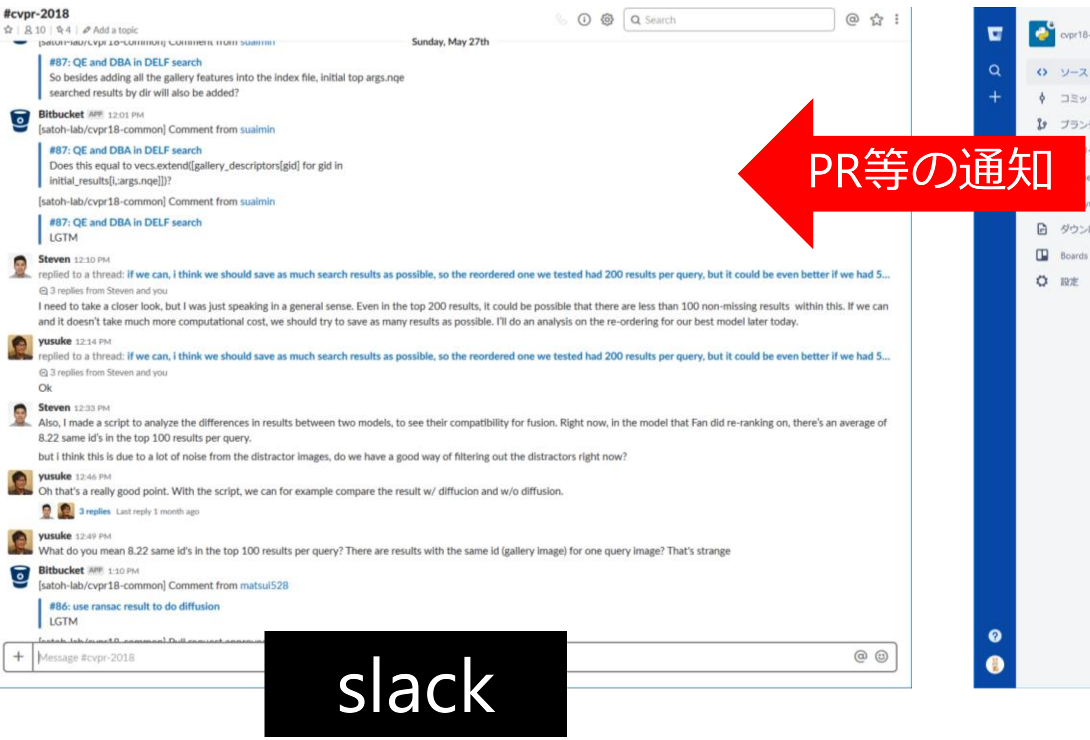
Your Best Entry

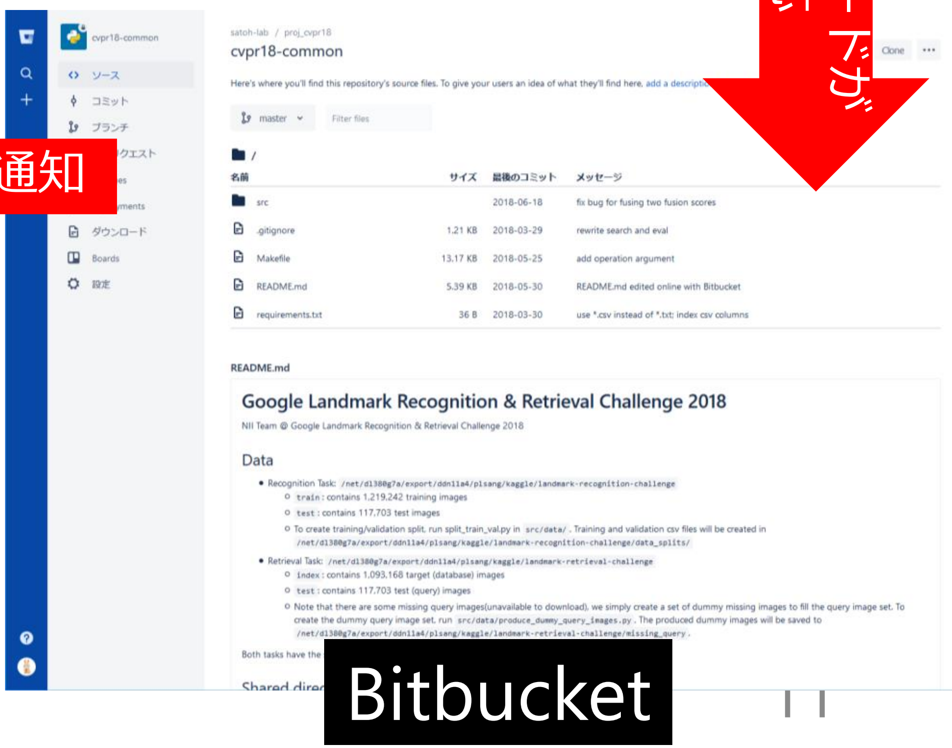
Your submission scored 0.599, which is not an improvement of your best score. Keep trying!

#	△1w	Team Name	Kernel	Team Members	Score	Entries	Last
4	▼1	Naver Labs Europe			0.597	138	9d
5	—	VRG Prague			0.592	197	10d
6	▲1	SevenSpace			0.586	138	10d

管理：

- Kanban (Trello) + Bitbucket + Slack でナウいチーム開発を目指した
- Github Flow
- なるべく処理をパイプライン化して分割することで、みんながコミットできるように
- （アカデミアにしては）実装を頑張ったほうだと思われる





マシン：

ラボGPUサーバ：
- V100 * 8, CPU 80 threads (DGX1)
- P100 * 4, CPU 28 threads
- K80 * 8, CPU 24 threads

ラボCPUサーバ：
- CPU 60 threads

6.4万円/日 🤗

AWS:
- x1e.32xlarge * 1 : vCPU 128, **3904** G mem, 26.7 USD/h
  CPU世代はHaswellで、これはdgx1のBroadwellより前
- m5.24xlarge * 10 : vCPU 96, 384G mem, 4.6 USD/h
  CPU世代は上より新しいSkylake

  1.1万円/日 🤗
- EFSでファイル共有（各ec2インスタンスから共有マウント）

マシン：

ラボGPUサーバ：

- V100 * 8, CPU 80 threads (DGX1)
- P100 * 4, CPU 28 threads
- K80 * 8, CPU 24 threads

ラボCPUサーバ：

- CPU 60 threads

6.4万円/日 🤗

学び：

- マシン世代が違うとCUDAバージョン問題などに気を使う必要がある
- DGX1はCPUもいいので、ついCPUもGPUも使うコード（e.g. faiss-gpu）を書いてしまいがちだが、後にAWSに移す可能性があるなら分けたほうがいい
- どうせ最後にAWSで気合ブーストするので、AWSに移せるように意識してコードを書くといいと思う

/日 🤗

マシン：

ラボGPUサーバ：
- V100 * 8, CPU 80 threads (DGX1)
- P100 * 4, CPU 28 threads
- K80 * 8, CPU 24 threads

ラボCPUサーバ：
- CPU 60 threads

学び：
- CPU世代が違うと速度がかなり違う。スレッド数だけでは決まらない
- 巨大なマシン一台より、ミドルレベル複数台のほうがコスト的にも総メモリ的にも楽
- EBSではEC2をポコポコ増やすときにデータ移動大変。EFS推奨。ただEFSは書き込みが遅いときがあった

AWS:
- x1e.32xlarge * 1 : vCPU 128, **3904** G mem, 26.7 USD/h CPU世代はHaswellで、これはdgx1のBroadwellより前

6.4万円/日 🤗

- m5.24xlarge * 10 : vCPU 96, 384G mem, 4.6 USD/h CPU世代は上より新しいSkylake

1.1万円/日 🤗

- EFSでファイル共有（各ec2インスタンスから共有マウント）

MIRU 2018 (北海道)で発表します

CONFIDENTIAL EXTENDED ABSTRACT.
DO NOT DISTRIBUTE ANYWHERE.

The 21st Meeting on Image Recognition and Understanding

Landmark Retrieval via Local Guidance and Global Expansion

Aimin Su[1,2] Steven Ly[3,2] Fan Yang[4,2] Ryota Hinami[4,2] Qier Meng[2] Sang Phan[2]
Yusuke Matsui[2] Zheng Wang[2] Shin'ichi Satoh[2,4]

Abstract

Google recently held a competition on the Google-Landmarks dataset, the largest worldwide dataset of images for large-scale image retrieval research. We proposed our own model, which incorporated different modern image retrieval techniques together, and finished in 7^{th} place in the competition. We present our full image retrieval model and its results, as well as go over the challenges we faced in the competition.

1. Introduction

With the explosive growth of multimedia data on the web, image retrieval is becoming a fundamental problem in computer vision. The task is straightforward: given a

3. Proposed pipeline

The pipeline of our image retrieval system is illustrated in Fig. 1. It consists of five key steps: **(1)** Deep local feature (DELF) search. **(2)** Sptatial verification (RANSAC) on top-100 results of (1). **(3)** Deep image retrieval (DIR) search with database-side feature augmentation (DBA). **(4)** Query expansion with top-5 results of (3) and the results of (2) with inlier > 40. **(5)** Re-ranking by regional diffusion. In the next sections, we go over the details of the different components of our pipeline, and also explain how they tie together.

3.1 Image Representation