Higher-order Statistical Modeling based Deep CNNs （Part-IV）

Code and Challenge Achievements

Qilong Wang, Jiangtao Xie

2018-11-23

Code and Challenge Achievements

- Code
  - MatConvNet toolkit
    - RAID-G
    - G2DeNet
    - MPN-COV
    - Fast MPN-COV (i.e., iSQRT-COV)
  - PyTorch
    - Fast MPN-COV
- Challenge Achievements
  - ALISC: Alibaba Large-scale Image Search Challenge (**2nd place**)
  - iNat Challenge@FGVC5 (**1st place**)

Code and Challenge Achievements

- Code
  - MatConvNet toolkit
    - RAID-G
    - G2DeNet
    - MPN-COV
    - Fast MPN-COV (i.e., iSQRT-COV)
  - PyTorch
    - Fast MPN-COV
- Challenge Achievements
  - ALISC: Alibaba Large-scale Image Search Challenge (**2nd place**)
  - iNat Challenge@FGVC5 (**1st place**)

Code and Challenge Achievements

- RAID-G
  - Paper:

    RAID-G: Robust Estimation of Approximate Infinite Dimensional Gaussian with Application to Material Recognition.
  - Language: Matlab
  - Framework: MatConvNet-1.0-beta20
  - With an open box demo. (FMD dataset)
  - For classification, we employ LibSVM package which was self-contained in our code.
  - More details: *www.peihuali.org/publications/RAID-G/RIAD-G_V1.zip*

Code and Challenge Achievements

- G^2DeNet
  - Paper:

    G^2DeNet: Global Gaussian Distribution Embedding Network and Its Application to Visual Recognition.
  - Language: Matlab
  - Framework: MatConvNet-1.0-beta20
  - With an open box demo. (3 fine-grained benchmarks, i.e., Birds, Cars and Aircrafts)
  - More details: *www.peihuali.org/publications/G2DeNet/G2DeNet-FGVC-v1.0.zip*

Code and Challenge Achievements

- MPN-COV and Fast MPN-COV
  - Paper:

    MPN-COV: *Is Second-order Information Helpful for Large-scale Visual Recognition?*

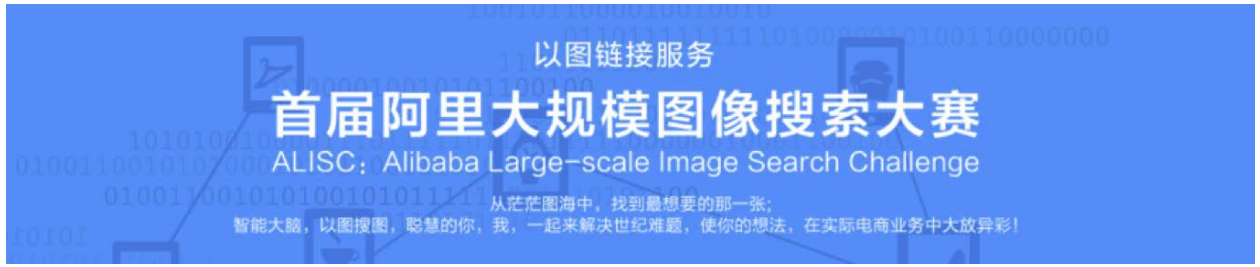
    Fast MPN-COV: *Towards Faster Training of Global Covariance Pooling Networks by Iterative Matrix Square Root Normalization.*
  - Language: CUDA C/C++, Python, Matlab
  - Framework: PyTorch-0.4.0 and MatConvNet-1.0-beta22
  - Demo for training from scratch on ImageNet-2012
  - Transfer learning demo for Fine-grained datasets
  - More details: *http://peihuali.org/iSQRT-COV/index.html*

Code and Challenge Achievements

- Code
  - MatConvNet toolkit
    - RAID-G
    - G2DeNet
    - MPN-COV
    - Fast MPN-COV (i.e., iSQRT-COV)
  - PyTorch
    - Fast MPN-COV
- Challenge Achievements
  - ALISC: Alibaba Large-scale Image Search Challenge (**2nd place**)
  - iNat Challenge@FGVC5 (**1st place**)

Code and Challenge Achievements

- ALISC (2nd place)



Alibaba Group

证 书

2015阿里巴巴大规模图像搜索大赛
ALISC:Alibaba Large-scale Image Search Challenge

参赛队伍
DLUT_VLG
参赛成员

曾辉 Hui Zeng　孙伟健 Weijian Sun　王旗龙 Qilong Wang

导师：李培华

在全球843支队伍中脱颖而出,荣获二等奖，总排名

第二名

阿里巴巴图像大赛组委会
2016年1月

Code and Challenge Achievements

- ALISC (2nd place)

排行榜

第 2 赛季排行榜　第 1 赛季排行榜

排名	参赛者	所在组织	评分	提交次数	最优成绩提交日
1	VIPL-3L	中国科学院	0.4929	24.00	2015-12-16
2	DLUT_VLG	大连理工大学	0.4916	27.00	2015-12-16
3	Choroi	复旦大学	0.4554	21.00	2015-12-16
4	Hitsz_BCC	哈尔滨工业大学	0.4195	21.00	2015-12-16
5	ToSsBoY	复旦大学	0.3801	8.00	2015-12-16
6	KNIGHT-BUPT	北京邮电大学	0.3720	23.00	2015-12-16
7	tjucs_lemon	天津大学	0.3579	14.00	2015-12-16

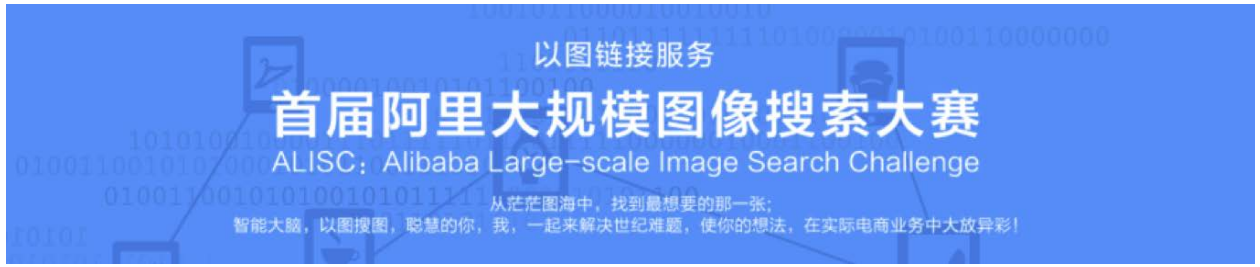
843支队伍参加

—国内高校：清华、北大、中科院、上交、复旦、浙大、中山、哈工大、大连理工、香港城市大学等

— 海外高校：NUS、UIUC、University of Sydney等.

Code and Challenge Achievements

- ALISC (2nd place)



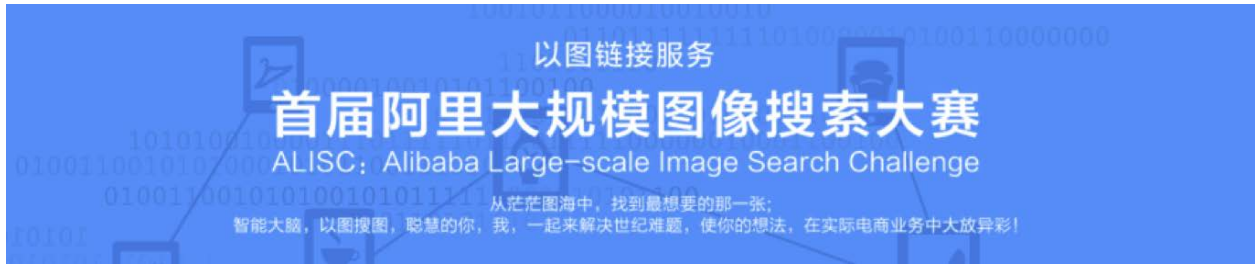
Alibaba Group

数量统计

		初赛	决赛
训练集		971465	1950558
标签	大类目	10	10
	小类目	559	604
	属性值种类	156	157
评测集		1069124	3195334
query集		1559	3567
验证query集		500	1417
代码验证集		210	210

Code and Challenge Achievements

- ALISC (2nd place)



Alibaba Group

赛制安排 (2015年)

初赛赛季，9月1日—10月31日，Top 20入围

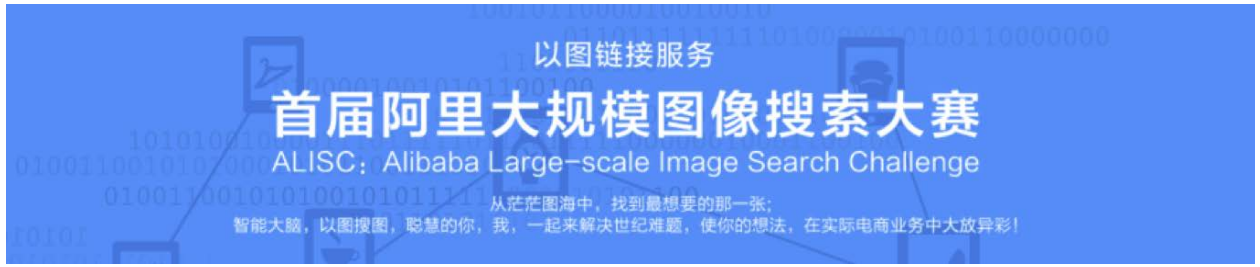
决赛赛季，11月9日-12月9日

数量统计

		初赛	决赛
训练集		971465	1950558
标签	大类目	10	10
	小类目	559	604
	属性值种类	156	157
评测集		1069124	3195334
query集		1559	3567
验证query集		500	1417
代码验证集		210	210

Code and Challenge Achievements

- ALISC (2nd place)



Alibaba Group

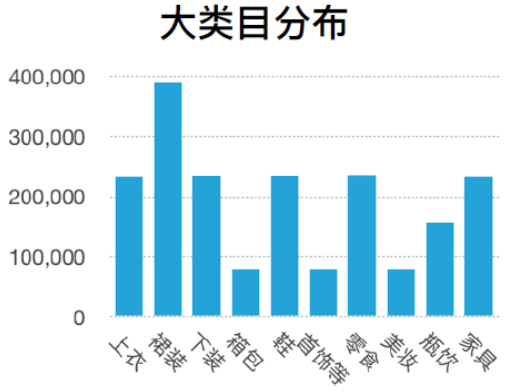
赛制安排 (2015年)

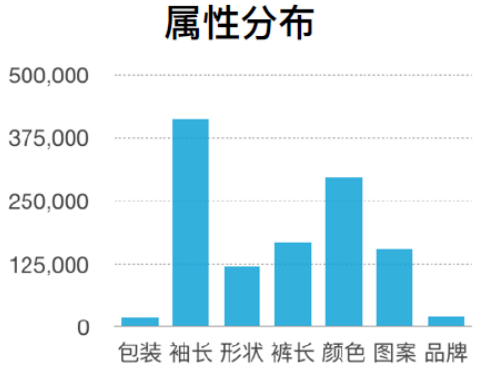
初赛赛季，9月1日—10月31日，Top 20入围

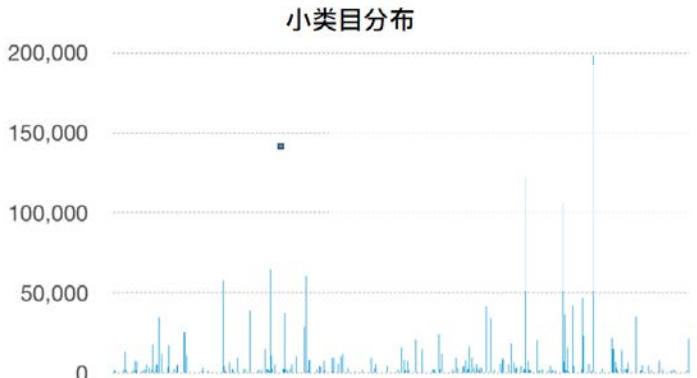
决赛赛季，11月9日-12月9日

数量统计

		初赛	决赛
训练集		971465	1950558
标签	大类目	10	10
	小类目	559	604
	属性值种类	156	157
评测集		1069124	3195334
query集		1559	3567
验证query集		500	1417
代码验证集		210	210





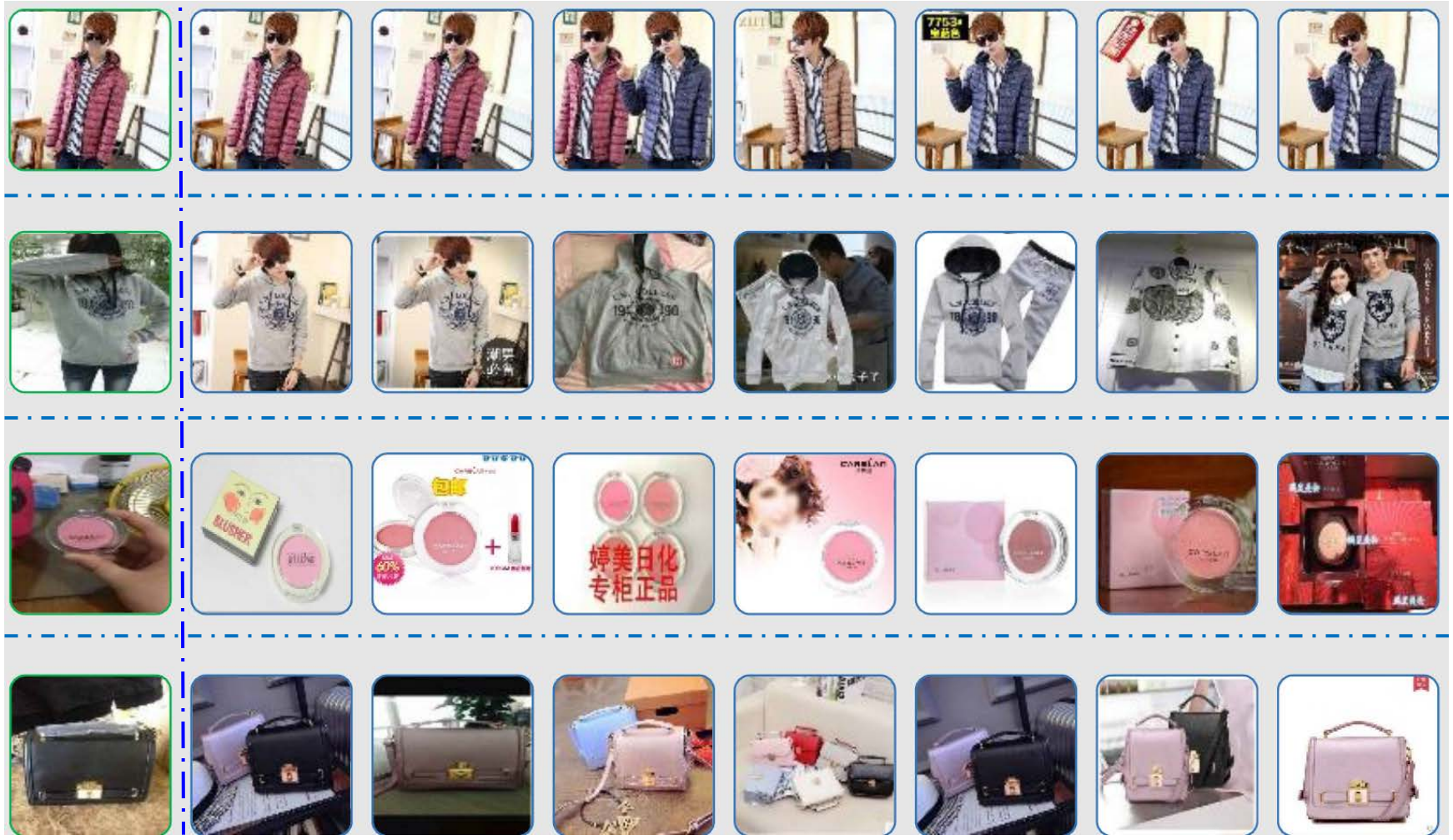


Code and Challenge Achievements

- ALISC (2nd place)

Query Images

Retrieved Images from Dataset



Code and Challenge Achievements

- ALISC (2nd place)

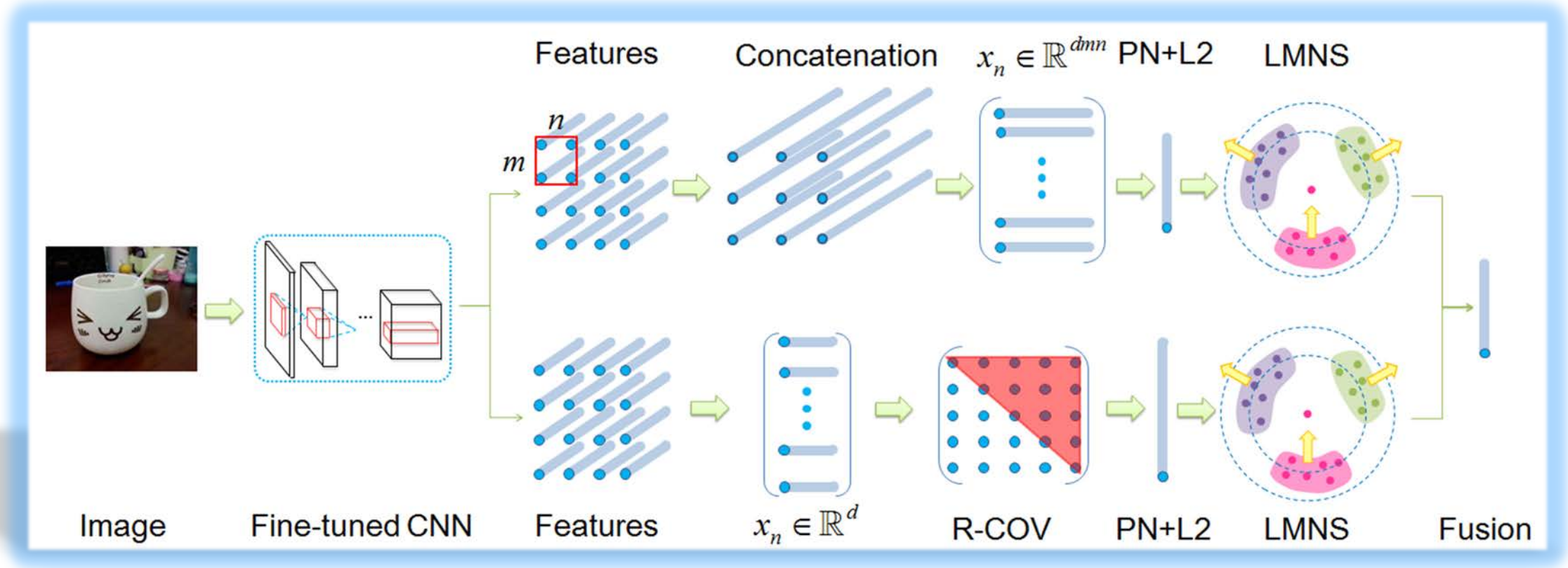
There are diversities between inter-class

There are differences between Query and test set

Code and Challenge Achievements

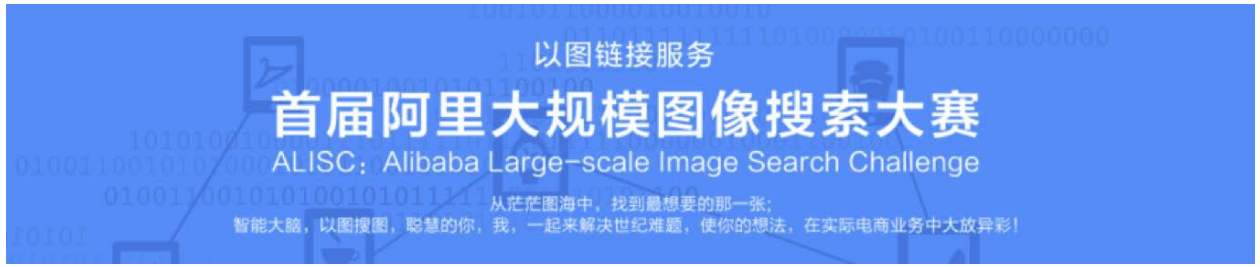
- ALISC (2nd place)

Method overview



Code and Challenge Achievements

- ALISC (2nd place)



Alibaba Group

Representative Publications

Qilong Wang, Peihua Li, Wangmeng Zuo, Lei Zhang. RAID-G: Robust Estimation of Approximate Infinite Dimensional Gaussian with Application to Materiel Recognition. In *CVPR*, 2016.

Zeng Hui, Qilong Wang, Peihua Li. Beyond NBNN: Large Margin Nearest Subspace Classification with Deep Representations. Submitted.

李培华, 王旗龙, 曾辉, 孙伟健, 鲁潇潇. 一种大规模数据背景下的快速有效的图像检索方法. 中国发明专利, 申请日:2016.5.23, 授权日：2018.2.6, 专利号：ZL201610340978.1.

Code and Challenge Achievements

- Code
  - MatConvNet toolkit
    - RAID-G
    - G2DeNet
    - MPN-COV
    - Fast MPN-COV (i.e., iSQRT-COV)
  - PyTorch
    - Fast MPN-COV
- Challenge Achievements
  - ALISC: Alibaba Large-scale Image Search Challenge (**2nd place**)
  - iNat Challenge@FGVC5 (**1st place**)

Code and Challenge Achievements

- iNat Challenge@FGVC5 CVPR 2018 SALT LAKE CITY • JUNE 18-22 (1st place)

Organized by

Google

Cornell University

Caltech

Supported by

Microsoft

https://sites.google.com/view/fgvc5/competitions/inaturalist

[MPN-COV] Peihua Li, Jiangtao Xie, et al. Is Second-order Information Helpful for Large-scale Visual Recognition? In ICCV, 2017.
[Fast MPN-COV] Peihua Li, Jiangtao Xie, et al. Towards Faster Training of Global Covariance Pooling Networks by Iterative Matrix Square Root Normalization. In *CVPR*, 2018.

Code and Challenge Achievements

- iNat Challenge@FGVC5 CVPR 2018 SALT LAKE CITY • JUNE 18-22 (1st place)

[MPN-COV] Peihua Li, Jiangtao Xie, et al. Is Second-order Information Helpful for Large-scale Visual Recognition? In ICCV, 2017.
[Fast MPN-COV] Peihua Li, Jiangtao Xie, et al. Towards Faster Training of Global Covariance Pooling Networks by Iterative Matrix Square Root Normalization. In *CVPR*, 2018.

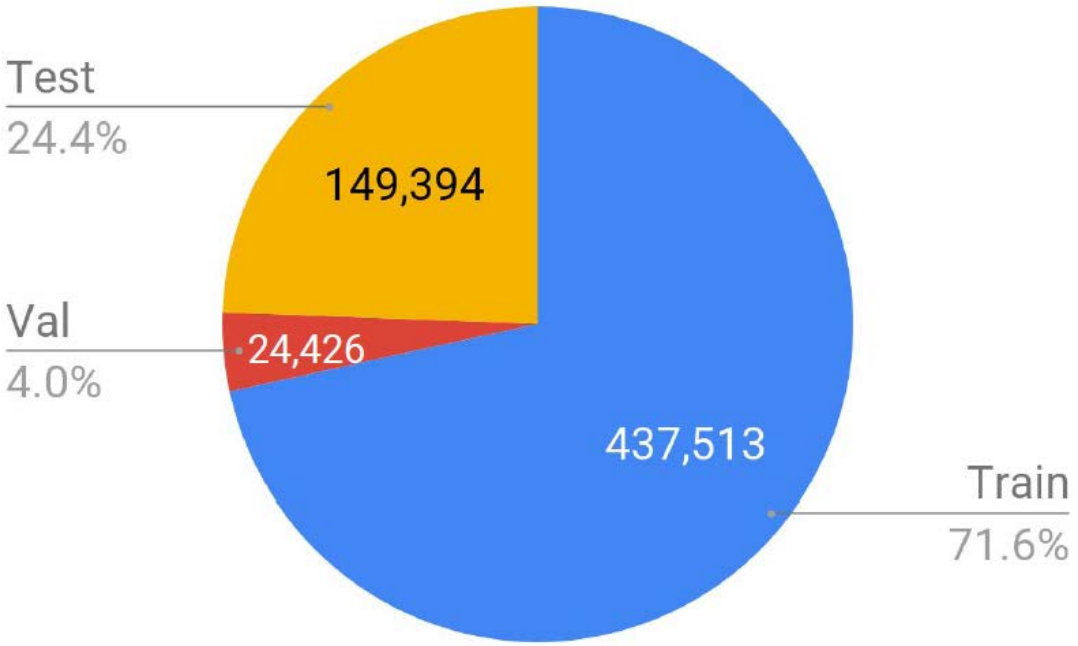
Code and Challenge Achievements

- iNat Challenge@FGVC5 

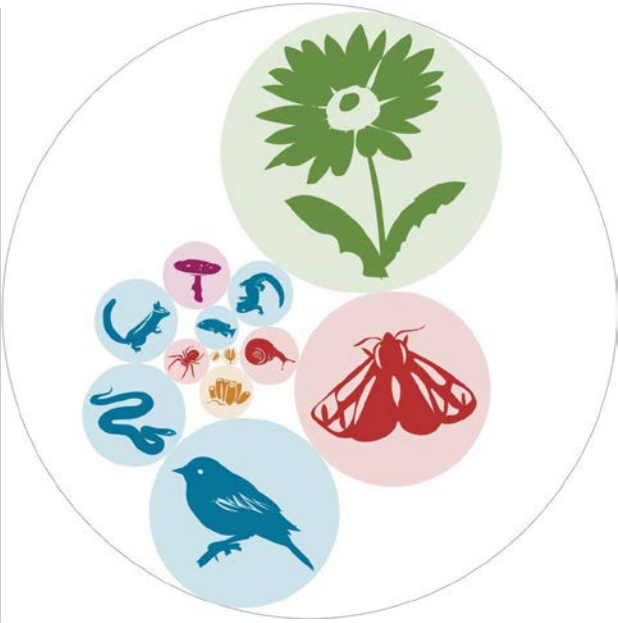
(1st place)

Number of Classes: 8,142

Number of Images: 611,333

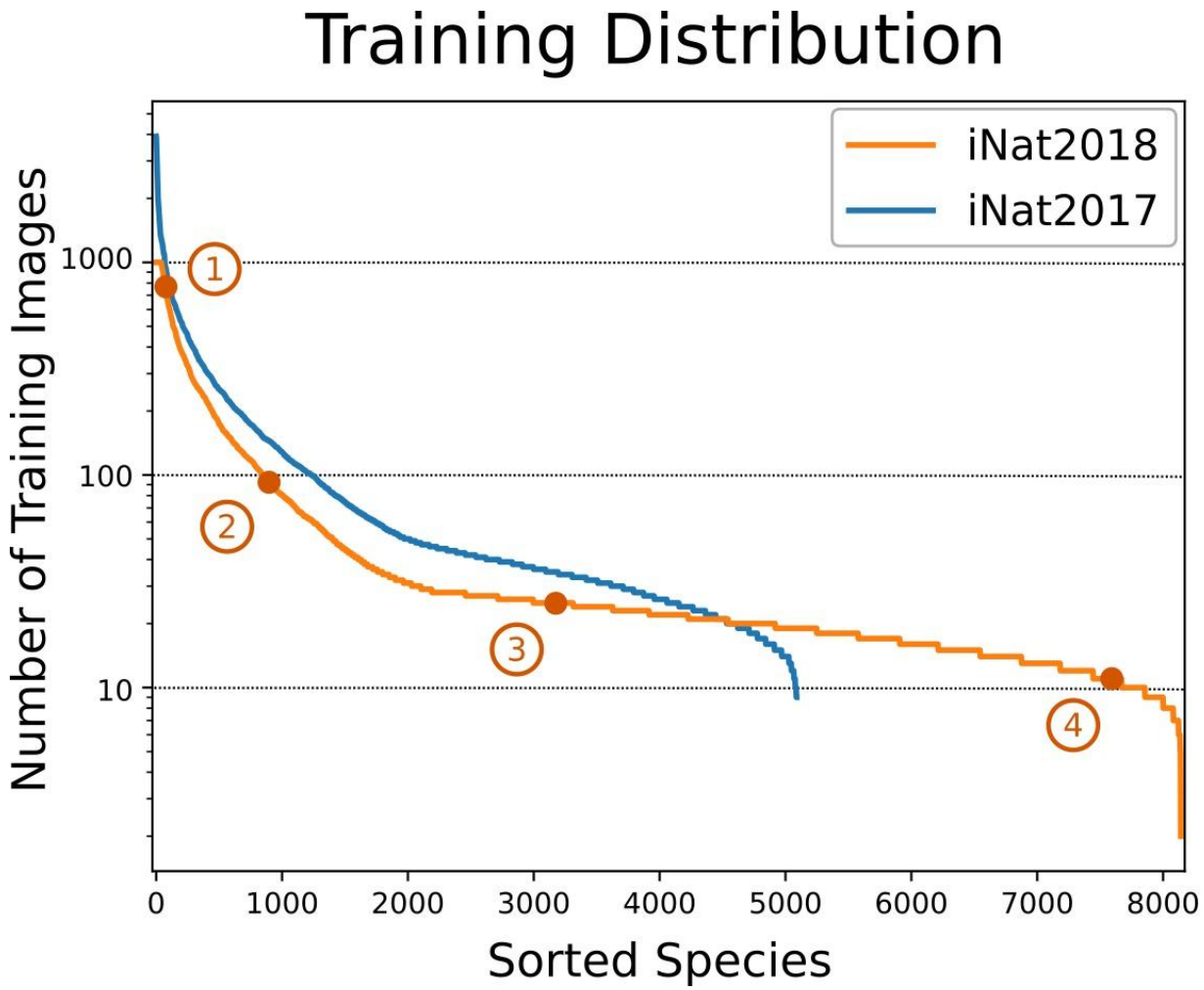


Super Cat	Num Classes
Bacteria	1
Protozoa	4
Insecta	2,031
Aves	1,258
Plantae	2,917
Chromista	25
Animalia	178
Arachnida	114
Fungi	321
Actinopterygii	369
Mammalia	234
Reptilia	284
Mollusca	262
Amphibia	144



Code and Challenge Achievements

- iNat Challenge@FGVC5 CVPR 2018 SALT LAKE CITY • JUNE 18-22 (1st place)



① Cooper's Hawk

② American Bison

③ Mallow Bindweed

④ Island Fox

Code and Challenge Achievements

- iNat Challenge@FGVC5 CVPR 2018 (1st place)

Aphonopelma hentzi

Aphonopelma chalcodes

Commonly confused classes Top1 Public for DLUT_VLG

Code and Challenge Achievements

- iNat Challenge@FGVC5 CVPR 2018 SALT LAKE CITY • JUNE 18-22 (1st place)

Acronicta lobeliae

Acronicta hasta

Commonly confused classes Top1 Public for DLUT_VLG

Code and Challenge Achievements

- iNat Challenge@FGVC5 CVPR 2018 SALT LAKE CITY • JUNE 18-22 (1st place)

Papilio thoas

Papilio rumiko

Commonly confused classes Top1 Public for DLUT_VLG

Code and Challenge Achievements

- iNat Challenge@FGVC5 

(1st place)

Lepus alleni

Lepus californicus

Commonly confused classes Top1 Public for DLUT_VLG

iNaturalist Competition 2018 - Results

Rank	Team	Affiliation
1st	Shuyu Ge, Qiule Sun Jiangtao Xie, Peihua Li	Dalian University of Technology, China
2nd	Jeremy Trammell, Priyanka Oberoi, John Kaufhold	Deep Learning Analytics
3rd	Kaiyu Yue, Ming Sun, Ti Bai, Xiao Tan, Canxiang Yan, Yuchen Yuan, Yingying Li, Xiao Liu, Feng Zhou	Baidu VIS

kaggle

Microsoft

Code and Challenge Achievements

- iNat Challenge@FGVC5 CVPR 2018 SALT LAKE CITY • JUNE 18-22 (1st place)

DLUT_VLG performs relatively 8% in top-3 error better than 2nd place

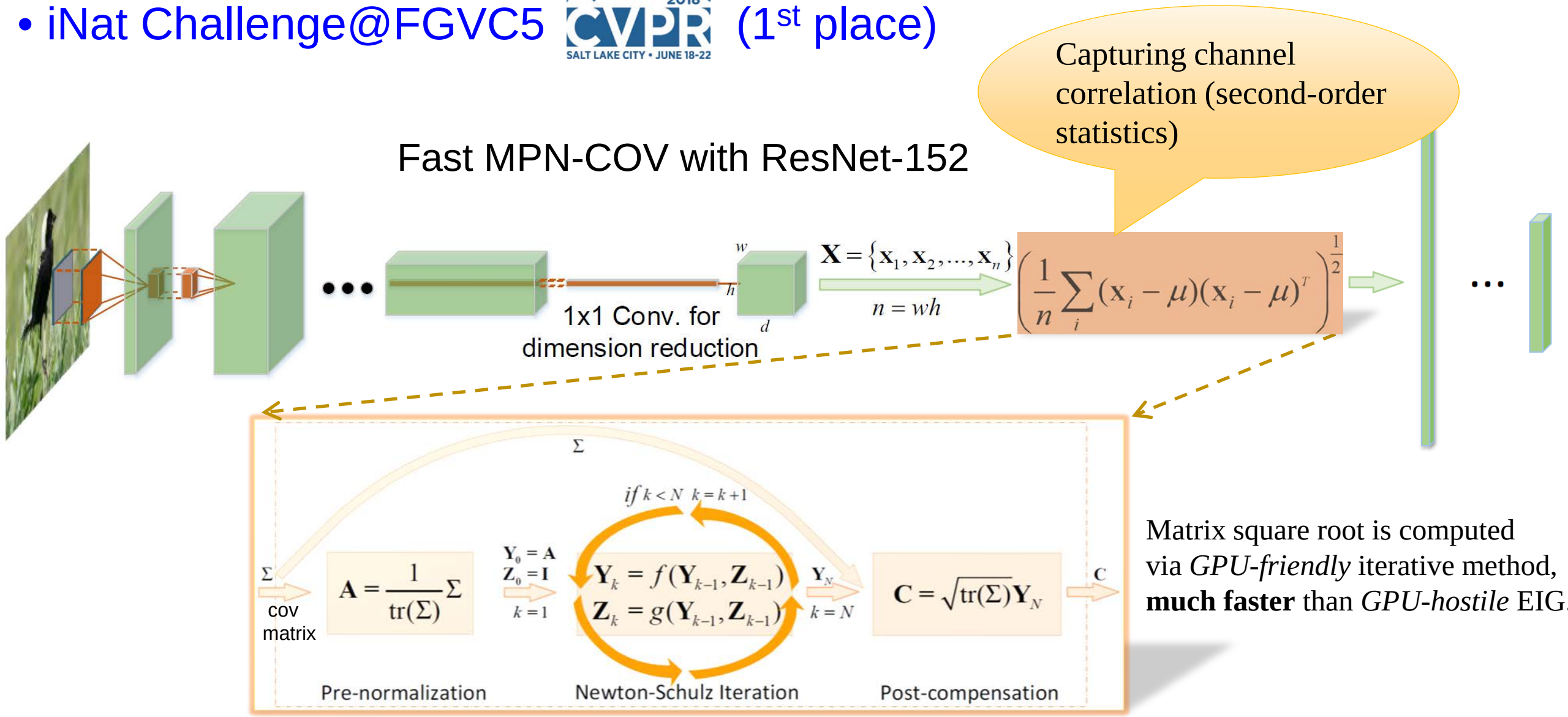
Public Leaderboard | Private Leaderboard

The private leaderboard is calculated with approximately 50% of the test data. Refresh

#	▵pub	Team Name	Kernel	Team Members	Score	Entries	Last
1	—	DLUT_VLG (Dalian University ...			0.12858	133	11d
2	—	Deep Learning Analytics			0.13981	93	11d
3	—	fadivugibs			0.14618	79	11d
4	—	CMP			0.16076	14	11d
5	▲1	fISHpAM			0.16892	3	23d
6	▼1	traveler			0.16988	30	11d
7	—	yen			0.17201	20	11d

Code and Challenge Achievements

- iNat Challenge@FGVC5 (1st place)



Matrix square root is computed via *GPU-friendly* iterative method, **much faster** than *GPU-hostile* EIG.

Code and Challenge Achievements

- iNat Challenge@FGVC5 

(1st place)

MPN-COV with ResNet-152 architecture, fine-tuned on iNat 2018

- Implementation
- Three useful tricks for performance boost
  - Exploit higher resolution images
  - Deal with long tailed distribution
  - Usage of iNat2017 dataset
- Results

[MPN-COV] Peihua Li, Jiangtao Xie, Qilong Wang and Wangmeng Zuo. Is Second-order Information Helpful for Large-scale Visual Recognition? In *ICCV*, 2017.

[Fast MPN-COV] Peihua Li, Jiangtao Xie, Qilong Wang and Zilin Gao. Towards Faster Training of Global Covariance Pooling Networks by Iterative Matrix Square Root Normalization. In *CVPR*, 2018.

Code and Challenge Achievements

- iNat Challenge@FGVC5 

(1st place)

➢Pre-trained ResNet-152 on **ImageNet-11k** and then finetuned on iNaturalist 2017 dataset

➢ Two stage MPN-COV training on iNaturalist 2018 dataset

- Firstly, fine-tune MPN-COV module–1x1 Conv. layer for dimension reduction, and 8142-way FC connecting normlized COV to output.
- Next, fine-tune final 9 residual blocks and the subsequent MPN-COV module.

➢Fast MPN-COV module

- Iteration number: 3
- Dimension of input: 160 (after a 1x1x2048x160 Conv. layer)
- Dimension of image representation: 12,880

➢Dense crop + multiple scales on test images for inference

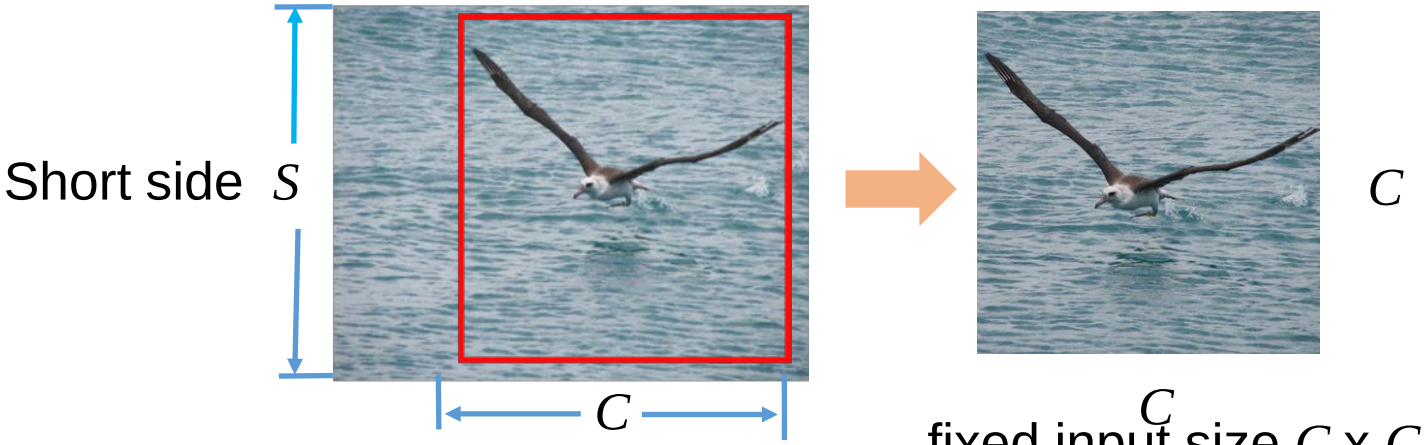
- Analogous to the method in Simonyan & Zisserman (ICLR 2015).

➢Data augmentation with MatConvNet default

Code and Challenge Achievements

- iNat Challenge@FGVC5 (1st place)
  - Randomly crop a *C*x*C* image from resized image with shorter side *S*,

    where C=224,320,392 and S=256, 360, 448, respectively



Top-3 errors (%) with single model on test set using varying input size

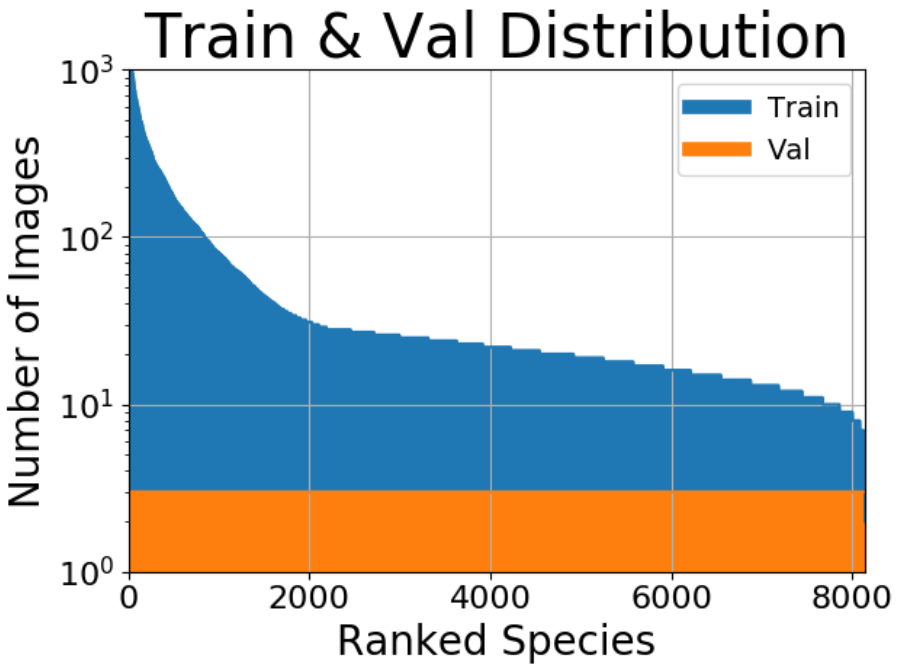
Input size *C*	MPN-COV with ResNet-152	Vanilla ResNet-152
320x320	**15.038**	16.623
392x392	**14.704**	16.024

Code and Challenge Achievements

- iNat Challenge@FGVC5 

(1st place)

Following Cui et al., we use validation set with balanced distribution to fine-tune with smaller learning rate (2e-4)



Top-3 errors (%) on test set (center 320x320 crop, *S*=360)

Finetuning on val. set	MPN-COV with ResNet-152	Vanilla ResNet-152
NO	**17.875**	18.770
YES	**15.038 (2.837 ↑)**	16.623 (2.147 ↑)

Yin Cui, Yang Song, Chen Sun, Andrew Howard, Serge Belongie. Large Scale Fine-Grained Categorization and Domain-Specific Transfer Learning. In *CVPR*, 2018

Code and Challenge Achievements

- iNat Challenge@FGVC5 

(1st place)

We fine-tune on iNat Challenge 2017 dataset before training on iNat Challenge 2018 dataset

Top-3 errors (%) on test set (center 224x224 crop, *S*=256)

Finetuning on iNat2017	MPN-COV with ResNet-152	Vanilla ResNet-152
NO	N/A	25.451
YES	N/A	24.660 (0.7911 ↑)

Code and Challenge Achievements

- iNat Challenge@FGVC5 

(1st place)

Evaluation on test set with *single model* using varying image resolution

Method	Input size *C*	Fusion scales *S*	Top-3 error (%)
MPN-COV with ResNet-152	320x320	360,480,512	15.038
	392x392	448,544,608	**14.704**
Vanilla ResNet-152	320x320	380,480	16.623
	392x392	480,576	16.024

Ensemble of three models with input size 392x392

Method	Top-3 error (%)
MPN-COV with ResNet-152	**13.499**
Vanilla ResNet-152	14.625
MPN-COV+Vanilla ResNet-152	**13.103**

Note: Fusion of 6 MPN-COV models (3 w/ *C*=392, 3 w/ *C*=320) and 6 vanilla ResNet-152 models (3 w/ *C*=392, 3 w/ 320), the error decreases to 13.068%.

Code and Challenge Achievements

- iNat Challenge@FGVC5 

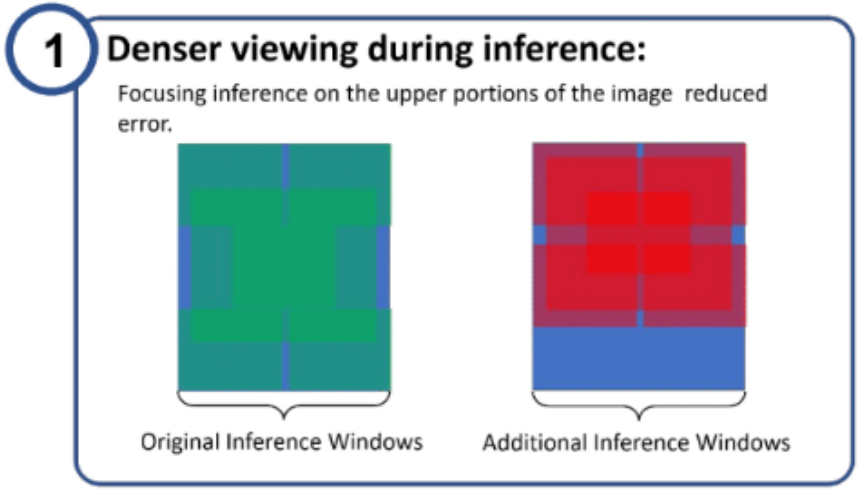
(1st place)
  - ➢ Matrix Power Normalized COVariance pooling (MPN-COV) networks are compelling for large-scale classification
  - ➢ **Potentially**
    - — Pretrained MPN-COV on *ImageNet* will generalize better
    - — Ensemble of MPN-COV with varying architectures (e.g. ResNet and Inception) will further improve
    - — Fast MPN-COV with more iterations (>3) will benefit
    - — Higher resultion images (>392) will be more helpful

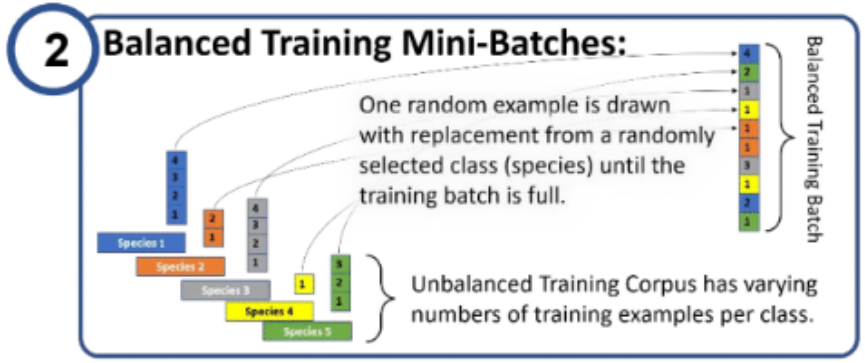
[MPN-COV] Peihua Li, Jiangtao Xie, Qilong Wang and Wangmeng Zuo. Is Second-order Information Helpful for Large-scale Visual Recognition? In *ICCV*, 2017.

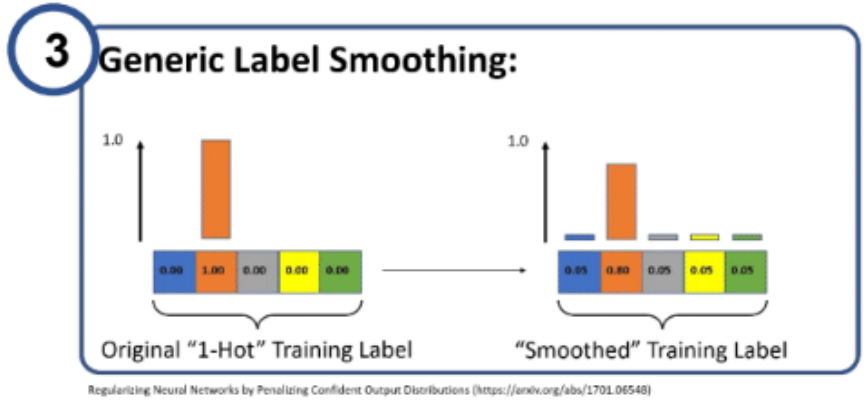
[Fast MPN-COV] Peihua Li, Jiangtao Xie, Qilong Wang and Zilin Gao. Towards Faster Training of Global Covariance Pooling Networks by Iterative Matrix Square Root Normalization. In *CVPR*, 2018.

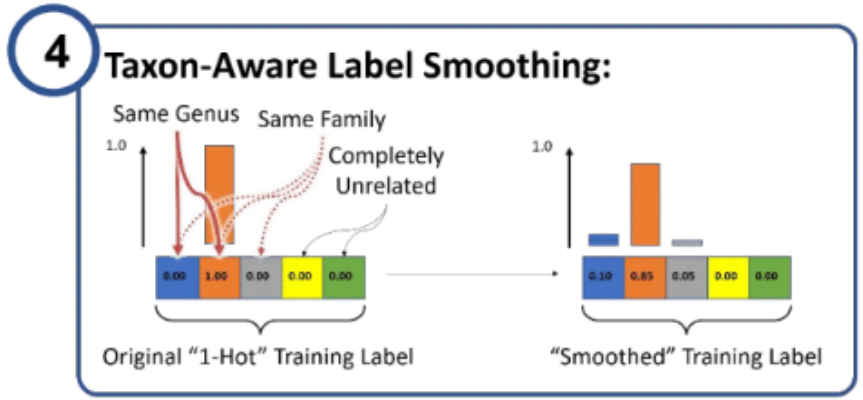
Code and Challenge Achievements

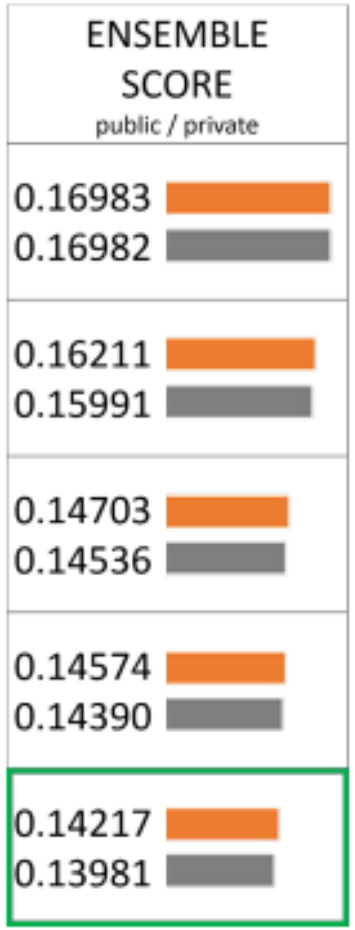
- iNat Challenge@FGVC5 (1st place)
- Method overview of other competitor (Deep Learning Analytics 2nd place)











Our single model result: 14.704 (public)

Thank you!