画像認識コンペティションの取り組み方

郁　青

The University of Tokyo

Self-introduction

- Name: 郁 青 Yu Qing
- Account: YK 
- Hometown: Shanghai, China
- Affiliation: Aizawa Laboratory, The University of Tokyo
- Grade: D1
- HP: https://yu1ut.com/

- Research interest: Image Recognition

Self-introduction

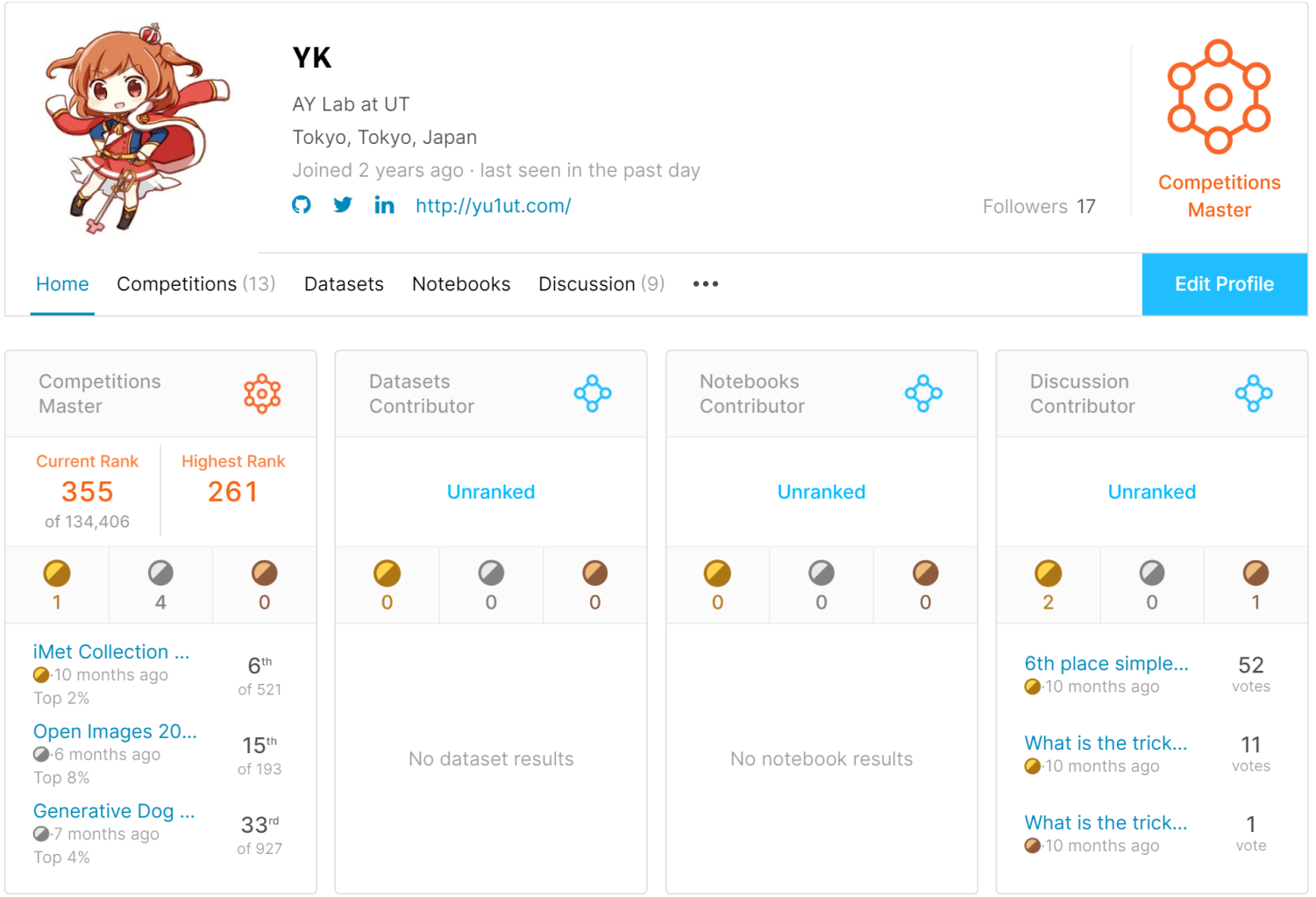
- Signate

コンペティション名	開催期間	賞金総額	戦績	
			投稿数	最終順位
The 1st Tellus Satellite Challenge	2018/10/16～2018/12/7	総額¥2,000,000	97件	1 位 / 174人投稿
ステアラボ メテオサーチチャレンジ	2018/3/3～2018/6/30	総額¥530,000	22件	14 位 / 83人投稿
産業技術総合研究所 衛星画像分析コンテスト	2018/2/27～2018/4/26	総額¥800,000	48件	1 位 / 66人投稿
JSAI Cup 2018 人工知能学会データ解析コンペティション	2018/1/19～2018/3/29	総額¥1,000,000	87件	1 位 / 121人投稿

Self-introduction

- Kaggle
  - 2019/06 : iMet Collection
  - 2019/08: Generative Dog
  - 2019/09: Recursion Cellular Image Classification

  -> Kaggle Master
  - 2019/10: Open Images 2019 - Instance Segmentation
  - 2020/03: Bengali.AI Handwritten Grapheme Classification



Kaggle (Image recognition competitions)

Your Competitions

Active | Closed | Pinned | Hosted

SIIM-ISIC Melanoma Classification Identify melanoma in lesion images Featured • 2 months to go • 726 Teams		**$30,000**
TReNDS Neuroimaging Multiscanner normative age and assessments prediction with brain function, structure, and connectivity Research • 21 days to go • 626 Teams		**$25,000**
ALASKA2 Image Steganalysis Detect secret data hidden within digital images Research • a month to go • 495 Teams		**$25,000**
Prostate cANcer graDe Assessment (PANDA) Challenge Prostate cancer diagnosis using the Gleason grading system Featured • a month to go • Code Competition • 549 Teams		**$25,000**
Global Wheat Detection Can you help identify wheat heads using image analysis? Research • 2 months to go • Code Competition • 752 Teams		**$15,000**
Open Images Object Detection RVC 2020 edition Detect objects in varied and complex images Playground • 2 months to go • 24 Teams		**Knowledge**
Open Images Instance Segmentation RVC 2020 edition Outline segmentation masks of objects in images Playground • 2 months to go • 5 Teams		**Knowledge**

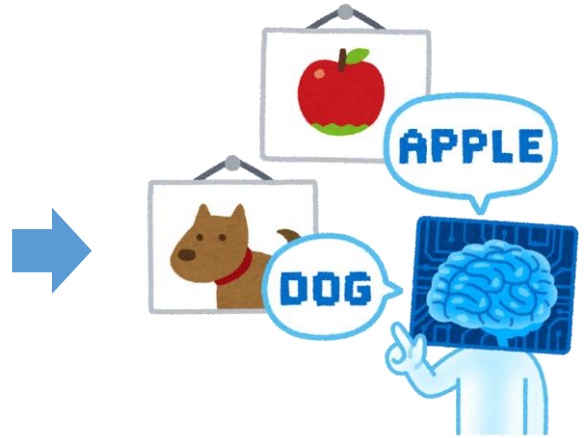
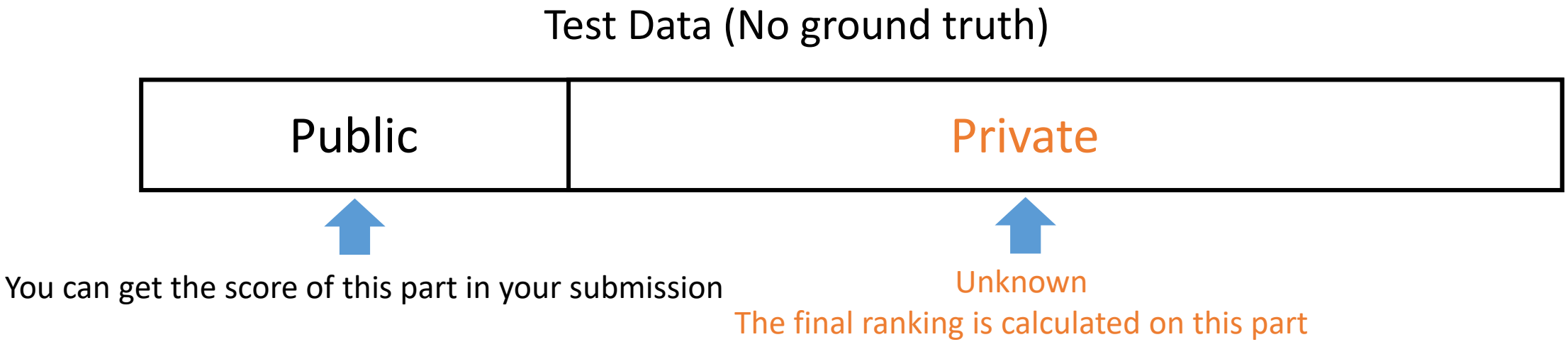


Image Recognition Model

Kaggle

- Evaluation of the model's performance:



Kaggle

- Competition Medals

	0-99 Teams	100-249 Teams	250-999 Teams	1000+ Teams
Bronze	Top 40%	Top 40%	Top 100	Top 10%
Silver	Top 20%	Top 20%	Top 50	Top 5%
Gold	Top 10%	Top 10	Top 10 + 0.2%*	Top 10 + 0.2%*

- Performance Tiers

Master

☑ 1 gold medal

☑ 2 silver medals

Grandmaster

☐ 5 gold medals

☑ Solo gold medal

Kaggle (Image recognition competitions)

Your Competitions

Active | Closed | Pinned | Hosted

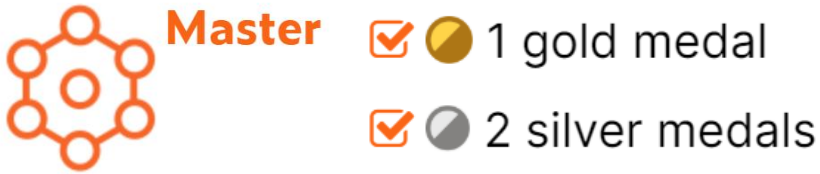
	**SIIM-ISIC Melanoma Classification** Identify melanoma in lesion images Featured • 2 months to go • 726 Teams	**$30,000**	Classification
	**TReNDS Neuroimaging** Multiscanner normative age and assessments prediction with brain function, structure, and connectivity Research • 21 days to go • 626 Teams	**$25,000**	Classification
	**ALASKA2 Image Steganalysis** Detect secret data hidden within digital images Research • a month to go • 495 Teams	**$25,000**	Classification
	**Prostate cANcer graDe Assessment (PANDA) Challenge** Prostate cancer diagnosis using the Gleason grading system Featured • a month to go • Code Competition • 549 Teams	**$25,000**	Classification
	**Global Wheat Detection** Can you help identify wheat heads using image analysis? Research • 2 months to go • Code Competition • 752 Teams	**$15,000**	Detection
	**Open Images Object Detection RVC 2020 edition** Detect objects in varied and complex images Playground • 2 months to go • 24 Teams	**Knowledge**	Detection
	**Open Images Instance Segmentation RVC 2020 edition** Outline segmentation masks of objects in images Playground • 2 months to go • 5 Teams	**Knowledge**	Segmentation

Kaggle (Image recognition competitions)

- Types of image recognition competition:
  - Image Classification
  - Image Retrieval
  - Object Detection
  - Segmentation
  - Image Generation

Today’s Goal

- Types of image recognition competition:
  - Image Classification
  - Image Retrieval
  - Object Detection
  - Segmentation
  - Image Generation



Road To Kaggle Master ①: The simplest method

- Team merge with Top Rankers/Kaggle Masters/Grandmasters

Cannot leave the team once merged!

Competition Start → (2-3 months) → Entry deadline → (0-1 week) → Team Merger deadline → (1 week) → Final submission deadline

Road To Kaggle Master ②: Just do it

- Prerequisites
  - GPU
  - GPU
  - GPU

Road To Kaggle Master ②: Just do it

- Prerequisites
  - GPU
  - GPU
  - GPU
- If you want to run deep learning on local machine
  -> NVIDIA GeForce RTX 2080/2080ti (One PC: ¥200k-300k)
- No local machine:
  -> Kaggle Notebook: NVIDIA P100 (30 hours/week)
  - Some competitions are notebook only.

My approach: Before join in the competition

- Check the task
  - The image type – RGB image or N-channel image
  - The method of evaluation – Accuracy or F-score or Quadratic weighted kappa
  - The difficulty of the task – The differences of scores in leaderboard

- Check the data size
  - The image size – 256x256, 512x512, 1024x1024
  - The dataset size – 10k ~ 10m

-> Is your time/GPU OK?

First Day

- Exploratory Data Analysis (EDA)
  - Check the images visually
  - Check the pairs of image and label
  - Check the distribution of labels
  - Check the distribution of train and test images
- Train a baseline model
  - Small model: resnet18, resnet34
  - Basic augmentation: Crop, Flip
  - Basic loss function: Cross Entropy

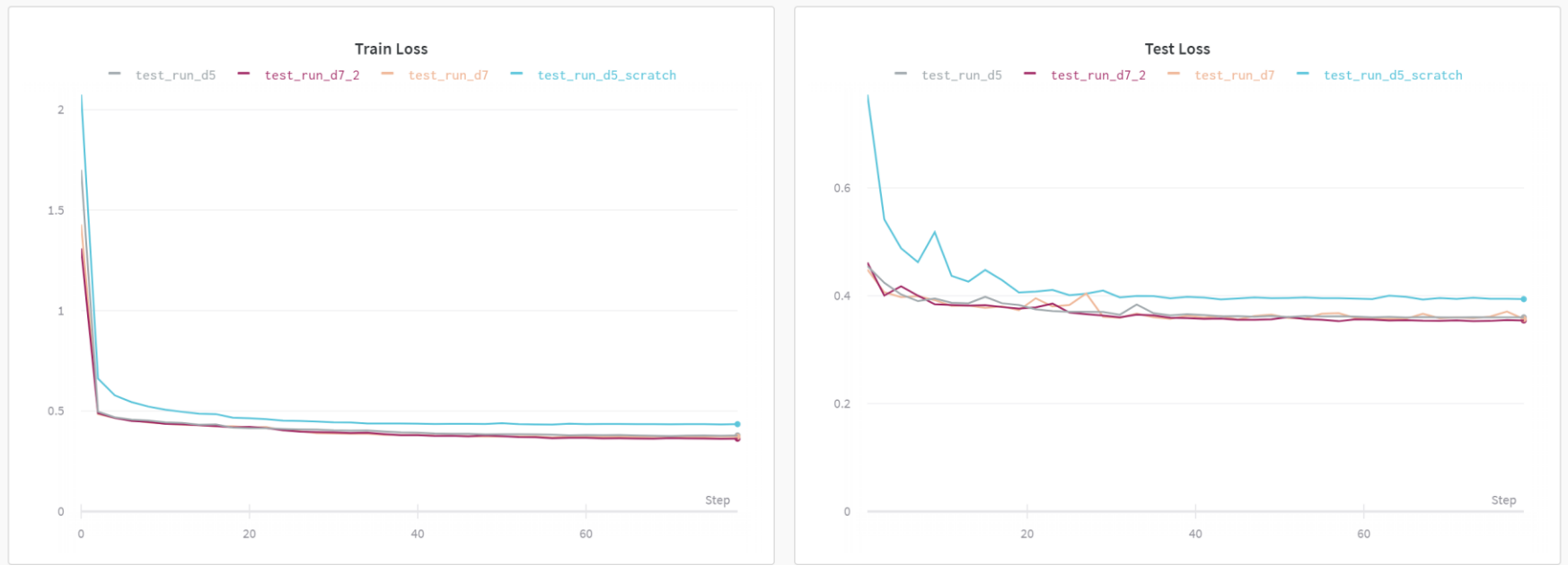
How to train the baseline

- Train and test a model with 100 samples

-> Overfitting ✔

How to train the baseline

- Monitoring the training process (using Tools like Weights & Biases)
  - Loss curve
  - Metric curve
  - Score (CV, Public LB)
  - Compare to the training history with different hyper-parameters



Code Management

- GitHub
  - Push before training
  - Commit history
  - Use Issues as Todo and Note

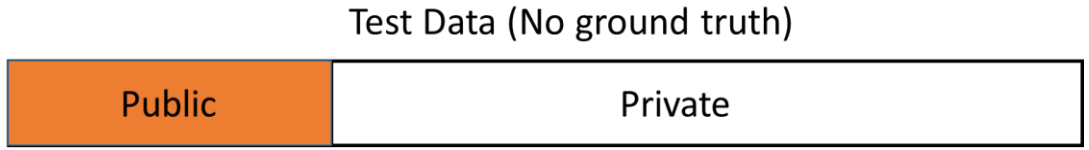
- Weights & Biases/Comet
  - Experiment tracking
  - Visualization of training history
  - Reproducibility (Hyper-parameter, git commit, training commend)

Second day ~ One week

- Find a proper way to evaluate the model

Second day ~ One week

- Find a proper way to evaluate the model
  - Leaderboard (LB) (Public Test Data):



  - Need to estimate the score on private dataset
  - ➢ ex) Cross Validation (CV)
    - Which metric? ex) Loss or Accuracy or Evaluation Metric
    - How to split validation data from training data?
    - How many data as validation?
  - Goal: CV ≈ LB

Second day ~ One week

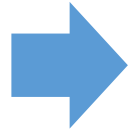
- Find a proper way to evaluate the model
- Survey
  - Solution of similar competitions
  - Papers related to the same task
  - SOTA methods
- Refine the baseline model
  - Try different model
  - Try different loss function
  - Add augmentation

Second day ~ One week

- Find a proper way to evaluate the model

- Survey
  - Solution of similar competitions
  - Papers related to the same task
  - SOTA methods

- Refine the baseline model
  - Try different model
  - Try different loss function
  - Add augmentation

Better baseline  Better final results

Refine the baseline model

- Model selection
  - pytorch-image-models (https://github.com/rwightman/pytorch-image-models)
  - pretrained-models.pytorch (https://github.com/Cadene/pretrained-models.pytorch)
  - Start from Inception, SeResNet, DenseNet
  - Also try EfficientNet [Tan+, ICML 2019], ResNeSt [Zhang+, Arxiv 2020]
  - (Modify the model if necessary – Pooling, Dropout, Dense block)
  - From shallow to deep

Refine the baseline model

- Model selection
- Optimizer
  - Adam / SGD with Nesterov momentum
- Loss Function
  - Cross Entropy Loss / Focal Loss [Lin+, ICCV 17] / ArcFace Loss [Deng+, CVPR 19]

Refine the baseline model

- Augmentation
  - Basic (Tool: albumentations)
    - Resize (Which interpolation? - bilinear or bicubic or others)
    - Padding (Which border mode? – constant or reflect or others)
    - Crop
    - Flip?
    - Rotation?
    - Normalize
    - Others
  - Advanced
    - Fast AutoAugment [Lim+, NeurIPS 2019]
    - RandAugment [Cubuk+, Arxiv 2019]
    - MixUp [Zhang+, ICLR 2018]
    - CutMix [Yun+, ICCV 2019]

	ResNet-50	Mixup [48]	Cutout [3]	CutMix
Image				
Label	Dog 1.0	Dog 0.5 Cat 0.5	Dog 1.0	Dog 0.6 Cat 0.4
ImageNet Cls (%)	76.3 (+0.0)	77.4 (+1.1)	77.1 (+0.8)	**78.6** **(+2.3)**
ImageNet Loc (%)	46.3 (+0.0)	45.8 (-0.5)	46.7 (+0.4)	**47.3** **(+1.0)**
Pascal VOC Det (mAP)	75.6 (+0.0)	73.9 (-1.7)	75.1 (-0.5)	**76.7** **(+1.1)**

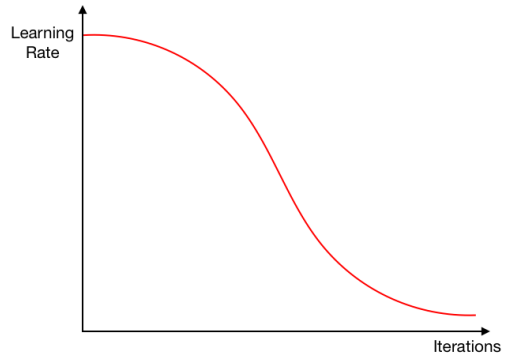
Refine the baseline model

- Hyper-parameter
  - Batch size
    - Larger is better -> Fill your GPU memory
  - Weight Decay
    - Standard weight decay :1e-4
    - Smaller dataset -> Larger weight decay: 1e-3
  - Momentum
    - Standard momentum: 0.9

Refine the baseline model

- Hyper-parameter
  - Learning rate
    - Try as many values as you can
    - Linear scale rule of SGD: batch size 128, lr 0.015 -> batch size 256, lr 0.03 [Goyal+, Arxiv 2017]

Refine the baseline model

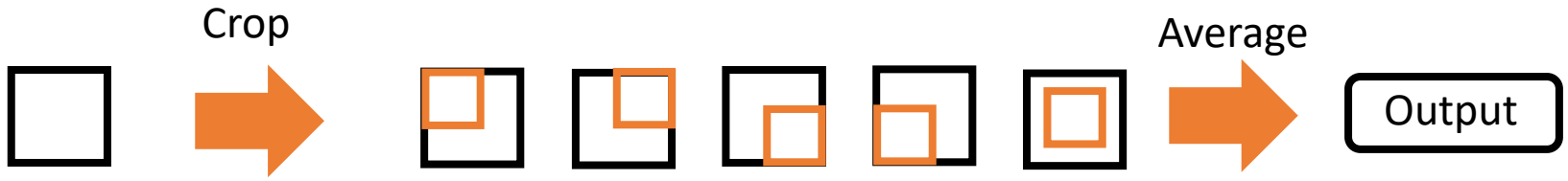
- Hyper-parameter
  - Learning rate
  - Learning rate scheduling
    - ReduceLROnPlateau
      - Reduce learning rate when a metric has stopped improving.
    - CosineAnnealingLR
    - 

Refine the baseline model

- Hyper-parameter
  - Learning rate
  - Learning rate scheduling
  - Tool
    - Automate hyperparameter search: Optuna **(**https://optuna.org/**)**

Refine the baseline model

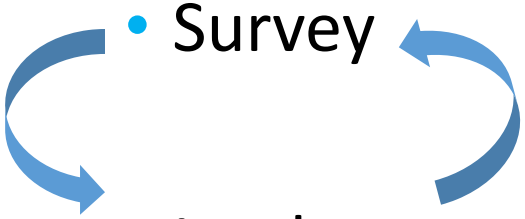
- Test Time Augmentation (TTA)



- Other tricks
  - Bag of Tricks for Image Classification [He+, CVPR 2019]

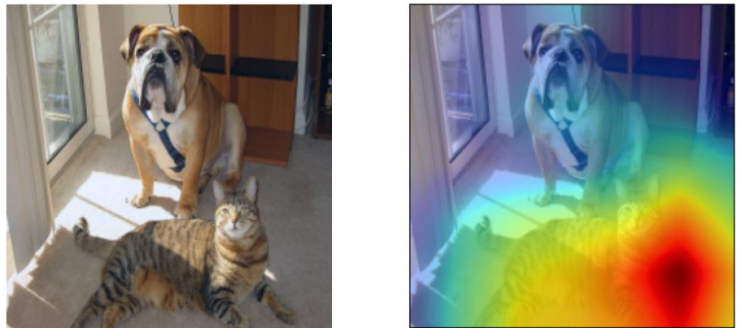
One week ~ Last week

- Survey



- Implement the ideas:
  - 90% fancy methods described in papers will not work 😢
  - Proper method should be chosen according to the data

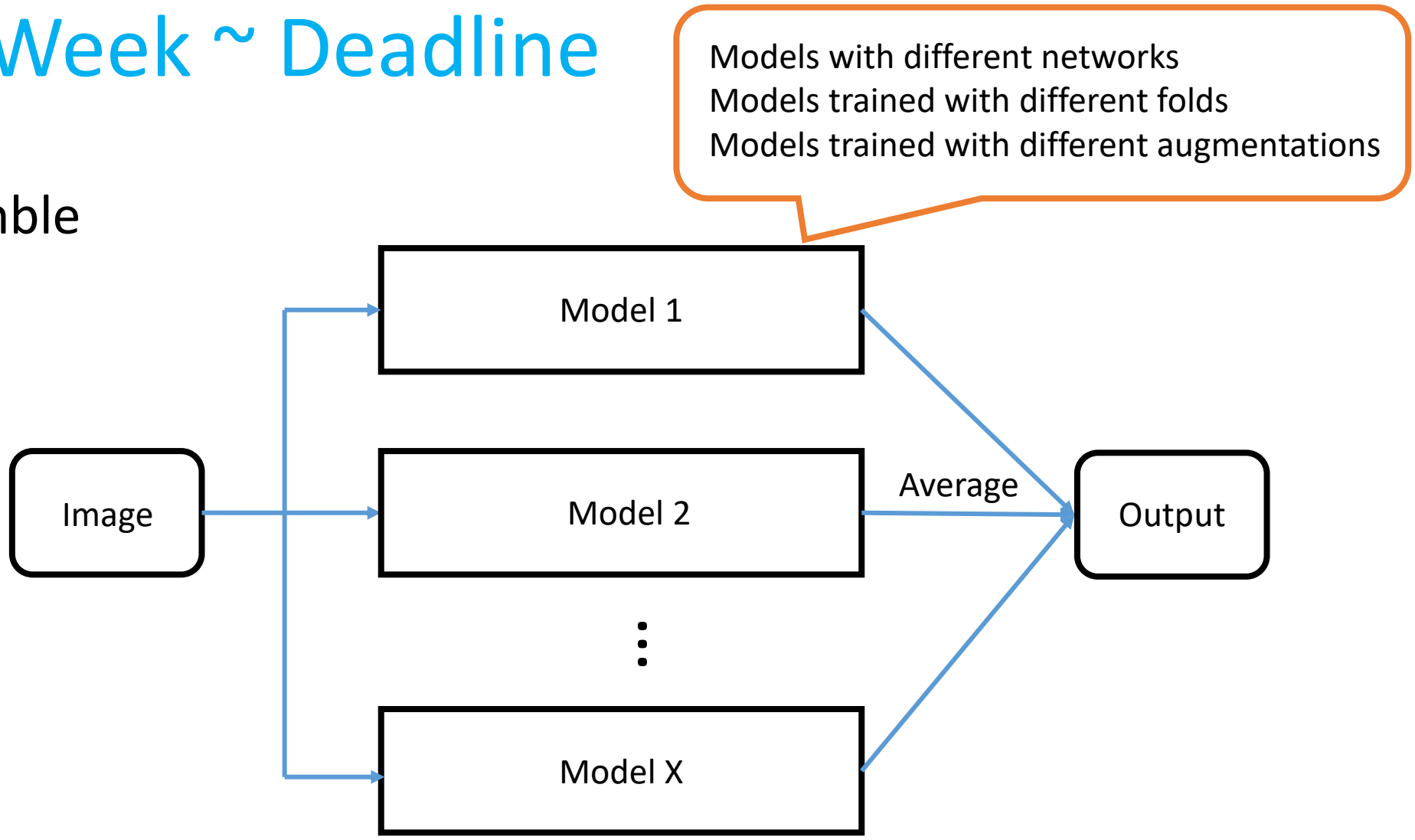
- Error analysis:
  - Check the error made by the model
  - Visualizing the activation can be helpful



Grad-CAM

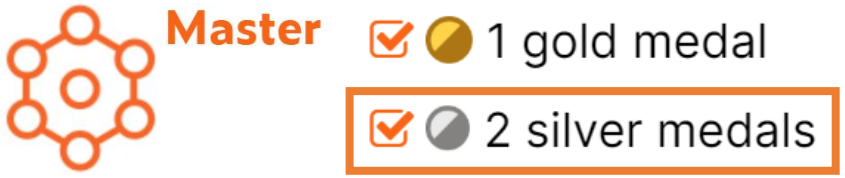
Last Week ~ Deadline

- Ensemble



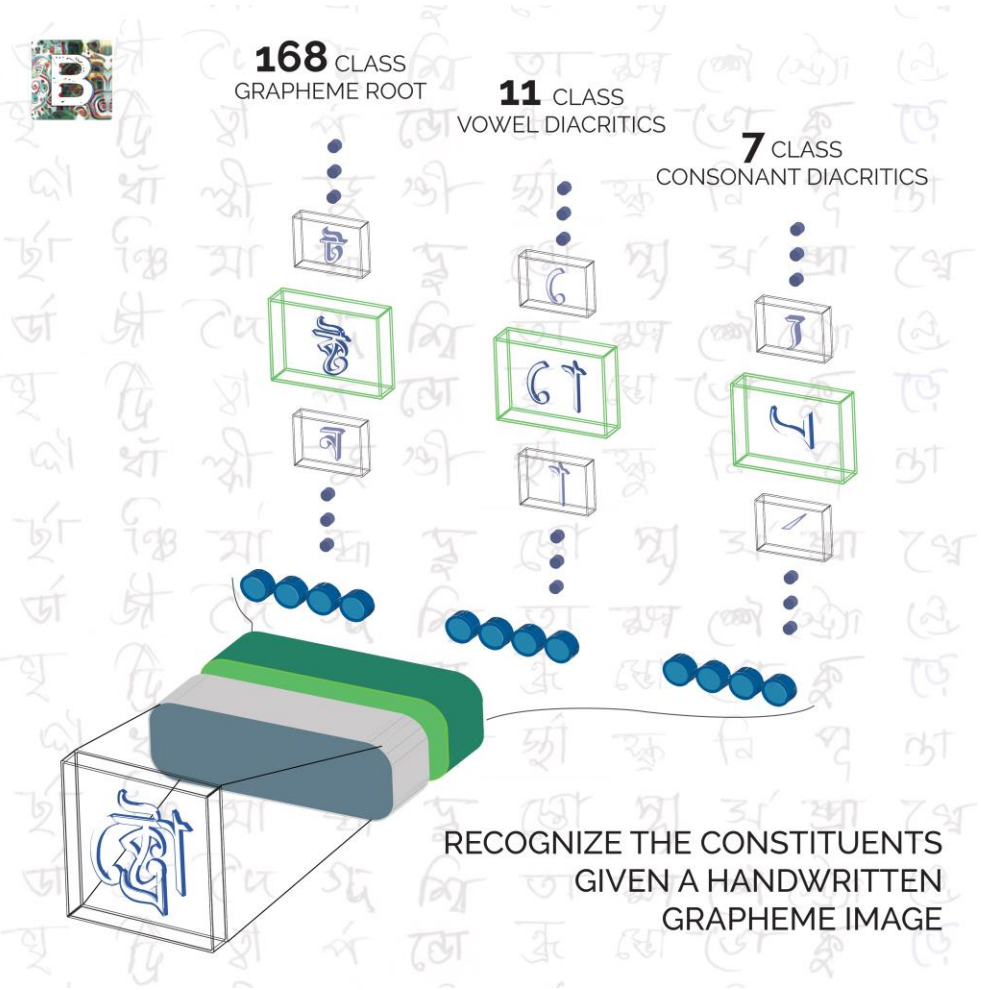
Today’s Goal

- Types of image recognition competition:
  - Image Classification
  - Image Retrieval
  - Object Detection
  - Segmentation
  - Image Generation



Case Study ①: Bengali.AI Handwritten Grapheme Classification

- Task: Image Classification X 3
- Data: 200,840 handwritten images with size 236 x 137 x 1
- Evaluation metric:
  weighted macro-averaged recall
  (2x on grapheme root)



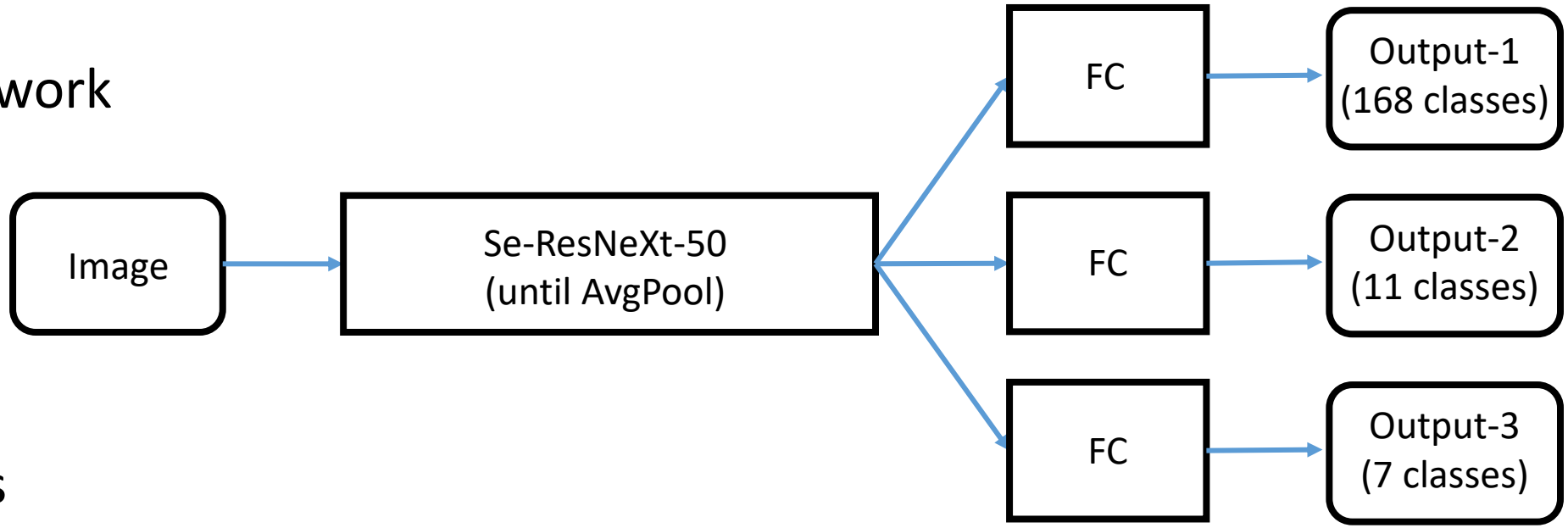
My Approach

- Preprocess
  - Multilabel Stratified Kfold -> 5 fold

- Augmentation
  - Resize(256, 256, interpolation=cv2.INTER_NEAREST)
  - RandomCrop(224, 224)
  - Rotate(10)
  - Normalize(mean=[0.5, 0.5, 0.5], std=[0.5, 0.5, 0.5])
  - CutMix [Yun+, ICCV 2019]

My Approach

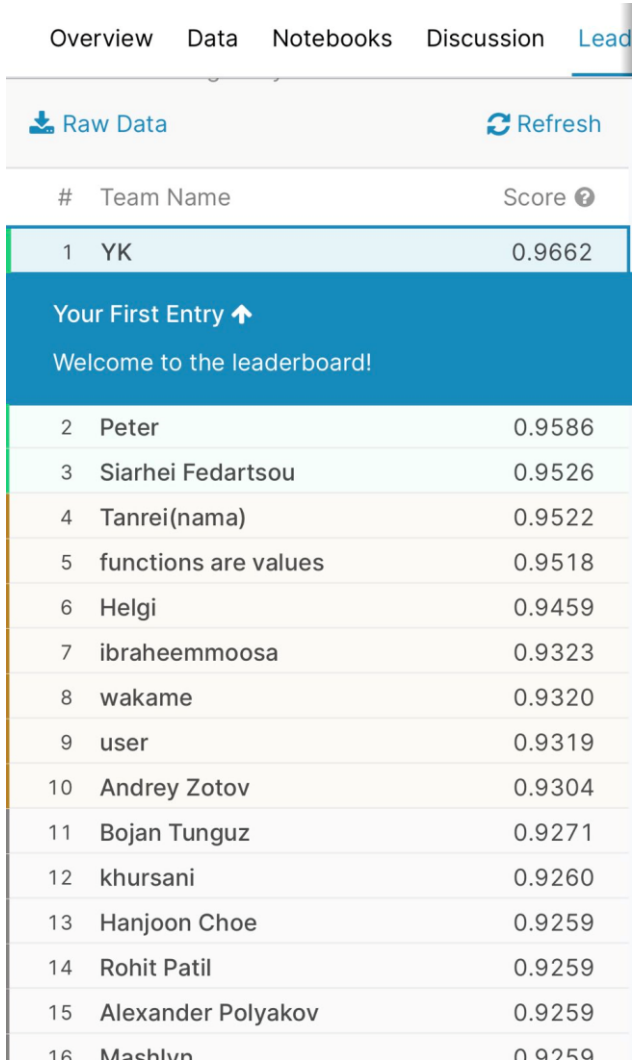
- Network



- Loss
  - Cross Entropy Loss x 3 (weight x 2 on grapheme root)
- Optimizer・Learning Rate
  - SGD・CosineAnnealingLR from 0.01

My Approach

- Ensemble
  - 5-fold ensemble
- Other hyper-parameters
  - Not carefully tuned … (due to ECCV deadline…)



Three months before deadline

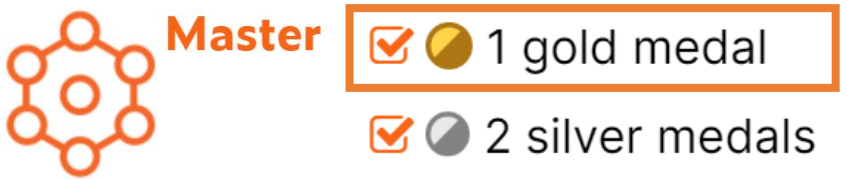
Result

- 81th/2059

77	▲ 1143	Georgi Pamukov	0.9360	14	4mo
78	▲ 853	Jo Tom	0.9360	20	3mo
79	▲ 56	AK-47	0.9359	54	3mo
80	▲ 1170	Hanjoon Choe	0.9359	9	5mo
81	▲ 122	YK	0.9359	22	5mo
82	▲ 475	Benjamin Warner	0.9358	16	3mo
83	▼ 50	YS	0.9357	112	3mo
84	▲ 1346	Subhajit Banerjee P.	0.9357	3	3mo

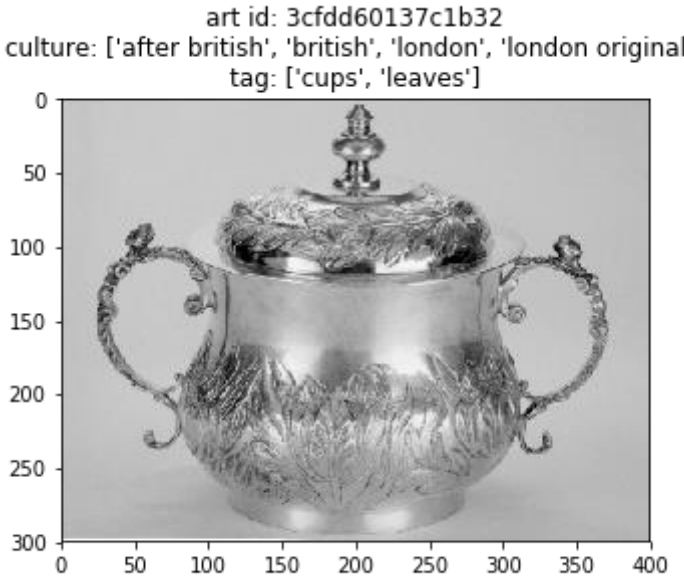
Today’s Goal

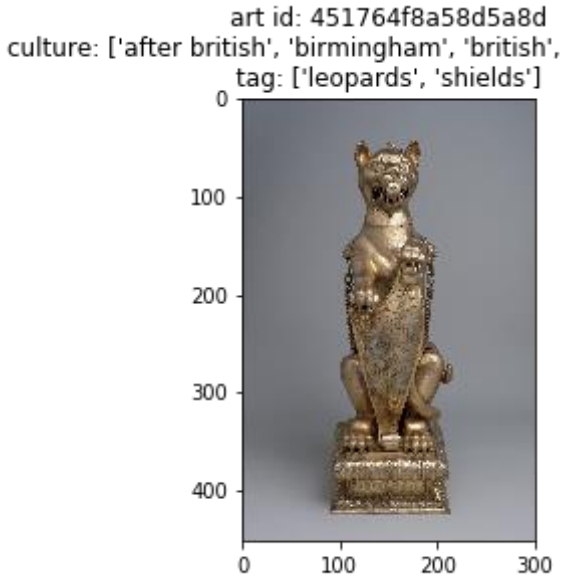
- Types of image recognition competition:
  - Image Classification
  - Image Retrieval
  - Object Detection
  - Segmentation
  - Image Generation

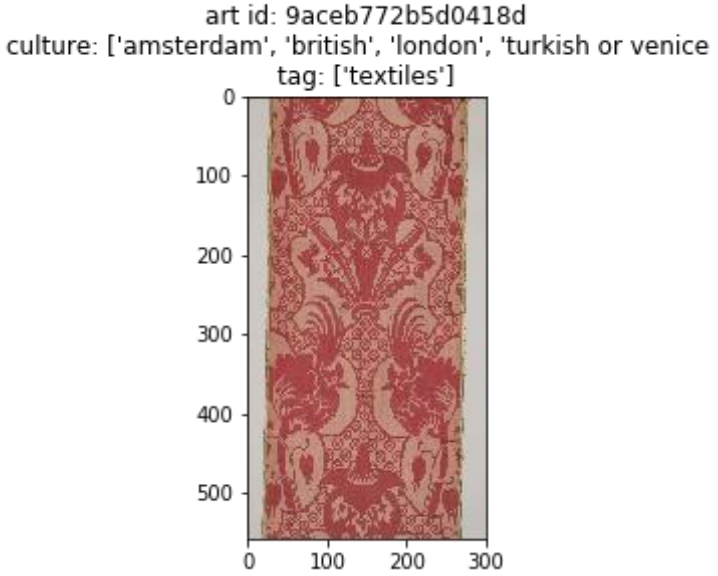


Case Study ②: iMet Collection 2019 - FGVC6 (CVPR workshop)

- Task:
  Multilabel Classification 1,108 classes (398 cultures, 705 tags)







Case Study ②: iMet Collection 2019 - FGVC6 (CVPR workshop)

- Task:
  Multilabel Classification 1,108 classes (398 cultures, 705 tags)

- Data:
  109,237 images with various size

- Evaluation metric:
  mean F2 score of each samples

$$\frac{(1+\beta^2)pr}{\beta^2 p + r} \text{ where } p = \frac{tp}{tp+fp}, \quad r = \frac{tp}{tp+fn}, \quad \beta = 2.$$

My Approach

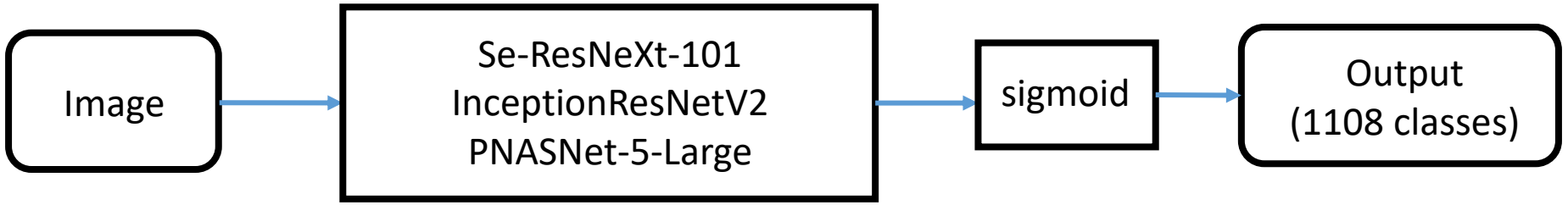
- Preprocess
  - Multilabel Stratified Kfold -> 40 fold

RandomErasing
[Zhong+, AAAI 2020]

- Augmentation
  - RandomCrop(320, pad_if_needed=True)
  - RandomHorizontalFlip()
  - Normalize(mean=[0.5949, 0.5611, 0.5185], std=[0.2900, 0.2844, 0.2811])
  - RandomErasing [Zhong+, AAAI 2020]
  - MixUp [Zhang+, ICLR 2018]

My Approach

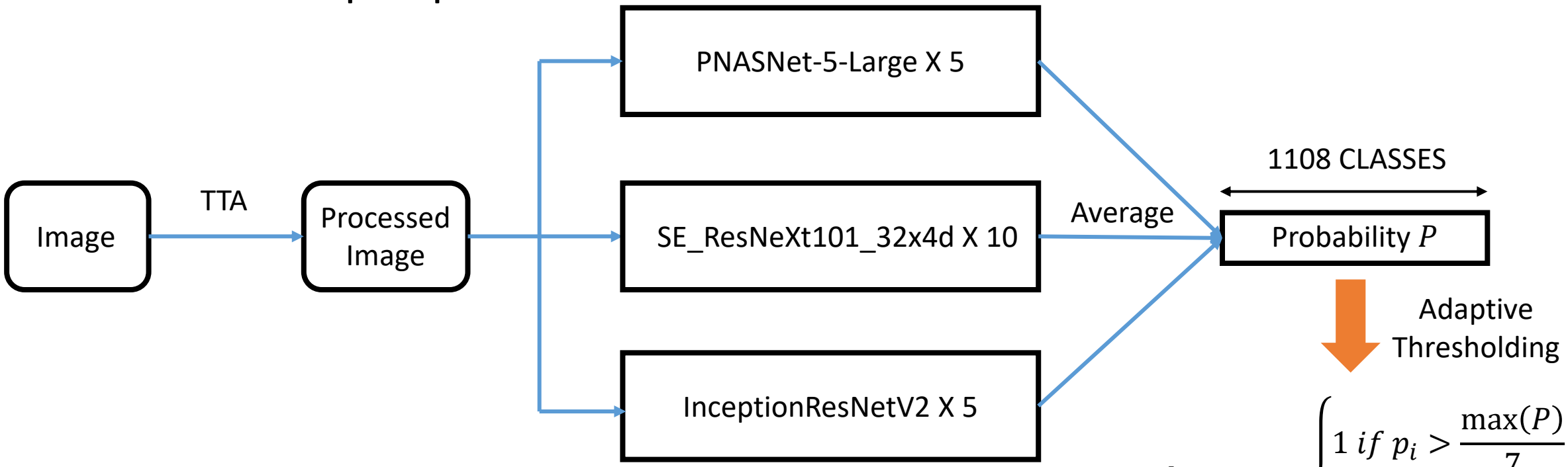
- Network



- Loss
  - Binary Cross Entropy Loss
- Learning Rate
  - ReduceLROnPlateau of validation loss from 0.0001 (Adam)

My Approach

- Ensemble and postprocess



Result

- 6th/441

#	△pub	Team Name	Kernel	Team Members	Score	Entries	Last
1	▲ 400	[ods.ai] Konstantin Gavrilchik			0.672	2	2d
2	▲ 425	ヽ(°ʖ°)つ—☆*:・°			0.667	1	2d
3	▲ 426	[ods.ai] Ilya Kibardin			0.664	2	2d
4	▲ 411	pudae			0.663	2	2d
5	▲ 349	[ods.ai] n01z3			0.662	2	2d
6	▲ 332	みんなをStarlightしちゃいます			0.660	2	2d
7	▲ 419	X5, we need an explanation f...			0.659	2	2d
8	▲ 425	X5, Best Russian Company			0.658	2	2d
9	▲ 321	Appian			0.658	2	2d
10	▲ 422	Alchemists' Creed: Obey the ...			0.655	2	2d
11	▲ 292	頼む！！！！！			0.654	2	2d
12	▲ 418	[ods.ai]X5.earhian 卢本伟没有...			0.653	2	2d

Conclusion

- The approach to an image competition on Kaggle
  - To win 🥈
    - Carefully tuned baseline
  - To win 🥇
    - Carefully tuned SOTA methods + Some small tricks
  - To win 🏆
    - Carefully tuned SOTA methods + Some original ideas
    - Please refer to winner solutions on Kaggle