Challenge Overview

webvision

Challenge Tasks

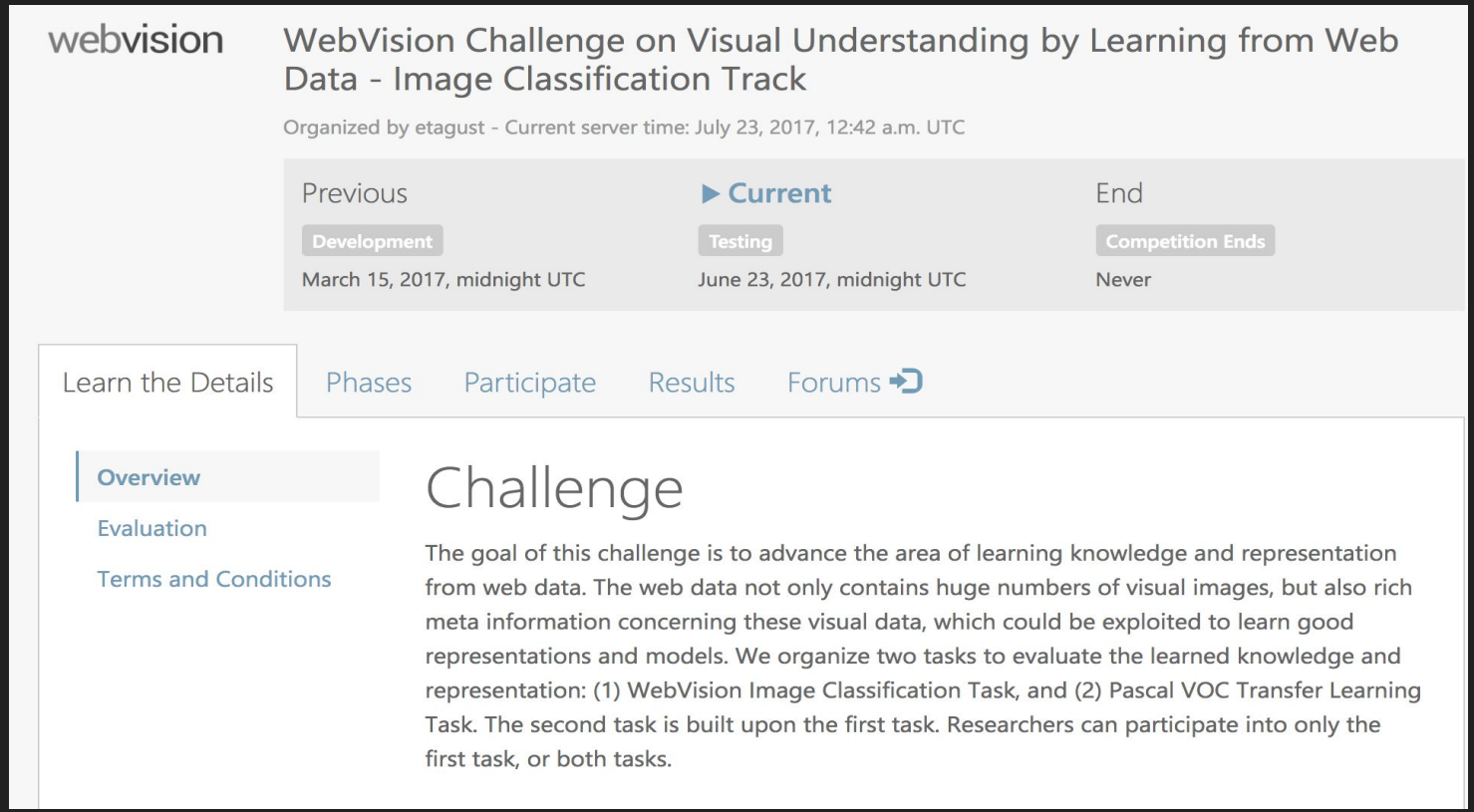
WebVision Image Classification Task

- Learn models on the WebVision train set and evaluate on the val and test set

PASCAL VOC Transfer Learning Task

- Verify the knowledge learned on WebVision dataset on Pascal VOC dataset

Challenge Platform



Challenge Schedule

Development

Start: March 15, 2017, midnight

Description: The Development Leaderboard is based on a fixed random subset of 50% of the test images. To submit, upload a .zip file containing a predictions.txt file with the prediction in the format used in the dev kit. (

Testing

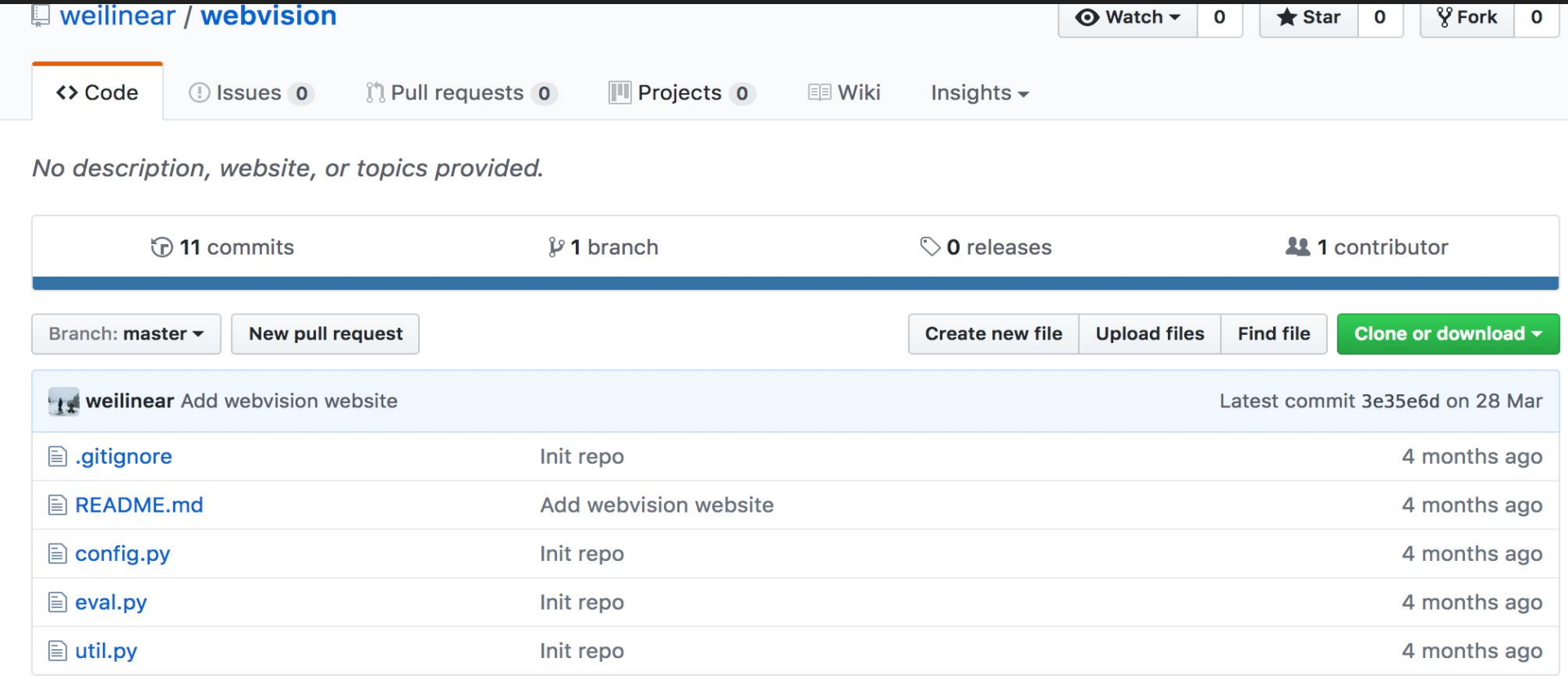
Start: June 23, 2017, midnight

Description: o submit, upload a .zip file containing a predictions1.txt, ..., predictions5.txt file with the prediction in the format used in the dev kit. The file with the best top-5 accuracy will be used to determine the winner. Please also include a readme.txt file with a description for your entry. An example submission file can be found at: https://data.vision.ee.ethz.ch/cvl/webvision/example_submission_classification.zip

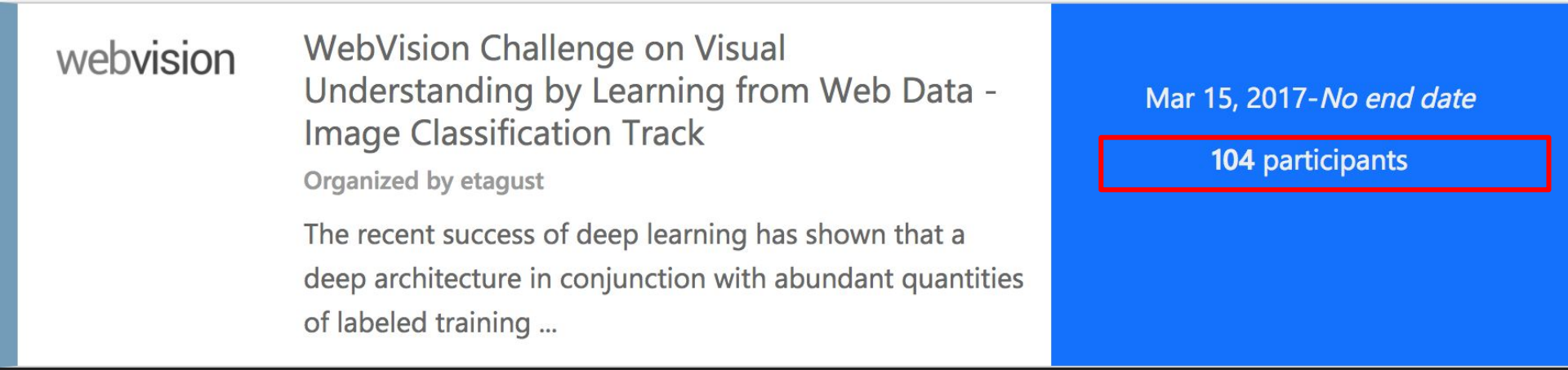
Submission Policies

- Each participant may have maximum 10 submissions during development phase.
- Each team may have 5 submissions during test phase.
- Learn vision models from noisy data (WebVision dataset).
- No extra data is allowed to use.

Provided Tools



Number of participants



We have 8 teams to submit results to image classification track.

We have 1 team to submit results to Transfer learning track.

Challenge Results: Classification Track

WebVision Image Classification Task

Rank	Team name	Run1	Run2	Run3	Run4	Run5
1	Malong AI Research	0.9358	0.9467	**0.9478**	**0.9478**	0.9470
2	SHTU_SIST	0.9223	**0.9225**	0.9218	0.9219	0.9216
3	HG-AI	**0.9189**	0.9152	0.9152	0.9189	0.9189
4	VISTA	0.8979	**0.9005**	0.8980	0.8992	0.8980
5	LZ_NES	**0.8853**	0.8758	0.8723	0.8504	0.8504
6	CRCV	0.8707	0.8717	0.8701	0.8712	**0.8721**
7	Chahrazad	**0.8705**	0.8705	0.8705	0.8705	0.8705
8	Gombru (CVC and Eurecat)	0.8475	0.8374	**0.8586**	0.8586	0.8586

Challenge Result: Transfer Learning Track

Pascal VOC Transfer Learning Task

Rank	Team name	mAP
1	Malong AI Research	0.90

Qualitative results: easy classes

Qualitative results: hard classes

Team: CVC and Eurecat

Modalities: Image, Query ID, Title, Tags

1. Similarities between image associated text and mean LDA topic distribution are used to re-weight samples.
2. A CNN (GoogLeNet) with 2 heads has been trained:
   a. One head for the groundtruth label
   b. One head for the LAD topic distribution of image associated text.

Team: Chahrazad

Modalities: Image, Query ID

Architectures: DenseNet-BC architecture with 121 layers and 32 growth rate

The network was trained with batchsize of 128 for 100 epochs.

Team: CRCV

Modalities: Image, Query ID

Architecture: Inception ResNet V2 from scratch (check point at 980k iterations)

Residual loss: L = CrossEntroy(y, softmax(w+u)) - CrossEntroy(y, softmax(w)) + Regularization(u)

Learning without forgetting: alternating fine-tuning between Google and Flickr without forgetting the knowledge in other dataset

Data clean: remove sample that has lowest prediction scores.

Team: LZ_NES

Modalities: Image, Query ID, Text

Architecture: ResNet-101

Use NLP tools (word2vec and gensim) to generate training labels (v1 and v2)

Three train label lists (query list, v1, and v2) are employed to train CNNs with different weights.

Team: VISTA

Modalities: Image, Query ID, Text

Architecture: Inception V3

Try different methods to deal with label noise: label smooth, label distillation, different bootstrap, different denoising ways etc.

Simplest way works the best.

Team: HG-AI

Modilites: Image, Query ID

Architectures: GoogLeNet, Inception-ResNet v2, ResNet 50

Different batchsize for different architectures.

First train on balanced data and then fine tune on all data.

Team: SHTU-SIST

Modalities: Image, Query ID

Architecture: Inception ResNet v2

Five base models are first trained from scratch on all training data.

Original data is cleaned by ensemble models to re-train five base models.

Team: Malong AI Research

Modalities: Image, Query ID

Architectures: ?

Dividing data into clean part and noisy part. First train on clean part and then fine tune on all parts

Data balance

Adaptive LR drop