AI6126 Project 1
CelebAMask Face Parsing

Project 1 Specification (Version 1.0. Last update on 18 February 2022)

Important Dates

Issued: 18 February 2022

Release of test set: 18 March 2022 12:00 AM SGT

Due: 25 March 2022 11:59 PM SGT

Group Policy

This is an individual project

Late Submission Policy

Late submissions will be penalized (each day at 5% up to 3 days)

Challenge Description

Face parsing assigns pixel-wise labels for each semantic components, e.g., eyes, nose, mouth. The goal of this mini challenge is to design and train a face parsing network. We will use the data from the CelebAMask-HQ Dataset [1] (See Figure 1). For this challenge, we prepared a mini-dataset, which consists of 5000 training and 1000 validation pairs of images, where both images and annotations have a resolution of 512 x 512.

The performance of the network will be evaluated based on the mIoU between the predicted masks and the ground truth of the test set (the ground truth of the test set will not be released).

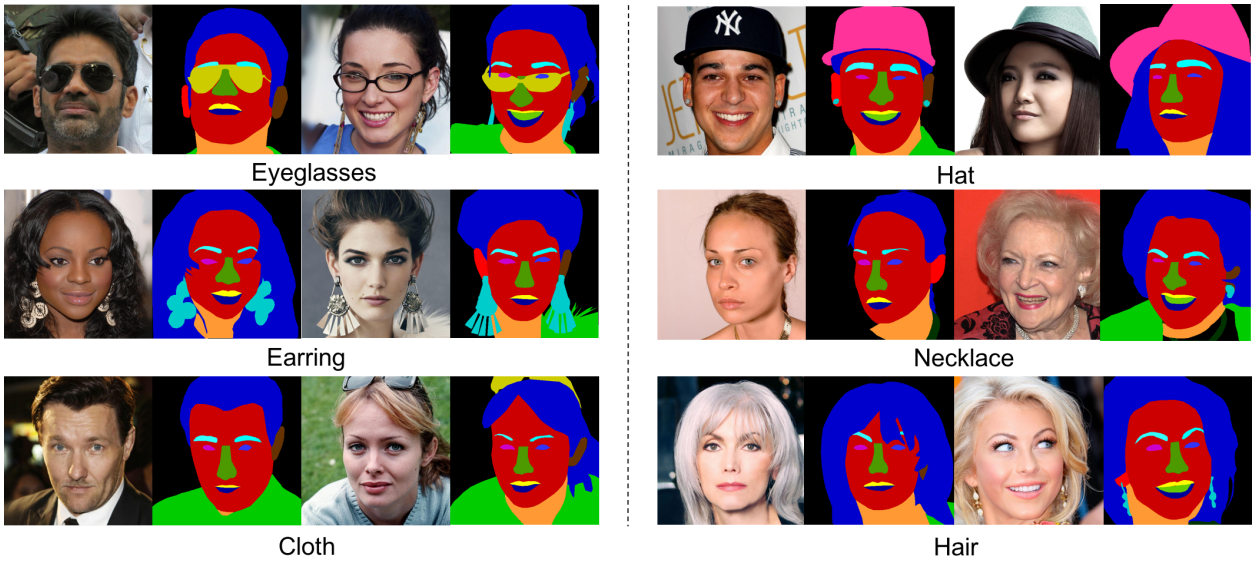


Figure 1. Sample images in CelebAMask-HQ

Assessment Criteria

We will evaluate and rank the performance of your network model on our given **1000** test images based on the mIoU.

The higher the rank of your solution, the higher the score you will receive. In general, scores will be awarded based on the Table below.

Percentile in ranking	≤ 5%	≤ 15%	≤ 30%	≤ 50%	≤ 75%	≤ 100%	*
Scores	20	18	16	14	12	10	0

Notes:

- We will award bonus marks (up to 2 marks) if the solution is interesting or novel.
- Marks will be deducted if the submitted files are not complete, e.g., important parts of your core codes are missing or you do not submit a short report.

Requirements

- Download dataset: https://entuedu-my.sharepoint.com/:f:/g/personal/yuming002_e_ntu_edu_sg/Ep5E8hYimVtFmW5DFi_DjjIBNc3GNJMsHqklyGhADbgESA?e=clEmok
- Train your network using our provided training set, you are allowed to use ImageNet pretrained models.
- Tune the hyper-parameters using our provided validation set.
- The test set will be available one week before the deadline (this is a common practice of major computer vision challenges).
- While ImageNet pretrained models are allowed, no external data are permitted in this mini challenge. If the use of any external data (e.g., using the full CelebAMaskHQ dataset rather than our provided dataset) is spotted, you will get a 0 mark for this project.
- You should not use an ensemble of models.

Submission Guidelines

Submitting Results on CodaLab

We will host the challenge on CodaLab. You need to submit your results to CodaLab. Please follow the following guidelines to ensure your results are successfully recorded.

- The CodaLab competition link: https://codalab.lisn.upsaclay.fr/competitions/1945?secret_key=a47e8cdb-3ad0-45cc-b182-13fa9ae81616
- Register a CodaLab account with your NTU email.
- After your registration, please fill in the username in the Google Form: https://docs.google.com/forms/d/e/1FAIpQLSd5ysD7ElcHlfDSUkcPWOMEUXyYwyHTftyqkDdzOtOxozACIw/viewform?usp=sf_link
- Submit the results (i.e., the predicted masks) from your model on the 1000 test images as a zip file. Put the results in a subfolder and use the same file name as the original test images. (e.g., if the input image is named as 100.jpg, your result should also be named as 100.png)
- You can submit your results multiple times but no more than 10 times per day. You should report your best score (based on the test set) in the final report.
- Please refer to Appendix A for the hands-on instructions for the submission procedures on CodaLab if needed.

Submitting Report on NTULearn

Submit the following files (all in a single **zip** file named with your matric number, e.g., **A12345678B.zip**) to NTULearn before the deadline:

- A short report in pdf format of not more than five A4 pages (single-column, single-line spacing, Arial 12 font, the page limit excludes the cover page and references) to describe your final solution. The report must include the following information:
    - the model you use
    - the loss functions
    - training curves (i.e., loss, mIOU curve for val set)
    - the number of parameters of your model. Use "sum(p.numel() for p in model.parameters() if p.requires_grad)"
    - Specs of your training machine, e.g., number of GPUs, GPU model.

  You may also include other information, e.g., any data processing or operations that you have used to obtain your results in the report.
- The best results (i.e., the predicted masks) from your model on the 1000 test images. We will use the results to check plagiarism.
- All necessary **codes** you used in this project. You don't need to submit the training or validation images.
- **Screenshot on Codalab of score achieved**
- The **model checkpoint (weights)** of your submitted model.
- A **Readme.txt** containing the following info:
    - Your matriculation number and your CodaLab username.
    - Description of the files you have submitted.
    - References to the third-party libraries you are using in your solution (leave blank if you are not using any of them).
    - Any details you want the person who tests your solution to know when he/she tests your solution, e.g., which **script to run**, so that we can check your results, if necessary.

Tips

- For this project, you can use the MMSegmentation codebase, which implements many popular segmentation methods with modular design and provides detailed documentation. Except for installation instructions, please also pay attention to customizing datasets, designing data pipelines, customizing modules, and customizing runtime. If MMSegmentation is too complicated for you, you can

also refer to this repo. **Be extra careful to **not** introduce pre-trained models other than ImageNet-pre-trained.

- For the calculation of mIoU, you can refer to 'evaluate.py' in the shared dataset folder. You should replace the corresponding path 'xxx' to your own path.
- The following techniques may help you to boost the performance:
    - Data augmentation
    - Regularization
    - Deeper model (but be careful on overfitting, since we only have 5000 training images in this project)
    - Hyper-parameters fine-tuning, e.g., choice of optimizer, learning rate, number of iterations
    - Think about what is unique to this dataset and propose novel modules.

Computational Resource

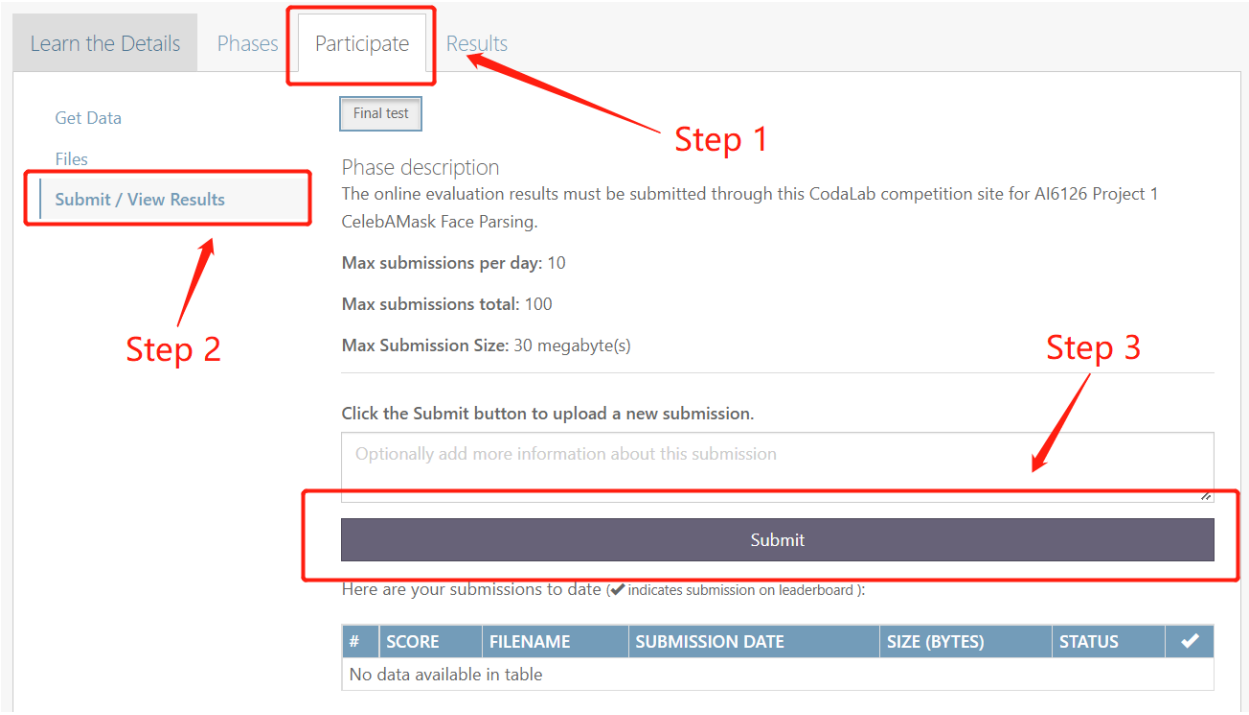
You can use the computational resources assigned by the MSAI course. Alternatively, you can use Amazon's EC2 or Google CoLab for computation. As a student, you can sign up to receive free $100 credit through the AWS Educate program. We encourage students to use *g2.2xlarge* instances running Ubuntu for maximal ease of installing. Note that $100 of Amazon credit allows you to run a *g2.2xlarge* GPU instance for approximately 6 days without interruption (you should keep it on only while using it).

References

[1] Cheng-Han Lee, Ziwei Liu, Lingyun Wu, Ping Luo, MaskGAN: Towards Diverse and Interactive Facial Image Manipulation, CVPR 2020

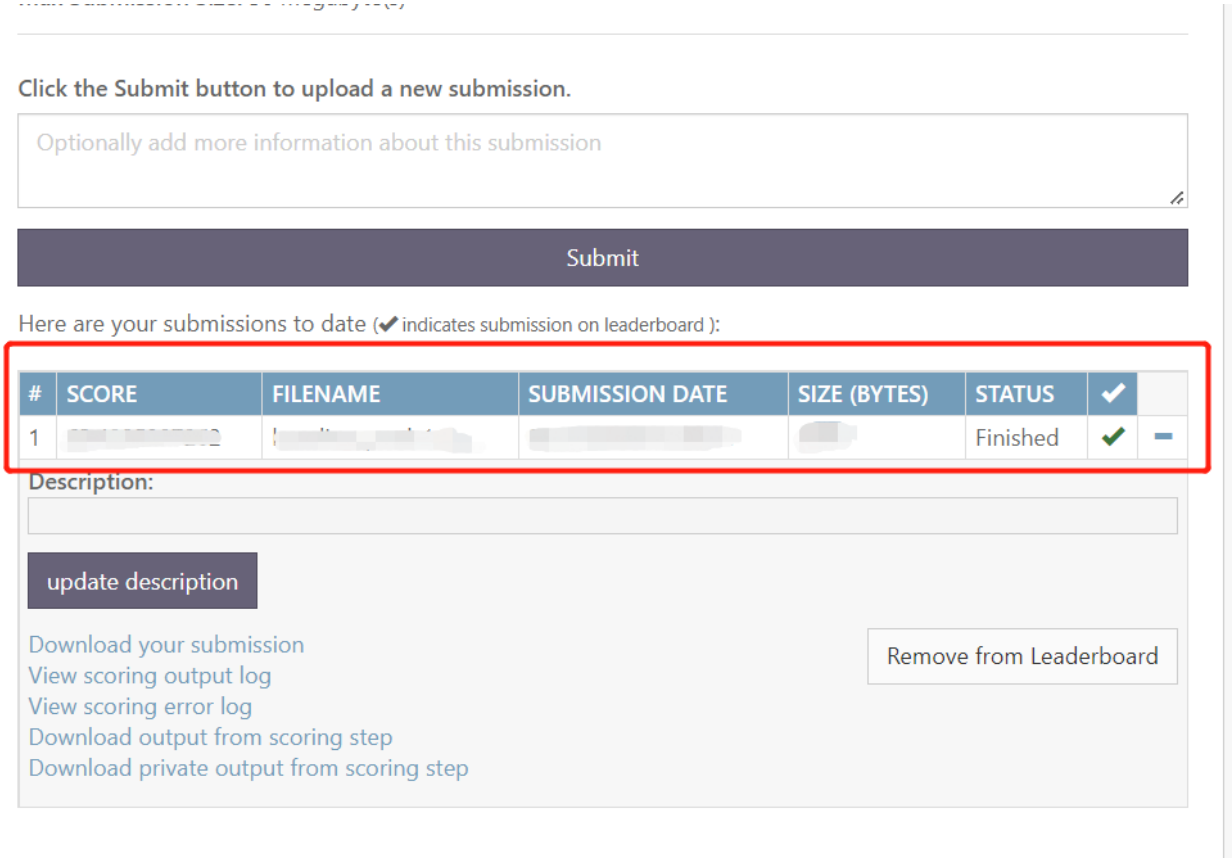
Appendix A Hands-on Instructions for Submission on CodaLab

After your participation to the competition is approved, you can submit your results here:



Then upload the zip file containing your results.

If the ‘STATUS’ turns to ‘Finished’, it means that you have successfully uploaded your result. Please note that this may take a few minutes.



Appendix B Definition of Mask Labels

0: background	4:left eye	8: left ear	12: lower lip	16: necklace
1: skin	5: right eye	9: right ear	13: hair	17: neck
2: nose	6: left brow	10: mouth	14: hat	18: cloth
3: eye glasses	7: right brow	11: upper lip	15: earring	