Institute of Electronics
National Yang Ming Chiao Tung University
Hsinchu, Taiwan

Midterm Project
Hand-pose estimation (Kaggle)

Architecture Development & Algorithm Refinement for Intelligent Computing

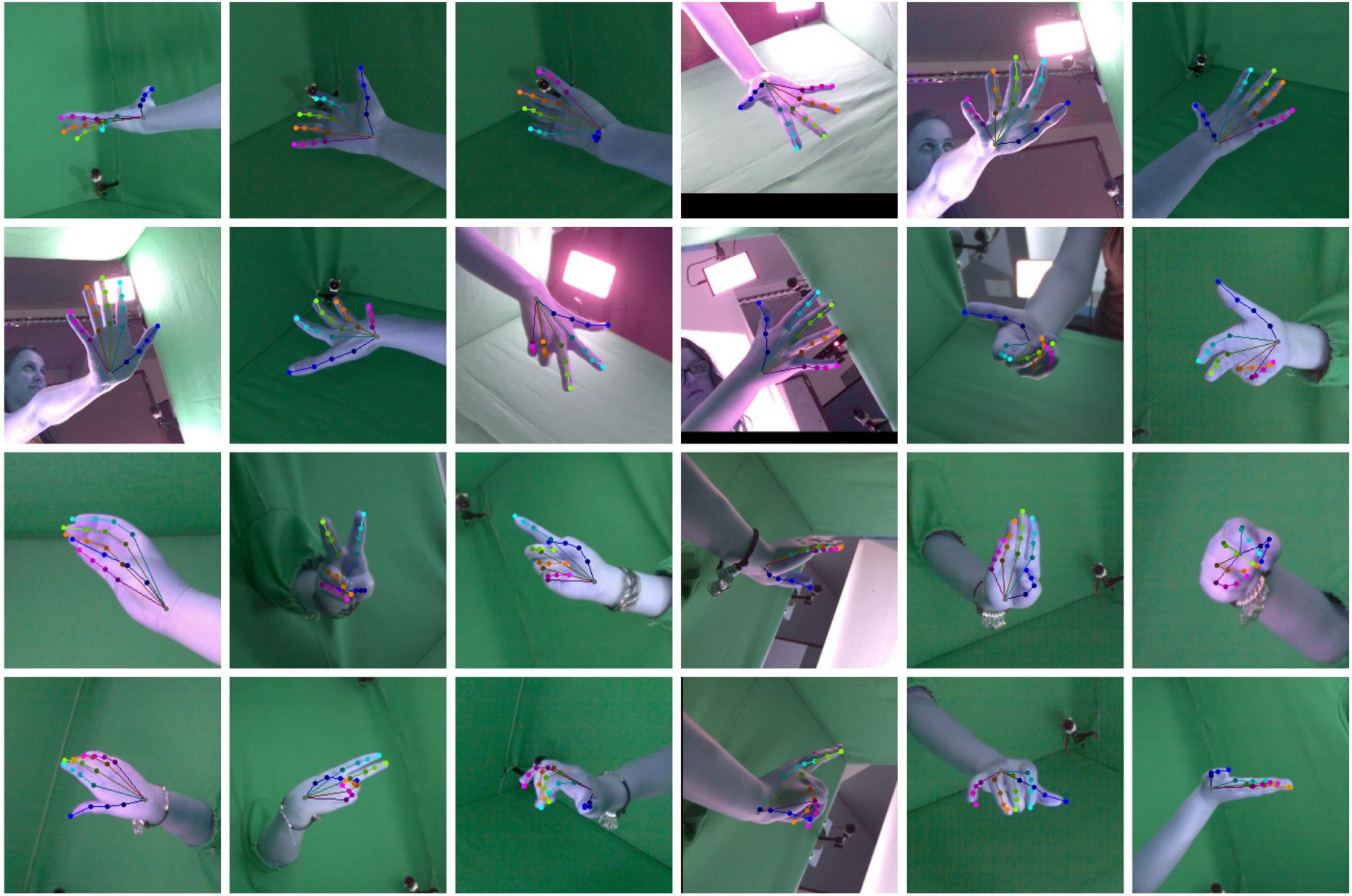
Presenter: Chih-Yao, Liang
Advisor: Juinn-Dar Huang, Ph.D.

July 25, 2024

Outline

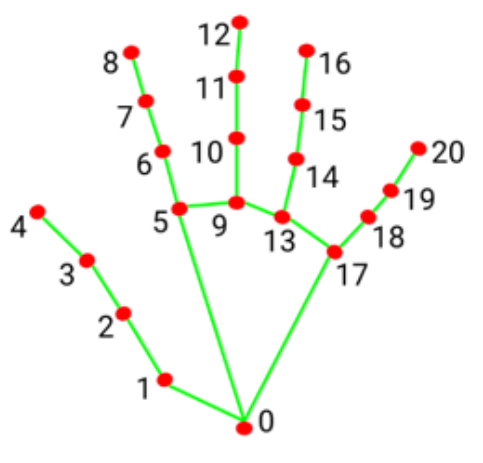
- Midterm project: Hand-pose estimation
  - Task Description
  - Grading and Requirement
  - Submission
  - Kaggle
  - Report
  - Reference
  - Kaggle usage

Task Description : Hand-pose estimation



Task Description: Hand-pose estimation

- The images are collected from FreiHAND dataset



- Dataset provided by TA:
  - Train: 32560 images
  - Test: 1000 images

- DO NOT use the test data, we will regard this as CHEATING.

Grading

- Deadline :
  - Kaggle: 8/8 (四) 23:59
- Grading :
  - Kaggle (40%) :
    - 10 ≤ Avg PixelDiff <15 (28%)
    - 7 ≤ Avg PixelDiff <10 (32%)
    - 5 ≤ Avg PixelDiff < 7 (35%)
    - Avg PixelDiff < 5 (40%)
  - Ranking (20% -> 依比例無條件進位)
  - Report (40%)
  - Bonus (10%)

Ranking formula

- Model Size [M-param] × Cost [G-FLOPs] × $e^{\text{Avg PixelDiff}}$
- Ex: 10 M-param * 10 G-FLOPs * e^{4} = 5459.815

- The smaller the better.

Requirement

- Rules:
  - You can discuss with others, but you should write your homework by yourself!
  - Your codes should be able to generate the results you submit to the leaderboard. (within 1% error)
  - Please build up the model by yourself. (reference some well known model structures are allowed)
  - DO NOT use the testing data in any way!

Submission

- Submit your code via email andy.312510138@gmail.com
  - <your English name>_midterm_project.zip
  - Include:

    > **Your codes (.ipynb)**

    > Model weight file (.pth)

    > **Report.pdf**

- **DO NOT**
  - Submit the dataset
  - Submit your Project in wrong format

    > like **.rar/.7z** or **report.doc**

Kaggle

- Register a Kaggle account by yourself
- Join our in-class competition
  - Link: https://www.kaggle.com/t/64315fed6af7429da41f074343b6a9b7
  - Maximum daily submission: **10**

- Rules:
  - Your team name should be: < your English name >

    > **E.g. Chih-Yao Liang**

    > Otherwise, your submission will NOT be graded.
  - You can NOT create multiple Kaggle accounts to submit more results.
  - You can NOT upload results identical to other people's.

Report (40%)

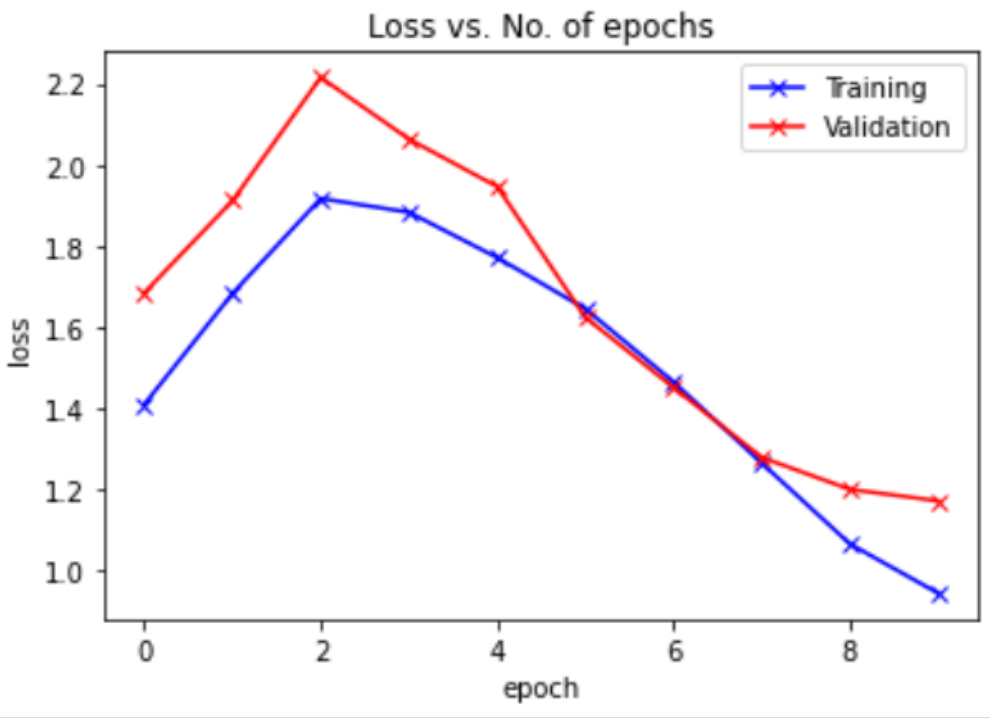
- Model Architecture and Training (20%)
  - Draw your model architecture using simple symbols. (5%)
  - Print learning curve and briefly state your observations. (loss / epoch) (5%)
    - Hint: you can try some different training methods.
  - What are the differences between training loss and validation loss? (10%)
    - Hint: plot the train loss and valid loss per epoch together.
    - Hint: explain the relationship between them.
- How to improve the accuracy (list your methods) (20%):
  - Loss function ?
  - Your network?
  - Activity function?
  - Training skill?
  - Etc…

Report example

Layer (type)	Output Shape	Params	FLOPs(M+A) #
Conv2d-1	[1, 64, 32, 32]	1792	3538944
BatchNorm2d-2	[1, 64, 32, 32]	256	131072
ReLU-3	[1, 64, 32, 32]	0	0
Conv2d-4	[1, 128, 32, 32]	73856	150994944
BatchNorm2d-5	[1, 128, 32, 32]	512	262144
ReLU-6	[1, 128, 32, 32]	0	0
MaxPool2d-7	[1, 128, 16, 16]	0	0
Conv2d-8	[1, 128, 16, 16]	147584	75497472
BatchNorm2d-9	[1, 128, 16, 16]	512	65536
ReLU-10	[1, 128, 16, 16]	0	0
Conv2d-11	[1, 128, 16, 16]	147584	75497472
BatchNorm2d-12	[1, 128, 16, 16]	512	65536
ReLU-13	[1, 128, 16, 16]	0	0
Conv2d-14	[1, 256, 16, 16]	295168	150994944
BatchNorm2d-15	[1, 256, 16, 16]	1024	131072
ReLU-16	[1, 256, 16, 16]	0	0
MaxPool2d-17	[1, 256, 8, 8]	0	0
Conv2d-18	[1, 512, 8, 8]	1180160	150994944
BatchNorm2d-19	[1, 512, 8, 8]	2048	65536
ReLU-20	[1, 512, 8, 8]	0	0
MaxPool2d-21	[1, 512, 4, 4]	0	0
Conv2d-22	[1, 512, 4, 4]	2359808	75497472
BatchNorm2d-23	[1, 512, 4, 4]	2048	16384
ReLU-24	[1, 512, 4, 4]	0	0
Conv2d-25	[1, 512, 4, 4]	2359808	75497472
BatchNorm2d-26	[1, 512, 4, 4]	2048	16384
ReLU-27	[1, 512, 4, 4]	0	0
MaxPool2d-28	[1, 512, 1, 1]	0	0
Flatten-29	[1, 512]	0	0
Dropout-30	[1, 512]	0	0
Linear-31	[1, 100]	51300	102400
ResNet18-32	[1, 100]	0	0

Total parameters: 6,626,020 6.6M
Trainable parameters: 6,621,540
Non-trainable parameters: 4,480
Total flops(M) : 379,684,864 379.7M
Total flops(M+A): 759,369,728 759.4M

Parameters size (MB): 25.28



Not necessary like this !

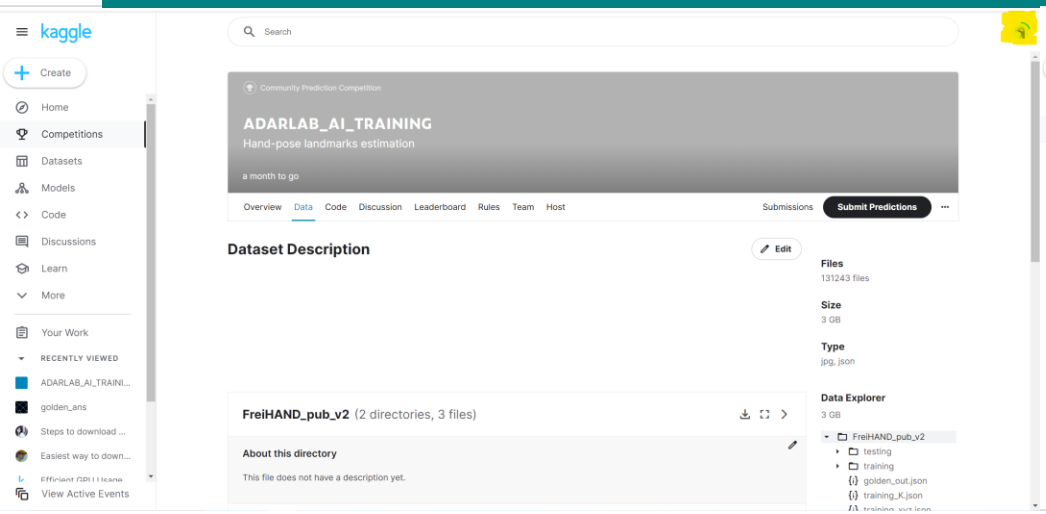
Reference

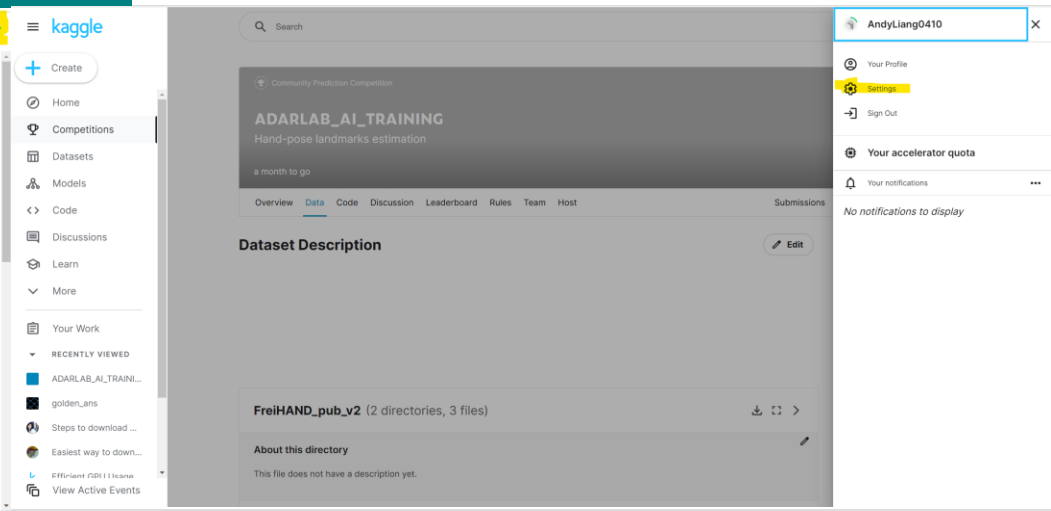
- Colab sample code and video introduction.

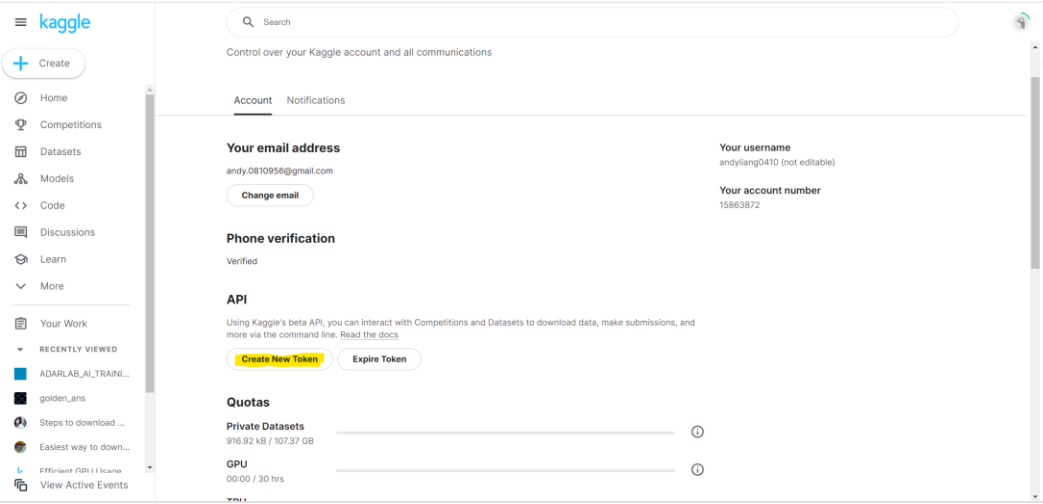
  https://drive.google.com/drive/folders/1tpLSH1wd476eeuwY7-tSZKDEey801Va-?usp=sharing

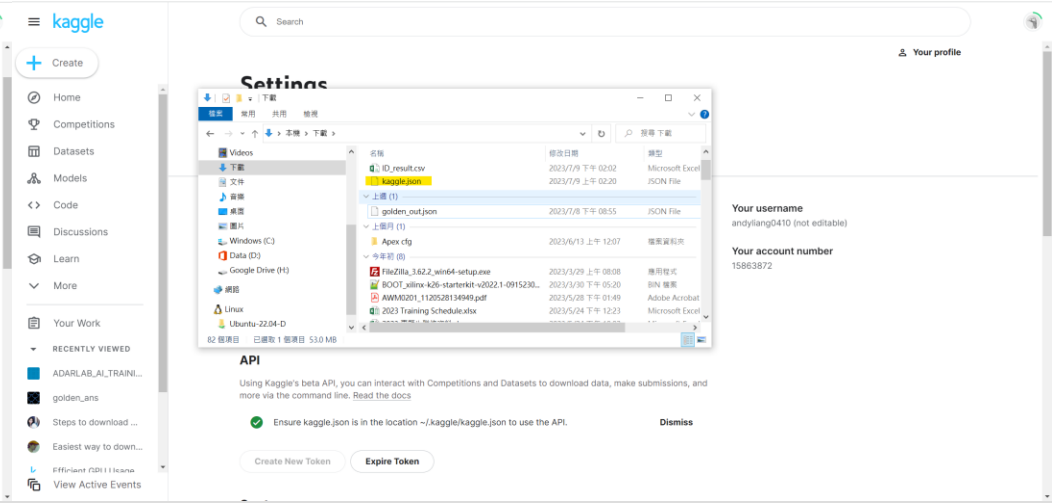
- 學習使用關鍵字搜尋任何你想問的問題

Kaggle usage









Thank you