Machine Learning HW2

MLTAs
ntueemlta2024@gmail.com

Links

- sample code
- kaggle
- report template
- math problem

Outline

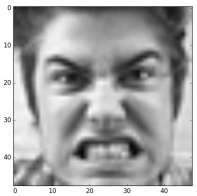
- HW2 - Facial Expression Classification
  - Dataset and Tasks Description
  - Sample Submission
- Kaggle
- Grading / Assignment Regulation

Task - Facial Expression Classification

本次作業為網路上收集到的人臉表情資料,

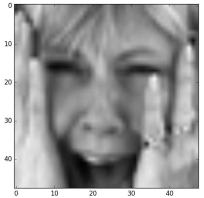
給定一張灰階的jpg照片 (1x64x64),

分辨出該照片中人物的表情屬於下列何者。

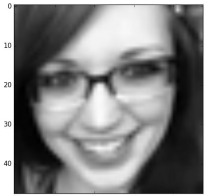


0(生氣)

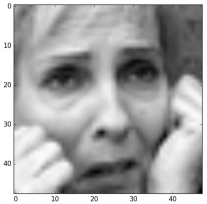
1(厭惡)



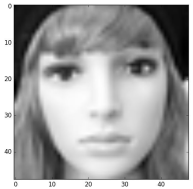
2(恐懼)



3(高興)



4(難過)



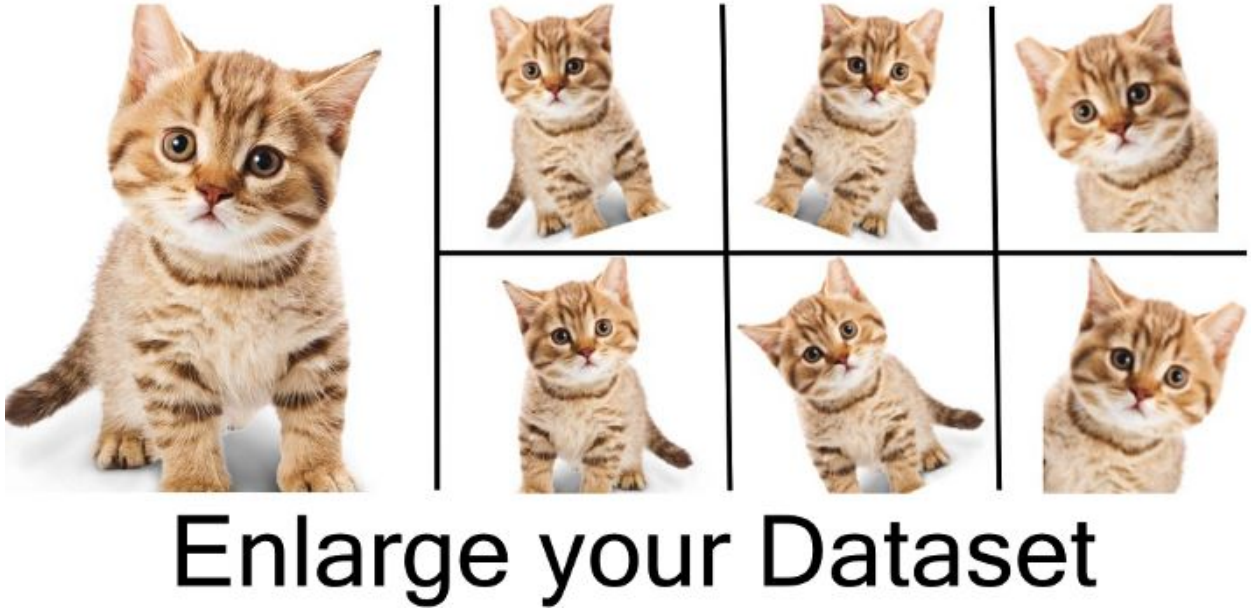
5(中立)

6(驚訝)

Task and Dataset

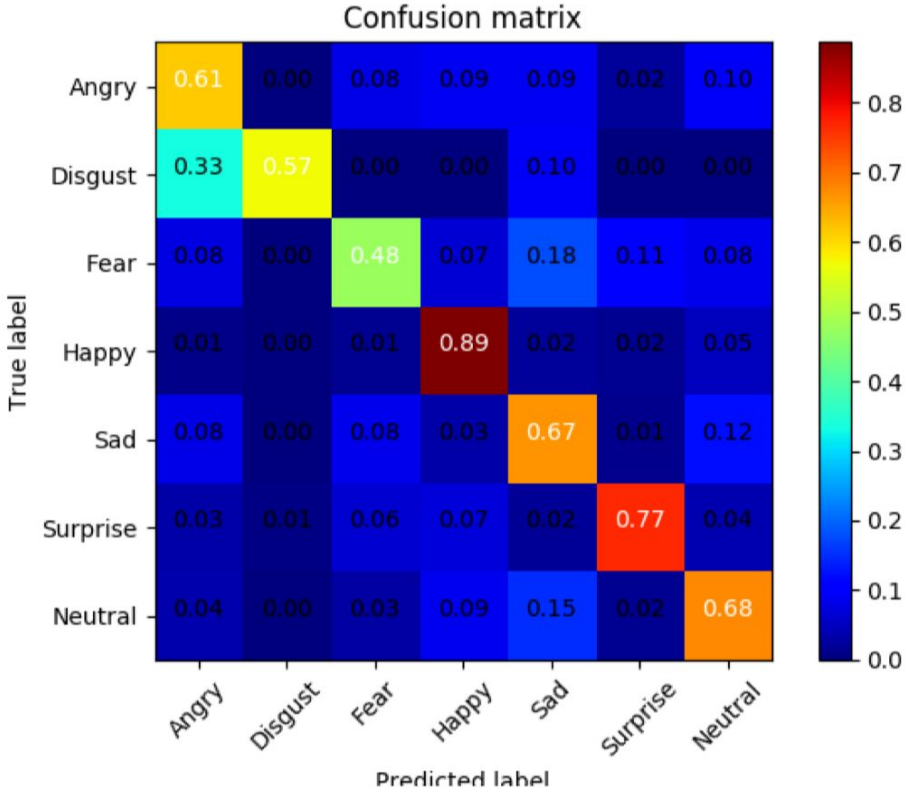
- Task : **CNN**
  - Build your own CNN model
    - You are welcomed to use pre-trained models
    - sample code
  - Report
    - Data Augmentation
    - Confusion Matrix
    - report template

Data Augmentation



ref: https://nanonets.com/blog/data-augmentation-how-to-use-deep-learning-when-you-have-limited-data-part-2/

Example - confusion matrix



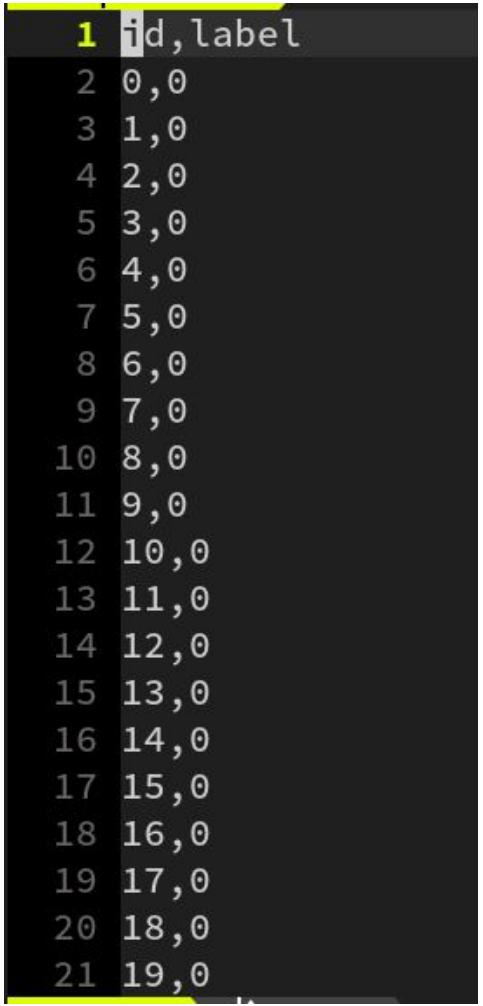
Kaggle Info & Deadline

- Link: https://www.kaggle.com/competitions/ml-2024-fall-hw2/overview
- 個人進行、不須組隊
- Team Name:
  - 修課學生: **學號_任意名稱 (ex: b09901105_謝博揚喜洋洋)**
  - 旁聽: 旁聽_任意名稱
- Maximum Daily Submission: 5 times
- Kaggle Deadline: 2024/10/18 23:59:59 (GMT+8)
- Cool Deadline: 2024/10/18 23:59:59 (GMT+8)
- 在Kaggle Deadline前可以選擇2份submission作為private score的評分依據。
  如果未勾選, 系統會自動選擇 Public Leaderboard中表現最佳的兩次。

Kaggle submission format

請預測test set中七千筆資料並將結果上傳Kaggle

1. 上傳格式為csv。
2. 第一行必須為id,label，第二行開始為預測結果。
3. 每行分別為id以及預測的label，請以逗號分隔。
4. Evaluation: Accuracy



Grading Criteria

- Kaggle - 4%
  - 超過public leaderboard的simple baseline分數： **1%**
  - 超過private leaderboard的simple baseline分數： **1%**
  - 超過public leaderboard的strong baseline分數： **1%**
  - 超過private leaderboard的strong baseline分數 : **1%**
- Programming report - 2%
  - report template
- Math problem - 6%
  - math problem
  - 若有和其他修課同學討論，請務必於題號前標明 collaborator（含姓名、學號）

Assignment Regulation

- 開放使用套件
  - numpy
  - pandas
  - pytorch
  - torchvision
  - cv2
  - pillow
  - 若需使用其他套件，請儘早寄信至助教信箱詢問，並請闡明原因。
- Please use CNN models
- No extra data allowed

Cool Submissions

在Cool上分別繳交以下檔案:

1. **report.pdf**
2. **math.pdf**
3. **code.ipynb**

其他規定 Other Policy

- Lateness
  - Cool 遲交每小時分數*0.95，兩天後歸0
  - 有特殊原因請找助教
- Runtime Error
  - 當程式錯誤，造成助教無法順利執行，請在公告時間內寄信向助教說明，修好之後重新執行所得kaggle部分分數將x0.5。

其他規定 Other Policy

- Cheating
  - 抄 code、抄 report（含之前修課同學）
  - 開設 kaggle 多重分身帳號註冊 competition
  - 於訓練過程以任何不限定形式接觸到testing data 的正確答案
  - 不得上傳之前的kaggle 競賽
  - 教授與助教群保留請同學到辦公室解釋coding作業的權利

References

1. https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/overview

2. https://nanonets.com/blog/data-augmentation-how-to-use-deep-learning-when-you-have-limited-data-part-2/