

Machine Learning HW2

MLTAs

ntueemlta2025@gmail.com

Links

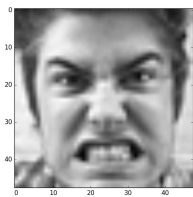
- [sample code](#)
- [kaggle](#)
- [report template](#)
- [math problem](#)

Outline

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

Task - Facial Expression Classification

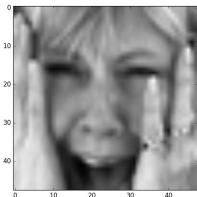
本次作業為網路上收集到的人臉表情資料，
給定一張灰階的jpg照片 (1x64x64)，
分辨出該照片中人物的表情屬於下列何者。



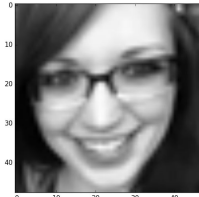
0(生氣)



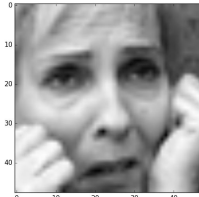
1(厭惡)



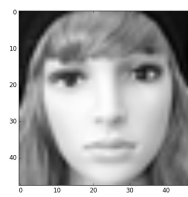
2(恐懼)



3(高興)



4(難過)



5(中立)

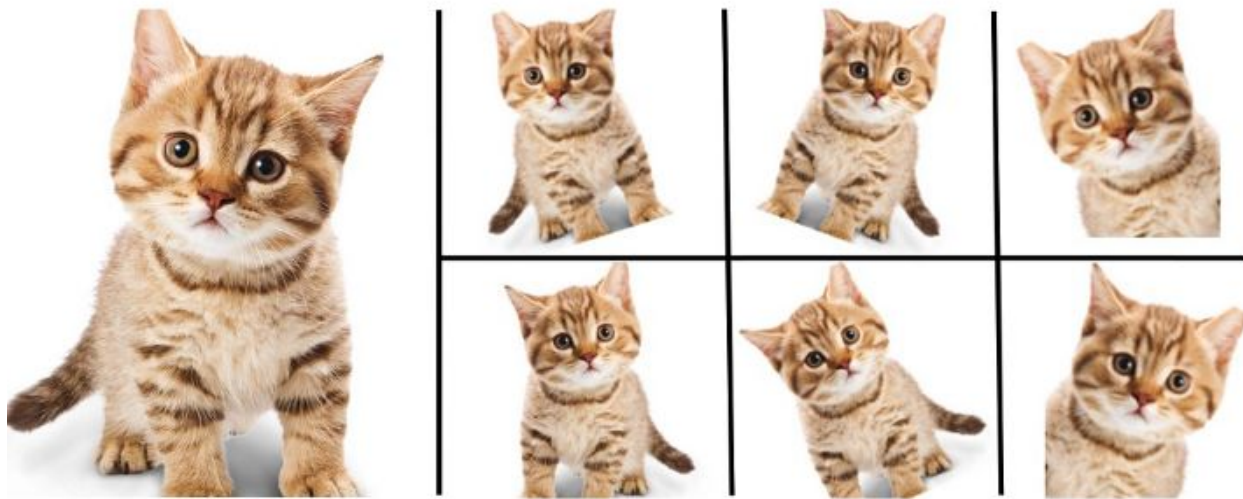


6(驚訝)

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](#)

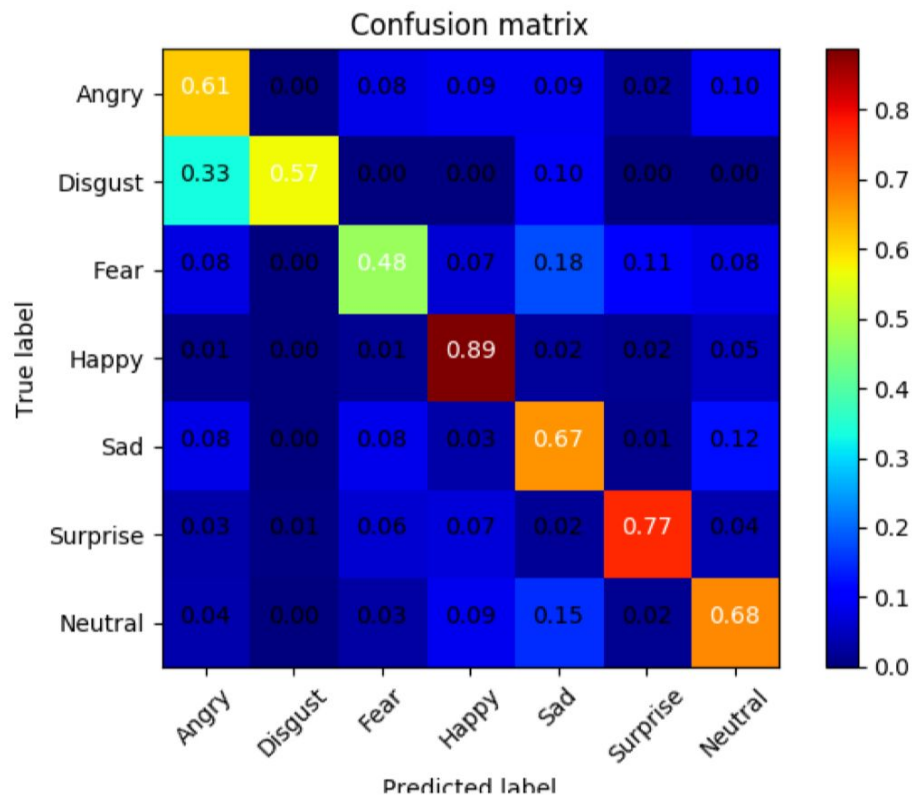
Data Augmentation



Enlarge your Dataset

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-2025-fall-hw-2>
- 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: 2025/10/17 23:59:59 (GMT+8)
- Cool Deadline: 2025/10/17 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

1	id, label
2	0,0
3	1,0
4	2,0
5	3,0
6	4,0
7	5,0
8	6,0
9	7,0
10	8,0
11	9,0
12	10,0
13	11,0
14	12,0
15	13,0
16	14,0
17	15,0
18	16,0
19	17,0
20	18,0
21	19,0

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**

Other Policy

- Lateness

- Cool 遲交
 - 一天: 以小時為單位線性遞減至七折
 - 兩天: 從七折線性遞減零分
- 不接受程式 or 報告單獨遲交
- 若有特殊原因請儘速聯絡助教

- Reproduce

- 請同學確保你上傳的程式所產生的結果, 會跟你在 kaggle 上的結果一致, 基本上誤差在 ± 0.5 之間都屬於一致, 若超過以上範圍, kaggle 將不予計分。

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

TA hour

- Mon 11:00~12:00 @總圖B1討論區
- Mon 13:00~14:00 @電二蟲洞
- Wed 12:00~14:00 @電二蟲洞
- Thurs 16:00~17:00 @電二蟲洞

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/>
3. <https://medium.com/jameslearningnote/%E8%B3%87%E6%96%99%E5%88%86%E6%9E%90-%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92-%E7%AC%AC5-1%E8%AC%9B-%E5%8D%B7%E7%A9%8D%E7%A5%9E%E7%B6%93%E7%B6%B2%E7%B5%A1%E4%BB%8B%E7%B4%B9-convolutional-neural-network-4f8249d65d4f>