

HW3 : Multi-Object Tracking

分組規則

- 每組1~3人
- 合作完成報告與實做成果
- 組內協調分工，確保每位同學都有參與
- 沒分組別到就自己一組

作業目標

在這次作業中，同學將透過實際操作一個開源的 Multi-Object Tracking (MOT) 模型，理解並體驗物體追蹤技術的整體流程，包括物體偵測 (Detection) 和關聯追蹤 (Association tracking) 等關鍵步驟。此作業的目的是讓同學深入瞭解影像處理領域中如何將偵測出的物體持續且穩定地追蹤於連續的畫面中，掌握目標定位與身份管理的核心技術。此外，同學將學習如何搭建並設定能夠穩定執行 MOT 模型的環境，包括軟體工具、套件安裝與必要的依賴設定，並且能夠成功執行現有的預訓練模型於指定的資料集上。同學也將實際操作資料處理、視覺化工具，並學習如何以清晰且標準化的方式呈現追蹤結果，包括產生視覺化的追蹤影片，以及輸出符合指定格式 (例如 CSV 或 TXT) 的資料檔案。透過這些流程的實踐，同學將不僅掌握 MOT 技術的核心原理，更培養分析與解決實際操作中可能遇到問題的能力，奠定未來在計算機視覺與影像分析相關領域研究或實務工作的良好基礎。

作業流程

1. 下載並整理資料集

根據以下網址下載對應的資料集：

- MOT17 : <https://motchallenge.net/data/MOT17/>
- Fisheye8k : <https://scidm.nchc.org.tw/en/dataset/fisheye8k>
- LOAF : loafisheye.github.io/download.html

可對資料集做預處理，不過須在報告中敘述做了什麼預處理。

2. 選擇並下載預訓練模型

由各組自行討論決定，推薦從 ByteTrack、DeepSORT、FairMOT (或其他公開開源可直接執行的 MOT 模型) 中擇一下載程式庫與預訓練權重。

3. 架設執行環境

根據所選擇的模型，在本地或伺服器上建立虛擬環境（Conda或 Docker 均可），並將安裝必要套件。

如果有 GPU/CUDA 需求，請註明版本並安裝對應的 PyTorch和CUDA版本。

4. 執行模型推理

因為本次作業不需要同學從頭訓練模型，只需要同學推理這些已經訓練好的模型，所以請各位同學將資料集中所有的資料(包括train和test資料夾的內容)都要輸入至模型中推理並得到結果。

5. 產生標準化輸出

每個資料集的結果 CSV 必須包含以下欄位：

```
frame_index, track_ID, x, y, w, h, scene_id
```

6. 將追蹤結果視覺化

視覺化影片請將每個 Bounding Box 括出並在左上角標上 ID，影片左上角標注場景名稱和frame index。

影片中也要畫出人物的移動軌跡，若是人物在畫面當中則畫出出現時間至當前時間點在畫面中的軌跡，反之，人物在畫面之外則不用呈現其軌跡。

可使用的MOT模型

建議使用以下開源的工具，或可選擇其他可執行的MOT模型：

- ByteTrack
 - paper : <https://arxiv.org/abs/2110.06864>
 - github : <https://github.com/ifzhang/ByteTrack.git>
- DeepSORT
 - paper : <https://arxiv.org/abs/1703.07402>
 - github : https://github.com/nwojke/deep_sort.git
- FairMOT
 - paper : <https://arxiv.org/abs/2004.01888>
 - github : <https://github.com/ifzhang/FairMOT.git>

報告撰寫

1. MOT模型介紹

- 說明模型的基本概念與運作流程
- 說明為何選擇該模型

2. 環境架設說明

必寫，助教將依此重現作業

請清楚列出環境架設流程,, 至少包括以下：

- 作業系統（推薦Ubuntu 22.04，其他也可以）
- Python版本
- 所需的套件和安裝方法（例如 `pip install -r requirements.txt`）
- 特殊依賴項（例如 PyTorch、CUDA 版本等）

可以使用colab或在windows環境下運行，只要附上詳細步驟以及確保能被正常執行即可。

3. MOT模型運行流程說明

必寫，助教將依此重現作業

說明在安裝完環境之後的步驟，內容包括在終端機執行的指令與程式架構。

4. 輸出結果說明

- 說明運行MOT模型後的結果
- 實做中遇到怎樣的困難？
- 結果的哪部份表現不佳？
- 可以如何改進？

5. 分工表

寫上同組組員各自負責的內容。

若是自己一組則不用寫這部分。

繳交項目

1. 作業報告 (20%)

2. 程式碼 (20%)

3. 追蹤結果 (.csv或.txt) (30%)

- 欄位須為 [frame index, track ID, x, y, w, h, scene id]
- 可儲存為.csv檔或.txt檔
- scene id 是各個不同場景的名稱:
 - MOT17 : [MOT17-01, MOT17-02, ..., MOT17-14]
 - Fisheye8k : [camera1_A, camera2_A, camera3_A, camera3_N, ...]
 - LOAF : [0000, 0001, 0002, ..., 0074]

4. 追蹤結果可視化影片 (.mp4) (30%)

- 利用追蹤結果用Bounding Box將人物在影片畫面中匡選出來並在Bounding Box左上角附上其ID, 以及在繪製出人物移動軌跡。
- 需要在影片的左上角標注是在哪個場景以及在第幾個frame的畫面。

繳交期限與方式

作業所有內容打包成HW3_<組長學號>.zip檔案，並應於5/23前上傳至E3教學平台。

(若為單人組，同學組長學號即為自己學號)

內容結構應為：

```
HW3_<組長學號>
|___HW3_report_<組長學號>.pdf
|
|___MOT17
| |___MOT17_codes
| |
| |___MOT17_track_result
| | |___train
| | | |___MOT17_train_result.csv
| | |___test
| | | |___MOT17_test_result.csv
| |
| |___MOT17_videos
| | |___train
| | | |___MOT17_train_video.mp4
```

```

|   |___test
|   |   |___MOT17_test_video.mp4
|
|___Fisheye8k
|   |___Fisheye8k_codes
|   |
|   |___Fisheye8k_track_result
|   |   |___train
|   |   |   |___Fisheye8k_train_result.csv
|   |   |
|   |   |___test
|   |   |   |___Fisheye8k_test_result.csv
|   |
|   |___Fisheye8k_videos
|   |   |___train
|   |   |   |___Fisheye8k_train_video.mp4
|   |   |___test
|   |   |   |___Fisheye8k_test_video.mp4
|
|___LOAF
|   |___LOAF_codes
|   |
|   |___LOAF_track_result
|   |   |___train
|   |   |   |___LOAF_train_result.csv
|   |   |___test
|   |   |   |___LOAF_test_result.csv
|   |
|   |___LOAF_videos
|   |   |___train
|   |   |   |___LOAF_train_video.mp4
|   |   |___test
|   |   |   |___LOAF_test_video.mp4

```

輸出結果範例

- 追蹤結果

frame	track_id	x	y	w	h	scene_id
00000	0	181.61117445618032	230.3300711117286	53.00703455793268	120.27481526793859	0059
00000	1	216.27006321075123	187.21647270272894	42.43173606412839	106.08745369247023	0059
00000	2	667.4229342626304	363.56452235433585	64.12172610406839	171.94599443098485	0059
00000	3	45.85372735174324	578.4960223294856	21.779282334215694	44.71166331304412	0059
00000	4	52.42528952230721	591.6633184919847	21.904636074251545	47.753070835881665	0059
00000	5	64.76716799072595	677.5932263480252	26.059807134043947	22.348065041023396	0059
00000	6	48.623496741342365	640.5304399437192	15.088006063651275	35.96725068647144	0059
00000	7	44.74310160516495	652.201245289541	20.777883460459172	38.64727156129265	0059
00000	8	476.54617185767876	964.3902077379657	31.224887671496997	34.84291188545307	0059
00000	9	824.6368363100189	824.8214321005198	16.744120159068835	32.50079840982267	0059
00000	10	789.9393286928228	897.189928041592	9.278119482775839	14.943216092390216	0059
00000	11	56.97416604715416	645.0809163406784	24.963615809395332	25.192849317811163	0059
00015	2	668.0349811340336	362.26392921573535	64.80910725044313	172.2585296423504	0059
00015	0	183.783362501333	236.98838918758776	52.394023862272626	112.62406688708862	0059
00015	1	207.10423774892612	192.12995982736962	55.72720714069329	104.30059421056394	0059
00015	12	200.41611979825385	205.13655944400887	32.687632726266266	81.57382817198354	0059
00015	9	807.7128177879348	830.5014585342507	21.153908604574426	36.050068276018614	0059
00015	8	528.7842104950712	965.2336330163295	17.496062537956135	33.77005126296222	0059
00015	13	428.75679851396086	961.891104314771	28.23580386221834	34.18683510769121	0059
00015	3	48.782256936119616	580.3605490568285	19.12595657500921	41.80323557587065	0059
00015	4	48.17678314778648	590.027606798244	25.73261248660707	50.930693344393596	0059
00015	5	64.76026096392171	677.3224942416582	25.286577585496435	23.105834198489912	0059
00015	7	45.819835241561165	652.7783093058526	20.920674955006866	40.239537929894	0059
00015	6	46.60665689991521	637.003285800429	19.14717937521128	38.67968459360808	0059
00015	11	57.20288816525129	645.4796078155729	23.4107844450732866	24.7036747266672	0059
00015	14	13.58854933861346	580.1830240966402	11.363748876517963	19.55548527132484	0059
00030	15	222.01391745247406	160.03466888919922	28.00050668503901	88.9239783413882	0059
00030	1	217.0071506303967	199.1065497424574	62.437294333903246	97.19678555602914	0059
00030	12	198.43227103240258	192.5934482818833	54.63305361310917	84.8596193979182	0059
00030	0	182.3264048323517	230.52871818556582	50.28941916179035	119.87557756539046	0059
00030	2	667.6879346168968	361.8449974995272	66.69821353353568	171.96517569237236	0059

- 追蹤影片 (範例為截圖)

