

roadmap 梳理成一份 詳細未來計劃表，包括階段、任務、交付物、時間建議同埋檢查點。咁樣你就可以清楚知道每一步做咩、幾時完成、點樣驗收。

---

### 📍 未來計劃表

#### Phase 0：模型基線確定 ✅

- 已完成：
    - Teacher model = (RN18 + B3) ensemble
    - Student model = ViT small (最佳結果)
  - 交付物：模型清單、版本鎖定、環境設定文檔
- 

#### Phase 1：核心 + 對照組實驗 ( Offline )

時間建議：1 – 1.5 週

- 任務：
  - 跑完核心 set ( CNN ensemble 、 CNN+ViT ensemble )
  - 跑完對照組 ( multi-teacher 分路、四師均權，精簡  $\beta=4$  )
- 交付物：
  - 統一格式結果表 ( Macro-F1 、 Minority F1 、 ECE ， mean/std across seeds )
  - 圖表 ( trend curves 、 bar chart )
- 檢查點：
  - $\beta^*$  已決定

- 主線配置明確（作為 demo 選型依據）
  - 報告：
    - 撰寫 offline 對照報告（方法、結果、初步討論）
- 

## Phase 2 : Demo Testing

時間建議：3 – 5 日

- 任務：
    - 用 demo script 測試三個模型：MNV3 最佳 student、ViT small 最佳 student、Teacher baseline
    - 覆蓋不同場景（光線、距離、角度）
  - 交付物：
    - Demo 記錄表（Latency、FPS、False accept/reject、Calibration observation）
    - Demo 場景矩陣（光線/角度/距離）
  - 檢查點：
    - 選出最適合 real-time demo 噢 student（通常係 MNV3）
  - 報告：
    - 撰寫 demo 報告（直觀觀察 + 數據化指標）
- 

## Phase 3 : 部署選型（ONNX / 其他）

時間建議：3 - 4 日

- 任務：

- 將最佳 student export 成 ONNX
- 驗證一致性 (ONNX vs PyTorch 輸出差異  $<1e-5$ )
- 嘗試優化 (simplify、FP16、INT8)
- 測試性能 (Latency、FPS、Memory)

- 交付物：

- ONNX 模型 (FP32/FP16/INT8)
- 部署測試報告 (性能對比表)

- 檢查點：

- 精度降幅  $\leq 0.5\%$ ，性能顯著提升
- 

Phase 4：中期報告撰寫

時間建議：1 週

- 任務：

- 撰寫完整報告 (方法、結果、demo、部署、討論)
- 整合 offline + demo + 部署結果

- 交付物：

- 中期報告初稿 → 內審 → 完整版本
  - 檢查點：
  - 報告結構完整，圖表齊全，結論清晰
-

## Phase 5：弱點分析與改進

時間建議：1 – 2 週

- 任務：

- 5.1 Data imbalance → 分析類別分佈，測試 class-balanced sampling/focal loss
- 5.2 Data quality → 檢查標註錯誤、模糊、遮擋，測試 CLAHE/denoise
- 5.3 外在影響 → 光線、膚色、年齡、鏡頭解像度、FPS 測試
- 5.4 硬件影響 → 不同 GPU/CPU、相機，記錄 FPS/latency
- 5.5 Demo 參數設計 → temperature、smooth、unknown-thresh、bbox-smooth
- 5.6 其他 → adversarial robustness、domain shift

- 交付物：

- 弱點分析報告（每項至少一個結論 + 改進建議）
  - 檢查點：
    - 弱點清單 + 下一步行動方案
-

## ⌚ 時間總覽（兩個半月內）

- Week 1 – 2 : Phase 1 (核心 + 對照完成，offline 報告)
  - Week 3 : Phase 2 (Demo 測試 + 報告)
  - Week 4 : Phase 3 (ONNX 部署 + 報告)
  - Week 5 : Phase 4 (中期報告撰寫)
  - Week 6 – 7 : Phase 5 (弱點分析)
  - Week 8 – 9 : 緩衝、補實驗、報告修訂
- 

呃樣你就有一個完整嘅 roadmap：由 offline → demo → 部署 → 報告 → 弱點分析，每一步都有交付物同檢查點，確保唔會迷失。

要唔要我幫你將呢個計劃表再轉成 Excel/CSV 欄位設計，方便你直接用嚟 tracking？

好，咁我幫你將「數據圖表部分」完整梳理，除咗你列出嘅 1 – 7，我再幫你 mark 低更多常見而且有用嘅圖表/表格，然後再設計一個 詳細記錄表 template，方便你逐個逐個填。

---

 數據圖表清單（建議完整版本）

## A. 數據分佈

## 1. 各類型 data 數量

- Training / Validation / Testing 數量

- 各 class 分佈 (bar chart)

## B. Teacher / Student 對比

## 2. 各類型對比表

- Single teacher、Top-3 ensemble teacher

- Top-3 CNN student、Top-3 ViT student

- 指標：Macro-F1、Minority F1、ECE、Latency/FPS (如有)

## 3. Ensemble 對比表

- Pre-ensemble vs Post-ensemble vs Multi-teacher split

- KD vs DKD ( $\beta = 4$ )

## C. 訓練過程

## 4. 成長曲線 (Learning curves)

- Macro-F1 / Loss / ECE 隨 epoch 變化

- KD vs DKD 對比

## D. 錯誤分析

## 5. Confusion matrix (至少 2 個：最佳 CNN student、最佳 ViT student)

## 6. Per-class F1 (bar chart, highlight minority class)

## 7. Per-class Pain Points

- 哪些 class consistently 低分
- 哪些 class 容易混淆 (由 confusion matrix 對應)

E. 進階補充 (我幫你 mark 低)

8. Calibration plot (Reliability diagram) → 檢查 ECE

9. ROC / PR curves (macro vs minority class)

10. Latency vs Accuracy trade-off (PyTorch vs ONNX FP32/FP16/INT8)

11. Demo 場景對比表 (光線/角度/距離 → 誤判率/FPS)

12. Ablation summary (不同  $\beta$ 、不同  $\lambda$  權重 → Macro-F1/ECE)

---

 詳細記錄表 Template

以下係一個 統一表格設計，你可以用 Excel/CSV/Notion 去 tracking，每個 run 或每個圖表都可以填：

Section	Sub-item	Model/Config	Seed	Metric1 (Macro-F1)	Metric2 (Minority F1)	Metric3 (ECE)	Latency/FPS s	Notes
Data 分佈	數據量	—	—	Train=xxxx	Test=xxxx	—	—	各 class 分佈圖
Teacher/Student 對比	Single teacher	RN18	—	78.5	65.2	0.12	—	baseline
Teacher/Student 對比	Ensemble teacher	RN18+B3	—	79.4	66.0	0.11	—	teacher final
CNN student	MNV3 best	1337	71.2	58.0	0.18	40 FPS	最佳 CNN	
ViT student	ViT small	2025	73.8	61.5	0.15	25 FPS	最佳 ViT	
Ensemble 對比	CNN+ViT ensemble	KD	1337	72.5	59.0	0.16	—	—
Learning curve	MNV3 KD	1337	—	curve	curve	curve	—	圖片連結
Confusion matrix	MNV3 best	1337	—	—	—	—	—	圖片連結
Per-class F1	ViT small	2025	—	class-wise F1	—	—	—	minority class 低分
Calibration	ViT small	2025	—	—	—	ECE plot	—	reliability diagram
Demo	MNV3 best	—	—	—	—	—	40 FPS	光線暗 → 誤判 ↑
Deployment	MNV3 ONNX FP16	—	—	71.0	57.8	0.16	80 FPS	精度 drop <0.5%

---

## ✓ 用法

- 每個 Section (Data / Teacher vs Student / Ensemble / Learning curve / Confusion matrix / Demo / Deployment) 都可以有多行記錄。
  - 圖表：用 Notes 或者加一列「File/Path」去存放圖片連結。
  - 報告引用：到時中期報告直接 copy 數據 + 插圖。
- 

👉 建議你先從 1 - 7 基本圖表開始，逐個完成；之後再加上我幫你 mark 低  
慨進階圖表 (Calibration、ROC/PR、Latency vs Accuracy、Demo 場景)。

要唔要我幫你將呢個 記錄表 template 再拆成「Offline 部分」、「Demo 部  
分」、「Deployment 部分」三個子表，方便你分階段填？