

一、實驗設計與結果

本章節中，根據第三章之研究方法以各項實驗驗證系統之設計，並介紹各實驗之目的與設計、評估方式以及結果與分析，透過以下五章子節進行說明：臉部偵測準確度實驗、臉部偵測執行時間實驗、臉部遮擋辨識實驗、姿勢辨識實驗及影片危險偵測實驗。

另外，本章各項實驗皆於相同硬體環境操作，詳細資訊如下：

- 作業系統：Windows 10
- CPU：Intel(R) Core(TM) i7-10700KF CPU @ 3.80GHz
- 記憶體：128GB
- GPU：NVIDIA GeForce GTX 1660

1.1 臉部偵測準確度實驗

在收集嬰兒臉部資料集時，需針對嬰兒影像擷取出臉部範圍，進而後續之臉部遮擋辨識階段。

為了使本系統擁有較佳的臉部偵測準確性且兼具執行效能，本文透過1.1節及1.2節之實驗，分別進行臉部偵測演算法準確度與執行時間之比較，進而驗證以下設計：先使用 SSD 演算法偵測嬰兒臉部，此方法之召回率雖低，但其準確度很高，故能利用此算法之時間優勢；而若 SSD 演算法找不到嬰兒面部時，則接續使用 RetinaFace 演算法，利用其正確率及準確率皆高之優點進行嬰兒臉部偵測。

1.1.1 實驗目的與設計

本實驗為計算人臉偵測演算法之嬰兒面部擷取準確度，使用??節的嬰兒姿勢資料集共 15416 張影像，分析 RetinaFace **deng_retinaface_2020**、MTCNN **zhang_joint_2016**、SSD **ye_face_2021**及 Haar cascade **goyal_face_2017**等演算法之偵測結果。

1.1.2 實驗評估方式

透過分類標註四項演算法偵測嬰兒臉部之結果影像，計算出各演算法的 accuracy、precision 及 recall。

1.1.3 實驗結果與分析

由表 1.1及表 1.2 可見 RetinaFace 與 MTCNN 之實驗結果。而表 1.3及表 1.4 為 SSD 與 Haar cascade 之實驗結果，可發現其將多數影像皆誤判為無臉（False），亦即影像中有嬰兒臉部畫面但演算法未偵測之，故此部分僅關注判斷為有臉（True）之數據統計。

表 1.1: RetinaFace **deng_retinaface_2020**偵測嬰兒臉部結果

	True（預測有臉）	False（預測無臉）
True （實際有臉）	12925	11
False （實際無臉）	33	2447
Total	12958	2458

再經計算後，四項演算法之 accuracy、precision 及 recall 值如下：

1. RetinaFace：accuracy 為 99.78%、precision 為 99.75% 以及 recall 為 99.91%。

表 1.2: MTCNN zhang_joint_2016偵測嬰兒臉部結果

	True (預測有臉)	False (預測無臉)
True (實際有臉)	9361	3399
False (實際無臉)	517	2140
Total	9877	5539

表 1.3: SSD ye_face_2021偵測嬰兒臉部結果

	True (預測有臉)	False (預測無臉)
True (實際有臉)	4830	8141
False (實際無臉)	5	2440
Total	4835	10581

2. MTCNN：accuracy 為 74.60%、precision 為 94.78% 以及 recall 為 73.36%。
3. SSD：僅關注判斷為有臉 (True) 之數據，其 precision 為 99.90%。
4. Haar cascade：僅關注判斷為有臉 (True) 之數據，其 precision 為 79.90%。

因此，透過本實驗結果可得出選用 RetinaFace 演算法行嬰兒臉部偵測，可擁有較佳的偵測準確度。

表 1.4: Haar cascade **goyal_face_2017**偵測嬰兒臉部結果

	True (預測有臉)	False (預測無臉)
True (實際有臉)	2882	9546
False (實際無臉)	725	2263
Total	3607	11809

1.2 臉部偵測執行時間實驗

本研究進行嬰兒臉部偵測除了考量準確度外，亦希望提升整體系統之執行效率。

1.2.1 實驗目的與設計

本實驗為計算人臉偵測演算法之執行時間，使用??節的嬰兒姿勢資料集共 15416 張影像，分析 RetinaFace **deng_retinaface_2020**、MTCNN **zhang_joint_2016**、SSD **ye_face_2021**及 Haar cascade **goyal_face_2017**等演算法之偵測結果。

1.2.2 實驗評估方式

透過計算四項演算法偵測完整資料集所花費之時間，計算各演算法平均偵測一張影像之執行時間。

1.2.3 實驗結果與分析

RetinaFace、MTCNN、SSD 及 Haar cascade 四項演算法偵測 15416 張影像之詳細實驗結果如下：

1. RetinaFace：共 5 小時 42 分 2.10 秒，平均每張影像需 1.33 秒。
2. MTCNN：共 2 小時 8 分 22.05 秒，平均每張影像需 0.50 秒。

3. SSD：共 9 分 17.26 秒，平均每張影像需 0.04 秒。

4. Haar cascade：共 18 分 01.78 秒，平均每張影像需 0.07 秒。

因此，透過本實驗結果可得出使用 SSD 演算法進行嬰兒臉部偵測，將可擁有較佳的偵測速度。

總結1.1節與1.2節之實驗結果，驗證本系統先使用 SSD 演算法偵測嬰兒臉部，未如期找到目標時，則改以 RetinaFace 演算法偵測，達成兼具準確性及執行效率之系統目標。

1.3 臉部遮擋辨識實驗

本研究中，利用深度學習技術辨識嬰兒臉部是否遭非奶嘴之異物遮蔽，進而判斷嬰兒是否處於危險情境中。

1.3.1 實驗目的與設計

本實驗為訓練針對嬰兒臉部遮擋辨識之模型，以 ResNet50 he_deep_2016訓練??節的嬰兒臉部資料集，並透過驗證集進行模型驗證。

程式實作中，網路訓練回合數為 20，設定影像資料大小為 224x224，包含三個類別（臉部無遮擋之安全狀態、使用奶嘴及面部遭異物遮蔽之警示狀態），且透過 data augmentation 技術生成訓練及測試資料，輸出層使用 softmax 作為激發函數，並使用 Adam 作為 optimizer 且將學習率設為 0.000001 以進行微調。

1.3.2 實驗結果分析

模型最終訓練準確率達 98.06%，而測試準確率達 99.43%，詳細訓練結果請見圖 1.1。

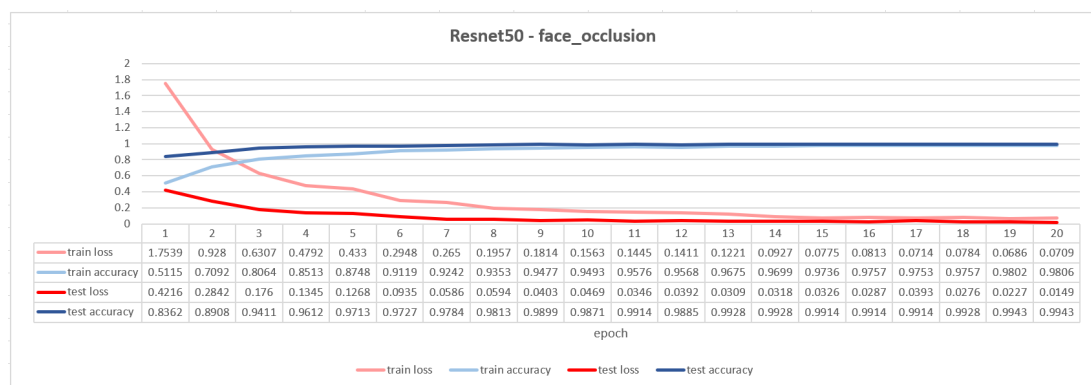


圖 1.1: 臉部辨識訓練及測試結果

我們使用 342 張之驗證集影像進行模型驗證，所有影像皆辨識正確。此模型之混淆矩陣如表 1.5，表中數字為各類影像之張數及比例。

表 1.5: 臉部遮擋辨識模型之混淆矩陣（單位：張 / 百分比）

		預測類別		
		安全	奶嘴	警示
實際類別	安全	120 (100.00)	0 (0.00)	0 (0.00)
	奶嘴	0 (0.00)	115 (100.00)	0 (0.00)
	警示	0 (0.00)	0 (0.00)	107 (100.00)

1.4 姿勢辨識實驗

本研究中，利用深度學習技術辨識嬰兒基礎姿勢，進而判斷嬰兒是否處於危險情境中。

1.4.1 實驗目的與設計

本實驗為訓練針對嬰兒姿勢辨識之模型，以 ResNet50 he_deep_2016 訓練節的嬰兒姿勢資料集，並透過驗證集進行模型驗證。

程式實作中，網路訓練回合數為 20，設定影像資料大小為 224x224，包含四個類別（正躺、趴躺、坐姿及站立），且透過 data augmentation 技術生成訓練及測試資料，輸出層使用 softmax 作為激發函數，並使用 Adam 作為 optimizer 且將學習率設為 0.000001 以進行微調。

1.4.2 實驗結果分析

模型最終訓練準確率達 99.45%，而測試準確率達 99.71%，詳細訓練結果請見圖 1.2。



圖 1.2: 姿勢辨識訓練及測試結果

我們使用 744 張之驗證集進行模型驗證，包含了五張類別辨識錯誤的影像，其中三張將坐姿誤判為趴躺姿勢，推測原因為嬰兒雖呈現坐姿，但上半身貼近其腿部（如圖 1.3），而導致誤判。此模型之混淆矩陣如表 1.6，表中數字為各類影像之張數及比例。



圖 1.3: 姿勢辨識錯誤之影像：坐姿誤判為趴躺

表 1.6: 姿勢辨識模型之混淆矩陣（單位：張 / 百分比）

		預測類別			
		正躺	趴躺	坐姿	站立
實際 類別	正躺	164 (100.00)	0 (0.00)	0 (0.00)	0 (0.00)
	趴躺	1 (0.52)	191 (99.48)	0 (0.00)	0 (0.00)
	坐姿	0 (0.00)	3 (1.50)	196 (98.00)	1 (0.50)
	站立	0 (0.00)	0 (0.00)	0 (0.00)	192 (100.00)

1.5 影片危險偵測實驗

本研究基於嬰兒影像進行臉部遮擋及姿勢辨識，透過讀取嬰兒影片達成危險監測之目標。

1.5.1 實驗目的與設計

本實驗為驗證此系統能基於嬰兒影像進行危險監測，利用網路之真實嬰兒影片，包含不同之拍攝視角、嬰兒樣貌及狀態等，實驗臉部遮擋辨識模型與姿勢辨識模型之準確性。

1.5.2 實驗評估方式

透過輸出每幀影像之臉部遮擋及姿勢辨識結果，計算其 accuracy、precision 及 recall，以驗證此二模型得以應用在監測嬰兒危險情境。

1.5.3 實驗結果分析

本實驗影片為嬰兒正躺於畫面中，並包含使用奶嘴及未使用奶嘴之情境，共切成 3374 幀影像，將未拍攝到嬰兒畫面之影像刪除後，剩餘 3307 張嬰兒影像進行辨識。

首先，姿勢辨識的部分，包含了 278 張誤判為趴躺姿勢的影像，推測原因為嬰兒身體遭棉被遮擋（如圖 1.4），而只拍攝到露出的嬰兒臉部，

故造成姿勢辨識錯誤。



圖 1.4: 姿勢辨識錯誤之影像：正躺誤判為趴躺

其次，臉部遮擋辨識的部分，包含 36 張嬰兒臉部為被偵測之影像（如圖 1.5），另有多張影像類別應為嬰兒正在使用奶嘴或安全狀態，但誤判為遭異物遮蔽之警示狀態，推測原因為影像中之奶嘴或嬰兒臉部遭手部等遮擋（如圖 1.6），而誤判類別。



圖 1.5: 未偵測嬰兒臉部之影像

兩部分之混淆矩陣如表 1.7及表 1.8，表中數字為各類影像之張數及比例。

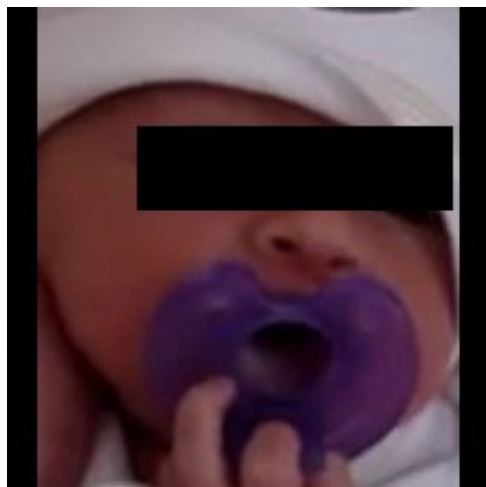


圖 1.6: 臉部遮擋誤判之為警示狀態

表 1.7: 實驗影片姿勢辨識之混淆矩陣（單位：張 / 百分比）

		預測類別			
		正躺	趴躺	坐姿	站立
實際類別	正躺	3029 (91.59)	278 (8.41)	0 (0.00)	0 (0.00)
	趴躺	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
	坐姿	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
	站立	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)

表 1.8: 實驗影片臉部遮擋辨識之混淆矩陣（單位：張 / 百分比）

		預測類別		
		安全	奶嘴	警示
實際類別	安全	1223 (83.82)	52 (3.56)	184 (12.61)
	奶嘴	17 (1.01)	1308 (77.72)	358 (21.27)
	警示	0 (0.00)	0 (0.00)	129 (100.00)