

# 一、實驗設計與結果

## 1.1 臉部偵測準確度實驗

### 1.1.1 實驗目的與設計

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 4.1 節及 4.2 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，以驗證以下設計：先使用 SSD 演算法偵測嬰兒臉部，此方法雖常未能如期找到嬰兒臉部範圍，但其準確度很高，故能利用此算法之時間優勢；而為補足 SSD 演算法召回率不足之缺失，若其找不到嬰兒臉部時，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。

故本實驗使用 3.3.1 節之嬰兒姿勢資料集，分析 OpenCV `goyal_face_2017`、SSD `ye_face_2021`、MTCNN `zhang_joint_2016` 及 RetinaFace `deng_retinaface_2016` 等臉部偵測演算法，計算其臉部擷取之準確度。

### 1.1.2 實驗評估方式

本實驗針對四項演算法之準確度進行比較，將嬰兒臉部偵測結果影像進行分類標註，以計算出各演算法之 accuracy、precision 及 recall。

### 1.1.3 實驗結果與分析

由表 1.1 及表 1.2 可見 MTCNN 與 RetinaFace 之實驗結果。而表 1.3 及表 1.4 為 OpenCV 與 SSD 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: 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	994
False (實際無臉)	517	4544
Total	<b>9878</b>	<b>5538</b>

表 1.2: 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	<b>12958</b>	<b>2458</b>

表 1.3: OpenCV **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	11809
False (實際無臉)	725	

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

1. OpenCV：由於多數影像皆誤判為無臉 (False)，故僅計算其 precision 為 79.90%。

表 1.4: 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	10581
False (實際無臉)	5	

2. SSD：由於多數影像皆誤判為無臉 (False)，故僅計算其 precision 為 99.90%。
3. MTCNN：accuracy 為 90.20%、precision 為 94.76% 以及 recall 為 90.93%。
4. RetinaFace：accuracy 為 99.78%、precision 為 99.75% 以及 recall 為 99.91%。

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。

## 1.2 臉部偵測執行時間實驗

### 1.2.1 實驗目的與設計

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。

故本實驗使用 3.3.1 節之嬰兒姿勢資料集，分析 OpenCV goyal\_face\_2017、SSD ye\_face\_2021、MTCNN zhang\_joint\_2016 及 RetinaFace deng\_retinaface\_2016 等臉部偵測演算法之執行時間，以驗證適合本系統之演算法。

### 1.2.2 實驗評估方式

本實驗針對四項演算法之執行速度進行比較，透過計算演算法偵測 15416 張影像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

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

1. OpenCV：共 18 分 01.78 秒，平均每張影像需 0.07 秒。
2. SSD：共 9 分 17.26 秒，平均每張影像需 0.04 秒。
3. MTCNN：共 2 小時 8 分 22.05 秒，平均每張影像需 0.50 秒。
4. RetinaFace：共 5 小時 42 分 2.10 秒，平均每張影像需 1.33 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 4.1 節與 4.2 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

### 1.3.1 實驗目的與設計

本系統為偵測嬰兒臉部是否遭非奶嘴之異物遮擋，使用 3.2.2 節之嬰兒臉部資料集以 ResNet50 he\_deep\_2016 訓練模型，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、臉部遭奶嘴遮擋及臉部遭異物遮擋之危險狀態），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

本實驗訓練之模型其最終訓練準確率達 98.06%，而測試準確率達 99.43%，詳細訓練結果請見圖 1.1。

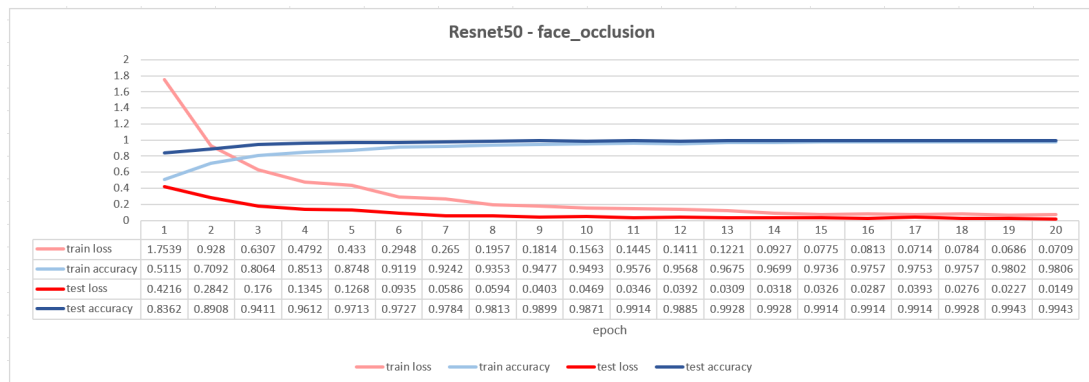


圖 1.1: 臉部辨識訓練結果

接著，再使用 342 張之驗證集影像進行模型驗證，所有影像皆辨識正確，其混淆矩陣如表 1.5，表中數字為各類影像之張數及占比。

表 1.5: 臉部遮擋辨識模型之混淆矩陣

		預測分類		
		無遮蔽	使用奶嘴	有異物
實際分類	無遮蔽	120 (1.00)	0 (0.00)	0 (0.00)
	使用奶嘴	0 (0.00)	115 (1.00)	0 (0.00)
	有異物	0 (0.00)	0 (0.00)	107 (1.00)

## 1.4 姿勢辨識實驗

### 1.4.1 實驗目的與設計

本系統為辨識嬰兒姿勢是否處於危險狀態，使用 3.3.1 節之資料集以 ResNet50 he\_deep\_2016 訓練模型，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

本實驗訓練之模型其最終訓練準確率達 99.45%，而測試準確率達 99.71%，詳細訓練結果請見圖 1.2。

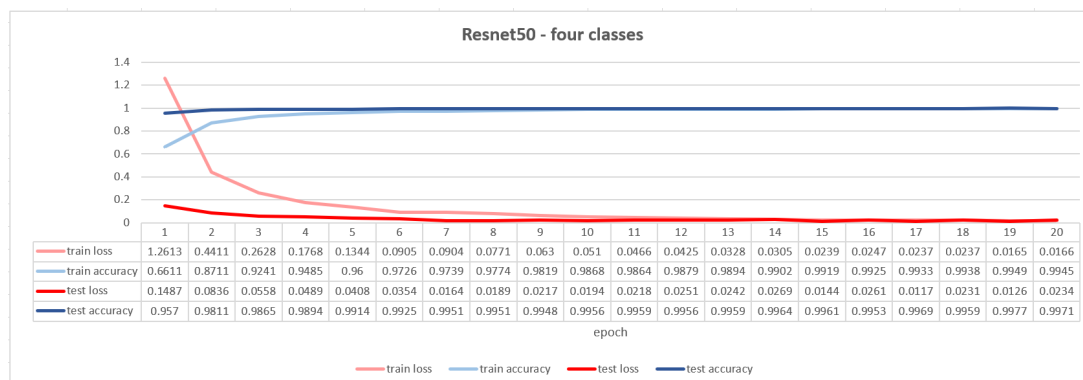


圖 1.2: 姿勢辨識訓練結果

接著，再使用 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占比。



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

		預測分類			
		正躺	趴躺	坐姿	站立
實際分類	正躺	164 (1.00)	0 (0.00)	0 (0.00)	0 (0.00)
	趴躺	1 (0.01)	191 (0.99)	0 (0.00)	0 (0.00)
	坐姿	0 (0.00)	3 (0.02)	196 (0.98)	1 (0.01)
	站立	0 (0.00)	0 (0.00)	0 (0.00)	192 (1.00)

## 1.5 影片危險偵測實驗

### 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

本實驗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。此部分之混淆矩陣如表 1.7，表中數字為各類影像之張數及占比。



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 張未拍攝至嬰兒臉部而未被偵測之影像。而其中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 1.5），而誤判為類別，此部分之混淆矩陣如表 1.8，表中數字為各類影像之張數及占比。



表 1.7: 實驗影片之姿勢辨識混淆矩陣

		預測分類			
		正躺	趴躺	坐姿	站立
實際分類	正躺	3029 (0.92)	278 (0.08)	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)

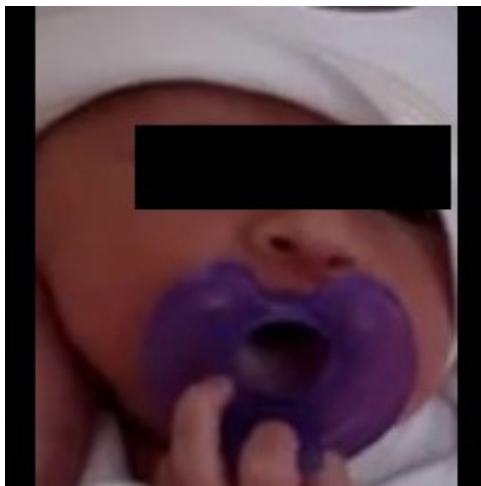


圖 1.5: 臉部遮擋誤判之為警示狀態

表 1.8: 實驗影片之臉部遮擋辨識混淆矩陣

		預測分類		
		無遮蔽	使用奶嘴	有異物
實際分類	無遮蔽	1223 (0.84)	52 (0.04)	184 (0.13)
	使用奶嘴	17 (0.01)	1308 (0.78)	358 (0.21)
	有異物	0 (0.00)	0 (0.00)	129 (1.00)