

國 立 中 央 大 學

資訊工程學系
碩士論文

基於深度學習之嬰兒危險監測系統

A Deep-learning-based Danger Monitoring System
for Infants

研 究 生：王佳君
指 導 教 授：蘇木春 博 士

中 華 民 國 一 百 一 十 一 年 六 月

國立中央大學碩士班研究生
論文指導教授推薦書

資訊工程學系碩士班 學系/研究所 王佳君 研究生
所提之論文 基於深度學習之嬰兒危險監測系統
係由本人指導撰述，同意提付審查。

指導教授 王佳君 (簽章)

111 年 6 月 20 日

基於深度學習之嬰兒危險監測系統

摘要

嬰兒照護者在照顧嬰兒時，可能發生無法隨時關注嬰兒狀態的情形，使得嬰兒因溢奶、翻身、趴睡等情形，致使呼吸不順而發生憾事。又因現有產品多用感測器偵測嬰兒狀態，功能單一且多有使用限制，便利性不佳。

因此，本論文提出基於深度學習技術，專注於嬰兒影像畫面之危險監測系統：首先，使用 SSD 演算法及 RetinaFace 演算法偵測嬰兒臉部；接著，利用 ResNet50 訓練模型，用以辨識嬰兒臉部是否遭非奶嘴之異物遮蔽，以及辨識四種基礎姿勢：正躺、趴躺、坐姿及站立。故當系統輸入嬰兒影片時，可透過模型辨識嬰兒的姿勢可能處於危險狀態或臉部遭異物遮擋，則可即時警示照護者。

本文中，利用 SSD 演算法偵測之時間優勢，偵測每張影像平均僅需 0.04 秒，且準確度達 99%；使用 RetinaFace 演算法偵測嬰兒臉部雖需較長時間，但其正確率、準確度及召回率皆達 99%；而危險偵測之臉部遮擋辨識模型及姿勢辨識模型，其準確度亦皆達 99%。

另外，由於目前未有公開之嬰兒資料集，本文中所使用的嬰兒照片皆為網路真實嬰兒圖片及影片，進行前處理收集而成。

關鍵字：嬰兒危險偵測，嬰兒臉部辨識，深度學習，嬰兒猝死症

A Deep-learning-based Danger Monitoring System for Infants

Abstract

When taking care of the infant, the baby caregiver may not be able to pay attention to the status of the baby at any time, which may cause the infant to suffer from unpleasant breathing due to overflowing milk, turning over, sleeping on the stomach, etc. In addition, the existing products use multiple sensors to detect the state of the infant, which has a single function and many restrictions on use, resulting in poor convenience.

Therefore, this paper proposes a danger detection system based on deep learning technology, focusing on baby images: First, the SSD algorithm and the RetinaFace algorithm is used to detect infant's face. Then, the ResNet50 is used to train the model to analyze whether the infant's face is obscured by foreign objects other than pacifiers, and identify the four basic postures: lying on the back, lying on the stomach, sitting and standing. Therefore, when the system inputs an infant video, the model can recognize that the infant's posture may be in a dangerous state or it's face is covered by a foreign object, , and the caregiver can be immediately alerted.

In this paper, we exploit the time advantage of SSD algorithm detection. On average, it takes only 0.04 seconds to detect each infant image with 99% accuracy. Although it takes a long time to detect an infant's face using the RetinaFace algorithm, its accuracy, precision, and recall rate are all as high as 99%.

The face occlusion recognition model and the posture recognition model for danger monitoring also have an accuracy of 99%.

Additionally, since there are currently no publicly available infant datasets, the infant data used in this research are all captured and preprocessed from online pictures and videos of real infants.

Keywords: Infant danger detection, Infant face detection, Deep learning, Sudden infant death syndrome

目錄

	頁次
摘要	ii
Abstract	iii
目錄	v
一、 緒論	1
1.1 研究動機	1
1.2 研究目的	2
1.3 論文架構	2
二、 相關研究	4
2.1 嬰兒猝死症	4
2.2 嬰兒監測系統	5
2.2.1 感測器偵測	5
2.2.2 影像式偵測	8
2.3 ResNet.....	11
2.4 人臉偵測演算法	13
2.4.1 MTCNN.....	13
2.4.2 RetinaFace.....	15
三、 研究方法	17
3.1 系統流程介紹	17

3.2	臉部遮擋辨識	18
3.2.1	嬰兒臉部偵測	19
3.2.2	嬰兒臉部資料集	19
3.2.3	模型訓練	20
3.3	姿勢辨識	20
3.3.1	嬰兒姿勢資料集	22
3.3.2	模型訓練	24
3.4	危險情境判斷方法	25
四、	實驗設計與結果	26
4.1	臉部偵測準確度實驗	26
4.1.1	實驗目的與設計	26
4.1.2	實驗評估方式	26
4.1.3	實驗結果與分析	26
4.2	臉部偵測執行時間實驗	28
4.2.1	實驗目的與設計	28
4.2.2	實驗評估方式	28
4.2.3	實驗結果與分析	29
4.3	臉部遮擋辨識實驗	29
4.3.1	實驗目的與設計	29
4.3.2	實驗結果分析	30
4.4	姿勢辨識實驗	31
4.4.1	實驗目的與設計	31
4.4.2	實驗結果分析	31
4.5	影片危險偵測實驗	32
4.5.1	實驗目的與設計	32
4.5.2	實驗評估方式	33
4.5.3	實驗結果分析	33

五、 結論與未來展望	36
5.1 結論	36
5.2 未來展望	37
參考文獻	38

圖 目 錄

頁 次

2.1 Linti 嬰用感測背心之穿脫示意圖 [23]	6
2.2 Ferreira 嬰用感測胸帶 [24]	6
2.3 Lin 系統架構圖 [25]	7
2.4 Ziganshin 系統之嬰兒呼吸及心跳圖 [26]	7
2.5 Fang 系統流程圖 [36]	8
2.6 Liu 系統設計圖 [37]	9
2.7 MARRSIDS 模型架構圖 [38]	9
2.8 Wang 系統之網路架構 [39]	10
2.9 Bharati 系統之嬰兒睡眠姿勢分類 [44]	10
2.10 Bharati 系統架構圖 [44]	11
2.11 網路深度與訓練誤差關係 [50]	12
2.12 殘差學習 [50]	12
2.13 在 ImageNet 上訓練之誤差曲線圖 [50]	13
2.14 在 CIFAK-10 上訓練之誤差曲線圖 [50]	13
2.15 MTCNN 架構圖 [51]	14
2.16 MTCNN pipeline [51]	14
2.17 MTCNN 偵測人臉之效果比較 [51]	15
2.18 MTCNN 人臉對齊之效果比較 [51]	15
2.19 RetinaFace 架構圖 [67]	16
2.20 RetinaFace 偵測人臉之效果比較 [67]	16

3.1 系統流程圖	17
3.2 臉部膚色偵測	18
3.3 臉部遮擋辨識流程圖	18
3.4 嬰兒臉部偵測結果	19
3.5 嬰兒臉部資料集	20
3.6 嬰兒平躺之骨架偵測結果	21
3.7 嬰兒趴躺之骨架偵測結果	21
3.8 不同視角之嬰兒骨架偵測結果 [79]	22
3.9 危險動作辨識流程圖	22
3.10 嬰兒腹面朝下之姿	23
3.11 嬰兒姿勢資料集	24
4.1 臉部辨識訓練結果	30
4.2 臉部遮擋辨識模型之混淆矩陣	30
4.3 姿勢辨識訓練結果	31
4.4 姿勢辨識模型之混淆矩陣	32
4.5 姿勢辨識錯誤之影像：坐姿誤判為趴躺	32
4.6 姿勢辨識錯誤之影像：正躺誤判為趴躺	33
4.7 實驗影片之姿勢辨識混淆矩陣	34
4.8 未偵測嬰兒臉部之影像	34
4.9 臉部遮擋誤判之為警示狀態	35
4.10 實驗影片之臉部遮擋辨識混淆矩陣	35

表 目 錄

頁次

4.1	OpenCV [80] 偵測嬰兒臉部結果	27
4.2	SSD [77] 偵測嬰兒臉部結果	27
4.3	MTCNN [51] 偵測嬰兒臉部結果	28
4.4	RetinaFace [67] 偵測嬰兒臉部結果	28

一、緒論

1.1 研究動機

根據衛生福利部統計處所發布的嬰兒主要死因統計 [1] 中，101年至105年間每年至少30位嬰兒死於嬰兒猝死症候群（Sudden infant death syndrome，簡稱 SIDS），106年至109年每年亦仍有超過20位嬰兒因此症狀逝世，是為嬰兒十大死亡原因之一。

三軍總醫院對於嬰兒猝死症的說明為：一個原本無異狀的嬰兒，突然且無法預期的死亡，常發生在嬰兒睡眠時，並在事後的屍體解剖檢查中找不到其真正致死原因。凡未滿一歲的嬰幼兒皆可能發生，其中二至四個月時期尤為常見，亦可能發生在嬰兒出生一至兩周內。醫界雖持續探討嬰兒猝死症的發生原因，但目前對於真正的成因仍不清楚，綜合醫界當前相關因素的研究中，包含了嬰兒因溢奶或嘔吐產生呼吸道緊縮反射及憋氣，或因翻身、趴睡致使呼吸困難，而窒息死亡等原因。

當照護者視線離開嬰兒，如泡奶、如廁等情境，無法即時關注嬰兒狀態。若此時嬰兒發生溢奶、物品遮蓋口鼻、自行翻身或站立等情形，將使嬰兒處於危險情境中，進而導致憾事發生。

國內外有許多為自動化監測嬰兒狀態之研究，主要包含兩種偵測方式：其一為使用感測器量測嬰兒之特定生理訊號，如：心率、呼吸頻率、體溫、身體位置或方向及嬰兒周圍之氣體濃度等，透過收集到的數值以判定所監測之嬰兒處於正常狀態與否；然而，使用此種監測方式具功能單一性，若欲偵測其他生理訊號，則需增設更多不同種類的感測器，不

僅可能影響嬰兒之活動，亦可能產生更多潛在的危險性，如：裝置纏繞嬰兒、孩童誤食裝置等。其二為透過電腦視覺偵測嬰兒影像，判定嬰兒是否處於危險狀態，而現有研究中多僅針對嬰兒之面部特徵或單一狀態進行偵測；然而，我們認為一張嬰兒影像包含了許多資訊得以應用，如：同時偵測嬰兒面部及姿勢等，則可透過影像進行更廣泛的嬰兒危險情境之監測。

因此，本論文透過嬰兒影像辨識其基礎姿勢與面部狀態，以監測嬰兒是否因姿勢不適當或面部遭異物遮擋，處於危險情境中而需警示照護者。此方法不僅可監測多種危險情境，也避免感測器影響嬰孩的問題發生，且對於增加其他監測功能有良好的擴充性。

1.2 研究目的

本論文基於深度學習技術，利用 ResNet50 進行嬰兒動作及臉部遮擋之模型訓練，且以 RetinaFace 等多種臉部偵測演算法進行嬰兒影像前處理以擷取其臉部畫面，對目標嬰兒進行危險監測。

本研究預計達成以下目標：

- 針對嬰兒姿勢部分，辨識嬰兒之正躺、趴睡、坐姿及站立之四項基礎姿勢，判斷嬰兒是否做出具危險之動作。
- 針對嬰兒臉部部分，判斷嬰兒是否因嘔吐物、毛巾等非奶嘴之外物遮蓋其面部，而可能使嬰兒發生窒息危險。

綜上目標，本論文將建構出一可對嬰兒姿勢及臉部遮擋進行危險監測之系統。

1.3 論文架構

本論文分為五個章節，其架構如下：

第一章、緒論，敘述本論文之研究動機、研究目的及論文架構。

第二章、相關研究，敘述嬰兒猝死症之定義，並探討近年與嬰兒監測相關之研究以及深度學習模型架構與面部辨識網路。

第三章、研究方法，說明本研究之詳細內容，包含完整系統之流程說明以及資料集之分類定義及前處理等。

第四章、實驗設計與結果，說明各項實驗設計內容以及評估方法，並對於實驗結果進行探討。

第五章、結論與未來展望，對於研究結果進行總結，並討論研究的未來展望。

二、相關研究

2.1 嬰兒猝死症

嬰兒猝死症（The Sudden Infant Death Syndrome, 簡稱 SIDS）[2] 其特徵為一位看似健康的嬰兒在睡眠期間突然死亡，其真正致死之原因尚不明確且非單一。

目前醫界對嬰兒猝死症之直接致死原因尚未有統一的定義，但可統整出多項促使此症發生之風險因素，主要可分為兩類：其一為外在因素，包含嬰兒因俯臥及側睡姿勢、蓋住臉部的床單、嬰兒睡在沙發或其他容易陷入的柔軟家具上等 [3]–[7]，致使嬰兒呼吸困難而死亡；其二則為內在因素，包含成長因素（如：早產 [8]）、遺傳因素（如：家族性之嬰兒猝死症 [9], [10]）、性別（男性比例為女性的兩倍）或種族 [11] 等。除此之外，嬰兒也可能因其他外在環境條件，如：產前或產後暴露於不良物質中（如：香菸煙霧、酒精或非法藥物等），而弱化嬰兒之內在條件。

在嬰兒猝死症研究中，有許多關於此症之死亡機制理論，其中心肺控制假說主導了多數研究，也造就了往後關於此症之探討多基於嬰兒呼吸或自主神經機制的缺陷。此論點主要包含了五個步驟 [2]：(1) 發生危及生命的事件（如：面部朝下 [12] 或面部遭遮蔽 [13]，造成反射性或阻塞性呼吸暫停），而將導致嬰兒窒息或腦部灌注不足，亦可能兩者皆發生。(2) 嬰兒無法自行轉頭，以應對窒息的情境，而導致無法從呼吸暫停中恢復。(3) 持續性窒息導致嬰兒失去意識或反射，即低氧昏迷。(4) 發生心率過緩及缺氧喘氣，此現象在嬰兒因嬰兒猝死症逝世前將明顯發

生。(5) 嬰兒的自主復甦能力受損，即因無效的喘氣而最終導致呼吸暫停及死亡。因此，由嬰兒猝死症之紀錄中，可看出此症並非一種突發疾病，而是在嬰兒死亡前，即會出現心率不正常或呼吸暫停之惡性循環現象。

另外，亦有研究人員使用 Triple-Risk Model 來解釋嬰兒猝死症 [14]，即嬰兒死於此症需同時包含以下三個因素：(1) 有風險的嬰兒：嬰兒含有一個未知的問題，可能是基因突變或腦部缺陷等，這使其面臨了嬰兒猝死症的風險；(2) 嬰兒發育的重要時期：嬰兒出生後的前六個月，將經歷許多快速成長的階段，此階段會改變身體控制和調節自身的能力，且嬰兒的身體也會於此時學習如何對環境做出應變；(3) 環境中的壓力源 [15]：即前述中之外在因素，包含嬰兒睡姿及接觸香菸煙霧等。若僅發生其中一項因素，將不足以導致嬰兒猝死症引發死亡。

因此，若能消除環境中的壓力源，將有利於嬰兒的生存。同時醫界亦發現俯臥睡姿勢將使嬰兒猝死症風險增加三倍以上 [3]，故在 1990 年代初期國際間即提倡嬰兒仰臥睡姿，此症之發病率也因此降低了 50% 以上，但仍為嬰兒主要死亡原因之一。

2.2 嬰兒監測系統

在照護嬰兒的過程中，由於嬰兒尚未發展出語言能力表達自己的不適，或尚無能力將自己避免於危險之外。因此，為了協助照顧者關注嬰兒狀態，現有許多為自動化監測嬰兒之研究，主要分為以感測器偵測生理訊號及以影像式偵測兩種方式。

2.2.1 感測器偵測

此種方式利用呼吸感測器、濕度感測器、溫度感測器、慣性感測器、三軸加速度計、一氧化碳感測器及二氧化碳感測器等多種不同設備，進

行嬰兒呼吸頻率、出汗狀況、體溫、心率、身體位置或方向、睡眠姿勢、周圍一氧化碳濃度及呼出的二氧化碳濃度變化等訊號之偵測，且多會透過物聯網技術開發出可穿戴式裝置 [16]–[25]，以下介紹幾項基於感測器式偵測嬰兒之研究。

Linti 等人 [23] 將多個感官元件融入紡織品以開發嬰兒感測背心，此裝置可用來量測嬰兒之呼吸、心率、溫度及濕度，其示意圖見圖 2.1。



圖 2.1: Linti 嬰用感測背心之穿脫示意圖 [23]

Ferreira 等人 [24] 將心律感測器、3D 加速度計、熱電堆感測器裝設於胸帶中（圖 2.2），以量測嬰兒之體溫、心率、呼吸頻率及身體位置，再透過 ZigBee 技術傳送數據至伺服器，並可由網頁查收訊息。

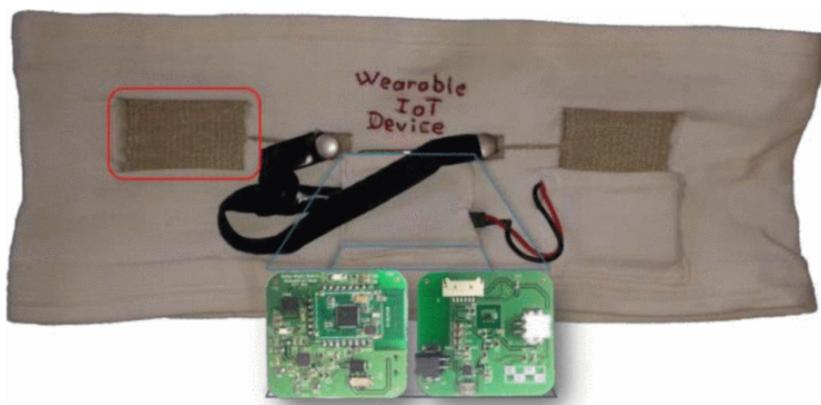


圖 2.2: Ferreira 嬰用感測胸帶 [24]

Lin 等人 [25] 於胸帶中嵌入三軸加速度計、溫度感測器及一氧化碳感測器，可用來量測嬰兒睡姿（面朝上、下、左或右）、體溫及其周圍一氧化碳濃度，且利用三軸加速度計 z 軸資訊可獲得嬰兒之呼吸頻率，最終驗證計算之呼吸頻率準確率達 100%。此系統架構圖見圖 2.3，

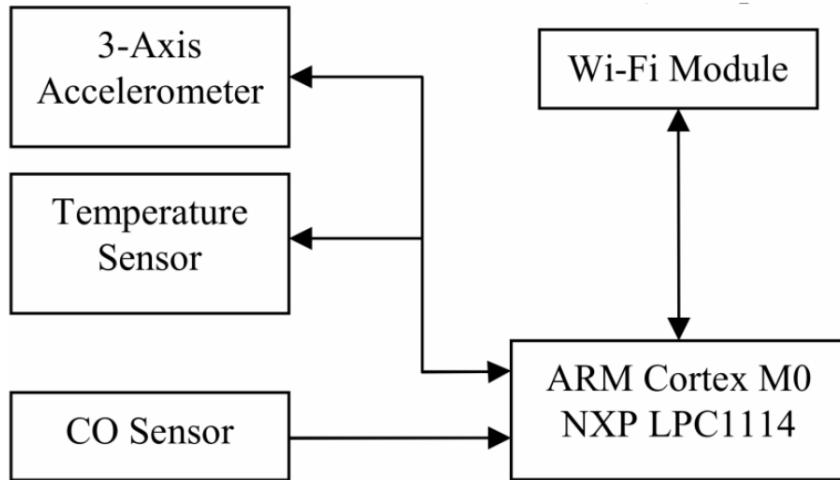


圖 2.3: Lin 系統架構圖 [25]

Ziganshin 等人 [26] 基於超寬頻技術 [27] 開發出可監測呼吸及心率之系統（如圖 2.4），以檢測嬰兒三種狀態：睡眠（無運動）、清醒（長期大幅度運動）及警示（偵測到呼吸頻率異常低或暫停）。

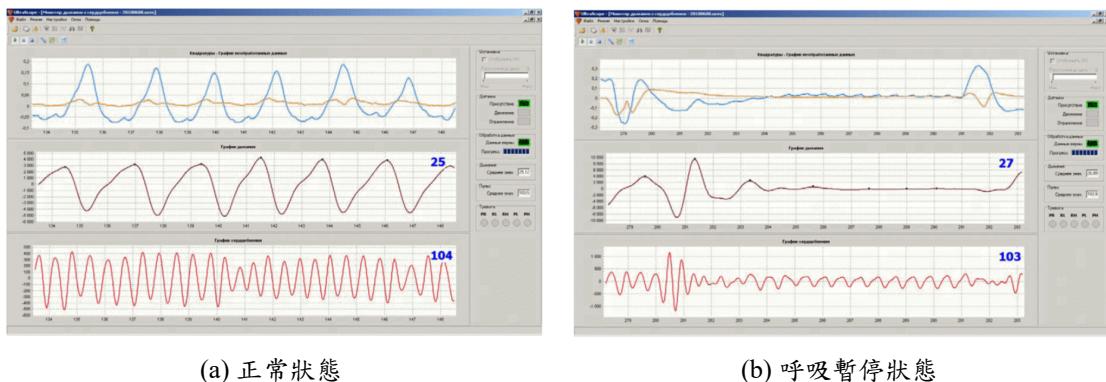


圖 2.4: Ziganshin 系統之嬰兒呼吸及心跳圖 [26]

此種利用感測器監測嬰兒的方法，雖然可直接量測其生理訊號以判斷狀態正常與否，但仍可能因硬體設備之缺陷無法準確量測，進而有失判斷準確性，亦或者因需額外穿戴裝置而造成嬰兒不適，進而影響活動或導致更多危險的發生。

2.2.2 影像式偵測

現有研究中，基於電腦視覺技術之監測系統雖日漸廣泛，但多針對小孩、成人或老人之照護進行開發 [28]–[31]，而少數應用於嬰兒的偵測系統中，又多關注於呼吸頻率、面部特徵及趴睡姿勢等，以下介紹幾項基於影像式偵測嬰兒之研究。

有別於往的接觸式或非視覺技術之呼吸頻率偵測系統 [32]–[35]，Fang 等人 [36] 基於視覺技術進行非接觸式之呼吸頻率偵測。此系統完整流程如圖 2.5a：先進行嬰兒運動偵測，其判斷流程見圖 2.5b，包含頭部、四肢及身體運動，但不包含因呼吸引起的輕微運動；若系統未偵測到運動情形，則透過空間特徵擷取候選呼吸點，再利用模糊積分技術選擇呼吸點，即可計算呼吸頻率，進而判斷嬰兒呼吸是否異常。

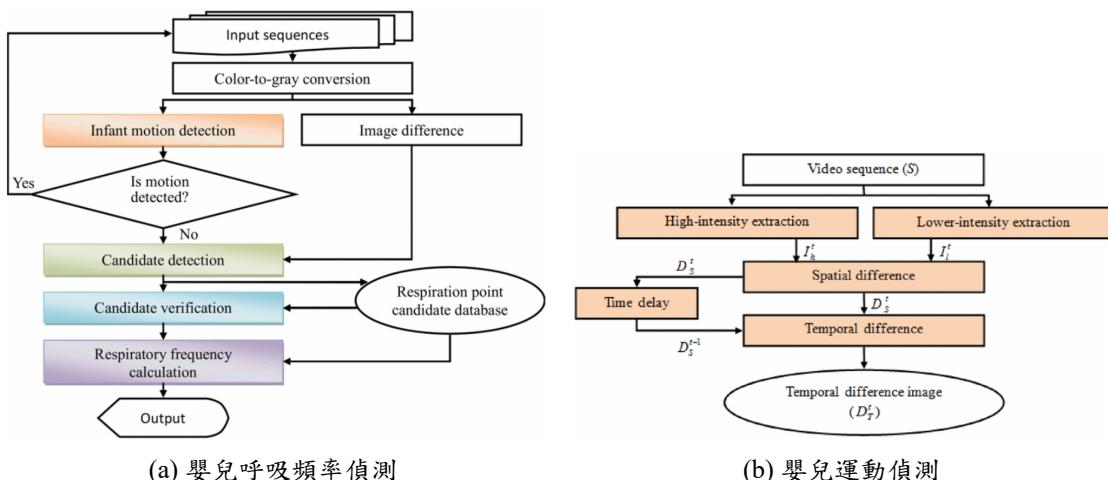


圖 2.5: Fang 系統流程圖 [36]

Liu 等人 [37] 利用夜視攝影機拍攝在嬰兒床內的嬰兒，判斷其呼吸頻率異常與否。此系統主要包含三個部分：(1) 影片收集：為圖 2.6 中的前兩項，透過夜視攝影機連接到 Artik 板以收集嬰兒影片；(2) 呼吸偵測演算法：為圖 2.6 中的第三及四項，包含了使用 Eulerian Magnification 技術放大影片中的細微運動，以監測拍攝對象之胸部運動，而若經正規化之像素差異值低於設定閾值，則判斷為呼吸頻率異常；(3) 警示：為圖 2.6 中的最後兩項，透過 Twilio 在演算法偵測到緊急狀況時，向父母手

機發出警報。

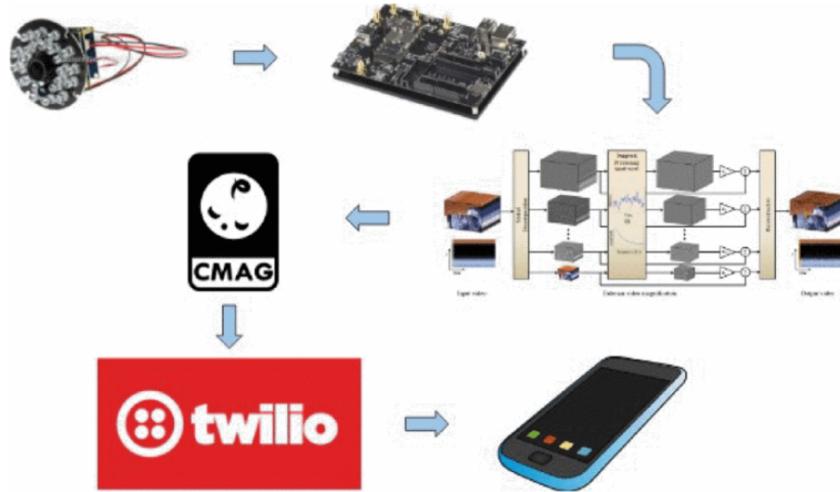


圖 2.6: Liu 系統設計圖 [37]

Gallo 等人 [38] 提出名為 MARRSIDS 的模型（架構如圖 2.7），其利用 OpenCV 之 Haar-Like Features 偵測嬰兒面部特徵，若未偵測到其臉部，則認為嬰兒可能位於不良姿勢，需發出警示；而若偵測到嬰兒臉部且為睜眼狀態，則代表其處於清醒狀態，並非處於風險中。

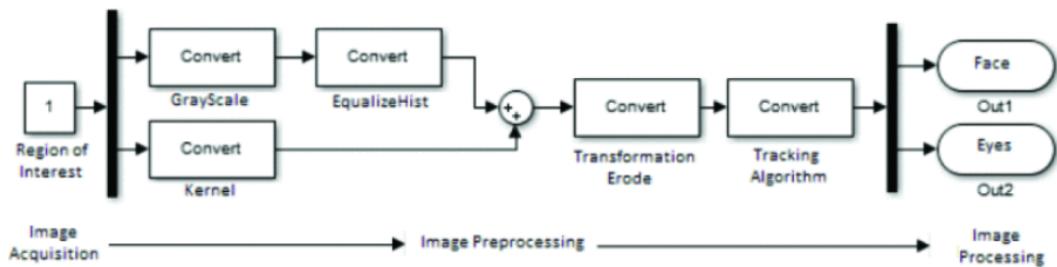


圖 2.7: MARRSIDS 模型架構圖 [38]

Wang 等人 [39] 提出用以偵測嬰兒臉部遮擋的貝氏深度神經網路架構（圖 2.8），共包含四項子任務：(1) 眼睛、鼻子或嘴巴是否可見，(2) 不可見的原因是否為被外物（如：枕頭）遮擋，(3) 眼睛睜開與否，及(4) 五個臉部座標之位置。此框架使用了損失函數進行訓練，並考慮不同偵測任務間的不確定性 [40]，且有別於現有不專注於嬰兒影像的研究 [41]–[43]，此方法使用 MobileNetV2 針對自行收集之 YunInfants 資料集進行嬰兒頭部影像分析。

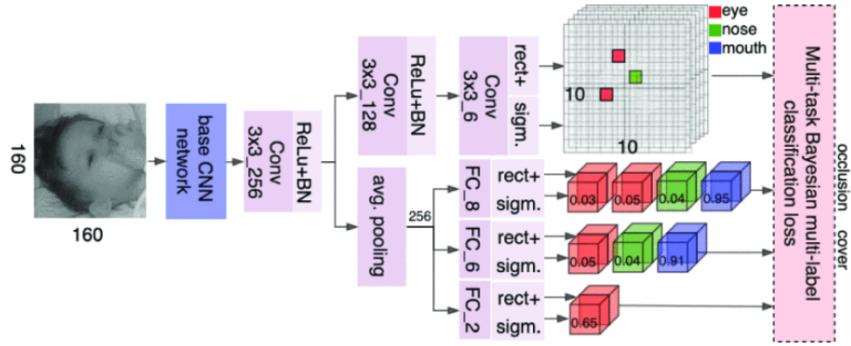


圖 2.8: Wang 系統之網路架構 [39]

Bharati 等人 [44] 開發基於卷積神經網絡之嬰兒睡眠姿勢電腦視覺系統，以評估仰臥（正常狀態）、從仰臥轉換到趴臥（警報狀態）及趴臥（危險狀態）三種姿勢（如圖 2.9），並於嬰兒呈現趴臥時，透過手機提醒照護人員。此系統完整架構見圖 2.10a，其中卷積神經網路部分之架構如圖 2.10b：首先，輸入經轉換的 2D 嬰兒灰階影像；接著，經過多次的卷積層及最大池化層；再傳入全連接層後，最終輸出嬰兒的三種睡眠姿勢機率值。由於目前未有公開之嬰兒姿勢資料集，故此研究所使用的影像資料為和真實嬰兒相同比例之娃娃。



圖 2.9: Bharati 系統之嬰兒睡眠姿勢分類 [44]

現有研究中，視覺化之嬰兒監測系統多關注其呼吸運動、面部特徵或單一姿勢，而未有辨識嬰兒常見動作之研究，故我們提出一可偵測嬰兒基礎姿勢及面部遮擋之危險監測系統。

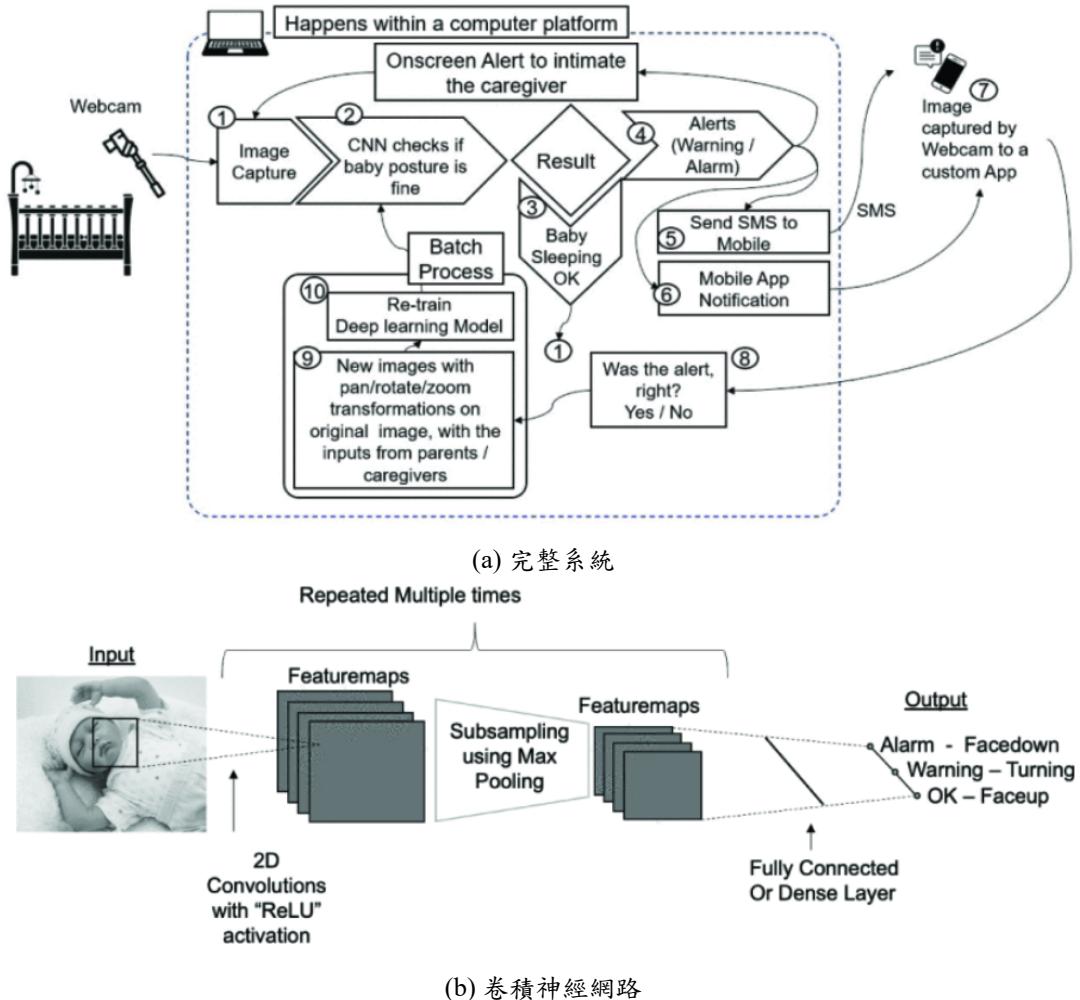


圖 2.10: Bharati 系統架構圖 [44]

2.3 ResNet

既有研究中 [45]–[47]，可知卷積神經網路其深度至關重要，在 ImageNet 資料集中結果表現良好的網路皆為十六至三十層之深層網路。

然而，當訓練更深層的神經網路時，卻會出現退化問題，亦即隨著網路深度的增加，準確率達飽和後，反而迅速下降，而這樣的結果並非因過度擬合所致 [48], [49]，如圖 2.11 可看到兩個不同層數的網路其訓練及測試誤差值。

因此，He 等人 [50] 提出使用深度殘差學習 (圖 2.12) 的網路架構，利用 shortcut connection 執行 identity mapping，且不需要增加額外的參數，亦即不增加計算複雜度。

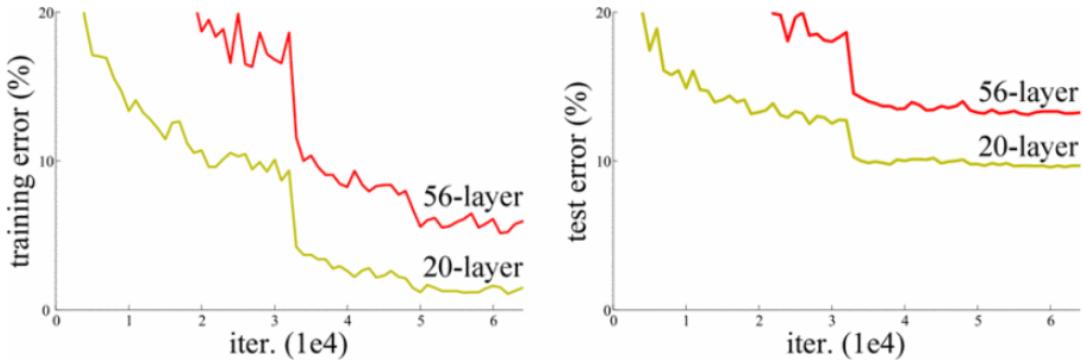


圖 2.11: 網路深度與訓練誤差關係 [50]

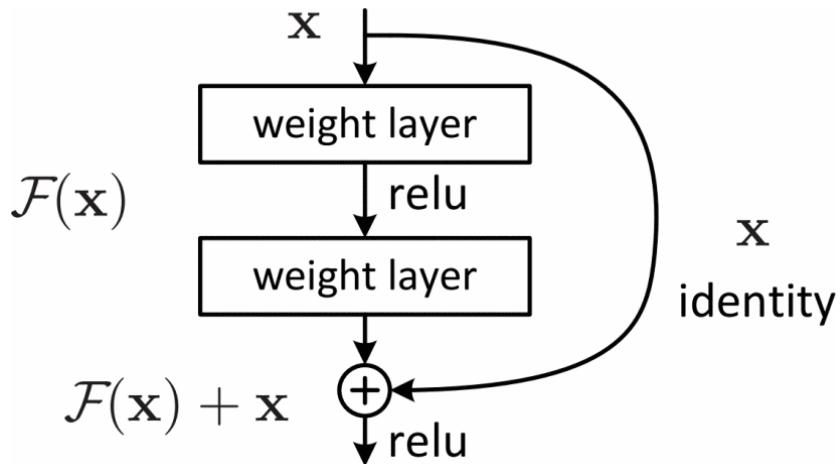


圖 2.12: 殘差學習 [50]

此研究中，通過訓練 ImageNet 評估不同層數之普通網路與殘差網路，由圖 2.13（左）可觀察到普通網路中 34 層卻比 18 層有更高的驗證誤差，而圖 2.13（右）則可看出殘差網路中 34 層相對於 18 層有較低的訓練誤差，此結果說明退化問題獲得了解決。而文中亦透過訓練 CIFAK-10 進行更多的比較，圖 2.14（左）為普通網路不同層數之誤差曲線圖，圖 2.14（中）為殘差網路不同層數其誤差曲線圖，圖 2.14（右）則為使用 110 層及 1202 層殘差網路之誤差曲線圖。

最終，本研究以 152 層之殘差網路在 ILSVRC 2015 中獲得第一名，此網路比 VGG 網路深八倍，卻仍擁有較低的複雜度。

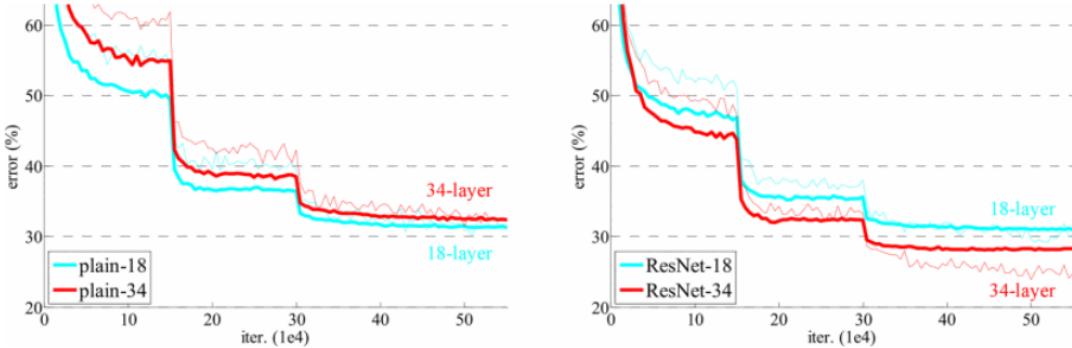


圖 2.13: 在 ImageNet 上訓練之誤差曲線圖 [50]

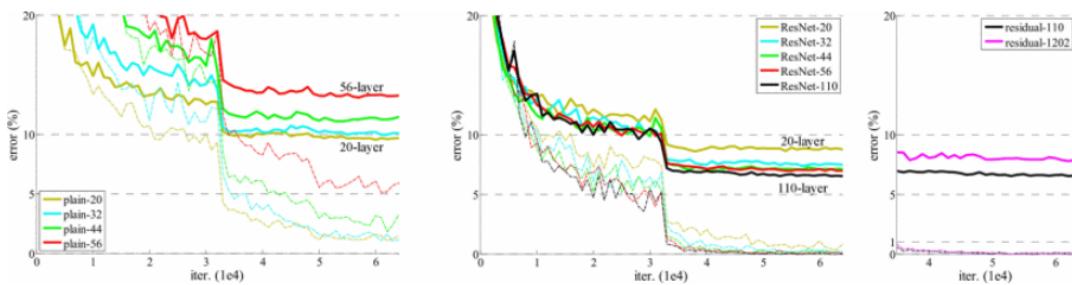


圖 2.14: 在 CIFAR-10 上訓練之誤差曲線圖 [50]

2.4 人臉偵測演算法

2.4.1 MTCNN

MTCNN [51] 是由 Zhang 等人於 2016 年提出的一種多任務級聯卷積神經網路，可以同時處理人臉偵測及對齊任務，並提出 online hard sample mining 策略以提升效能。

此演算法為包含三階段級聯架構的深度卷積網路，架構如圖 2.15：第一階段為全卷積網路構成之 proposal network (P-Net)，用來獲得人臉區域的候選窗口及其邊界框回歸向量，根據此估計回歸向量校準候選者，再以非極大值抑制合併高度重疊的候選者；而第二階段，所有候選者皆饋送至另一個稱為 refine network (R-Net) 的卷積神經網路，以進一步拒絕大量錯誤候選者，並使用邊界框回歸進行校準及非極大值抑制；第三階段中，則利用 output network (O-Net) 輸出五個臉部的座標位置，其類似於第二階段，但目標為識別受更多監督的人臉區域。由圖 2.16，可清

楚看出此網路是以粗到細的方式預測人臉及座標位置。

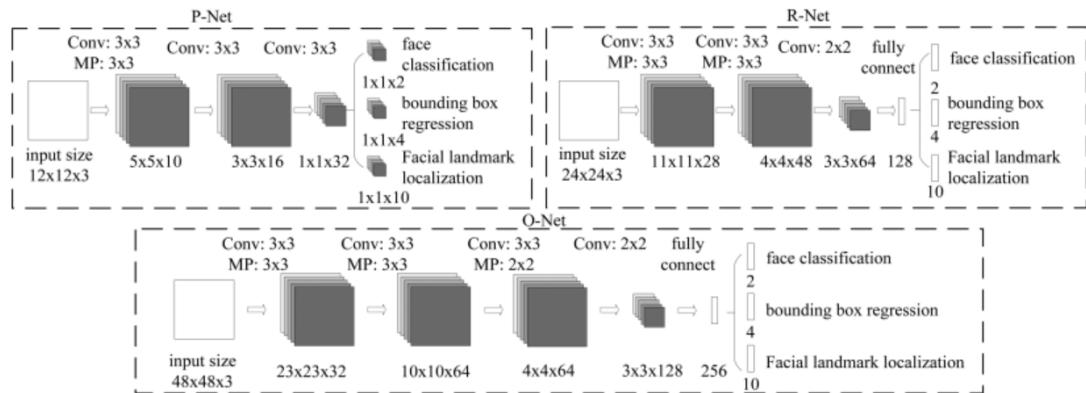


圖 2.15: MTCNN 架構圖 [51]

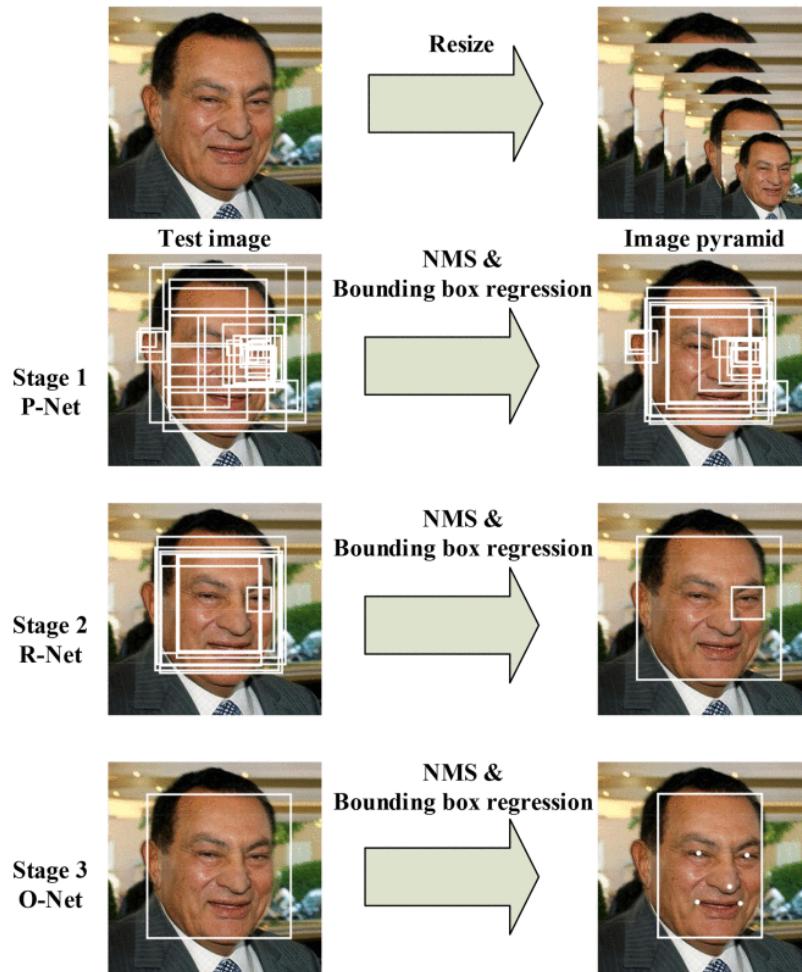


圖 2.16: MTCNN pipeline [51]

為了評估人臉偵測效果，此網路與 FDDB 及 WIDER FACE 中的其他方法 [52]–[60] 進行比較，結果如圖 2.17，可看出此演算法確實優於其他

方法。而為了評估人臉對齊效果，此網路亦與多種方法 [61]–[66] 進行比較，結果如圖 2.18，可看出此演算法雖在嘴角定位較不具優勢，但其他部分仍優於多數方法。

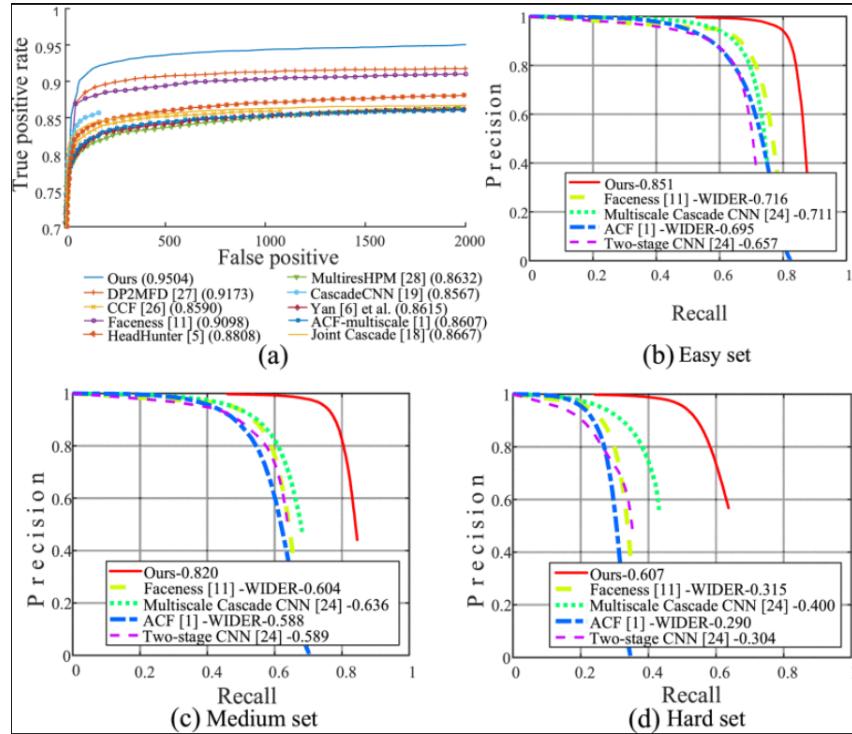


圖 2.17: MTCNN 偵測人臉之效果比較 [51]

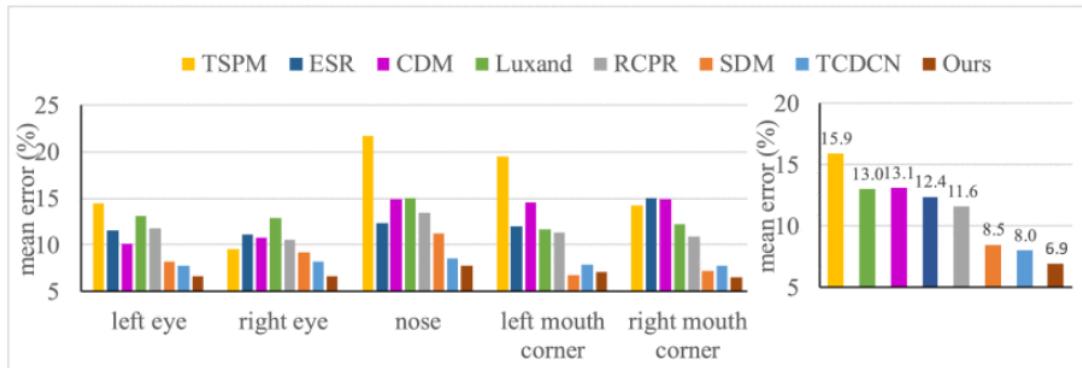


圖 2.18: MTCNN 人臉對齊之效果比較 [51]

2.4.2 RetinaFace

由 Deng 等人於 2020 年提出的 RetinaFace [67]，為一種 single-shot、multi-level 人臉定位方法，其基於影像平面之點回歸整合了人臉框預測、

2D 人臉標示定位及 3D 頂點回歸。

此模型架構（見圖 2.19）中，主要由三個部分組成：(1) feature pyramid network、(2) context head module 及 (3) cascade multi-task loss。首先，由 feature pyramid network 輸入影像，並輸出五個不同比例的特徵圖；接著，context head module 獲得這些特徵圖以計算多任務的損失。亦即第一個模組會從一般的 anchor 預測範圍框，而後第二個模組利用第一個模組迴歸出的 anchor 以預測更精準的範圍框。

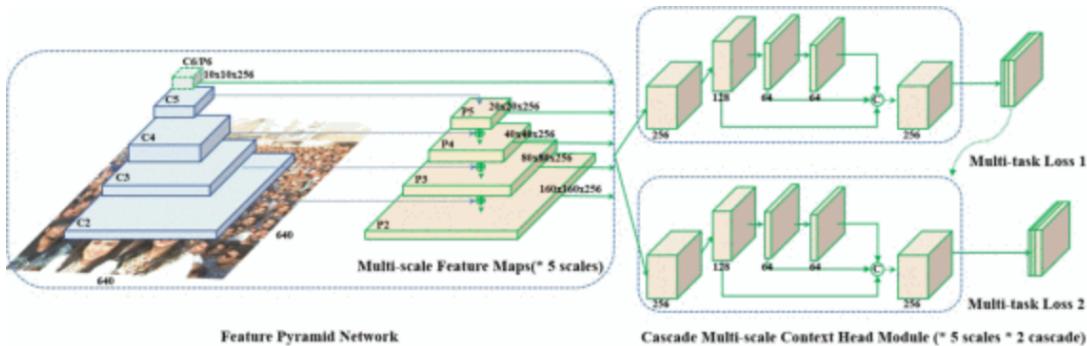


圖 2.19: RetinaFace 架構圖 [67]

本篇論文展示了 RetinaFace 和其他 29 種人臉偵測演算法 [68]–[72] 平均準確度之比較，如圖 2.20 所示，此演算法擁有 91.7% 的良好結果。

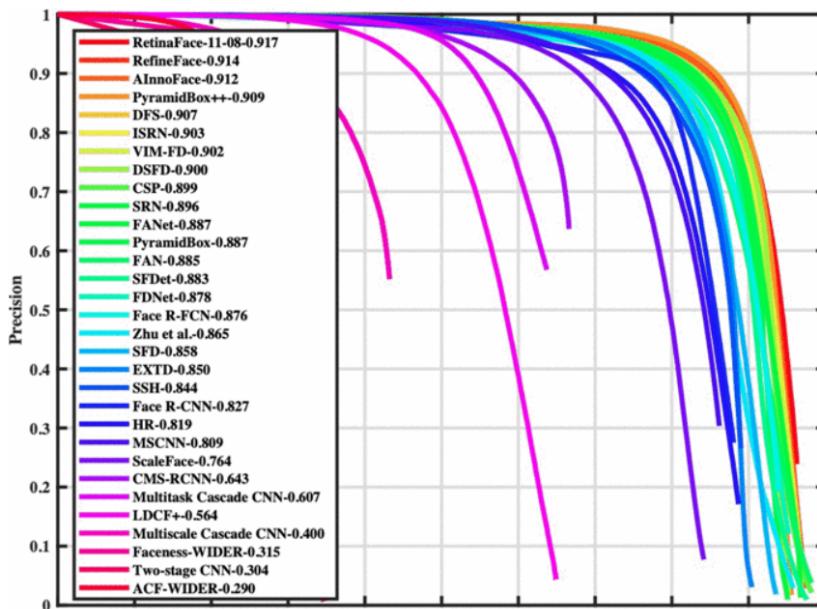


圖 2.20: RetinaFace 偵測人臉之效果比較 [67]

三、研究方法

3.1 系統流程介紹

本論文所開發之嬰兒危險監測系統，針對嬰兒影像畫面進行辨識，以判斷其是否處於危險狀態，而須提醒照護者。

系統之完整流程為：首先，讀取一段待觀測之嬰兒影片，將影片切成數幀影像，並判斷影像存在與否，若不存在將發出異常警告，反之則開始對該影像畫面進行危險偵測判斷。針對每幀嬰兒影像，系統對其臉部遮擋及姿勢進行辨識，若透過模型分析為警示狀態，則再經後續步驟判斷是否提醒照護者；而若分析為安全狀態，則可接續下一幀之影像進行偵測。系統完整流程圖，請見圖 3.1。

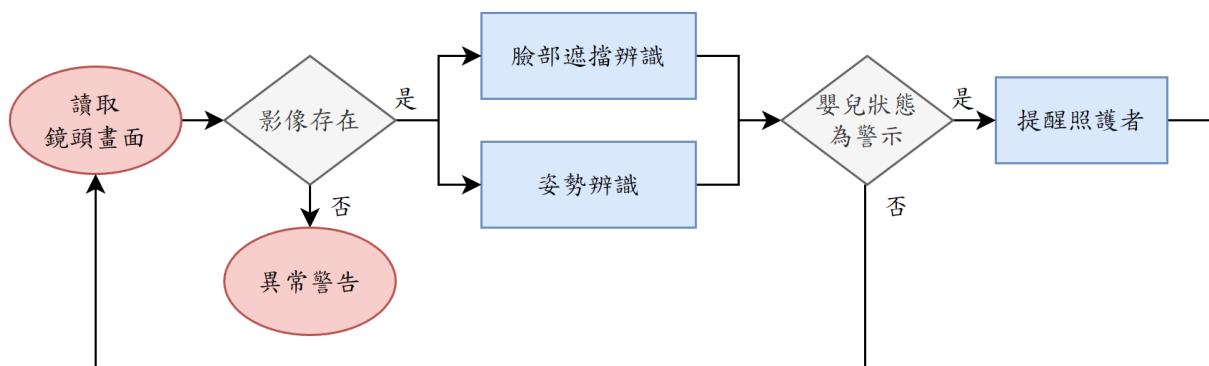


圖 3.1: 系統流程圖

本系統中，包含兩項危險辨識模型：(1) 嬰兒臉部遮擋辨識：先將嬰兒畫面擷取出僅含臉部範圍之影像，再透過此模型判斷其臉部是否遭非奶嘴之異物遮蔽，若是，則嬰兒為警示狀態；(2) 嬰兒危險動作辨識：將拍攝之嬰兒全身影像透過此模型進行辨識，判斷嬰兒為正躺或坐姿之安

全狀態，或為需警示的趴躺及站立姿勢。而若兩模型結果皆為安全，則系統會判斷嬰兒狀態為安全，否則，嬰兒狀態則為警示。此二部分辨識之詳細方法，將於 3.2 節及 3.3 節進行介紹。

3.2 臉部遮擋辨識

如前言所述，目前醫界對於嬰兒猝死症之相關因素研究提出，注意嬰兒臉部是否遭遮蔽，將有助於降低此症的發生；另亦有研究發現嬰兒使用奶嘴，對於預防嬰兒猝死症有幫助。因此，本文對於嬰兒臉部遮擋辨識將排除使用奶嘴之情境。

起初，基於電腦視覺及影像處理技術，例如：利用 Cb , Cr 色彩空間及 ellipse clustering [73] [74] [75] [76] 等偵測膚色，判斷嬰兒臉部是否出現非膚色之區塊，以進行臉部遮擋辨識，其效果如圖 3.2。

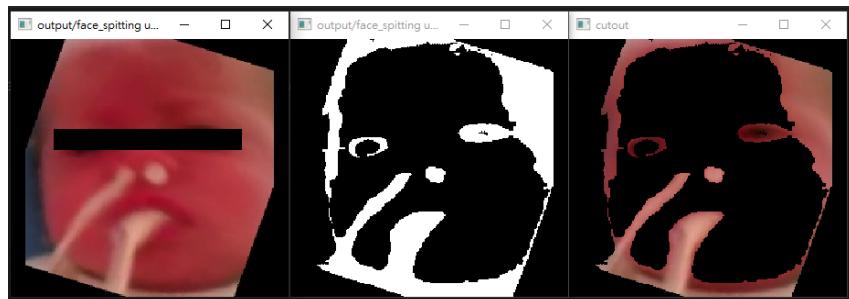


圖 3.2: 臉部膚色偵測 [76]

而後考量能有較佳的推廣性，因此，本研究改為使用深度學習技術進行臉部遮擋辨識，針對嬰兒面部影像收集資料，以訓練可辨識三種嬰兒臉部狀態之模型。本部分之流程圖，請見圖 3.3。

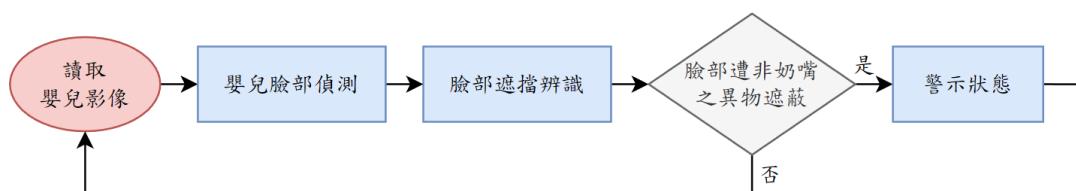


圖 3.3: 臉部遮擋辨識流程圖

3.2.1 嬰兒臉部偵測

嬰兒臉部遮擋辨識僅需關注臉部畫面，故本文會先透過人臉偵測演算法進行前處理，以獲得只涵蓋嬰兒面部範圍之影像。經多方實驗現有人臉偵測演算法後，同時考量嬰兒臉部偵測之正確率及執行時間，最終本研究選用 RetinaFace [67] 及 SSD [77] 等演算法進行嬰兒臉部偵測，其效果如圖 3.4。



圖 3.4: 嬰兒臉部偵測結果

3.2.2 嬰兒臉部資料集

本論文將嬰兒臉部狀態分為三類，各類定義如下：

1. 臉部無遮蔽：嬰兒五官皆未被遮擋，為安全狀態，如圖 3.5a。
2. 臘部遮蔽物為奶嘴：嬰兒正在使用奶嘴，為安全狀態，如圖 3.5b。
3. 臘部遮蔽物非奶嘴：嬰兒臉部因溢奶遭嘔吐物遮蔽，或被毛巾等其他外物遮蓋，而可能造成窒息危險，為警示狀態，如圖 3.5c。

本資料集包含嬰兒之正臉及側臉影像，總共 3475 張照片。我們將所有影像分為訓練、測試及驗證集，各部分占比為 70%、20% 及 10%，即各有 2436 張、697 張及 342 張影像。

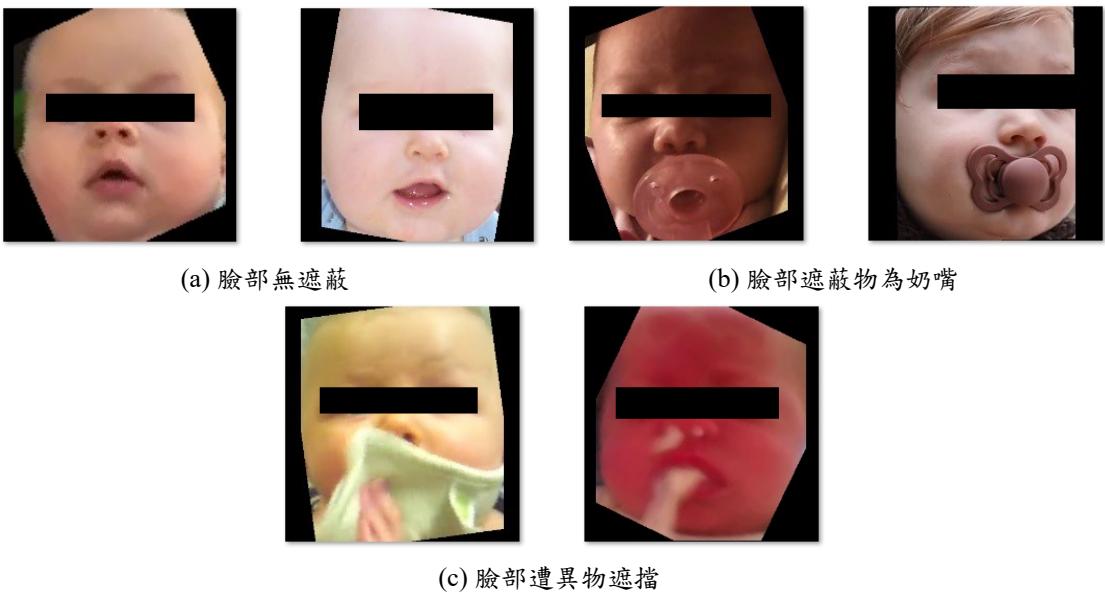


圖 3.5: 嬰兒臉部資料集

3.2.3 模型訓練

本論文使用 3.2.2 節之嬰兒臉部資料集，以 ResNet50 [50] 進行臉部遮擋辨識模型之訓練，最終達成辨識三種嬰兒臉部狀態：安全、使用奶嘴及警示。

3.3 姿勢辨識

承前言所述，除了臉部遮蔽可能造成嬰兒猝死症外，嬰兒做出不適當的動作也常為意外發生原因，例如：嬰兒側躺或趴睡時，因頸部肌肉較弱等原因，無力自行將臉移開，造成呼吸困難而窒息死亡；或者當嬰兒自行站立，而有可能爬落嬰兒床等，亦可能使嬰兒處於危險情境中。

在現有成人動作辨識的研究中，多會以骨架偵測開始著手。我們使用 OpenPose [78] 及 MediaPipe Pose [79] 等演算法進行嬰兒骨架之偵測，發現嬰兒需在特定情境下才能有較佳的偵測效果，結果如圖 3.6 及圖 3.7，可看到嬰兒平躺之偵測效果尚可，但趴躺等其他姿勢之效果則不如預期。又因本研究目標為能從非限定視角辨識嬰兒動作，然而嬰兒骨架圖在俯

視角與平視角中多有相似之處（如圖 3.8），而無法達到良好的應用結果。

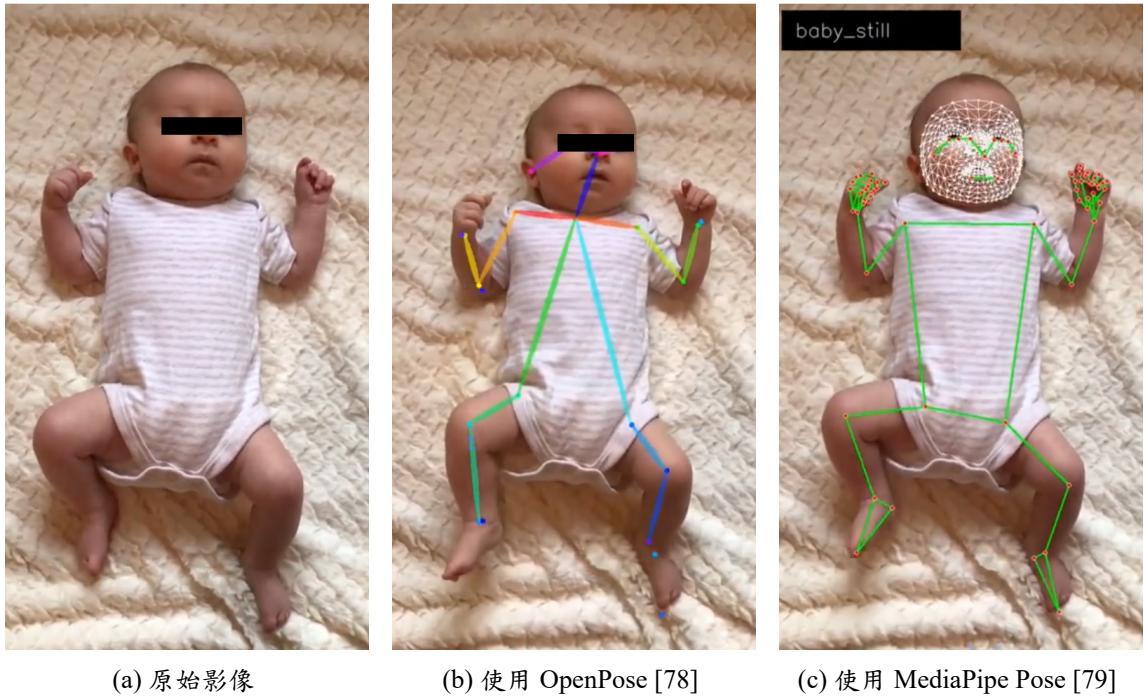
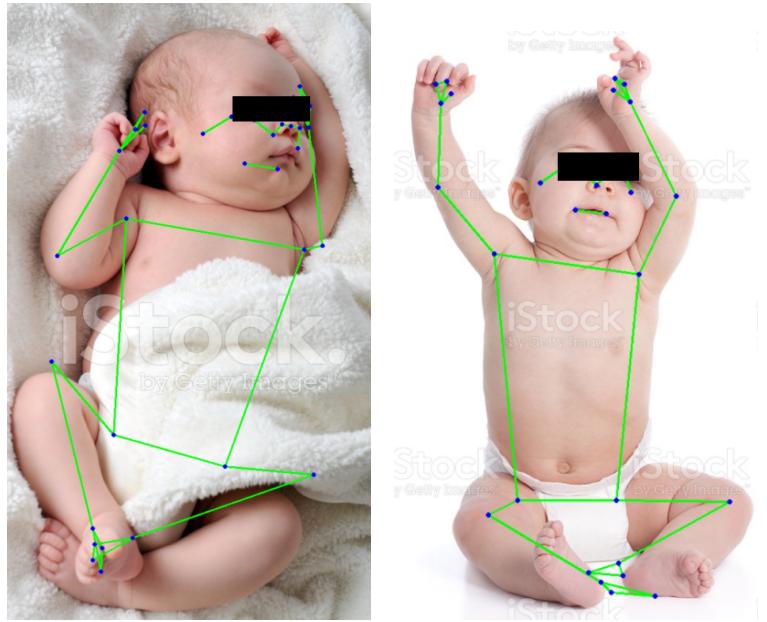


圖 3.6: 嬰兒平躺之骨架偵測結果



圖 3.7: 嬰兒趴躺之骨架偵測結果

因此，本文最終使用深度學習技術進行嬰兒動作辨識，使用自行收集之嬰兒影像資料集（於 3.3.1 節進行詳細介紹），訓練可辨識四種嬰兒基礎姿勢之模型。本部分之流程圖，請見圖 3.9。



(a) 俯視嬰兒躺姿

(b) 平視嬰兒坐姿

圖 3.8: 不同視角之嬰兒骨架偵測結果 [79]

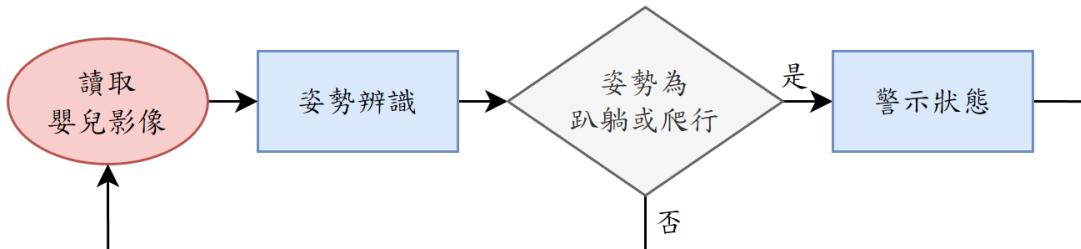


圖 3.9: 危險動作辨識流程圖

3.3.1 嬰兒姿勢資料集

起初，將嬰兒姿勢分為五類：正躺、趴睡、爬行、坐姿及站立，而趴睡及爬行二類時常發生互相誤判，致使辨識錯誤率高。我們推測因為此二類嬰兒皆呈現腹面朝下之姿，不同處在於四肢及軀體是否貼地，如圖 3.10 所示，但若接續細分姿勢，將導致動作分類過細。

因此，最終本論文將嬰兒姿勢分成基礎四類，包含正躺（腹面朝上）、趴躺（腹面朝下）、坐姿及站立，以供辨識嬰兒大部分之姿。對於此四類姿勢之詳細分類定義為：

1. 正躺：嬰兒腹部面朝上，背部貼於水平面，而頭部及四肢位置不



圖 3.10: 嬰兒腹面朝下之姿

限，如圖 3.11a。

2. 趴躺：嬰兒腹部面朝下，包含趴睡及爬行等多動作，而頭部及四肢位置不限，如圖 3.11b。
3. 坐姿：嬰兒臀部貼於水平面，且背部未貼於同一平面，而頭部及四肢位置不限，如圖 3.11c。
4. 站立：嬰兒腳掌貼於水平面，且腹部和背部皆未平行於此水平面，而頭部及上肢位置不限，如圖 3.11d。

本研究為了能有較廣泛的使用情境，所收集之嬰兒影像不限定拍攝視角，包含俯視及平視等，總共 15416 張照片。我們將所有影像分為訓練、測試及驗證集，各部分占比為 70%、25% 及 5%，即各有 10815 張、3857 張及 744 張影像。

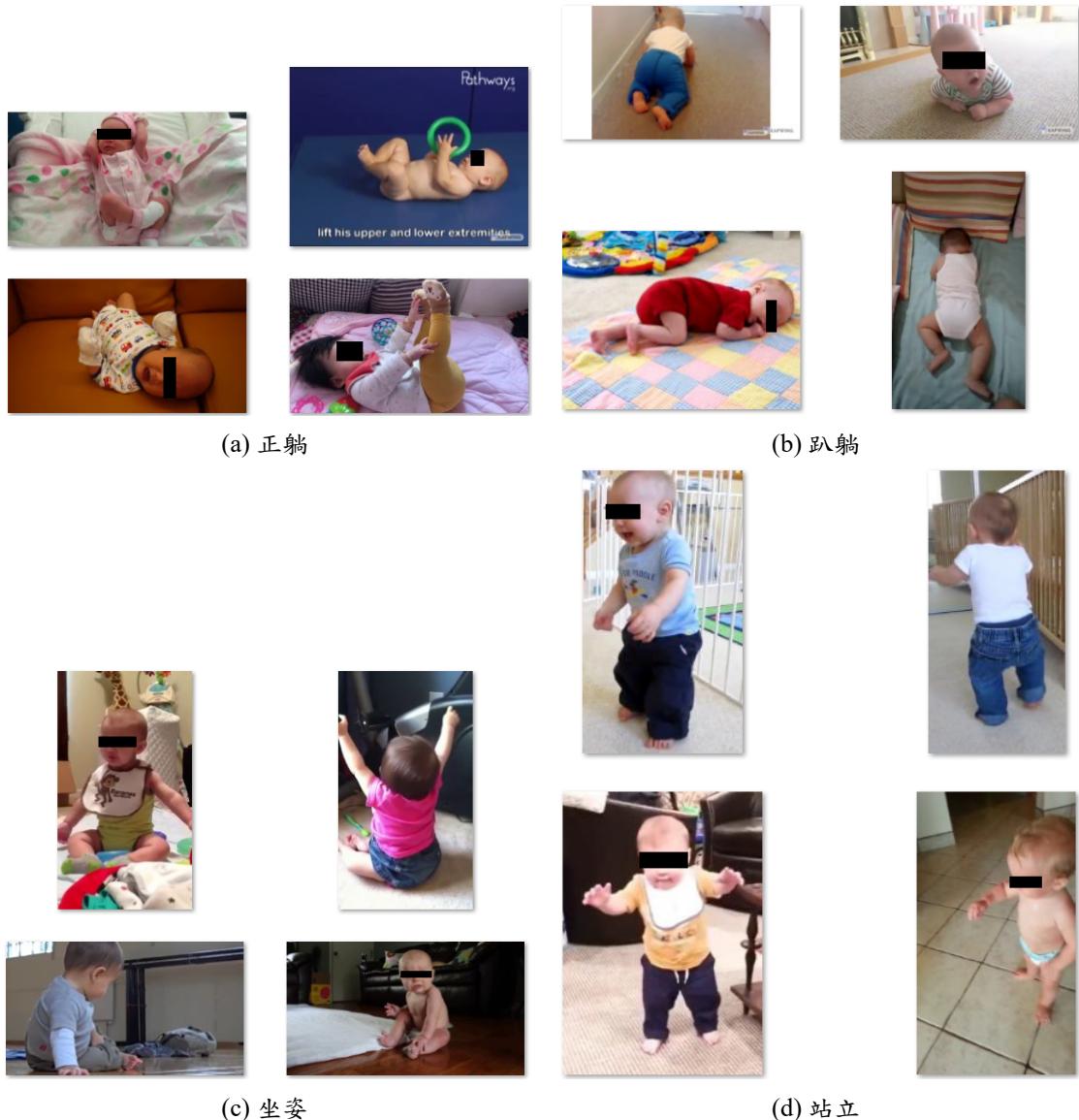


圖 3.11: 嬰兒姿勢資料集

3.3.2 模型訓練

本論文使用 3.3.1 節之嬰兒姿勢資料集，以 ResNet50 [50] 進行姿勢辨識模型之訓練，最終達成辨識四種嬰兒姿勢：正躺、趴躺、坐姿及站立。

3.4 危險情境判斷方法

在實際情境中，當嬰兒做出具危險性之行為時，需持續一段時間才會導致危險發生，並不須判斷一幀畫面為警示狀態，就立即通知照護者。

因此，本系統使用一變數累積模型判斷嬰兒狀態為警示之幀數，當此變數超過設定閥值時，系統才會發出警示提醒照護者。此步驟不但更符合實際使用情境，同時亦可減少因模型辨識錯誤而誤判及誤發警報的情形。

四、 實驗設計與結果

4.1 臉部偵測準確度實驗

4.1.1 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。

因此，本實驗使用 3.3.1 節之嬰兒姿勢資料集，分析 OpenCV [80]、SSD [77]、MTCNN [51] 及 RetinaFace [67] 等臉部偵測演算法其臉部擷取之準確度，以驗證可應用於本系統之演算法。

4.1.2 實驗評估方式

本實驗為驗證臉部偵測演算法應用於本系統之可行性，針對四項演算法之偵測準確度進行比較，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。

4.1.3 實驗結果與分析

根據實驗結果，OpenCV [80] 及 SSD [77] 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，詳細實驗結果請見表 4.1 及表 4.2。而 MTCNN [51] 及 RetinaFace [67] 之實驗結果，請見表 4.3 及表 4.4。

表 4.1: OpenCV [80] 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	

表 4.2: SSD [77] 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	

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

1. OpenCV [80]：由於多數影像皆誤判為無臉 (False)，故僅計算其 precision 為 79.90%。
2. SSD [77]：由於多數影像皆誤判為無臉 (False)，故僅計算其 precision 為 99.90%。
3. MTCNN [51]：accuracy 為 90.20%、precision 為 94.76% 以及 recall 為 90.93%。
4. RetinaFace [67]：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 [67] 進行嬰兒臉部偵測，可擁有較佳的偵測準確度。

表 4.3: MTCNN [51] 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	9878	5538

表 4.4: RetinaFace [67] 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

4.2 臉部偵測執行時間實驗

4.2.1 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速度。

因此，本實驗使用 3.3.1 節之嬰兒姿勢資料集，分析 OpenCV [80]、SSD [77]、MTCNN [51] 及 RetinaFace [67] 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。

4.2.2 實驗評估方式

本實驗為驗證臉部偵測演算法應用於系統之可行性，將針對四項演算法之執行時間進行比較，透過計算演算法偵測共 15416 張影像之資料集所花費的時間，得以算出各演算法平均偵測一張影像之執行時間。

4.2.3 實驗結果與分析

OpenCV [80]、SSD [77]、MTCNN [51] 及 RetinaFace [67] 四項演算法偵測 15416 張影像之詳細實驗結果如下：

1. OpenCV [80]：共 18 分 01.78 秒，平均每張影像需 0.07 秒。
2. SSD [77]：共 9 分 17.26 秒，平均每張影像需 0.04 秒。
3. MTCNN [51]：共 2 小時 8 分 22.05 秒，平均每張影像需 0.50 秒。
4. RetinaFace [67]：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 [77] 進行嬰兒臉部偵測，將可擁有較佳的偵測速度。

而綜觀 4.1 節及 4.2 節之實驗結果，若系統欲擁有較佳的執行速度又兼具偵測準確度，可得出以下結論：先使用 SSD 演算法 [77] 偵測嬰兒臉部，雖然此方法在許多狀況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；而為補足此方法召回率不足之缺失，若 SSD 演算法 [77] 找不到嬰兒臉部時，則接續使用 RetinaFace 演算法 [67]，利用其正確率及準確率皆高之優點進行嬰兒臉部偵測。

4.3 臉部遮擋辨識實驗

4.3.1 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 [50] 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。

4.3.2 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 4.1。

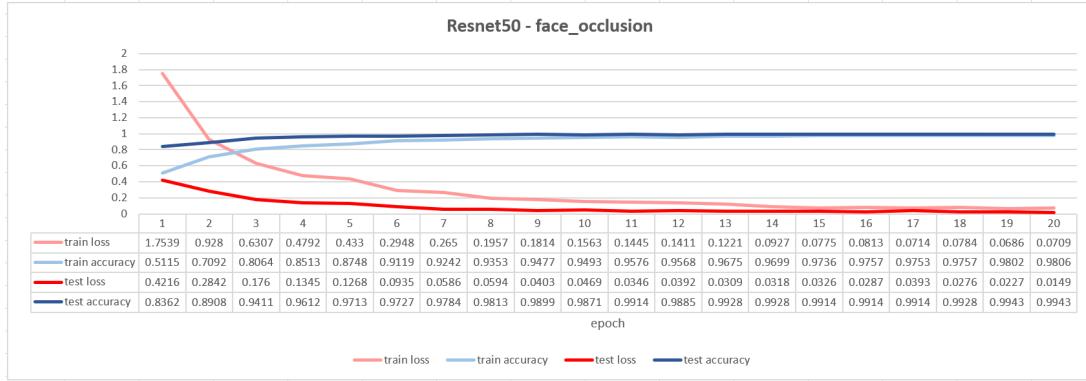


圖 4.1: 臘部辨識訓練結果

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，其混淆矩陣如圖 4.2。

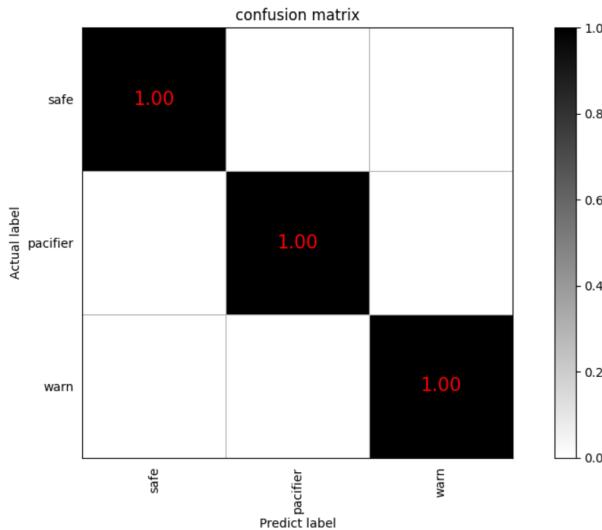


圖 4.2: 臥部遮擋辨識模型之混淆矩陣

4.4 姿勢辨識實驗

4.4.1 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 [50] 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。

4.4.2 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 4.3。



圖 4.3: 姿勢辨識訓練結果

接著，再使用 744 張之驗證集影像進行模型驗證，包含了五張類別辨識錯誤的影像，其混淆矩陣如圖 4.4。辨識錯誤之五張影像中，有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 4.5），而導致誤判。

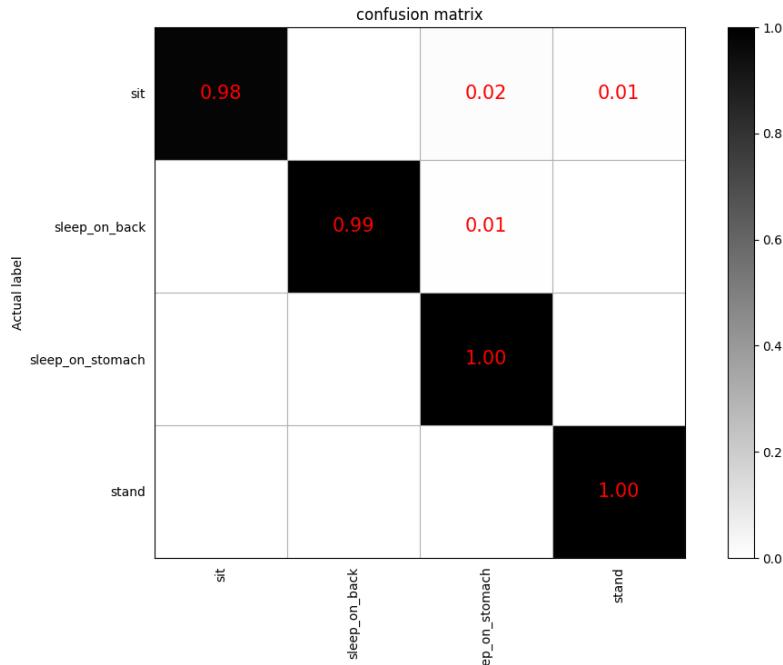


圖 4.4: 姿勢辨識模型之混淆矩陣



圖 4.5: 姿勢辨識錯誤之影像：坐姿誤判為趴躺

4.5 影片危險偵測實驗

4.5.1 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。

4.5.2 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。

4.5.3 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 4.6），而只拍攝到露出的嬰兒臉部，故造成姿勢辨識錯誤，其混淆矩陣如圖 4.7。



圖 4.6: 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 張嬰兒臉部未被偵測之影像（如圖 4.8），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

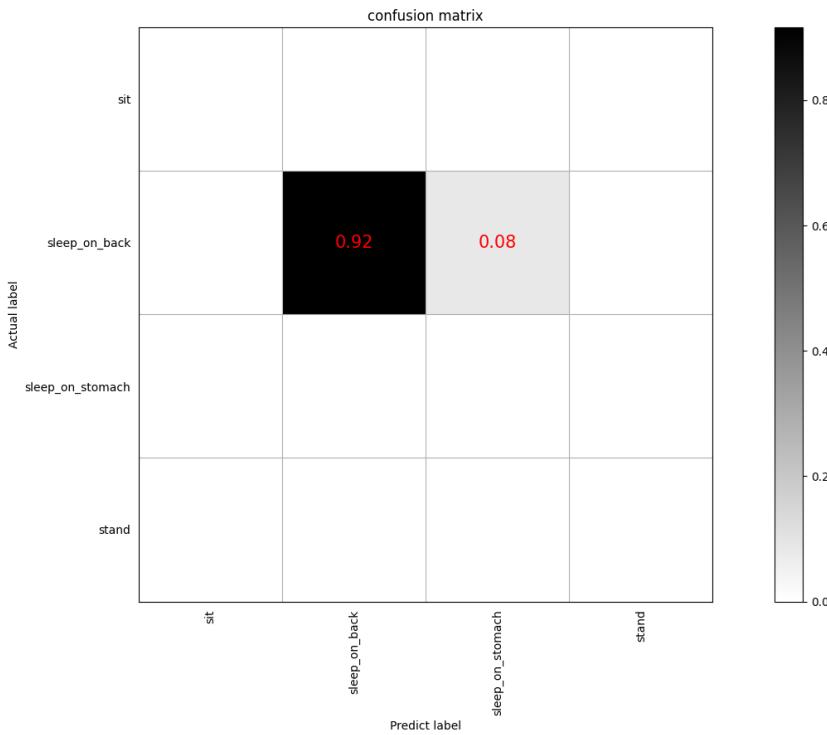


圖 4.7: 實驗影片之姿勢辨識混淆矩陣

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 4.9），而誤判為類別，其混淆矩陣如圖 4.10。



圖 4.8: 未偵測嬰兒臉部之影像

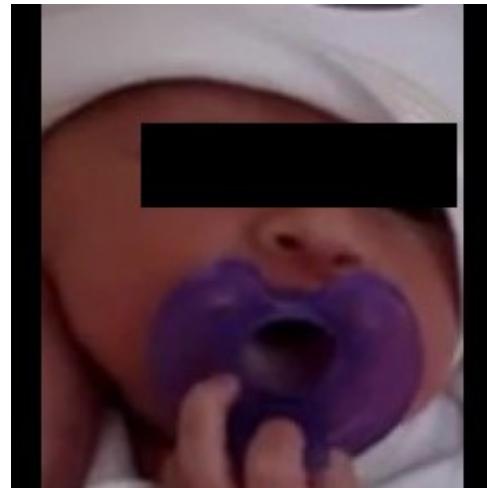


圖 4.9: 臉部遮擋誤判之為警狀狀態

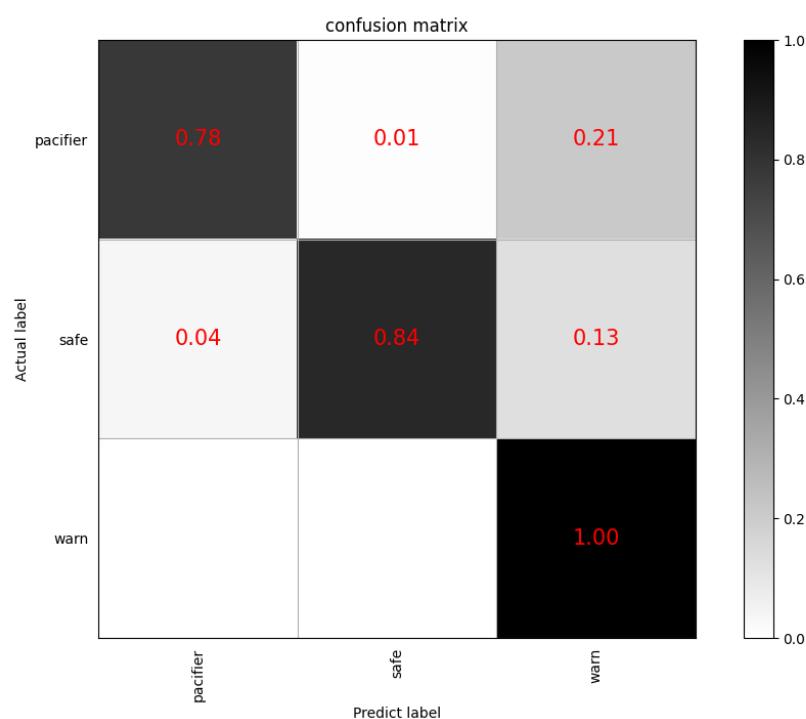


圖 4.10: 實驗影片之臉部遮擋辨識混淆矩陣

五、結論與未來展望

5.1 結論

本論文基於深度學習技術，透過嬰兒影像畫面進行危險偵測，目前可進行兩大部分之偵測：(1) 嬰兒臉部遮擋辨識、及(2) 嬰兒姿勢辨識，其訓練及測試準確率皆達 98% 以上。

本系統優於過往感測器式偵測之功能單一性及不便性，也不同於既有之影像式偵測僅關注嬰兒呼吸或單一動作之研究，而提供了關注於嬰兒臉部及動作之危險監測系統，將有助於協助照護者，並降低嬰兒猝死症發生風險。

另外，由於目前未有公開之嬰兒資料集，故本論文使用之所有嬰兒影像，皆收集自網路上實際嬰兒照片或影片擷取，再經前處理及分類標示而成。

5.2 未來展望

本論文中，目前僅辨識四項嬰兒姿勢，若在偵測姿勢時加入時間資訊，預期得以判斷更多嬰兒行為，如：翻身及爬行等動作，即可監測更多危險情境；而除辨識嬰兒臉部遭異物遮蔽外，若加入偵測面部表情等其他資訊，亦可更詳盡監測嬰兒狀態，以提醒照護者；此外，亦可提供更多嬰兒情境之危險偵測，則使用場景將可更廣泛。

而系統實作方面，未來可提供設定觀測之年齡區間，即可針對不同之特定年齡嬰幼兒警示其具危險性之動作，以達到更符合實際使用情境的危險偵測；亦可結合通訊社群軟體等，如：Line 或 Telegram 等，進行即時之推播訊息以通知照顧者。

參考文獻

- [1] 統計處, 歷年統計, 中文, 文字, Mar. 2021.
- [2] H. C. Kinney and B. T. Thach, “The Sudden Infant Death Syndrome,” *The New England Journal of Medicine*, vol. 361, no. 8, pp. 795–805, Aug. 2009.
- [3] M. Willinger, H. J. Hoffman, and R. B. Hartford, “Infant sleep position and risk for sudden infant death syndrome: Report of meeting held January 13 and 14, 1994, National Institutes of Health, Bethesda, MD,” eng, *Pediatrics*, vol. 93, no. 5, pp. 814–819, May 1994.
- [4] E. A. Mitchell, J. M. Brunt, and C. Everard, “Reduction in mortality from sudden infant death syndrome in New Zealand: 1986-92,” eng, *Archives of Disease in Childhood*, vol. 70, no. 4, pp. 291–294, Apr. 1994.
- [5] S. Iyasu, L. L. Randall, T. K. Welty, *et al.*, “Risk factors for sudden infant death syndrome among northern plains Indians,” eng, *JAMA*, vol. 288, no. 21, pp. 2717–2723, Dec. 2002.
- [6] J. S. Kemp, B. Unger, D. Wilkins, *et al.*, “Unsafe sleep practices and an analysis of bed-sharing among infants dying suddenly and unexpectedly: Results of a four-year, population-based, death-scene investigation study of sudden infant death syndrome and related deaths,” eng, *Pediatrics*, vol. 106, no. 3, E41, Sep. 2000.
- [7] A. L. Ponsonby, T. Dwyer, L. E. Gibbons, J. A. Cochrane, and Y. G. Wang, “Factors potentiating the risk of sudden infant death syndrome associated with the prone position,” eng, *The New England Journal of Medicine*, vol. 329, no. 6, pp. 377–382, Aug. 1993.
- [8] R. S. C. Horne, “Effects of prematurity on heart rate control: Implications for sudden infant death syndrome,” eng, *Expert Review of Cardiovascular Therapy*, vol. 4, no. 3, pp. 335–343, May 2006.
- [9] H. Ce, “Sudden infant death syndrome and other causes of infant mortality: Diagnosis, mechanisms, and risk for recurrence in siblings,” en, *American journal of respiratory and critical care medicine*, vol. 164, no. 3, Jan. 2001.
- [10] N. Oyen, R. Skjaerven, and L. M. Irgens, “Population-based recurrence risk of sudden infant death syndrome compared with other infant and fetal deaths,” eng, *American Journal of Epidemiology*, vol. 144, no. 3, pp. 300–305, Aug. 1996.

- [11] F. R. Hauck, S. M. Herman, M. Donovan, *et al.*, “Sleep environment and the risk of sudden infant death syndrome in an urban population: The Chicago Infant Mortality Study,” eng, *Pediatrics*, vol. 111, no. 5 Pt 2, pp. 1207–1214, May 2003.
- [12] J. S. Kemp and B. T. Thach, “Sudden death in infants sleeping on polystyrene-filled cushions,” eng, *The New England Journal of Medicine*, vol. 324, no. 26, pp. 1858–1864, Jun. 1991.
- [13] B. T. Skadberg and T. Markestad, “Consequences of getting the head covered during sleep in infancy,” eng, *Pediatrics*, vol. 100, no. 2, E6, Aug. 1997.
- [14] *What causes SIDS?* en.
- [15] R. Y. Moon and Task Force on Sudden Infant Death Syndrome, “SIDS and Other Sleep-Related Infant Deaths: Expansion of Recommendations for a Safe Infant Sleeping Environment,” *Pediatrics*, vol. 128, no. 5, e1341–e1367, Nov. 2011.
- [16] T. Klingenberg and M. Schilling, “Mobile wearable device for long term monitoring of vital signs,” en, *Computer Methods and Programs in Biomedicine*, Design of Environments for Ageing, vol. 106, no. 2, pp. 89–96, May 2012.
- [17] C. Zhou, C. Tu, J. Tian, J. Feng, Y. Gao, and X. Ye, “A low power miniaturized monitoring system of six human physiological parameters based on wearable body sensor network,” *Sensor Review*, vol. 35, no. 2, pp. 210–218, Jan. 2015.
- [18] S. Bouwstra, W. Chen, L. Feijis, and S. B. Oetomo, “Smart Jacket Design for Neonatal Monitoring with Wearable Sensors,” in *2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks*, Jun. 2009, pp. 162–167.
- [19] K. Malhi, S. C. Mukhopadhyay, J. Schnepper, M. Haefke, and H. Ewald, “A Zigbee-Based Wearable Physiological Parameters Monitoring System,” *IEEE Sensors Journal*, vol. 12, no. 3, pp. 423–430, Mar. 2012.
- [20] S. González-Valenzuela, M. Chen, and V. C. M. Leung, “Mobility Support for Health Monitoring at Home Using Wearable Sensors,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 4, pp. 539–549, Jul. 2011.
- [21] A. Darwish and A. E. Hassanien, “Wearable and Implantable Wireless Sensor Network Solutions for Healthcare Monitoring,” en, *Sensors*, vol. 11, no. 6, pp. 5561–5595, Jun. 2011.
- [22] J. Ko, C. Lu, M. B. Srivastava, J. A. Stankovic, A. Terzis, and M. Welsh, “Wireless Sensor Networks for Healthcare,” *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1947–1960, Nov. 2010.
- [23] C. Linti, H. Hörter, P. Osterreicher, and H. Planck, “Sensory baby vest for the monitoring of infants,” in *International Workshop on Wearable and Implantable Body Sensor Networks (BSN’06)*, Apr. 2006, 3 pp.–137.

- [24] A. G. Ferreira, D. Fernandes, S. Branco, *et al.*, “A smart wearable system for sudden infant death syndrome monitoring,” in *2016 IEEE International Conference on Industrial Technology (ICIT)*, Mar. 2016, pp. 1920–1925.
- [25] W. Lin, R. Zhang, J. Brittelli, and C. Lehmann, “Wireless Infant Monitoring Device for the prevention of sudden infant death syndrome,” in *2014 11th International Conference Expo on Emerging Technologies for a Smarter World (CEWIT)*, Oct. 2014, pp. 1–4.
- [26] E. G. Ziganshin, M. A. Numerov, and S. A. Vygolov, “UWB Baby Monitor,” in *2010 5th International Conference on Ultrawideband and Ultrashort Impulse Signals*, Sep. 2010, pp. 159–161.
- [27] E. Staderini, “UWB radars in medicine,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 17, no. 1, pp. 13–18, Jan. 2002.
- [28] Y.-M. Kuo, J.-S. Lee, and P.-C. Chung, “A Visual Context-Awareness-Based Sleeping-Respiration Measurement System,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 255–265, Mar. 2010.
- [29] T. Yiping, J. Shunjing, Y. Zhongyuan, and Y. Sisi, “Detection Elder Abnormal Activities by using Omni-directional Vision Sensor: Activity Data Collection and Modeling,” in *2006 SICE-ICASE International Joint Conference*, Oct. 2006, pp. 3850–3853.
- [30] O. Rudovic, Y. Utsumi, J. Lee, *et al.*, “CultureNet: A Deep Learning Approach for Engagement Intensity Estimation from Face Images of Children with Autism,” in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct. 2018, pp. 339–346.
- [31] Z. Liu, P. Luo, X. Wang, and X. Tang, “Deep Learning Face Attributes in the Wild,” in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015, pp. 3730–3738.
- [32] D. Mack, J. Patrie, P. Suratt, R. Felder, and M. Alwan, “Development and Preliminary Validation of Heart Rate and Breathing Rate Detection Using a Passive, Ballistocardiography-Based Sleep Monitoring System,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 1, pp. 111–120, Jan. 2009.
- [33] T. Berset, I. Romero, A. Young, and J. Penders, “Robust heart rhythm calculation and respiration rate estimation in ambulatory ECG monitoring,” in *Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics*, Jan. 2012, pp. 400–403.
- [34] H. Cao, L.-C. Hsu, T. Ativanichayaphong, J. Sin, and J.-C. Chiao, “A non-invasive and remote infant monitoring system using CO₂ sensors,” in *2007 IEEE SENSORS*, Oct. 2007, pp. 989–992.
- [35] L. Scalise, I. Ercoli, P. Marchionni, and E. P. Tomasini, “Measurement of respiration rate in preterm infants by laser Doppler vibrometry,” in *2011 IEEE International Symposium on Medical Measurements and Applications*, May 2011, pp. 657–661.

- [36] C.-Y. Fang, H.-H. Hsieh, and S.-W. Chen, “A Vision-Based Infant Respiratory Frequency Detection System,” in *2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, Nov. 2015, pp. 1–8.
- [37] X. Liu, K. Takeuchi, T. Ogunfunmi, and S. Mathapathi, “Video-based IoT baby monitor for SIDS prevention,” in *2017 IEEE Global Humanitarian Technology Conference (GHTC)*, Oct. 2017, pp. 1–7.
- [38] X. L. Gallo, S. Lechón, S. Mora, and D. Vallejo-Huanga, “MARRSIDS: Monitoring Assistant to Reduce the Risk of Sudden Infant Death Syndrome,” in *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, Apr. 2019, pp. 1–4.
- [39] T.-J. Wang, J. Laaksonen, Y.-P. Liao, B.-Z. Wu, and S.-Y. Shen, “A Multi-Task Bayesian Deep Neural Net for Detecting Life-Threatening Infant Incidents From Head Images,” in *2019 IEEE International Conference on Image Processing (ICIP)*, Sep. 2019, pp. 3006–3010.
- [40] A. Kendall and Y. Gal, *What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?* Oct. 2017.
- [41] V. Viitaniemi, M. Karppa, J. Laaksonen, and T. Jantunen, “Detecting Hand-Head Occlusions in Sign Language Video,” en, in *Image Analysis*, J.-K. Kämäärinen and M. Koskela, Eds., ser. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 2013, pp. 361–372.
- [42] S. Ge, J. Li, Q. Ye, and Z. Luo, “Detecting Masked Faces in the Wild with LLE-CNNs,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 426–434.
- [43] G. Ghiasi and C. Fowlkes, “Using Segmentation to Predict the Absence of Occluded Parts,” en, in *Proceedings of the British Machine Vision Conference 2015*, Swansea: British Machine Vision Association, 2015, pp. 22.1–22.12.
- [44] V. Bharati, “An Efficient Edge Deep Learning Computer Vision System to Prevent Sudden Infant Death Syndrome,” in *2021 IEEE International Conference on Smart Computing (SMARTCOMP)*, Aug. 2021, pp. 286–291.
- [45] K. Simonyan and A. Zisserman, *Very Deep Convolutional Networks for Large-Scale Image Recognition*, Apr. 2015.
- [46] C. Szegedy, W. Liu, Y. Jia, *et al.*, “Going deeper with convolutions,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 1–9.
- [47] S. Ioffe and C. Szegedy, *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*, Mar. 2015.
- [48] K. He and J. Sun, “Convolutional neural networks at constrained time cost,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 5353–5360.

- [49] R. K. Srivastava, K. Greff, and J. Schmidhuber, *Highway Networks*, Nov. 2015.
- [50] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778.
- [51] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, “Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks,” *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016.
- [52] B. Yang, J. Yan, Z. Lei, and S. Z. Li, *Aggregate channel features for multi-view face detection*, Sep. 2014.
- [53] M. Mathias, R. Benenson, M. Pedersoli, and L. Van Gool, “Face Detection without Bells and Whistles,” en, in *Computer Vision -ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., ser. Lecture Notes in Computer Science, Cham: Springer International Publishing, 2014, pp. 720–735.
- [54] J. Yan, Z. Lei, L. Wen, and S. Z. Li, “The Fastest Deformable Part Model for Object Detection,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2014, pp. 2497–2504.
- [55] S. Yang, P. Luo, C.-C. Loy, and X. Tang, “From Facial Parts Responses to Face Detection: A Deep Learning Approach,” in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015, pp. 3676–3684.
- [56] D. Chen, S. Ren, Y. Wei, X. Cao, and J. Sun, “Joint Cascade Face Detection and Alignment,” en, in *Computer Vision -ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., ser. Lecture Notes in Computer Science, Cham: Springer International Publishing, 2014, pp. 109–122.
- [57] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, “A convolutional neural network cascade for face detection,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 5325–5334.
- [58] B. Yang, J. Yan, Z. Lei, and S. Z. Li, “Convolutional Channel Features,” in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015, pp. 82–90.
- [59] S. S. Farfade, M. J. Saberian, and L.-J. Li, “Multi-view Face Detection Using Deep Convolutional Neural Networks,” in *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, ser. ICMR ’15, New York, NY, USA: Association for Computing Machinery, Jun. 2015, pp. 643–650.
- [60] S. Yang, P. Luo, C. C. Loy, and X. Tang, *WIDER FACE: A Face Detection Benchmark*, Nov. 2015.
- [61] X. Cao, Y. Wei, F. Wen, and J. Sun, “Face Alignment by Explicit Shape Regression,” en, *International Journal of Computer Vision*, vol. 107, no. 2, pp. 177–190, Apr. 2014.

- [62] X. Yu, J. Huang, S. Zhang, W. Yan, and D. N. Metaxas, “Pose-Free Facial Landmark Fitting via Optimized Part Mixtures and Cascaded Deformable Shape Model,” in *2013 IEEE International Conference on Computer Vision*, Dec. 2013, pp. 1944–1951.
- [63] X. Xiong and F. De la Torre, “Supervised Descent Method and Its Applications to Face Alignment,” in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2013, pp. 532–539.
- [64] X. P. Burgos-Artizzu, P. Perona, and P. Dollár, “Robust Face Landmark Estimation under Occlusion,” in *2013 IEEE International Conference on Computer Vision*, Dec. 2013, pp. 1513–1520.
- [65] X. Zhu and D. Ramanan, “Face detection, pose estimation, and landmark localization in the wild,” in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2012, pp. 2879–2886.
- [66] Z. Zhang, P. Luo, C. C. Loy, and X. Tang, “Facial Landmark Detection by Deep Multi-task Learning,” en, in *Computer Vision -ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., ser. Lecture Notes in Computer Science, Cham: Springer International Publishing, 2014, pp. 94–108.
- [67] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, “RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild,” in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, pp. 5202–5211.
- [68] S. Zhang, L. Wen, H. Shi, Z. Lei, S. Lyu, and S. Z. Li, “Single-Shot Scale-Aware Network for Real-Time Face Detection,” en, *International Journal of Computer Vision*, vol. 127, no. 6, pp. 537–559, Jun. 2019.
- [69] X. Tang, D. K. Du, Z. He, and J. Liu, “PyramidBox: A Context-Assisted Single Shot Face Detector,” en, in *Computer Vision -ECCV 2018*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds., ser. Lecture Notes in Computer Science, Cham: Springer International Publishing, 2018, pp. 812–828.
- [70] M. Najibi, P. Samangouei, R. Chellappa, and L. S. Davis, “SSH: Single Stage Headless Face Detector,” in *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct. 2017, pp. 4885–4894.
- [71] S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, and S. Z. Li, *S\$^3\$FD: Single Shot Scale-invariant Face Detector*, Nov. 2017.
- [72] J. Li, Y. Wang, C. Wang, *et al.*, *DSFD: Dual Shot Face Detector*, Apr. 2019.
- [73] H.-k. Tang and Z.-q. Feng, “Hand’s Skin Detection Based on Ellipse Clustering,” in *2008 International Symposium on Computer Science and Computational Technology*, vol. 2, Dec. 2008, pp. 758–761.
- [74] W. Li, Q. Yang, and X. He, “Face detection algorithm based on double ellipse skin model,” in *2011 IEEE 2nd International Conference on Software Engineering and Service Science*, Jul. 2011, pp. 335–339.

- [75] *Python opencv 膚色檢測的實現示例* _ 程式設計 _ 程式人生.
- [76] WalkonNet, *Python opencv 膚色檢測的實現示例-WalkonNet*, zh-TW.
- [77] B. Ye, Y. Shi, H. Li, L. Li, and S. Tong, “Face SSD: A Real-time Face Detector based on SSD,” in *2021 40th Chinese Control Conference (CCC)*, Jul. 2021, pp. 8445–8450.
- [78] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, *OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields*, May 2019.
- [79] *Pose*, en-US.
- [80] K. Goyal, K. Agarwal, and R. Kumar, “Face detection and tracking: Using OpenCV,” in *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)*, vol. 1, Apr. 2017, pp. 474–478.