

# 一、相關研究

## 1.1 嬰兒猝死症

嬰兒猝死症 (The Sudden Infant Death Syndrome, 簡稱 SIDS) [kinney\\_sudden\\_2009](#) 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 [willinger\\_infant\\_1994](#), [mitchell\\_reduction\\_1994](#), [iyasu\\_risk\\_2002](#), [kemp\\_unsafe\\_2000](#), [ponsonby\\_factors\\_1993](#)，致使嬰兒呼吸困難而死亡；其二則為內在因素，包含發展因素(如：早產 [horne\\_effects\\_2009](#))、遺傳因素(如：家族性之嬰兒猝死症 [ce\\_sudden\\_2001](#), [oyen\\_population-based\\_1996](#))、性別（男性比例為女性的兩倍）或種族 [hauck\\_sleep\\_2003](#)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 [kinney\\_sudden\\_2009](#)：(1) 發生危及生命的事件 (如：面部朝下 [kemp\\_sudden\\_1991](#)或面部遭遮蔽 [skadberg\\_consequences\\_1997](#)，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

## 一、相關研究

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 **noauthor\_what\_nod** 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 **moon\_sids\_2011**：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 **willinger\_infant\_1994**，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。

## 1.2 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。

### 1.2.1 感測器偵測

此種方式利用多種不同感測器進行生理訊號之偵測，包含利用呼吸感測器、濕度感測器、溫度感測器、非接觸式紅外溫度感測器、三軸加速度計、慣性感測器、一氧化碳感測器、二氧化碳感測器等，分別量測嬰兒之呼吸頻率、出汗狀況、體溫、心率、身體位置或方向、睡眠姿勢、嬰兒周圍的一氧化碳濃度、呼出的二氧化碳濃度的變化等，且多會透過物聯網技術開發出可穿戴式裝置之系統 **klingeberg\_mobile\_2012, zhou\_low\_2015, bouwstra\_smart\_2009, malhi\_zigbee-based\_2012, gonzalez-valenzuela\_mobility\_2011, darwish\_wearable\_2010**。

如：Linti 等人 **linti\_sensory\_2006** 所開發的嬰兒感測背心（圖 1.1），其將多個感官元件融入紡織品中以用來量測嬰兒之呼吸、心率、溫度及濕度；Ferreira 等人 **ferreira\_smart\_2016** 開發了將心律感測器、3D 加速



圖 1.1: 嬰兒多感測器背心之穿脫示意圖 **linti\_sensory\_2006**

度計 **hung\_estimation\_2008, pitts\_respiratory\_2013, bates\_respiratory\_2010**、熱電堆感測器裝設於胸帶中（圖 1.2），而得以量測嬰兒之體溫、心率、呼吸頻率及身體位置，並透過 ZigBee 技術將收集到的數據傳送至伺服

## 一、相關研究

器，用戶則可透過醫療網頁介面（圖 1.3）進行查看及收到緊急訊息；Zi-

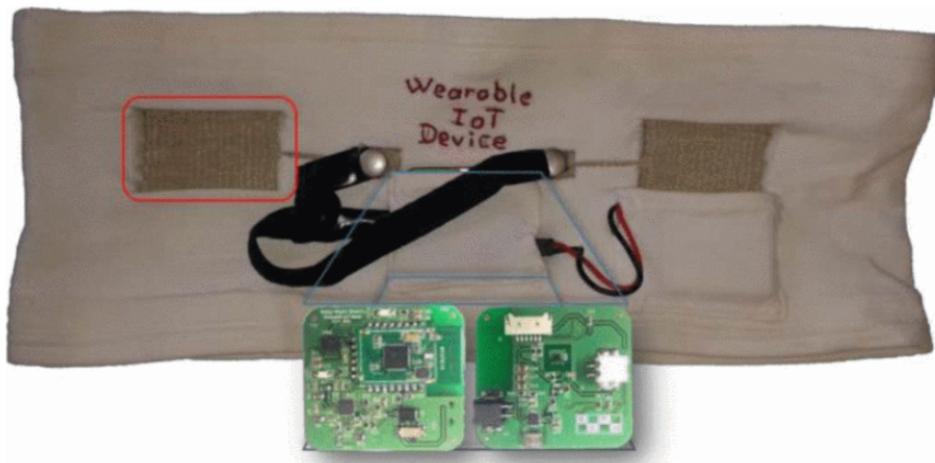


圖 1.2: 嬰兒穿戴式感測器裝置圖 ferreira\_smart\_2016

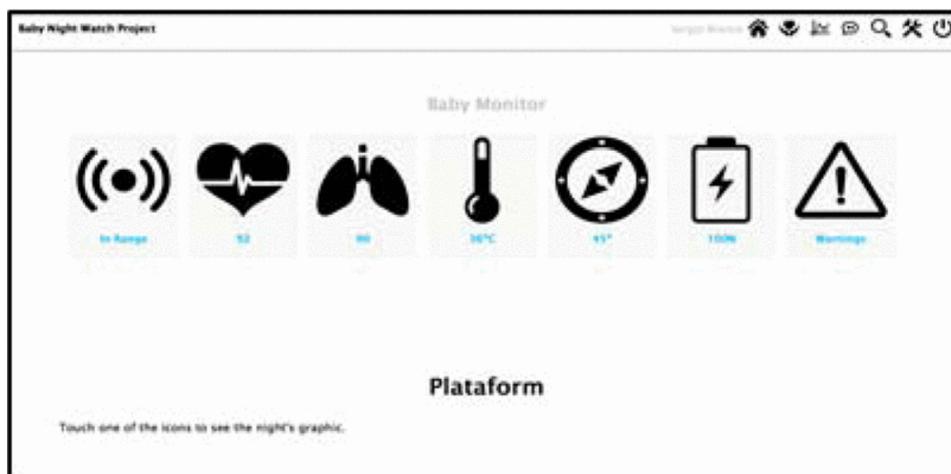


圖 1.3: 醫療網頁介面 ferreira\_smart\_2016

ganshin 等人 **ziganshin\_uwb\_2010** 基於超寬頻技術 **staderini\_uwb\_2002** 開發出可監測嬰兒呼吸及心率之系統，其可檢測嬰兒之睡眠狀態（圖 1.4）、清醒狀態及警示狀態（圖 1.5）；Lin 等人 **lin\_wireless\_2014** 開發出一套在嬰兒胸帶上嵌入了三種不同感測器的系統（圖 1.6），透過三軸加速度計可確定嬰兒睡姿（面朝上、下、左或右）與計算 Z 軸資訊得出呼吸頻率（圖 1.7）、利用溫度感測器量測體溫以及使用一氧化碳感測器偵測嬰兒周圍之一氧化碳濃度，再藉由 WiFi 模組傳送收集之生理資料至伺服器，而其驗證所計算之呼吸頻率準確率達 100%。

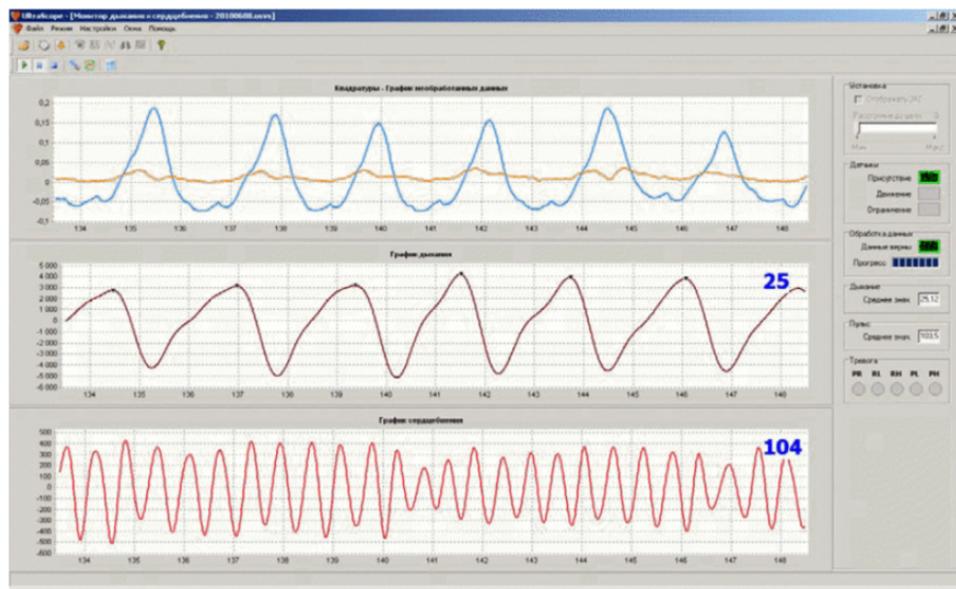


圖 1.4: 嬰兒正常狀態之呼吸及心跳圖 ziganshin\_uwb\_2010

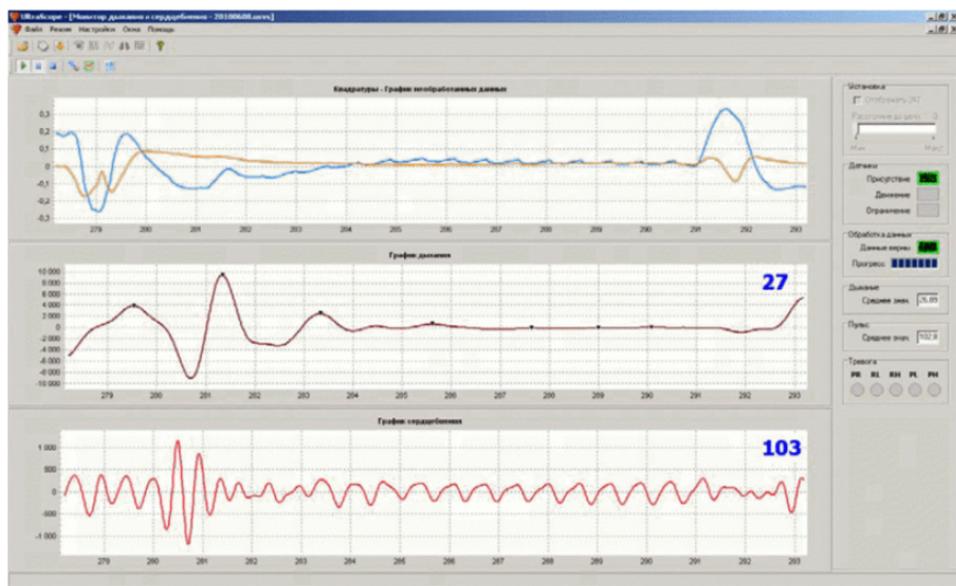


圖 1.5: 嬰兒呼吸暫停狀態之呼吸及心跳圖 ziganshin\_uwb\_2010

此種利用感測器監測嬰兒的方法，雖然可直接量測嬰兒之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，亦或者因嬰兒需額外穿戴裝置而造成不適，進而影響嬰兒活動或導致更多危險的發生。

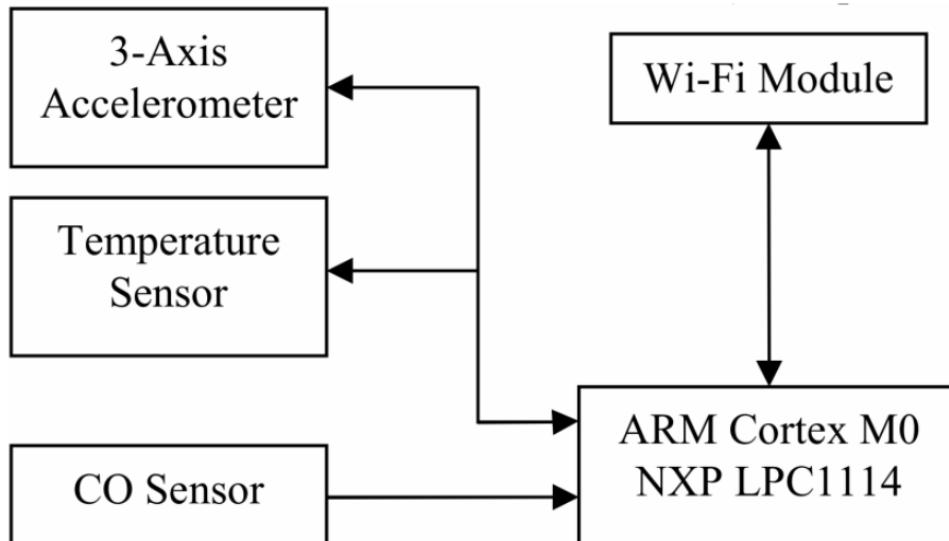


圖 1.6: 嬰兒監測系統架構圖 lin\_wireless\_2014

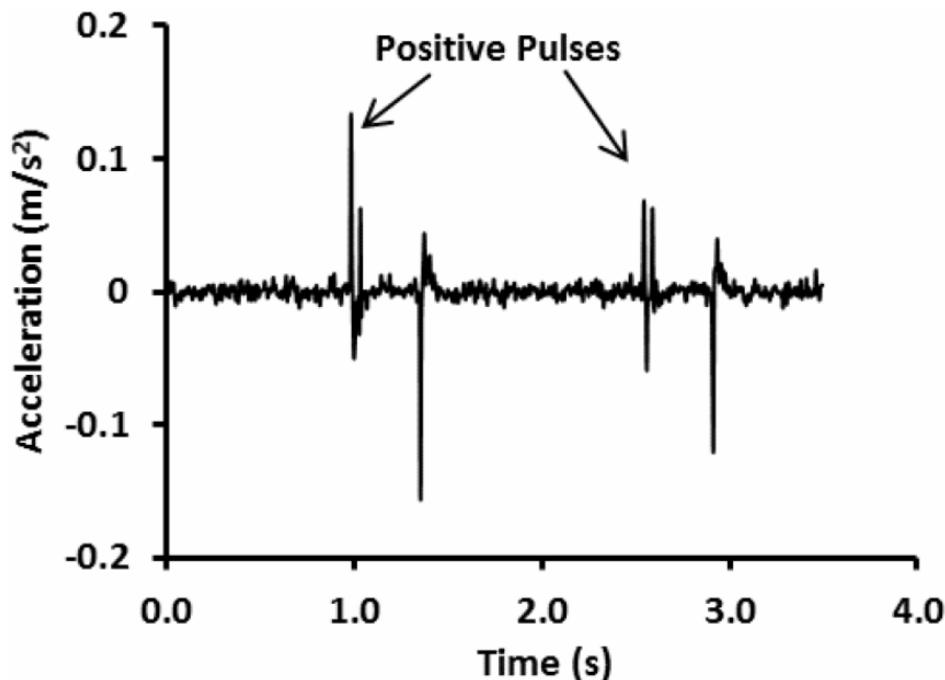


圖 1.7: z 軸顯示吸氣及呼氣脈衝 lin\_wireless\_2014

### 1.2.2 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 [kuo\\_visual\\_2010](#), [yiping\\_detection\\_2006](#), [rudovic\\_culturenet\\_2018](#), [liu\\_deep\\_2015](#)，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。

Fang 等人 **fang\_vision-based\_2015** 開發了基於視覺之非接觸式呼吸頻率偵測系統，其有別於過往接觸式呼吸頻率偵測系統，如利用呼吸電感體積描記法 **mack\_development\_2009** 及心電圖 **berset\_robust\_2012** 等；或使用非視覺技術之非接觸式系統，如透過二氧化碳感測器 **cao\_non-invasive\_2009**、雷射都卜勒測振計 **scalise\_measurement\_2011** 等進行偵測。此系統完整流程為（見圖 1.8）：先判斷嬰兒是否正在運動，即包含頭部、四肢及身體運動，但不包含因呼吸引起的輕微運動，此部分判斷流程圖見圖 1.9；若系統未偵測到嬰兒運動，則開始計算嬰兒之呼吸頻率：首先，透過空間特徵擷取呼吸之候選點；接著，利用模糊積分技術選擇呼吸點；最終，得以計算嬰兒的呼吸頻率，進而判斷是否發生呼吸異常之情形。

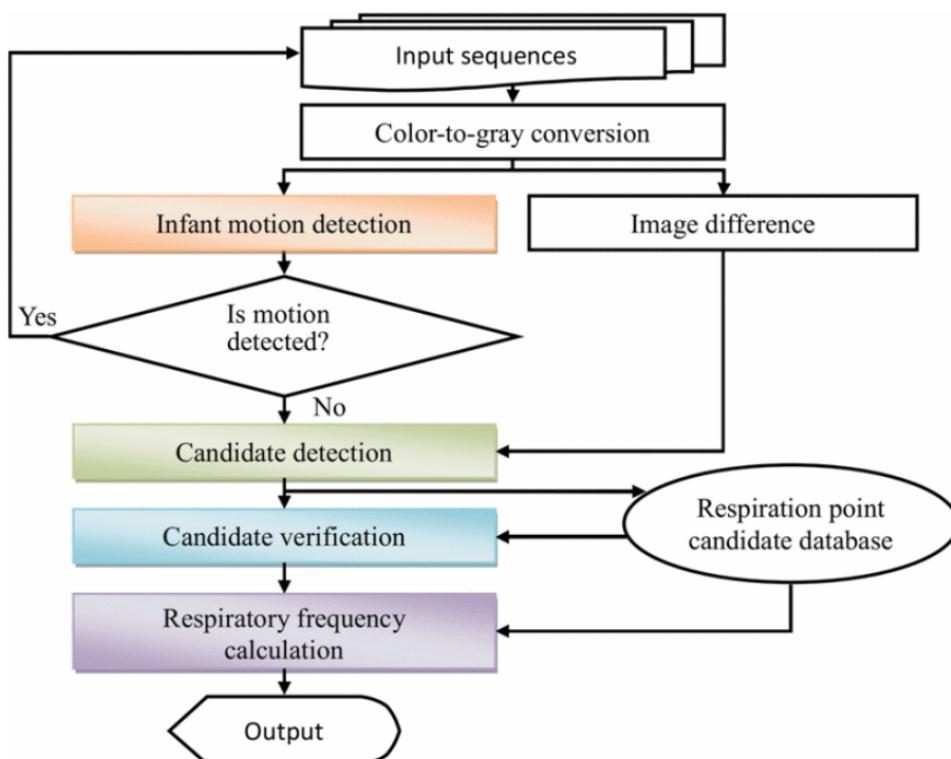


圖 1.8: 嬰兒呼吸頻率偵測系統流程圖 **fang\_vision-based\_2015**

Liu 等人 **liu\_video-based\_2017** 利用夜視攝影機拍攝在嬰兒床內的嬰兒，以判斷其呼吸頻率異常與否。此系統設計圖見圖 1.10，主要包含了三個部分：(1) 影片收集：為圖 1.10 中的前兩項，透過夜視攝影機連接

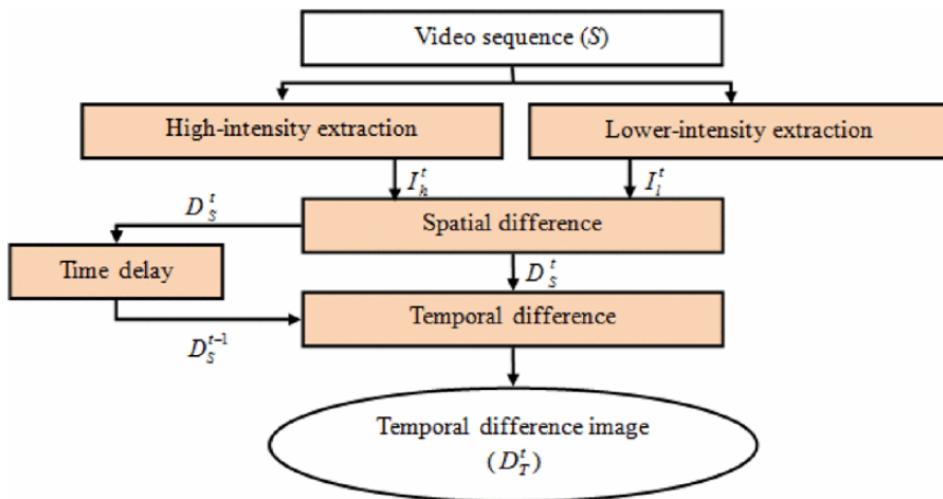


圖 1.9: 嬰兒運動偵測流程圖 fang\_vision-based\_2015

到 Artik 板以收集嬰兒影片；(2) 呼吸偵測演算法：為圖 1.10 中的第三及四項，包含了使用 MIT 所提出之 Eulerian Magnification 技術，用來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閥值，則判斷為呼吸頻率異常；(3) 警示：為圖 1.10 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。

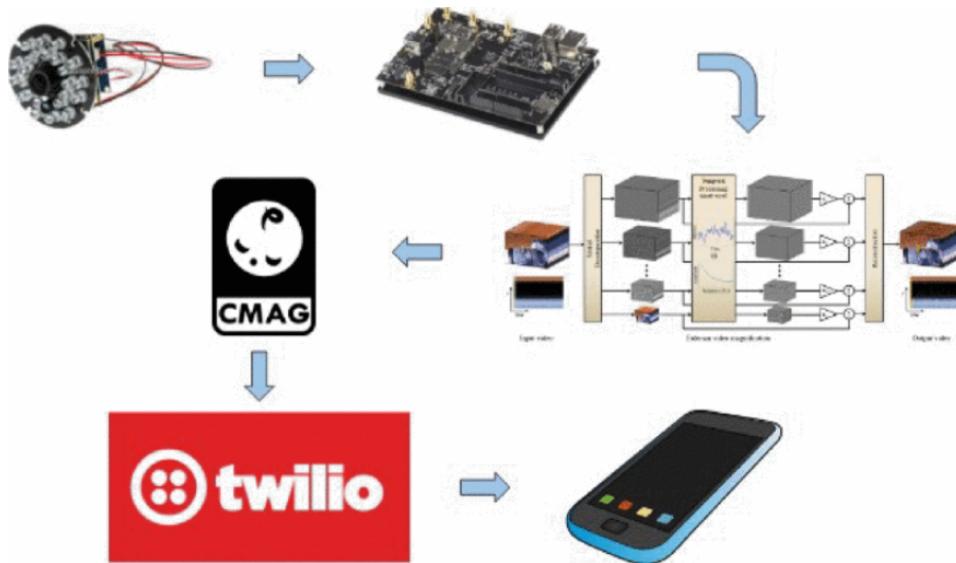


圖 1.10: 系統設計圖 liu\_video-based\_2017

Gallo 等人 gallo\_marrsids\_2019 提出一名為 MARRSIDS 的模型，其利用 OpenCV 之 Haar-Like Features 偵測嬰兒之面部特徵。系統透過嬰兒

臉部辨識與否及睜眼狀態，判斷其是否處於危險情境中，而需發出聲音警示：若臉部未被偵測，則認為嬰兒可能位於不良姿勢需發出警示；而若嬰兒為睜眼狀態，則代表嬰兒處於清醒狀態，並非處於風險中。系統之硬體架構圖見圖 1.11，人工視覺模型架構見圖 1.12。

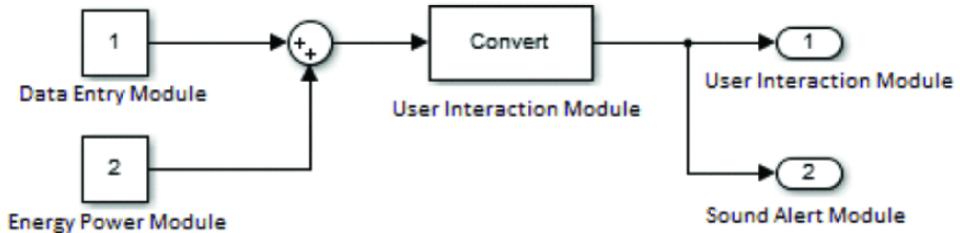


圖 1.11: 硬體架構圖 gallo\_marrsids\_2019

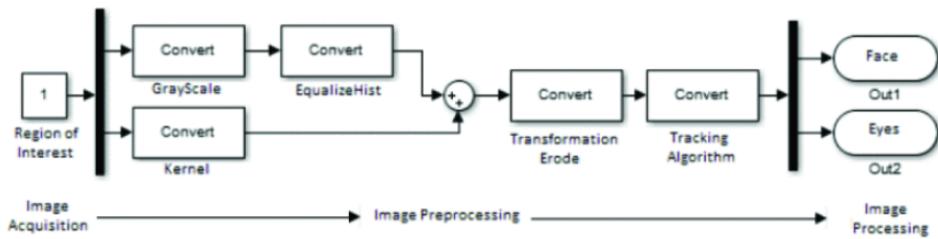


圖 1.12: 人工視覺模型架構圖 gallo\_marrsids\_2019

Wang 等人 wang\_multi-task\_2019 提出用以偵測嬰兒臉部遮擋的貝氏深度神經架構（圖 1.13），其包含四項子任務：(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，並考慮了不同偵測任務間的不確定性 kendall\_what\_2017，且有別於現有不專注於嬰兒影像的研究 viitaniemi\_detecting\_2013, ge\_detecting\_2017, ghiasi\_using\_2015，其使用 MobileNetV2 針對自行收集之 YunInfants 資料集進行嬰兒頭部影像分析。

Bharati 等人 bharati\_efficient\_2021 提出基於卷積神經網絡之嬰兒睡眠姿勢電腦視覺系統，其可評估仰臥（正常狀態）、從仰臥轉換到趴臥（警報狀態）及趴臥（危險狀態）三種睡姿（如圖 1.14），並於嬰兒呈現趴臥姿勢時，透過手機提醒照護人員，此系統架構圖見圖 1.16。其中的

## 一、相關研究

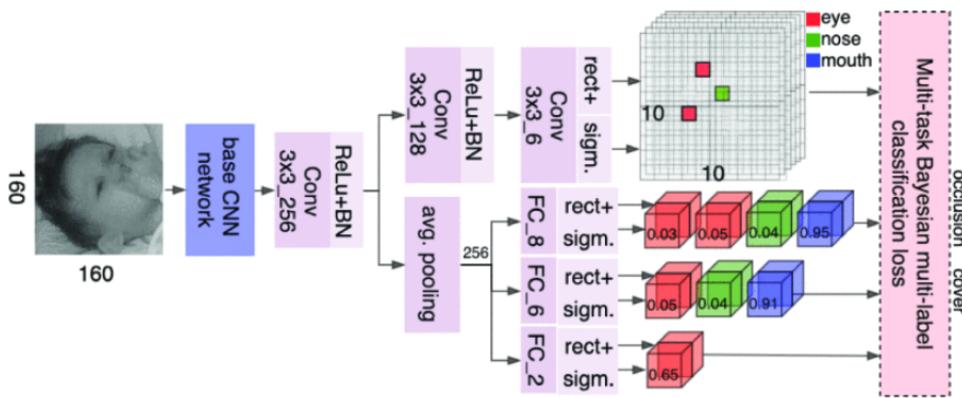


圖 1.13: 網路架構 wang\_multi-task\_2019

CNN 架構（圖 1.15）為：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，最終輸出嬰兒的三個睡眠姿勢機率值。由於目前未有公開之嬰兒姿勢資料集，故本文所使用的影像資料為和真實嬰兒相同比例之娃娃。



圖 1.14: 睡姿分類 bharati\_efficient\_2021

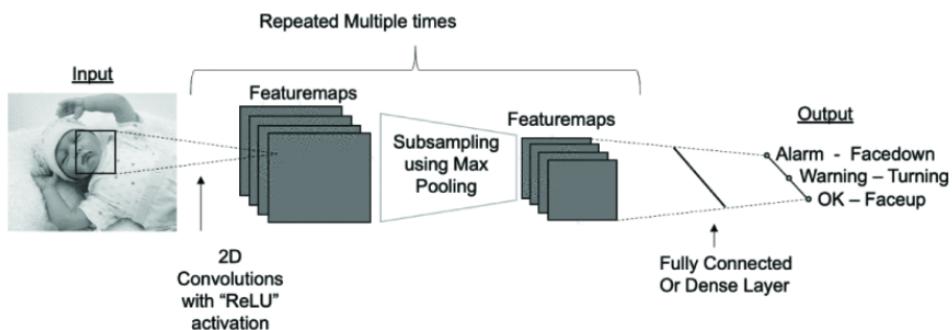


圖 1.15: CNN 架構圖 bharati\_efficient\_2021

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，而尚未有對於嬰兒常見動作之辨識模型，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。

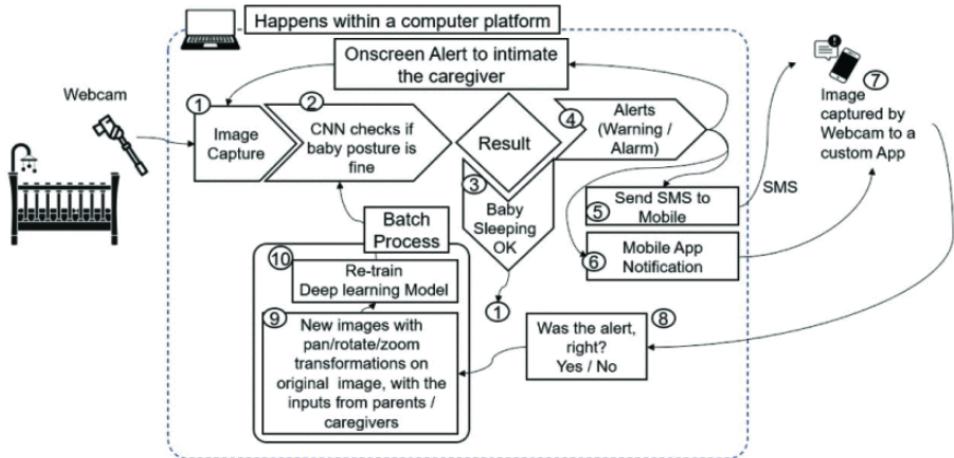


圖 1.16: 系統架構圖 bharati\_efficient\_2021

## 1.3 ResNet

既有的研究中，可知卷積神經網路其深度至關重要 simonyan\_very\_2015, szegedy\_going\_2015，在 ImageNet 資料集中結果表現良好的網路，皆利用了深度十六層 simonyan\_very\_2015 至三十層 ioffe\_batch\_2015 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 he\_convolutional\_2015, srivastava\_highway\_2015，如圖 1.17 可看到兩個不同層數的網路其訓練誤差值。

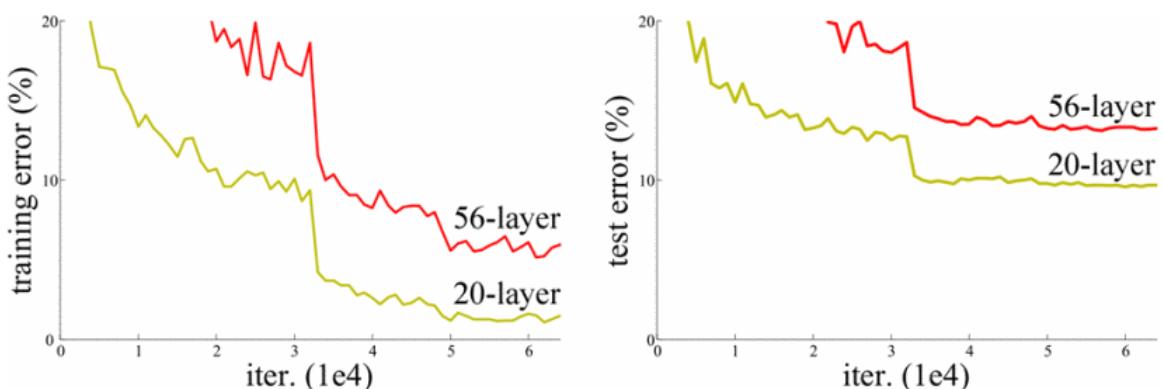


圖 1.17: 網路深度與訓練誤差關係

因此，He 等人 **he\_deep\_2016** 提出使用深度殘差學習（圖 1.18）的網路架構，其利用 shortcut connection 執行 identity mapping，且不需要增加額外的參數，即不增加計算複雜度。在此研究中，通過訓練 ImageNet 評估 18 層及 34 層之普通網路與殘差網路，由圖 1.19（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 1.19（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 1.20（左）為不同層數之普通網路其誤差曲線圖，圖 1.20（中）為使用殘差網路之不同層數其誤差曲線圖，圖 1.20（右）則為使用 110 層及 1202 層殘差網路之誤差曲線圖。

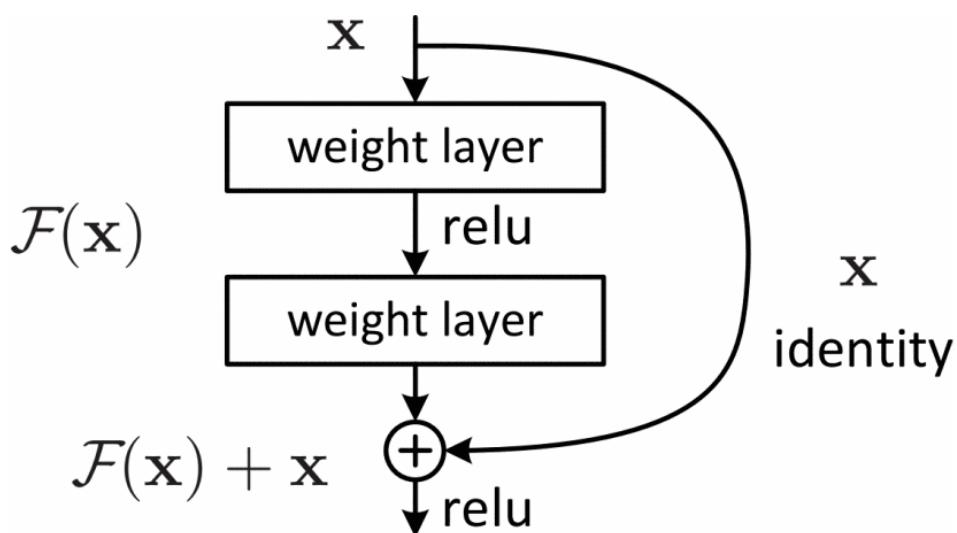


圖 1.18: 殘差學習 **he\_deep\_2016**

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。

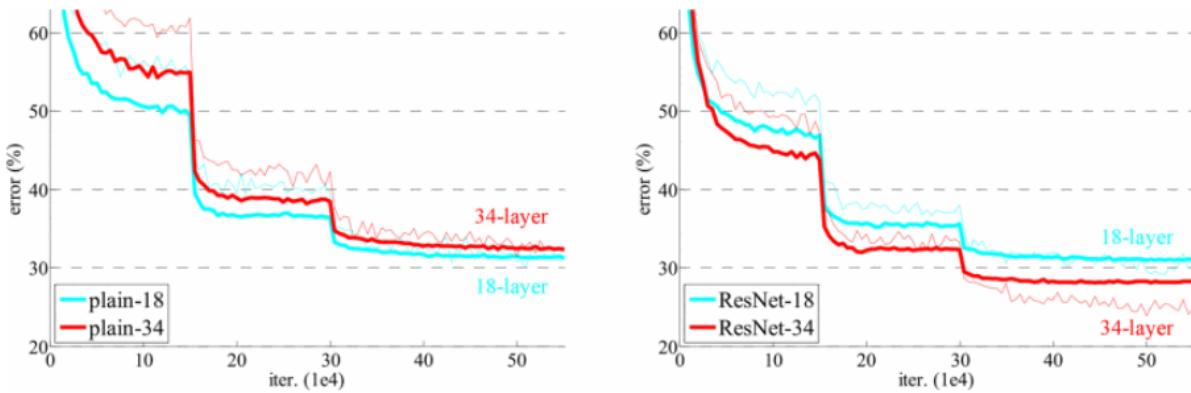


圖 1.19: 訓練 ImageNet 之訓練及驗證誤差曲線圖 he\_deep\_2016

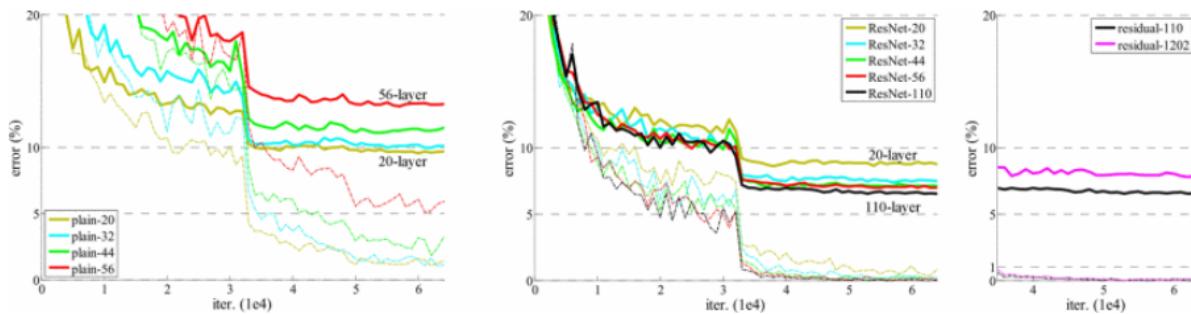


圖 1.20: 訓練 CIFAK-10 之訓練及驗證誤差曲線圖 he\_deep\_2016

## 1.4 人臉偵測演算法

### 1.4.1 MTCNN

MTCNN `zhang_joint_2016` 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，而使用此方法與否之效能差距如圖 1.21。

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 1.23（圖中”MP”為 max pooling，”Conv”為 convolution，池化層及卷積層的步長分別為 2 及 1）：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 nonmaximum suppression (NMS) 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

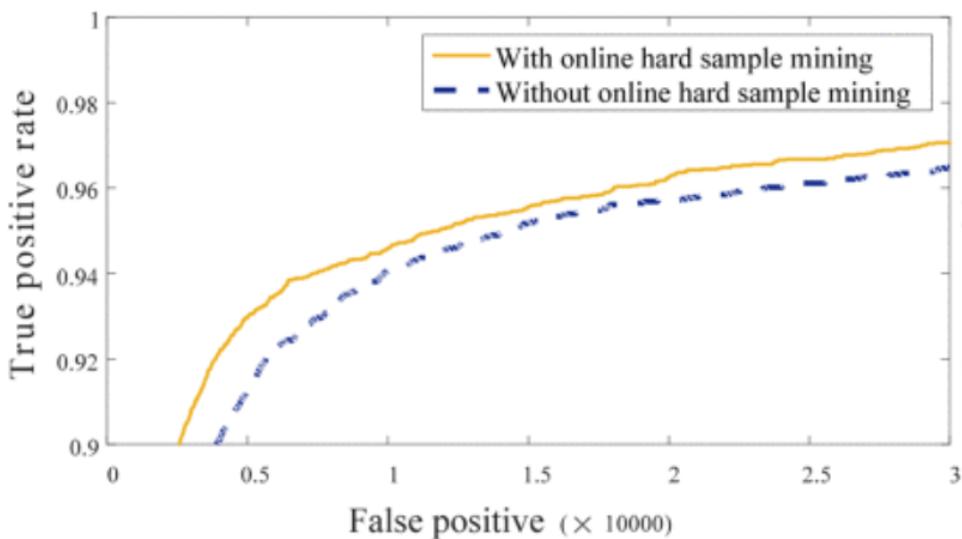
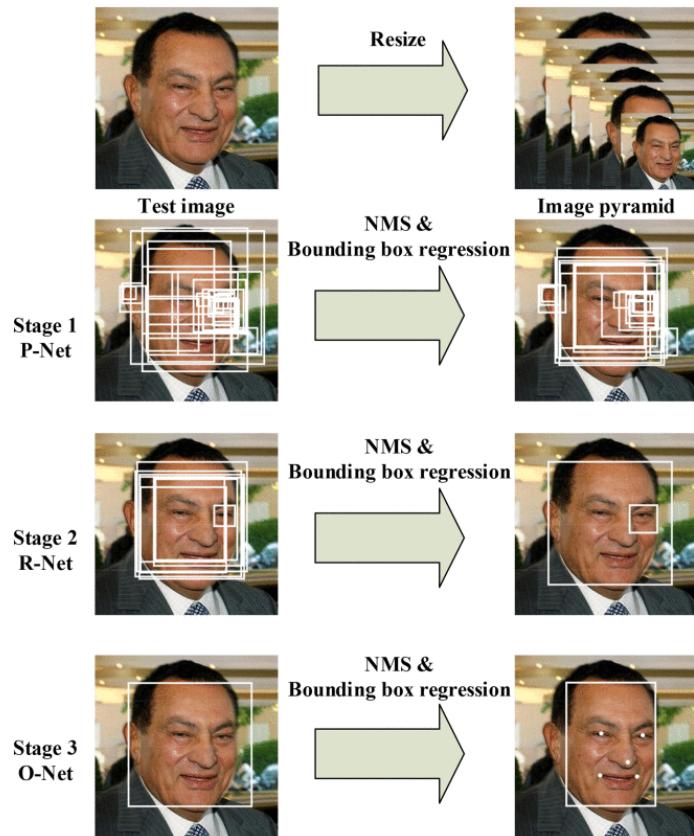
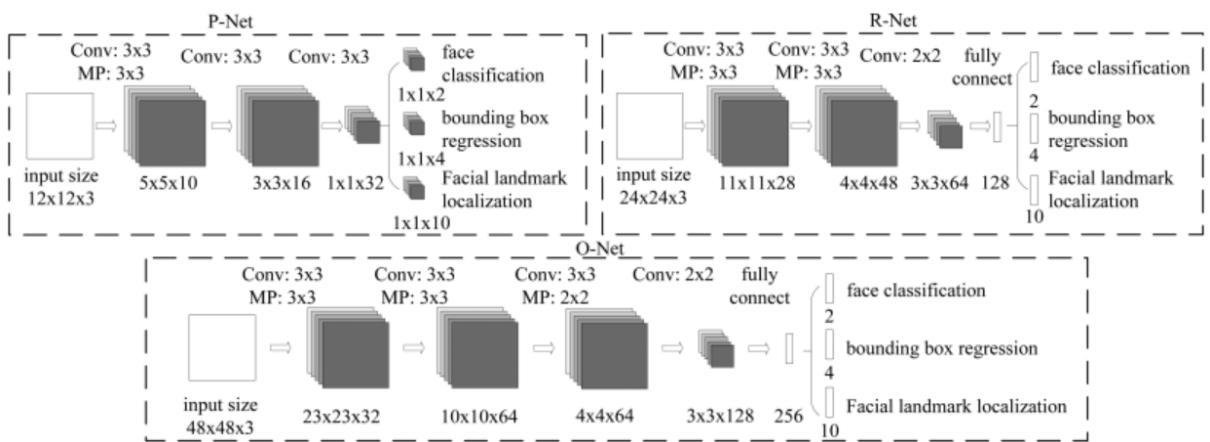


圖 1.21: 使用 online hard sample mining 策略效能比較 `zhang_joint_2016`

network (R-Net) 的 CNN，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 NMS；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 1.22 可看出此網路是以粗到細的方式預測人臉及座標位置。

為了評估人臉偵測的效果，此網路與 FDDB 及 WIDER FACE 中的其他方法 `yang_aggregate_2014`, `mathias_face_2014`, `yan_fastest_2014`, `yang_facial_2015`, `chen_joint_2014`, `li_convolutional_2015`, `yang_convolutional_2015`, `farfade_multi-view_2015`, `yang_wider_2015` 進行比較，結果如圖 1.24，可看出此演算法確實優於其他方法。而為了評估人臉對齊的效果，此網路亦與多種方法進行比較 `cao_face_2014`, `yu_pose-free_2013`, `xiong_supervised_2013`, `burgos-artizzu_robust_2013`, `zhu_face_2012`, `zhang_facial_2014`，結果如圖 1.25，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 1.22: MTCNN pipeline `zhang_joint_2016`圖 1.23: MTCNN framework `zhang_joint_2016`

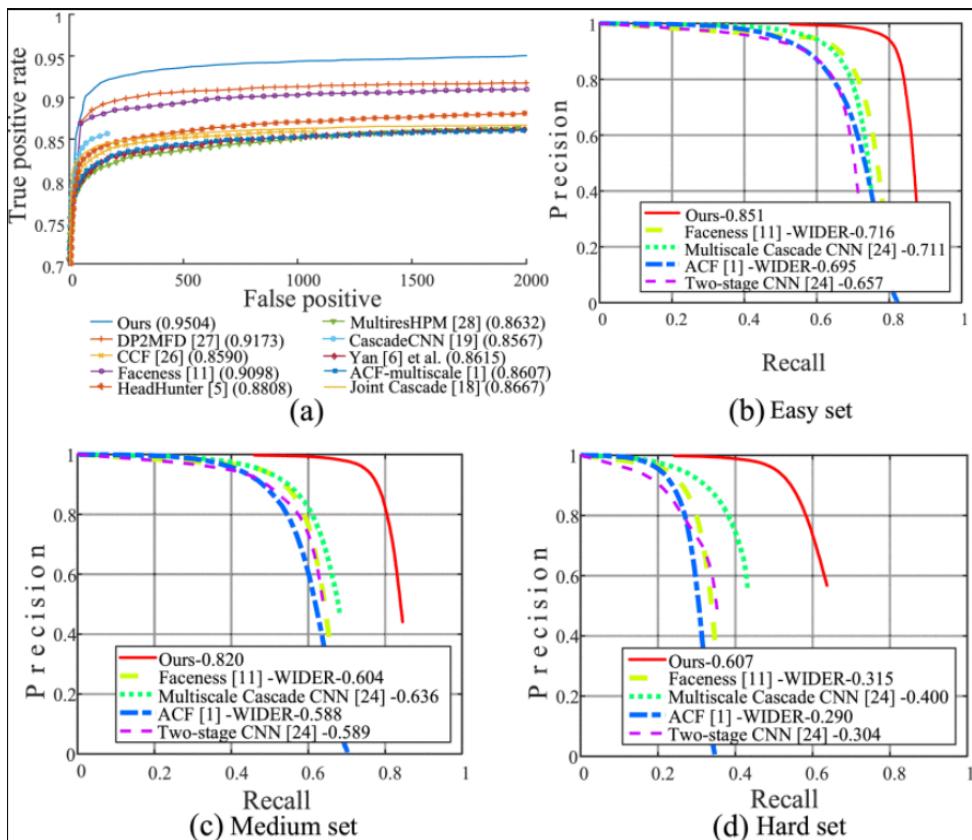


圖 1.24: MTCNN 及其他方法偵測人臉之結果比較 zhang\_joint\_2016

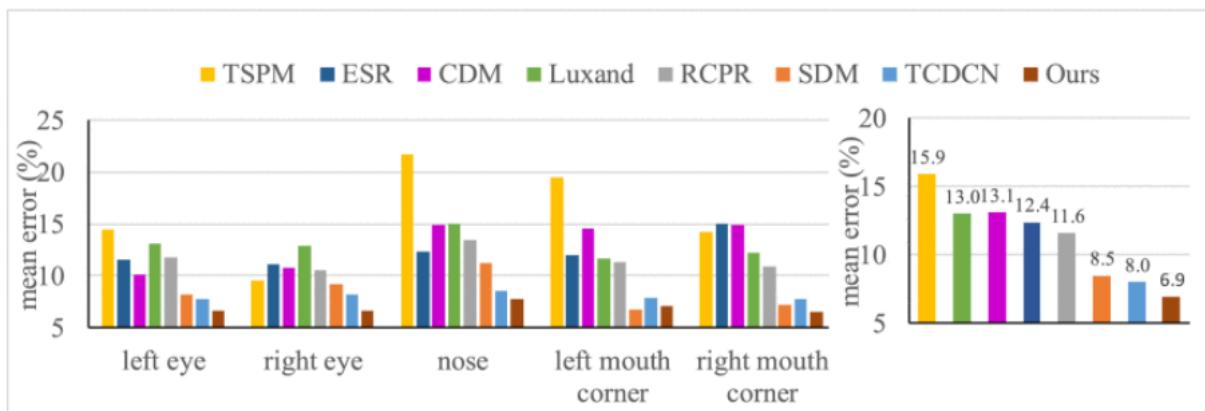


圖 1.25: MTCNN 及其他方法人臉對齊之結果比較 zhang\_joint\_2016

## 1.4.2 RetinaFace

由 Deng 等人於 2020 年提出的 RetinaFace **deng\_retinaface\_2020**，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 1.26）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，feature pyramid network 獲得輸入影像，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。

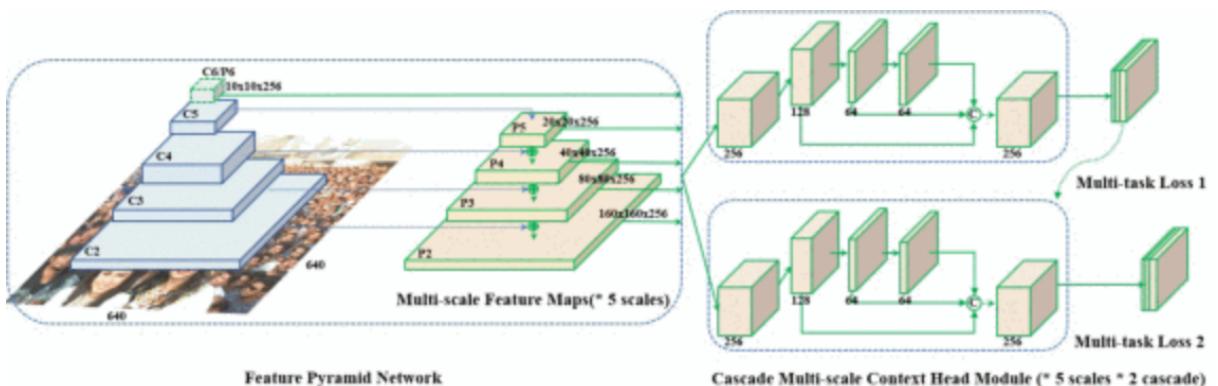


圖 1.26: RetinaFace 架構 **deng\_retinaface\_2020**

本篇論文展示了 RetinaFace 和其他 29 種人臉偵測演算法 **zhang\_single-shot\_2019**, **tang\_pyramidbox\_2018**, **najibi\_ssh\_2017**, **zhang\_s3fd\_2017**, **li\_dsf\_2019** 之平均準確度（Average Precision）比較，如圖 1.27 所示，此演算法擁有 91.7% 的良好結果。

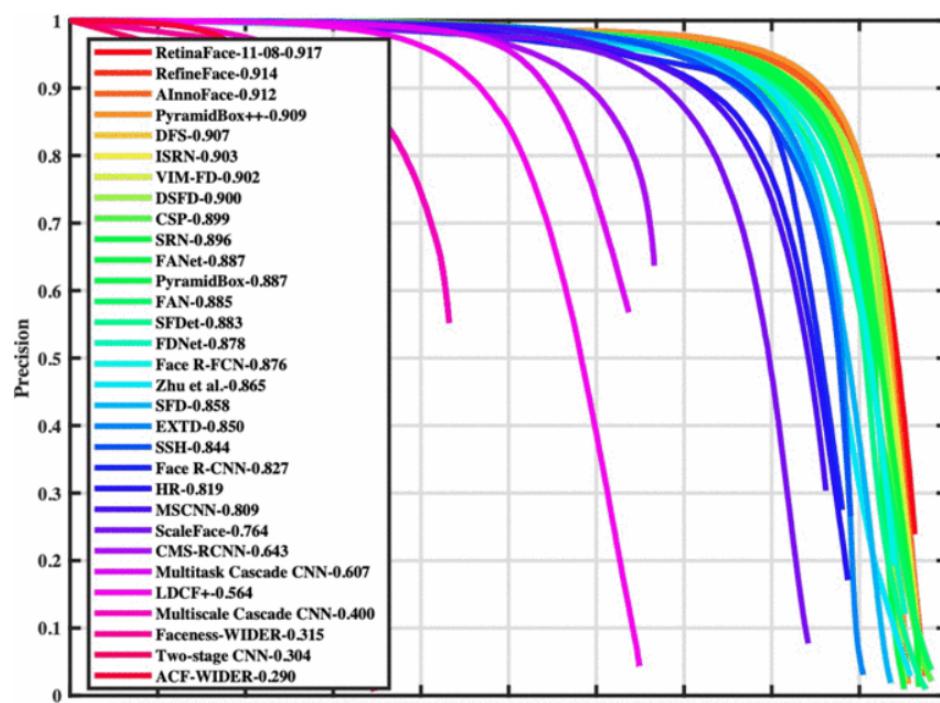


圖 1.27: RetinaFace (ResNet-152) 在 WIDER FACE 測試集之 Precision-Recall 曲線 deng\_retinaface\_2020