

國立中央大學

通訊工程學系
碩士論文

基於時序卷積網路之單 FMCW 雷達應用於非接
觸式即時生命特徵監控

研究生：林以恩

指導教授：張寶基、陳永芳博士

中華民國一一二年六月

基於時序卷積網路之單 FMCW 雷達應用於非接觸式即時生命特徵監控

摘要

生命特徵偵測在醫學和臨床領域中扮演著重要的角色。往年靠著接觸式裝置帶在人體上偵測並分析關鍵的生理參數，包括心率、呼吸率、心電圖等，用來協助醫生判斷病情且做出最適當的治療。在臨床診斷和監護中，生命特徵偵測的應用範圍廣泛且多樣。然而，生命特徵偵測技術也面臨著挑戰和問題。例如，穿戴裝置造成的不適感以及隱私保護等方面需要進一步的改進和研究。於是，近年開始提出非接觸式裝置檢測人體生命特徵，除了能避免長期穿戴醫療設備所帶來得不適外，比起使用 RGB 相機等設備，連續調變波雷達(Frequency modulated continuous waveform radar, FMCW)也能提供有效的隱私保護。

本篇論文的目的為非接觸式的生命特徵監控。我們提出的方法為以時序卷積網路 (Temporal Convolutional Network, TCN) 為主體之模型用以提取生命特徵。TCN 的特點為使卷積神經網路 (Convolutional Neural Network, CNN) 也能具有時序序列建模的能力。搭配具有注意力增強功能的子系統，提供充足有效的訊息，能使模型更有效地監控生命特徵，從而提升準確度。

關鍵字：深度學習、毫米波雷達、時序卷積網路

Non-contact Real-Time vital sign detection Using single FMCW Radar based on Temporal Convolutional Networks

Abstract

Detecting vital signs plays a crucial role in the medical and clinical fields. In the past, vital signs were typically measured and analyzed using contact-based devices, including heart rate monitors, respiratory rate monitors, and electrocardiography (ECG) systems, to assist doctors in diagnosing conditions and making informed treatment decisions. Vital sign detection has a wide range of applications in clinical diagnosis and monitoring. However, there are challenges and issues associated with traditional contact-based methods, such as discomfort caused by wearing devices and concerns regarding privacy protection. As a result, non-contact methods for monitoring vital signs have been proposed in recent years. In addition to avoiding the discomfort of long-term device wear, non-contact methods, particularly Frequency Modulated Continuous Wave (FMCW) radar, offer effective privacy protection compared to devices that utilize RGB cameras.

The objective of this thesis is non-contact monitoring of vital signs. We propose a model primarily based on Temporal Convolutional Networks (TCN) to extract vital sign features. TCN enables Convolutional Neural Networks (CNN) to effectively model temporal sequences. By incorporating a subsystem with attention mechanisms, which provides relevant and informative cues, the model can monitor vital signs more effectively, thereby improving accuracy.

Keywords: Deep Learning, mmWave, FMCW radar, TCN

誌謝

在我作為碩士研究生的旅程中，經歷了無數的波折和轉變，有著平靜和激烈、歡笑和堅毅的時刻。這段過程中，我得到了無數人的幫助，他們是我完成學業和論文研究的不可或缺的支持者。首先，我要深深感謝我的導師張寶基教授，他在我的研究旅程中提供了無私的指導，耐心解答我的各種疑問，分享了寶貴的專業經驗，讓我能夠堅定地面對困難。另一方面，張寶基教授也很重視我們的生活，時常帶著我們外出爬山郊遊，增加彼此之間的感情，亦師亦友。

此外，我也要衷心感謝資工系的王家慶教授，在我的論文研究中提供了重要的支持，啟發我以不同的視角看待問題，增強了我的問題解決能力。

實驗室的同學們是我學術旅程中的摯友，他們的協助和陪伴令我感激不盡。特別要感謝凱雯學姊，她在我的求學過程中給予了許多寶貴的建議和幫助，確保我們順利實現了目標。同樣，我要感謝實驗室的學長姐們，包括彥廷、楓錡、品萱、祐昇，他們在忙於自己的論文研究的同時，仍然願意提供幫助。在我初次進入實驗室時給予了我很多支持，嚴肅對待修課和研究的問題，為我們奠定了良好的基礎，同時也讓我們享受著快樂和歡笑。

同屆的同學們，崇訓、瑤萱，也一直與我並肩前行。我們一同完成了

修課，一同面對計劃中的挑戰，。在我進行論文研究時，我們總是可以共同探討問題，互相幫助。這些人的存在讓我的碩士生涯更加豐富和有趣，我將永遠珍惜他們的友誼和支持。最誠摯的謝意，

目錄

摘要	1
Abstract	2
誌謝	3
目錄	5
圖目錄	7
表目錄	8
第一章 緒論	9
1-1 研究背景	9
1-2 研究動機與目的	11
1-3 論文架構	13
第二章 傳統非接觸式生命特徵偵測方法	14
2-1 FMCW 雷達介紹	16
2-1-1 FMCW 雷達系統架構	16
2-1-2 混波器	18
2-1-3 數位訊號處理器	19
2-2 FMCW 雷達測距原理	19
2-2-1 距離推導	19
2-2-2 Range FFT	20
2-2-3 靜態雜波濾除	20

2-3 相位提取	21
2-3-1 相位關係式	22
2-3-2 相位計算	22
2-3-3 心率及呼吸率估計	24
第三章 深度學習相關介紹	25
3-1 時序卷積網路	26
3-1-1 因果卷積	26
3-1-2 擴張卷積	27
3-1-3 殘差模塊	28
3-2 注意力機制	29
3-2-1 編碼器	31
第四章 提出之架構	34
4-1 系統架構	34
4-2 訓練模型階段	35
4-2-1 資料前處理	36
4-2-2 神經網路模型	39
4-2-3 反向傳播及更新梯度	39
4-3 測試階段	40
第五章 實驗結果與分析討論	41
5-1 實驗環境介紹	41

5-2 實驗結果比較與討論	42
5-2-1 時序卷積網路以不同訊號時長之比較	42
5-2-2 時序卷積網路模型參數之比較	43
第六章 結論與未來展望	44
參考文獻	45

圖目錄

圖 1.FMCW 雷達系統架構	16
圖 2.調變波種類	17
圖 3.因果卷積	26
圖 4.填充示意圖	27
圖 5.擴張卷積	28
圖 6.殘差模塊流程圖	29
圖 7.TRANSFORMER 架構圖	31
圖 8.TRANSFORMER 編碼器	31
圖 9.自注意力機制架構圖	33
圖 10.多頭注意力架構圖	33
圖 11.提出之系統架構圖	34
圖 12.FMCW 雷達	36
圖 13. MASIMO PULSE OXYGEN	37

圖 14.實驗場景圖	38
------------------	----

表目錄

表 1.訓練模型階段參數設定列表	36
表 2.實驗環境軟硬體之配置表	41
表 3.不同訊號時長之心率及呼吸率比較表	42
表 4.模型參數之比較表	43

第一章 緒論

1-1 研究背景

隨著科技的高度發展，人們越來越注重自身的健康，許多有著健康監測的產品推陳出新，受到大家的歡迎。在健康監測的產品裡，最常見的莫屬於心率及呼吸率，兩項數據都是最直接影響到個人健康的重要指標。目前市面上有包含這兩項數據的產品，可以大略區分為接觸式及非接觸式。接觸式裝置通常以穿戴的形式使用，常見的有，智能手錶/智能手環：許多智能手錶和智能手環配備了光學心率監測器，可以實時監測使用者的心率。這些設備通常使用 LED 光源和光學感測器來檢測皮膚上的血液流動。胸帶式心率監測器：這些設備通常是透過佩戴在胸部的帶子來監測心率。胸帶內部嵌入有心率感測器，可以感應到心臟跳動時的電氣信號，並將數據傳送到連接的手機或其他設備上。這種類型的心率監測器通常被運動員和運動愛好者使用。衣物式心率監測器：這些產品通常是以衣物的形式設計，例如運動上衣、運動背心等。它們內部嵌入有心率感測器，可以實時監測使用者的心率。這種設計使得監測更加隱蔽和舒適，適合進行運動或日常使用。即時這些穿戴式裝置已經發展的相當輕便，長時間穿戴能會造成一定程度的不適，尤其是本身有皮膚疾病的人，因此近年來逐漸發展出非接觸式感測器來替代接觸式的穿戴式裝置，減輕長期穿戴造成的影響。目前所知的非接觸式裝置有，無線雷達：利用射頻信號穿透皮膚並反射回來，藉此檢測心跳。雷射測距儀：通

過照射脈搏部位，使用雷射光測量皮膚的微小震動，從而推斷心跳頻率。熱

成像攝影: 熱成像攝像頭可以檢測物體的紅外輻射，並顯示為熱圖像。通過

熱成像攝像頭可以從人的臉部或其他部位捕捉到微小的心跳變化，從而推

斷心率。

1-2 研究動機與目的

在非接觸式感測器中，雷達具有較高的穿透能力，訊號能穿透覆蓋在人體身上的衣物，同時雷達相較於其他產品能進行更遠更久的監控。在雷達感測器中，又以超寬頻(Ultra Wide-Band, UWB)雷達及連續調變波(Frequency Modulated Continuous Waveform, FMCW)雷達[1]最為常見。UWB 雷達使用非常短脈衝(通常短於 1 納秒)的高頻信號來進行雷達掃描。通常採用大帶寬的信號，可以利用許多不同的頻率和脈衝形狀，以及其信號的時域特性來提取目標物的距離、速度和方向等信息。具有高分辨率和精確度的目標檢測和定位能力以及抗干擾能力。然而 UWB 雷達有著功率頻譜密度的大小限制使得量測距離只在約一公尺內有效，導致實際應用上有著眾多限制。因此我們選用了能量測較遠距離的 FMCW 雷達作為實驗感測器。FMCW 雷達是一種連續發射和連續接收的雷達系統，它持續發射一個固定頻率的連續波信號，並通過接收到的反射信號來測量目標物的距離和速度。在實際監控生命特徵時，與心跳及胸壁的移動相比，身體突然的大幅度運動很容易影響雷達信號的反射，再受到身體位移的影響下，很難從中提取出心跳及胸壁的位移。因此，本論文研究的目的為從受到身體位移影響的信號中，將心跳及胸壁的運動頻率提取出來並估計出心率及呼吸率。為了達到此目的，我們以時序卷積網路[2]以及注意力模塊[3]為主體架構模型，將雷達得時域訊號及頻域訊號作為輸入分別放入(Convolutional Layer) 卷積層中提取心率及呼吸率

的特徵，再將兩者輸出的結果串聯放入(Attention Layer)注意力層，由於心跳運動是連續性的，透過注意層參考過去的結果並提高當前的生命特徵。最後經過兩個回歸器分別估計出心率及呼吸率。

1-3 論文架構

本篇論文分成六個章節，第一章為緒論，介紹本論文的研究背景、動機和目的。第二章為傳統非接觸式生命特徵偵測方法，介紹 FMCW 雷達基本原理以及訊號前處理。第三章為深度學習相關介紹，介紹本論文應用的學習網路，包含 TCN、注意力機制(Attention Mechanism) 和回歸器(Regression)。第四章介紹本篇論文所提出之系統架構。第五章為實驗結果的比較與討論。第六章為結論與未來展望。

第二章 傳統非接觸式生命特徵偵測方法

近年來，隨著毫米波和第五代移動通信（5G）的迅速發展，FMCW 雷達技術以其在距離、速度測量以及目標檢測和成像方面的優異表現而受到廣泛關注。在地震監測、智能家居、無人機、車用電子、醫療應用等等各大領域皆有良好的成果，因此在近年成為許多研究者投入研究的主題。

FMCW 雷達可以追溯到 20 世紀早期。最早由一位德國科學家首次提出使用連續發射的高頻訊號(CW Radar)來探測目標物體，二戰時開始發展出頻率調變連續波（FMCW）雷達，使用頻率連續變化的連續波信號，可以同時測量目標的距離和速度。1980 年代至今：隨著科技的不斷發展，FMCW 雷達技術在多個應用領域中取得了突破。在軍事、航空、交通、環境監測、醫療等領域中。同時，隨著微波和毫米波技術的進步，FMCW 雷達的性能和應用領域也在不斷擴展。

連續波雷達可因發射訊號的形式不同，分成都普勒雷達以及調頻連續波雷達：

都普勒雷達基於都普勒效應，用於測量目標物體的速度。當目標物體相對於雷達系統運動時，反射回來的信號的頻率會發生變化。這種頻率變化與目標物體的速度成正比。利用這一原理，通過測量回波信號的頻率變化來計算目標物體的速度。主要用於測量目標物體的速度、運動方向和速度分量等，最常見的測速照相機就是使用這一原理。

調頻連續波雷達發射信號的頻率會隨時間連續變化。當發射信號與反射信號進行混頻後，得到的中頻信號中包含了目標物體的距離信息。通過分析中頻信號的時域和頻域特徵，可以測量目標物體的距離和其他特性。調頻連續波雷達具有較高的距離分辨率和測量精度，並且適用於需要同時測量多個目標物體距離的場景。

本章首先介紹 FMCW 雷達的基本概念和應用[4]。2-1 節介紹 FMCW 雷達系統架構；2-2 節介紹 FMCW 雷達訊號處理；2-3 節介紹生命特徵估計。

2-1 FMCW 雷達介紹

目前所知 FMCW 雷達系統的架構包含訊號產生器、頻率調變器、發射天線、接收天線、放大器、混波器、數位訊號處理器等等。其架構如圖 1 所示。接下來會逐一介紹各個組件的功用。

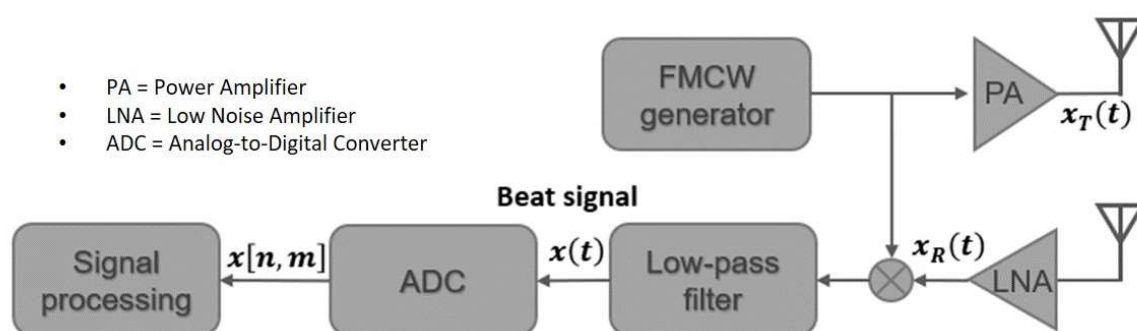


圖 1. FMCW 雷達系統架構

2-1-1 FMCW 雷達系統架構

在 FMCW 雷達中最廣泛應用鎖相迴路（Phase-Locked Loop, PLL）作為訊號產生器。PLL 是一種回授控制系統，通過監測和調整輸出信號的相位，使其與參考訊號保持恆定的相位差。FMCW 雷達系統裡，透過不斷調整 PLL 裡的壓控振盪器(Voltage-Controlled Oscillator, VCO) 的頻率來生成頻率連續調變的信號，使雷達能夠提供高分辨率的距離測量和測速功能。此外，經由此方法所發出一次完整的頻率調變訊號，稱之為啁啾(chirp)。

PLL 能產生出三種不同調變的連續電池波，其發射頻率皆隨著時間呈線性變化，分別有:三角波調變、鋸齒波調變以及正弦波調變。產生好的訊

號會經由發射天線傳出，天線具有指向性，訊號會朝天線所指的方向成圓錐狀發射。為了確保能將訊號傳到更遠的距離，還會經過功率放大器將訊號功率增強，增強後的訊號打到物體後再反射回雷達。

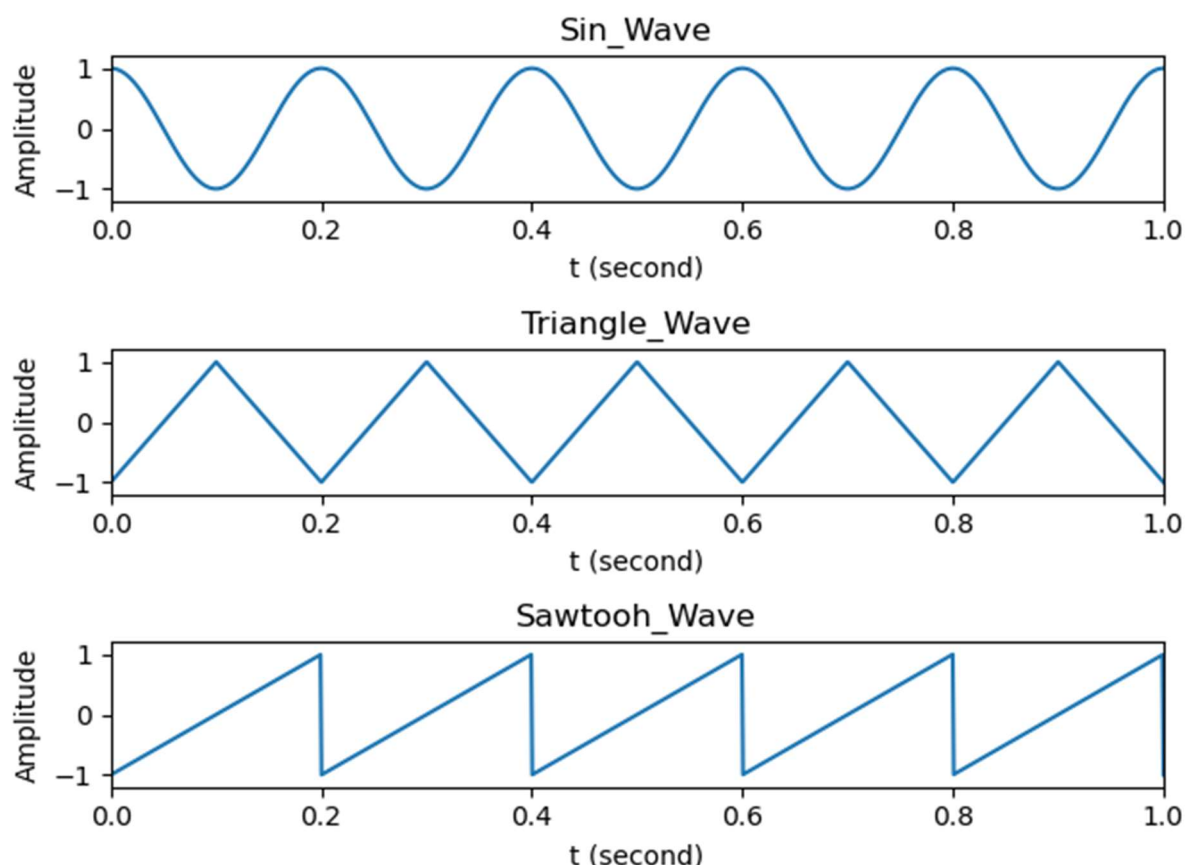


圖 2. 調變波種類

反射後的雷達訊號藉由接收天線回收，接收到的信號包含了目標物體反射回來的能量，以及可能存在的干擾和雜訊。為了提取出目標信號並抑制干擾，通常利用低雜訊放大器（LNA）來加強訊號強度。發射信號的表示如下：

$$x_T(t) = A_T \cos(2\pi f_T(t)t + \phi_T(t)) \quad (2-1)$$

其中 A_T 為發射的傳輸能量大小， $f_T(t)$ 為發射隨時間變化的頻率。接收信號

的表示如下：

$$x_R(t) = A_R \cos(2\pi f_R(t)t + \phi_R(t)) \quad (2-2)$$

A_R 為接收的傳輸能量大小， $f_R(t)$ 為接收隨時間變化的頻率。

2-1-2 混波器

FMCW 雷達訊號發射及接收的訊號使用頻段落在 60~80GHz，屬於毫米波的範圍。在訊號處理上，高頻訊號相較基頻訊號有更高的採樣率及更短的時間間隔，運算較複雜且時間也較久。因此，透過混頻器，我們能將 LNA 加強後的訊號與發射訊號做混頻處理，經過混頻後的訊號稱為拍頻訊號，其表示如下：

$$x_T(t) \cdot x_R(t) = A_T \cos(2\pi f_T(t)t + \phi_T(t)) \cdot \quad (2-3)$$

$$A_R \cos(2\pi f_R(t)t + \phi_R(t))$$

將混頻完的訊號取低頻部分，即為我們所需的基頻訊號。為達到此目的，只需加上低頻濾波器即可，低頻濾波後的訊號結果如下：

$$\begin{aligned} x(t) &= \frac{1}{2} A_T A_R \cdot \cos(2\pi(f_T(t) - f_R(t))t + (\phi_T(t) - \phi_R(t))) \quad (2-4) \\ &= A \cos(2\pi f_b(t)t + \phi_b(t)) \end{aligned}$$

其中 $f_b(t)$ 為拍頻訊號頻率隨時間的變化， $\phi_b(t)$ 為拍頻訊號相位隨時間的變化。

2-1-3 數位訊號處理器

在做數位訊號處理之前，需先將上述的拍頻類比訊號透過取樣轉取數位訊號。

$$x[n, m] = A \cos \left(2\pi f_b (nT_f + mT_s) nT_f + \phi_b (nT_f + mT_s) \right) \quad (2-5)$$

拍頻訊號的取樣分兩種時間，一個是針對每個 chirp 裡的時間做取樣，一個 chirp 裡的時間相較於每個 chirp 的時間較短，所以稱為快速時間(Fast-time) 用 T_f 表示;第二個則是針對不同 chirp 之間的時間間隔做取樣，此時間與 chirp 內的時間相比較久，我們稱為慢速時間(Slow-time)用 T_s 表示

轉換成數位訊號以後，便可以透過快速傅立葉轉換(Fast Fourier Transform, FFT)及後續處理計算出目標的距離以及相位資訊。

2-2 FMCW 雷達測距原理

上一節介紹了 FMCW 雷達如何發射及接收訊號，並將接收到的訊號經過混頻器轉成拍頻訊號。本節將繼續說明如何從拍頻訊號中取得目標距離資訊。

2-2-1 距離推導

FMCW 雷達發射訊號的頻率會隨著時間變化，假設一個 chirp 的起始頻率為 f_s 、截止頻率為 f_e 、週期為 T_s 、頻寬為 B 、斜率為 S ，則可以表示成以下式子：

$$B = f_e - f_s \quad (2-6)$$

$$S = \frac{B}{T_s} \quad (2-7)$$

毫米波以光速 C 傳遞，因此可以透過前面得到的拍頻訊號經由下式推導出距離 $d(t)$

$$t_d = \frac{2d(t)}{C} \quad (2-8)$$

$$f_b = S t_b = \frac{B}{T_s} \cdot \frac{2d(t)}{C} \quad (2-9)$$

2-2-2 Range FFT

上節介紹了 f_b 與 $d(t)$ 之間的關係，這節將介紹如何利用一維快速傅立葉轉換從拍頻訊號中取得 f_b 。在 2-1-4 節我們得到經取樣過的拍頻訊號後，將每一個快速時間取樣點做 Range FFT，經過 FFT 的表示如下：

$$X_m[K] = \sum_{n=0}^{N-1} x[n, m] e^{-j \frac{2\pi n k}{N}}, k = 0, \dots, N-1 \quad (2-10)$$

其中 K 為 FFT 指標(index)， N 為每個 chirp 裡的取樣數。當第 K 個達到最大值時，即可透過距離分辨率換算物體距離。其換算的公式如下：

$$d = \frac{t_d \cdot C}{2} = f_b T_c \cdot \frac{C}{2B} \quad (2-11)$$

其中 $C = 3 \times 10^8$ 為光速， B 為 chirp 之頻寬。 $\frac{C}{2B}$ 定義為距離分辨率。

2-2-3 靜態雜波濾除

使用雷達偵測生命特徵時，不單單只會接收到從人體身上反射的訊號，環境背景裡的物件，例如：桌子、床、牆壁…等等，也會經由反射被

雷達所接收。這些靜態的背景物件訊號相較於動態的生命特徵訊號還要來的大，以致於我們感興趣的動態訊號被靜態訊號蓋過。為了解決此問題，需進行靜態雜波濾除(Clutter removal)。

靜態訊號不會隨著時間變化，利用此特性我們能透過平滑處理將動態訊號造成的峰值變化濾除，剩下的靜態訊號圖稱為靜態雜波圖(Clutter Map)。平滑處理表示式如下：

$$X'_m[k] = \sigma X_m[k] + (1 - \sigma)X_{m-1}[k], \quad 0 \leq \sigma \leq 1 \quad (2-12)$$

其中 σ 為平滑系數，介於 0 到 1 之間。 σ 越大，平滑效果越好，但較多的目標訊號也會被平滑掉；反之亦然， σ 越小，平滑效果越差，較多的靜態雜波會被保留，使得靜態雜波濾除的效果不好。

最後，將平滑處理後的靜態雜波圖與原本未經處理的訊號相減，便能得到濾除靜態雜波後的結果。其表示如下：

$$Y_m[K] = X_m[k] - X'_m[k] \quad (2-13)$$

此時的 $Y_m[K]$ 訊號已經濾除了大部分靜態雜波造成的峰值，剩下的峰值多為動態物體所造成的。從這些 $Y_m[K]$ 找出最大值的頻率，結合 2-2-2 節推導的距離分辨率(式 1-1)便能計算出物體距離，其表示如下：

$$k_{max} = \arg_k \max |Y_m[K]| \quad (2-14)$$

$$d = f_b T_c \cdot \frac{c}{2B} = k_{max} \cdot \frac{c}{2B}$$

2-3 相位提取

2-3-1 相位關係式

拍頻訊號裡除了包含距離資訊外，還藏有心跳及呼吸的胸壁振幅資訊。因為心跳及呼吸而造成的胸壁運動振幅會隨著時間改變，藉由 2-1-2 節所推導的拍頻訊號公式，可以表示成以下：

$$\begin{aligned} x(t) &= A \cos(2\pi f_b(t)t + \phi_b(t)) \\ &= A \cos\left(2\pi\left(\frac{2B}{T_s C} \cdot (d_0 + \Delta d(t))\right)t + (\phi_0 + \Delta\phi(t))\right) \end{aligned} \quad (2-15)$$

其中 d_0 及 ϕ_0 為胸壁距離雷達的初始位置及初始相位， $\Delta d(t)$ 及 $\Delta\phi(t)$ 為心跳及呼吸而造成的胸壁運動振幅與相位。此外，還可將 $\Delta d(t)$ 表示成以下：

$$\Delta d(t) = A_{hr} \sin(2\pi f_{hr} t) + A_{rr} \sin(2\pi f_{rr} t) \quad (2-16)$$

A_{hr} 與 A_{rr} 分別為心跳及呼吸之振幅， f_{hr} 與 f_{rr} 分別為心跳頻率及呼吸頻率。

人體的心跳及呼吸造成的胸壁起伏非常微小，約為 1~2 公分。這樣的位移小於 FMCW 雷達距離分辨率，因此無法有效測量到 $\Delta d(t)$ ，因此需透過拍頻訊號的相位差來提取心跳及呼吸頻率。其關係如下所示：

$$\Delta\phi(t) = 4\pi \frac{\Delta d(t)}{\lambda} \quad (2-17)$$

2-3-2 相位計算

根據 2-2-3 節找到的 k_{max} 可以取得胸壁的位置。若想取得其相位資訊，需取出 $Y_m[k_{max}]$ 的實部及虛部做 \arctan ，其表示如下：

$$\phi_m = \arctan\left(\frac{\text{Im}(Y_m[k_{max}])}{\text{Re}(Y_m[k_{max}])}\right) \quad (2-18)$$

需要注意的是，上式(2-18)的相位只會在 $-\pi \sim \pi$ 之間變化，因此兩個連續相位差的結果不能超過 $-\pi \sim \pi$ 之間，超過 $-\pi \sim \pi$ 之間的現象稱為相位落差(Phase Jump)，是不可接受的範圍。解決相位落差需要運用到相位展開技術(Phase Unwrapping)，當計算出超過 $-\pi \sim \pi$ 之間的相位差時，補上 $-2\pi \sim 2\pi$ 的落差值，將相位差範圍限制在 $-\pi \sim \pi$ 之間，其表示如下：

$$\phi_{b,m} = \begin{cases} \phi_m + 2\pi, & \phi_m - \phi_{m-1} > \pi \\ \phi_m - 2\pi, & \phi_m - \phi_{m-1} < -\pi \\ \phi_m, & \text{otherwise} \end{cases} \quad (2-19)$$

取得相位展開後的相位差值，可以與式(2-16)結合成以下：

$$\phi_{b,m} = \frac{4\pi}{\lambda} [A_{hr} \sin(2\pi f_{hr} t) + A_{rr} \sin(2\pi f_{rr} t)] \quad (2-20)$$

此時的 $\phi_{b,m}$ 為胸壁造成的相位差，其包含了呼吸及心率兩種相位，如何有效的分離兩種相位是生命特徵監控的一大考題。我參考[4]提出的方法，該作者串聯兩個雙二階的 IIR 濾波器，每個雙二階的 IIR 濾波器的結構為直接 II 階型(Direct-form II)，其差分方程如下：

$$\begin{aligned} y[n] &= b_0 \omega[n] + b_1 \omega[n-1] + b_2 \omega[n-2] \\ \text{where } \omega[n] &= x[n] - a_1 \omega[n-1] - a_2 \omega[n-2] \end{aligned} \quad (2-21)$$

該作者只針對呼吸率做研究，將截止頻率設定為頻率(0.1-0.6 Hz)較符合人體呼吸的合理範圍頻率，則可得到呼吸相位 $\Phi_{rr,m}$ ；同理，也能將截止頻率設定為頻率(0.8-2 Hz)較符合人體心跳的合理範圍頻率，也能得到心率相位 $\Phi_{hr,m}$ 。

2-3-3 心率及呼吸率估計

在呼吸率估計上，[4]蒐集一段時間的 $\Phi_{rr,m}$ ，並放入長度為 L 的循環暫存器(Circular Buffer)內，確保每次取樣皆能更新呼吸率。將循環暫存器內所有相位做 L 點的 FFT 即可得到呼吸頻率 f_{rr} ，其表示如下：

$$\Phi_{m,rr}[l] = \sum_{n=0}^{L-1} \Phi[n] e^{-j\frac{2\pi nl}{L}}, l = 0, \dots, L-1 \quad (2-22)$$
$$f_{rr} = \text{arc}_l \max |\Phi_{m,rr}[l]| \cdot \frac{1}{T_s L}$$

在心率估計上，則同理呼吸率，其表示如下：

$$\Phi_{m,hr}[l] = \sum_{n=0}^{L-1} \Phi[n] e^{-j\frac{2\pi nl}{L}}, l = 0, \dots, L-1 \quad (2-23)$$
$$f_{hr} = \text{arc}_l \max |\Phi_{m,hr}[l]| \cdot \frac{1}{T_s L}$$

第三章 深度學習相關介紹

科技不斷進步，世界上不斷涌現各種研究技術，同時過去的技術也在不斷改進。人工智慧技術的發展對科技進步有著重要的推動作用。越來越多不同領域的學者和專家紛紛引入人工智慧技術，以改善或解決問題，許多研究取得了良好的成果，因此成為近年來許多研究者投入研究的熱門主題。

人工智慧相關技術在 1950 年代開始逐步被開發出來，其概念啟發自生物的神經系統。這些技術模仿神經系統的結構，由許多神經元組成，構建類似神經元的單元，然後將大量單元組織成神經網絡，以模擬生物神經系統的架構。研究人員期望這些神經網絡模型能具有類似於人類大腦的功能。然而，由於早期電腦硬體的計算能力有限，神經網絡的計算需要大量的運算和時間成本，因此在發展上受到較大的限制。然而，近年來，隨著電腦計算能力的提升，這類計算以及更深層和非線性的深度學習相關研究和模型訓練變得可行且越來越普及，許多研究如雨後春筍般湧現出來。

本章節將介紹本論文使用的深度學習技術，以時序卷積網路及自注意力為主。3-1 節介紹時序卷積網路；3-2 節介紹注意力機制；3-3 節介紹多頭回歸器。

3-1 時序卷積網路

以往序列訓練的任務，最常用的便是遞迴神經網路(recurrent neural networks, RNN)及其變種 LSTM(Long short-term memory)、GRU(Gated Recurrent Unit)等。即使這些模型在序列訓練的任務有很好的表現，但仍存在著一些問題。例如:梯度爆炸/消失或需使用大量的暫存記憶體。基於上述問題，Shaojie Bai[2]等人便提出了時序卷積網路(Temporal Convolutional Network, TCN)，是一種具有相同輸入及輸出長度的擴張(Dilation)、因果(Causal)的一維卷積層(1D Convolution)組成。作者展示了 TCN 解決了 RNN 的缺點，相較於 LSTM 與 GRU 也有著更好的表現。

3-1-1 因果卷積

因果卷積(Causal Convolutions)使卷積神經網路(Convolution Neural Network, CNN)時序特徵，所謂因果即對於 t 時刻的輸出 y_t 來說，其輸入只能為 t 或 t 以前的時刻，如 $x_0, x_1 \dots x_t$ ，其表示如下圖 3:

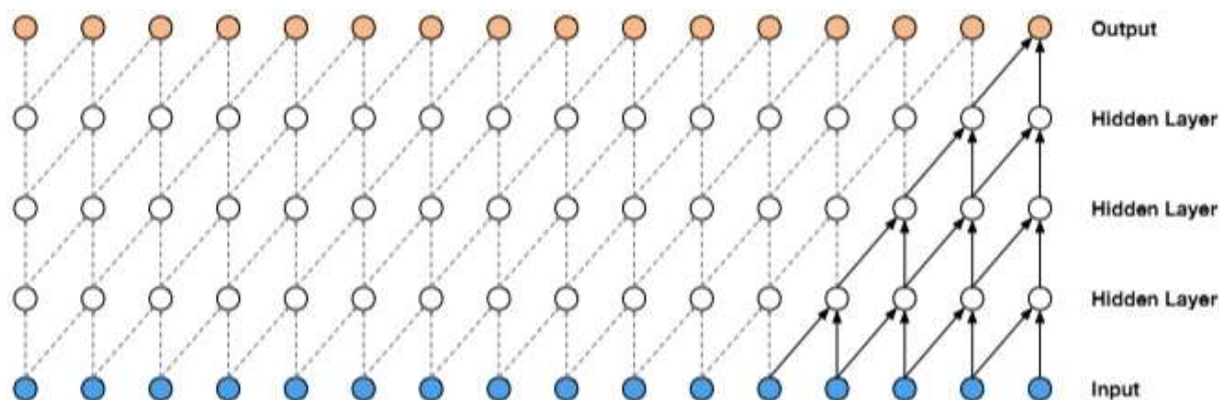


圖 3.因果卷積

每個時刻的輸入會和前一刻的輸入做卷積傳輸到下一層。為了確保每一層輸入與輸出有著相同的長度，我們需在最後一個時刻後面補上數值，這個動作稱為填充(Padding)，通常會補上 0。圖 4 為填充的示意圖。

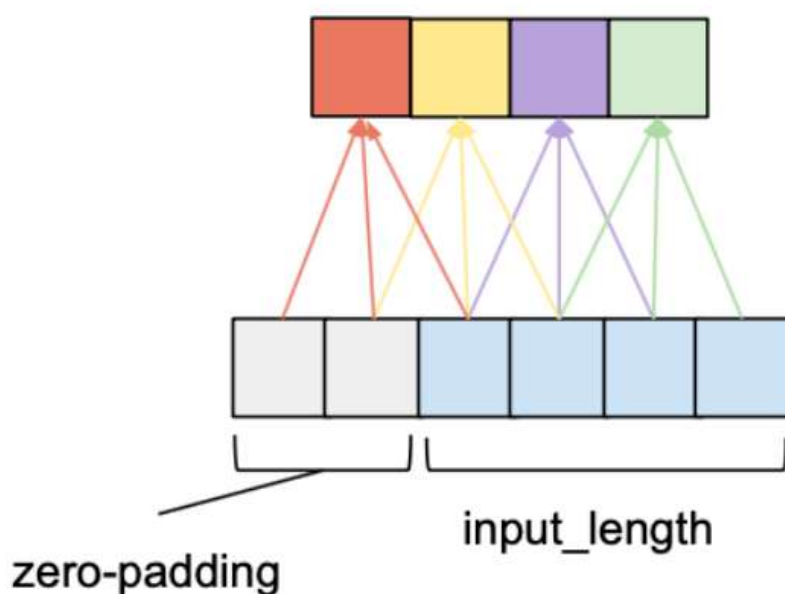


圖 4.填充示意圖

通過觀察圖 2 可以發現，如果一次只與前一刻的輸入做卷積，但輸入特徵可能與較遠的時刻有關連，要取得聯繫需加深 TCN 的深度。考慮到太遠的時刻，則會使 TCN 的層數太過龐大，因此可以利用膨脹卷積(Dilated Convolutions)來解決此問題。

3-1-2 擴張卷積

為了有效聯繫較深較遠的特徵，作者引入了擴張因子 d (dilation factor)。可以透過控制該因子來調整相鄰多少間隔的時段做卷積。

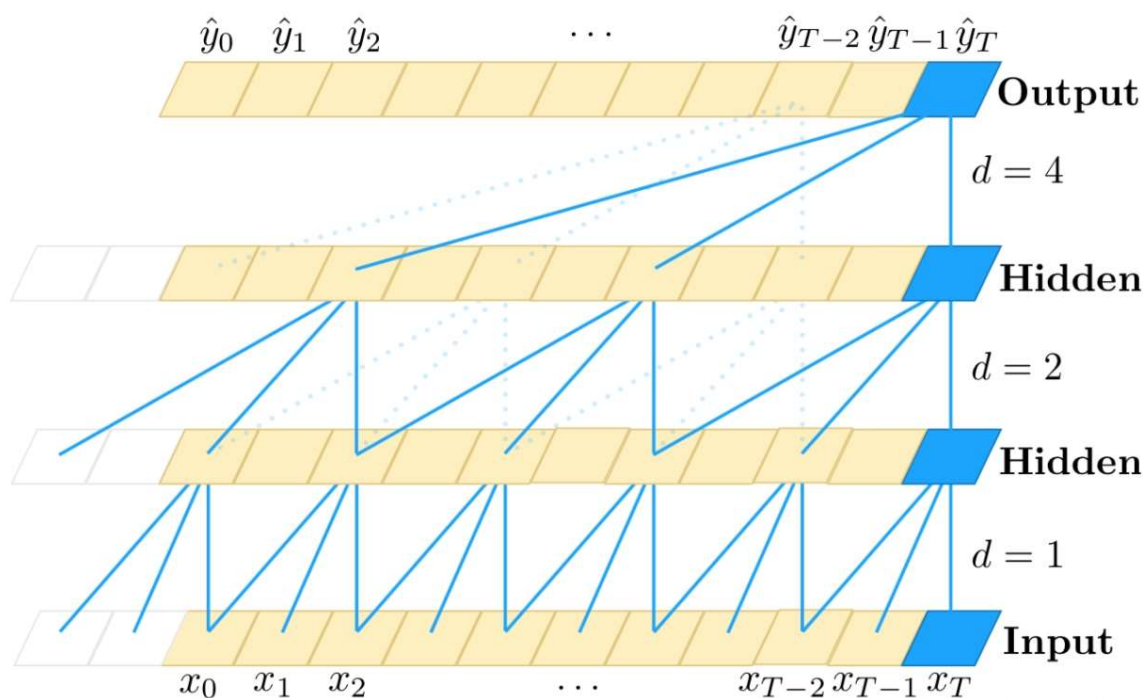


圖 5.擴張卷積

$$d = 2^{i-1}, i = 0, 1, \dots, n \quad (3-1)$$

根據式 3-1，將 d 設計為 2 的指數次方遞增，隨著層數增加卷積的間隔也會變大，如圖 5 所示。較深的層所產生的空洞也越大，所以也可以稱為空洞卷積。透過擴張卷積可以大幅減少 TCN 層數的同時也能取得較長訊號的特徵資訊。

3-1-3 殘差模塊

即使使用了擴張卷積使 TCN 的層數大幅降低，但仍然會需要較深模型的訓練任務，然而較深的網路結構可能引起梯度消失/爆炸等問題存在，為了避免梯度消失/爆炸，作者加入了類似於 ResNet 中的殘差模塊(Residual block)，使 TCN 架構具有更好的泛化能力。

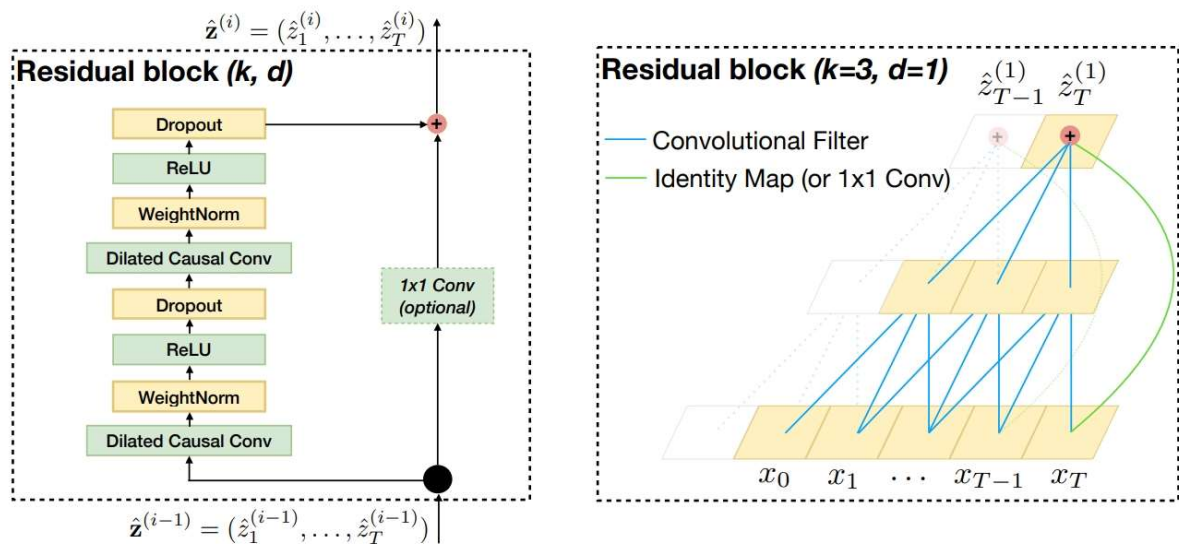


圖 6. 殘差模塊流程圖

圖 5 為殘差模塊的架構圖。殘差模塊取代了原本 TCN 各層之間簡單的連接，加入了權重歸一化 WeightNorm、激活函數 ReLU、隨機神經元關閉 Dropout 三種模塊，這些模塊在深度學習裡常常被使用來解決梯度消失/爆炸，因此透過殘差模塊的加入，可以使較深層的 TCN 模型也能避免梯度消失/爆炸的問題。

3-2 注意力機制

注意力機制 (Attention mechanism)最早可以追溯到 90 年代的視覺圖像領域上，近期才被應用在深度學習的領域上，主要用來改善序列問題。序列學習的最大問題是當訊息過長時，容易丟失訊息內容，因此提出注意力機制來解決此問題。舉例來說，在一段話或文章裡真正重要的關鍵字可能在前段及後段，傳統的遞迴神經網路(Recurrent Neural Networks, RNN)無法將前後的關鍵字連接在一起。而 Attention 可以透過時間權重來注意每個單

詞彼此之間的關聯性，幫助模型找出值得注意的地方，使輸出的結果更準確，也因此命名為注意力機制。然而，2017 年發布一篇 Attention Is All You Need [3]，提出了基於 Sequence to Sequence (Seq2Seq) + 自注意力機制 (self-attention mechanism) 並命名為 Transformer。接下來我們將著重 Encoder 的部分介紹。

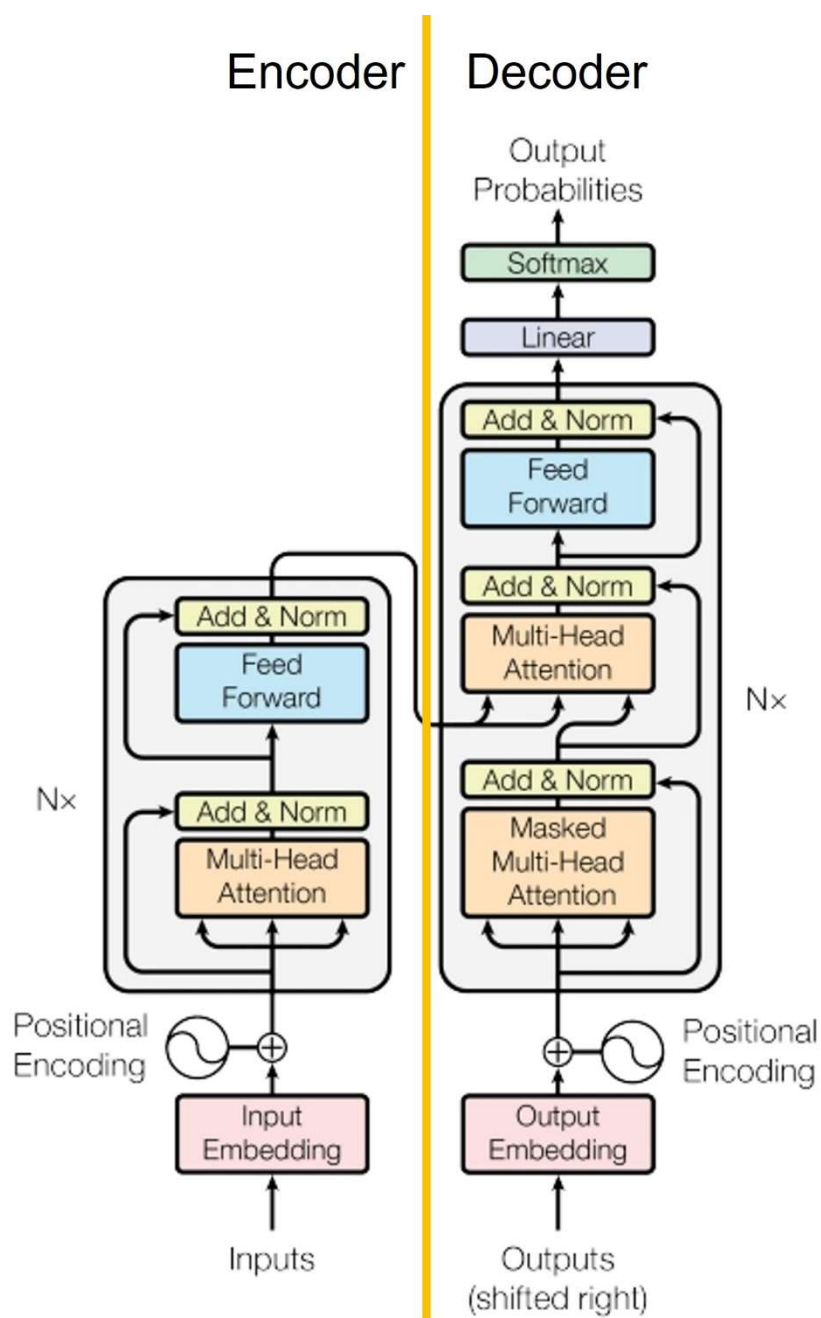


圖 7.Transformer 架構圖

3-2-1 編碼器

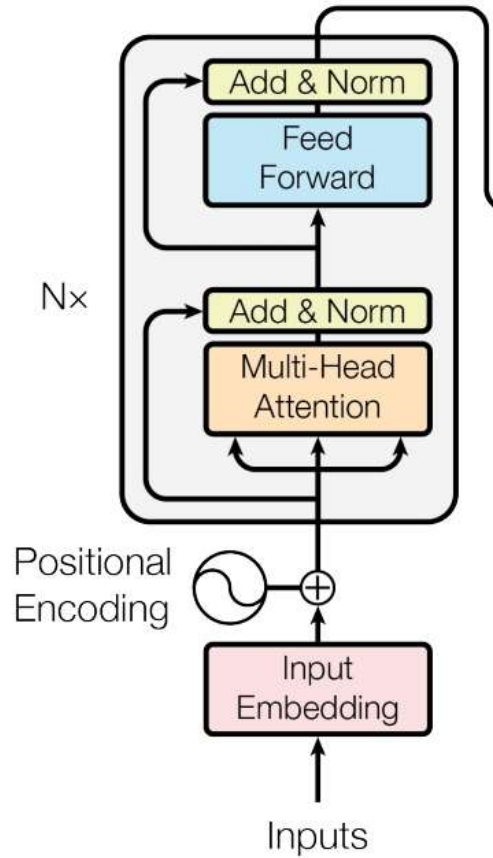


圖 8.Transformer 編碼器

圖 8 為 Transformer 的編碼器(Encoder)，主要分成兩個部分:位置編碼(Positional Encoding, PE)以及自注意力機制模塊(Self-Attention mechanism)。輸入會先經過嵌入層(Embedding layer)轉換成向量作為位置編碼的輸入，位置編碼的目的在於讓模型考慮單詞之間的順序。其表示如下:

$$PE_{(pos,2i)} = \sin\left(\frac{pos}{10000^{\frac{2i}{d_{model}}}}\right) \quad (3-2)$$

$$PE_{(pos,2i+1)} = \cos\left(\frac{pos}{10000^{\frac{2i}{d_{model}}}}\right)$$

其中 pos 為單詞在序列中的位置， $2i, 2i + 1$ 為該單詞的指標(index)， d_{model} 為位置編碼的維度，通常與遷入層轉換出來的向量維度相同。編完位置編碼後，自注意力機制會將每個單詞分別乘上三個權重生成出三種參數。乘上 W_q 可得參數 q (query)，其代表當前的單詞向量，用來對每個 key 做匹配程度的打分；乘上 W_k 可得參數 k (key)，其代表序列中所有的單詞向量；乘上 W_v 可得參數 v (value)，其代表實際的序列內容。將每個單詞分別與序列中其他的單詞向量做匹配打分，找出彼此的關聯性，其表示如下：

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

其中除以 $\sqrt{d_k}$ 是避免 QK^T 的值過大使梯度不穩定。

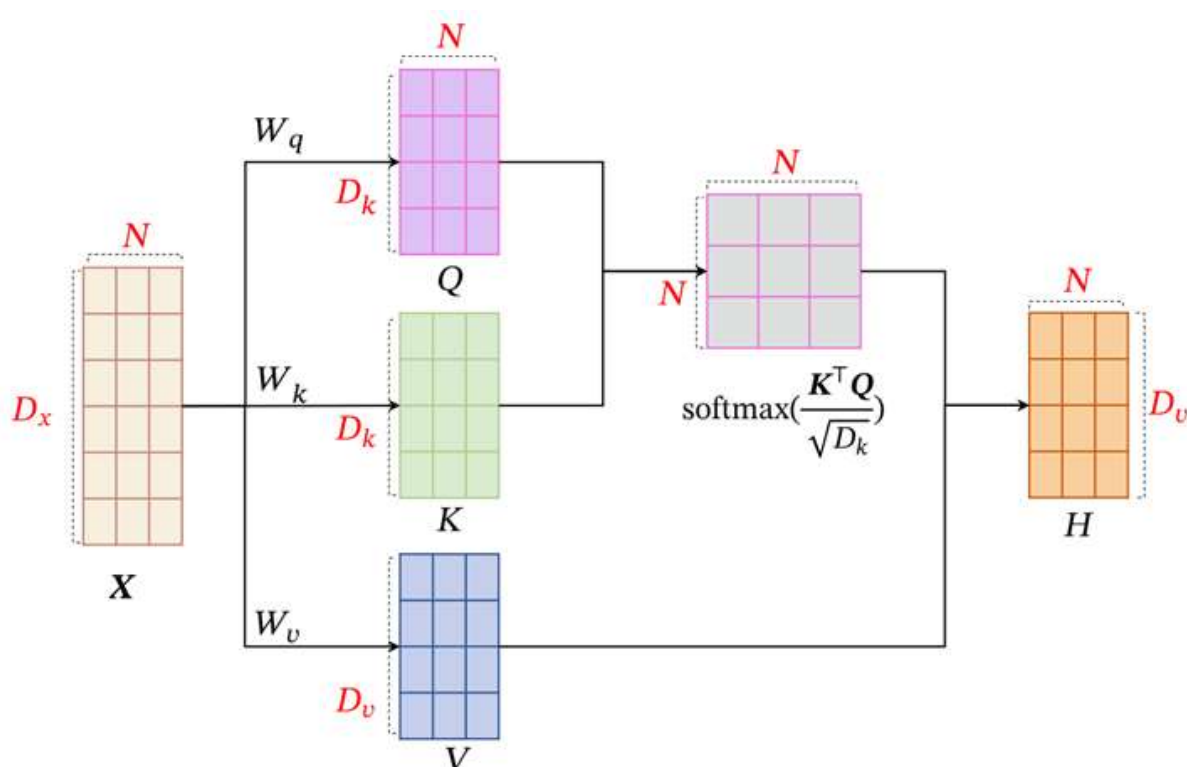


圖 9.自注意力機制架構圖

多頭注意力機制(Multi-Head Attention)運算方式皆與自注意力機制相同，差別在於多頭注意力機制會先將參數 q, k, v 拆分成多個低維度向量，最後再把這些輸出串連起來做一次線性計算，然而這樣的好處是能夠利用多頭的特性關注不同的資訊，例如有的關注局部資訊有的關注全局資訊等。

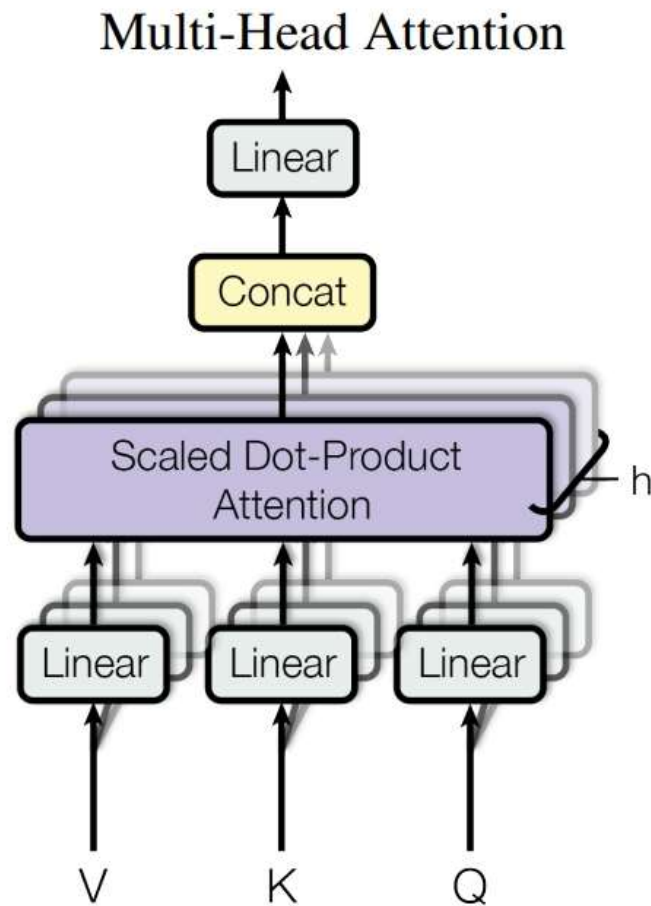


圖 10.多頭注意力架構圖

第四章 提出之架構

在本章節中將介紹我們所提出的系統架構，以時序卷基網路為主體提取生命特徵。其中 4-1 節介紹所提出之系統架構，在 4-2 節介紹訓練模型階段，4-3 節介紹測試模型階段

4-1 系統架構

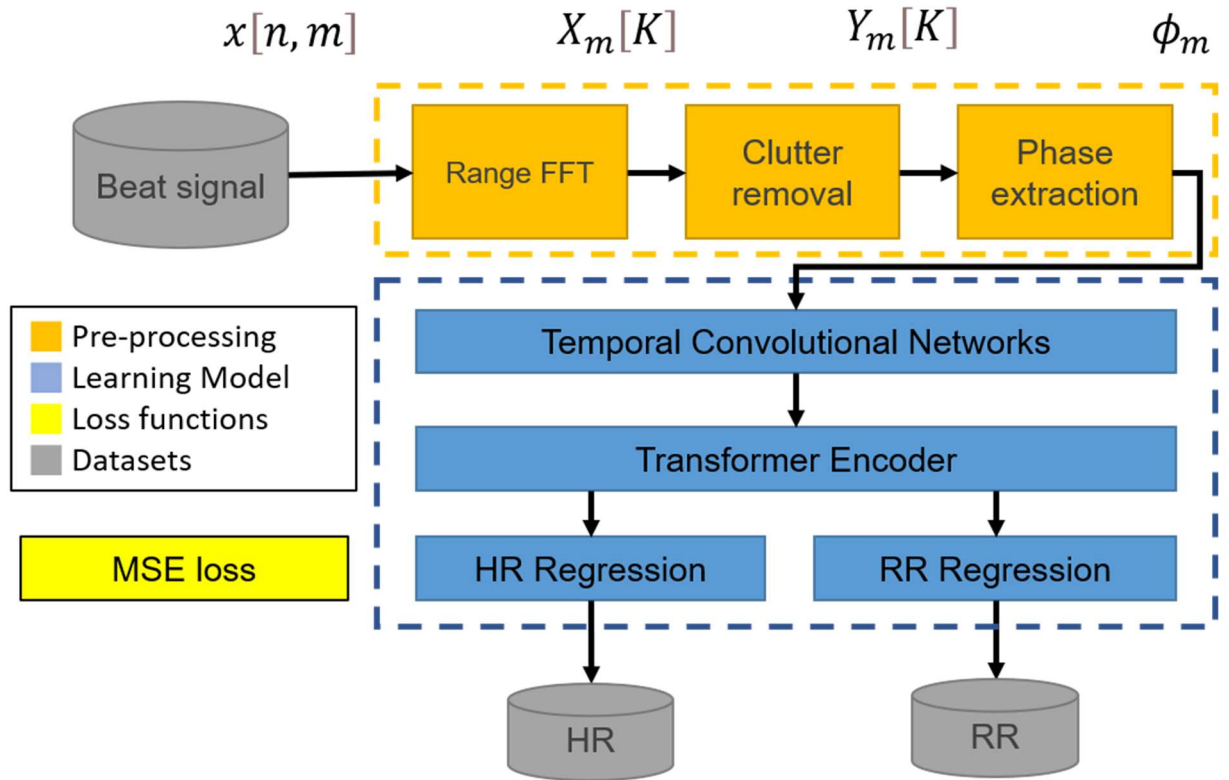


圖 11. 提出之系統架構圖

上圖展示了本篇論文所提出的整體系統架構。在此系統架構中分成兩個階段:訓練模型階段與測試模型階段。在訓練模型階段中，依照順序主要分為資料前處理(Data Pre-processing)，神經網路模型(Neural Network Module)，反向傳播(Backpropagation)及更新梯度(Gradient)。將接收到的雷達訊號經

由前處理去除雜訊及計算相位，接著放入以時序卷基網路為主的模型提取生命特徵，最後進入回歸器模型中將這些特徵回歸至一個值，而這個值便是我們所估計出的心率及呼吸率。我們將血氧儀(Oximeter)測量出的結果當作實際答案(Ground True)，與模型估計的結果做比較，計算損失函數(Loss Function)，再藉由反向傳播(Backpropagation)來更新梯度(Gradient)，用來降低損失函數。同時透過驗證來調整學習模型的超參數，以利模型有效學習。在模型測試階段，流程皆與模型訓練階段相同，唯一不同的是，測試階段不會進行反向傳播及損失函數計算。測試的結果會以一些客觀的評估指標來衡量此模型的品質與效能。

4-2 訓練模型階段

延續上一節所提到的，在訓練模型階段將以三個小節依序介紹。資料前處理(Data Pre-processing)介紹雷達訊號如何轉換成訓練資料集(Train Dataset)，神經網路模型(Neural Network Module)介紹整個模型的流程，從時序卷積網路到回歸器模型，最後反向傳播(Backpropagation)及更新梯度(Gradient)將介紹損失函數及優化器的使用，以此來更新網路架構的內部參數。表 1 是訓練模型的階段之參數設置表。

表 1. 訓練模型階段參數設定列表

批次處理量 Batch size	64
學習率 Learning rate	0.0001
學習率衰減 Learning rate decay	0.0001
迭代次數 Epoch	100
優化器 Optimizer	Adam optimizer W

4-2-1 資料前處理

在本篇研究中，我們使用了 infineon FMCW radar(BGT60TR13C)[5]作為本次實驗的傳感器，該雷達以 10 赫茲的頻率發射訊號至目標。並從美國食品與藥物管理局批准的醫療設備 Masimo Pulse Oxygen[6]測量出的結果當作 Ground Truth。

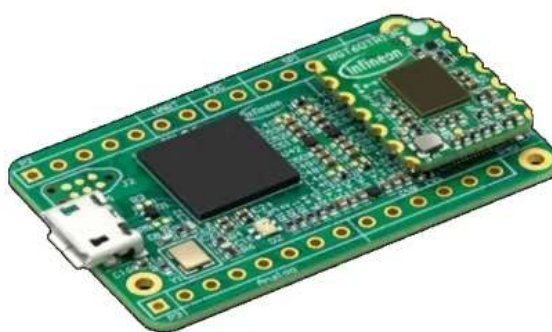


圖 12.FMCW 雷達



圖 13. Masimo Pulse Oxygen

參考[7]使用以 1 秒為單位的滑動窗口(sliding windows)，針對不同的時間長度做比較。將實驗場景設定在辦公室環境中，使用電腦的日常行為，雷達架設在距離人體胸前約 60 公分的螢幕下方。我們總共邀請了 7 位志願者來協助蒐集資料，每個人花了約 1 個小時，其中包含日常使用電腦的行為(打字、移動滑鼠)、飲食、揮動手臂...等等隨機行為。此外，還允許志願者身體可以前傾及後仰這類大幅度位移，增加訓練時的挑戰性。訓練數據集(train dataset)約有 20000 筆、驗證資料集(validation dataset)有 1000 筆、測試資料集(test dataset)約有 1000 筆。

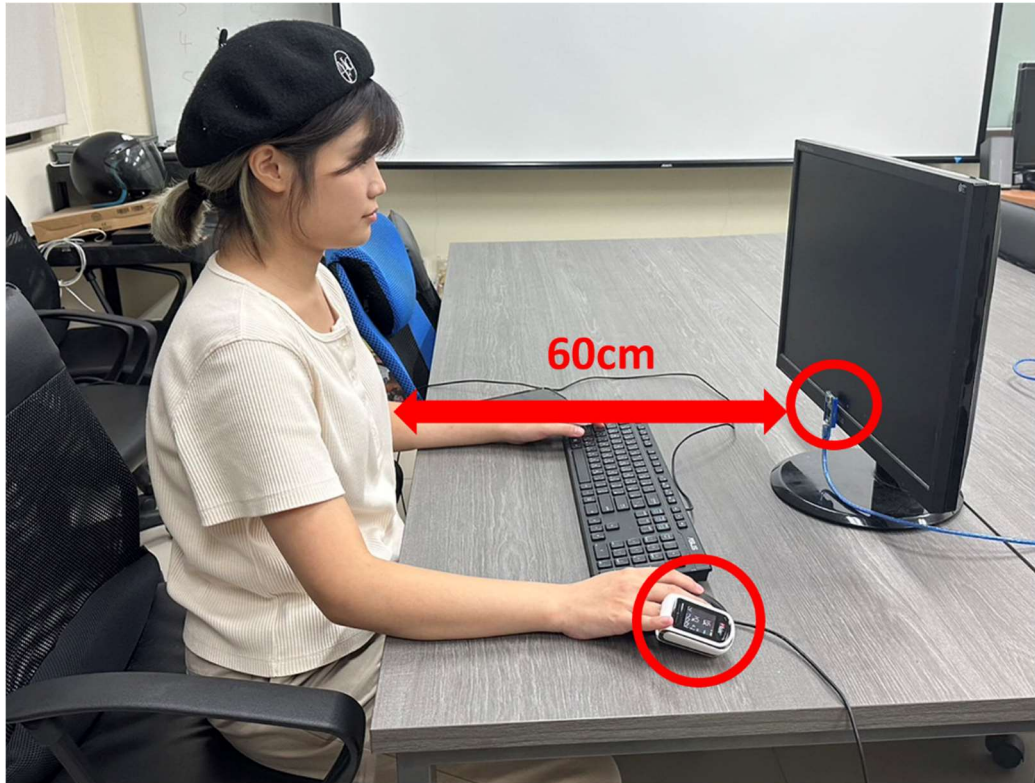


圖 14.實驗場景圖

本論文學習任務的目標為從人體相位中做生命特徵的提取，因此在接收到雷達訊號後，需經由前處理將原始訊號轉成相位訊號。圖 8 為前處理流程圖，反射訊號經由天線給接收後，與發射訊號混頻可得拍頻訊號。對拍頻訊號取樣轉換成數位訊號，才能進行後續的數位訊號處理。此時的訊號仍包含靜態的背景雜訊以及動態的生命訊號，因此我們利用靜態雜波濾除移去背景大部分的靜態雜訊。接者便能透過快速傅立葉變換計算出目標與雷達的距離，透過相位提取、相位展開、相位差計算將每幀轉換成相位資訊。最後依照不同實驗的時間長度存成一筆筆資料集，放入學習模型訓練。

4-2-2 神經網路模型

本論文神經網路主要目標是根據輸入人體相位資訊估計心率及呼吸率，從上一小節取得經前處理的相位結果作為輸入，放入以時序卷積網路為主體設計的模型。整體流程為，輸入至 TCN 提取局部特徵，再經由 Transformer encoder 加入全局觀視野，最後放入兩個回歸器分別估計心率及呼吸率。學習模型的詳細架構可以參考第三章深度學習相關介紹。

4-2-3 反向傳播及更新梯度

在模型訓練的流程中，損失函數的計算是不可或缺的，透過損失函數計算出的結果來進行反向傳播，更新模型裡的參數權重。在回歸的任務中都希望預測出來的結果能與實際結果一樣，但現實是不可能達到完美的訓練。因此，可以透過預測結果與實際結果的落差當作評斷標準，這個落差也稱為損失。損失函數越小，表示預測結果與實際結果的落差也越小，反向傳播及更新梯度的目的在於降低損失函數，達到訓練模型的效果。本論文採用均方誤差(Mean square error, MSE)作為模型訓練的損失函數，在回歸學習相當常見的一種損失函數。均方誤差的計算方法就是求「預測值與真實值之間差異的均方值」，其表示如下：

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (4-1)$$

其中 \hat{y}_i 為估計的結果， y_i 為血氧儀測量的結果。

4-3 測試階段

在測試模型階段，我們使用測試數據集來評估訓練好的神經網路模型的成效。測試階段的流程與訓練階段類似，差別在於測試階段不會計算損失函數，也不會反向傳波及更新參數，輸入相位資訊使模型產生估計的心率及呼吸率。後續我們選用 MSE 當作客觀的評斷標準，估計出來的心率以每分鐘的跳動次數(Beat Per Min, BPM)為單位，呼吸率則以每分鐘吸吐次數(Respiratory Per Min, RPM)為單位。

第五章 實驗結果與分析討論

本章節將介紹對此次提出之系統進行的實驗結果，並對此進行討論分析及效果比較。主要分為三個部分，5-1 節將對實驗環境的軟硬體進行介紹，5-2 節會對實驗結果進行相關的比較和討論。

5-1 實驗環境介紹

採用 pytorch 框架來搭建整體神經網路模型，並在訓練模型階段以 GPU GeForce GTX 2080 設備加速神經網路收斂，縮減網路訓練時間。下表 4 為實驗環境相關配置規格。

表 2. 實驗環境軟硬體之配置表

CPU	Intel Core i7-9700k @ 3.60GHz
GPU	RTX 2080
RAM	64GB DDR4-3200 MHz
OS	Ubuntu 18.04
Software language	Python3.6
Neural Network tool	Pytorch

5-2 實驗結果比較與討論

本節將呈現本論文提出架構針對不同長度訊號 MSE 的誤差結果，以及參數比較，並對結果進行討論。

5-2-1 時序卷積網路以不同訊號時長之比較

本實驗分別採用不同的訊號時長作為輸入訊號訓練模型，比較訊號長度對整體效果的影響，此外，加入 Deep VS[7]的結果作為參考基準(Baseline)。不同的是，Baseline所使用的雷達為超寬頻雷達(Impulse radio ultra-wideband, IR-UWB)，資料前處理以及學習模型也與我們提出的不同，訓練資料採用一秒為增量的平滑窗口做堆疊，其訊號長度為 60 秒。因此，此結果僅供參考基準，不列入比較。

表 3. 不同訊號時長之心率及呼吸率比較表

	HR_error (bpm)	RR_error (rpm)
Deep VS [7]	7.40	4.90
TCN(60 sec)	8.62	4.00
TCN(30 sec)	7.69	3.53
TCN(20 sec)	6.90	3.49
TCN(10 sec)	6.67	3.36
TCN(5 sec)	6.98	3.44

經由表 3 所示，以 TCN 為主體的模型在輸入為 10 秒時心率有著 6 bpm 的最低錯誤值，呼吸率則有 3bpm 的最低錯誤值。藉由表 3 推斷，過長的訊號經由擴張卷積提取特徵時，會將許多關聯較低的特徵一起被提取出來；反之，過短的訊號經由擴張卷積提取特徵時，提取太少有關聯的特

徵，導致整體效果降低，所以 10 秒的時間長度對於生命特徵監測來說較為合適。

5-2-2 時序卷積網路模型參數之比較

考慮完整體模型效果，接下來我們針對模型參數量進行分析。參數量的多寡可以代表模型的複雜度，複雜度較高的模型通常有著較多的參數量，伴隨著較高的運算量;複雜度較低的模型通常有著較少的參數量，伴隨著較低的運算量。較低的模型能夠降低訓練時間，為此希望模型越輕便越好，下表為參數量之比較:

表 4. 模型參數之比較表

	parameter
Deep VS [7]	160k
TCN	101k

由表 4 所示，我們所提出的架構與 baseline 相比少了 60k 參數量，推斷我們的架構更輕便且有較少的運算時間。

第六章 結論與未來展望

本篇論文提出了基於時序卷積網路為主體的神經網路模型，可以應用於辦公環境之生命特徵監控。在傳統的生命特徵監控中，通常利用訊號處理的方式，對於靜態的目標來說十分有效，但當訊號受到身體運動所干擾時，便需要增加更多的運算來解決被破壞的訊號，這會使計算量變大且整體表現的效果降低。為此後續的一些研究改以深度學習的方式來克服人體運動所造成訊號破壞的問題。

我們提出的方法基於前人的研究及經驗，採用時序卷積的方法，在一般卷積時加入時序因果性，使特徵裡包含局部的時間觀，加強相鄰時刻訊號的關係。除了局部的時間觀，我們認為全局觀的變化對於生命特徵監控也是有幫助的，因此加入了注意力機制，在 TCN 取得局部特徵後，將這些特徵做全局的匹配。此方法在實驗結果中展示了運算參數的降低，也能保持較低的錯誤值，並在 MSE 的錯誤計算上有更好的效果。

在未來的方向上，實際辦公環境裡可能同時存在好幾位受試者，因此系統應該要具備應對在多人環境的能力與持續提高準確度為主要課題。在輕型模型上，期望能夠將模型套用在許多輕便的嵌入式裝置上使用，使此系統能夠更為普及。我們希望找出改良方法使系統可以更普及的被應用在更複雜的情境中保持良好的效果。

參考文獻

- [1] D. Wang, S. Yoo, and S. H. Cho, “Experimental Comparison of IR-UWB Radar and FMCW Radar for Vital Signs,” *Sensors*, vol. 20, no. 22, p. 6695, Nov. 2020, doi: 10.3390/s20226695.
- [2] S. Bai, J. Z. Kolter, and V. Koltun, “An empirical evaluation of generic convolutional and recurrent networks for sequence modeling,” *arXiv preprint arXiv:1803.01271*, 2018.
- [3] C. Subakan, M. Ravanelli, S. Cornell, M. Bronzi and J. Zhong, "Attention Is All You Need In Speech Separation," *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, 2021, pp. 21-25, doi: 10.1109/ICASSP39728.2021.9413901.
- [4] Y. H. Shen (2020).Non-Contact Real-Time Respiration Rate Monitoring Using FMCW Radar ◦ (master's thesis ◦ National Ilan University)
- [5] [n. d.]. infineon FMCW radar(BGT60TR13C).
<https://www.infineon.com/cms/en/product/sensor/radar-sensors/radar-sensors-for-iot/60ghz-radar/>
- [6] [n. d.]. Masimo - MightySatRx.
<https://www.masimo.com/products/monitors/spot-check/mightysatrx/>
- [7] Z. Xie, H. Wang, S. Han, E. Schoenfeld, and F. Ye. 2022. “DeepVS: a deep learning approach for RF-based vital signs sensing”. In *Proceedings of the 13th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (BCB '22)*.