

國立臺灣師範大學跨域科技產業創新研究學院

AI 跨域應用研究所

碩士論文

Graduate Institute of AI Interdisciplinary Applied Technology, College of
Interdisciplinary Industry-Academia Innovation

National Taiwan Normal University

Master's Thesis

以毫米波雷達為基礎的手勢辨識之研究

Hand Gesture Recognition Based on Millimeter-Wave Radar



Sheng-Kai Lin

指導教授：黃文吉 博士

Advisor: Wen-Jyi Hwang, Ph.D.

中華民國 114 年 07 月

July 2025

誌謝

回首這段碩士求學的歷程，從毫無基礎開始接觸毫米波雷達，到能夠獨立完成論文研究與系統實作，過程中所經歷的挑戰與成長，感謝師長與親友的支持與陪伴。在此，謹向所有曾在我學習與研究過程中提供協助的人致上最深的感謝。

首先，我要感謝我的指導教授黃文吉教授。在我對毫米波雷達技術毫無概念之初，教授透過清晰的指引與觀念解析，引導我快速掌握關鍵知識，為我後續的研究奠定堅實基礎。透過教授的安排，我得以前往開酷科技實習，獲得實務經驗並加深對研究主題的理解。教授在研究方向上的引導，使我能逐步釐清問題與目標，朝明確方向前進，深感謝意。

接著，我要深深感謝我的家人，尤其是父母與外婆。謝謝你們一直以來無條件地支持與陪伴，是我面對困難時最堅強的後盾。你們的理解與鼓勵，讓我能安心地投入研究，專注地完成每一階段的任務。每一次的進展與突破，都離不開你們在背後默默的守護。

我也想特別感謝開酷科技的經理與實習期間協助我的同仁們。在這段寶貴的實習經驗中，您們不僅傳授我各項雷達操作與資料處理的實用技術，進一步讓我學會如何將理論知識有效應用於實務情境中，也讓我更加確信未來的方向。

此外，我要感謝 MC2 實驗室的夥伴們，宇恩、玳維、昱誠、琮達、承翰與子堯，謝謝你們在研究過程中的協助與陪伴。從彼此討論問題到一起加班的時光，這段共同奮鬥的回憶將會是我人生中珍貴的一頁。

也感謝幾位學弟—彥儒、子晏、信中、紹詞、政昇與大有，在資料收集、系統交接與經驗分享上的幫助。你們的協助讓研究推進順利，成果也更為完善。

最後，感謝所有在我研究路上給予幫助與鼓勵的人，是你們的支持讓我在迷惘與挫折中重新站穩腳步。這篇論文的完成，不只是個人成果，更是大家共同成就的累積。謝謝你們，讓這段碩士旅程充滿溫暖與力量。

摘要

隨著人機互動技術的快速發展，毫米波雷達因具備隱私性高、不受光線干擾、可穿透遮蔽物等優勢，已逐漸成為手勢辨識應用之新興感測技術。本文採用開酷科技所開發之 60GHz 毫米波雷達，並搭配其專用的視覺化平台 Ksoc Tool，此工具為專門配合該雷達設計，具備資料收集與標註功能。透過 Ksoc Tool 完成原始資料擷取與資料標註後，進一步進行資料前處理、模型訓練與即時辨識顯示，建構出一套具備即時性與彈性的手勢辨識系統流程。

在影像資料方面，本文深入說明兩種常見雷達影像格式：RDI (Range-Doppler Image) 與 PHD (Phase Difference Map)，並透過圖像與實例詳細解析其物理意義與應用情境。為有效處理動態手勢資料，系統採用滑動視窗機制切割連續序列，並透過高斯函數生成 soft label，提升標註於手勢邊界區域的過渡敏感性。模型部分則採用三維卷積神經網路 (3D CNN) 以同時擷取空間與時間特徵，並搭配均方誤差 (MSE) 作為損失函數進行監督式訓練。

為強化手勢段落之區分能力，本文提出雙門檻後處理機制，透過進入與離開閾值協助界定動作啟始與終止點，並架設圖形介面，實現雷達資料的即時推論與手勢顯示。實驗結果顯示，本系統可正確辨識包含背景、PatPat、Come 與 Wave 四類別手勢，整體準確率達 95.8%，展現本研究於準確性、即時性與可擴展性三方面之應用潛力。

關鍵詞：毫米波雷達、手勢辨識、三維卷積神經網路、滑動視窗

Abstract

With the rapid advancement of human-computer interaction technologies, millimeter-wave radar has emerged as a promising sensing modality for gesture recognition due to its high privacy, resistance to lighting conditions, and ability to penetrate occlusions. This study utilizes a 60 GHz mmWave radar developed by Kaiku Technology, along with its dedicated visualization platform, the Ksoc Tool—designed specifically for the radar—to perform data collection and annotation. Based on the collected data, we implement a full pipeline encompassing preprocessing, model training, and real-time recognition display.

The proposed system incorporates two types of radar images—Range-Doppler Images (RDI) and Phase Difference Maps (PHD)—with detailed analysis of their physical meaning and usage scenarios. A sliding window mechanism is employed to segment continuous gesture data, and soft labels are generated using Gaussian functions to enhance sensitivity at gesture boundaries. A 3D Convolutional Neural Network (3D CNN) is adopted to capture spatiotemporal features, trained with Mean Squared Error (MSE) loss. To further improve gesture segmentation, a dual-threshold post-processing mechanism is introduced to detect gesture start and end points. A graphical user interface is implemented to support real-time inference and visualization.

Experimental results demonstrate that the system accurately recognizes four gesture types—Background, PatPat, Come, and Wave—with an overall accuracy of 95.8%, highlighting its potential in real-time performance, accuracy, and scalability.

Keywords: millimeter-wave radar; gesture recognition; 3D CNN; sliding window

目錄

誌謝.....	i
摘要.....	ii
Abstract.....	iii
目錄.....	iv
圖目錄.....	vi
第一章 緒論.....	1
1-1 論文背景.....	1
1-2 論文動機.....	2
1-3 論文目的.....	4
第二章 毫米波雷達介紹.....	5
2-1 K60168A Dongle 套件.....	5
2-2 雷達可視範圍.....	6
2-3 毫米波雷達參數介紹.....	8
2-3-1 逕向距離.....	8
2-3-2 逕向速度.....	10
2-3-3 入射角.....	12
2-3-4 逕向距離、逕向速度和入射角關聯.....	13
2-3-5 入射角計算方式.....	14
第三章 毫米波雷達開發工具.....	16
3-1 Ksoc Tool 平台簡介.....	16
3-2 雷達影像類型與物理意義.....	18
3-2-1 Range-Doppler Image (RDI).....	18
3-2-2 Range-Doppler Image 參數與解析度說明.....	20
3-2-3 單一目標於 RDI 上的基本成像行為.....	22
3-2-4 小手勢之垂直方向動作範例.....	26
3-2-5 小手勢之水平方向移動動作範例.....	28
3-2-6 大手勢之水平方向移動動作範例.....	30
3-2-7 Phase Difference Map (PHD).....	32
3-3 資料擷取與標註流程.....	33
3-4 資料格式與儲存方式.....	35
第四章 基礎應用實例及結果.....	37
4-1 毫米波雷達手勢辨識之侷限.....	37
4-2 手勢定義與資料蒐集方式.....	38
4-3 標註資料處理與 Ground Truth 建構.....	40
4-3-1 標註格式設計.....	40

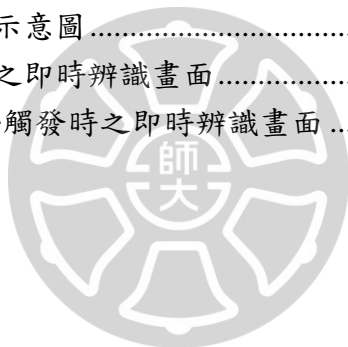
4-3-2 Soft Label 建構方法.....	42
4-3-3 滑動視窗切片機制	44
4-3-4 Ground Truth 對應機制.....	46
4-4 模型架構與訓練流程.....	48
4-5 預測結果與後處理機制.....	49
4-6 即時辨識系統設計	53
第五章 結論.....	56
參考文獻.....	57



圖目錄

圖 1-1 傳統影像感測器與毫米波雷達在手勢辨識應用之比較.....	2
圖 2-1 K60168A Dongle 套件.....	6
圖 2-2 K60168A 雷達模組之可視範圍示意圖.....	7
圖 2-3 相同逕向距離下不同入射角目標於空間中的示意圖.....	9
圖 2-4 運動速度與其逕向與切線速度分量關係示意圖.....	10
圖 2-5 不同空間位置下相同運動速度目標的逕向速度變化示意圖.....	11
圖 2-6 左至右移動時之入射角變化示意圖.....	12
圖 2-7 右至左移動時之入射角變化示意圖.....	13
圖 2-8 具有相同逕向距離與逕向速度之不同位置目標示意圖.....	14
圖 2-9 入射角與距離差計算原理示意圖.....	15
圖 3-1 Ksoc Tool 操作介面圖.....	17
圖 3-2 Range-Doppler Image 座標示意圖.....	19
圖 3-3 Range-Doppler Image 各參數說明圖.....	20
圖 3-4 Ksoc Tool 設定檔選擇介面.....	21
圖 3-5 各觀測設定之參數對應圖.....	22
圖 3-6 目標沿雷達軸線方向遠離時於 RDI 上之成像位置.....	23
圖 3-7 不同逕向速度對應於 RDI 上不同位置.....	23
圖 3-8 目標沿雷達軸線方向靠近時於 RDI 上之成像位置.....	24
圖 3-9 逕向速度為實際速度與雷達視線夾角投影於 RDI 上之成像位置.....	24
圖 3-10 多目標具相同逕向距離與速度時於 RDI 上之成像重疊.....	25
圖 3-11 手部直線接近雷達時於 RDI 上的動態變化軌跡.....	26
圖 3-12 手部遠離雷達時於 RDI 上的動態變化軌跡.....	27
圖 3-13 手部靠近後再遠離時於 RDI 上的完整移動軌跡.....	27
圖 3-14 左滑動作之 RDI 變化軌跡.....	28
圖 3-15 右滑動作之 RDI 變化軌跡.....	29
圖 3-16 等速橫向左移物體之 RDI 投影路徑示意圖.....	30
圖 3-17 等速橫向右移物體之 RDI 投影路徑示意圖.....	31
圖 3-18 小手勢與大手勢於雷達視野內之運動軌跡示意圖.....	32
圖 3-19 Phase Difference Map (PHD) 坐標示意圖.....	33
圖 3-20 Ksoc Tool 資料擷取與即時標註操作畫面.....	34
圖 3-21 RDI 與 PHD 資料儲存格式示意圖.....	36
圖 4-1 三類動作手勢定義與示意圖.....	39
圖 4-2 各手勢類別與背景資料之蒐集者分布統計.....	39
圖 4-3 時間軸標註示意圖.....	41
圖 4-4 雷達資料與標註結構示意圖.....	41

圖 4-5 類高斯分布於手勢區段中的 soft label 配置範例	42
圖 4-6 多類別 One-Hot 手勢標註對應結果	43
圖 4-7 類高斯 Soft Label 權重分布範例	44
圖 4-8 滑動視窗切片示意圖	45
圖 4-9 視窗落於背景區段之 Ground Truth 對應機制示意圖	46
圖 4-10 視窗涵蓋手勢主體區段之 Ground Truth 對應機制示意圖	47
圖 4-11 視窗位於手勢結束後區段之 Ground Truth 對應機制示意圖	47
圖 4-12 本系統所採用之雙通道輸入影像結構	48
圖 4-13 三維卷積神經網路 (3D CNN) 模型架構	49
圖 4-14 Ground Truth 與預測結果之時間序列對比圖	50
圖 4-15 雙閾值狀態轉移流程圖	51
圖 4-16 雙閾值機制應用於預測輸出之示意圖	51
圖 4-17 單筆手勢分類之混淆矩陣	53
圖 4-18 即時手勢辨識系統流程架構圖	54
圖 4-19 暫存區更新機制圖	54
圖 4-20 暫存區更新程式碼示意圖	54
圖 4-21 系統於背景狀態下之即時辨識畫面	55
圖 4-22 系統於 Come 手勢觸發時之即時辨識畫面	55



第一章 緒論

本章節將說明本論文之背景、動機與目的。隨著智慧科技日新月異，非接觸式操作逐漸取代傳統按鍵與觸控式互動，成為人機介面設計的主流趨勢。手勢辨識技術因其自然直觀的操作方式，在物聯網裝置與智慧生活應用中扮演重要角色。然而，現行主流的影像式手勢辨識受限於環境光源與隱私風險，應用場景仍存諸多限制。為克服此問題，毫米波雷達技術以其穿透性強、抗光性佳與不涉及隱私等優勢，逐漸成為手勢辨識領域中的替代方案。

1-1 論文背景

隨著物聯網 (IoT) 與人工智慧 (AI) 技術的持續進展，非接觸式操作介面廣泛應用於智慧家庭、車載系統、醫療照護與穿戴裝置等領域。其中，手勢辨識技術透過感測人體動作進行指令輸入[1]，具備直觀操作、高自由度與衛生安全等優點，逐漸成為提升用戶體驗的重要手段。

目前常見的手勢辨識多以攝影機影像為主，這些系統仰賴明確的視覺資訊進行分析，因此對光照條件高度敏感，在昏暗或背光環境[2]下辨識效果大幅降低。此外，影像感測涉及使用者臉部、肢體等個人隱私資訊，易產生資安與倫理爭議，導致其在醫療、浴室、臥室等高隱私場域應用受限。

相較之下，毫米波雷達 (Millimeter-Wave Radar) 採用主動式訊號發射與接收機制，可在無光照或遮蔽環境中穩定運作，並具備解析動態距離與速度的能力。由於其僅蒐集反射訊號特徵，並不具備可辨識人臉等影像資訊，故在保障使用者隱私方面表現優異。

下圖比較了傳統影像感測器與毫米波雷達在應用上的差異：

	傳統影像感測器	毫米波雷達
光源影響	較明顯	無影響
隱私風險	較高	較低
適用場景	較侷限	較廣泛

圖 1-1 傳統影像感測器與毫米波雷達在手勢辨識應用之比較

由表中可見，毫米波雷達在光照不佳或需重視隱私的場域具有明顯優勢，特別適合應用於智慧居家、醫療監控與沉浸式互動裝置等環境。

1-2 論文動機

毫米波雷達近年來因具備抗光線干擾、隱私保護佳、解析度高等優勢，逐漸被應用於非接觸式手勢辨識領域。然而，在實務開發過程中，從資料理解到系統設計仍面臨多項挑戰，使其在人工智慧應用中的實作與推廣受限。以下說明本論文的三項主要動機：

1. 雷達資料難以如相機影像般直觀理解

在與手勢辨識相關的深度學習應用[3]中，開發者普遍採用相機擷取的影像資料作為模型輸入[4]。這類影像資料可呈現手部的外觀形狀、動作變化與背景環境等資訊，其所反映的視覺內容與人眼所接收到的感知訊息相近。因此，開發者能夠直接透過圖像內容進行手勢類別定義、資料標註與樣本篩選，操作上具備高度的直觀性與便利性，亦有助於前期系統開發流程的推展。

相較之下，毫米波雷達所提供的資料形式與相機影像截然不同。其輸出影像如 Range-Doppler Image (RDI) 與 Phase Difference Map (PHD)，所反映的並非真實畫面中物體的外觀形狀，而是目標物與雷達之間的物理特徵。RDI 描述物體在不同「逕向距離」與「逕向速度」下的反射強度分布，PHD 則對應「逕向距離」與「入射角」的相位差資訊。這些資料雖然包含與動作相關的豐富訊號，但其影像內

容並不具備人眼熟悉的形狀、邊界或外形資訊，開發者難以直觀地從畫面中辨識手勢動作的特徵區段或運動方向。由於這類資料需要結合雷達原理與訊號處理概念進行解讀，對未具備相關背景知識的開發者而言，往往難以掌握資料與動作之間的關聯性，進而在資料標註、樣本挑選與模型錯誤分析等階段面臨較高的操作困難。整體而言，毫米波雷達雖提供可支援動作辨識的重要特徵資訊，但其資料形式的理解門檻，已成為手勢辨識系統實作與應用推廣上不可忽視的挑戰之一。

2. 現有開發工具尚未充分支援 AI 應用

目前可取得的毫米波雷達開發工具大多著重於硬體驅動、資料擷取與影像顯示等基本功能，其設計初衷主要為測試雷達效能與進行基礎觀測，對於人工智慧應用導向的開發流程支援相對有限。在實際的深度學習模型建構過程中，常需針對原始雷達輸出進行多階段的資料處理作業，包括資料擷取、資料標註、資料格式的轉換與正規化等。然而，這些關鍵步驟在現行開發工具中大多未被納入，開發者須額外撰寫大量的自定義程式碼進行補強。

3. 具實作性的應用案例較少

儘管毫米波雷達在學術研究中已廣泛應用於目標追蹤、姿態估測與手勢辨識等任務，並具備許多潛在優勢，但目前相關文獻多著重於演算法設計或模型準確率的驗證，實際將研究成果落實為可操作系統的案例仍相對稀少。在多數研究中，僅以靜態資料集驗證辨識性能，缺乏完整整合資料擷取、即時推論與應用互動的系統實作流程。

1-3 論文目的

本論文旨在探討毫米波雷達於近距離手勢辨識任務中的應用可行性，並透過現有工具實作資料處理與模型應用流程，以回應現行雷達資料不易理解、工具支援有限與缺乏實務應用案例等問題。具體論文目的如下：

1. 基礎毫米波雷達影像資料解析說明

本論文首先針對毫米波雷達所產生的影像資料進行結構性解析，說明 Range-Doppler Image (RDI) 與 Phase Difference Map (PHD) 中所蘊含的物理意義與資料特徵，分別對應逕向距離、逕向速度與入射角等參數。透過視覺化示意，協助開發者建立對雷達資料的基本認知，有助於後續進行有效的特徵擷取與動作理解。

2. 應用毫米波雷達資料處理流程於模型辨識任務

本論文採用現有的雷達資料擷取與處理工具，取得毫米波雷達影像資料與對應的標註資訊，並轉換為可供深度學習模型使用的輸入格式。相關處理流程，包括 RDI 與 PHD 影像的擷取、標註資料的對應整理，皆由既有工具與程式實現，本論文聚焦於說明如何實際操作這些工具以完成資料處理，說明各處理步驟於資料前處理階段中的功能與實際應用，協助使用者理解整體流程如何銜接至後續模型建構工作。

3. 提供雷達實作手勢辨識應用之範例

本論文提供一個完整應用流程的實作範例，作為整體系統可行性驗證之基礎。應用設計涵蓋三種近距離動態手勢，並透過離線資料測試與即時操作情境下的驗證，評估系統在手勢辨識準確率方面的表現。藉由本實作範例，說明毫米波雷達結合人工智慧模型於近距離手勢辨識任務中的可行性，並提供後續系統設計與應用開發之參考依據。

第二章 毫米波雷達介紹

本章節將說明本論文所使用之毫米波雷達模組與其相關量測原理，作為後續資料處理與模型設計的基礎。本論文所採用之毫米波雷達模組為 K60168A Dongle [5]，為開酷科技（Kaikutek Technology）所開發之毫米波雷達動作感測開發套件，具備一個發射天線與三個接收天線，可支援 1 至 7 GHz 的可調頻寬，其觀測範圍涵蓋水平 ± 45 度、最遠達 240 公分，適合用於手勢等短距離動作的觀測任務。

毫米波雷達所產生的訊號，經處理後可反映三項物理特徵參數，包括逕向距離（Radial Distance）、逕向速度（Radial Velocity）與入射角（Angle of Arrival）。這些參數可進一步轉換為影像資料形式，並作為後續深度學習模型之輸入依據。藉由掌握各項特徵的意義與空間分佈方式，能有效理解雷達資料在手勢辨識任務中的應用價值。

2-1 K60168A Dongle 套件

本論文所採用之毫米波雷達模組為 K60168A Dongle，一款專為近距離動作感測設計之開發套件。其核心為一顆 60 GHz 毫米波感測晶片，適用於多種近距離感測應用，包括動作識別、物體存在檢測、呼吸監測與手勢辨識等。該套件內部配置一個發射天線與三個接收天線（1T3R），可透過反射訊號擷取動作資訊，適用於多種非接觸式近距離感測場景。

K60168A 具備中短距離觀測能力，最大可偵測距離為 240 公分，支援 ± 45 度的水平視角，可涵蓋一般人體動作與手部活動所需的感測範圍。圖 2-1 為實際模組照片，圖中紅色框線標示為發射天線，藍色框線則為三個接收天線，構成 1T3R 的天線配置。模組透過 Type-C 介面進行資料傳輸，具備體積小巧、接線簡便等優點，便於進行實驗架設與資料擷取作業。

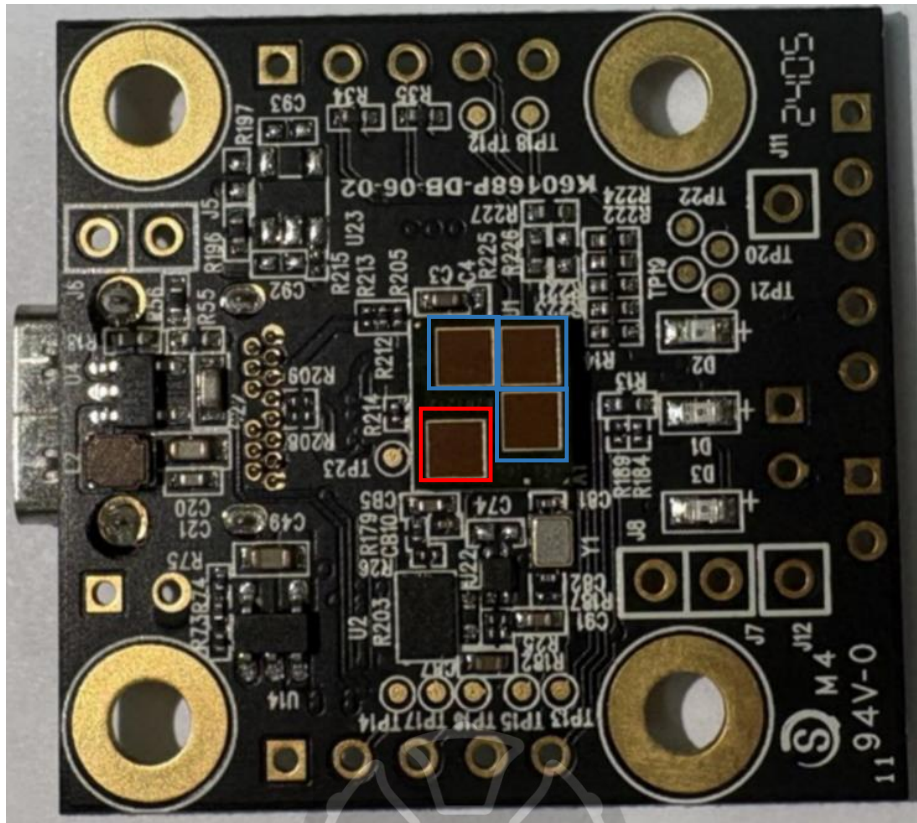


圖 2-1 K60168A Dongle 套件

模組支援 1 至 7 GHz 之可調式頻寬設定，頻寬大小將直接影響後續影像資料中可分辨的物體細節程度與偵測距離。下一節將進一步說明雷達在空間中可觀測的範圍，包括從雷達發射點擴展出去的視角結構，並介紹頻寬設定如何影響影像解析度與偵測距離，幫助讀者理解這些參數設定對後續辨識任務的實際影響。

2-2 雷達可視範圍

毫米波雷達的感測範圍，主要取決於其發射訊號所涵蓋的角度範圍與所設定的頻寬參數。K60168A 模組在水平方向的觀測角度為 ± 45 度，而整體視野則呈現一個向前擴展的錐狀空間，涵蓋了水平方向與垂直方向的擴散範圍，如圖 2-2 所示。圖中左側為立體視角，可見雷達訊號從模組正前方呈錐形擴散，右側則為俯視圖，顯示其在水平方向的涵蓋範圍。這樣的空間結構決定了在裝置固定不動的情況

下，雷達能夠偵測到的動作區域。

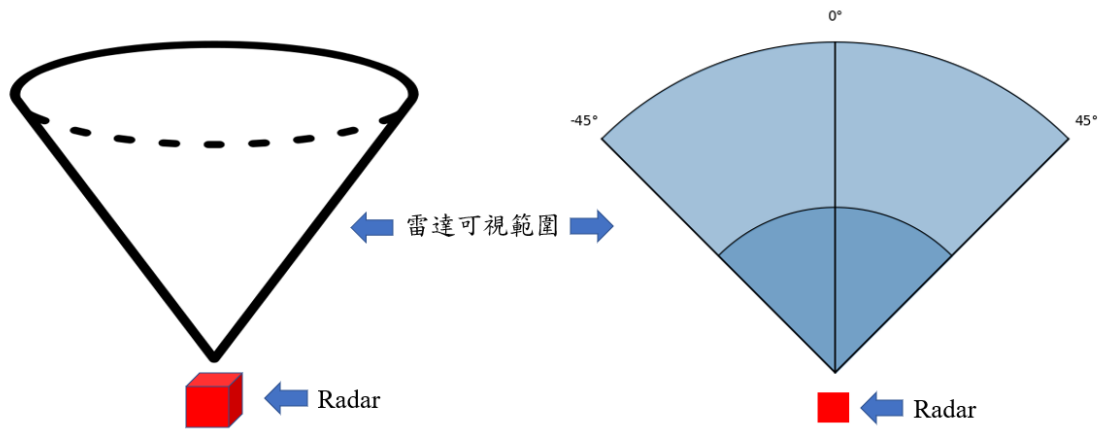


圖 2-2 K60168A 雷達模組之可視範圍示意圖（左：立體視角圖；右：俯視圖）

然而，實際雷達的「最大可偵測距離」與「距離解析度」並非僅由視角決定，而是高度依賴頻寬設定。K60168A 模組支援 1 至 7 GHz 之間的頻寬設定，其中，頻寬設定的大小會直接影響雷達可分辨的距離精細程度與可偵測的最遠距離。當頻寬較大時，雷達能夠更細緻地分辨不同距離的物體，例如在 7 GHz 的頻寬設定下，可達到約 2.14 公分的距離解析度，適合用於精細的手部動作辨識。不過此時可偵測的最遠距離會相對縮短。反之，若頻寬較小，雖然距離解析度下降，但觀測範圍會擴大，更適合用於偵測遠距離的目標。實務應用中，需根據實驗需求與場域條件，在解析能力與偵測距離之間取得適當平衡。

2-3 毫米波雷達參數介紹

毫米波雷達透過發射與接收訊號，量測目標物體與雷達之間的距離、移動速度與方向等資訊，其觀測資料可對應三項關鍵參數，分別為：逕向距離（Radial Distance）、逕向速度（Radial Velocity）與入射角（Angle of Arrival, AoA）。這些參數為毫米波雷達判讀目標物在空間中位置、運動狀態的依據，亦是本論文中後續手勢辨識模型輸入資料的重要來源。

2-3-1 逕向距離

逕向距離為雷達觀測中最基本的空間參數之一，此參數可顯示目標物目前距離雷達多遠，對於辨識其在空間中的前後位置分佈具有參考價值。然而，單一時刻的逕向距離僅能反映靜態位置，若要判斷物體是否正在靠近或遠離雷達，則需觀察連續時間下的距離變化。

此外，逕向距離無法區分物體是出現在雷達正前方偏左還是偏右的位置，也無法判斷其移動方向，因此需要結合其他參數（如逕向速度與入射角）來補足空間資訊。

如圖 2-3 所示，若兩個目標分別位於雷達左右兩側，若其與雷達的距離皆為 40 公分，則其逕向距離相同。此例說明逕向距離僅與雷達與目標之間的直線長度有關，無法反映橫向位置或方位角。

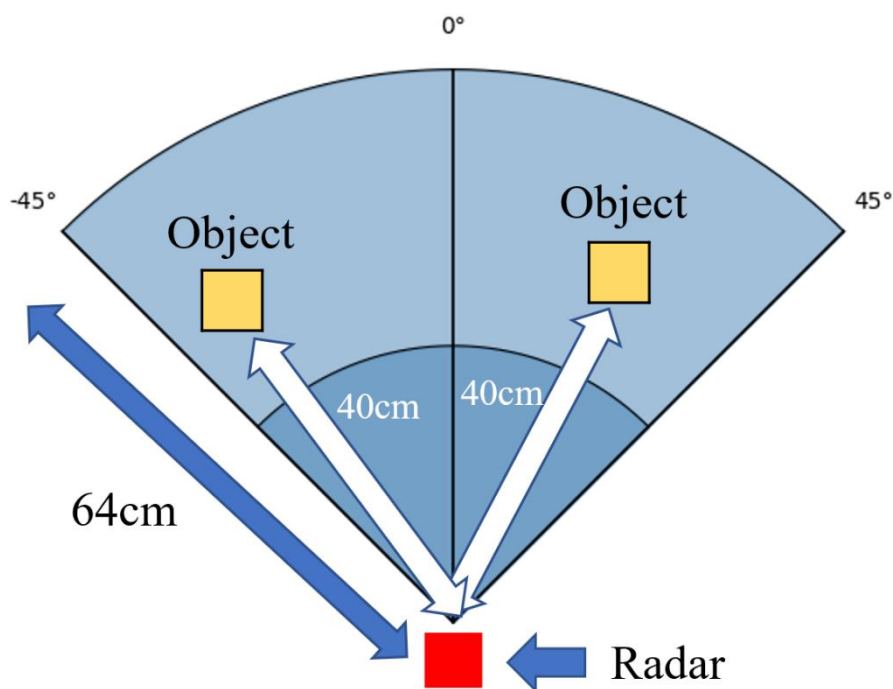


圖 2-3 相同逕向距離下不同入射角目標於空間中的示意圖

逕向距離本身僅反映物體在單一時刻與雷達之間的直線距離，無法直接判斷移動方向或趨勢。然而，當逕向距離隨時間連續觀測時，即可形成一組具變化趨勢的時間序列，進而反映目標在前後方向上的動作軌跡。例如，在「往後收手」的動作中，手部距離雷達逐漸拉遠，對應的逕向距離數值隨時間穩定上升；反之，「伸手靠近」則呈現逐步下降的趨勢。這種具方向性與節奏性的距離變化，有助於系統辨識深度移動類型的動作，並強化對動態特徵的掌握能力。

2-3-2 逕向速度

逕向速度表示目標物沿著雷達方向的移動速率，數值為負代表物體靠近雷達，為正則代表遠離。此參數能反映物體在不同時間點的運動方向與速度變化，有助於判斷目標是否為靜止、持續移動，或在進行往返運動。透過逕向速度的觀測，可補足僅有位置資訊所無法呈現的動態特性，使雷達能對移動行為進行更全面的判讀。

如圖 2-4 所示，物體的總運動速度可分解為兩個分量：一為沿雷達方向的逕向速度，另一為垂直方向的切線速度。雷達只能觀測到其中的逕向速度。圖中紅色箭頭代表物體的總運動速度 V ，黃色箭頭為可被雷達偵測到的逕向速度 $V * \cos(\theta)$ ，綠色箭頭則為雷達無法觀測的切線速度 $V * \sin(\theta)$ 。此分解方式有助於說明目標移動時，雷達所接收到的逕向速度變化。

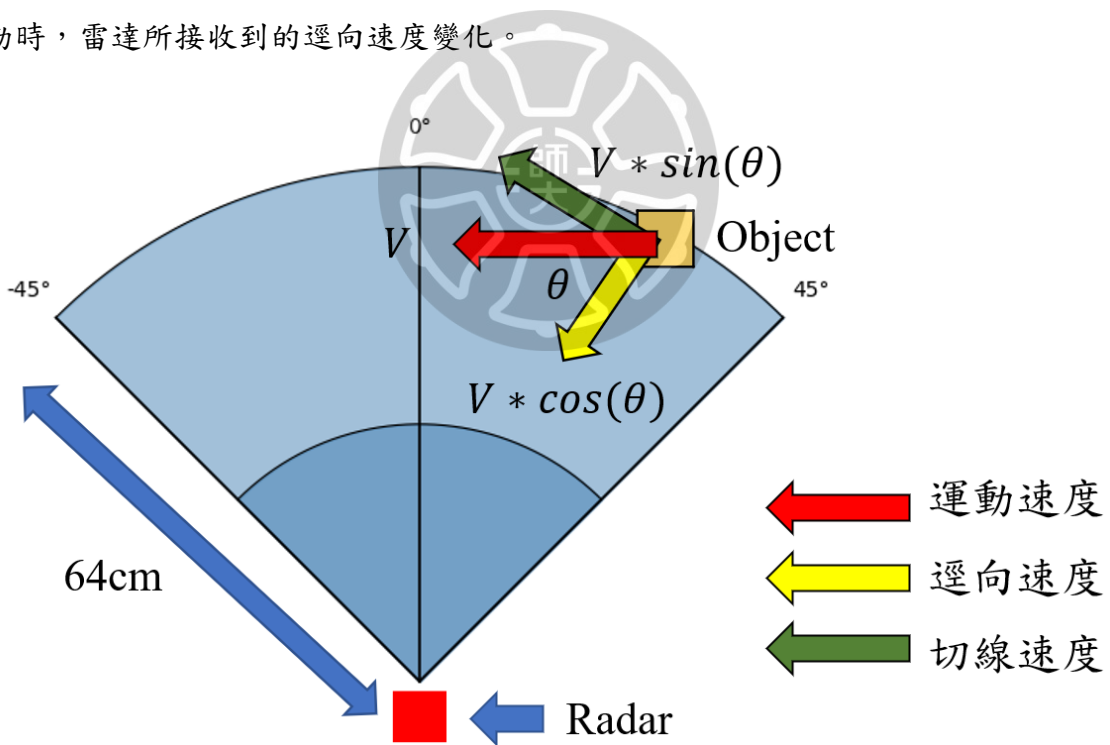


圖 2-4 運動速度與其逕向與切線速度分量關係示意圖

如圖 2-5 所示，當目標物以相同速率（10 m/s）在不同水平位置移動時，其可被雷達觀測到的逕向速度會因角度不同而產生變化。圖中以紅色箭頭表示總運動速度、黃色箭頭為逕向速度、綠色箭頭則為與雷達方向垂直的切線速度。

從圖右側的物體開始觀察，其移動方向與雷達方向成 60 度角，逕向速度為 5m/s，僅為總速度的一半。當物體的運動方向與雷達方向垂直時（如位於畫面中央並沿水平方向移動），逕向速度為 0 m/s。繼續往左觀察，另一側物體雖然與右側對稱，其與雷達方向同樣呈 60 度角，因此逕向速度同樣為 5 m/s，但方向相反。

由此可見，逕向速度的大小會隨著物體運動方向與雷達方向之間的夾角而變化。當物體的移動方向與雷達方向一致時（夾角為 0 度），逕向速度最大；當夾角變大時，逕向速度則逐漸減小。這反映了雷達僅能測量物體在其正前方方向上的速度分量，無法觀測垂直方向的運動。因此，若需完整掌握物體的實際移動軌跡，仍需搭配其他參數（如入射角）進行整合分析。

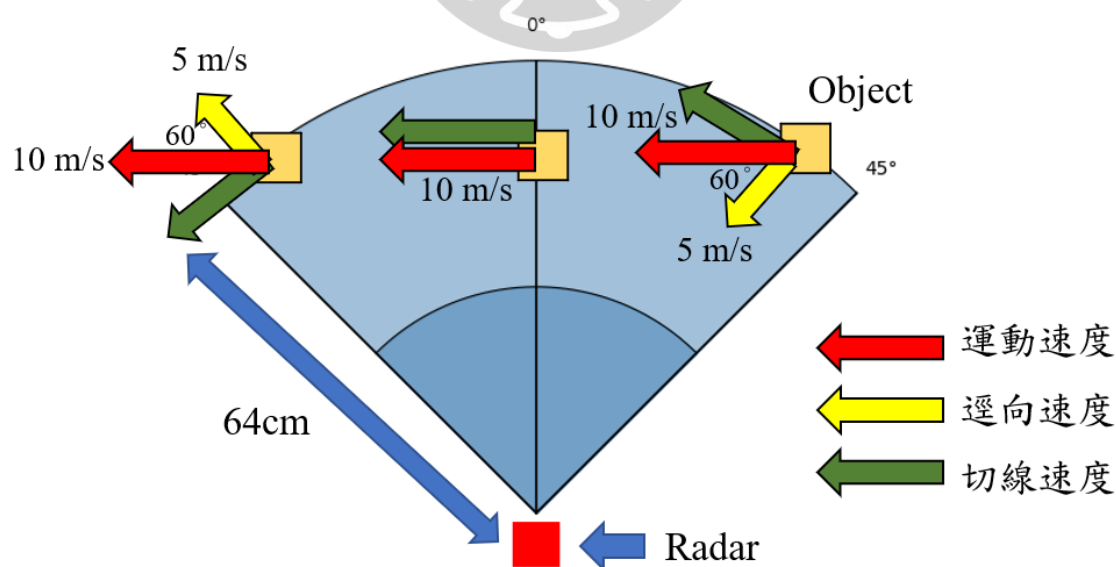


圖 2-5 不同空間位置下相同運動速度目標的逕向速度變化示意圖：圖中展示三個目標雖具相同運動速度，但因位置不同與雷達方向形成不同夾角，導致逕向速度不同。

2-3-3 入射角

入射角 (Angle of Arrival, AoA) 描述反射訊號自物體返回至雷達天線時的夾角，是用來判定物體在水平視野中相對方位的重要依據。透過多接收天線配置，毫米波雷達能夠量測反射波在不同天線間的逕向距離差，進而推估目標的入射角度。

如圖 2-6 與圖 2-7 所示，當目標物於水平視野中不同位置移動時，其反射訊號進入接收天線的角度，即入射角 (Angle of Arrival, AoA)，亦隨之變化。圖 2-6 示意物體由視野左側向右側移動時，入射角由負變正；圖 2-7 則呈現物體自右向左移動的情形，其入射角由正變負。此圖說明，透過入射角變化，可明確掌握目標於水平方向上的相對位置與移動方向。

相較之下，前述的逕向距離與逕向速度雖能分別反映目標距離遠近與前後運動趨勢，但皆無法判別物體是位於雷達視野的左側、中間或右側，亦即無法提供橫向位置資訊。因此，入射角所補足的角度資訊對於目標定位與動作方向判斷具有關鍵作用，能顯著提升雷達系統對目標空間行為的整體掌握能力。

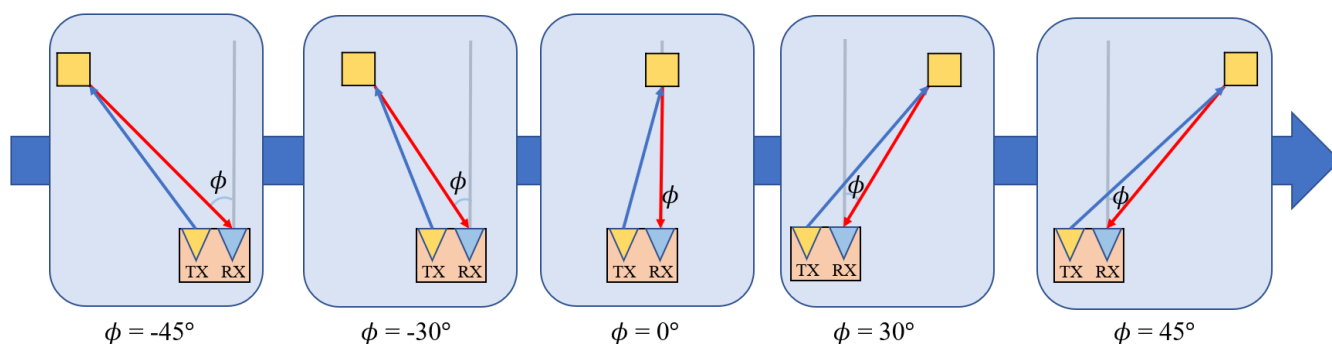


圖 2-6 左至右移動時之入射角變化示意圖：圖中展示物體自畫面左側移動至右側過程中，與接收天線之間所形成的入射角 ϕ 逐漸由負轉正，顯示其在水平視野中移動所對應的角度變化情形

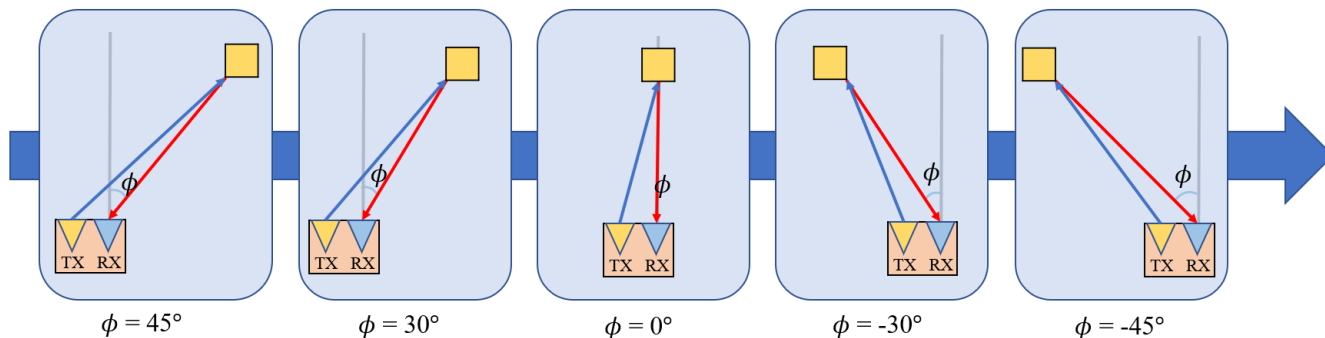


圖 2-7 右至左移動時之入射角變化示意圖：圖中展示物體自畫面右側移動至左側過程中，與接收天線之間所形成的入射角 ϕ 由正轉為負，說明可藉由入射角判斷其水平位置變化。

2-3-4 逕向距離、逕向速度和入射角關聯

毫米波雷達藉由逕向距離、逕向速度與入射角三項物理量，能夠建立物體在空間中的定位與動作輪廓。然若僅具備逕向距離與逕向速度資訊，實際上仍難以完整還原目標的空間位置。

雖然逕向距離與逕向速度可分別提供目標與雷達之間的直線距離與前後移動速度，然而這兩項參數皆僅反映目標在雷達直線方向上的變化，無法判斷其於水平方向上的相對位置。如圖 2-8 所示，左圖與右圖的兩個物體雖具有相同的逕向距離（40 公分）與逕向速度（4 m/s），實際上卻分別位於雷達視野的中間與右側。若僅依賴距離與速度資訊，系統將無從分辨其位於空間的哪一側，導致對目標空間行為的理解不完整。

入射角作為反映目標水平方位的關鍵參數，能有效補足逕向距離與逕向速度所無法涵蓋的橫向空間資訊。藉由入射角的判定，雷達得以在平面視野中區分目標位於左側、正前方或右側，並進一步建立具空間解釋力的角度座標系統。當逕向距離、逕向速度與入射角三者整合後，雷達系統便能同時掌握目標在空間中的相對位置、移動趨勢與方位方向，構成「位置—速度—方向」的完整觀測架構。此三參數結構為實現動態手勢辨識、目標追蹤與行為分析等應用的基礎條件，亦為毫米波雷達系統進行空間理解與動作感知之核心能力。

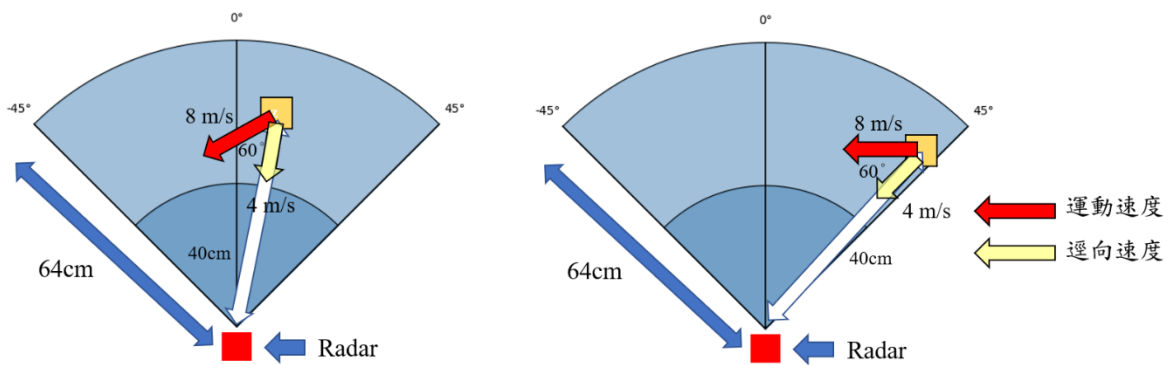


圖 2-8 具有相同徑向距離與徑向速度之不同位置目標示意圖

2-3-5 入射角計算方式

為了推算目標反射波進入雷達的入射角，雷達系統需具備至少兩個以上的接收天線。這是因為入射角的估算仰賴接收天線間所觀測到的反射訊號的徑向距離差。若系統僅具備單一接收天線，將無法比較不同位置的接收結果，自然也無從判斷訊號的來向，亦即無法估算目標在水平方向的相對方位。因此，多接收天線的配置是進行入射角估算的基本前提。

為了計算反射訊號的入射角，除了需具備至少兩個接收天線外，還需建立一項幾何上的重要假設：目標物與雷達之間的距離需足夠遠。在此條件下，反射訊號從目標返回時，對兩個接收天線而言，可假設其反射訊號以近似方向傳播至兩接收天線，即兩者所接收到的訊號具有近似的入射角 φ 。這使得我們可以利用兩接收天線之間所產生的徑向距離差 ΔL ，推回反射訊號的入射角度。

如圖 2-9 所示，假設觀測目標距離足夠遠時，當雷達中的發射天線 (TX) 發射訊號並遇到目標物產生反射後，訊號會以角度 φ 回傳至接收天線。若兩個接收天線 RX1 與 RX2 之間間距為 R ，根據幾何原理，回波會在 RX1 與 RX2 之間產生一段距離差 ΔL ，此值與角度間的關係可表示為公式(1)：

$$\Delta L = R * \sin(\varphi) \quad (1)$$

透過量測兩接收天線之間的距離差 ΔL ，即可反推出反射波入射的角度。此方法常作為毫米波雷達角度估算的基礎。是空間定位中的一項關鍵技術。

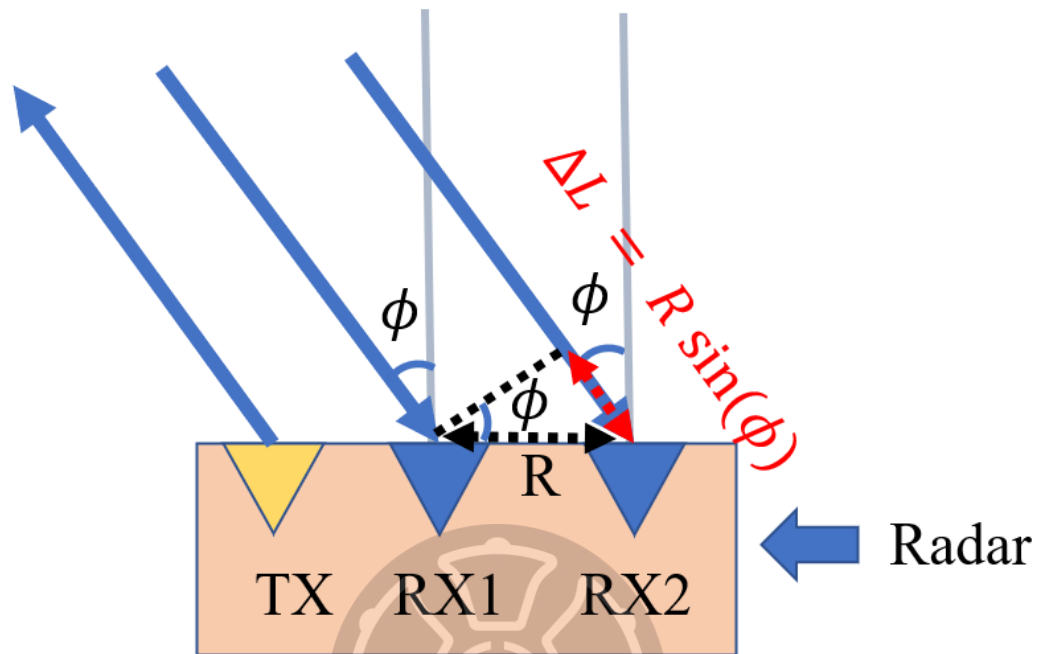


圖 2-9 入射角與距離差計算原理示意圖：TX 為發射天線，負責發射訊號；RX1 與 RX2 為接收天線，接收目標反射訊號。透過天線的幾何配置，可量測徑向距離差 ΔL ，進而推算入射角 ϕ

第三章 毫米波雷達開發工具

為有效推動毫米波雷達於近距離感測場域中的應用開發，開發工具的設計與操作效率扮演關鍵角色。本章所探討之 Ksoc Tool[5] (KaiKuTek system on Chip Tool) 為開酷科技 (KaiKuTek) 專為 K60168A 雷達模組所開發之圖形化操作平台，整合了雷達資料擷取、即時影像呈現、資料標註與儲存等多項功能，可支援完整的資料建構流程。透過簡潔明確的介面設計與高彈性的操作流程，使用者可快速完成目標資料的蒐集與管理，並針對各類應用需求調整參數設定與擷取方式。

Ksoc Tool 提供兩種主要的雷達影像類型：Range-Doppler Image (RDI) 與 Phase Difference Map (PHD)。RDI 影像揭示目標物體的逕向速度與逕向距離資訊，能即時反映物體運動狀態在時間序列中的變化；PHD 則透過多接收天線回波所形成的距離差，推估目標物在水平方向上的入射角度資訊，補足 RDI 無法分辨左右方位的限制。兩者結合後可形成具有空間與動態維度的雷達影像序列，成為應用開發中不可或缺的基礎特徵來源。

本章將詳述 Ksoc Tool 之操作介面與模組功能，並解析其所生成影像類型的空間架構與物理意義。透過參數設定、影像解析度與數據格式等層面說明工具在資料處理上的彈性與效能，進一步奠定後續演算法訓練與實作的基礎。

3-1 Ksoc Tool 平台簡介

Ksoc Tool[5] 為開酷科技 (KaiKuTek) 所研發之圖形化操作平台，專為其毫米波雷達模組 K60168A Dongle 所設計，提供完整的資料擷取、即時顯示與標註功能。該平台整合多種功能模組，包含雷達影像視覺化、錄製流程管理、自訂手勢標註與資料儲存選項等，支援開發者進行雷達資料的蒐集與標記工作。

圖 3-1 為 Ksoc Tool 的操作介面，介面左側為「資料擷取設定區」，使用者可

在此設定錄製幀數、模型名稱、手勢所屬者、使用手別與標註數量等參數，並透過下方按鈕進行錄製啟動或儲存控制。

介面右上方則為統計資訊區，會即時顯示目前已錄製的幀數與標註段落，讓使用者能清楚掌握資料錄製的進度與標註情形。

右下方為雷達影像顯示區，呈現兩種即時特徵圖(Feature Map)。左圖為 Range-Doppler Image (RDI)，其橫軸為逕向速度，縱軸為逕向距離，顯示各位置上的雷達反射能量分佈，反映物體的移動方向與前後距離。右圖則為 Phase Difference Map (PHD)，是由兩個接收天線的回波資訊進行比對，藉由估算各距離單位的逕向距離差，推測對應的入射角資訊，用以補足 RDI 所無法分辨的水平方向位置變化。

透過此操作介面，使用者不僅能同步擷取與觀測影像資料，亦能搭配即時標註機制，精確記錄動作發生時間與範圍。此平台具備完整的資料處理與管理功能，為本論文實作毫米波雷達手勢辨識系統的重要基礎工具。

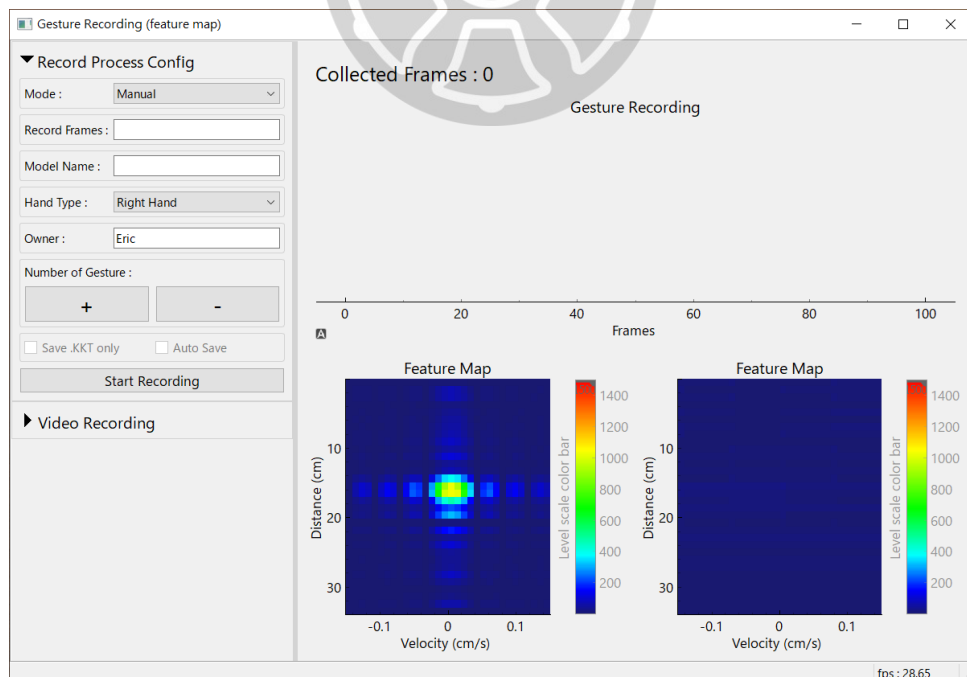


圖 3-1 Ksoc Tool 操作介面圖：左側為資料擷取設定區，右上方為錄製進度資訊，右下方為即時影像顯示區，其中左圖為 RDI，右圖為 PHD，呈現資料收集與影像觀測之整合操作介面

3-2 雷達影像類型與物理意義

毫米波雷達經由信號處理後可生成多種影像資料，用以反映目標物在空間中的相對位置與運動特性。為協助開發者觀測與理解這些雷達影像資訊，Ksoc Tool 提供即時顯示功能，可同步呈現兩種影像格式：Range-Doppler Image (RDI) 與 Phase Difference Map (PHD)。本節將介紹這兩種影像的結構與物理意涵，作為後續應用設計與資料處理的基礎。

3-2-1 Range-Doppler Image (RDI)

Range-Doppler Image 是由雷達反射訊號經快速傅立葉轉換 (FFT) 處理後所生成的二維影像，其橫軸對應逕向速度 (Radial Velocity)，縱軸對應逕向距離 (Radial Distance)，而影像中每一像素的強度代表該位置的反射能量大小。此影像能夠同時呈現目標物與雷達之間的距離變化與速度資訊。

如圖 3-2 所示，RDI 的縱軸為逕向距離，單位為公分 (cm)，由上至下分別對應從雷達最近處 (0 cm) 到最遠處 (64 cm)；橫軸則為逕向速度，單位為公尺每秒 (m/s)，由左至右從 -8 m/s 至 $+8\text{ m/s}$ ， 0 m/s 位於中央，代表靜止目標。左側為負速度，代表物體正向雷達靠近；右側為正速度，代表物體遠離雷達。此速度分量為「逕向」方向，即沿雷達觀測軸線的投影分量，無法反映橫向移動的真實速度大小。

此圖中的每個格點皆代表一個特定的距離與速度組合區塊，其值即為該區塊的反射能量強度。在實際應用中，當一個物體靠近雷達，其反射訊號會在圖的左半部 (負速度區) 且靠近上方 (近距離) 的位置出現；反之，當物體遠離時，則會出現在右半部 (正速度區) 且逐漸往下方延伸的位置。

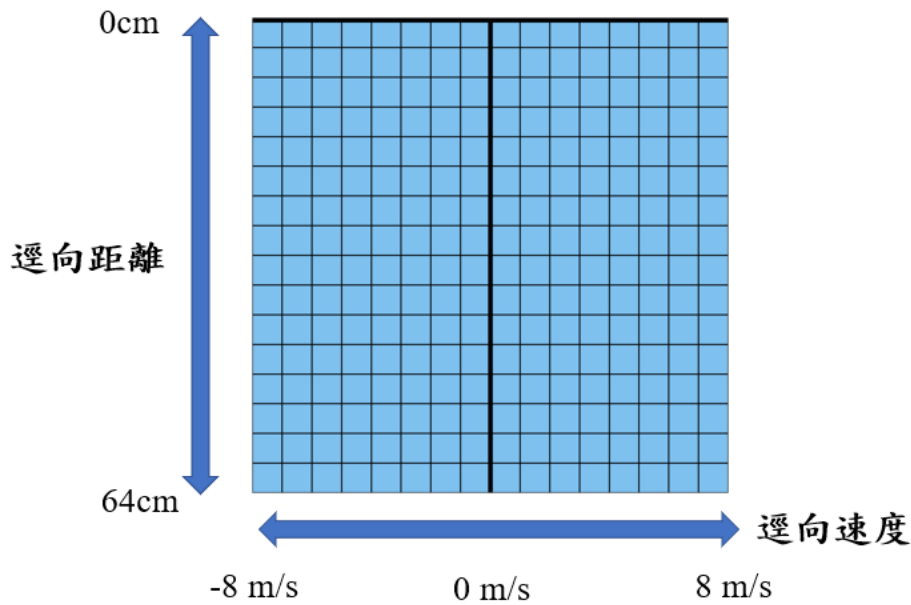


圖 3-2 Range-Doppler Image 座標示意圖：橫軸為逕向速度（-8 至 +8 m/s），縱軸為逕向距離（0 至 64 cm），左側代表逕向速度靠近雷達，右側代表遠離，此為舉例用之示意圖，座標設定與實際 RDI 不同，僅供此章節說明使用

本論文後續將以圖 3-2 所示之 RDI 座標架構作為範例解釋基礎，逐步說明不同移動方式或動作情境下所對應的影像變化。為簡化後續圖示與說明，本論文範例皆以 $16 * 16$ 的影像維度進行繪製與分析；此尺寸符合常見 FFT 處理中採用之 2^n 維度格式，有助於清楚觀察各區塊變化。然而，在實際使用 Ksoc Tool 進行資料擷取與模型訓練時，系統預設輸出的 RDI 維度為 $32 * 32$ ，能提供更細緻的距離與速度分佈資訊。

3-2-2 Range-Doppler Image 參數與解析度說明

在觀測與分析 Range-Doppler Image 時，除了掌握其基本座標意義外，更需進一步理解其影像範圍與解析度設定。不同的距離與速度範圍，會直接影響雷達可以看多遠、分多細，因此在進入動作範例之前，需先解釋這些影像的基本設定。

如圖 3-3 所示，RDI 為一張橫軸為速度、縱軸為距離的二維影像。圖中橫軸的速度範圍為 -8 至 $+8$ 公尺/秒，表示系統可觀測的最大接近與遠離速度皆為 8 公尺/秒；縱軸距離範圍則從 0 至 64 公分，代表此設定下最遠可感測至 64 公分的目標。整張圖共劃分為 16×16 個格點，因此可知每個水平方向的格子間距對應 1 公尺/秒的速度變化，垂直方向則對應 4 公分的距離變化。這兩個數值即為此設定下的速度與距離解析度，分別代表系統可區分的最小速度與距離。解析度越細，圖像越能呈現物體移動與位置的細節，提升觀測精度；但若解析度較粗，則容易將鄰近的訊號混合，使得動作變化不易判讀，影響辨識表現。因此，解析度不僅決定了每格所代表的物理範圍，也直接影響整體影像對物體運動的呈現能力。

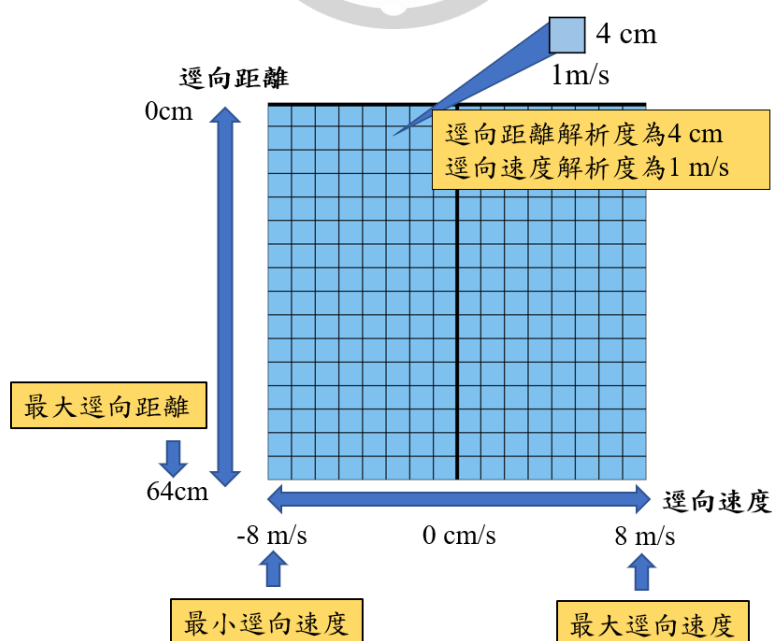


圖 3-3 Range-Doppler Image 各參數說明圖：展示 RDI 中速度與距離軸的範圍與每格代表的解析度，此為說明用示意圖，與實際 RDI 尺寸與解析度不同

在 Ksoc Tool 的操作介面中，使用者可切換不同的設定類型，以選擇符合需求的觀測模式。這些設定的主要差異來自於所使用的頻寬大小 (B)，頻寬的變化會直接影響系統的距離解析能力與可偵測的最大距離。

如圖 3-4 所示，Ksoc Tool 提供多組觀測設定可供選擇，使用者可透過操作介面中「Hardware Setting」的下拉選單切換不同的設定檔。這些設定檔對應的詳細參數如圖 3-5 所列，共包含四組設定，其主要差異在於頻寬 (B) 大小對距離相關參數的影響。頻寬越大時，距離解析度越細 (如設定 1 的 2.14 cm)，但可偵測距離較短 (34.29 cm)；反之，頻寬越小，則解析度降低但最大偵測距離增加 (如設定 4 可達 240 cm)。

在所有設定中，速度軸的參數 (最大速度與速度解析度) 保持一致，因此主要影響的是 RDI 影像中的距離軸比例與成像範圍，使用者可依據實際應用場景選擇最合適的設定檔。

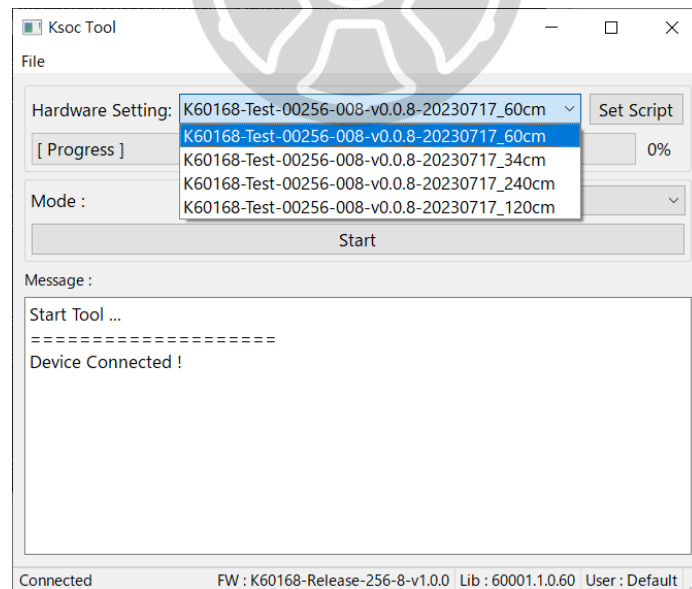


圖 3-4 Ksoc Tool 設定檔選擇介面：使用者可於「Hardware Setting」選單中選取不同觀測設定，依應用需求調整最大感測距離與距離解析度

設定種類	頻寬 B (GHz)	最大逕向距離 (cm)	逕向距離解析度 (cm)	最大逕向速度 (m/s)	逕向速度解析度 (m/s)
1	7.0	34.29	2.14	6.104	0.381
2	4.0	60	3.75	6.104	0.381
3	2.0	120	7.5	6.104	0.381
4	1.0	240	15.0	6.104	0.381

圖 3-5 各觀測設定之參數對應圖：列出四組可選擇的設定檔所對應之頻寬、最大逕向距離與距離解析度，其中速度相關參數在各組設定中保持一致

從圖中可見，四組設定中，速度相關的參數（最大逕向速度與速度解析度）保持一致，僅距離相關的範圍與精度隨頻寬而調整。本論文採用的為設定 1（頻寬 7 GHz），其具有最小的距離解析度（2.14 cm）與最短的最大逕向距離（34.29 cm），非常適合用於近距離的精細動作辨識，例如手部動作與手勢變化。此設定能完整捕捉手勢在空間中的變化，有助於提昇手勢辨識的準確性與穩定性。

3-2-3 單一目標於 RDI 上的基本成像行為

本節將依圖 3-2 所建立的 RDI 座標架構，舉例說明不同動作情境下在 RDI 圖上所顯示的變化趨勢。範例皆基於 16×16 的影像維度繪製，並簡化每個目標在 RDI 上僅對應為單一亮點顯示，藉以觀察速度與距離的變化如何對應至 RDI 圖上的訊號位置與形狀。

如圖 3-6 所示，目標物位於雷達正前方偏右上位置，實際距離為 42 公分，且以每秒 4.5 公尺的速度沿其移動方向前進。在此情境下，該速度的方向與雷達視線完全一致，因此此處的移動速度即為逕向速度，無需進行角度投影轉換。根據這樣的空間位置與速度關係，該目標於 RDI 上的成像將落在距離約 42 公分、速度約 +4.5 公尺/秒的位置，對應於圖像右下區域偏右方的單一亮點。此範例清楚展現當物體的運動方向與雷達視線一致時，其逕向速度與實際移動速度相等，RDI 上的定位也最為直接與準確，有助於進行清楚的目標識別與動態分析。

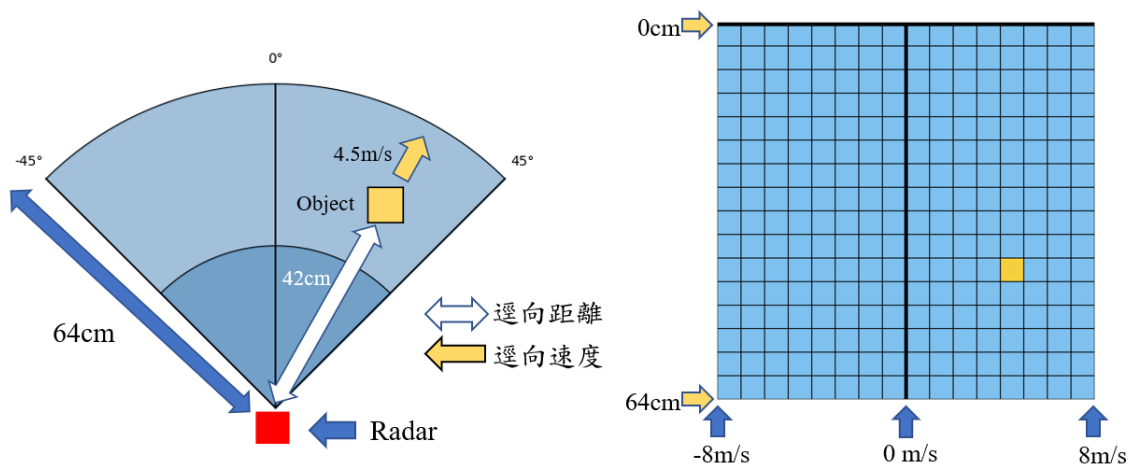


圖 3-6 目標沿雷達軸線方向遠離時於 RDI 上之成像位置：模擬目標以 4.5 m/s 速度由距離 42 cm 處遠離雷達時的 RDI 成像，亮點出現在右下區域，為速度與雷達視線一致下的遠離情境範例

如圖 3-7 所示，目標物與圖 3-6 中相同，皆位於雷達右前方，距離約為 42 公分，移動方向亦朝右上方。然而在此範例中，物體的移動速度為每秒 1.5 公尺，遠低於圖 3-6 所示的 4.5 公尺/秒。由於兩者位置與移動方向相同，因此其逕向距離相同，對應至 RDI 上的距離軸位置亦一致；但逕向速度不同，使得訊號亮點在速度軸上的位置明顯改變，從原本偏右的高速度區域，轉為靠近中央的低速度區域。此範例說明在距離不變的情況下，目標物若速度改變，其在 RDI 上的成像位置會發生平移，速度越快則亮點越遠離中央，速度越慢則越靠近 0 m/s 區域。

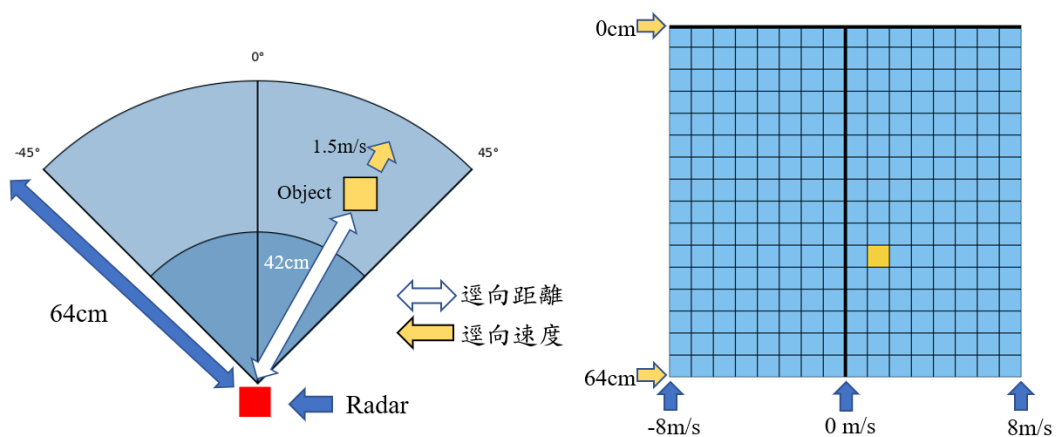


圖 3-7 不同逕向速度對應於 RDI 上不同位置：模擬目標以較低速度 (1.5 m/s) 由距離 42 cm 處遠離雷達，其在 RDI 上成像位置較圖 3-6 更接近中央，顯示在相同距離下，速度越慢，成像越靠近 0 m/s 區域

如圖 3-8 所示，當目標物正對雷達方向以每秒 5.5 公尺的速度接近，且與雷達的距離為 42 公分時，由於其移動方向與雷達視線完全一致，因此其實際速度即為逕向速度，無需進行任何投影計算。此時在 RDI 上，該目標的亮點將落在負速度區域，代表靠近雷達，並且位於距離軸中段偏下的位置，對應其實際距離。與前述遠離情境相對，當逕向速度為負時，亮點會往圖像左側偏移，且速度越快，偏移程度越大。本範例說明當物體沿雷達軸線方向靠近時，其在 RDI 上的成像會出現在負速度範圍，與遠離方向明確區分。

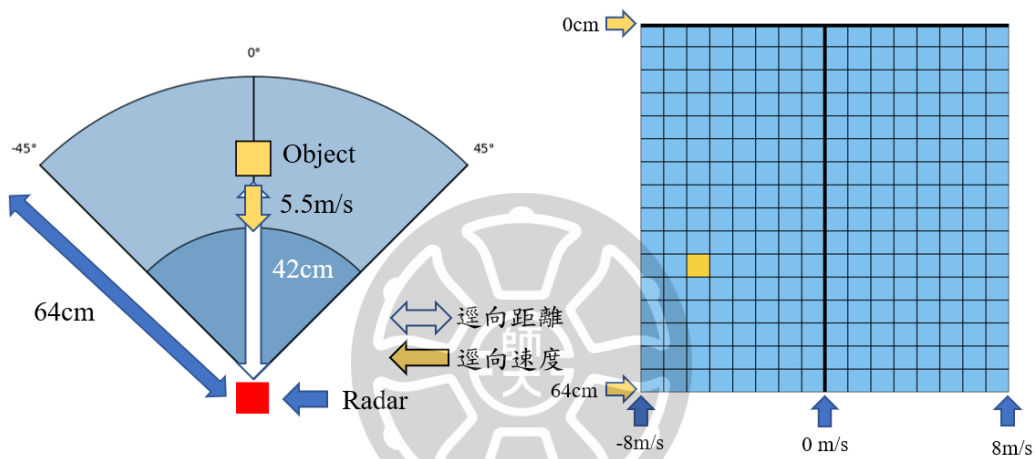


圖 3-8 目標沿雷達軸線方向靠近時於 RDI 上之成像位置：模擬目標以 5.5 m/s 速度由距離 42 cm 處靠近雷達，其 RDI 成像落於左下區域，顯示當逕向速度為負時亮點位於負速度範圍

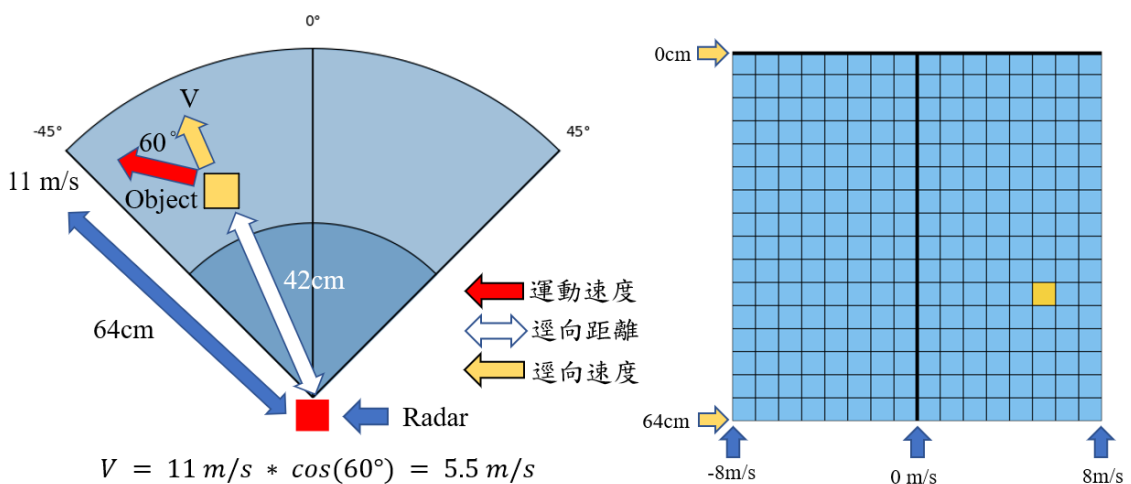


圖 3-9 逕向速度為實際速度與雷達視線夾角投影於 RDI 上之成像位置：模擬目標以 11 m/s 速度、與雷達視線夾角 60 度方向遠離，實際速度經投影後逕向速度約為 5.5 m/s

如圖 3-9 所示，目標物以每秒 11 公尺的速度朝左上方移動，與雷達視線形成 60 度夾角。儘管實際移動速度較快，但雷達僅能觀測其速度在視線方向上的分量，也就是逕向速度。此範例中，逕向速度為 $11\text{m/s} \times \cos(60^\circ) = 5.5\text{m/s}$ ，且因移動方向為遠離雷達，故為正值。由於物體與雷達之間的距離仍為 42 公分，因此在 RDI 上的亮點位置將落於距離軸中段偏下，速度軸上則對應約 +5.5 公尺/秒的位置。此圖明確說明毫米波雷達所呈現的速度為逕向速度，並無法反映目標實際移動方向與完整速度大小，僅能捕捉與雷達正前方方向相關的速度分量。

如圖 3-10 所示，三個目標物皆位於與雷達距離相同（約 42 公分）的位置，並以相同的速度（每秒 3.5 公尺）向外移動，但其運動方向分別對應雷達視野的左側、正前方與右側。由於雷達僅能觀測物體在其正前方方向上的速度分量，因此這三個目標在 RDI 上的成像位置將完全相同。儘管它們實際的移動方向不同，逕向速度卻一致，皆為約 +3.5 公尺/秒，距離亦相同。故在 RDI 影像上，這三個目標所對應的亮點將重疊在同一個位置，位於距離約 42 公分、速度約 +3.5 公尺/秒的區塊中。此範例說明當目標具有相同的逕向速度與逕向距離時，無論其實際運動方向為何，在 RDI 上皆無法區分，強調了 RDI 對於橫向空間辨識的限制。

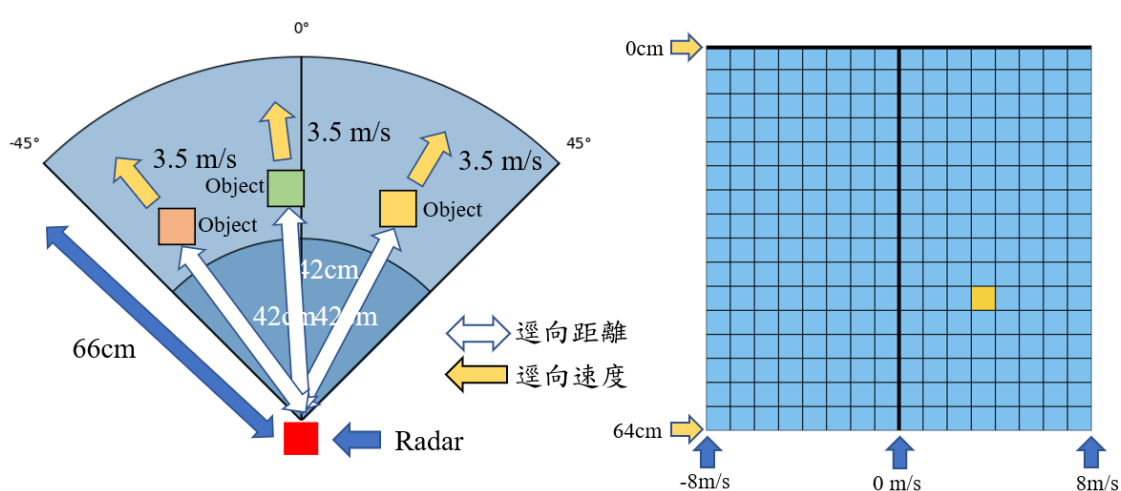


圖 3-10 多目標具相同逕向距離與速度時於 RDI 上之成像重疊：模擬三個目標分別自左側、正前方與右側以相同速度與距離移動，RDI 成像亮點完全重疊，顯示 RDI 無法區分不同橫向方向的目標運動

3-2-4 小手勢之垂直方向動作範例

本節分析手部在靠近雷達、遠離雷達與往返三種動作情形，其在 RDI 上呈現具規律性的速度與距離變化軌跡。

如圖 3-11 所示，當使用者手部從遠處向雷達方向直線移動時，手掌與雷達之間的逕向距離逐漸縮短，且移動方向與雷達視線完全對齊。此過程通常從靜止開始，隨後逐漸加速移動，接近雷達後再逐漸減速並停止。因此，手部的逕向速度從 0 開始，變為較大負值後再回到 0。對應在 RDI 上，會形成一條由下方往左上彎曲上升的半圓形路徑。此軌跡反映出距離與速度的同步變化，圖中標示出該變化軌跡，可作為動態動作分析與時間序列建模的重要依據。

如圖 3-12 所示，當使用者手部從近處向遠端方向移動時，手掌與雷達之間的逕向距離逐漸拉遠，且移動方向與雷達視線完全對齊。此時逕向速度為正，距離則隨時間增加，對應在 RDI 上將呈現一條由上往右下彎曲下降的訊號軌跡。隨著手部遠離，亮點由圖像上方向下移動，並從中央略為右移再回到中央，反映速度與距離的同步變化。由於手勢從靜止開始、加速、再減速至靜止，整體軌跡呈現半圓形。

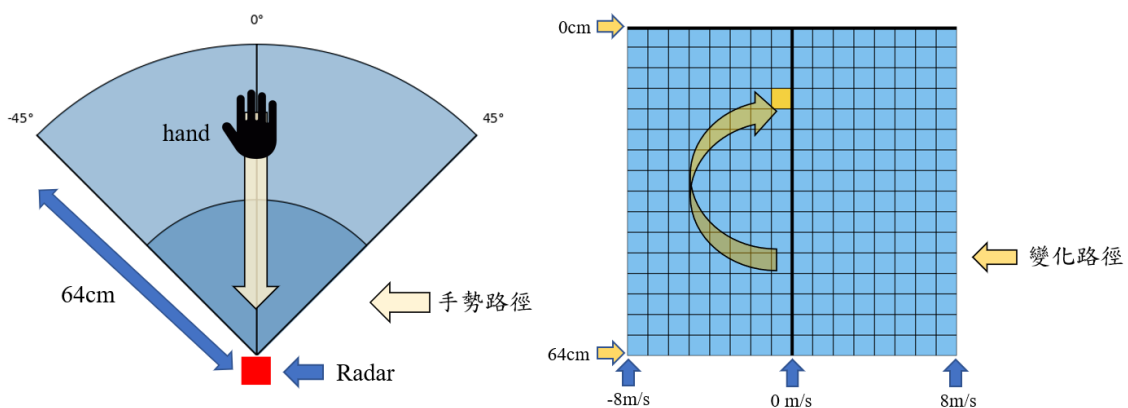


圖 3-11 手部直線接近雷達時於 RDI 上的動態變化軌跡：模擬手部從遠處靜止出發，加速向雷達靠近後再減速至靜止，RDI 成像呈現由下往左上方上升的半圓形軌跡，反映距離逐漸縮短與逕向速度先增後減的變化過程

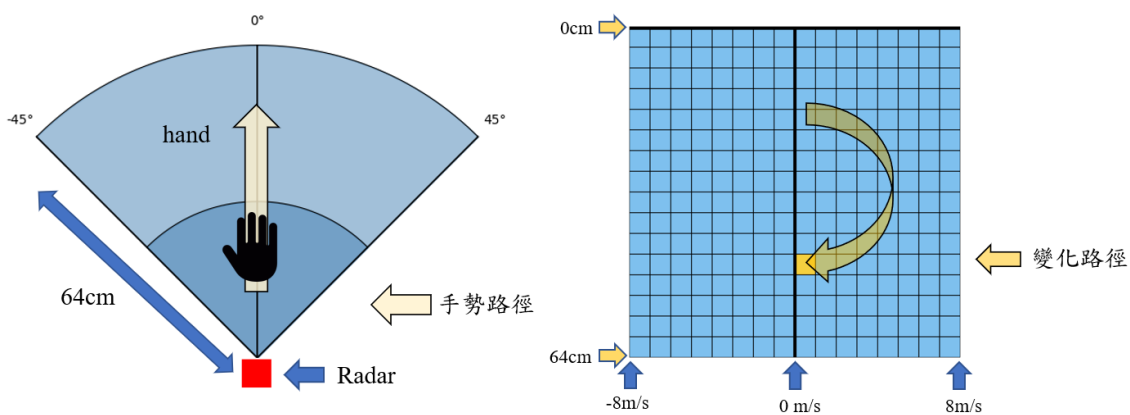


圖 3-12 手部遠離雷達時於 RDI 上的動態變化軌跡：模擬手部從近距離靜止開始，加速向遠處移動後再減速至靜止，RDI 成像呈現由上往右下方移動的半圓形軌跡，反映距離增加與逕向速度先增後減的遠離動作特徵

如圖 3-13 所示，當手部先向雷達靠近，再往後遠離時，RDI 上的訊號亮點將呈現圈形運動軌跡。這表示手部經歷了一段完整的往返移動過程：靠近時亮點由下往上、左移，遠離時則由上往下、右移，最終回到接近原點位置。此種亮點路徑變化可視為動作週期性的反映，亦有助於後續系統分析手勢的起始與結束階段。

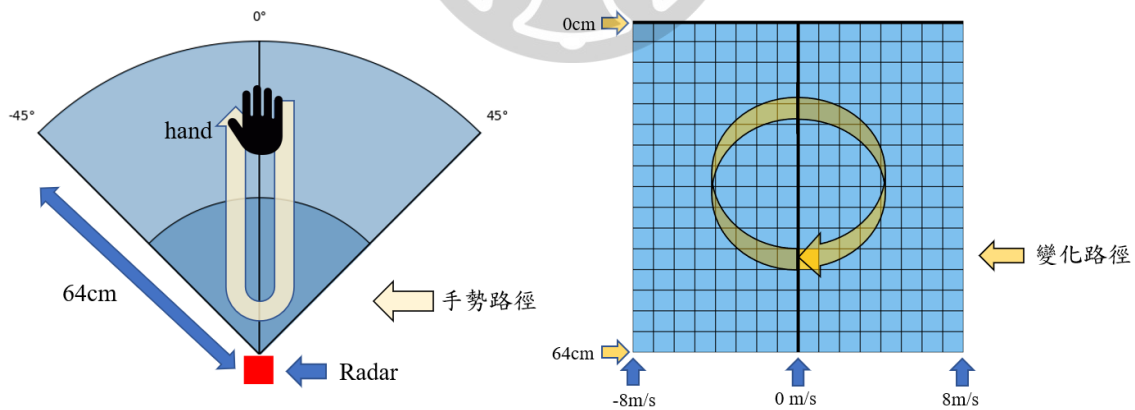


圖 3-13 手部靠近後再遠離時於 RDI 上的完整移動軌跡：模擬手部先加速靠近雷達再遠離，形成先左移後右移的動態亮點變化，RDI 成像呈現近似圓形的閉合軌跡，反映完整的往返移動過程

3-2-5 小手勢之水平方向移動動作範例

本節探討手部於視野範圍內，由靜止啟動再回到靜止的水平滑動行為，RDI 上呈現橢圓形軌跡，難以區分移動方向。

如圖 3-14 所示，當手部從右側向左側滑動時，首先觀察其逕向距離變化：由於雷達位於畫面下方中央，手部在起始與結束位置（兩側）時距離較遠，中間通過正前方時則距離最短，因此 RDI 亮點會呈現由下往上、再往下的弧形路徑，反映出一個對稱的逕向距離變化。

再觀察逕向速度變化：手部一開始處於靜止，雖與雷達夾角約 45 度，投影量 ($\cos \theta$) 尚大，但因實際速度小，逕向速度仍接近 0；隨後加速前進，實際速度逐漸增加，然而夾角也同時變大， \cos 值逐漸減小，導致逕向速度在中段雖然速度最大，投影量卻最小，整體變化呈現先變快（負值增加）、再變慢至 0 的特性。手部通過畫面中央後進入遠離階段，情形對稱但符號相反，最終回到遠處並靜止。

因此，整體 RDI 上的亮點軌跡呈現橢圓形，正是逕向距離與逕向速度雙重變化共同作用下的結果，為橫向滑動動作的重要影像特徵。變化特性。

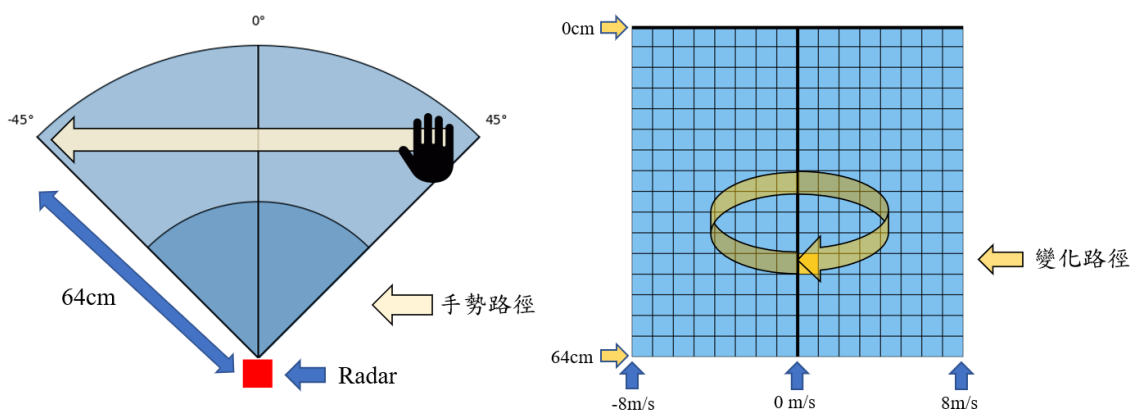


圖 3-14 左滑動作之 RDI 變化軌跡：模擬手部從右側靜止開始，向左側加速滑動後再減速至靜止，因距離先縮短再拉遠、逕向速度先變負再回到零，RDI 成像呈現橢圓形軌跡

如圖 3-15 所示，手部從左側滑向右側，其移動方向與圖 3-14 相反，但由於整體移動過程中逕向距離與逕向速度的變化趨勢一致，因此在 RDI 上同樣呈現橢圓形的變化軌跡。具體來說，手部從遠端（夾角約 45° ）出發，靠近中央時逕向距離逐漸縮短，而逕向速度則因實際速度與角度變化而先增加再減少，最終再次遠離雷達並回到遠端。此種速度與距離變化模式與圖 3-14 完全相同，導致兩種反向手勢在 RDI 上呈現幾乎一致的圖形特徵，進而造成無法僅依據 RDI 判斷物體移動方向的限制。這也說明了在橫向滑動情境中，需額外結合其他資訊才能正確辨識動作方向。

此類動作對應於「左滑」與「右滑」等水平方向的小手勢，但由於 RDI 無法單獨分辨方向，因此實際辨識上須結合 PHD 或其他時序特徵。

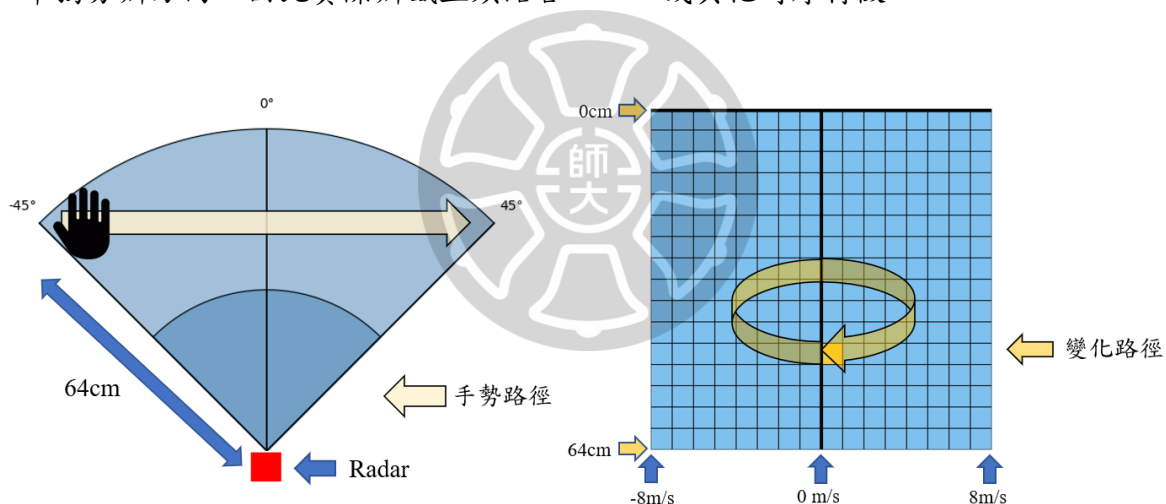


圖 3-15 右滑動作之 RDI 變化軌跡：模擬手部從左側靜止開始，向右側加速滑動後再減速至靜止，過程中距離先縮短再拉遠、逕向速度先為負再回到零，RDI 成像與左滑動作相同，顯示 RDI 難以判斷橫向滑動的方向

3-2-6 大手勢之水平方向移動動作範例

本節說明當手勢以固定速度穿越雷達視野的水平滑動情境，其 RDI 成像會因入射角變化而呈現特定軌跡模式。

如圖 3-16 所示，此範例展示一個以等速從右側移動至左側的物體，其在不同位置皆維持相同的運動速度（紅色箭頭）。由於移動方向與雷達視線夾角會隨位置變化，因此對應的逕向速度（黃色箭頭）也會改變。物體在起始與結束階段夾角約為 45 度， \cos 值較大，所對應的逕向速度最大；而移動至中央時與雷達視角垂直，夾角為 90 度， \cos 值為 0，逕向速度亦為 0，形成一段從負值逐漸趨近於零再轉為正值的速度變化過程。

如圖 3-17 所示，物體以等速從左側向右側移動，與圖 3-16 的移動方向相反。儘管其實際運動路徑呈現對稱關係，但由於物體的運動速度為定值，且其在移動過程中的夾角變化一致，因此兩者的逕向速度變化相同：皆為從負值逐漸趨近 0，再轉為正值。

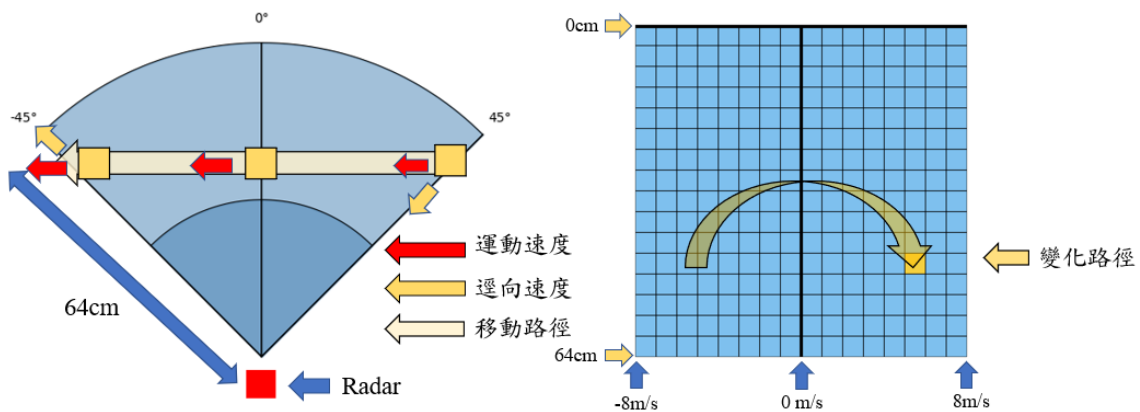


圖 3-16 等速橫向左移物體之 RDI 投影路徑示意圖：模擬物體以固定速度從右側向左側移動，因與雷達視線夾角變化導致逕向速度從負值逐漸趨近零再轉為正值，RDI 成像呈現由左至右的橢圓形軌跡，反映速度方向投影隨位置變化而改變的現象

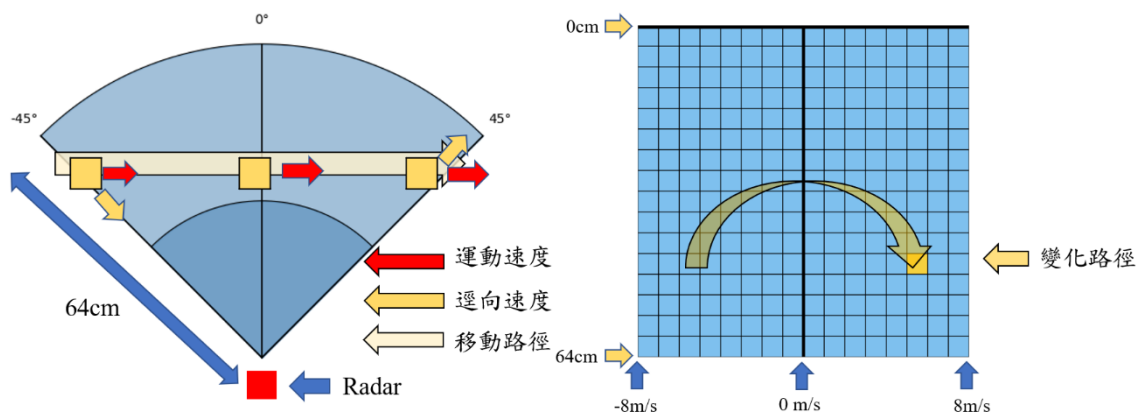


圖 3-17 等速橫向右移物體之 RDI 投影路徑示意圖：模擬物體以固定速度從左側向右側移動，徑向速度隨與雷達視線夾角變化，從負值逐漸趨近零再轉為正值，RDI 成像與圖 3-16 相同，呈現一致的橢圓形軌跡，顯示 RDI 無法辨別等速橫向移動的方向

由於徑向距離與徑向速度變化一致，圖 3-17 在 RDI 上所呈現的亮點軌跡與圖 3-16 完全相同，皆為自左至右的半橢圓形曲線。此結果說明，僅觀察 RDI 影像，無法區分實際移動方向，需結合其他影像資訊（如 PHD）才能完整解析物體運動的真實行徑。

此類動作對應於從 FOV 外快速進入並離開的「揮手」或「掃過」等大手勢，其 RDI 軌跡與小手勢特徵不同。兩者關鍵差異在於：小手勢的起始與結束皆發生於 FOV 內，具備完整的靜止、加速、減速階段；大手勢則在 FOV 外完成啟動與停止，於視野內僅短暫高速通過，成像時間短、速度變化劇烈。圖 3-18 進一步以空間軌跡呈現兩者在雷達視野中的運動差異，小手勢具備完整動作歷程，大手勢則快速掃過，無明顯起止階段。

此外，毫米波雷達對整個手掌的動作變化較為敏感，能有效觀察大幅度運動所對應之 RDI 軌跡；但對於手指細部的靜態變化則較不敏銳。像「比 2」「比 5」等靜態手勢，由於未產生明顯的距離或速度變化，這類靜態手勢的手指與手掌重疊在同距離的 RDI 上，在 RDI 上僅呈現模糊亮點，難以與其他部位區分，無法形成穩定可辨識特徵。此限制也顯示本系統較適合辨識動態特徵明顯的手勢。

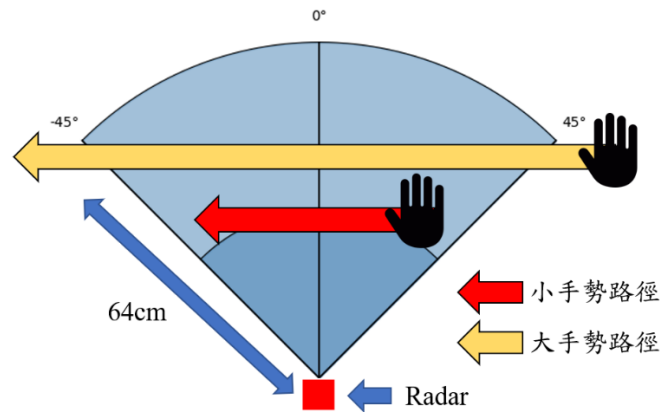


圖 3-18 小手勢與大手勢於雷達視野內之運動軌跡示意圖：小手勢從 FOV 內靜止起始並於 FOV 內靜止結束，動作歷程包含加速與減速階段；大手勢則於 FOV 外啟動並完成靜止動作，僅在 FOV 中短暫高速通過，整體為連續快速移動

3-2-7 Phase Difference Map (PHD)

Phase Difference Map (PHD) 是一種能補足 RDI 限制的角度資訊圖像，其核心依據是物體與兩接收天線之間的「逕向距離差」。當物體位於雷達正前方時，至兩天線的距離相等；若偏向左側或右側，則會產生細微的距離差異，透過這些資訊可判定物體在水平面上的入射角。

Ksoc Tool 中的 PHD 計算方式，是透過比較物體在兩個接收天線上所產生的逕向距離差，來推估其入射角位置。最終生成如圖 3-19 所示的影像。圖像的橫軸為入射角軸，中央為 0° ，向左為負入射角（物體偏左），向右為正入射角（物體偏右）；縱軸則為逕向距離軸，自上而下由近至遠。每個像素格即代表某一入射角與某一距離位置上的目標反射強度。

使用者可觀察不同時間點的 PHD 影像，分析亮點的橫向變化趨勢，進一步推論物體在空間中的移動方向。此功能特別適用於判斷物體是從左往右，還是從右往左運動，彌補 RDI 僅提供速度與距離資訊，無法分辨左右方向的缺陷。

透過 RDI 與 PHD 的互補，整體雷達影像能提供更完整的動作資訊，對後續動作辨識與手勢分析應用具有關鍵價值。

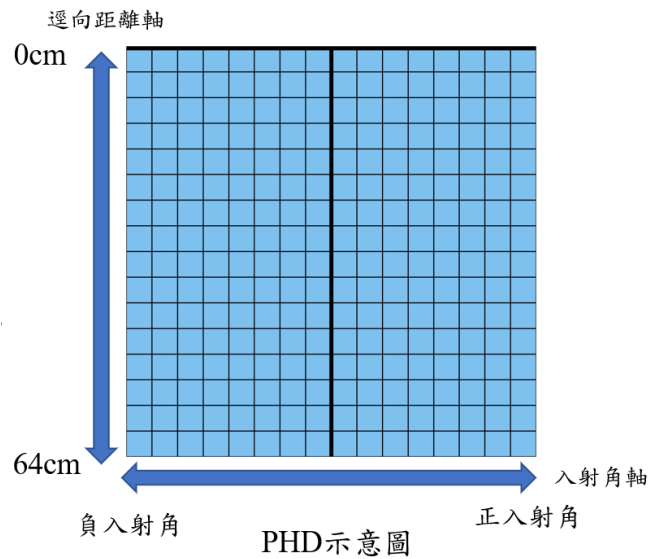


圖 3-19 Phase Difference Map (PHD) 坐標示意圖：橫軸為入射角方向，左側為負入射角（目標偏左）、右側為正入射角（目標偏右），縱軸為逕向距離，用於表示目標在不同水平位置與距離下的反射能量分佈，補足 RDI 無法辨別左右方向的限制

3-3 資料擷取與標註流程

在建立雷達影像資料集的過程中，資料標註的效率與準確性對後續的應用與分析品質具有關鍵影響。Ksoc Tool 提供一套整合式的資料擷取與即時標註介面，能夠同步記錄影像幀與標註資訊，大幅提升資料處理流程的整體效率與一致性。

圖 3-20 展示該介面的運作情形。左側為資料擷取剛開始的畫面，此時尚未接收到任何資料，畫面右上方的影像仍為空白；右側則為資料擷取完成後的畫面，已記錄完指定的 100 張影像幀。畫面右上方的條狀圖代表每一幀的標註狀態，其中紅色區段即為使用者即時標註的位置，表示該段資料被指定為標註區間，其餘藍色部分則代表未標註區段。

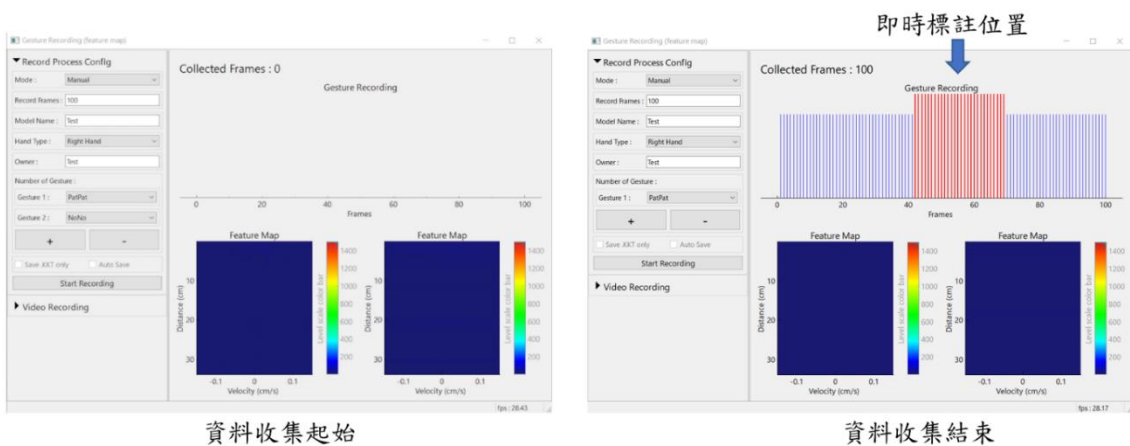


圖 3-20 Ksoc Tool 資料擷取與即時標註操作畫面：左圖顯示資料擷取初始狀態，右圖為完成後的畫面，右上方條狀圖標示各影像幀的標註情況，其中紅色區段為使用者即時標註之動作區間，藍色為背景區段

在進行資料擷取前，使用者可透過介面預先設定欲錄製的影像幀數、標註類別與標註次數等參數，並輸入相關資料名稱與記錄屬性。完成設定後，點選開始按鈕即開始進行雷達資料收集。錄製過程中，畫面會即時顯示所擷取之影像資訊，使用者可根據實際動作發生的時間點，操作標註鍵切換狀態，將某一時間區段標記為特定類別。系統將自動依據使用者操作，在長條圖上對應區段以紅色區塊表示，未標註部分則保持藍色。整體流程設計使得擷取與標註可於同一階段內完成，無需仰賴事後手動裁切或對齊時間，能有效提升標註效率並減少潛在誤差。

介面提供即時影像回饋功能，可協助使用者觀察當前所擷取的雷達特徵影像是否如預期，進一步提升資料品質的可控性。配合圖形化操作介面與直觀的流程設計，Ksoc Tool 提升使用便利性，對資料品質管理與後續分析流程具有實質助益。

3-4 資料格式與儲存方式

Ksoc Tool 在完成資料擷取與即時標註後，會自動將所有雷達影像與對應標註資訊儲存為結構化的 HDF5 格式檔案（副檔名為 .h5）。此格式具備良好的儲存效率與彈性，適合處理大量連續性的高維影像資料，同時能夠將資料與標註資訊整合於同一檔案中，便於後續模型訓練與分析流程使用。

雷達影像資料會以一組四維陣列的形式儲存，維度為 $(2, 32, 32, N)$ ，其中第一個維度大小為 2，分別代表 RDI 與 PHD 兩種影像類型；第二與第三維則為影像的空間解析度，預設為 32×32 ；最後一個維度 N 則對應於實際錄製的影像幀數，代表時間軸上的連續影像序列。

與影像資料對應的標註資訊則儲存在一個長度為 N 的一維陣列中，每個位置對應一幀影像，並記錄該影像是否屬於標註區段。標註採整數編碼方式表示，其中 0 表示背景區段，代表該幀未被標註為特定動作；1 則表示標註區段，對應使用者在錄製過程中所標記的有效資料範圍。由於影像資料與標註資料在時間軸上完全對齊，系統能夠確保每一幀影像皆具備明確標籤資訊，無需額外對齊或後處理。

圖 3-21 示意整體資料儲存結構，RDI 與 PHD 各自形成一組尺寸為 $32 \times 32 \times N$ 的影像序列， N 表示資料錄製的時間長度（即影像幀數）。兩組影像依序堆疊形成雙通道的時間序列輸入，並與一組長度為 N 的一維標註向量對齊。該標註向量中，每一元素皆對應至影像序列中同一時間點的資料，其中 0 表示背景區段，1 表示動作區段。

此結構設計具備良好擴充性，能有效支援批次讀取、訓練前處理與模型驗證等需求。透過此格式，使用者可快速建立具時間連續性與標註完整性的雷達影像資料集，提供深度學習模型一套整合、清晰且高品質的輸入來源，進一步提升整體資料處理與訓練流程的效率與穩定性。

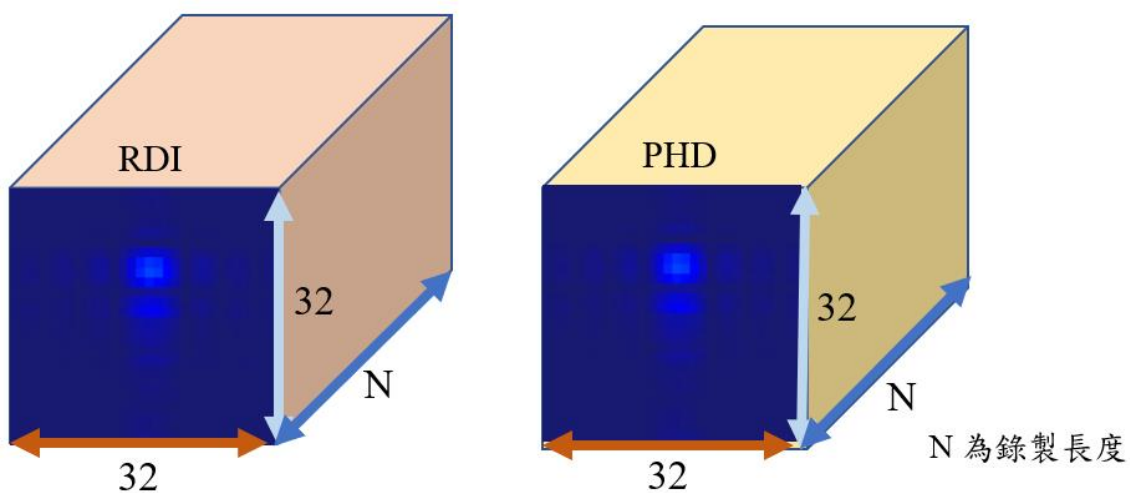


圖 3-21 RDI 與 PHD 資料儲存格式示意圖：展示資料以四維陣列儲存為雙通道時間序列，其中尺寸為 $(2, 32, 32, N)$ ，2 代表 RDI 與 PHD 兩個通道， 32×32 為影像空間尺寸， N 為影像幀數，標註資訊以一維陣列同步記錄每幀資料的類別，用於支援後續模型訓練與分析流程



第四章 基礎應用實例及結果

本章介紹本論文設計之毫米波雷達手勢辨識系統的應用實例與實驗結果。首先說明現有雷達手勢辨識系統的侷限，以 Google Soli 雷達[6]為例，指出其手勢類別固定、無法擴充及缺乏起訖判斷等問題，進而對比本系統在彈性與即時性上的優勢。接著依序說明手勢定義與資料蒐集、標註與 soft label 生成、資料切片與 Ground Truth 對應策略，並詳述所採用之 3D CNN 模型架構與訓練流程。

為驗證系統效能，本章亦展示預測結果與後處理方式，包含時間序列預測變化、雙閾值設計與分類準則，以及最終於測試資料上的混淆矩陣分析。最後，說明本系統如何整合雷達裝置與模型推論流程，實現即時手勢辨識系統，並透過圖形化介面即時回饋使用者操作結果，驗證系統於互動應用中的實用性與準確性。

4-1 毫米波雷達手勢辨識之侷限

現有毫米波雷達於手勢辨識應用中，雖已展現不受光源干擾、具穿透性與隱私保護等優勢，但目前主流技術在即時推論 (online inference) 方面仍普遍存在限制。

文獻[7]所提出之方法能有效從完整手勢片段中萃取高區辨性特徵，卻需待整段資料收集完成後才能啟動分類，缺乏對連續資料流的低延遲設計，因而難以滿足動態互動場景中手勢接續與快速回饋的需求。

Soli 雷達[6]採用預先定義的單一完整手勢為推論單位，雖能在靜態、分段手勢上達到高準確度，但在手勢過渡區段與背景動作等即時變化情境下仍難以維持穩定辨識。

文獻[8][9]則提出空間-時序混合模型，針對預先分割的手勢片段進行離線訓練與分類，並達到高準確率，但文獻中未提及支援線上推論的資料流處理機制，無法直接應用於即時偵測或互動系統中。

相較之下，本論文提出之手勢辨識系統不僅具備即時推論能力，並透過滑動視窗與雙閾值判斷機制，能夠精準辨識出每段手勢的進入與退出時機（詳見第 4-5 節）。此外，本系統可依據不同需求靈活新增、調整手勢類別與訓練資料，解決手勢固定與辨識範圍有限之困境，提升應用彈性與辨識精度。

4-2 手勢定義與資料蒐集方式

本論文針對毫米波雷達動作辨識任務所需資料，設計一組具代表性且結構化之手勢資料集。該資料集涵蓋三類常見手勢動作：拍拍 (PatPat)、左右揮 (Wave) 與 招手 (Come)，以及一類背景動作 (Background)，共計四個分類。

如圖 4-1 所示，各類手勢皆設計為重複兩次之動作循環，藉此強化手勢在雷達影像中的時序特徵，提升辨識準確度與穩定性。相較於單次動作，重複動作能在時間序列中形成更加明顯且一致的特徵模式，有助於深度學習模型辨識與分類。

其中，拍拍 (PatPat) 為手掌朝向雷達、前推出再收回的往返動作；左右揮 (Wave) 為手掌朝向雷達、左右擺動的水平動作；招手 (Come) 則為手掌朝上、手指彎向自己的招引動作。每一類手勢皆由靜止狀態開始與結束，強調動作的完整性與可分辨性，並利於資料標註與模型訓練。

在資料蒐集方面，本研究邀請六位受測者（標記為 A 至 F）執行上述三種手勢與背景靜止動作，各自蒐集 100 筆影像序列，唯受測者 A 額外多蒐集 100 筆背景與手勢資料，以補足訓練需求。最終每個類別均蒐集 700 筆資料，共計 2800 筆，如圖 4-2 所示。所有資料皆採用統一的設定條件錄製，包括固定雷達位置、相同視野範圍與姿勢流程，以確保訓練樣本間的動作規則性與資料品質。

本資料集具備樣本數平衡、受測者多樣性與動作設計明確等優勢，作為後續模型訓練、驗證與即時辨識測試的基礎，並於後續進一步說明其應用與成效。

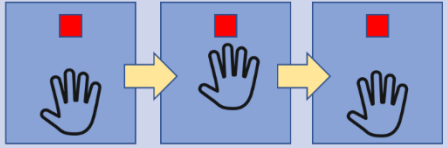
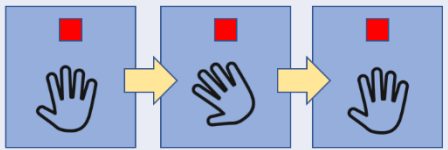
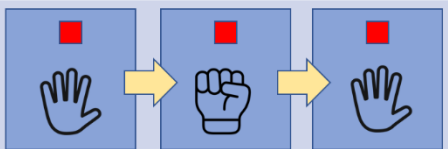
手勢名稱	手勢說明	動畫圖解
拍拍(PatPat)	手掌朝向雷達，向前推出再收回，重複兩次	
左右揮(Wave)	手掌朝向雷達，左右揮，重複兩次	
招手(Come)	手掌朝上，手指彎曲向自己，重複兩次	

圖 4-1 三類動作手勢定義與示意圖：本研究定義三種常見動作手勢類型，分別為拍拍 (PatPat)、左右揮 (Wave) 與招手 (Come)。各手勢皆設計為重複兩次，以增強特徵表現與穩定辨識效果。圖中紅色方塊為毫米波雷達位置，手掌與手指方向皆依實際動作流程繪製，提供手勢執行之動態

變化參考

收集者	拍拍	左右揮	招手	背景
A	200	200	200	200
B	100	100	100	100
C	100	100	100	100
D	100	100	100	100
E	100	100	100	100
F	100	100	100	100
總共	700	700	700	700

圖 4-2 各手勢類別與背景資料之蒐集者分布統計：展示本研究資料集中各手勢類別與背景資料的蒐集數量，並標示六位參與者 (A-F) 分別負責蒐集的比例。每類手勢與背景皆包含 700 筆影像序列資料，總計 2800 筆，確保數據分布均衡，有利於模型訓練之泛化能力

4-3 標註資料處理與 Ground Truth 建構

在動態手勢辨識任務中，準確且結構清晰的標註資料扮演關鍵角色。為了有效訓練深度學習模型，本論文針對毫米波雷達所擷取之時間序列影像資料，設計出具有一致性與可擴充性的標註格式，並進一步導入 soft label 機制，強化模型對過渡區段的感知能力。除此之外，為配合模型訓練需求，本論文亦設計滑動視窗式的資料切片策略，並定義各切片對應之 ground truth 的選取方式，確保資料處理流程與模型輸出具有高度對齊性。

本節將依序說明標註格式的設計、soft label 構建邏輯、滑動視窗切片機制與 ground truth 對應方式，並搭配圖示輔助說明，以呈現完整的資料處理流程與設計理念。

4-3-1 標註格式設計

本論文所使用的雷達影像資料均透過 Ksoc Tool 所提供的圖形化平台進行收集與即時標註。使用者可在資料擷取過程中同步標記動作發生的時間區段，使標註結果與影像幀在時間軸上精確對齊。透過此整合式操作流程，不僅提升了資料製作的效率，也確保了標註準確度與一致性。

標註內容以動作的實際起始與終止時間為界，將每一筆資料劃分為「手勢區間」與「背景區間」。其中，手勢區間指的是定義手勢從動作開始到動作完成的全程，而背景區間則包含三種情境：動作準備過程、非定義手勢（如臨時動作或轉場移動），以及靜止狀態或畫面中無任何活動的時段。這樣的區分方式有助於強化模型對有效手勢的學習，同時降低雜訊對辨識準確率的影響。

如圖 4-3 所示，整體時間軸上明確標示手勢區段與背景區段，利於後續視覺化與模型設計。圖 4-4 為實際資料格式範例，所使用之 .h5 檔案包含 DATA 與

LABEL 兩個主要欄位，其中 DATA 為形狀為 $(2, N, 32, 32)$ 的四維陣列，表示雙通道 (RDI 與 PHD) 序列影像；LABEL 為長度為 N 的一維整數向量，標記每一幀影像的類別，0 表示背景，1 起依序對應不同手勢類型。

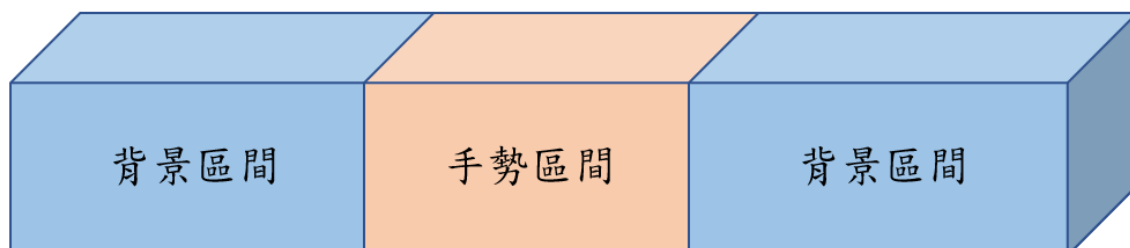


圖 4-3 時間軸標註示意圖：本圖展示 Ksoc Tool 於雷達資料擷取期間所對應之標註格式與時間軸分布情形。橫軸為影像序列的時間幀，紅色區段為使用者標註之手勢動作範圍，藍色區段則為背景資料，包括靜止狀態、準備動作與非定義動作。此標註方式可與影像資料完全同步，提升訓練資料的一致性與正確性

```
Function Load_H5_Data(h5_path):
    with h5py.File(h5_path, 'r') as f: # 開啟 .h5 檔案
        data = f['DATA'][:] # 讀取RDI 以及 PHD資料，shape : (2, N, 32, 32)
        labels = f['LABEL'][:] # 讀取標註資料，shape : (N,)
```

圖 4-4 雷達資料與標註結構示意：圖中展示實際儲存的資料結構，包含影像資料區 (DATA) 與標註向量區 (LABEL)。其中影像資料為大小為 $(2, N, 32, 32)$ 的四維張量，分別對應 RDI 與 PHD 雙通道影像；標註資料為長度為 N 的一維整數向量，用以標示每一幀所對應的類別。當值為 0 時，代表該幀為背景區段；值為 1 則代表該幀屬於使用者明確標註的有效手勢區段

4-3-2 Soft Label 建構方法

在動態手勢辨識任務中，資料標註往往存在模糊的邊界，特別是在動作開始與結束的過渡區段。傳統的 one-hot 編碼方式雖能提供清楚的類別定義，但對於手勢起始與收尾的模糊地帶較為敏感，容易造成模型於邊界處的判斷不穩定。為了解決此問題，本論文設計一套 soft label 標註策略，透過類高斯分布模擬手勢動作強度，強化模型對於動作邊界的辨識能力。

圖 4-5 展示整體 soft label 建構流程，起始資料為 Ksoc Tool 所輸出的標註向量，其中 0 表示背景區段，1 表示使用者標註的手勢區段。系統會依序搜尋連續為 1 的區段，並對每段手勢資料進行獨立轉換。轉換時，系統將整段資料視為對稱分布，根據其長度設定中心位置與標準差，再以類高斯函數對每個時間點賦予對應的權重。這些權重會經過正規化，使中間點為 1，兩側則逐漸遞減，呈現出類似動作強度的時間變化曲線。

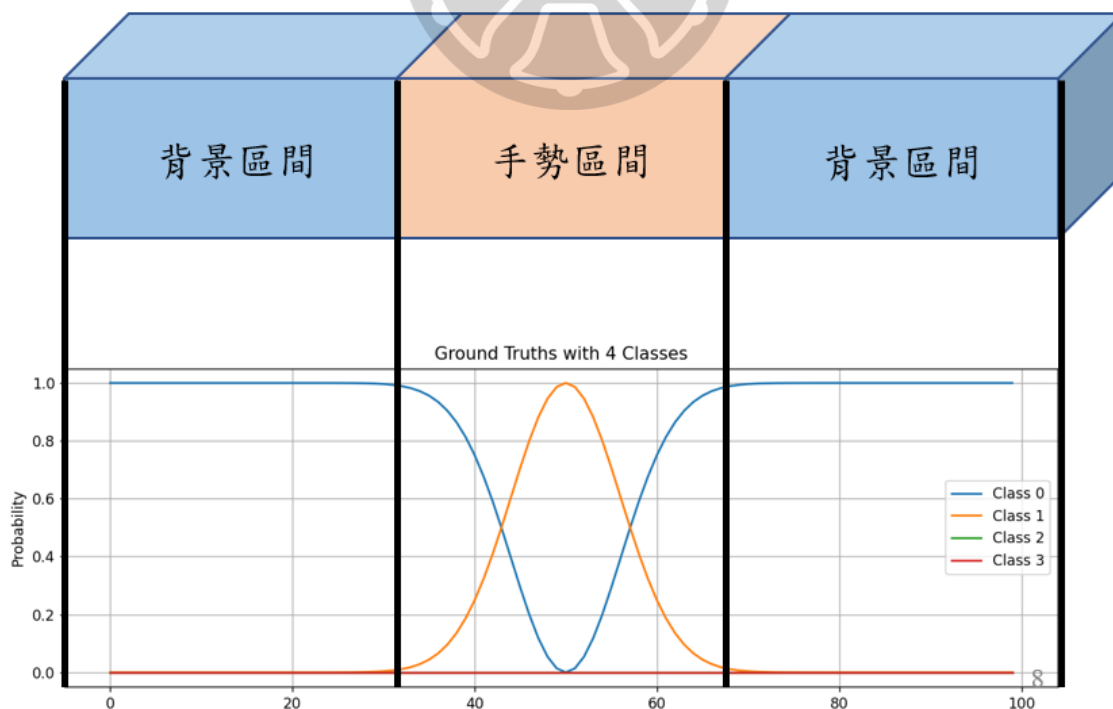


圖 4-5 類高斯分布於手勢區段中的 soft label 配置範例：本圖展示一段手勢區間的 soft label 分布，以手勢中心為對稱軸，向兩側以類高斯形式遞減，反映出動作強度的時間變化特性

圖 4-6 顯示了將原始的標註向量對應至多類別手勢的轉換結果。此階段主要針對每段由使用者標記為有效動作的區段，將其對應至對應的手勢類別索引，並將該區段在該類別通道設為 1，其餘設為 0，以構成多類別 one-hot 編碼格式。在此表示方式下，標註資料能清楚標示出各時間點所屬的手勢類別，同時保留背景類別 (index 0) 作為非動作區段的標記基準。這為後續 soft label 轉換提供了明確依據，也確保手勢標註的語義完整性。

圖 4-7 則說明了如何進一步將 one-hot 格式轉換為 soft label 分布。針對每段標註為手勢的連續區段，系統以其長度估算類高斯曲線的中心與展寬，並為該類別生成一段類高斯權重曲線。轉換後的分布呈現中間幀權重最高、兩側逐漸遞減的連續變化。為保持整體標註向量的邏輯一致性，背景類別的標註值設為 1 減去所有手勢類別權重總和，使得每一幀的標註總和恆為 1，達成與多類別分類任務相符的監督訊號設計。這樣的轉換方式有助於提升模型對於邊界區段的學習能力，並降低硬性標註所帶來的過度擬合風險。

Frame	手勢	Frame	背景	手勢 1	手勢 2	手勢 3
0	0	0	1	0	0	0
1	0	1	1	0	0	0
2	1	2	0	1	0	0
3	1	3	0	1	0	0
4	1	4	0	1	0	0
5	1	5	0	1	0	0
6	1	6	0	1	0	0
7	0	7	1	0	0	0
8	0	8	1	0	0	0
9	0	9	1	0	0	0

圖 4-6 多類別 One-Hot 手勢標註對應結果：針對使用者標註的手勢區段，將其轉換為對應手勢類別的 one-hot 編碼形式，明確劃分各時間點所屬類別，作為 soft label 建構的前置步驟



圖 4-7 類高斯 Soft Label 權重分布範例：將 one-hot 編碼轉換為具有連續性的 soft label，圖中展示類高斯分布於手勢區段的權重變化，並搭配背景補償機制確保各幀標註總和為 1

4-3-3 滑動視窗切片機制

為了將連續的毫米波雷達影像資料轉換為模型可接受的固定長度輸入，本論文採用滑動視窗[10] (Sliding Window) 方式進行資料切片。此機制能將一段長度不一的雷達序列，轉換為多筆具時間連續性的等長樣本，利於模型學習時間序列中的動態特徵。

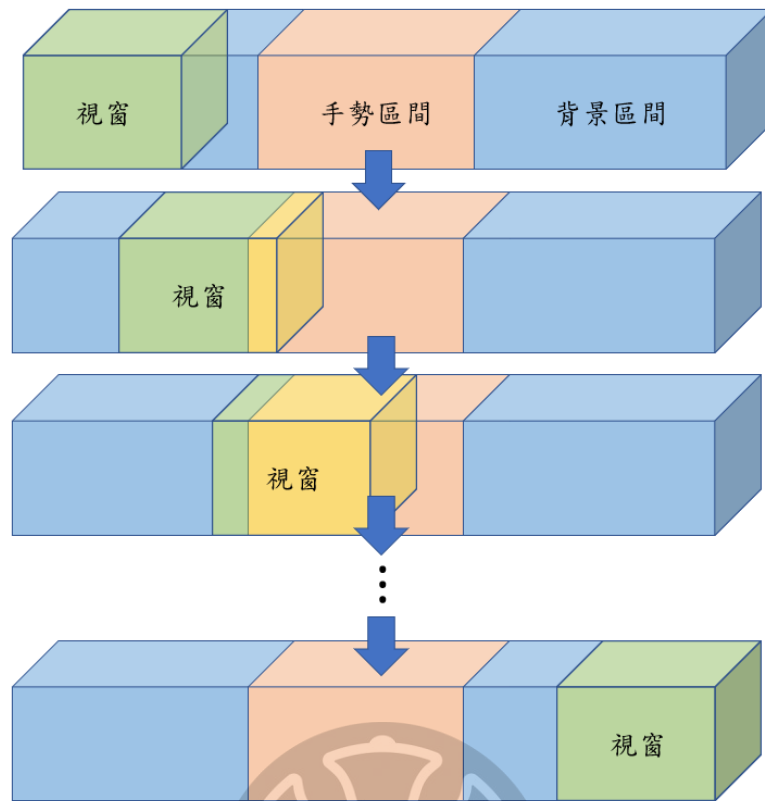


圖 4-8 滑動視窗切片示意圖：將整段手勢資料以固定大小的視窗切片，視窗以設定步長逐格滑動，產生連續且重疊的樣本資料，提升模型的時間序列學習能力與樣本多樣性

實作上，視窗大小設定為 30 幀，對應實際時間約為一秒，此設定是根據本系統設計之手勢動作時間皆約落在一秒內所調整。如此可確保每筆視窗資料完整涵蓋一個手勢的主要區段。視窗的滑動步長 (stride) 設為 1，代表每次滑動一幀，能最大程度擷取資料細節，增加訓練樣本的數量與時間解析度，有助於強化模型對手勢邊界變化的感知能力。

圖 4-8 展示滑動視窗切片過程。當雷達資料中存在手勢區段與背景區段時，視窗將以設定的長度與步長沿時間軸滑動。每個視窗內含的連續 30 幀資料即構成一筆訓練樣本，並於下一節對應其 Ground Truth 標籤。由圖中可見，視窗在滑動過程中可能橫跨不同類型的區段，例如由背景進入手勢、或手勢結束過渡至背景，這樣的切片方式能使模型學習不同狀態轉換的特徵，有助於提升模型於真實情境下的辨識準確性與穩定性。

4-3-4 Ground Truth 對應機制

在本論文的資料切片流程中，每筆訓練樣本皆為滑動視窗分割後取得之固定長度連續影像序列。由於模型訓練需對應單一標籤向量，本論文借鑑 CenterNet 的「中心點標註」概念[11]，採用中心幀標籤法，將每個視窗中間幀所對應的標註機率分布，作為整個視窗的 Ground Truth。

這種做法的出發點，在於視窗中間位置通常最能反映整段序列的主要動作特徵。當視窗完全落於純背景區段時，中間幀的標註會傾向為背景類別(如圖 4-9 所示，機率為 $[1, 0, 0, 0]$)；而當視窗涵蓋手勢主體區段時，其中心幀常落在手勢動作最明顯的時刻，對應的 soft label 會呈現手勢類別機率較高的情形(如圖 4-10 所示，機率為 $[0.3, 0.7, 0, 0]$)，提升模型學習動作關鍵區域的能力。當視窗逐漸滑出手勢範圍並回到背景段落，其中心幀標註也會重新轉為背景主導(如圖 4-11 所示，機率為 $[1, 0, 0, 0]$)，代表手勢已結束並過渡回穩定狀態。這樣的設計能有效協助模型辨別手勢的起始與終止，有助於提升推論階段的穩定性與準確性。

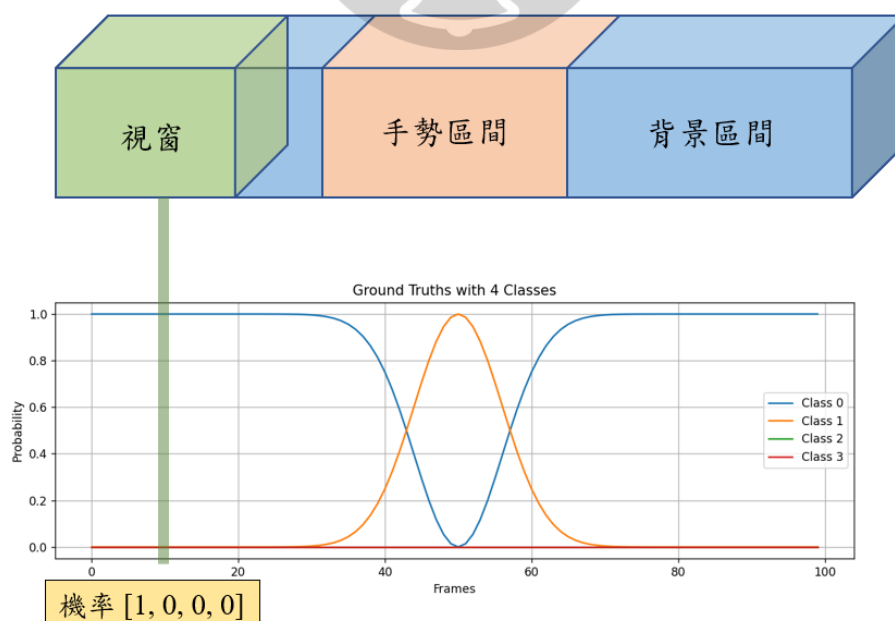


圖 4-9 視窗落於背景區段之 Ground Truth 對應機制示意圖：此圖展示當滑動視窗完全落在背景區間時，其中心幀(圖中綠線所指位置)對應的 soft label 為 $[1.0, 0.0, 0.0, 0.0]$ ，即背景類別的機率為 1，其他手勢類別為 0。由於此區段未涵蓋任何定義手勢動作，因此系統正確地將整段視窗標註為背景，確保訓練時模型能學習背景的特徵表現

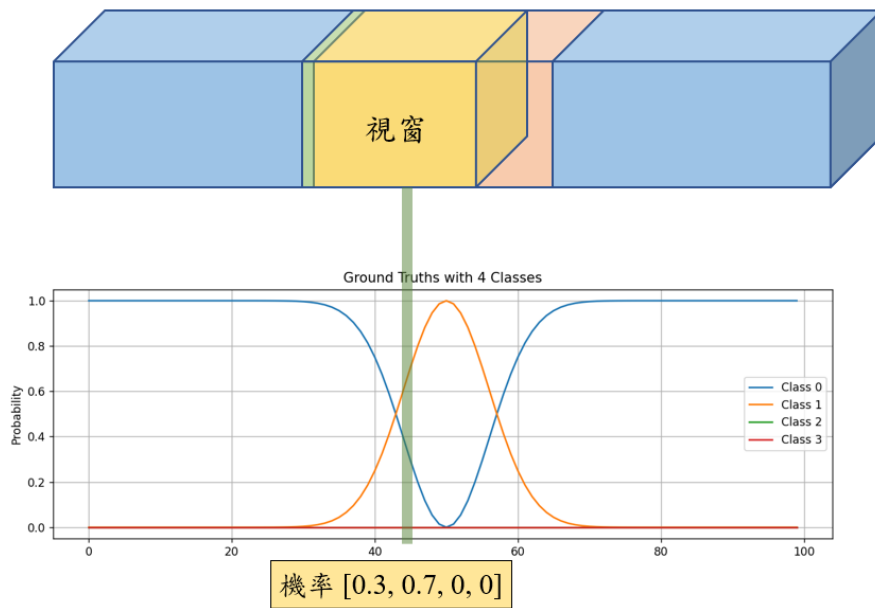


圖 4-10 視窗涵蓋手勢主體區段之 Ground Truth 對應機制示意圖：此圖說明滑動視窗部分涵蓋手勢區段，且其中心幀（綠線）位於手勢動作較明顯的區域。該幀對應的 soft label 為 $[0.3, 0.7, 0.0, 0.0]$ ，代表手勢 1 的機率達 0.7，背景為 0.3，顯示由手勢類別主導。這種設計使模型能更精準地學習手勢主體特徵與邊界過渡，有助於辨識性能的提升

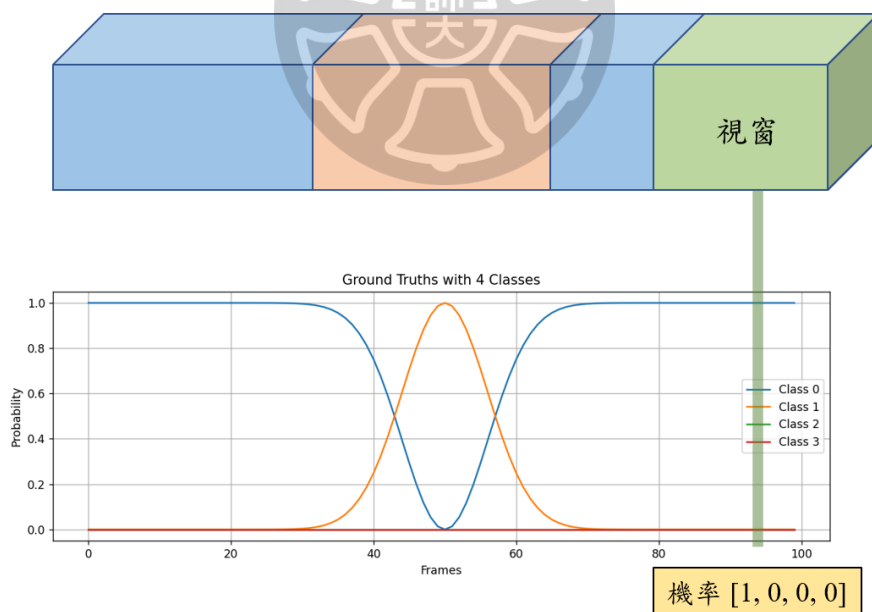


圖 4-11 視窗位於手勢結束後區段之 Ground Truth 對應機制示意圖：此圖顯示滑動視窗已超出手勢主體範圍，覆蓋背景段落為主。綠線所示中心幀對應的 soft label 為 $[1.0, 0.0, 0.0, 0.0]$ ，顯示系統將該視窗視為純背景。此機制有助於模型學習手勢終止後的背景特徵，提升手勢起訖辨識的明確性

4-4 模型架構與訓練流程

本論文所提出之手勢辨識系統，採用三維卷積神經網路[12] (3D Convolutional Neural Network, 3D CNN) 作為主要模型架構，藉由同時擷取時間與空間特徵，有效強化對手勢變化的理解能力。相較於傳統二維 CNN [13] 僅能處理單幀影像，3D CNN 可處理時間序列資料，使模型能觀察動作在時間軸上的演進，進而提升對手勢的辨識準確性與穩定性。

本系統輸入資料為連續 30 幀雷達影像，涵蓋約一秒的手勢動作時間，輸入張量維度為 $(2, 30, 32, 32)$ ，其中 2 表示 RDI 與 PHD 兩個通道，30 為時間軸長度， $32 * 32$ 為單幀影像的解析度 (如圖 4-12 所示)。RDI 描述目標在距離與速度上的變化，PHD 則補充方位角度資訊，兩者結合可提供完整的空間特徵；而 30 幀的時間維度則紀錄動作發展的過程，讓模型能觀察手勢如何隨時間展開與結束。



圖 4-12 本系統所採用之雙通道輸入影像結構：輸入資料包含 Range-Doppler Image (RDI) 與 Phase Difference (PHD) 兩組影像，各通道長度為 30 幀，並具備 $32 * 32$ 空間解析度，形成三維時間影像結構

如圖 4-13 所示，模型由三層 Conv3D 模組組成，每層皆配有 ReLU 激活函數、Batch Normalization (BN) 與 MaxPooling 操作，逐層抽取時空特徵並進行下採樣。其後經由 Global Average Pooling 將特徵壓縮，再通過兩層全連接層與 Dropout 正則化處理，最終由 Softmax 輸出四分類的手勢預測結果。

在訓練過程中，為強化模型對邊界過渡區的感知能力，我們使用 Soft Label 作為 Ground Truth，搭配 MSE (Mean Squared Error) 作為損失函數 (Loss Function)。此舉相較於傳統 one-hot 編碼配合交叉熵 (Cross-Entropy) 的方式，能更平滑地引導模型學習手勢轉換的連續變化，特別有助於改善手勢起始與結束區段的辨識穩定性，並提升模型在非理想輸入條件下的泛化能力。

由於手勢是一種具備「時間演變」的動作，若僅以單一幀為輸入，將無法掌握手勢的起始與結束時機，容易導致誤判。透過訓練時間序列的輸入資料，模型得以學習動作的連貫變化，例如手從遠處移近、掌心旋轉等時間性資訊，進而提高對不同手勢的辨識準確度與泛化能力。這也使得 3D CNN 成為處理雷達手勢影像的理想選擇。

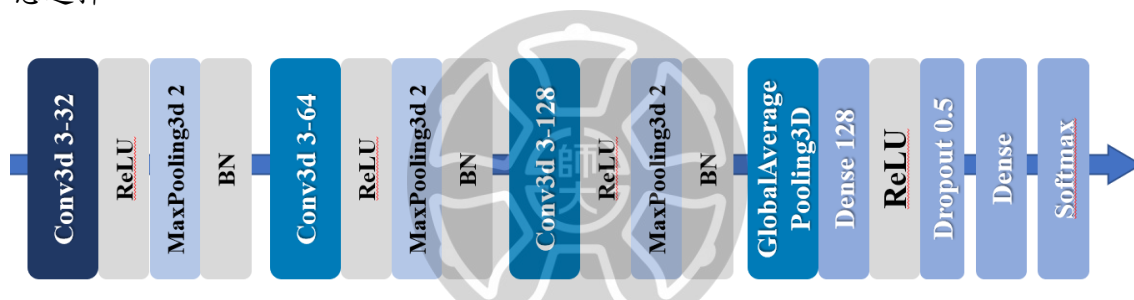


圖 4-13 三維卷積神經網路 (3D CNN) 模型架構：模型接收雙通道影像序列作為輸入，經多層時空卷積、特徵壓縮與全連接層，最終輸出手勢四分類的預測機率向量

4-5 預測結果與後處理機制

為驗證所提出之手勢辨識系統之有效性，本節將展示實驗結果與分析方式，包含模型輸出與 Ground Truth 之對比圖、雙閾值機制[14]設計流程、時序變化示意，以及最終分類表現的混淆矩陣。

圖 4-14 展示模型在測試資料中某一手勢片段的預測分數變化。上方為 Ground Truth，顯示類高斯型態的手勢 1 分佈；下方為實際預測結果。可觀察到預測分數與標註在主體結構上相當吻合，並可有效捕捉手勢開始與結束的漸變區段。

圖 4-15 則說明系統採用的雙閾值分類機制。考量連續資料中手勢動作之模糊

邊界與分類穩定性問題，本系統以狀態機設計方式引入兩組判斷門檻：進入手勢狀態的門檻稱為進入閾值 α ，離開手勢狀態則使用較低的離開閾值 β 。當預測中任一手勢類別之機率大於進入閾值 α ，系統即進入手勢狀態，反之，若所有手勢類別機率皆低於離開閾值 β ，則回到背景狀態。此設計可有效降低邊界抖動與誤判現象。

圖 4-16 將雙閾值策略應用於連續輸出範例，顯示進入與退出手勢狀態的時機。橫軸為時間軸，縱軸為每類別的預測機率。可見透過雙閾值策略，系統可在動作發生區段內穩定維持手勢分類狀態，而在動作前後準確返回背景狀態，提升分類決策的準確性與穩定性。

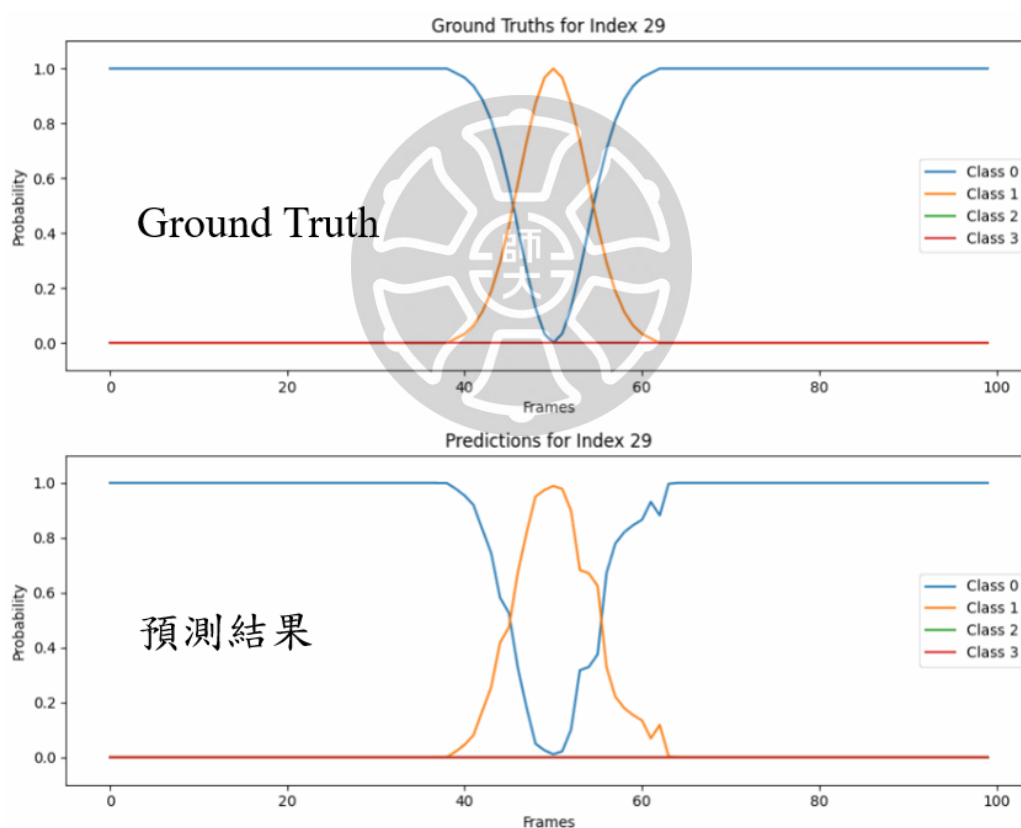


圖 4-14 Ground Truth 與預測結果之時間序列對比圖：比較模型對於測試片段的預測分數與標註真值，可見其在手勢區段具有良好對應性

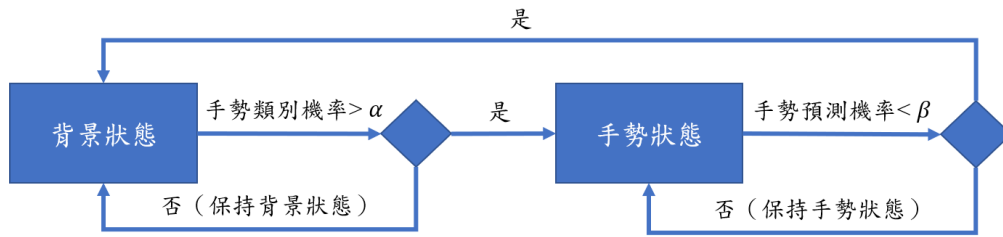


圖 4-15 雙閾值狀態轉移流程圖：透過設計進入閾值 α 與離開閾值 β 來穩定手勢與背景的判斷流程

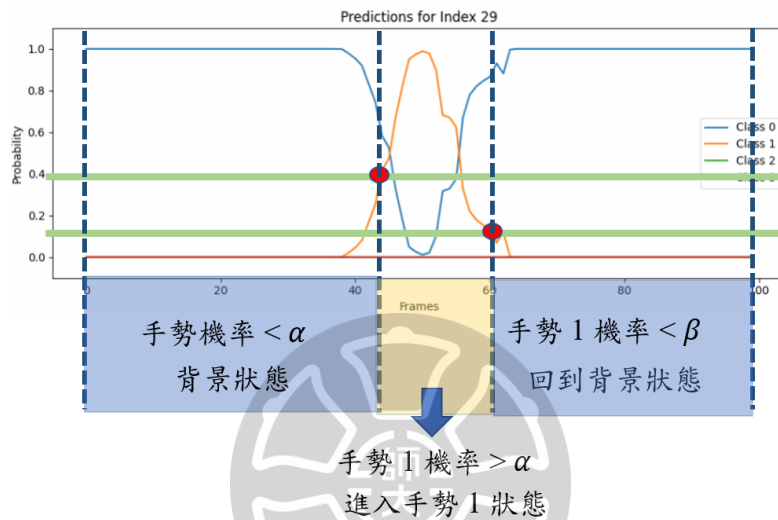


圖 4-16 雙閾值機制應用於預測輸出之示意圖：藉由雙閾值設計，準確判斷手勢進入位置及退出位置

在進行混淆矩陣分析前，需先說明本論文中對單筆手勢資料的分類判斷方式。系統採用雙閾值設計進行辨識，進入閾值 α 設為 0.4，退出閾值 β 設為 0.1，當任一非背景類別的預測機率高於進入閾值 α ，即視為進入手勢狀態；而當所有非背景類別的預測機率皆低於退出閾值 β 時，則回到背景狀態。根據此機制，每筆資料僅對單一動作進行辨識，若資料中完全未觸發手勢狀態，即歸類為背景；若出現一次完整的「進入→退出」過程，則視為有效手勢，整比手勢機率皆低於 α 則視為背景，並以該期間內機率最高者作為預測類別；若系統於單筆資料中偵測到兩次以上不同手勢進入事件，則視為 Multi-Gesture；若出現手勢進入但未完成退出，則歸為 Incomplete。

圖 4-17 為本系統在完整測試資料上的混淆矩陣，分類結果依上述規則統計。由圖可見，系統整體分類表現良好，準確度為 95.8%，展現出良好的辨識能力。

本系統設計與評估以右手單手操作為主，辨識的三類手勢皆屬動態幅度明顯者，包含 PatPat、Wave 與 Come。雖整體成效穩定，然部分 Wave 與 Come 樣本在速度或距離變化上不夠明顯，導致特徵與背景相近，較易誤判為背景，可能為準確度略低的原因之一。此外，所設計手勢皆為重複兩次的動作循環。即使使用者實際執行三次手勢，系統採用滑動視窗切片方式，單一視窗僅涵蓋其中兩次連續片段，故不會因次數增加而誤判為背景，僅會增加手勢持續時間。然而，若多次手勢之間具明確區隔，亦可能被系統判定為兩個獨立手勢事件。

本論文未與 Google 所提出之 Soli 雷達系統進行直接性能比較，主因在於其所採用之硬體架構和資料型態與本研究存在差異。Soli 雷達系統為 2T4R 的天線配置，即具備兩個發射天線與四個接收天線，整體感測能力在反射訊號的收集與空間解析度上均優於本研究所使用的 1T2R 配置。因此雙方系統在資料品質與解析能力上存在本質差異，無法進行直接比較。

Soli 提供的資料集[15]為單一手勢執行完成後所裁切之片段，每筆資料僅對應一種手勢類別，並以整段資料進行標註，無法涵蓋連續手勢輸入時可能出現的背景干擾、過渡動作等狀況，亦不支援本研究所採用之平滑式 soft label 標註方式。

Soli 系統並未提供原始雷達資料擷取與感測參數設定等相關工具鏈，因此無法於相同場域與感測條件下，重新蒐集與本研究一致之手勢資料。由於缺乏可建立對照組之平台一致性與資料重現能力，實作層面難以達成公平且有效之性能比較。

綜合而言，Soli 系統在雷達解析度、資料結構與標註方式等面向皆與本研究設計不同，難以進行一致的性能對照與實作比較。

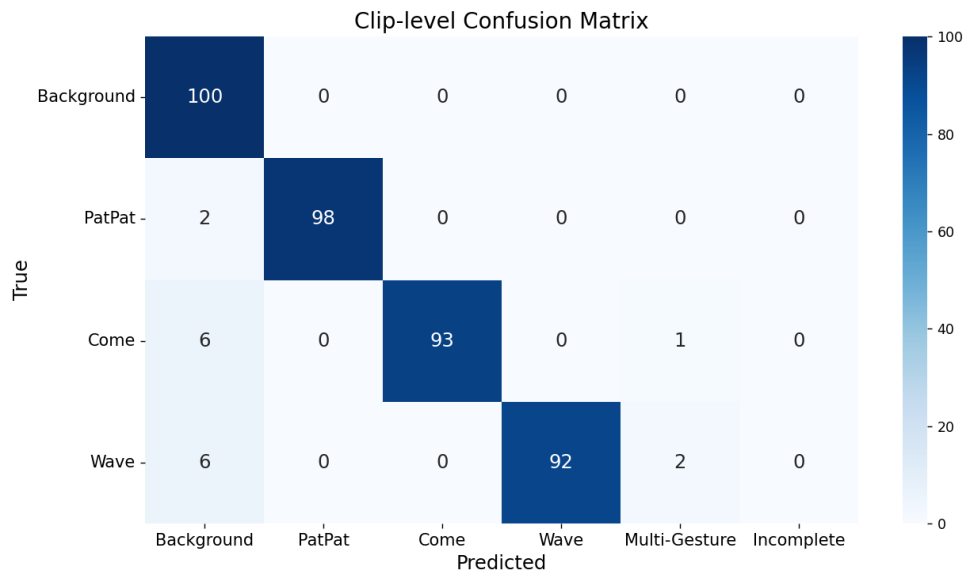


圖 4-17 單筆手勢分類之混淆矩陣 (Clip-level Confusion Matrix)：本圖為系統在完整測試資料上，以單筆手勢為單位所計算之分類混淆矩陣。分類結果依據雙閾值判斷機制進行定義：當任一手勢類別之預測機率超過進入閾值 α 時，視為進入該手勢狀態；當預測機率低於退出閾值 β 時，視為離開手勢狀態。若過程中僅進入一次並成功退出，視為有效手勢，整比手勢機率皆低於進入閾值 α 則視為背景；若包含兩次以上不同進入點，歸為多重手勢 (Multi-Gesture)；若進入後未再次退出，視為不完整手勢 (Incomplete)；若全程未達進入門檻，則視為背景。由圖可見，系統整體分類準確率良好

4-6 即時辨識系統設計

本論文所設計之即時手勢辨識系統流程如圖 4-18 所示，整體架構可分為五個階段：雷達資料擷取、暫存區儲存、模型預測、後處理，以及結果可視化。其中雷達資料係由開酷科技所提供之 K60168A Dongle 雷達裝置以 Ksoc Tool API 即時傳入，格式為 $(2, 32, 32)$ ，分別對應兩通道接收天線所產生之 Range-Doppler Image (RDI) 與 Phase-Difference Map (PHD)。

雷達資料進入後，系統需先建立一個時間視窗大小為 30 的暫存區 (如圖 4-19 所示)，以維持連續性的時間序列輸入。此暫存區實作上為一先進先出 (FIFO) 佇列，當新的影格加入後，若超過 30 幅，則會移除最早的一幀資料，確保模型每次輸入皆為 30 幀之時間連續資料。

圖 4-19 展示了此機制的四個步驟：首先建立一個空的資料區，當新的影格進入時會依序填入；當緩衝區達到最大長度後，後續新資料將推進整個緩衝區，並移除最舊的一筆資料，以此維持固定視窗長度。實作程式如圖 4-19 所示，透過 Python 程式不斷輪詢 API 並更新緩衝內容。



圖 4-18 即時手勢辨識系統流程架構圖：本系統流程包含雷達資料擷取、暫存區儲存、模型預測、後處理與結果可視化等五個階段，資料自 K60168A 雷達即時傳入後，經滑動視窗機制整理成連續輸入，送入訓練完成之模型進行即時推論與顯示

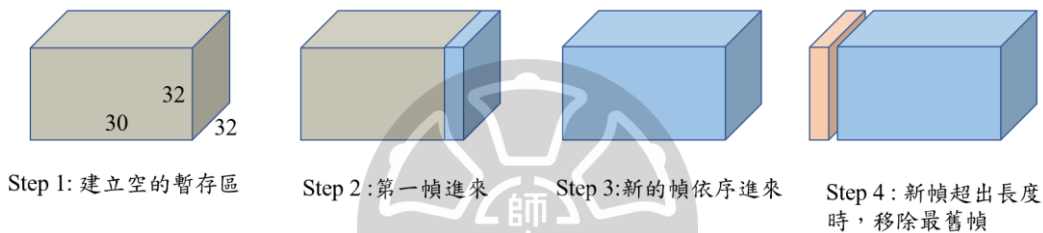


圖 4-19 暫存區更新機制圖：視窗長度為 30 幀，當超出長度限制時，移除最舊資料，維持固定長度之時間序列輸入

```
buffer = [] # Step 1: 建立空的暫存區

window_size = 30 # 設定模型需要的時間視窗長度 (這裡是 30 幀)

def append_to_buffer():
    new_frame = ksoc_api.get_data() # 使用 Ksoc Tool 的 api 即時取得新幀資料，形狀為 (2, 32, 32)
    buffer.append(new_frame) # Step 2-3: 加入新的幀進入暫存區
    if len(buffer) > window_size:
        buffer.pop(0) # Step 4: 超出長度時，移除最舊幀

while True:
    append_to_buffer() # 不斷接收新資料並更新暫存區
```

圖 4-20 暫存區更新程式碼示意圖：視窗長度為 30 幀，以 Python 撰寫之暫存區管理程式，使用 Ksoc API 擷取雷達資料，將資料以先進先出方式儲存至 buffer，超過視窗長度時自動移除最舊資料，以維持連續輸入供模型進行推論

暫存區填滿後，資料即被送入訓練完成之 3D CNN 模型，產生四類別的機率輸出結果。接著系統將此預測結果套入後處理流程，包括雙閾值判斷(詳見 4-5 節)

以區分有效手勢區段與背景段落，最後進行可視化顯示，即時回饋使用者目前所辨識之手勢結果，如圖 4-21 與圖 4-22 所示，分別呈現背景狀態與手勢觸發時的即時畫面。

透過上述機制，本系統可達成雷達影像的即時收集、推論與辨識，有效應用於手勢控制、穿戴裝置操作等人因互動場景。

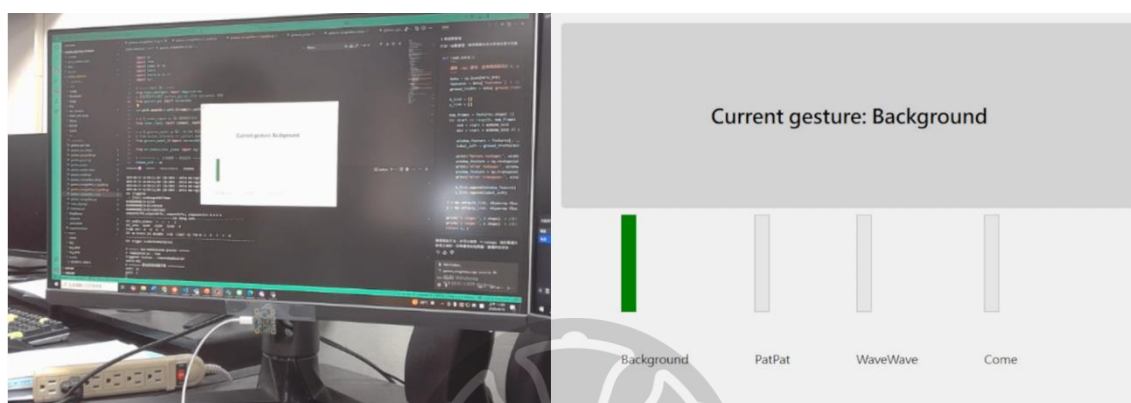


圖 4-21 系統於背景狀態下之即時辨識畫面：左圖為本系統實際執行畫面，可見雷達模組連接至主機，並透過圖形介面即時顯示辨識結果；右圖為介面中所呈現之手勢機率條狀圖，顯示目前手勢為 Background，各手勢之預測機率皆未超過進入閾值，系統正確維持背景狀態

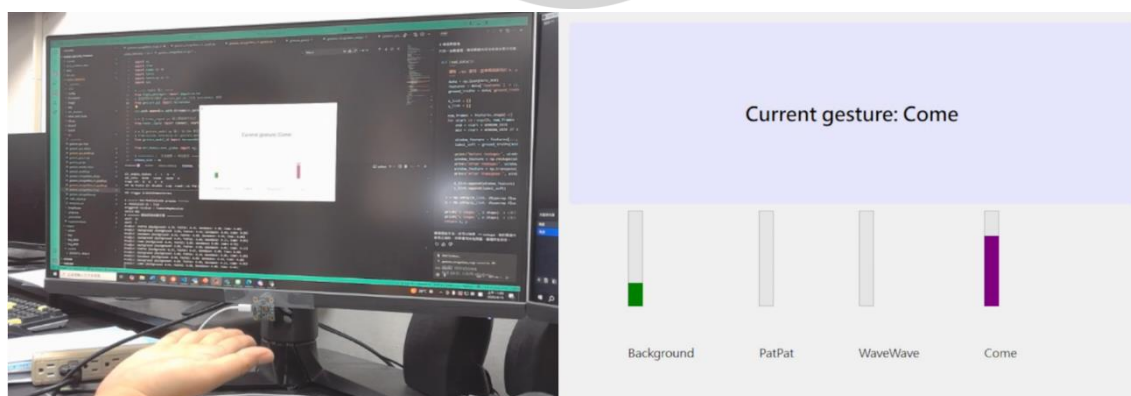


圖 4-22 系統於 Come 手勢觸發時之即時辨識畫面：左圖顯示使用者於實際操作中執行 Come 手勢，雷達即時接收手部動作資料；右圖之手勢機率條圖顯示「Come」類別機率高於進入閾值，其餘手勢機率則低於門檻，系統正確辨識並即時回饋「Current gesture: Come」結果

第五章 結論

本論文以毫米波雷達手勢辨識為核心，探討如何結合現有雷達平台與深度學習技術，建構具備即時辨識能力之系統。研究中採用開酷科技所開發之 K60168A 毫米波雷達晶片與 Ksoc Tool 平台，作為雷達資料擷取與標註的操作環境，並透過其圖形化介面完成手勢區段之標註，建立對應的雷達影像資料集。

在資料處理部分，本文針對雷達所產生之 RDI(Range-Doppler Image)與 PHD(Phase Difference) 影像進行分析，說明其物理意涵與應用價值，建立對雷達影像結構之理解。此外，透過滑動視窗切片機制與類高斯 soft label 設計，使模型能更精確捕捉手勢的時序變化與邊界過渡，進而提升學習效率與辨識精度。

模型設計上，採用三維卷積神經網路(3D CNN)架構，結合時間序列特性，強化對動態變化之掌握；並搭配均方誤差(MSE)作為損失函數，達到更平滑且具連續性的分類效果。於推論階段，導入雙門檻判斷機制，有效排除不穩定預測與背景干擾，強化對手勢起始與終止時機之判斷能力。

根據測試資料之評估結果，本系統在四類別(背景、PatPat、Come、Wave)之分類任務中達成 95.8% 的整體準確度，顯示其具備高度辨識效能與實用潛力。系統亦可即時接收雷達資料、進行推論與可視化顯示，成功實現毫米波雷達手勢辨識之完整應用流程。

總結而言，本文結合現有雷達平台，整合並應用於手勢辨識，具備即時性、彈性與高準確率，未來可作為人因互動、穿戴裝置操控等應用場景之基礎架構，亦提供後續研究在雷達影像學習與系統應用上的重要參考依據。

參考文獻

- [1] A. Dix, J. Finlay, G. D. Abowd, and R. Beale, “Human-computer interaction (3rd ed.),” Pearson Educ., 2004.
- [2] P. Molchanov, S. Gupta, K. Kim, and J. Kautz, “Hand gesture recognition with 3D convolutional neural networks,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), pp. 1–7, 2015.
- [3] J. Brownlee, “Deep learning for computer vision: Image classification, object detection and face recognition in Python,” Mach. Learn. Mastery, 2019.
- [4] F. Al Farid, N. L. Hashim, J. Abdullah, M. R. Bhuiyan, W. N. S. M. Isa, J. Uddin, M. A. Haque, and M. N. Husen, “A structured and methodological review on vision-based hand gesture recognition systems,” J. Imaging, vol. 8, no. 6, p. 153, 2022.
- [5] KaikuTek Technology, “K60168A Radar User Manual & Ksoc Tool Technical Guide,” [Online]. Available: <https://www.kaikutek.com/zh-tw/products/detail/42>, 2024.
- [6] S. Wang, J. Song, J. Lien, I. Poupyrev, and O. Hilliges, “Interacting with Soli: Exploring fine-grained dynamic gesture recognition in the radio-frequency spectrum,” Proc. ACM Symp. User Interface Softw. Technol. (UIST), pp. 851–860, 2016.
- [7] A. Santra, S. Hazra, L. Servadei, T. Stadelmayer, M. Stephan, and A. Dubey, “Deep metric learning,” in Methods and Techniques in Deep Learning: Advancements in mmWave Radar Solutions, Wiley-IEEE Press, ch. 2.6, pp. 73–86, 2023.
- [8] Q. Fang, Y. Yan, and G. Ma, “Gesture recognition in millimeter-wave radar based on spatio-temporal feature sequences,” arXiv preprint arXiv:2309.09528, Sep. 2023.
- [9] Q. Chen, Y. Li, Z. Cui, and Z. Cao, “A hand gesture recognition method for mmWave radar based on angle-range joint temporal feature,” Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS), pp. 2650–2653, 2022.
- [10] J. Ortiz-Laguna, A. G. Olaya, and D. Borrajo, “A dynamic sliding window approach for activity recognition,” Lect. Notes Comput. Sci., vol. 6787, Springer, pp. 219–230, 2011.
- [11] X. Zhou, D. Wang, and P. Krähenbühl, “Objects as points,” arXiv preprint arXiv:1904.07850, Apr. 2019.
- [12] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, “Learning spatiotemporal features with 3D convolutional networks,” Proc. IEEE Int. Conf. Comput. Vis. (ICCV), pp. 4489–4497, 2015.
- [13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Proc. Adv. Neural Inf. Process. Syst. (NIPS), pp.

1097–1105, 2012.

- [14] J. He, W. Yu, X. Liu, Y. Yang, and J. Ma, “Real-time gesture recognition using a dual-threshold method on Doppler images,” *IEEE Sens. J.*, vol. 17, no. 10, pp. 3103–3110, 2017.
- [15] C. Gupta, “Google Project Soli: RADAR Gesture Recognition,” Kaggle, [Online]. Available: <https://www.kaggle.com/datasets/chandragupta0001/soli-data>, 2016.

