

國立台北大學資訊工程學系專題報告

A Device-Free Interactive Quiz Game Using Deep Learning And Channel State Information Based Gesture Recognition

具深度學習結合無線網路訊號實現動作辨識之無穿戴式裝置答題遊戲應用專題
專題編號:PRJ-NTPUCSIE-109-003

執行期間:2020 年 9 月 至 2021 年 6 月

2. 簡介

本專題計畫將利用類神經網路 Convolutional Neural Network 結合 WiFi Channel State Information 以此建構出一款不需穿戴任何裝置的互動式答題遊戲，再架好天線的環境當中，將不需要架設攝像機，利用玩家的手勢對 WiFi CSI 造成的干擾對其進行 CNN 類神經網路辨識，且玩家將不需攜帶任何搖桿或穿戴任何裝置就能進行遊戲，希望能在遊戲的同時能夠兼顧到隱私權的問題。

由於現今市面上所有的互動式遊戲都需要玩家手持搖桿裝置或穿戴式裝置，又或是需要架設一個攝像頭來對玩家的動作進行辨識，除了隱私權的問題外，我們想要找尋是否有方法能夠讓使用者不需穿戴或手持裝置就能進行遊戲。故我們利用架設 WiFi 網卡的方式對得到的頻譜圖進行辨識。

3. 專題進行方式

本專題計畫可分為三個部份來做概述，第一個是使用者端，第二個是 AI model 辨識端，最後則是遊戲應用端：

(一) 玩家端

1. WiFi 天線架設：在網卡上加裝天線以讓接收到的資料更準確，並將天

1. 摘要

隨著科技的日新月異，人們的生活越來越便利，隱私權的問題也就油然而生。像是安養中心在年長者的房間內安裝監視器來確保沒有突發狀況發生，但同時也可能侵犯到了他人的隱私權；以及臉部辨識的普及，其衍伸出的隱私權問題更是層出不窮。

以及近年來無線網路訊號的普及，低成本及不侵犯隱私性的優勢，利用無線網路訊號實現的動作辨識在這幾年蓬勃發展，通道狀態訊息 CSI(Channel State Information) 在無線通訊中是指一個通訊鏈路的已知通訊特徵，記錄了綜合所有因素對訊號的影響，例如：散射、衰弱等等。利用環境對 WiFi 訊號的影響來產生一張類似頻譜圖的訊息，以此來記錄當前環境的狀態。

互動式遊戲這些年來逐漸盛行，像是很受歡迎的 VR(Virtual Reality) 類遊戲，能使玩家身歷其境享受不一樣的遊戲體驗、或是健身環的崛起，讓不愛出門的人在家中就能夠擁有足夠的運動量。不同於以往的桌機遊戲或手機遊戲，互動式遊戲能將玩家的遊戲體驗提升到另一個層次。

線架設在固定範圍內，讓玩家能在架設好天線的環境內進行遊戲。

2. 玩家手勢種類：玩家將能以 4 種不同的手勢進行答題遊戲，分別是上揮、下揮、左揮、右揮。

(二) AI model 辨識端

1. AI 模型訓練：利用我們人工蒐集到的 4000 筆資料集對 CNN 模型進行訓練。

2. 玩家手勢辨識：將玩家當前做出的動作頻譜圖傳送到我們事先訓練好的 CNN 模型中進行辨識，接著將結果傳送到我們的遊戲應用端。

(三) 遊戲應用端

1. 題目資料庫建立：利用現有的題庫建立成資料庫。

2. 答案選擇：玩家在想好答案之後能利用上揮、下揮、左揮、右揮 4 種手勢來選擇 題目所給予的四個答案選項。

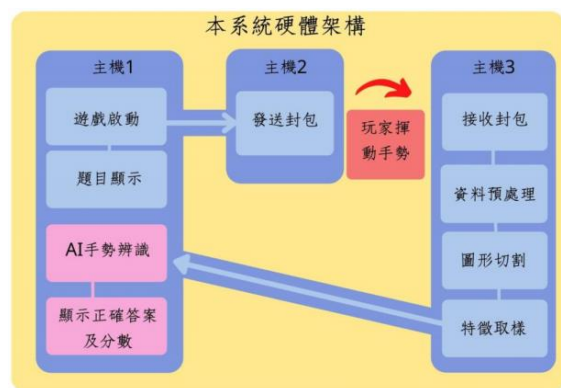


Fig.1. 系統架構

4. 可能遇到之問題與解決方法

(一) 系統三大階段問題探討

1. CSI WiFi 封包資料預處理：

(1) 發送封包時，如何解決封包遺失與因噪聲產生的極值問題？

當我們在發送封包時，接收端的天線沒辦法全數完整的接收到，大概會有 3%到 5%的封包遺失，而遺失的

封包將會影響到我們資料的正確性。因此，我們採用線性插值的方法，利用一條子載波前後的值取平均，並且比對上下兩條子載波在同一時間所收到的值，以此才能解決實際波型在一瞬間的幅值浮動，因為當人體干擾波型時，幅值的浮動往往在一瞬間就有劇烈變化。對於因噪聲產生的極值問題，在 WiFi CSI 人體感知中，低頻訊號相當重要，常常蘊含著人體對訊號的干擾，高頻部分則能顯示出訊號裡的細節差別，因此，本計畫採用離散小波變換的方法進行分析與濾波，透過兩個相互濾波器產生兩個訊號，盡可能的保留下動作所產生的資訊，最後得出一個時頻圖。

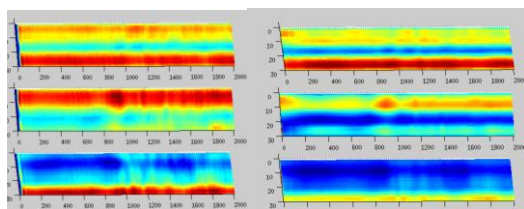


Fig.2. 不同濾波器下的波型圖

(2) 如何將接收到的封包轉成可視化的圖片？

收到封包資料後，我們只靠封包中的數據無法輕易看出資料的好壞，也不易對其進行分析，因此我們將收到的封包利用 Matlab 讀取，接著利用 Matlab 中的 API 函式 mesh，將我們收到的封包轉換成彩色的 3D 圖形，X 軸代表時間、Y 軸代表子載波、而顏色則代表訊號的幅值大小。如此一來，我們能夠輕易地看出濾波的效果是否有達到我們的預期。

(3) 該如何對資料進行裁切取樣而不破壞資料完整性，以減少原先過大的資料量？

如果以原始的 raw data 直接進

到 AI model 的話，訓練時間將會要花上好幾天，因此，我們將原始的 90x2000 的矩陣裁切成三等份，然後往下堆疊，形成一個 270x660 的矩陣，多於的一小部分則直接捨去。接著，我們將子載波的長度取樣成原先的一半，也就是將總共有 2 秒的資料取樣成 1 秒，利用每隔 1 毫秒就刪除一筆資料的方式，以此來取樣奇數秒的資料，能將原本的資料大小縮減成原先的一半，例如，一個 90x2000 的矩陣能減少成 90x1000。至此我們能得到一個 270x330 的矩陣，但因為我們選用的是 CNN 模型，較適合正方形的圖片進行訓練，所以我們將對最後的矩陣進行插值以及取樣，使其大小變成我們要的 256x256，能使我們的 AI 模型提升準確率以及訓練速度。

2. AI 模型訓練：

(1) 並不存在現成的資料集可用，如何蒐集訓練的資料以及確保資料的品質？

CSI WiFi 目前在網路上的可用資源並不多，也沒有現成的資料集能讓我們使用。因此，我們自行蒐集總共 4000 筆的資料，在收資料的過程中，可能會因為人的不同以及環境的不同對資料造成不一樣的干擾。所以在蒐集過程中，我們以每 100 筆為單位，每蒐完一個單位的資料就進行資料檢視，如此來確保每一類的動作波形圖差異不要太大且能將外在因素對訊號波型的影響降到最低。

(2) AI 模型如何提升準確率及速度？

如果以原始的 raw data 直接進到 AI model 的話，訓練時間將會要花上好幾天，因此，我們將子載波的

長度取樣成原先的一半，也就是將總共有 2 秒的資料取樣成 1 秒，利用每隔 1 毫秒就刪除一筆資料的方式，以此來取樣奇數秒的資料，能將原本的資料大小縮減成原先的一半，例如，一個 90x2000 的矩陣能減少成 90x1000。在減少了圖片大小後，訓練時間將能大幅縮減，在取樣過程中可能對準確率有所影響，因此我們利用 CNN 類神經網絡來增加神經網絡的深度，以此提高準確率。

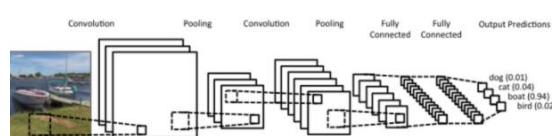


Fig. 3. CNN 類神經網路架構

3. 遊戲應用：

(1) AI 辨識結果與遊戲端如何連結？

遊戲開始後，總共會有 5 道題目，一題 20 分，總分 100 分，每一題會有 20 秒的時間作答，每一道題目將會有四個不同的選項，四個選項分別放置在畫面的上下左右，玩家在閱讀完題目後能夠依照自己想選的答案進行上揮、下揮、左揮或著右揮的動作選擇答案。玩家動作結束後，將 AI 辨識完的結果傳送到遊戲端，分別代表四種不同的控制訊號，也就是四個答案選項，依照玩家作答正確與否進行畫面顯示以及分數計算。

(2) 如何正確判別當前接收到的資料是玩家的答案而不是空資料？

因為我們無法預期玩家揮動手勢的時間，所以沒辦法準確知道要將哪一段的資料進到 AI 進行辨識，這將會嚴重的影響到我們辨識的準確率。因此，我們利用遊戲端來控制玩家進行作答的時間，在每一道題目顯示出

來後，將會有 10 秒的時間讓玩家思考答案，接著會出現【請作答】的指示，玩家則將在 3 秒之內進行作答，否則將以 答案錯誤為結果，【請作答】指示出現的同一時間，我們將立即發送 CSI WiFi 訊號，以此方法來獲取最準確的 CSI WiFi 波型圖。

5. 主要成果與評估

最後在 AI 模型辨識的部分，我們使用訓練及測試的準確率能達到 95% 而使用全新的 RAW DATA 進行模型評估時，準確率大概能到 65%，考量到不同環境的影響會對波形圖造成很大的改變，以及我們人工收集的資料集總數有點不夠，會使模型過擬合，因此我們希望未來能增加資料集的大小以及減少不同環境造成的影響。

6. 結語與展望

未來我們期望在濾波器的部分能將波形濾的噪聲更低，以及 AI MODEL 的部分能加深神經網路的難度以及層數，從而提升辨識的準確率。

7. 銘謝

老師對於專題的研究進行經驗豐富，對於研究計畫的執行速度和其他事項的掌握，需要藉由老師的指導與規範，使得本計畫能夠處於掌握之中，並能如預期完成，方法為每兩週開會與老師報告研究進度，而當我們對研究有疑問時，能在開會中與老師討論以及老師會指導我們之後繼續研究的方向與可行性。

8. 參考文獻

[1] Siamak Yousefi, Hirokazu Narui, Sankalp Dayal, Stefano Ermon, Shahrokh Valaee, " A Survey on Behavior Recognition Using WiFi Channel State Information", IEEE

Communications Magazine ,Vol. 55, Issue 10, pp. 98-104, Oct. 2017.

[2] Han Zou, Yuxun Zhou, Jianfei Yang, Hao Jiang, Lihua Xie, Costas J. Spanos, " DeepSense: Device-Free Human Activity Recognition via Autoencoder Long-Term Recurrent Convolutional Network", IEEE International Conference on Communications (ICC), Kansas City, MO, USA, May. 2018.

[3] Zhenguo Shi, J. Andrew Zhang, Richard Xu, Qingqing Cheng, " Deep Learning Networks for Human Activity Recognition with CSI Correlation Feature Extraction ", IEEE International Conference on Communications (ICC), Shanghai, China, May. 2019.

[4] Xiang Li, Daqing Zhang, Qin Lv, Jie Xiong and Shengjie Li, "IndoTrack: Device-Free Indoor Human Tracking with Commodity WiFi", Proceedings of the ACM on Interactive Mobile Wearable and Ubiquitous Technologies, vol. 1, no. 3, pp. 1-22, 2017

[5] Chao Peng, Jeffrey Hansberger, Vaidyanath Areyur Shanthakumar, Sarah Meacham, Victoria Blakley, Lizhou Cao, "A Case Study of User Experience on Hand Gesture Video Games", IEEE International Conference on Communications (ICC), Galway, USA, Aug. 2018.

