

輔仁大學資訊工程學系110學年度 畢業專題

基於肢體與語音辨識之運動遊戲

Sport Game Based on Body and Speech Recognition

組別：A09 指導教授：王國華

407261099 資工四甲 陳禹辰 407261128 資工四甲 康智絜

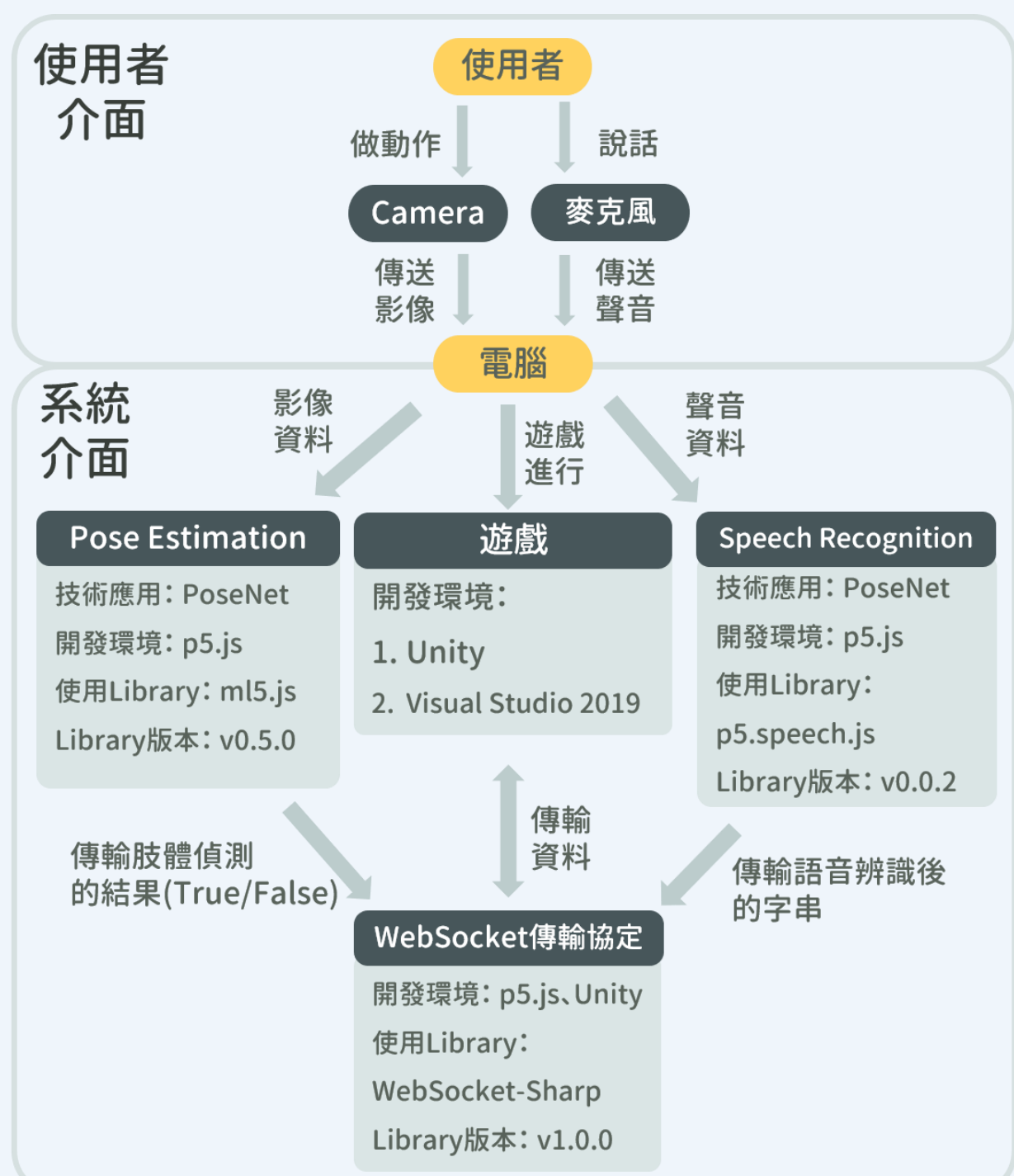
407261415 資工四甲 陳思璇 407261439 資工四甲 張鈞奕

摘要

由於新冠肺炎的影響，多數人在家中的時間增加，較不方便去戶外運動或去健身房。本專題開發了一個運動遊戲，旨在建立一個僅需一台電腦及可連接到電腦的鏡頭，即可偵測肢體動作的運動健身遊戲，希望滿足想在室內運動的人的需求並兼具趣味性。

本專題使用Unity製作遊戲，利用WebSocket傳輸協定將p5.js和Unity作連接，使用p5.js中訓練完成的模型判斷玩家是否有做指定動作的辨識結果，以及語音辨識產生的字串回傳給Unity，讓Unity將對應結果呈現在遊戲畫面中。透過p5.js建置PoseNet來辨識玩家身體的17處關節所在位置，作為模型訓練之依據，利用深度學習來訓練模型辨識動作，並設計一個簡單的二元分類法即時用於運動遊戲中，以辨識玩家動作是否符合遊戲規定。本專題在p5.js建立語音辨識系統，利用語音辨識來控制遊戲開始前的關卡選擇與遊戲進行，讓玩家不需額外器材，即可操作遊戲畫面。

系統架構



開發技術

遊戲環境架設

使用Unity來架設整個遊戲的環境，呈現遊戲中的畫面與功能給玩家。

角色模型

使用Blender來製作胡蘿蔔和甜甜圈怪物的角色建模，在遊戲畫面中呈現角色的畫面，並製作甜甜圈怪物噴水柱的動畫特效。若甜甜圈怪物發動攻擊會噴出水柱，玩家的HP值會減少。

辨識身體關節點

在p5.js中使用PoseNet的pose estimation model 收集玩家的身體關節點。

訓練模型

訓練模型：利用ml5.js將收集的身體關節點資料傳給神經網路進行訓練。

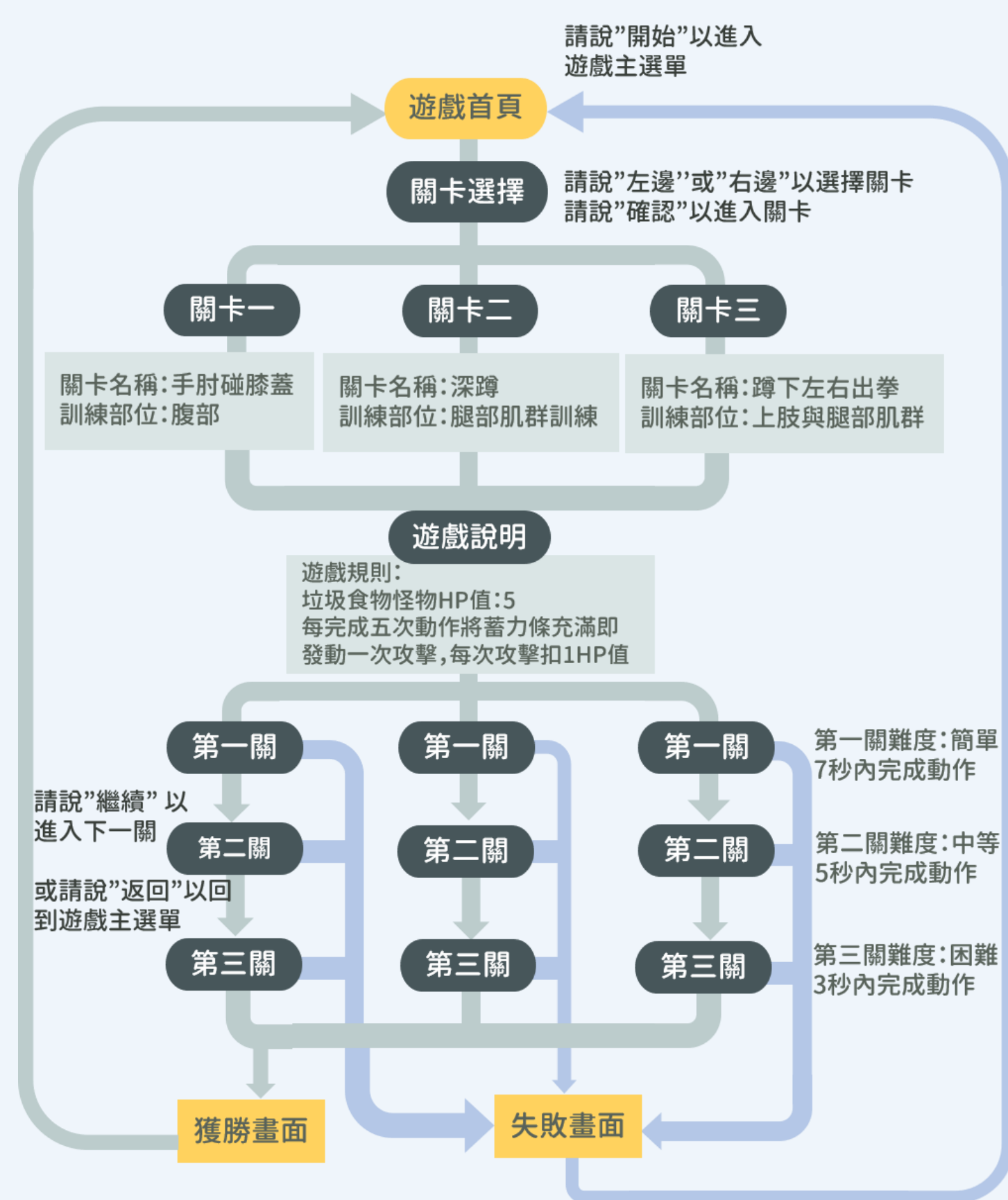
伺服器與瀏覽器的連接

利用WebSocket讓瀏覽器與Unity進行互動通訊，將神經網路模型的辨識結果傳給Unity。

語音辨識

透過p5.js所提供的語音辨識函式庫p5.speech.js，將接受到的語音轉成字串。

遊戲流程



肢體辨識技術

收集資料

首先站在Webcam前2.5公尺遠，做出欲被辨識的訓練動作，例如深蹲或是出拳，並大約維持此動作五分鐘左右，則ml5.js會每0.01秒擷取相機上的影像，辨識該影像中人體各個部位的特徵點，若要得到更多訓練資料加長時間即可，ml5.js會去辨識影像中人體各個部位的特徵點，例如：眼、耳、鼻、肩、臀，並將這些特徵點在平面上的x、y座標中做輸出。

訓練模型

將多筆資料的x、y座標餵入神經網路的Input layer，經過多次的運算來調整Hidden layer中各個神經元的weight。訓練所有weight後便可得到可以辨識特定肢體動作的神經網路。

辨識動作

每次讀取動作影像時，ml5.js轉出的xy座標乘上神經網路的weights所產生的一個數值能幫助我們辨識是否為想要的肢體動作。此數值介於0至1之間，每當數值越接近1，程式便會告訴使用者所做的動作是正確的，反之則不正確。

