前言

此文件包含 4 個部分，是關於大學期間的專題與研究的報告。

時間	名稱	貢獻度
2021~2022 年	中山大學專題(計畫名稱: 三維 CT 影像冠狀動脈鈣化自動標註)	25%
2023 年	中山大學專題(計畫名稱: 基於持續性深度學習檢測釣魚網頁並防護)	60%
2020 ~ 2023 年	手機遊戲逆向工程 App 漏洞(包含 Unity-IL2CPP、Cocos2dJs)	20% ~ 100%
	手機遊戲逆向工程封包漏洞利用(以 WPF、C# .net 開發應用)	30%
	逆向工程結合 Discord bot 將手機遊戲聊天訊息發送到 Discord 進行儲存。	100%
2023 年	艾法科技 iCtrl Pro 開發藍芽功能 (芯片: RTL8720CF、RTL8722DM)	藍芽功能: 40%

專題報告書

三維 CT 影像冠狀動脈鈣化自動標註

專題組員	貢獻範圍	貢獻比例
B083022014 吳怡萱	人工標註，分析實驗結果，成果報告	25%
B083022020 唐若婷	人工標註，分析實驗結果，成果報告	25%
B083022051 李奕勳	建立與修改模型	25%
B083022053 黃啟桓	程式撰寫	25%

指導教授：嚴成文 教授

嚴成文

(一)　摘要

心血管疾病(Cardiovascular Disease, CVD)長年高佔台灣十大死因第二名，僅次於癌症，在這之中其主因為冠狀動脈鈣化(Coronary Artery Calcification, CAC)，鈣化現象通常在動脈粥狀硬化早期就會發生，血管狹窄程度將與鈣化程度成正比。因此當冠狀動脈鈣化越嚴重，發生冠心病(Coronary Artery Disease, CAD)的可能性就越大。

為檢查心血管疾病， 通常以標準劑量電腦斷層掃描影像(Standard-Dose Computed Tomography, SDCT)，如圖 1.1，為醫師判斷依據。然而，其花費金額較高且輻射劑量高易對身體造成不良影響。而低劑量電腦斷層掃描影像(Low-Dose Computed Tomography, LDCT) ，如圖 1.2，為檢查肺炎、肺癌常用之輔助影像，其掃描出來的畫質較 SDCT 低，不易用於檢驗心臟。而年紀增長與抽菸習慣為心血管疾病與肺癌的共同危險因子，若可以達到同時檢查肺部與心臟的效果，不僅可大大降低醫療資源的浪費與病人的花費且能在檢查肺部疾病的同時檢驗心臟 是否有冠狀動脈鈣化的徵兆。

因此本計畫為使用高雄榮民總醫院醫師合作的病患資料，設計數位濾波器將雜訊多的 LDCT 影像還原至仿 SDCT 的標準影像，並將仿 SDCT 放入自動鈣化標註系統(DeepMedic)中針對鈣化區域做估算冠狀動脈鈣化分數(Agatston 分數)。

研究成果可快速且簡易判斷該病患是否有冠狀動脈鈣化及其嚴重程度，以協助醫師進行初期診斷。

圖 1.1 標準劑量電腦斷層掃描影像

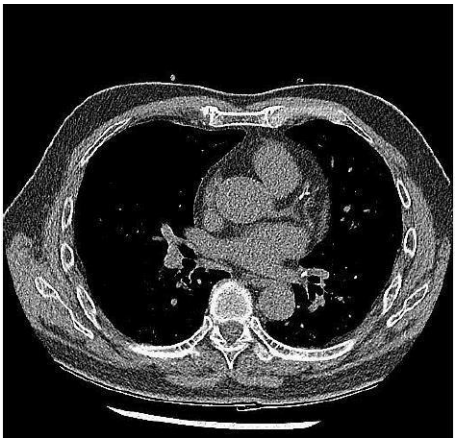


圖1.2 低劑量電腦斷層掃描影像

(二)　前言

冠狀動脈遍布於心臟表面，主要是提供心肌血液氧氣使心臟能正常收縮與舒張，並將加壓的血液透過血管輸送到全身各器官。位置起於主動脈根部，其中包含三條主要的動脈，右冠狀動脈、左前降支動脈及左迴旋支動脈，如圖 2.1。

冠狀動脈鈣化是由血管上班塊，也就是脂肪沿血管壁沉積所造成，當其累積到一定程度時，血管壁將增厚造成粥狀動脈硬化，而血管將會部份或完全阻塞，使供給心肌的氧氣和養分減少，引起心臟不適，如圖 2.2。

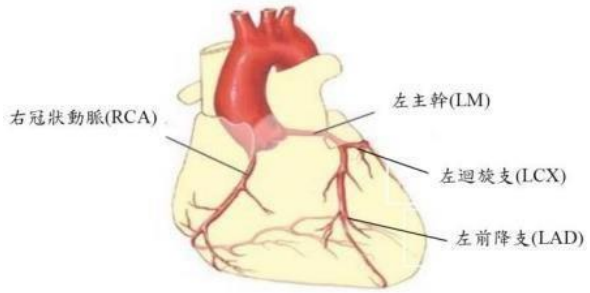


圖2.1 心臟及冠狀動脈立體示意圖[1]

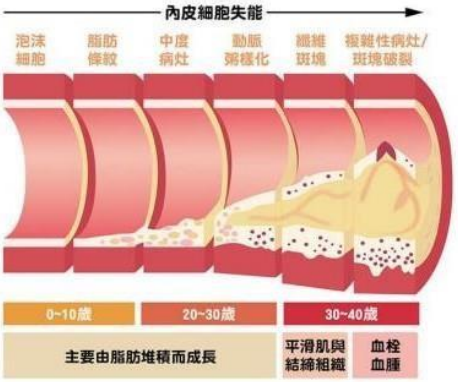


圖2.2 粥狀動脈硬化情形[2]

心肌梗塞發生情形有兩種。一為冠狀動脈發生堵塞，造成血液與養分不易通過因此導致心臟缺氧引發心絞痛，進而發生心臟肌肉壞死，出現急性心肌梗塞、心臟衰竭甚至猝死的現象；二為瓣膜出現鈣化、增厚的情形，使原本防止血液倒流的功能失效，無法完全撐開及閉合，造成血液從心臟輸出時就變得困難。[3]

許多患者在六、七十歲就開始有冠狀動脈與瓣膜鈣化的狀況但未被發現，等到七、八十歲時出現症狀，已相當嚴重。雖然目前醫療體系越來越進步，不再對這些疾病束手無策，常見的治療方式有藥物治療、心導管治療或是心臟血管繞道手術，但侵入式治療並不是我們所樂見的。

一般檢查 CAC 最直接的方式就是照射標準劑量電腦斷層掃描(SDCT)就能清楚地看到硬化情形，但有研究指出，經常重複接受電腦斷層掃描的輻射線 X 光的檢查者有一定程度的罹癌風險，所以醫生在判斷是否做電腦斷層掃描時， 除非是冠狀動脈心臟病的高危險群，否則不輕易讓患者使用，因此更難以在發生初期時及早診斷。而低劑量電腦斷層掃描(LDCT)為輻射量較低的電腦斷層 掃瞄，所使用的輻射劑量指有標準劑量的 1/4，已廣泛地用地早期肺癌篩檢上，但相較於標準劑量影像較模糊且雜訊也多。

(三)　研究動機

通常在做胸部低劑量電腦斷層掃描檢查肺癌時也會掃描到心臟區域，故本計畫以「預防勝於治療」作為出發點，目標是透過深度學習的方法去研究低劑量電腦斷層掃描影像，先將影像還原到與標準劑量相幫的清晰度，並分析心臟冠狀動脈鈣化的情形，以協助醫師對患者進行初期評估，使病人也只需做一次低劑量電

腦斷層掃描就能同時檢查兩個部位，不只降低輻射劑量的危害、減少醫療成本及時間，還達到「及早發現、及早治療」的效果。

(四)　文獻探討

4.1　卷積神經網路

卷積神經網路(Convolution Neural Network, CNN)是各種深度學習模型中，最具代表性的人工神經網路，許多影像辨識的模型都是以 CNN 的架構為基礎再進行延伸。CNN 的概念圖，如圖 4.1。

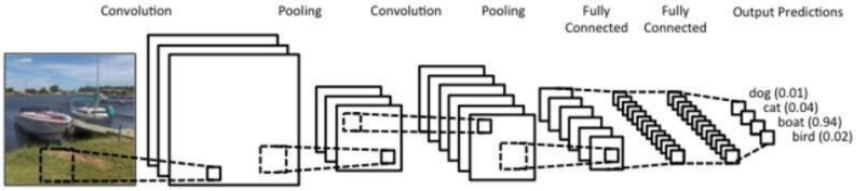


圖 4.1 卷積神經網路概念圖

4.2　高斯濾波[4]

高斯濾波，也叫高斯模糊，是圖像處理軟體中廣泛使用的處理效果，通常用它來**減少圖像雜訊以及降低細節層次**。高斯濾波對於圖像來說就是一個**低通濾波器**，可用於電腦視覺演算法中的預先處理階段，以增強圖像在不同比例大小下的圖像效果。

從數學的角度來看，圖像的高斯濾波過程就是圖像與常態分布做卷積。由於常態分布又叫作「高斯分布」。圖像與圓形方框模糊做卷積將會生成更加精確的焦外成像效果。

高斯濾波在二維空間的定義為:

$$G(u,v) = \frac{1}{2\pi\sigma^2} e^{-\frac{u^2+v^2}{2\sigma^2}}$$

其中 r 是模糊半徑($r^2 = u^2+v^2$)，σ 是常態分布的標準偏差。

分布不為零的像素組成的卷積矩陣與原始圖像做變換。每個像素的值都是周圍相鄰像素值的加權平均。原始像素的值有最大的高斯分布值，所以有最大的權重，相鄰像素隨著距離原始像素越來越遠，其權重也越來越小。這樣進行模糊處理比其它的均衡模糊濾波器更高地保留了邊緣效果。

圖 4.2、原圖(左)、均值濾波(中)和高斯濾波(右)

4.3　二值化(Binarization)

二值化是圖像分割的一種最簡單的方法。二值化可以把灰度圖像轉換成二值圖像。把大於某個臨界灰度值的像素灰度設為灰度極大值，把小於這個值的像素灰度設為灰度極小值，從而實現二值化。

4.4　八方向連通元件標記法(8-connected-components，又稱九宮格法)[5]

針對二值化後的影像去分區塊，每一個區塊給一個標籤，把所有相鄰區塊標示為同一個標籤，最後計算出整張影像每個像素是屬於哪個標籤，如果出現像素太少的標籤，就很有可能就是雜訊，在最後判斷為雜訊時就可以將此標籤上的所有像素都清除。

4.5 DeepMedic[6]

DeepMedic 是一個以多重尺度的 3D Deep Convolution Neural Network 為基礎的 3D 影像分割的軟體，此系統目前在掃描腦部的領域已有優異的表現，常被應用於診斷腦部腫瘤、腦部缺血性中風等腦部受損患者。

本研究採用 DeepMedic 自動標註鈣化系統，以深度學習的方式利用三維醫學影像分割分析鈣化情形，改善以往半自動計算鈣化分數所需大量人力選取鈣化區域。從篩選 HU 值大小可去除不必要的雜訊，使影像分為四大類，分別為背景、骨頭、冠狀動脈鈣化以及瓣膜鈣化部分，在這之中取最大連接物件只保留所需的冠狀動脈鈣化範圍套用於鈣化分數計算系統中，並依影像掃描的切片厚度做校正就可得到所要結果。相較過去使用的機器學習，利用深度學習的方式能更自動化且效果更好。

4.6 冠狀動脈鈣化分數評估

冠狀動脈得評估方法最早是由 Agatston[7]所提出，因此冠狀動脈鈣化分數又稱為 Agatston score，其分數高低對應相關風險如表 4.1 所示。

鈣化分數	鈣化斑塊程度	CAD 可能性	冠心病風險
0	無	極低	極小
0~10	微小	極低	小
11~100	輕度	低	中
101~400	中度	低~中	中高
401~1000	廣泛	中~高	高
>1000	非常廣泛	高	極高

表 4.1、鈣化分數對應相關風險

4.6.1　計算 Agatston score 之要點:

1. 鈣化物質於攝影像中的 HU 值會大於 130，故以 HU 值130 為分界，凸顯可能的鈣化區域。
2. 鈣化有可能分布在不同部位，選出第 i 個鈣化區域，為我們所關心發生在冠狀動脈區域的鈣化，將其設定為 Region of interest(ROI)。

3. 區域中最大的 HU 值決定CT_i^{max}，也決定權重值w_i，ROI 的面積為A_i，兩者相乘可計算出 Agatston score (CS_i)，如下方公式。

$$CS_i = w_i \times A_i$$

其中

$$w_i = 1 \ if \ 130\ HU \leq CT_i^{max} < 200\ HU$$
$$w_i = 2 \ if \ 200\ HU \leq CT_i^{max} < 300\ HU$$
$$w_i = 3 \ if \ 300\ HU \leq CT_i^{max} < 400\ HU$$
$$w_i = 4 \ if \ 400\ HU \leq CT_i^{max}$$

4. 將所有 ROI(整個心臟冠狀動脈)的 Agatston score 相加可得 Total Calcium Score (TCS) ，如下方公式。

$$TCS = \sum_{all\ ROIs} CS_i$$

(五) 研究方法及步驟

5.1 電腦斷層掃描影像篩選

本研究之電腦斷層掃描影像資料將由高雄榮民總醫院提供，資料分為兩個不同項目做訓練。(1) 成對的 LDCT 與 SDCT 影像，(2) 全部的 SDCT 影像

在篩選資料的過程中，由於許多資料的完整度不足，其影像有過度模糊或過度曝光的現象，或是也有些病人的安裝心臟支架情形過度嚴重，導致人工點選鈣化區域做黃金標準時造成判斷誤差。因此將這些資料經過人工篩選後，

(1) 完整成對資料有 300 筆。標準劑量電腦斷層掃描影像共有 17693 張；低劑量電腦斷層掃描影像共有 21419 張。

(2) 全部的 SDCT 有 1608 筆，共 88716 張。

5.2 高斯濾波架構

(1) LDCT圖檔可分為原圖層以及遮罩層(Mask)兩種

(2) 將LDCT的原圖層nii檔經過高斯濾波(設定sigma = 2, kernal_size = 5)做雜訊去除。

(3) 再將遮罩層進行二值化(以灰階值 0~255而言，閾值設定為100)，使灰階值100~255的數值調整為100，而灰階值段落控制在0~100，方便九宮格判斷。

(4) 再丟入九宮格中，設定九宮格區域周圍8個數值中超過(含)5個調整過的灰階值100(白)，將其判斷為非雜訊做保留;反之則被判斷為雜訊去除。

(5) 最後將高斯濾波過後的原圖層與做二值化及九宮格判斷後的遮罩層合併後，生成輸出圖檔(簡稱為LDCT*)。

5.3 人工標註驗證

使用LIFEX[8]此軟體將上述濾波過後的LDCT*做人工標註血管位置。標註方法如下所述，從第一張影像開始看，能先看到綠色動脈漸漸往右上方移動、黃色動脈逐漸往右下方移動，接著觀察到紅色動脈從左上方出現往下與黃色動脈會合。

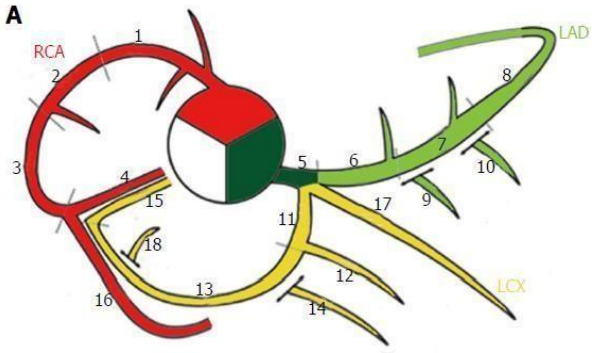


圖5.1、冠狀動脈標註示意圖

接著將人工標註過後的LDCT*利用文獻探討中介紹到的Agastone Score計算，比較其鈣化分數與SDCT分數判斷以上的高斯濾波成效如何。

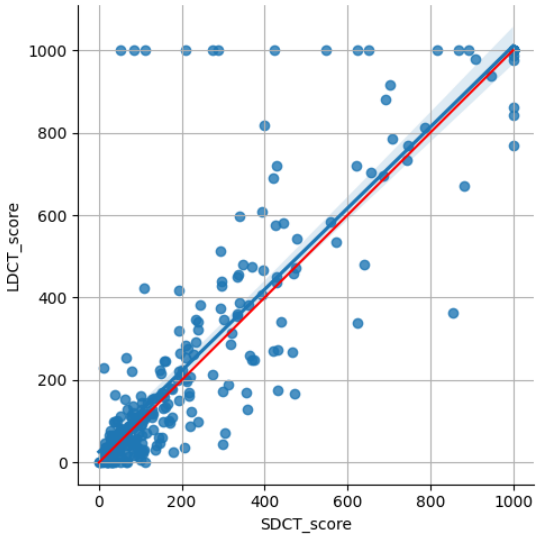


圖5.2、299筆SDCT與LDCT*回歸圖

5.4 DeepMedic訓練

本研究將總共1608筆的SDCT資料數，分為訓練資料1220筆、驗證資料149筆以及測試資料239筆，放進DeepMedic中訓練。從測試結果可得如下表5.1所示，模型在自動標註鈣化區域所算出的分數與SDCT分數接近，在鈣化分數較大時(大於100)基本上誤差不大，但在104筆分數為0的測試資料中，僅有40筆鈣化分數被判斷為0，甚至有7筆被判斷為100以上。顯現模型在判斷鈣化分數低的情況下，仍有改善的空間。

表5.1、測試資料的預測與真實分數比較表

預測分數 / 真實分數	0	1~10	11~100	101~400	401~1000	1000 以上
0	40	37	20	5	2	0
1~10	0	12	7	3	0	0
11~100	0	1	43	6	0	0
101~400	0	0	0	37	1	0
401~1000	0	0	0	0	15	0
1000 以上	0	0	0	0	0	10

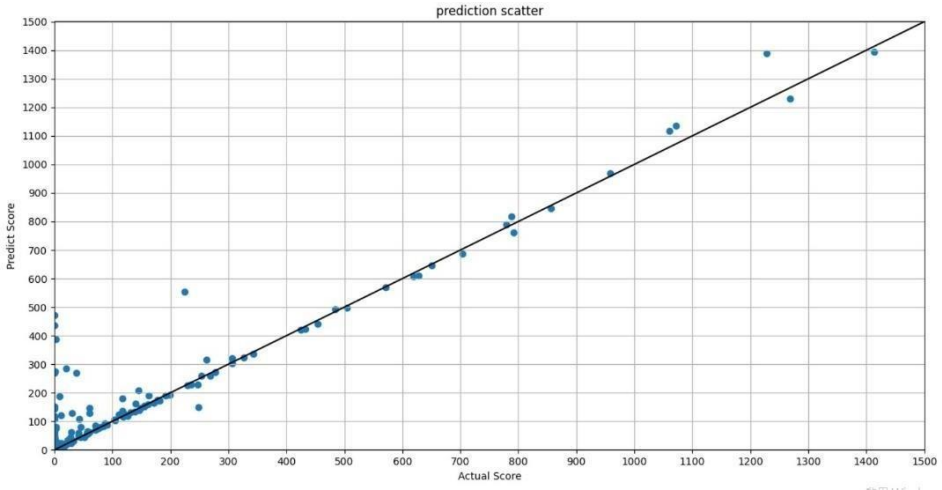


圖5.2、測試資料的預測與真實分數比較圖

5.5 LDCT自動鈣化標註

使用 5.4 中訓練出的最佳化模型，將經過高斯濾波後的 300 筆 LDCT*代入此模型中進行鈣化區域判斷並與所對應的 SDCT 分數做比較。

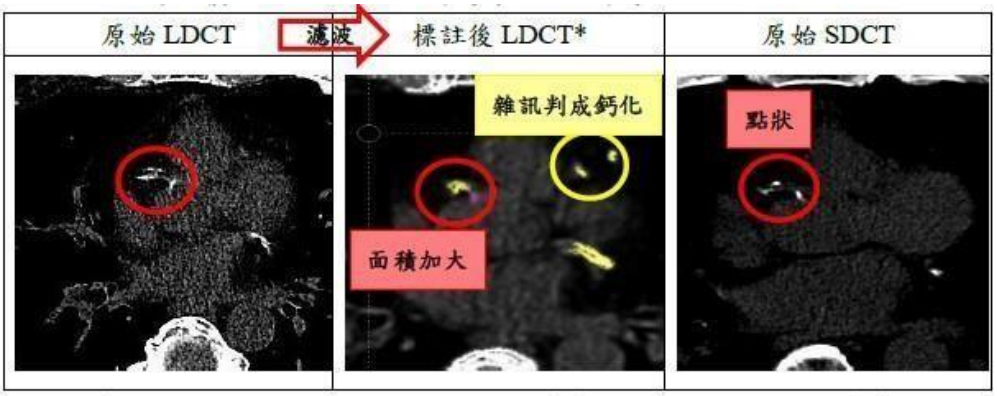
(六) 研究結果與討論

6.1 高斯濾波轉換影像結果

6.1.1 案例分析

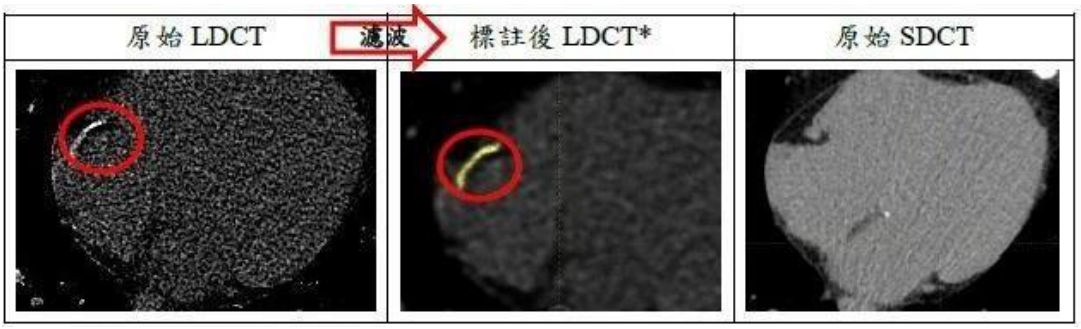
以案例 15197827 為例分析，計算出的 SDCT 鈣化分數為 1471，LDCT*鈣化分數為 2504；計算鈣化面積大小，SDCT 鈣化區面積為 149，LDCT*鈣化區面積為 383。

可以看出濾波後的 LDCT*和原始 SDCT 與LDCT 相比，經人工標註後有面積加大與雜訊被誤判成鈣化區的問題發生。



以案例 16385210 為例分析，計算出的 SDCT 鈣化分數為 894，LDCT*鈣化分數為 1545；計算鈣化面積大小，SDCT 鈣化區面積為 121，LDCT*鈣化區面積為 233。

將原始 LDCT 與標註後 LDCT*比較，可發現標註後 LDCT*因上述濾波設計中使用的八方向連通元件標記法導致鈣化區面積加大，因此被誤判成鈣化區。再和原始 SDCT 相比，由於切片密度的不同，在原始 SDCT 中並沒有看到相同的區域出現雜訊。



6.1.2 總結

以相同方法分析其餘案例，造成鈣化分數誤差的原因可分為以下幾點:

(1) **切片密度問題**: SDCT 和 LDCT 切片密度不同。例如:同一鈣化區域在 LDCT 可能有 5 張，但在 SDCT 只有 2 張，因而造成誤差。

(2) **濾波問題**: LDCT 本身鈣化區不易判斷又經濾波而導致被微幅加大，當 SDCT 鈣化區是由多個小鈣化區組成且在鄰近區域，LDCT 在濾波後由於高斯模糊導致被鈣化區面積被加大。

(3) **人工標註的失誤**: 在不同人判斷 LDCT*可能導致雜訊與鈣化區的判斷不同。

(4) **本身 SDCT 分數有誤**: 標準 SDCT 在早期標註時出現失誤，需再經過修正。

(七)　參考文獻

[1] Kamnitsas, K., Ledig, C., Newcombe, V., Simpson, J. P., Kane, A. D., Menon, D. K., Rueckert, D., & Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical image analysis, 36, 61–78.*

[2] 彭幸茹、許嘉真（2017）。心血管專題有關血管的二三事。
檢自：https://heho.com.tw/archives/1808

[3] 衛生福利部國民健康署(2004)。認識冠心病。
檢自: https://www.hpa.gov.tw/Pages/Detail.aspx?nodeid=632&pid=1188

[4] 台部落(2018)。圖像處理之均值濾波，高斯濾波(高斯模糊)，中值濾波，雙邊濾波
檢自: https://www.twblogs.net/a/5bb29e4d2b71770e645df36a

[5] 拿著放大鏡看自己(2014)。影像處理:Component Labeling(標號)。檢自: http://mermerism.blogspot.com/2014/05/component-labeling.html

[6] DeepMedic 官方網站。
檢自: https://deepmedic.org/

[7] Agatston, A. S., Janowitz, W .R., Hildner, F. J., Zusmer, N. R., Viamonte, M. Jr., & Detrano, R. (1990). Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol, 15(4), 827-832

NSTC 國家科學及技術委員會
National Science and Technology Council

學術補助獎勵查詢

回國科會首頁

綜合查詢 | 補助 | 獎勵

訪客人次：4318313

大專學生研究計畫

年度 ： 110 ~ 111

執行機關： 全部

學生姓名：　　計畫名稱：

學門 ： 全部

排序 ： ◉依年度 ○依機關 ○依姓名　　查詢

設定每頁顯示筆數： 10

年度	學生姓名	執行機關	內容
111	吳怡萱	國立中山大學機械與機電工程學系（所）	計畫名稱：三維CT影像冠狀動脈鈣化自動標註 計畫編號：111-2813-C-110-068-E 成果報告：無電子檔 執行起迄：2022/07/01~2023/02/28 指導教授：嚴成文 核定金額：48,000元

共1頁(共1筆)，目前在 第1頁

(本查詢結果僅供參考，實際補助結果以本部正式核定通知為準。)

科技部
111年度大專學生研究計畫核定名冊

申請機構：國立中山大學

序號	計畫編號	指導教授	職稱	申請機構	學生姓名	年級	就讀學校及科系	專題計畫名稱	歸屬司處	研究助學金（元）	耗材費(元)	合計(元)
1	111-2813-C-110-001-E	楊政融	助理教授	國立中山大學人文暨科技跨領域學士學位學程	黃靖珈	2	國立中山大學人文暨科技跨領域學士學位學程	3D列印再生塑膠線材在機械性質和成本的最佳製造參數設計	工程司	48,000	10,000	58,000
2	111-2813-C-110-002-M	楊弘敦	教授	國立中山大學物理學系（所）	張宇皓	3	國立中山大學物理學系（所）	探討MnGeTeO6為新多鐵性材料之可能性	自然司	48,000	0	48,000
3	111-2813-C-110-003-M	周雄	教授	國立中山大學物理學系（所）	呂佳恩	3	國立中山大學物理學系（所）	探討Cu2V2O7的α、β相轉變和鐵彈性的關係	自然司	48,000	0	48,000
4	111-2813-C-110-004-M	盧怡穎	助理教授	國立中山大學物理學系（所）	張名涵	3	國立中山大學物理學系（所）	介電層的屏蔽效應對InSe 金屬絕緣體相變的影響	自然司	48,000	0	48,000
5	111-2813-C-110-005-M	張鼎張	教授	國立中山大學物理學系（所）	邵奎祐	3	國立中山大學物理學系（所）	在真空下Schottky AlGaN/GaN HEMT 熱載子測試異常的劣化現象探討	自然司	48,000	0	48,000
6	111-2813-C-110-006-M	郭建成	副教授	國立中山大學物理學系（所）	施柏安	3	國立中山大學物理學系（所）	水自行催化於鎵砷共存之砷化鎵表面結構及電性量測	自然司	48,000	0	48,000
7	111-2813-C-110-007-M	洪玉珠	副教授	國立中山大學光電工程學系	陳致文	3	國立中山大學光電工程學系	基於五氧化二鉭薄膜實現非線性多模干涉全光開關之研究	自然司	48,000	0	48,000
8	111-2813-C-110-009-H	洪世謙	教授	國立中山大學哲學研究所	顏巍一	2	國立中山大學政治經濟學系	Between Violence and the State: Hannah Arendt's Thought on Violence（暴力與國家之間：漢娜鄂蘭關於暴力的思考）	人文司	48,000	5,000	53,000
9	111-2813-C-110-010-M	林伯樵	教授	國立中山大學化學系（所）	李怡樺	3	國立中山大學化學系（所）	新穎化學策略於醣蛋白微陣列晶片之應用	自然司	48,000	0	48,000
10	111-2813-C-110-011-M	邱政超	助理教授	國立中山大學化學系（所）	簡奕先	3	國立中山大學化學系（所）	自動辨別週期性化學結構之演算法開發	自然司	48,000	0	48,000

63	111-2813-C-110-066-E	洪裕涵	助理教授	國立中山大學光電工程學系	黃懷樘	3	國立中山大學光電工程學系	類骨質3D列印材料的開發應用於顱骨手術模擬器	工程司	48,000	10,000	58,000
64	111-2813-C-110-068-E	嚴成文	教授	國立中山大學機械與機電工程學系（所）	吳怡萱	3	國立中山大學機械與機電工程學系（所）	三維CT影像冠狀動脈鈣化自動標註	工程司	48,000	0	48,000
65	111-2813-C-110-069-B	黃淑萍	助理教授	國立中山大學生物科學系（所）	呂紹辰	3	國立中山大學生物科學系（所）	評估脫水程度對蜥蜴血液滲透壓改變的影響	生科司	48,000	0	48,000
66	111-2813-C-110-070-B	劉世慧	助理教授	國立中山大學生物科學系（所）	高瑄蔚	3	國立中山大學生物科學系（所）	好茶部落固有小米之分類研究	生科司	48,000	10,000	58,000
67	111-2813-C-110-071-B	王亮鈞	助理教授	國立中山大學海洋生物科技暨資源學系（所）	呂嘉芸	2	國立中山大學海洋生物科技暨資源學系（所）	利用魚皮模組研究黏膜共生菌對於魚皮免疫反應之影響	生科司	48,000	10,000	58,000
68	111-2813-C-110-072-B	劉莉蓮	教授	國立中山大學海洋科學系	高立容	3	國立中山大學海洋生物科技暨資源學系（所）	小琉球芋螺的食性和營養位階探討	生科司	48,000	0	48,000
69	111-2813-C-110-073-B	李昆澤	教授	國立中山大學生物科學系（所）	陳孟云	3	國立中山大學生物科學系（所）	常壓高氧對頸部脊髓挫傷後呼吸功能與脊髓神經發炎之影響	生科司	48,000	0	48,000
70	111-2813-C-110-074-B	許晉銓	教授	國立中山大學生物醫學研究所	呂天馨	3	國立中山大學生物科學系（所）	子宮抹片臨床研究	生科司	48,000	0	48,000
71	111-2813-C-110-075-E	王郁仁	副教授	國立中山大學機械與機電工程學系（所）	江冠霆	3	國立中山大學機械與機電工程學系（所）	微小力量感測器	工程司	48,000	0	48,000
72	111-2813-C-110-076-E	胡龍豪	副教授	國立中山大學機械與機電工程學系（所）	許紘睿	3	國立中山大學機械與機電工程學系（所）	結合深度學習和控制系統的視覺反應系統	工程司	48,000	10,000	58,000
73	111-2813-C-110-077-E	許煜亮	助理教授	國立中山大學機械與機電工程學系（所）	江杰飛	3	國立中山大學機械與機電工程學系（所）	適路性自動化車燈系統之研製	工程司	48,000	10,000	58,000
74	111-2813-C-110-078-B	吳長益	副教授	國立中山大學生物科學系（所）	曾品心	3	國立中山大學生物科學系（所）	Rasl11b對斑馬魚胚胎血管發育的影響（The effects of Rasl11b on vascular development in zebrafish）	生科司	48,000	0	48,000
75	111-2813-C-110-079-E	程啟正	教授	國立中山大學機械與機電工程學系（所）	蘇怡綾	3	國立中山大學機械與機電工程學系（所）	具上下坡功能之仿生機器蛇	工程司	48,000	10,000	58,000

團隊成員貢獻說明

專題名稱(中文)：基於持續性深度學習檢測釣魚網頁並防護

專題名稱(英文)：Detect and protect against phishing sites based on continual deep learning

※ 請詳述團隊分工情形：

	個人完成內容	貢獻比例	組員簽名
1.	程式撰寫、報告撰寫、建立及修改模型	60%	黃啟植
2.	程式撰寫、報告撰寫	25%	鍾多捷
3.	程式撰寫	15%	黃柏翔

※ 本作品若以其他原著作品為基礎，經大幅度修正或改進者，請詳述本作品與原著作品之關連性及不同之處：

註：如不敷填寫，可另以附件呈現。

指導教授簽名：

日期：112 年 9 月 26 日

基於持續性深度學習檢測釣魚網頁並防護

Detect and protect against phishing sites based on continual deep learning

國立中山大學資訊工程學系

111 學年度大學部專題製作競賽

組員: B083022053 黃啟桓

組員: B093040040 鍾名捷

組員: B093040042 黃柏翔

指導教授 : 徐瑞壕教授

摘要

隨著網路的蓬勃發展，網路詐欺、竊取資料等攻擊層出不窮，不僅是利用人性弱點，而擅長去偽裝釣魚網站的手法也非常逼真。無形中讓受害者交出自己的個人資料、財產安全、裝置權限。這些隱私資料對受害者的影響非常重大，可能面臨財產或身分遭盜用的結果。不只是個人用戶，企業方也深受其擾，並損失慘重。使得網路安全成為一個不可忽視的課題。

基於人工智慧的釣魚網頁複合式檢測，利用包含 URL 地址比對和網頁特徵提取，並同時兼顧防禦以及用戶隱私問題，高度的彈性以檢測多樣與多變的釣魚手段，我們也希望可以將「持續性深度學習(continual-learning)」加入此專題中，學習新的技能或是任務時不會將過去學習而來的知識忘記，就像人類可以不斷學習新知識，並同時利用舊的知識，以適應多變及多樣的釣魚網站。本專題將重點聚焦於網頁的檢測，以簡易的工具來面向廣大的瀏覽器用戶。

關鍵字: 釣魚檢測、釣魚防禦、Phishing detection、人工智慧

目錄

圖次

第一章 緒論

第一節 研究動機

一、釣魚網頁的危害

釣魚攻擊已經成為網路使用者面臨的重大安全威脅之一。釣魚攻擊通常是指詐騙者通過製作虛假網站或電子郵件，以欺騙網路使用者提供其個人敏感信息，如密碼、帳號、信用卡號碼等。釣魚攻擊的目的是盜取個人信息或進行其他惡意行為，這對個人隱私和金融安全構成嚴重威脅。

釣魚攻擊的方式愈來愈隱蔽，以至於許多網路使用者往往無法分辨真假網頁或電子郵件，因此，建立一種有效的檢測釣魚網頁的方法是非常必要的。此外，釣魚攻擊也對企業和組織的安全造成了嚴重威脅，因為釣魚攻擊通常針對企業和組織的員工，試圖從中獲取敏感信息。

因此，針對釣魚攻擊問題進行研究和開發釣魚網頁檢測工具是十分必要的。這將有助於保障網路使用者和企業的安全，減少釣魚攻擊對社會帶來的傷害。

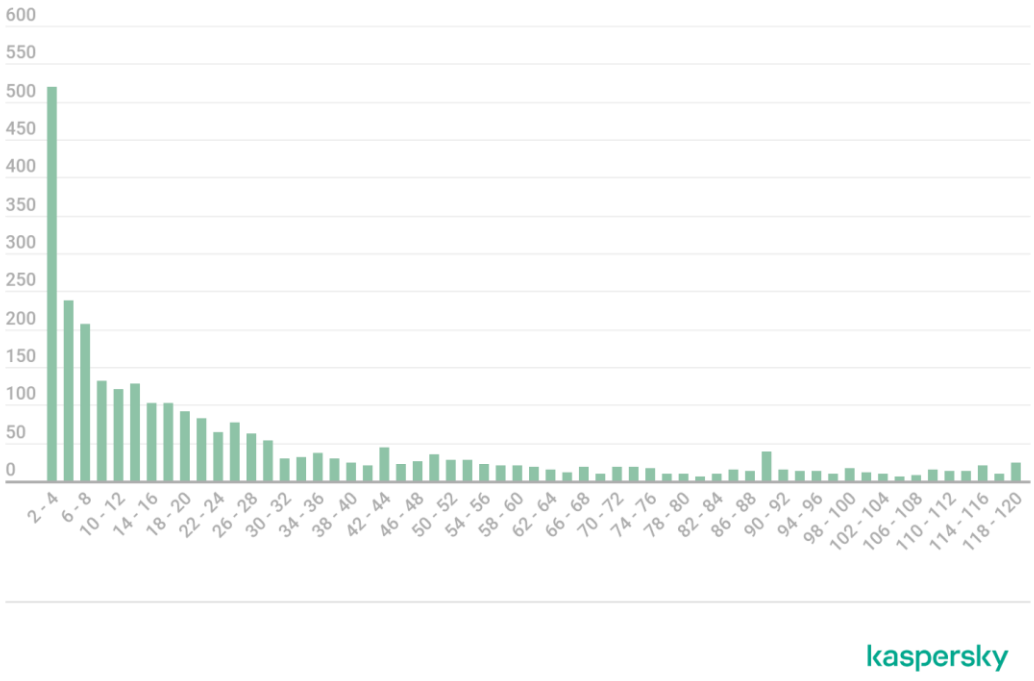
二、釣魚網頁發展

釣魚網頁的發展可以追溯到 20 世紀 90 年代，當時互聯網使用率開始普及，網路使用者逐漸增加，網路安全問題也開始受到關注。當時，黑客和駭客開始利用釣魚網頁來進行攻擊，盜取網路使用者的帳號和密碼等個人信息。

在 2000 年代初期，釣魚攻擊成為了一種主要的網路安全威脅，並且不斷發展和演進。釣魚攻擊的方式也變得更加隱蔽和精細，例如製作更真實的假冒網站、使用社交工程技術、利用電子郵件等途徑進行攻擊等。

隨著網路技術的不斷發展和安全技術的提升，釣魚攻擊的形式也越來越多樣化。現代釣魚攻擊不僅限於假冒網站和電子郵件，還包括社交媒體帳號、手機應用程式等。

根據釣魚網站(https://securelist.com/phishing-page-life-cycle/105171/)所提供的資料顯示，大部分的釣魚網站活躍的時間在 48 小時內。而要如何應對釣魚攻擊的不斷變化，就是我們本專題目標。



圖一、 根據活躍小時數對釣魚鏈接分類圖。該圖顯示了每個鏈接生命週期前五天的數據。

三、釣魚網頁-惡意軟件分發

釣魚網站中的惡意軟件分發已成為當今網絡安全領域中一個嚴重的問題。釣魚網站通過欺騙手段和惡意意圖，試圖引誘用戶下載和執行惡意軟件，從而對用戶和組織的信息安全造成嚴重威脅。

這種釣魚網站中的惡意軟件分發行為對用戶和組織的安全和隱私構成了重大威脅。這些惡意軟件可能包含病毒、間諜軟件、勒索軟件等，可以竊取用戶敏感信息、破壞系統功能、散播廣告或進行其他惡意活動。此外，這些惡意軟件的分發也對網絡生態系統的穩定運行產生了不良影響。

因此，本論文的動機在於解決釣魚網站中的惡意軟件分發問題。我們希望提出一種有效的方法，能夠檢測和對抗釣魚網站中的惡意軟件分發行為，以保護用戶免受惡意軟件的威脅。

通過研究和解決這個問題，我們能夠提高用戶和組織的網絡安全水平，減少釣魚攻擊對用戶的損害。同時，我們的研究還有助於維護網絡生態系統的健康運行，減少惡意軟件對整個網絡社區的影響。

五、釣魚網頁-自動及惡意的網頁跳轉

釣魚網站作為一種網絡安全威脅，已經對用戶和組織的信息安全造成了重大影響。釣魚網站使用各種欺騙手段來引誘用戶提供個人敏感信息，從而進行詐騙、竊取身份信息或散播惡意軟件。其中，釣魚網站使用自動及惡意的網頁跳轉技術，進一步增加了攻擊的隱匿性和成功率。

這種自動及惡意的網頁跳轉行為使得釣魚網站可以將用戶引導到意想不到的網頁，進而進行欺騙、攻擊和數據竊取等惡意活動。這對用戶和組織的信息安全構成了嚴重威脅，並對網絡生態系統的穩定運行產生了不良影響。

因此，本論文的動機在於解決釣魚網站使用自動及惡意的網頁跳轉的問題。我們希望提出一種有效的方法，能夠檢測和阻止釣魚網站中的自動及惡意的網頁跳轉行為，以保護用戶免受釣魚攻擊的危害。

通過研究和解決這個問題，我們能夠提高用戶和組織的網絡安全水平，減少釣魚攻擊對用戶的損害。同時，我們的研究也有助於維護網絡生態系統的穩定運行，減少惡意行為對整個網絡社區的影響。

因此，本論文的動機在於研究並提出有效的方法來解決釣魚網站使用自動及惡意的網頁跳轉的問題，以提高網絡安全性並維護網絡生態系統的健康運行。

第二節 研究問題

一、 哪一種應用是對於大眾而言最便利的

由於此專題著重於網頁的安全，所以相較於開發新的應用程式來保護使用者使用瀏覽器時面對的釣魚威脅，不如配合瀏覽器自帶的擴充功能使得安裝、卸載更加方便，並且只要瀏覽器啟動，其擴充功能也會跟著啟動，使得用戶在使用瀏覽器時全程受到保護。

二、 如何建立高度彈性的防護

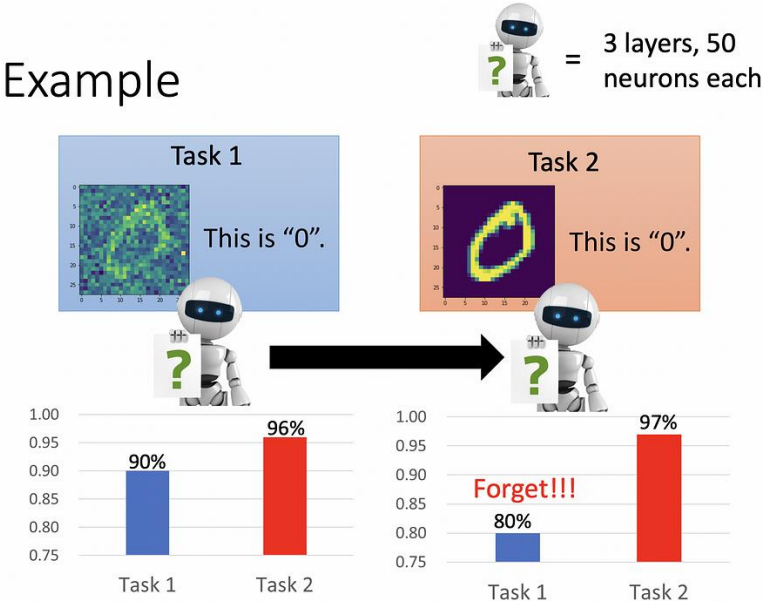
對於 URL、域名(domain)、主機(hosts)在網頁載入前與黑名單進行比對，黑名單來源每天更新一次。對於未在黑名單的釣魚網頁在網頁載入中進行 html 結構分析與 javascript 分析，載入完成後以網頁圖像進行深度學習預測是否為釣魚網頁。預測釣魚網站的深度學習模型以增量學習的概念，使得模型可以不斷的更新，以面對多變的釣魚網頁。

三、 人工智慧如何使檢測釣魚網站的準確率及效率更加提升

傳統的檢測方式仰賴龐大的資料庫，需要將標記為釣魚網站的資料紀錄到黑名單資料庫中。而檢測過程關係到搜索黑名單，將目標網站比對黑名單判斷是否為釣魚網站，執行速度將隨著黑名單的增加變得遲緩。且因釣魚網站推陳出新，在第一時間無法透過黑名單檢測成功。透過人工智慧的協助，只需要將標記過的資料訓練出能夠偵測釣魚網站的模型，即可讓訓練模型舉一反三地辨別出釣魚網站，提高準確率。而且只需要將目標網站代入模型算出結果，比傳統的資料庫檢索更加快速。

四、 克服「災難性遺忘」

災難性遺忘是指人工神經網絡在學習新資料時突然而崩潰性的忘記先前學習的資料的趨勢。具體來說，這些問題指的是製作對新信息敏感但不受新資料干擾的人工神經網絡的挑戰。藉由 Synaptic Intelligence(SI)的模型，能夠更長期持續學習新的訓練數據，避免未來的訓練模型變得無法解決先前的訓練數據。



圖二、 災難性遺忘的概念圖。

第三節 文獻回顧與探討

一、 文獻回顧

閱讀數篇關於釣魚網頁檢測的論文，了解釣魚網頁的特徵。

釣魚網址的案例:

1. 網址混淆:

   例如: http://mail-google.com、http://www.goog1e.com (goog1e 的 l 是數字一)

2. 二級域名混淆:

   http://mail.google.com.xyz，也與 google 本身沒有關係

3. 縮網址:

   縮網址服務會遮蔽連結的網址，使用者無法有效的判斷是否為釣魚網址。

4. 網址嫁接:

   DNS 伺服器被入侵並趁機竄改設定。

5. 網頁偽裝:

   使用相近的網頁 html 結構及圖像來混淆視覺。

傳統的網頁釣魚檢測主要基於用戶的回報，依賴於黑名單。近年來，由於釣魚網站的快速生成，短壽命的特性，使得傳統的網頁釣魚檢測缺乏立即性，未能提供有效的防護功能。

而相關的網頁釣魚檢測技術隨著機器學習蓬勃發展，開枝散葉。本計畫希望結合數種網頁釣魚檢測工具。以增強面對釣魚網站的防禦力，比較單一網頁釣魚檢測工具與本計畫的檢測效果進行比對。並且以實際的行為防禦來自釣魚網頁的攻擊。

查閱大量來自 github 的原始碼，大多專案對於釣魚網站的偵測及防禦沒有整合在一起。而使用瀏覽器擴充功能的軟體大多使用黑名單的機制來判斷釣魚網頁，缺乏彈性。因此此研究希望透過將這些功能進行整合，以保護使用者使用瀏覽器上網的安全。

二、 文獻探討

使用機器學習來判斷釣魚網站中的眾多特徵值，以權重及分類找出特徵值。並使用決策樹、隨機森林、類神經網路、倒傳遞類神經網路等演算法，但是因為在演算法中當訓練集資料比例達到一定的比例或參數過多時，便產生「過度配適」的風險，可以使用「交叉驗證」解決，將訓練資料進行分組，一部分做為訓練子集來訓練模型，另一部分做為驗證子集來評估模型。

[16]如果直接將訓練好的模型再次以新的資料訓練，則有可機會發生災難性遺忘，再也無法解決舊的資料。於是採用 Synaptic Intelligence(SI)的模型，能夠更長期持續學習新的訓練資料，避免未來的訓練模型變得無法解決先前的訓練資料。SI 的原則是學習新的資料前，先計算目前模型中節點的重要性，若改變節點的值越不會改變準確性，則代表節點的重要性越低。訓練新的資料時只要改變重要性較低的節點，就能訓練出不但能解決新的資料，也不會遺忘舊的數據的資料。

第四節 研究方法及步驟

一、 研究方法

透過閱讀數篇關於釣魚網頁檢測的論文，並進行實踐，分析各種檢測方法的優缺點及效果，取長補短以達到檢測效果的提升。並建立一套防禦機制，阻止釣魚網頁的運作，並回報給釣魚網頁回報的機構，以降低釣魚網頁的危害，增加使用網路的安全性。

釣魚網頁分析的特徵[14]:

1. 域名的特徵:
    a. 域名是否有 IP 地址
    b. 域名的長度
    c. 域名的子域名數量
    d. 域名中是否包含 "@"符號
    e. 域名中是否包含 "-"符號
    f. 域名的進行重定向 (Redirection)
    g. 域名是否安全 “http/https”
    h. 域名是否為縮網址服務
    i. 域名中英文的比例
    j. 域名中數字的比例
    k. 域名中符號的比例
2. 外部提供的特徵:
    a. 域名是否在 whois 註冊
    b. 域名的瀏覽排名
    c. 域名年齡(age)
    d. 域名的過期時間

3. HTML 和 JavaScript 的特徵:

   a. IFrame 重定向

   b. 狀態欄 onmouseover 事件

   c. 禁用右鍵點擊

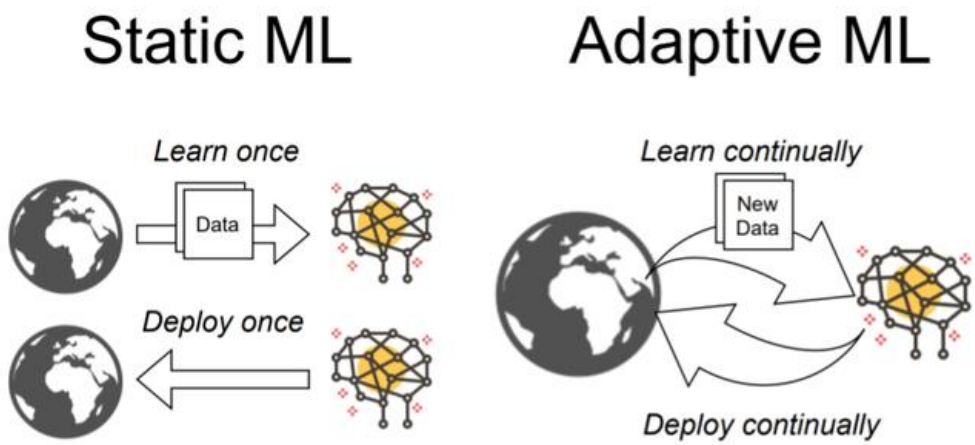
   d. 網站重定向次數

持續性深度學習:

傳統機器學習通常首先使用數據中的所有可用示例學習模型，然後部署以供實際使用。這意味著無論何時模型完成學習，在實踐中使用時它都保持不變。該模型的靜態性質存在問題，因為它不適合我們不斷變化的世界。這讓許多現代機器學習方法無所適從，因為部署時的靜態模型無法使用永無止境的數據流。持續學習旨在通過仔細研究動態機器學習模型、靈活且持續地適應數據來解決這個問題。

「連續學習系統可定義為一種自適應算法，能夠從連續的信息流中學習，這些信息隨著時間的推移逐漸變得可用，而要學習的任務數量（例如分類任務中的成員類別）並不是預先定義的。重要的是，新信息的納入應該在不產生災難性遺忘或干擾的情況下進行。」[17]

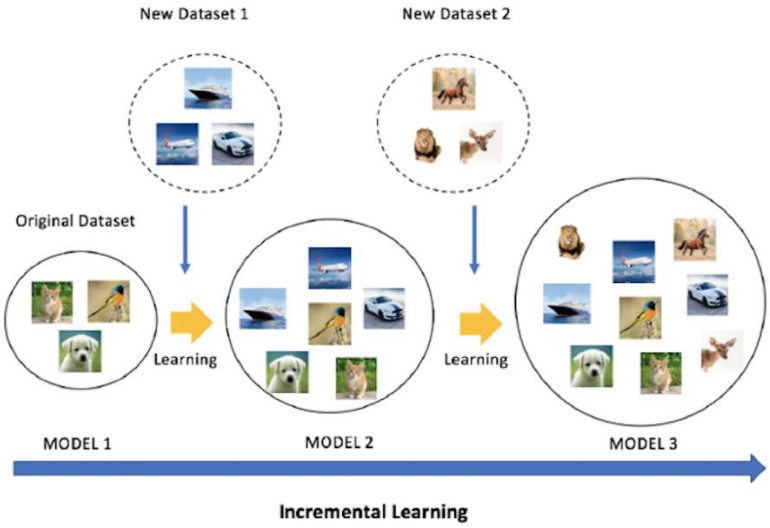
因此，在持續學習場景中，隨著任務分佈在其生命週期內動態變化，學習模型需要逐步構建和動態更新內部表示。理想情況下，此類內部表示的一部分將具有通用性和不變性，足以在類似任務中重複使用，而另一部分應保留和編碼特定於任務的表示。



圖三、 傳統機器學習與自適應性機器學習的差別

增量學習:

增量式學習主要的特性及應用在**動態式資料庫**，一般而言，資料量是逐漸增加的，在面臨新的數據時，演算法要能應對以訓練好的模型進行一些改動，以學習新的資料具備的知識，但又不應該過分更改模型，喪失已從舊有的資料所學習到的知識，而**對一個訓練好的模型進行修改的時間成本通常遠低於重新訓練一個模型所花費的成本**，這也更加符合人類的思維原理。

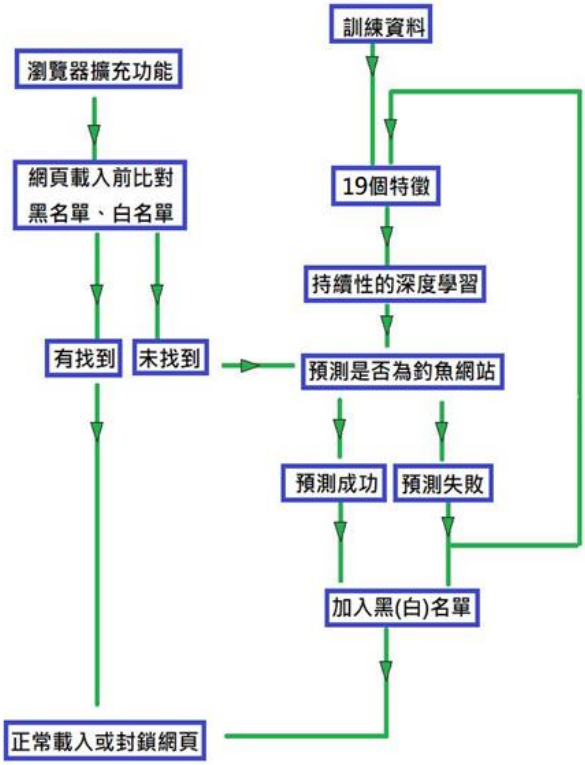


圖四、 增量學習的概念圖

二、 研究步驟

1. 抓取網路上公開的釣魚網站名單(https://gitlab.com/malware-filter/phishing-filter)以及正常網站名單(https://www.kaggle.com/datasets/aman9d/phishing-data)進行資料的蒐集。
2. 進行資料的前處理，包含清除無法訪問的網站。以及取得所有域名的19 個特徵。
3. 對於論文中釣魚網頁檢測的檢測方法進行實踐，例如:分析域名、數據分析，對照論文的結果與自身實作的結果，並進行修正，以提升單個檢測技術的檢測能力。
4. 結合多個釣魚網頁檢測技術，取長補短以達到檢測效果的提升。例如:傳統的釣魚網站檢測缺乏保護機制。

5. 建立一套釣魚網頁防禦機制的軟體，以阻止釣魚網頁的運作，自動暫停載入數據、自動清空 cookies、自動封鎖網站使用 JavaScript、自動封鎖網站背景執行。
6. 分析軟體的優缺點，包含檢測的速度、準確性，阻斷釣魚網頁的執行速度等方面，盡可能的達到最好。
7. 結合增量學習的技術作為深度學習的工具，使其具有持續學習的特徵。



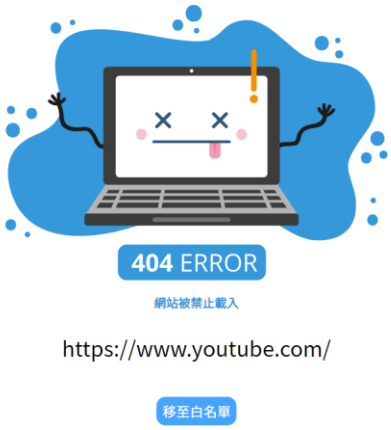
圖五、 運作流程圖。左邊為客戶端的運作模式，右邊為伺服器運作模式

第五節 研究成果

一、 Client 框架

框架採用 Chromium 基礎瀏覽器通用的擴充功能(extension)。

預測到該網站包含惡意軟件分發的時，將進行網頁的封鎖。以防止危害的發生。也可以自行針對個別網站加入黑名單。



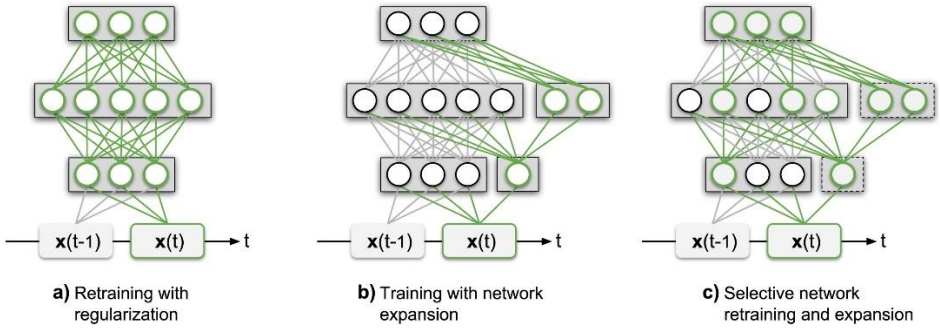
圖六、 嘗試將 youtube 網頁加入黑名單的結果

二、 Server 架構

使用 cookie 儲存每個使用者的最近期的訓練模型，在請求新的訓練資料時能透過持續性深度學習進行模組的更新，並且回饋給客戶端，使得客戶端能更新自己量身定做的模型。

三、 持續性深度學習

持續性深度學習模組，此模組在訓練模型生成時，會讀取舊版的訓練模型，並長出旁支、更新訓練模型，以此來避免「災難性遺忘」的問題。



圖七、 持續性深度學習的概念圖[17]。

四、 增量學習應用於釣魚網站的訓練

◆ 實驗方法

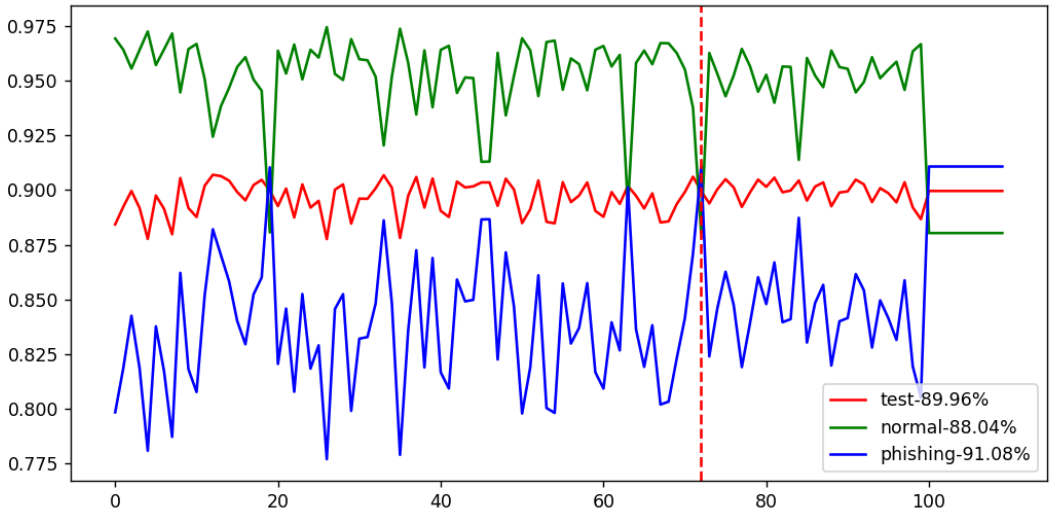
訓練資料: 5000 筆釣魚網站特徵及 5000 筆正常網站特徵隨機打亂。

測試資料: 5000 筆釣魚網站特徵及 5000 筆正常網站特徵隨機打亂。

訓練方法: 先使用 sklearn 所提供的隨機梯度下降法擬合線性模型 (SGD) 一般的深度學習(fit)，進行 1000 次迭代(Iteration)，再將訓練資料拆分成 100 組小資料集，每次以一小組增量的情況下進行增量訓練，每次訓練後，比對一下前一次儲存的訓練模型的精確度(專注於釣魚網站的精確度)，如果精確度上升，則儲存此模型，反之，則不做任何動作就進行下一組的增量訓練。

◆ 實驗結果

由於經過了一般的深度學習，所以此模型的具有高 robust，使得小數量的增量學習中精確度變化不大，測試資料的精確度穩定在 88%。在第 84 次增量學習的訓練模型中，釣魚網站的精確度具有高 91.08%，而測試資料的精確度高達 89.96%，正常網站的精確度高達 88.04% 。



圖八、 增量學習結合 SGD 實驗結果圖。

◆ **實驗討論**

以實驗結果顯示，即使使用增量學習的概念，還是會遇到災難性遺忘的問題(有顯著的波動，並非穩定的提升精確度)。但是使用實驗發現得益於使用 sklearn 所提供的 SGDClassifier 的 fit 函數，使得災難性遺忘的崩潰程度顯著的降低(圖九)，相當於純使用增量學習的 10~15 倍的效果(圖十)。

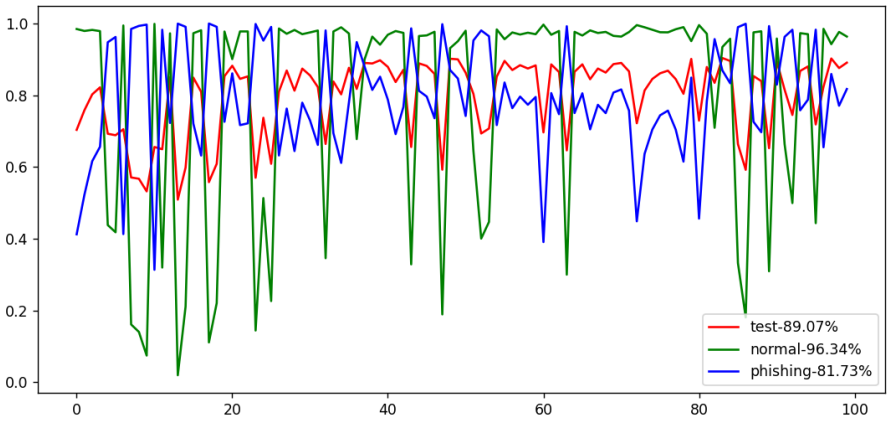
➢ **實驗方法的優點**

Robust: 一般深度學習(sklearn 所提供的 SGDClassifier 的 fit 函數)相較於純增量學習的方法，其本身已經進行了預設 1000 次迭代，所以其訓練模型更加強健。並且在更短的時間成本獲取相當的效果

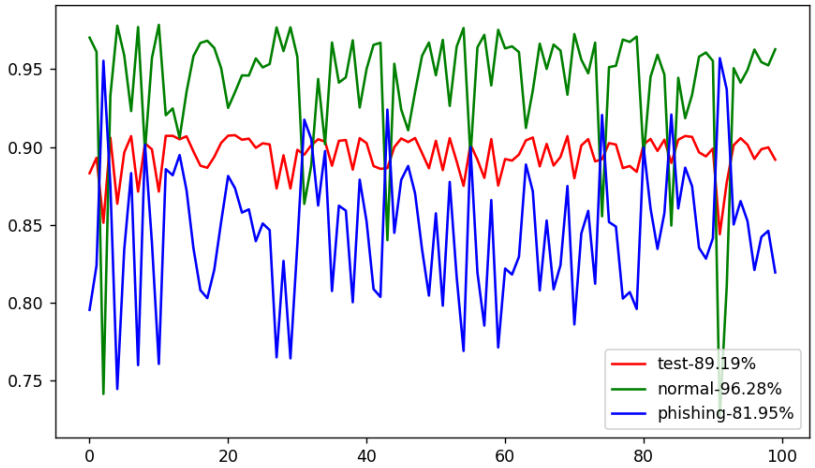
一般深度學習(fit)的 robust、節省時間(一次 fit 相當於 15 倍增量學習的效果)、災難性崩潰的影像降低(降幅變小)。

➢ **實驗方法的優點**

隨機梯度下降法。每次反覆運算都隨機從訓練集中抽取出 1 個樣本，在樣本量極其大的情況下，可能不用抽取出所有樣本，就可以獲得一個損失值在可接受範圍之內的模型了。缺點是**由於單個樣本可能會帶來雜訊，導致並不是每次反覆運算都向著整體最優方向前進。**



圖九、 僅使用增量學習完成 100 組訓練資料的精確度



圖十、 僅使用增量學習完成 100 組每組 15 次的訓練資料的精確度(最後 100 組的結果)

第六節 參考文獻

[1] Kondracki, B., Azad, B. A., Starov, O., & Nikiforakis, N.. (2021). *Catching Transparent Phish: Analyzing and Detecting MITM Phishing Toolkits*. https://doi.org/10.1145/3460120.3484765

[2] [3] Nguyen, L. D., Le, D.-N., & Vinh, L. T.. (2014). *Detecting phishing web pages based on DOM-tree structure and graph matching algorithm*. https://doi.org/10.1145/2676585.2676596

[4] Lomonaco, V., & Rish, I. (2021, July 19). *Continual Learning with Deep Architectures*. Continual Learning with Deep Architectures. https://icml.cc/virtual/2021/tutorial/10833

[5] Aaron, G. (2014, September 25). *Global Phishing Survey: Trends and Domain Name Use in 1H2014*. APWG. https://docs.apwg.org/reports/APWG_GlobalPhishingSurvey_1H2014.pdf

[6] Aaron, G. (2018, July 31). *APWG report*. APWG report. https://docs.apwg.org/reports/apwg_trends_report_q1_2018.pdf

[7] 洪慕藍（2022 年 1 月 17 日）。以機器學習演算法探討網路釣魚網站之特徵值。南臺科技大學。https://hdl.handle.net/11296/ga7s7r

[8] Di Leom, M. (2023, January 16). *phishing-filter*. GitLab. https://gitlab.com/malware-filter/phishing-filter

[9] 曾黎明、黃克仲、陳天豪（2007 年 5 月 8 日）。以 URL 資訊為基礎之網路釣魚偵測系統。TANET2007 臺灣網際網路研討會論文集。http://itech.ntcu.edu.tw/tanet%202007/5/430.pdf

[10] Zhang, Y., Hong, J. I., & Cranor, L. F.. (2007). *Cantina: a content-based approach to detecting phishing web sites*.
https://doi.org/10.1145/1242572.1242659

[11] Moghimi, M., & Varjani, A. Y. (2016). New rule-based phishing detection method. Expert Systems with Applications, 53, 231-242.
https://doi.org/10.1016/j.eswa.2016.01.028

[12] 蔡淑靜（2017 年 1 月 17 日）。基於支援向量機與整合式特徵抽取方法之釣魚網站偵測機制之研究。國立高雄第一科技大學。
https://hdl.handle.net/11296/ed3gke

[13] 羅正漢（2018）。重新認識釣魚郵件威脅。iThome。
https://www.ithome.com.tw/news/120507

[14] Sundari, S. G. (2020, May 11). *Phishing Website Detection by Machine Learning Techniques*. GitHub.
https://github.com/shreyagopal/Phishing-Website-Detection-by-Machine-Learning-Techniques

[15] von Oswald, J., Henning, C., Grewe, B. F., & Sacramento, J. (2022, April 11). *Continual learning with hypernetworks*. arXiv.
https://arxiv.org/abs/1906.00695

[16] Lee, H. Y. (2021, June 5). 【機器學習 2021】機器終身學習(*Life Long Learning, LL*) (一) - 為什麼今日的人工智慧無法成為天網？災難性遺忘 (*Catastrophic Forgetting*). YouTube.
https://www.youtube.com/watch?v=rWF9sg5w6Zk

[17] German I. Parisi , Ronald Kemker , Jose L. Part , Christopher Kanan , Stefan Wermter (May 2019) Continual lifelong learning with neural networks: A review.
https://www.sciencedirect.com/science/article/pii/S0893608019300231

【標題】

逆向工程應用專案經驗報告

【作者】

姓名: 黃啟桓

信箱: ascv0228@gmail.com

【摘要】

本報告描述了我在手機遊戲外掛領域的經驗，包括我參與的專案、遇到的挑戰、解決方案和個人成長。通過這份報告，我將分享我的經驗，並反思在這個領域中所學到的重要教訓。

【報告描述】

我參與了 4 個具有代表性的手機遊戲外掛專案，分別是遊戲 A 和遊戲 B。這些專案包括以下主要活動：

一、 遊戲 A 外掛(Android App)[**個人貢獻: 20%**]

遊戲說明：

遊戲 A 是日本 No.1 人氣動作 RPG 手遊，單手操作、直覺化戰鬥、無體力限制系統的特性。以 Unity-IL2CPP 開發。

契機：

這是我第一次開發遊戲外掛，是由一位前輩教導我的，當時我正好也有玩此遊戲，並且對於外掛的開發有興趣。

目標：

此專案的目標是為了學習外掛的相關技能，外掛的成果是次要的。

技術說明：

此專案利用了 il2cppDumper、Android-Mod-Menu、ApkSignatureKiller、il2cppApi、ida pro 等資源。

il2cppDumper 提供的是從 apk 中 libil2cpp.so、global-metadata.dat 兩個檔案還原成 DLL 檔，並輸出 dump.cs (包含所有類別、屬性、方法的 offset)、以及多個用於其他逆向工程程式的腳本。

Android-Mod-Menu 是在 App 上額外增加一個浮動的 UI 畫面，在 apk 的 AndroidManifest.xml 中改變 App 主要活動的標籤而啟動 UI 畫面。

ApkSignatureKiller 可以在應用層以 hook 函數的方式讓簽名檔檢測返回的數值改變。

il2cppApi 提供以類別、屬性、方法的名字查詢 offset，使得外掛程式碼不用因為遊戲版本更新而重新編寫 offset。

ida pro 一個強大的逆向工程工具，可以將二進制反彙編成組合語言或 C 語言，也能在上面編程。

開發過程中，最困難的是找到需要的函數的地址，這需要花大量的時間來看 ida pro 的結果，很多時候都是無用的。

成果：

此專案算是撿前輩的成品來改進，看前輩的程式碼，以及編寫的風格，對我影響深遠，在此之前我沒有學任何系統性的程式碼教學，程式風格很醜，效率也比較低。而這次的嘗試額外的收穫是使我熟練 C++和 Java，並且更好的了解一個遊戲 App 的運作原理。並且通過這次的經驗，使我面對 Unity-IL2CPP 所開發的遊戲外掛更加容易上手。

連結：

展示: https://youtu.be/CJvxspnVaro

圖一、 在遊戲 A 上啟動浮動窗 UI 用於控制

二、 遊戲 A 外掛(C#、WPF) [**個人貢獻: 30%**]

遊戲說明：

遊戲 A 是日本 No.1 人氣動作 RPG 手遊，單手操作、直覺化戰鬥、無體力限制系統的特性。以 Unity-IL2CPP 開發。

契機：

實作這個專案最大的契機是因為隨著遊戲道具、角色、裝備的增加，手機執行某些動作時需要確認全道具的數量，而「確認全道具、角色、裝

備的數量」這個回覆所消耗的時間已經高達 20 秒以上，對於遊戲體驗很差。由於很多動作會發送確認全道具的數量，這件事是很多情況不在乎的，所以製作此專案。

目標：

這個專案的終極目標是「完成所有手機上能操控的事」，所以花了很多時間去實作很多細微的功能，包含自動通關、裝備合成、裝備穿搭、禮物盒領取...，當時有幾乎完成 8 成的功能，實現了不開手機玩遊戲。

技術說明：

這次外掛是製作成電腦桌面應用，以脫機、修改封包的方式來完成。而抓取封包是使用 charles 進行，但是由於 google 政策，所以無法從 android 7 以上的裝置抓取封包，所以使用模擬機安裝遊戲 A。

這個困難點在於遊戲的封包是加密的，所以需要找到加密以及解密的方式，而封包加密、解密的密鑰來自於使用者 Hash 值，這來源也是加密的，以 AES 加密，透過反彙編的方式，找到並且實作出加密、解密的函數。

成果：

以成果來說，某些動作節省了 98%的時間。但麻煩的是，每次遊戲大版本更新都要重新檢查所有功能的封包內容是否有改動。並且做出相應的調整。當我對於遊戲沒興趣時，就會使我停止更新。

連結：

原始碼:

https://drive.google.com/drive/folders/1WTAzgf8Wzh8k9oJi04neaoEAnxS_WW5x

圖二、 開發桌面應用程式，用於輔助遊戲 A

三、遊戲 B 外掛(Android App) [**個人貢獻: 100%**]

遊戲說明：

此遊戲是由 Cocos Creator 所開發的(cocos2d-js)，策略性攻防戰鬥，冒險。

契機：

為了進行遊戲方面的修改，原先以為該遊戲也是由 Unity-IL2CPP 開發，在深入了解過後，發現是由 Cocos Creator 所開發，對於沒嘗試過的東西，勇於挑戰。最後成為我認為的佳作

目標：

網路上的資源比較少，原本是預期做出一個浮動的 UI 窗口，可以進行外掛功能開關，但是網路上的資源太少了，所以退而求其次，先求有功能再說。目標是可以快速通關。

技術說明：

此專案的困難點在於：要看的懂程式碼中的變數混淆、修改的程式碼要保留或覆蓋。

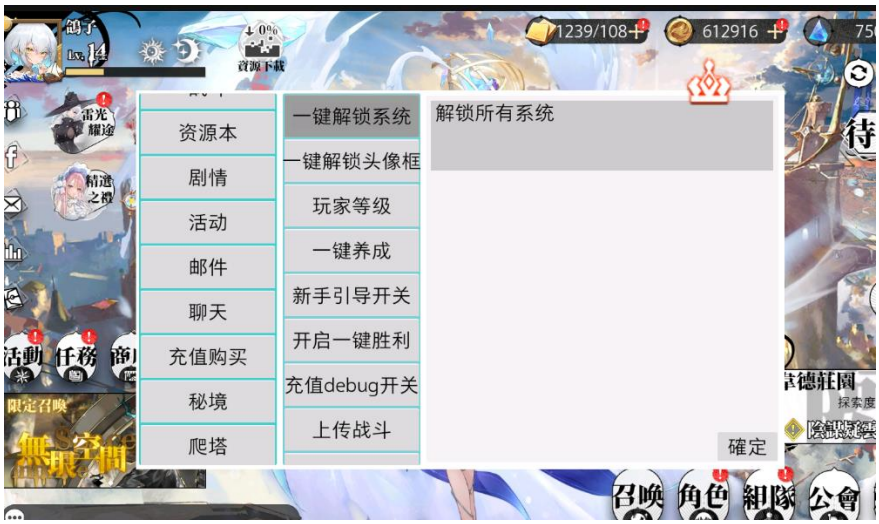
首先原始檔案副檔名是 jsc，要解密需要先知道 key 是多少，就要作逆向分析，但是方法非常簡單，一個能解碼 16 進制的程式讀取 libcocos2djs.so 就能找到 key 是多少。而解密完後的主程式檔是明文，除了變數經過混淆、程式碼經過轉譯、module 沒有分檔之外，其他都沒問題。而 JavaScript 有一個全域變數 console，一般而言，console.log 輸出的訊息會傳送到系統訊息，可以利用 Android Studio 的「Logcat」看到，但是 console.log 在程式碼中被覆蓋掉了，所以第一步要先讓 console.log 不要被覆蓋，才能更好的調適。

而這個遊戲的 Apk 內容總是舊版，在安裝 Apk 時會連接伺服器重新安裝最新版本的，所以在 Apk 中覆蓋主程式檔是無法被保留的，或是需要更麻煩的步驟才能保留。所以我是在 root 的情況下找到原始檔位置，直接覆蓋主程式檔，才能保留。

而連網遊戲都是需要傳送封包，所以我直接修改傳送封包的函數，讓我在戰鬥的時候，戰鬥失敗時也傳送戰鬥成功的封包，並且修改其中的內容，使這個封包盡量偽裝的像正常封包。

成果：

學會了針對於 cocos2d-js 相關的逆向技術，包含加密、解密，在長時間閱讀經過轉譯的程式碼，使我對於反轉譯的技巧加深，能夠更輕易的閱讀，並且通過這次的經驗，使我面對 cocos2d-js 所開發的遊戲外掛更加容易上手。遺憾的是可控制的浮動窗 UI 沒有實現，只能寫死在主程式裡。



圖三、 觸發遊戲開發者工具，可以作為浮動 UI 的參考

四、 遊戲 B 外掛(Discord bot) [**個人貢獻: 100%**]

遊戲說明：

此遊戲是由 Cocos Creator 所開發的(cocos2d-js)，策略性攻防戰鬥，冒險。

契機：

由於該遊戲的聊天內容是發送到手機上，由手機進行快取儲存。當手機重新啟動遊戲，聊天內容就會消失，非常討厭。所以我才使用 Discord 當作儲存裝置，幫我儲存訊息。

目標：

遊戲頻道聊天內容由 Discord bot 傳輸到 Discord 上。

技術說明：

我的 Discord bot 使用 TypeScript 編寫而 cocos2d-js 使用 JavaScript 編寫，所以只要移植程式碼到 Discord bot 上面就可以了，當然還是需要進行修改，把 cocos2d-js 特有的代碼轉換或刪除，然後還要對於每個 module 進行分檔，使程式碼更好看。並註冊一個遊戲新帳號，並以該帳號的 token 進行。

由於遊戲使用 socket 傳送封包，所以在知道 Host 和 Port 的情況下，就能進行傳輸，正式傳輸前使用以「登入的方式」來進行 Server 的分配，訊息打包經過 RSA 加密，但是在知道程式的明文的情況下，RSA 的公鑰是直接可以找到的。另外接收訊息的解密只需要使用一般的 msgpack-lite 就能完成了。

成果：

比較遺憾的是，遊戲內聊天頻道可以發送戰鬥影片(由代碼生成動畫)，但我無法生成動畫並轉成 mp4，所以我就忽略影片，雖然戰鬥影片也是

很寶貴的資源。

圖四、 Discord Bot 接收到遊戲聊天訊息轉發到 Discord 上

【個人成長】

技術技能提高：參與外掛開發使我能夠深入研究遊戲和安全技術，提高了我的技術技能。

道德和法律意識：我更加了解了道德和法律問題，並學會在這個領域中保持合法和道德的行為。

團隊合作：與團隊合作是成功的關鍵，我學會了如何有效地與不同專業的人合作。

【結論】

手機遊戲外掛領域是一個充滿挑戰和機會的領域，我在這個領域中獲得了寶貴的經驗。儘管我們面臨了許多挑戰，但這些挑戰使我更堅韌、更具技術能力，並加強了我的道德和法律意識。這個經驗對我個人和專業的成長都具有重要意義。

艾法科技股份有限公司離職證明書

編號:11209111

姓名	黃啟桓										出生年月日	民國 90 年 03 月 22 日
身分證字號	G	1	2	2	4	2	8	1	6	1	性別	■男 □女
服務部門	工程部										職稱	韌體工程師
擔任工作內容	程式撰寫與測試。											
任職日期	自民國 112 年 07 月 04 日起至 112 年 08 月 11 日止											
服務年資	服務期間共計滿 零 年 1.3 月											
在職情形	現已離職											
注意事項	服務單位亦可提供自訂在職證明書之格式,但務必包含本證明書之全部內容。											
備註	自願離職。											

證明機構(全銜):艾法科技股份有限公司

負責人:倪文祿

機構地址:高雄市鳳山區八德路二段230號1樓

電話:07-7777128

營利事業統一編號:84217382

中華民國一一二年九月十一日

(以上資料,如有不實證明,願負法律責任)

姓名	黃啟桓
公司	艾法科技股份有限公司
統編	84217382
公司所在地	高雄市鳳山區文德里八德路二段 230 號
負責人	倪文祿
服務部門	工程部
職稱	韌體工程師

【簡介】

此報告詳述本人黃啟桓工作經驗

【求職契機】

我是一位有自主想法的人，在我大二時我就計畫大學畢業就直接去工作，而在大四暑假時，考慮到我下學期每週只有兩堂課程，所以我毅然決然的在畢業前就求職。求職期間遇到很多問題包含「限碩士學歷以上應徵」、「無法接受工讀生」...，投遞多封求職信，只有一成的公司願意面試。為我的心中埋下一顆攻讀研究所的種子。

【面試過程】

在初次面試時，該公司希望我能做出一個使用 Flutter 開發的 App 使其進行 IOT 的功能。我在 6 月 20 號左右開始製作此小專案，此專案功能是使用「App 更改電腦桌面畫面」。我首先想好並拆分各個小部件，一一攻克，在進行組合。把重點放在 App 的製作，並寫了一個簡單的 Server 進行通訊，最後花費 12 天左右完成，最麻煩的部分是通訊的內容「要傳多少資料，才能滿足所有功能的需求」這部分改了很久。

在最後面試時，我的效率令老闆刮目相看，最後成功錄取軟體工程師的職位，並同意我開學偶爾到校上課。

【工作描述】

當時我以為我是要做 Flutter 開發手機的部分，但是因為我有 C 語言經驗所以老闆認為我有實力做韌體開發。因為我是第一次接觸韌體開發，完全不知道怎麼辦，後來經過前輩的指導，看得懂技術文件、安裝所需環境、了解如何編譯。那時我是作為協助的角色，處理 RTL8722DM、RTL8720CF 芯片的藍芽功能開發，以及解決前輩某些無法處理的問題。

當時藍芽功能的開發延宕了幾個月，經過和 Ameba 開發商的溝通得以在 15 天內完成藍芽功能開發。其後也解決的包含 mac number 重複問題(經了解是

前輩呼叫錯誤的函數當作 mac number 使用)，然後以 esp32c3 開發 homekit 的前端，因為我近期規劃要離職並攻讀研究所，沒有了解 iCtrl Pro 產品的程式碼，所以沒有進一步完善功能。

【個人成長】

技術提高：非常感謝公司的各位員工的幫助，在接觸新領域的時候，可以快速掌握相關技能，並且提供了不少好用的應用程式進行產品測試、分析有。

團隊合作：在職期間學習到溝通技巧、任務分配、問題解決，對於職業發展非常有價值。通過積極參與團隊合作，你可以提高自己的專業能力。

【結論】

在這份報告中，我分享了自己的工作經驗，從求職契機到工作描述，再到個人成長的方面都有所涵蓋。這段工作經驗對我來說是一次寶貴的學習機會，有幾個主要的收穫和體會：

我的工作描述涉及了不同的技術領域，從手機應用到韌體開發。雖然我一開始對韌體開發一無所知，但通過前輩的指導和學習，我成功地應對了挑戰，並解決了一系列技術問題。這個過程中，我學到了技術文件閱讀、編譯、解決問題的能力，以及如何協作處理複雜的工作。我在這份工作中不僅提高了技術能力，還學到了團隊合作的重要性。我學會了與不同背景和技能的人合作，並有效地溝通、分配任務和解決問題。這將對我的職業發展產生深遠的影響。