

智慧化企業整合 Final Project

利用 Robotic Process Automation + CRNN 進行驗證碼辨
識並自動登入流程

馮品叡

指導教授: 邱銘傳

利用 Robotic Process Automation + CRNN 進行驗證碼辨識並自動

登入流程

清華大學工業工程與工程管理學系 馮品叡

一、摘要

現今業務面臨的問題主要與高數據輸入量、高錯誤率、多個未集成的舊系統以及高重複/低增值等活動。Robotic Process Automation(RPA)恰可解決此一問題，RPA 的一個優點是，用戶可以對其進行培訓，以執行結構化的、可重複的計算機任務，同時與多個系統進行交互(非侵入式系統)，同時根據算法執行複雜的決策。通過這種方式，機器人可以進行重複性流程，以更快速的方式來縮短時間，也可以 24/7 全天候運行。若結合 Artificial Intelligence (AI)可將 RPA 使用性更為廣泛，該計畫使用 UiPath Studio Community 版本並結合 Colab 來進行驗證碼辨識並自動登入流程。

關鍵字: Robotic Process Automation、UiPath、AI

二、介紹

PwC 2018 年報告指出，RPA 指的是針對那些花人力時間處理、跨系統作業、重複性的工作，運用機器流程自動化來解決這些問題。根據 Forrester 在 2017 年排名 UiPath 跟其他公司相比更是名列第一(如圖一)，所以本研究使用 UiPath Studio Community 版本來做測試。然而 RPA 可處理的工作眾多，若該流程符合:複製黏貼、網路抓取資訊、模仿人類行為、高度手動和重複的過程、基於規則的流程、成熟穩定的流程等，均可透過 RPA 來解決。

RPA 具有以下優勢:降低成本提升品質、非侵入式軟件(可與舊系統相容)、全天性工作、穩定度高、效率好等優勢。其應用的範圍也相當廣泛，目前使用的公司有:阿里雲、日本財產保險公司(Sompo Japan Insurance)、富士房屋株式會社等，且在成效上均有顯著的優化。在 Grand View Research 進行的一項研究中“ RPA 的市場目前價值 11 億美元，預計從 2020 年到 2027 年的複合年增長率為 33.6%。”因此 RPA 的未來十分可期。

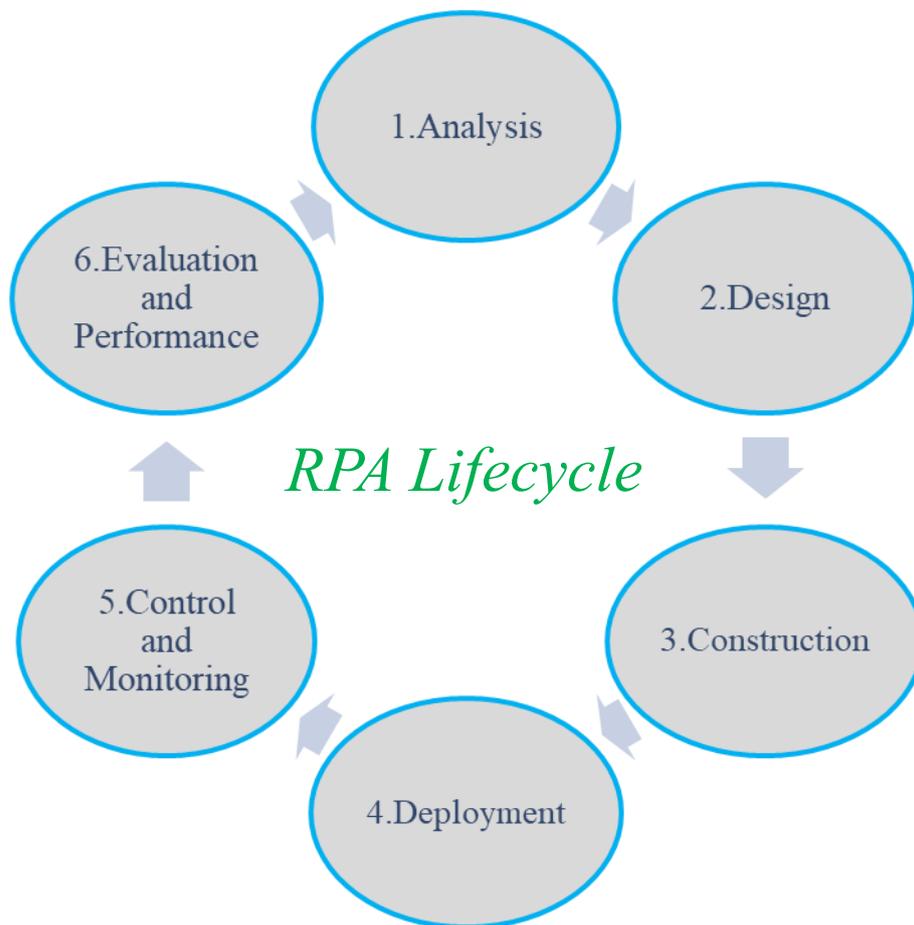
Technology Category	Uiipath	Automation Anywhere	NICE	EdgeVerve Systems	Blue Prism	Kryon Systems	WorkFusion	Pega Systems	Redwood Software	Contexor	Kofax	Softomotive
Bot development and core functions	3.25	3.70	3.60	2.85	2.50	3.45	2.30	3.25	2.35	3.10	1.65	2.85
Control room, system mgmt, reporting & resilience	3.80	2.80	3.25	3.25	3.80	3.25	3.45	2.45	3.50	2.25	2.75	2.85
RPA analytics	3.66	3.66	2.33	3.34	2.00	2.33	2.85	3.00	1.34	1.66	2.00	2.00
Architecture	3.99	4.33	4.33	3.67	3.66	3.66	2.99	3.34	2.67	3.00	2.67	1.99
Deployment, governance, and security	3.66	3.66	3.99	3.00	4.00	2.99	3.68	2.98	3.00	2.33	2.67	1.66
Total RPA Technology Score	3.67	3.63	3.50	3.22	3.19	3.14	3.05	3.00	2.57	2.47	2.35	2.27

圖一、RPA vendors ranked by technology capabilities

Source: <https://go.forrester.com/>.

三、方法

本研究首先考慮 RPA Lifecycle(如圖二)步驟如下: Step1 考慮並分析自動登入流程可行性及特性; Step2 詳細說明 RPA 流程中必須執行的操作動作和執行活動等; Step3 建構流程構成要素; Step4 給 robot 一個環境執行工作; Step5 控制和監視每個機器人的性能; Step6 評估機器人的表現。



圖二、RPA Lifecycle

本研究採用 CRNN 進行驗證碼辨識，資料來源為 GitHub(https://github.com/YenLinWu/RPA_UiPath/tree/master/OpenCV%20x%20Tesseract%20x%20UiPath)，CRNN 可解決驗證碼動態變化且破壞 Rule-Based 的問題。

本研究也使用 RPA 來進行抓取 Test Data 的流程，抓取 100 張驗證碼圖形只需 3-4 分鐘，若此動作由人工操作，可能須進行約 20-30 分鐘不等，由此可知 RPA 在處理基於規則的問題可節省大量的時間成本。

本研究在 UiPath 中使用以下 Activities: Use Application/Browser、Take screenshot、Save Image、Click、Get Text、Type Into、Do while、Message box...等。

四、研究結果

本研究融合 RPA+AI 可達到以下兩點效益:1. RPA 作為 AI 模型前的資料蒐集與處理達到非常高的效率，也替使用者大幅節省時間；2. 此專案目的為達成帳號登入，RPA 在執行重複性動作更是專精，因此可解決 AI 模型追求高準確率的問題，下述算式表示準確率在 0.85 與 0.99 在執行五次後完成登入，其成功率差異並不大，算式假設如下：

$$1-(1-0.99)^5=0.9999999999$$

$$1-(1-0.85)^5=0.9999240625$$

五、結論

本研究因驗證碼格式隨時變動，相較固定且規則化問題，需增加辨識的功能，若未來實務上格式規則化的狀況下則更為容易，而本研究在準確率上達 96.78%，在速度方面也節省約 8 成的時間。未來方向有以下三點:1. 組織將機器人技術和先進的數字技術結合在一起，平衡了人力和機器人的勞動力，並實現了智能自動化。2. 結合 AI 人工智能，以便 RPA 機器人能夠了解問題何時發生並做出決定以解決問題。3. Analytics (分析) 是一種被人類使用的工具，但最終將被機器人等所取代。

參考文獻

1. UiPath Forum

<https://forum.uipath.com/>

2. RPA 中國

<http://www.rpa-cn.com/>

3. UiPath

<https://www.uipath.com/>

4. Forrester

<https://go.forrester.com/>

5. Robotic Automation Process – The next major revolution in terms of back office operations improvement Sorin ANAGNOSTE The Bucharest University of Economic Studies, Bucharest, Romania
sorin.anagnoste@fabiz.ase.ro

6. RPA Challenge

<http://www.rpachallenge.com/>