Regular Paper

Communications of the CCISA
Vol. 30 No. 1 Feb. 2024

車載網路資安實驗平臺之開發與模擬攻擊檢測

王平[1*]、陳佳鴻[2]
崑山科技大學資訊管理學系
[1] pingwang@mail.ksu.edu.tw、[2] s106000750@g.ksu.edu.tw

摘要

隨著汽車技術的快速發展，汽車提供多元聯網需求大增，也引發了網路入侵等安全與隱私問題。目前車載網路(In-Vehicle Network, IVN)系統導入控制器局域網(Controller Area Network, CAN)匯流排系統提供多個電子控制單元(Electronic Control Units, ECUs)之間的通信，但 ECUs 可能存在資安漏洞，CAN 系統攻擊方式與傳統網際網路通信方式不同，傳統入侵檢測系統(Intrusion Detection System, IDS)將不易辨認車載網路威脅並作有效防禦。

本研究聚焦於運用深度學習為基礎的入侵檢測模式的發展，車載網路攻擊參考韓國 Hacking and Countermeasure Research Lab(HCRL)實驗室 2021 年開放資料集，進行竊聽攻擊(Sniffer)、汽車阻斷服務攻擊(Denial of Service, DoS)、模糊攻擊(Fuzzy)、欺騙駕駛設施(Spoofing Gear)和欺騙轉速表(RPM)和溫度攻擊(Temperature)等。執行車載網路資安實驗，發現國內缺乏一個實體環境以產出攻擊資訊以進行入侵檢測分析及性能驗證，因此本研究設計一個 CAN 資安實驗平臺，運用實體車載網路硬體及 Qt 模擬軟體工具，產出六項模擬攻擊資訊，搭配 VGG-16 卷積神經網路進行攻擊行為特徵學習，並通過實驗數據對模擬攻擊訊息進行分類。其中，VGG-16 作為一種高效的 CNN 架構，已在多個分類任務中表現出色。實驗結果證明，VGG-16 檢測模式對二分類與六分類測試準確性分別為 100%和 99%，CAN 攻擊偵測模式可以透過學習攻擊行為特徵，協助防禦者在車載網路發現相關威脅，以利制定防護措施。

關鍵字：深度學習、車載網路、入侵檢測、VGG-16、卷積神經網路

* 通訊作者 (Corresponding author.)

Regular Paper

Communications of the CCISA
Vol. 30 No. 1 Feb. 2024

Development of a Security Experimental Platform and Simulated Attack Detection for In-Vehicle Networks

Ping Wang[1*], Jia-Hong Chen[2],
Department of Information Management, Kun Shan University, Tainan, Taiwan
[1] pingwang@mail.ksu.edu.tw、[2] s106000750@g.ksu.edu.tw

Abstract

With the rapid advancement of automotive technology, the demand for diverse connectivity in vehicles has increased significantly, leading to security and privacy concerns such as network intrusions. In-Vehicle Networks (IVN) currently employ the Controller Area Network (CAN) bus system to facilitate communication between various Electronic Control Units (ECUs). However, ECUs may have security vulnerabilities, and the attack methods on CAN systems differ from traditional internet communication. Conventional Intrusion Detection Systems (IDS) may struggle to identify and defend against threats in in-vehicle networks effectively. This research focuses on the development of intrusion detection models based on deep learning. The study uses the 2021 dataset from the Korean Hacking and Countermeasure Research Lab (HCRL) for in-vehicle network attacks, including eavesdropping (Sniffer), Denial of Service (DoS), Fuzzy attacks, Spoofing Gear, RPM, and Temperature attacks. To conduct in-vehicle network security experiments, the study addresses the lack of a practical environment for generating attack message and conducting intrusion detection analysis and verification in the domestic context. Thus, the research utilizes physical in-vehicle network hardware to establish a CAN security experimental platform, which generates six real attack scenarios. These scenarios are then used for attack behavior feature learning and recognition using the VGG-16 convolutional neural network where VGG-16, as an efficient CNN architecture, has demonstrated excellent performance in multiple classification missions. The experimental results demonstrate that the VGG-16 intrusion detection model achieves 100% accuracy in binary classification and 99% accuracy in six-class classification tests. The CAN intrusion detection model, with the help of behavior feature learning, assists defenders in identifying relevant threats in in-vehicle networks, facilitating the development of protective measures.

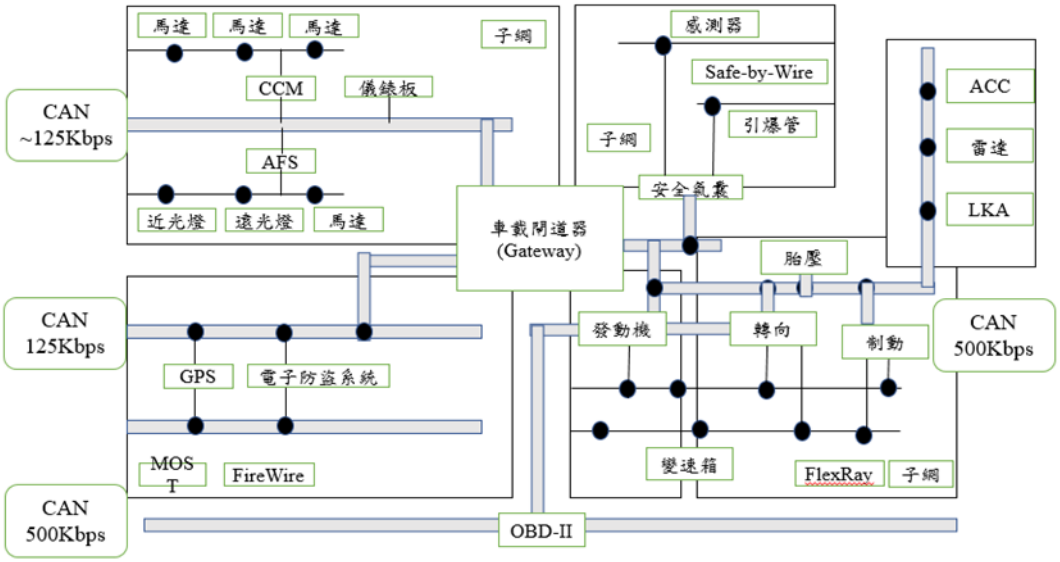
Keywords: Deep Learning, In-Vehicle Network, Intrusion Detection, VGG-16, Convolutional Neural Network

壹、緒論

1.1 研究背景

近年來，隨著汽車銷售量持續增加，車載網路（In-Vehicle Network, IVN）領域正在快速發展。然而，汽車本身對車載網路安全的考慮並不充分，並且面臨著許多挑戰。汽車的計算能力有限，同時在面對複雜的環境下，汽車內部部署了大量的電子控制單元（Electronic Control Units, ECUs），並使用各種外部通訊協定，如藍芽、GPS 和 3G/4G 等。根據 Strategy Analytics 的報告，過去十年內，汽車內部 ECUs 的數量快速增長，現在平均每輛汽車擁有數十個 ECUs，而高級汽車可能達到上百個，其程式碼行數達到數億行。(Charette, 2009)這種情況增加了網路漏洞和潛在攻擊的風險。

汽車的開放性和車載網路的複雜性使得汽車面臨了嚴重的安全威脅。其中，以德國博世公司（Bosch）開發的控制器局域網（Controller Area Network, CAN）為車載網路的主要代表，其安全性受到質疑。特別值得關注的是在 2013 年，兩位安全專家 Charlie Miller 和 Chris Valasek(2013) [2]成功實施了對 Jeep 汽車的 CAN 攻擊，控制了方向盤、制動和儀表等功能。這一事件突顯了 CAN 攻擊在 IVN 中的重要性，因為 CAN 控制著汽車內部的重要部件，如引擎、變速箱和儀表。一旦受到攻擊，將對駕駛和乘客造成嚴重後果。如圖一所示。



圖一：CAN 車載網路系統連接多個控制單元

資料來源: (Charette, 2009) [3]

然而，CAN 車在網路的設計初衷主要是為了實現汽車內部的即時通信。當 CAN 協議最初被引入時，汽車被視為封閉系統，且車輛的電子化水準相對較低。然而，隨著現代汽車的電子化程度不斷提升，CAN 協議的原始設計已經無法應對當前的安全需求。

以下是 CAN 協議存在的安全問題：

1. 訊息無加密： CAN 協議雖然能夠確保通信的正常進行，但它缺乏適當的安全和隱私保護機制。這意味著攻擊者可以輕易竊聽 CAN 訊息，從中獲取敏感的用戶數據，造成嚴重的隱私侵犯。儘管有些車廠提供通訊矩陣來實現保密性，但這種機制並不是足夠安全的，容易被攻破，無法滿足現代的保密標準。
2. 訊息無認證：任何連接到 CAN 網絡的設備都可以對總線執行讀寫操作，而 CAN 協議對這種行為幾乎沒有規定。它缺乏身份驗證或訪問控制的機制。此外，CAN 訊息的標識符中並不包含源位址和目的地址，這意味著無法對節點的合法性進行驗證，並且允許惡意節點對其他節點進行攻擊。
3. 無法確認訊息的真實性： CAN 協議無法區分真實的錯誤訊息與攻擊者製造的錯誤訊息。這意味著無法確定汽車中的特定設備（例如車載娛樂系統）是否處於故障狀態，還是受到了攻擊，進而導致該設備被迫關閉。

1.2 研究動機

國外學者和專家在車載網路安全研究方面起步較早，特別是對 CAN 架構的安全性研究，美國、德國和日本等傳統工業強國在此領域具有深厚的理論和實踐基礎。為了應對車載網路的資安問題,這些專家提出了多種措施,包括通訊訊息加密及訊息認證技術、防火牆隔離方法以及網路入侵檢測系統（Intrusion Detection System, IDS）的開發和部署。

1. 通過訊息加密和訊息認證技術來確保其私密性和完整性[4] (Zhou et al., 2018)
2. 使用防火牆方式將車載網路與攻擊介面隔離[5] (Zhou et al., 2018)
3. 為車載網路開發和部署入侵檢測系統(IDS)。

總結來說,這些方法各有優缺點,但車載入侵檢測系統 (IDS for In-Vehicle Network)為一種有效應對車載網路安全挑戰的方式，它能夠提供即時的保護和有效的威脅檢測，同時在成本上也相對較為實際。這是重要的動機來推進相關研究。

基本上,入侵偵測系統傳統上分為兩種形式：異常偵測(anomaly detection)與濫用偵測(misuse detection)。

1. 異常偵測

Larson et al. (2008) [6]提出了一種基於規範的攻擊異常檢測方法，先定義正常行為，如果訊息未遵循協議級別或 ECUs 行為規範，則將其視為攻擊。 Muter and Asaj (2011) [7]提出了一個基於熵的異常檢測方法，定義了 CAN 訊息熵，通過將該測量得到熵值與之前設定的參考值進行比對來檢測出攻擊。Lee et al. (2017) [8]提出利用 CAN 協議的遠程框架來分析遠程框架的響應時間和補償率，測量這些指標在攻擊狀態與正常狀態之間的偏差來描述 ECUs 的行為，並提出了對應的數據集。但是，這些入侵檢測方法僅對特定的威脅模型或在特定前提下有效，例如前提是 CAN 消息具有週期性，威脅模

型是攻擊者大量注入具有特定 CAN 訊息 ID 的消息以及違反 CAN 協議規範。過去的車載入侵偵測方法包括基於規範的攻擊檢測、基於熵的異常檢測，以及基於 CAN 協議的遠程框架分析。然而，這些方法對特定的威脅模型或前提下有效，且存在局限性。

2. 濫用偵測

濫用偵測主要特點是基於已知的攻擊特徵或簽名來識別惡意行為。這種方法假定攻擊可以通過比較網路流量與已知攻擊模式來識別。濫用偵測三種常用方法為特徵為基礎偵測(signature-based detection)，其依賴於一個廣泛的攻擊特徵資料庫，特徵集合定義了已知的惡意活動模式，當網路連線之內容與資料庫中的某個特徵匹配時，系統會發出警報。防毒軟體經常使用特徵基礎方法來識別和隔離已知的病毒和木馬。這些軟體會定期更新其病毒定義數據庫，以包括新發現的惡意軟體簽名。然而，對於檢測先前未知的攻擊和零日漏洞則有一定的局限性。

為提高車載網路入侵分類的精確性，學者們探索了機器學習（Machine Learning, ML）和深度學習（Deep Learning, DL）技術於濫用偵測的應用。ML 技術通過訓練模型學習正常和不同類別攻擊的行為模式，以識別多種網路攻擊類別。例如，Taylor et al. (2016) [9]使用長短期記憶（LSTM）網路模型，但其模型參數過多，增加了計算成本。Seo et al. (2018)[10]首次引入生成對抗網路（GAN）到車載 IDS 領域，但其方法過於複雜不利於即時應用。Song et al. (2020) [11] 提出了基於卷積神經網路（Convolutional Neural Networks, CNN）的入侵檢測方法，其將一個 CAN ID 訊息編碼成 29 位的二進制向量，然後再將連續的 29 個 CAN ID 組成一個序列，形成 29 × 29 的影像矩陣，形成二進制特徵圖。執行數據標註時，在序列中含有異常 ID 即標註該特徵圖為異常樣本。最後使用簡化的 Inception-ResNet 模型(Szegedy *et al.* 2016)[12]對這些特徵圖進行分類，但存在檢測延遲和架構改進的需求。

1.3 研究目的

本研究的主要目的是建立一個 CAN 車載網路的資安實驗平臺，參照 Lee et al. (2017) [8] 實驗設計的訊息格式，生成五項模擬的攻擊資訊，這些實驗資訊用於後續 VGG-16 卷積神經網路[13] 之偵測模式的學習和識別。VGG-16 作為一種高效的 CNN 架構，其網路架構包括其多個卷積層和全連接層的設計，以及這種設計有助於捕捉和學習圖像的複雜特徵，並已在多個分類任務中表現出色。

以下是本研究的具體目的：

1. 使用 Qt 模擬工具開發 CAN 資安實驗平臺，針對 CAN 協議的各種攻擊和正常情況進行模擬。通過使用真實車用儀表系統，搭建一個儀表平臺，再利用恩智浦（NXP）的嵌入式電路板和 Arduino 等硬體，在此基礎上建立一個可收發訊息的 CAN 實驗平臺。這個平臺通過收發器讀取 CAN ID 和資料區，實現各種 CAN 攻擊訊息的產生。
2. 提出使用 VGG-16 深度學習模型來進行車載網路的濫用偵測。為實現即時反應速度，

將對卷積神經網路（CNN）模型進行精簡，以最大限度地提高準確性並減少延遲。
3. 驗證入侵檢測模型在車載網路安全性方面的性能，特別是在威脅分類問題上。在實驗中對測試用例的異常檢測能力進行評估，包括分類的準確性、精確率 (Precision)、召回率(Recall)和 F1 值(F1-score)等指標。

1.3 研究限制

針對 CAN 通信協議的安全問題，研究學者陸續提出了多種應對惡意攻擊的方法。相對於加密、認證以及隔離等方法，入侵偵測系統作為增強網路安全性的方式，但是車載入侵偵測系統的設計和實施面臨各種挑戰和限制，主要體現在以下方面：
1. 硬體限制：主要採用 32 位嵌入式處理器，如飛思卡爾（NXP_Freescale）和瑞薩電子（Renesas）等。這些處理器的計算性能和存儲資源有限，因此入侵偵測系統的設計受到計算能力和存儲容量的限制。
2. 成本限制：考慮到汽車是大規模生產的產品，為了降低硬體成本和提高穩定性，才能為製造商帶來利潤。然而，為車載網路引入安全措施通常需要修改所有電子控制單元（ECUs）的硬體，這會顯著增加製造商的開發成本，並可能面臨推廣和採用的問題。因此，成本限制也是入侵偵測系統設計必須考慮的重要因素之一。

1.4 論文章節架構

本研究的流程分為兩個主要階段，每個階段的流程如下所述。第一階段是在確定研究主題和目標後，制定研究計劃，整理相關文獻，綜合資料後，提出研究架構模型並確定所需的開發軟硬體和工具。第二階段基於先前提出的研究架構模型進行模式設計和開發，包括以下工作：(1) 建立 CAN 資安實驗平臺；(2) 開發 VGG-16 分類器；首先，對 CAN 資安實驗平臺生成的數據進行預處理，然後進行模型的訓練和優化，接著進行研究模型的測試驗證。最後，根據分析結果進行模型性能比較分析，並撰寫研究成果報告。

貳、文獻回顧

目前，自駕車技術正快速發展，但與之同步發展的是「汽車駭客」的風險。汽車控制器局域網（CAN）是建立在資訊導向傳輸協議的基礎上，它使用了一種廣播機制，通過唯一的訊息標誌（Message Identifier）來識別內容並定義訊息的優先順序，以實現資訊傳遞。然而，CAN 協議在與外部通信時並未提供使用者身份驗證或訊息加密功能，這使得汽車在與外界進行資訊交流時存在安全性漏洞，這些漏洞可能被惡意攻擊者利用，

危害駕駛安全。因此，如何評估並保護 CAN 的網路漏洞成為了一個重要的安全問題。

為應對這一問題，於 2019 年 7 月，由十一家汽車專業公司組成的聯盟發布了一份名為「自動駕駛的安全第一安全白皮書」（Safety First for Automated Driving, SaFAD）的檔，旨在提供開發、測試和驗證自駕車安全性的框架。這份白皮書成為資訊安全專家和相關企業在研究車載網路（IVN）領域時的重要參考文獻。參與編撰這份白皮書的公司包括奧迪、Aptiv、百度、寶馬、戴姆勒、克萊斯勒汽車、Here、英飛凌、英特爾、大眾等。此外，以下整理了 2013 年至 2021 年期間引人注目的車載網路入侵事件，詳情請參見表一。

表一：汽車入侵事件大事紀

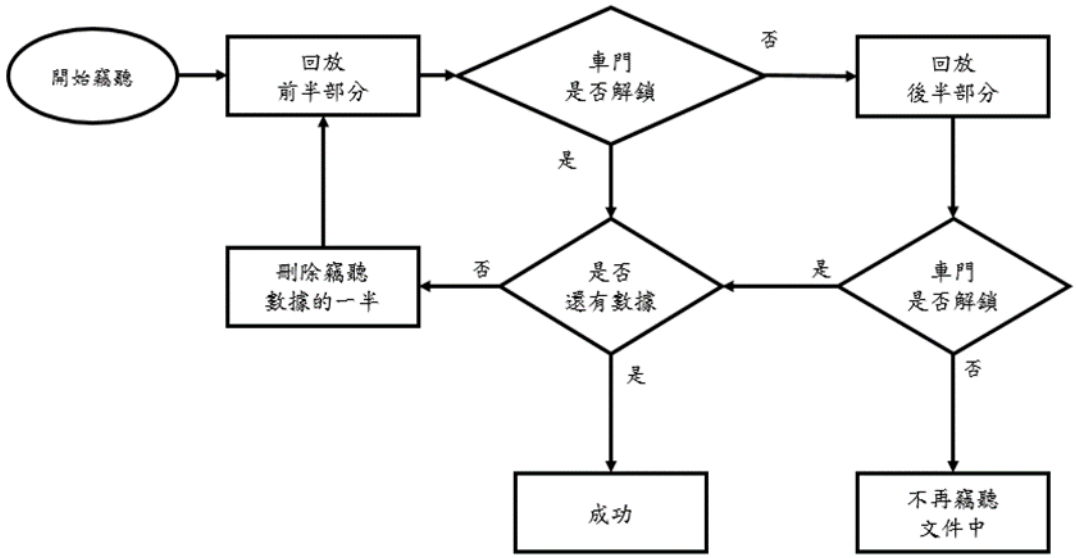
作者	攻擊目標	研究發現	攻擊方式
Miller and Valasek (2015) [14]	Jeep Cherokee	UConnectm 娛樂系統 Wi-Fi 漏洞入侵，與 D-Bus 服務溝通，篡改瑞薩控制晶片韌體，掌控汽車控制權。	透過車上 Uconnect 娛樂系統漏洞，使用者可以操作資訊娛樂/導航、電話功能，甚至還能設定車上 Wi-Fi 熱點，最後車廠全部召回 140 萬台 Jeep。
Kamka (2015) [15]	Generic Motor	通用汽車 OnStar 系統漏洞，讓使用者在 Android 及 iOS 手機上透過取得 RemoteLink 提供的權限，連接並操作車輛功能。	RemoteLink App 漏洞，可以用手機來尋找汽車位置、解鎖車門，甚至發動車輛。
Hunt (2016) [16]	Nissan Leaf	透過 Nissan Connect 電動車，由車輛辨識號(Vehicle Identification Number, VIN)辨識，而 VIN 數據位於 URL，會向 Nissan 發出請求，來控制車輛。	Nissan Connect EV App 漏洞，可以打開氣候控制，查看電池壽命或是通過運行加熱或冷卻系統來使駕駛員擱淺，從而遠程耗盡電池。
騰訊科恩實驗室 (2017) [17]	Telsa Model X	通過 Wi-Fi 晶片與行動網路連接情況下實現了對車載網路系統的破解，遠程式控制制剎車、車門、後備箱，操縱車燈以及廣播。	通過對汽車商晶片逆向工程，來進行遠程攻擊。 1.Wi-Fi 攻擊。 2.行動網路攻擊。 3.攻擊無線韌體更新代碼簽名。

比利時魯汶大學 Wouters et al. (2021) [18]	Telsa Model X	1.發現特斯拉 Model X 的無鑰匙(keyless)系統存在重大安全性漏洞。 2.研究人員開發的駭客攻擊套件，可以騙取密鑰卡中新改進的加密技術。	攻擊只需要 90 秒，利用特斯拉 keyless 系統漏洞，擋風玻璃讀取到汽車儀表板上的車輛識別號碼，然後為盜版車身控制器建立代碼，在駕駛要使附近提取射頻密碼，成功將車輛解鎖。

本研究針對 Song 等人（2020 年）[11] 的研究，對 CAN 協議的正常運作和異常情況進行了詳細的數據分析。研究過程中，從真實的車載儀表系統中錄製了數據，以識別相應的 CAN 消息內容。此外，我們對儀錶平臺進行了五種典型攻擊，並分析了這些攻擊對 CAN 系統運作的影響。這五種攻擊方式包括竊聽攻擊、汽車阻斷服務攻擊、模糊攻擊、溫度攻擊和欺騙設備攻擊。

（一）竊聽攻擊(Sniffer)

這是最複雜的一種攻擊方式。在這種攻擊中，我們對 CAN 流量數據進行分析，尤其關注仲裁 ID 內的資料區。這個步驟相當繁瑣，因為一旦駭客成功鎖定了仲裁 ID 並破解了資料區，他們就能針對 CAN 系統發動有效且具有致命性的攻擊，這種攻擊可能對車輛造成極大危害。如圖二所示。



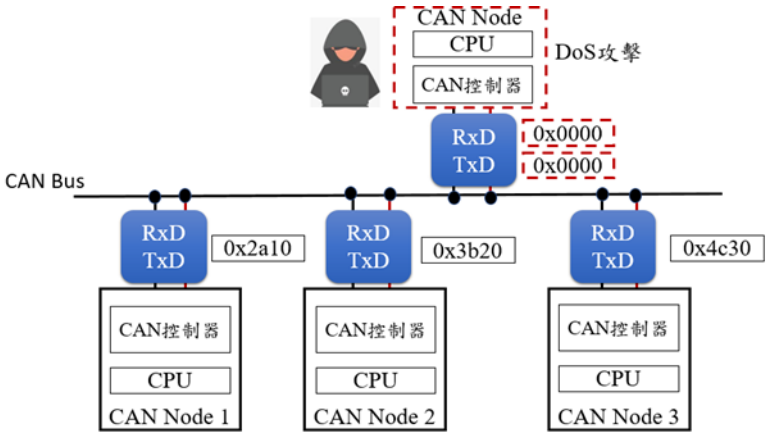
圖二：Sniffer 攻擊示意

資料來源：(Craig & Chris 2017) [19]

（二）汽車阻斷服務攻擊（DoS）

在此攻擊中，攻擊者利用某些 ECUs 系統漏洞連接到 CAN Bus，然後發起 DoS（Denial of Service）氾洪攻擊。在攻擊過程中，攻擊者在非常短的時間內（例如 0.3 毫秒）向 Bus

中注入最高優先級的訊息（標誌符 0x000）。由於所有節點共用相同的 Bus，惡意節點不遵守仲裁協議規則，突破了未經授權的訪問權，這將導致優先權較低的訊息發送延遲，佔據 Bus 並阻礙正常訊息的發送，導致交通堵塞，進而導致車輛無法及時響應駕駛員的命令。這種攻擊方式非常簡單，容易被檢測。因此，我們將其歸類為低級別的檢測難度，如圖三所示。



圖三：DoS 攻擊示意

DoS 攻擊可能導致的安全事件表，如表二所示。

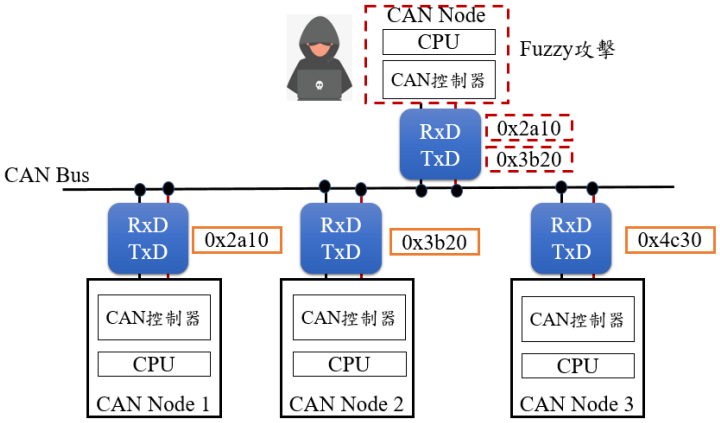
表二：常見 DoS 攻擊手法

DoS 攻擊	攻擊手法
先進駕駛輔助系統(ADAS)攻擊	先進駕駛輔助系統是當駕駛發生緊急狀況，主動介入相對應的動作，當被 DoS 氾洪攻擊時，會佔用 Bus 資源，引發故障，導致 ADAS 系統無法做動。
節氣門攻擊	利用 DoS 洪氾攻擊阻止駕駛控制節氣門，控制車輛行駛，阻止車輛行駛。
汽車門鎖攻擊	主要通過其他設備或是車載診斷系統(OBD-II)做為跳板，只需按車門鎖定/解鎖，就能對應一組固定標示符，以氾洪反覆式對車門鎖定與解鎖，幹擾駕駛者。

資料來源：陳力(2020) [20]

（三）模糊攻擊（Fuzzy）

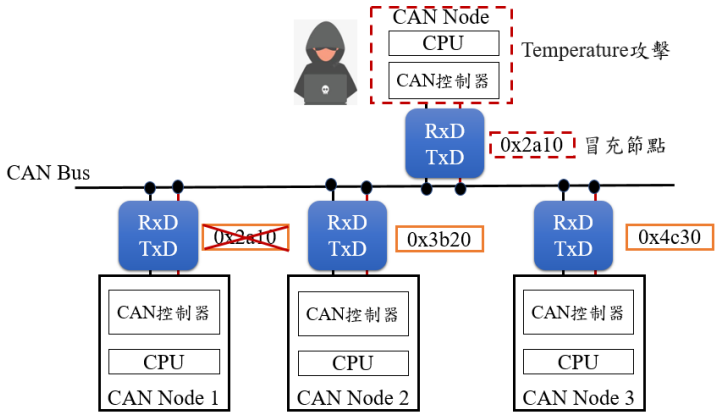
通常，Fuzzy 攻擊用於瞭解 ECUs 如何對特定數據類型作出反應。在 Fuzzy 攻擊中，攻擊者隨機模擬其他 ECUs，發送包括隨機仲裁 ID 和片段訊息的錯誤訊息。這將導致網路節點接收大量毫無意義的訊息，進而使特定車輛 ECUs 的功能受損。當受到 Fuzzy 攻擊時，攻擊者將錯誤訊息插入欺騙性指令中，這可能導致 ECUs 嚴重故障，例如改變換檔行為、方向盤劇烈擺動、不規則閃爍的方向燈，或者儀表板上的故障指示燈閃爍。與 DoS 攻擊不同，Fuzzy 攻擊僅導致特定車輛 ECUs 的功能受損，而不會佔用匯流排資源並延遲正常訊息的發送，如圖四所示。



圖四：Fuzzy 攻擊示意

（四）溫度攻擊（Temperature）

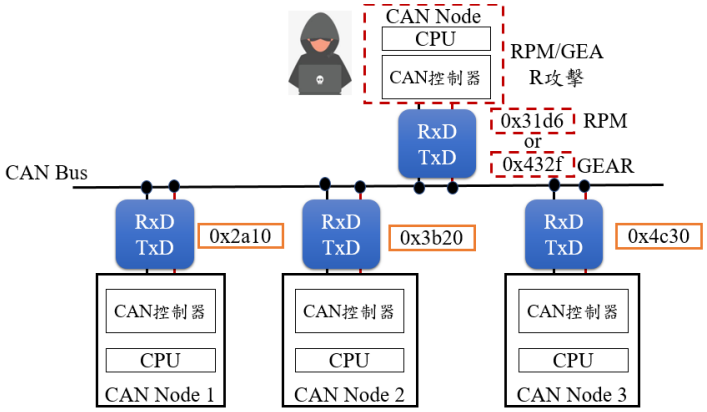
溫度攻擊是一種攻擊方式，駭客首先通過控制溫度節點停止訊息的發送，然後模擬節點，竄改訊息。當被攻擊的節點停止傳輸訊息時，如果某個 ECUs 接收到遠端訊息，它將立即回應。當對方節點未收到回應時，將認定對方 ECUs 處於損壞狀態。攻擊者可以冒充模擬節點，回應遠端請求。因此，模擬節點會定期冒充廣播資料訊息，並回應其他節點的遠端訊息，如圖五所示。



圖五：Temperature 攻擊示意

（五）欺騙駕駛設備/轉速表攻擊(Spoofing Gear/ RPM)

欺騙攻擊通常採用快速(例如每 1 ms 毫秒)注入與假冒 CAN ID 發出駕駛設備或車速表的訊息(RPM/Gear)相關的訊息，造成駕駛設備失效(Spoofing the Drive GEAR)或車速表失誤(Spoofing the RPM gauge) ，如圖六所示。



圖六：RPM/GEAR 攻擊示意

參、CAN 入侵檢測系統設計

本章節介紹了 CAN 資安實驗平臺模型的設計，包括兩個主要項目：(1) 構建 CAN 資安實驗平臺和 (2) 使用 VGG-16 分類器。

3.1 CAN 資安實驗平臺的系統設計

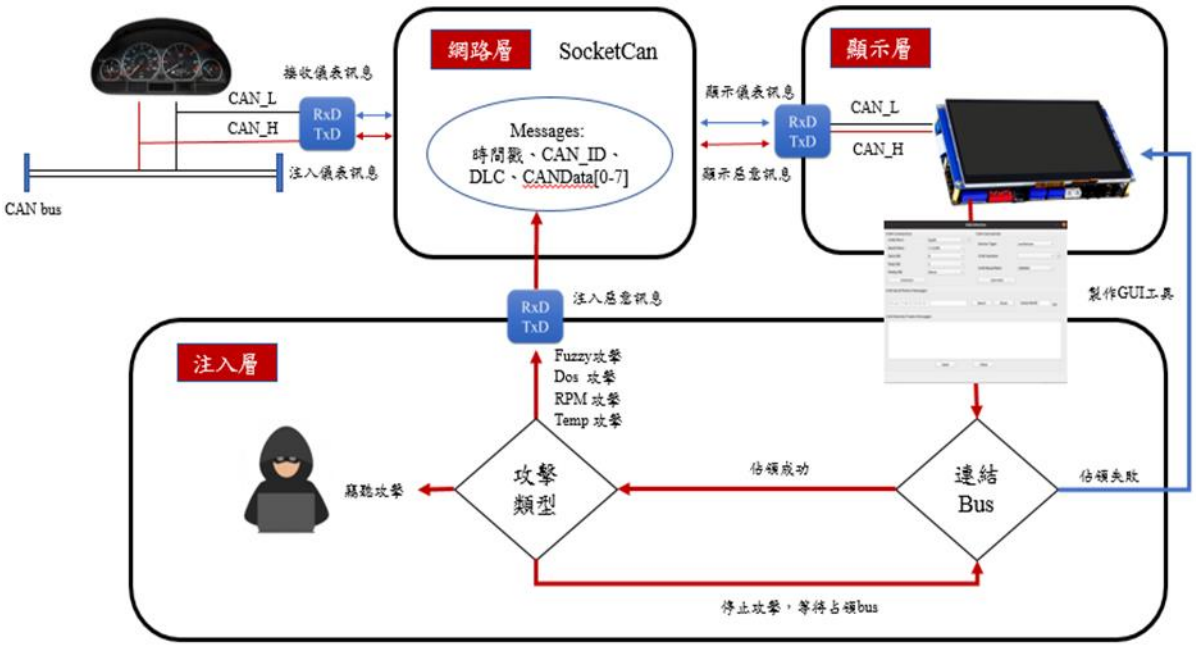
為了支援入侵檢測分析和性能驗證所需的攻擊資訊，本研究採用了模擬工具 QT 5.12.9 平臺、NXP 的嵌入式電路板、Arduino 和 BMW E-46 儀表板，建立了一個 CAN 資安實驗平臺，以執行攻擊訊息的模擬和錄製。CAN 資安實驗平臺的系統架構如圖 8 所示。CAN 車載網路之資安實驗平臺的系統架構設計包括以下的三個層級：(1)注入層、(2)網路層和(3)顯示層。

1. 注入層：使用恩智浦嵌入系統(NXP)的嵌入式板子，搭配 QT 跨平臺和 C++應用開發框架，開發 GUI 程式，以執行攻擊指令注入。攻擊指令經由 NXP 的嵌入式板子內的 TJA1050 晶片傳送到儀表系統。
2. 網路層：接收來自儀表系統內部 CAN 命令，並透過 CAN 通訊板發送至顯示層。
3. 顯示層：通過 Arduino MEGA 2560 控制板，執行儀表驅動和訊息顯示，使用儀表內的 K-bus 進行控制。

開發步驟如下：(如圖七)

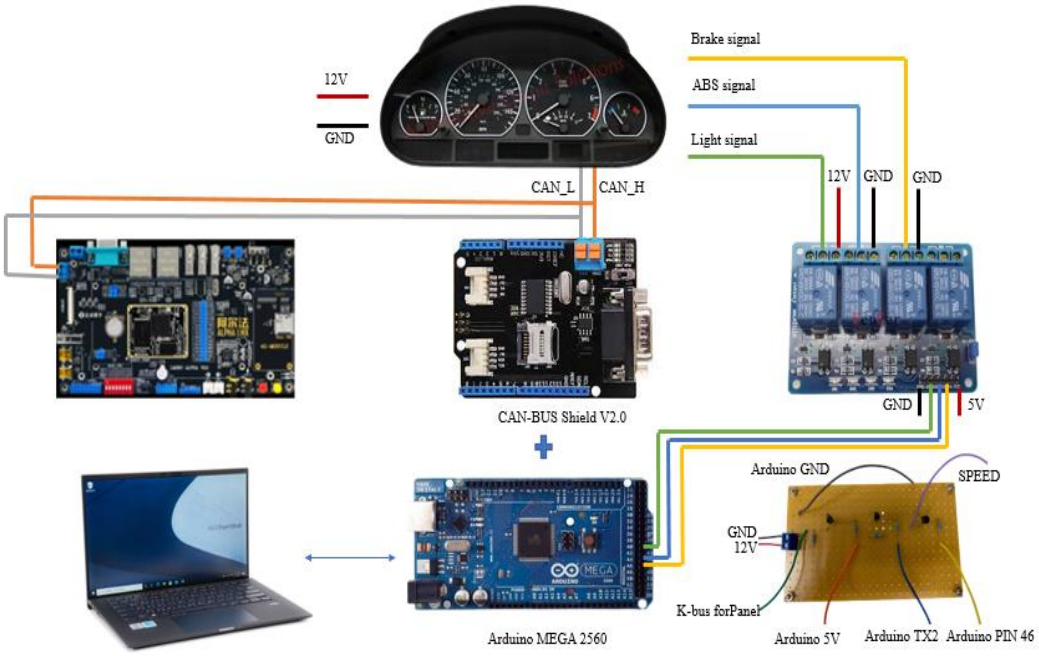
1. 本研究採用模擬工具 Qt 5.12.9 的 Qt CAN Bus 函數庫，開發一個注入層的人機介面應用程式。
2. 在模擬中，駭客發送攻擊命令，這些命令透過 NXP 的嵌入式板傳送到網路層。

3. 在網路層，CAN 接收來自 NXP 嵌入式電路板的 CAN 框架格式。
4. 接著，使用 Arduino MEGA 2560 CAN-BUS 通訊介面板來驅動儀表板。實驗中，我們透過模擬 ABS、手剎、背光連接 4 路繼電器模組，並連接至 K-bus。BMW E-46 儀表板使用 K-BUS 進行驅動。
5. 在顯示層，CAN-BUS 通訊介面板收集的資訊透過 NXP 的嵌入式板的 LCD 圖形介面進行顯示，以呈現惡意攻擊的信息。



圖七：CAN 資安實驗平臺的系統架構

接下來，將資安實驗平臺的系統架構透過電子零件加以實現，完成 CAN 平臺每個單元的 I/O 介面連線圖如圖八。



圖八：單元連線圖

3.2 VGG 神經網路模型設計

經過十多年的發展，卷積神經網路 (CNN) 隨著軟體和硬體的進步，以及越來越多的實驗數據集的建立，不斷深化和複雜化，從而提高了圖像識別的品質。但是，複雜的入侵檢測模型需要高性能的處理器和顯示卡，以滿足其運算需求，這對於計算能力和存儲受限的嵌入式系統提出了挑戰。因此，未來車載網路資安發展的趨勢之一將是設計適用於車載網路的模型。接下來，介紹了入侵偵測 VGG 模型，並在模型基礎上對參數和結構進行了優化，以實現車載網路的入侵檢測。

VGG 模型是由 Simonyan *et al*. (2014) [13] 與 DeepMind 公司共同研發的，並以所在的實驗室 Visual Geometry Group 命名。當年，該模型在 2014 年 ILSVRC 競賽中獲得了分類項目的第二名和定位項目的第一名。以下是 VGG 模型的步驟。

步驟 1：資料預處理

圖像的圖元值範圍為 0 到 255，因此網路數據也應進行正規化，使範圍保持在 0 到 255 之間。有兩種主要的正規化方法，分別是最小-最大（min-max）正規化和分位數正規化，它們都能夠將不同數據值轉換為相同的範圍。然而，最小-最大正規化不夠有效處理異常值，可能導致大多數數據樣本具有極小的值。因此，本研究選擇了分位數正規化 (Hussain et al. 2020) [21]來處理 CAN 正常和攻擊訊息，這一方法在處理極端值時較為穩妥。詳細內容參見表三。

表三：不同特徵正規化方法

方法	公式	描述	研究限制
最小值最大值 (MinMax normalisation)	$\frac{x_i^j - min^j}{max^j - min^j}$	一種資料預處理技術，常應用於特徵縮放，使得所有特徵值均在特定的範圍內[0, 1]。	不能處理離群 (outlier)
分位數 (Quantile normalisation)	$inf\{x \in \mathbb{R} : p \geq F(x)\}$	實現非線性變換，其中每個特徵的連續隨機變量的函數映射到相同的分配。	數據計算量大。

步驟 2. 建立 VGG 模型基本架構

1. VGG 每個網路輸入為 224 x 224 的 RGB 圖片。
2. 網路架構含權重值的層數可以從 11 層到 19 層。
3. 最小卷積核(1 x 1 和 3 x 3)和最小池化核(3 x 3)
4. 使用了兩層神經元數量為 4096 的全連接層，以及 1000 個神經元的全連結層進行分類任務，如表四所示。

表四：VGG 模型架構及參數

卷積網路配置					
A	A-LRN	B	C	D	E
11 權重層	11 權重層	13 權重層	11 權重層	11 權重層	11 權重層
輸入圖片(224 x 224 RGB)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
最大池化層					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
最大池化層					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
最大池化層					

conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
最大池化層					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
最大池化層					
FC-4096					
FC-4096					
FC-1000					
Softmax					

資料來源: (Simonyan, K., Zisserman, 2014) [13]

步驟 3. VGG 模型參數設定

完成 CNN 模型框架的設置後，接下來需要對模型進行參數設定和初始化，以提高模型的穩健性。

1. 損失函數：

損失函數用於通過梯度下降和反向傳播演算法（Back Propagation, BP）來調整模型參數以最小化損失，實現數據擬合。BP 演算法通過正向傳播和反向傳播誤差信號，使用梯度下降演算法調整神經網路的權重值（卷積核）。在初始狀態下，模型的預測值和真實標籤之間存在較大的差距，因此需要訓練模型，逐漸調整每一層的權重參數，使預測值更接近真實標籤，以達到最佳狀態。本研究選擇了交叉熵（cross-entropy）作為損失函數，計算公式如下（1），其中 y 表示真實標籤值，ŷ 表示預測值。

$$f(y,\hat{y}) = -\sum ylog(\hat{y}) \tag{1}$$

2. 隱藏層啟動函數：

在神經網路中，每個神經元節點接收來自上一層神經元的權重輸出值作為輸入，然後將這個值傳遞給下一層。這兩層之間的關係由一個函數稱為啟動函數所描述。當兩層之間沒有使用啟動函數時，模型將變成線性模型，僅能逼近線性函數，對特徵的表達能力受限。啟動函數的加入引入了模型的非線性因素，使模型能夠逼近任意函數。

在梯度反向傳播過程中，使用飽和的啟動函數會導致梯度消失（Vanishing Gradient）和梯度爆炸（Exploding Gradients）等問題，特別是在訓練深度神經網路時。為了避免

這些問題，CNN 模型的隱藏層通常使用修正線性單元（ReLU）作為啟動函數。ReLU 函數計算簡單，梯度不會飽和。其計算如下（2）：

$$Relu = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (2)$$

3.3 性能檢測指標

性能檢測指標

入侵分類器的性能評估通常透過以下指標：精確度（Accuracy）、精準率（Precision）、召回率（Recall）和 F1 值（F1-score）。這些指標在機器學習中被廣泛使用，尤其在我們關注多個類別而且想要提高模型預測能力的情況下。

1. 精確度（Accuracy）：

精確度是一個用來衡量分類模型整體性能的指標，計算公式如下（3）：

$$ACR = \frac{TP+TN}{TP+FP+FN+TN}, \quad (3)$$

其中，真陽性（True Positive, TP）代表預測為陽性且實際為陽性的案例，偽陽性（False Positive, FP）代表預測為陽性但實際為陰性的案例，偽陰性（False Negative, FN）代表預測為陰性但實際為陽性的案例，真陰性（True Negative, TN）代表預測為陰性且實際為陰性的案例。

2. 精準率（Precision）：

精準率衡量模型在所有預測為陽性的案例中，實際為陽性的比率，計算如下：

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

3. 召回率（Recall）：

召回率也稱為真陽性率，它衡量模型在所有實際陽性案例中成功檢測到的比率，計算如下：

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

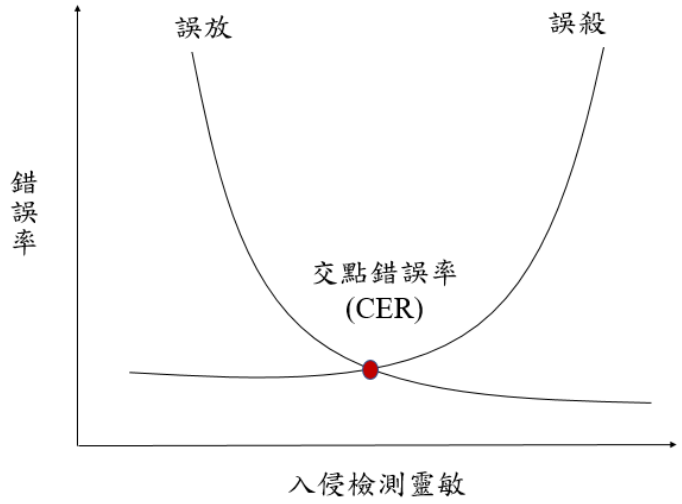
4. F1 值（F1-Score）：

F1 值綜合了精確率和召回率，用於平衡這兩者之間的性能。當不同類別的樣本數量

差異較大時，F1 值通常用來評估模型的性能。計算如下：

$$F1 - Soure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

入侵檢測的目標是盡可能檢測出攻擊行為，同時降低誤漏報率（False Negative, FN）和誤報率（False Positive, FP）。調整入侵檢測的靈敏度可以改變這兩種錯誤率，而最佳靈敏度通常出現在兩個錯誤率曲線的交點，這稱為交點錯誤率（Crossover Error Rate, CER），如圖九所示。



圖九：入侵檢測交點錯誤率

資料來源：潘天佑(2012) [22]

肆、 系統實作與數據分析

4.1 CAN 資安實驗平臺之實作

為了理解 CAN 車載網路運作與威脅的真實情況，本研究透過架設真實 CAN 車載網路及模擬各項攻擊，實現了一個基於 CAN 的入侵偵測系統，在實驗中涉及三個不同功能的層：注入層、網路層和顯示層，系統完成如圖十所示。資安實驗平臺可以模擬 CAN 入侵情境並產出各項惡意訊息，並進行訊息錄製以訓練 VGG-16 的模型，開發環境的軟硬體如表五。CAN 資安實驗平臺可支援錄製訊息、標籤和模型練訓，其輸出資料及提供 VGG-16 模型學習攻擊訊息和正常訊息，對車載網路的威脅進行分類與識別。

圖十：CAN 資安實驗平臺之實現

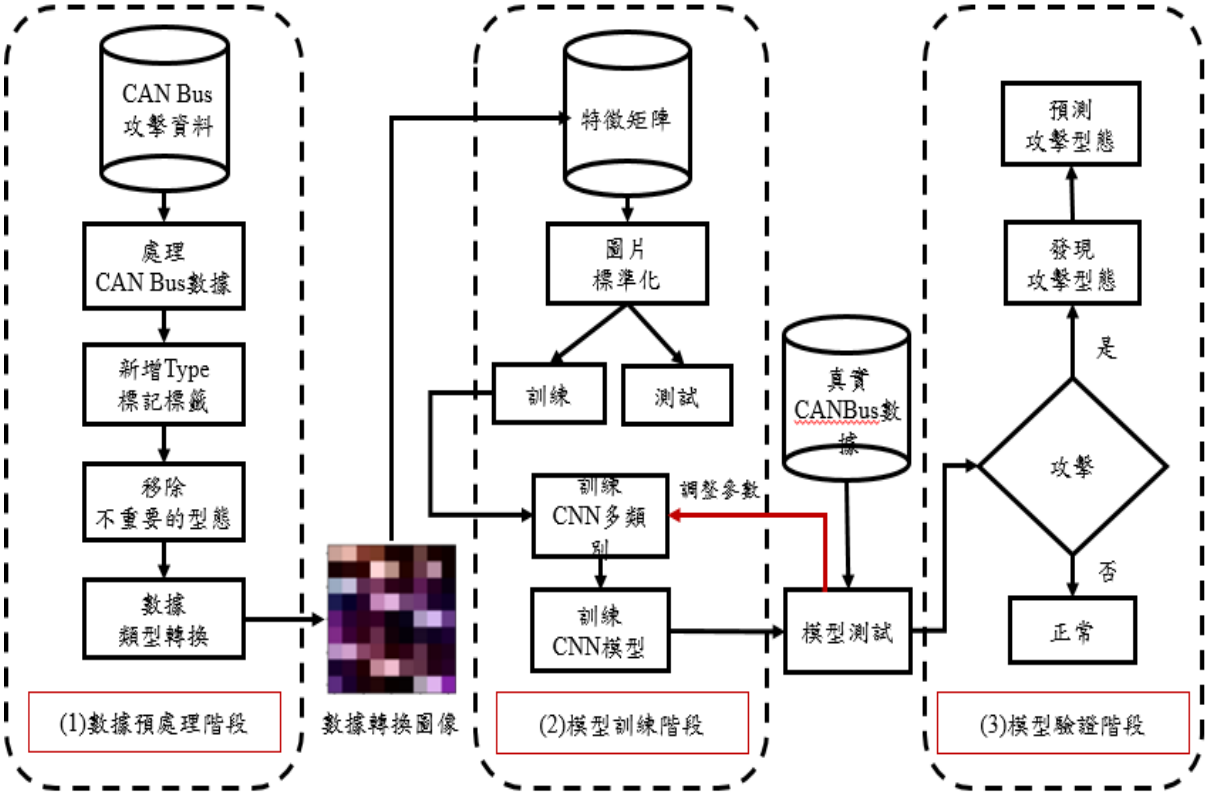
表五：實現 CAN 資安實驗平臺之軟硬體環境

中央處理器(Central Processing Unit, CPU)	Intel i5-12500【6 核/12 緒】3.0GHz(4.6G)
隨機存取記憶體(Random-access memory , RAM)	64GB
圖形處理器(Graphics Processing Unit, GPU)	NVIDIA RTX3050 GAMING OC 8G 顯示卡
作業系統(Operating System, OS)	Windows 10 專業版
DL 框架	Tensorflow 2.5.0、Scikit-learn
Python 套件	Opencv、Numpy、Scipy、Pandas

4.2 入侵偵測數據分析

模型的運作包括三個子階段：(1)數據預處理階段，(2)模型訓練階段，(3)模型驗證階段，如圖十一所示。

在特徵提取階段，本研究實驗攻擊的 CAN 數據格式參考來自 HCRL 汽車駭客資料集的格式[8]。本實驗使用真實儀表系統錄製 40～50 分鐘的 CAN 訊息，總共約有 300～400 萬條訊息，如表六。



圖十一：使用 CNN 進行車載入侵檢測

表六：CAN 訊息格式

編號	名稱	描述
1	CAN ID	CAN 標識符(ex. 545)
2	DLC	資料長度
3	DATA[0-7]	十六進制資料
4	Flag	T 或 R，T 代表注入，R 代表正常
5	label	標籤(ex. 0 = DoS, 1 = Fuzzy)
6	Type	類型(ex. DoS, Fuzzy)
7	Attack	攻擊方式

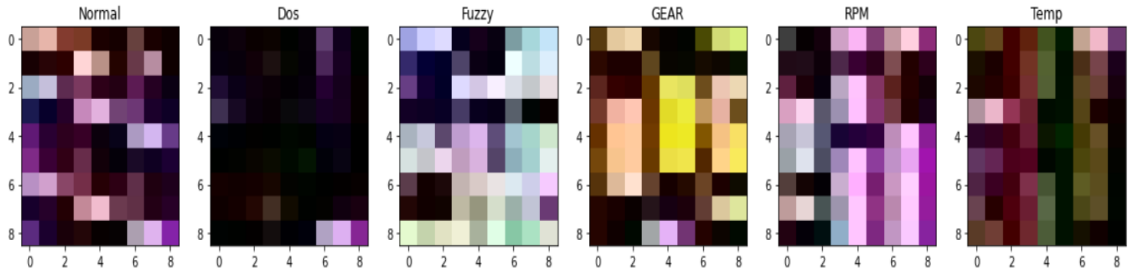
對於本研究方法，使用包含五種數據攻擊類型和正常數據：DoS 攻擊、Fuzzy 攻擊、Temperature 攻擊、Spoofing Gear / RPM 攻擊和正常數據集，將六個 CSV 文件的實驗合併，總共 24,000,000 條訊息，篩選在正常資訊中被注入攻擊，並平均資訊，如表七所示。

表七：實驗數據

攻擊類型	正常的訊息	注入的訊息數	總共的訊息
正常	4,000,000	-	4,000,000
DoS 攻擊	2,000,000	2,000,000	4,000,000
Fuzzy 攻擊	2,000,000	2,000,000	4,000,000
Spoofing Gear 攻擊	2,000,000	2,000,000	4,000,000
RPM 攻擊	2,000,000	2,000,000	4,000,000
Temperature 攻擊	2,000,000	2,000,000	4,000,000

步驟 1 資料預處理

本研究使用分位數轉換(Quantile Transformer, QT)搭配 Scikit-Learn 涵式庫 將攻擊訊息轉換成圖片，QT 轉換是一種非參數數據轉換技術，可將數值或數據分佈轉換為遵循一定的數據分佈(通常是高斯分佈(正態分佈))，可以將數據轉換為正態分佈或均勻分佈。分位數正規化方法變換特徵分佈到正態分佈並重新計算所有正態分佈的特徵值，所以大多數變量值接近中值，這在處理異常值方面是有效的。數據樣本被轉換根據時間戳記和特徵大小分成網路流量數據集。對於實驗數據集，因為它有 9 個重要特徵(CAN ID 和 DATA[0]-DATA[7])，每個具有 9 個特徵的 27 個連續樣本的流量(27 x 9 = 243 總特徵值)轉換為形狀圖像(9, 9, 3)。因此，每個轉換後的圖像都是正方形顏色具有三個通道(紅色、綠色和藍色)的圖像，如圖十二所示。



圖十二：訊息的可視化圖像(正常+攻擊)

步驟 2. 模型訓練和優化階段

本研究使用 VGG 模型的 VGG-16，因任務的不同，本研究對傳統的 VGG-16 架構進行修改，在原本架構進行簡化和改良，如表八所示。主要修改說明如下：

1. 根據惡意攻擊數據十六進制轉換整數型態，並使用分位數歸一化方法轉換圖片，對輸入圖片尺寸為(150 x 150 x 3)。
2. 減少每個卷積核內卷積層的濾波數量。
3. 加入 Dropout 和 Dense 層。

4. 實驗時將使用數據共有 5 種攻擊類型和 1 種正常類型，因此將 VGG-16 的輸出層增加為個神經元(含正常)。
5. 使用 softmax 在 CNN 上設定多分類的模型參數、優化器(optimizers)使用是 Adam。學習結果作為模型驗證的模型參數的基礎，如 weight 矩陣，batch_size，batches_per_epoch 和分類準確性。

表八：優化 VGG-16 模型

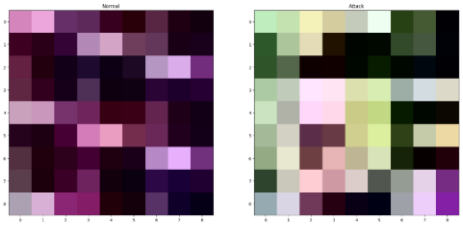
層	過濾器數量	過濾器大小	特徵圖生成
卷積層 C_1×2	64	3 x 3 x 3	150 x 150 x 64 150 x 150 x 64
池化層 P_1	1	2x2	75 x 75 x 64
卷積層 C_2×2	128	3 x 3 x 64	75 x 75 x 128 75 x 75 x 128
池化層 P_2	1	2x2	37 x 37 x 128
卷積層 C_3×3	256	3 x 3 x 128	37 x 37 x 256 37 x 37 x 256
池化層 P_3	1	2x2	18 x 18 x 256
卷積層 C_4×3	512	3 x 3 x 256	18 x 18 x 512 18 x 18 x 512
池化層 P_4	1	2x2	9 x 9 x 512
卷積層 C_5×3	512	3 x 3 x 512	9 x 9 x 512 9 x 9 x 512
池化層 P_5	1	2x2	4 x 4 x 512
最大池化層 P_6	1	2x2	2 x 2 x 512
Dropout (rate=0.5)	-	-	256
Dense 層	-	-	256
分類(Softmax)	-	-	6 分類

步驟 3. 模型驗證階段

入侵檢測準確率(Accuracy)計算是從分類器正確的預測資料佔所有樣本預測資的比例，此外，本研究搭配使用精準率(Precision)、召回率(Recall)和 F1 值(F1-score)來進行檢測模型性能的好壞。

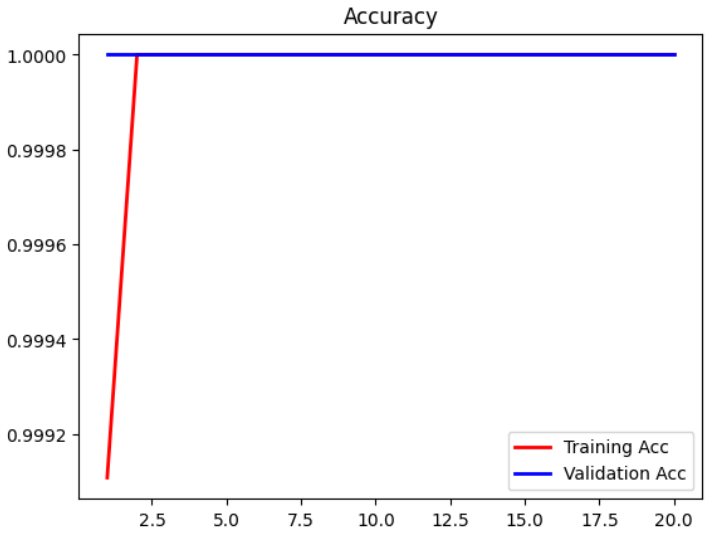
實驗一：二元分類(Binary classification)

將資訊預處理，二類圖片總共 4,000,000 張，將二類圖片分訓練和測試為 80%和 20%，轉換後的圖像作為 VGG-16 分類模型訓練資料，如圖十三所示。

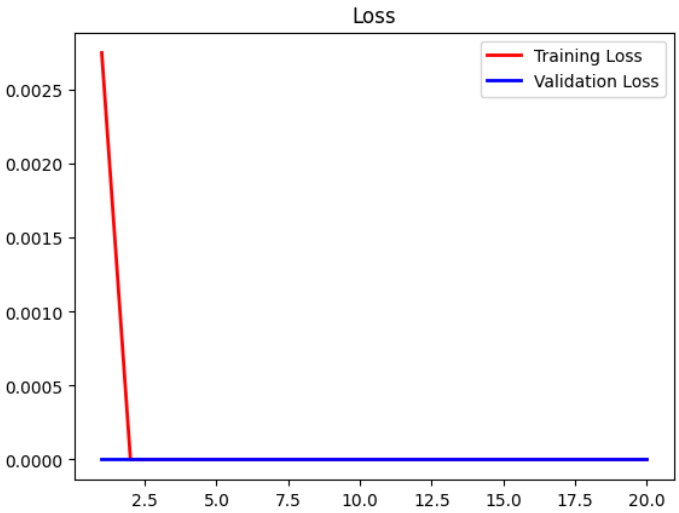


圖十三：二元分類資訊轉為圖像

用 VGG-16 模型進行驗證，如圖十四和表九所示。



準確率(a)



損失(b)

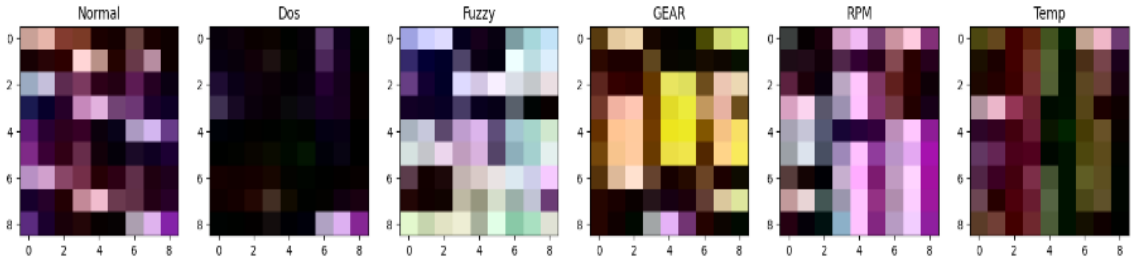
圖十四: 二分類的訓練和驗證準確率

表九：二分類模型驗證

	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
訓練	100%	100%	100%	100%
測試	100%	100%	100%	100%

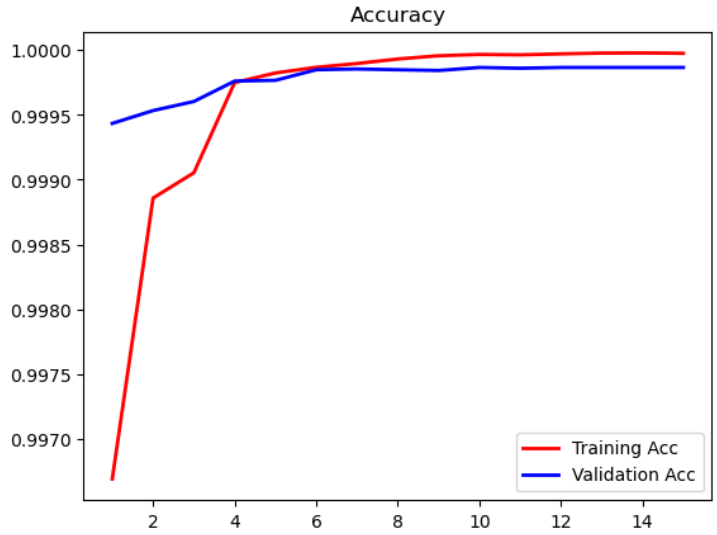
實驗二：多元分類(Multi-classification)

將資訊預處理，六類圖片總共 24,000,000 張，將六類圖片分訓練和測試分別為 80%和20%樣本，轉換後的圖像作為 VGG-16 分類模型訓練資料，如圖十五所示。

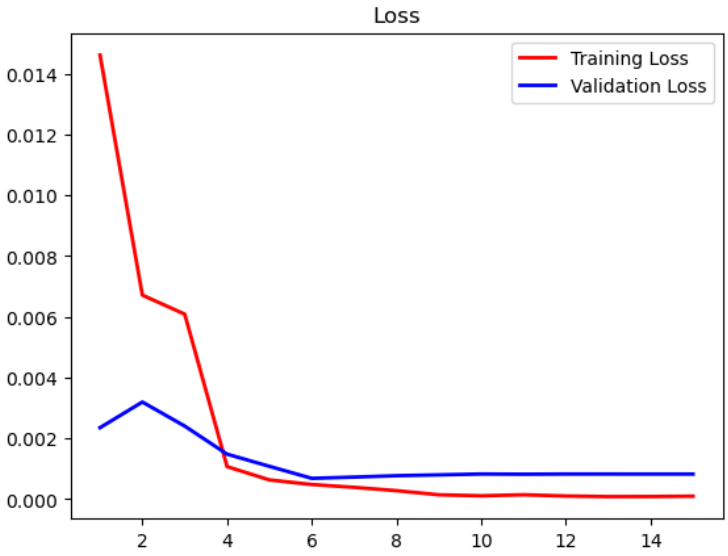


圖十五：六分類之資訊轉為圖像

用 VGG-16 模型進行驗證性能，實驗結果，如圖十六和表十，六分類混淆矩陣，如圖十七所示。



準確率(a)

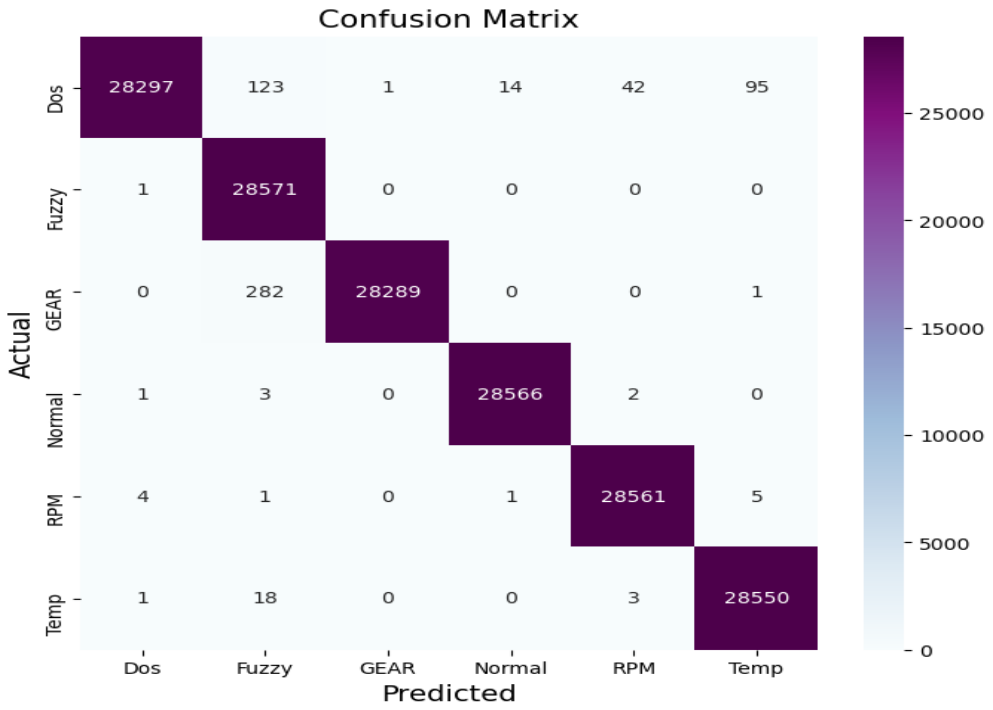


損失(b)

圖十六：六分類的訓練和驗證準確率(a)與損失(b)

表十：模型性能驗證

	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
訓練	99%	99%	99%	99%
測試	99%	99%	99%	99%



圖十七：六分類混淆矩陣

4.3 偵測模式性能比較

為了分析本研究使用的模型性能，參考 Lee et al.[8]提出車載網路的誤用訊息的偵測方法，在該數據及的分類性能指標與本研究結果進行比較。將對各種攻擊的檢測指標與本研究的檢測指標，結果進行比較，如表十一所示。從表十一看出，本研究使用分位數正規化方法轉換，結合簡化的 VGG-16 模型在針對 5 類攻擊和正常數據的檢測率，平均高達 99%以上，故本研究的方法在 5 類攻擊和正常數據中，都具有較高檢測能力。

表十一：模型性能指標比較

DoS	Precision	Recall	F1
VGG-16	**1.0**	**0.99**	**1.0**
ResNet	1.0	0.99	0.99
Fuzzy	Precision	Recall	F1
VGG-16	**0.99**	**1.0**	**0.99**
ResNet	0.99	0.99	0.99
Gear Spoofing	Precision	Recall	F1
VGG-16	**1.0**	**0.99**	**1.0**
ResNet	0.99	0.99	0.99

Tempature	Precision	Recall	F1
VGG-16	**1.0**	**1.0**	**1.0**
ResNet	0.99	0.99	0.99
RPM	Precision	Recall	F1
VGG-16	**1.0**	**1.0**	**1.0**
ResNet	0.99	0.99	0.99

伍、結論與未來研究方向

車載入侵偵測系統是目前解決車載網路資訊安全問題的手段之一，可完成車載網路訊息認證、加密、危害隔離等任務，不需要修改車載網路架構，不佔用汽車 CAN 車網路的有限資源，只需要在 CAN 上增加一個網路閘道器做為檢測與分析點，可以即時監控 CAN 流量，對入侵訊息進行有效檢測。

本研究運用 VGG-16 分類完成車載網路之模擬攻擊訊息的偵測，經實驗證明 VGG-16 分類模型可有效偵測 5 種常見入侵手段與查出對應系統弱點，協助防止駭客對車載網路的危害。

本研究實驗的入侵檢測方法針對單一 BMW-E46 汽車的儀表系統進行主要攻擊，無法代表能完整將攻擊訊息傳至所有汽車儀表，由於自身能力及諸多因素的限制，研究成果尚有許多不足處，將做未來車網路及車聯網(IoV)研究的方向。

[誌謝]
本研究承蒙教育部計畫（編號：MOE 2000-109CC5-001）、國科會計畫(NSTC 112-2410-H-168-002）經費補助，謹此致謝。

參考文獻

[1] R. Charetter, "This Car Runs on Code," in IEEE Spectrum, vol. 46, no. 3, pp. 3, 2009.

[2] M. Charlie and V. Chris, "Adventures in automotive networks and control units," DEFCON, Las Vegas, CA, USA, Aug. 1-4, 2013.

[3] Renesas, "CAN 入門教程." [Online]. Available: https://www.renesas.cn/cn/zh, 2022.

[4] E. Wang, W. Xu, S. Sastry et al., "Hardware module-based message authentication in intra-vehicle networks," in Proc. ACM/IEEE 8th Int. Conf. Cyber-Phys. Syst. (ICCPS), New York, NY, USA, Apr. 18-20, 2017, pp. 207-216.

[5] X. Zhou, Z. Pan, G. Hu, S. Tang, and C. Zhao, "Stock market prediction on high-frequency data using generative adversarial nets," Mathematical Problems in Engineering, vol. 2018, Article ID 4907423, 2018. [Online]. Available: https://doi.org/10.1155/2018/4907423

[6] U. E. Larson, D. K. Nilsson, and E. Jonsson, "An approach to specification-based attack detection for in-vehicle networks," in 2008 IEEE Intelligent Vehicles Symposium, Eindhoven, Netherlands, Jun. 4-6, 2008, pp. 220-225.

[7] M. Müter and N. Asaj, "Entropy-based anomaly detection for in-vehicle networks," in IEEE Intelligent Vehicles Symposium (IV), Baden-Baden, Germany, Jun. 5-9, 2011, pp. 1110-1115.

[8] H. Lee, S. H. Jeong, and H. K. Kim, "OTIDS: A novel intrusion detection system for in-vehicle network by using remote frame," in 15th Annual Conf. on Privacy, Security and Trust (PST), Calgary, AB, Canada, Aug. 28-30, 2017, pp. 5709-5757.

[9] A. Taylor, S. Leblanc, and N. Japkowicz, "Anomaly Detection in Automobile Control Network Data with Long Short-Term Memory Networks," in IEEE Int. Conf. on Data Science and Advanced Analytics (DSAA), Montreal, QC, Canada, Oct. 17-19, 2016. DOI: 10.1109/DSAA.2016.20

[10] E. Seo, H.M. Song, and H.K. Kim, "GIDS: GAN Based Intrusion Detection System for In-vehicle Network," 16th IEEE Annual Conf. on Privacy, Security and Trust (PST), Belfast, Northern Ireland, UK, Aug. 30, 2018, arXiv:1907.073-77.

[11] H.M. Song, J. Woo, and H.K. Kim, "In-vehicle network intrusion detection using deep convolutional neural network," Vehicular Communications, vol. 21, pp. 100-198, 2020.

[12] C. Szegedy, S. Ioffe, V. Vanhoucke, et al., "Inception-v4, inception-resnet and the impact of residual connections on learning," arXiv:1602.07261, 2016.

[13] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Computer Vision and Pattern Recognition (cs.CV), arXiv:1409.1556, 2014.

[14] C. Miller and C. Valasek, "Remote exploitation of an unaltered passenger vehicle," Black Hat USA, 2013.

[15] Zdnet, "OwnStar: Unlock and track any GM OnStar connected car for $100," 2015. [Online]. Available: https://www.zdnet.com/article/ownstar-the-gm-onstar-connected-cars-worst-security-nightmare/

[16] iThome, "資安大會焦點直擊智慧汽車漏洞現形記," Sep. 10, 2015. [Online]. Available: https://www.ithome.com.tw/news/98538

[17] 騰訊科恩實驗室, "最新研究成果:2017 再一次實現對特斯拉的無物理接觸遠程攻擊," 2017. [Online]. Available: https://reurl.cc/Q6D6zO

[18] L. Wouters, B. Gierlichs, and B. Preneel, "My other car is your car: compromising the Tesla Model X keyless entry system," in IACR Transactions on Cryptographic Hardware and Embedded Systems, vol. 2021, no. 4, pp. 149–172, 2021.

[19] S. Craig and C. E., The Car Hacker's Handbook, 杜靜, 李博, 敖富江譯, 清華大學出版社, 2017.

[20] 陳力, "基於深度學習的車載匯流排網路入侵檢測方法研究," 浙江科技學院, Dec. 18, 2020.

[21] F. Hussain, S. G. Abbas, M. Husnain, U. U. Fayyaz, F. Shahzad, and G. A. Shah, "IoT DoS and DDoS Attack Detection using ResNet," in Proceedings of the 2020 IEEE 23rd International Multitopic Conference (INMIC), Bahawalpur, Pakistan, Nov. 5-7, 2020, pp. 1–6.

[22] 潘天佑, 資訊安全概論與實務, 第三版, ISBN：9789862766644, 碁峰出版社, 2012.

[作者簡介]

王平，交通大學資訊管理研究所博士，現為崑山科技大學資訊管理系教授，同時兼任副研發長，研究方向為深度學習神經網路、資訊安全、網路服務及技術創新與專利佈局之研究。

陳佳鴻，現為崑山科技大學資訊管理系研究生，研究專長為電腦病毒特徵分析、網路入侵偵測與系統弱點掃描之研究。