

基於深度學習的神經網路架構故障診斷與容錯控制

Fault diagnosis and fault tolerant control of neural network architecture based on deep learning

壽鶴年

Ho-Nien Shou

空軍航空技術學院通訊電子系

Department of Aviation & Communication Electronics, Air Force Institute of Technology

摘要

本文提出深度學習中兩大神經網路演算法應用於微衛星姿態故障診斷(fault detection)與容錯控制(fault-tolerant control)。基於深度學習(deep learning)的故障診斷方式，以微衛星姿態故障小波頻譜作為主要依據，結合生成對抗網路(Generative Adversarial Networks, GAN)識別圖像的功能，在半監督學習(semi-supervised learning)生成對抗網路(Semi-Supervised Learning with Context-Conditional Generative Adversarial Networks, SSL)的基礎上對損失函數(loss function)進行最佳化，指導生成器和鑑別器的訓練，作為系統即時故障診斷與分類。當系統故障診斷確定故障發生時，小腦神經網路(Cerebellar Model Articulation Controller, CMAC)介入容錯控制。小腦神經網路是一種局部學習網路，是基於表格查詢式輸入輸出的局部逼近神經網路，提供了一種從輸入到輸出的多維非線性映射能力。利用生成對抗網路半監督學習能力與分別解決了樣本資料不足和樣本標籤問題，使用小腦神經網路為局部學習網路，具泛化能力(generalization)強收斂速度快，易於軟硬體實現。模擬結果表示，生成對抗網路故障診斷結合小腦神經網路容錯控制方法與其他方法相比能達到較高準確率與強健性。

關鍵字：深度學習、生成對抗網路、故障診斷、小腦神經網路、容錯控制

Abstract

In the paper, two neural network algorithms in deep learning are applied to micro satellite attitude fault detection and fault tolerant control respectively. Firstly, in the aspect of fault diagnosis, a method of satellite attitude fault diagnosis based on the improved Generative Adversarial Networks (GAN) is proposed. The wavelet spectrum of microsatellite attitude fault is taken as the main basis, combining the ability of continuous wavelet transform to process non-stationary signal and the Semi-supervised Generative Adversarial Networks(SGAN) can process and recognize images. Based on the SGAN, it introduces the condition model and optimizes the loss function. In the process, it will guide the training of generator and discriminator. In the event of actuator failure, Cerebellar Model Articulation Controller Neural Network(CMAC) intervenes in fault-tolerant control to restore the satellite attitude with the shortest time and the least energy. CMAC is a kind of local learning network, which is based on the local approximation neural network of table query input and output, and provides a multi-dimensional nonlinear mapping ability from input to output. The simulation results show that the fault diagnosis method based on CWT and semi supervised generation network can achieve higher accuracy than other mainstream diagnosis methods. When the system is normal or fault, cerebellar neural network can provide robustness and adaptability for the system.

Keywords: deep learning, fault detection, fault-tolerant control, generative adversarial networks,

semi-supervised generative adversarial networks, cerebellar model articulation controller neural network

一、前言

微衛星擁有機械結構簡單、可靠性高、價格低廉、研究人員與設備精簡等優點。所以設計成本低、製作時間短、及高精度的微小衛星便成為各大學及小型科技機構主要研究之領域。我國第三期太空計畫預計以自主發展微衛星為主要里程碑，並期望使用自主載具進行發射。對於低軌道、低費用的微衛星，由於重量和功率的限制，在保證姿態控制精度的情況下，需要盡可能地減少不必要的姿態感測器及控制器，採用最少的硬體構成姿態確定系統，微衛星工程是未來的趨勢。微衛星在軌任務過程中制動器(actuator)與感測器(sensor)出現故障的頻率相對較高，造成其故障的原因及類型多樣，實在無法逐一列舉與避免，故將最為常見類型種類分類多樣，因此對於微衛星的故障診斷具有理論和實際意義。

近年對於電機軸承故障診斷研究被學者工程師廣泛關注，電機軸承故障診斷軸承常用振動分析法，對於採集的軸承振動信號進行分析和分類。然而軸承振動信號具有非平穩性等特點因此對其進行特徵提取有相當難度[1]。連續小波變換(Wavelet Transform)作為一種時頻分析方法能夠有效提取非平穩信號的時頻特徵。小波變換的作用是以一維信號作為輸入對時頻分析結果本質上是反映信號在不同時間和頻率的能量強度二維圖譜，輸出以時頻圖(time-frequency diagrams)的形式展現。目前連續小波變換在機械故障診斷領域均得到廣泛應用[2-3]。在文獻[2]提出小波變換分支撐向量機(support vector machine)與卷積神經網路(convolutional neural networks, CNN)相結合，文獻[3]以小波變換與倒傳遞神經網路(back propagation neural network, BP)實現不同電機的故障診斷。上述文獻應用中特徵向量是以時頻圖作為基礎而人為構造的而且所選神經網路特徵學習表達能力有限[4]。

生成對抗網路(Generative Adversarial Network, GAN)於 2014 年被提出[5]，生成對抗網路利用生成器(Generator, G)和鑑別器(Discriminator, D)的相互對抗不斷優化網路達到更好的效果。生成對抗網路具有多個隱藏的深度學習網路模型具有良好的特徵學習能力，生成對抗網路最初主要應用於資料生成以及圖像生成用於解決樣本資料不足的問題。基於不同研究目的的生成對抗網路的衍生模型不斷被提出，其中以半監督學習為基礎提出半監督生成對抗網路與原先生成對抗網路結構類似，生成器結構保持不變鑑別器在二分類器的基礎上添加輸出層達到多分類功能[6]。

在 2016 提出半監督生成對抗網路(Semi-supervised Generative Adversarial Network, SGAN)[7]，強制鑑別器輸出類標籤，實現了在只有少量標籤資料的情況下，大量未標記資料能輔助鑑別器提取不同類別圖像的特徵，從而影響鑑別器的分類效果及泛化能力。

自從半監督生成對抗網路提出後被應用於語音的生成與識別圖像識別以及樣本分類等領域[8]。將半監督生成對抗網路應用於人體動作識別將視頻以幀為單位截取圖片，作為網路輸入進行動作識別[9]。在上述應用中半監督生成對抗網路達到了識別效果。

目前對於微衛星姿態故障診斷與容錯控制的研究，被工程師與研究員廣泛關注，基於各種智慧化方法的衛星相關故障診斷研究層面，無法像電機軸承故障診斷大多數採用監督學習的方式並且所需樣本量較大。本文對半監督生成對抗網路進行改進和優化提出一種基於連續小波變換和改進故障診斷方法採用一維振動信號作為原始資訊經過預處理後生成時頻圖以此作為網路的輸入進行致動器/感測器故障診斷。在衛星診斷判別出致動器/感測器故障後，將由小腦神經網路提供容錯控制，小腦神經網路來學習非線性系統的

輸出入函數，相較於一般類神經網路 (Artificial Neural Network, ANN) 而言，不但學習收斂速度快，同時亦具有良好的學習效能，相當適合於線上的即時控制系統，也較容易將演算法轉化成硬體實現。

二、生成式對抗網路

生成式對抗網路已經成為人工智慧學界一個熱門的研究方向。生成式對抗網路的基本思想源自對局理論(game theory) 的零和二人對局遊戲(zero-sum two-person game)，也是一個最小最大化策略 (Minmax Strategy) 問題。生成式對抗網路由一個生成器(Generator, G)和一個鑑別器(Discriminator, D)構成，通過對抗學習的方式來訓練，目的是估測資料樣本的潛在分佈並生成新的資料樣本，生成式對抗網路目前正在被廣泛應用在圖像和視覺計算、語音和語言處理、資訊安全、棋類比賽等領域研究上。[9-12]

生成對抗網路是一種生成器 (generative model)，其基本架構是從訓練庫裡獲取很多訓練樣本，從而學習這些訓練案例生成的機率分佈。實現的方法是讓兩個網路相互競爭，生成器 G ，它不斷從訓練庫裡真實資料的機率分佈，將輸入的隨機雜訊 (Random Noise) 轉變成新的樣本 (也就是假樣本)。而鑑別器 D 從觀察真實和假造的資料當中，判斷這個資料到底是不是真的。

鑑別器 D 實際上是對資料的來源進行一個判別：究竟這個資料是來自真實的資料分佈 p_{data} ，還是來自於一個生成器 G 所產生的一個資料分佈 p_g 。鑑別器 D 的訓練目的就是要儘量最大化自己的判別準確率。當這個資料被判別為來自於真實資料時(標注為 1)，至於生成資料時(標注為 0)。而與這個目的相反的是，生成器 G 的訓練目標，就是要最小

化鑑別器 D 的判別準確率。在訓練過程中，生成式對抗網路採用了一種非常直接的交替最佳化方式，它可以分為兩個階段，第一個階段是固定鑑別器 D ，然後最佳化生成器 G ，使得鑑別器的準確率儘量降低。而另一個階段是固定生成器 G ，來提高鑑別器的準確率。圖 1 所示為生成對抗網路方塊圖。

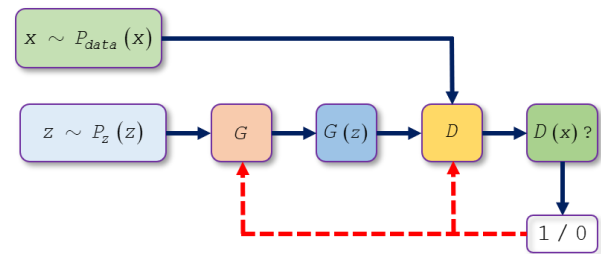


圖 1 生成對抗網路方塊圖

給定生成器 G 的情況下，我們考慮最佳化鑑別器 D ，和一般基於 S 形曲線 (Sigmoid) 的二分類模型訓練一樣，其損失函數為：

$$\begin{aligned} \text{Obj}^D(\theta_D, \theta_G) = & E_{x: p_{data}(x)} [\log D(x)] \\ & + E_{z: p_z(z)} [\log(1 - D(g(z)))] \end{aligned} \quad (1)$$

其中， x 取樣於真實資料分佈 $p_{data}(x)$ ， z 取樣於先驗分佈 $p_z(z)$ (例如高斯雜訊分佈)，

E 表示計算期望值。這裡實際訓練時和常規二值分類模型不同，鑑別器的訓練資料集來源於真實資料集分佈 $p_{data}(x)$ (標注為 1) 和生成器的資料分佈 $p_g(x)$ (標注為 0) 兩部分。

給定生成器 G ，我們需要最小化式(1) 來得

到最佳解，在連續空間上，式(1)可以寫為如下形式：

$$\begin{aligned} Obj &= \int_x p_{data}(x) \log(D(x)) dx \\ &+ \int_z p_z(x) \log(1-D(g(z))) dz \\ &= \int_x \left[p_{data}(x) \log(D(x)) \right. \\ &\quad \left. + p_g(x) \log(1-D(x)) \right] dx \end{aligned} \quad (2)$$

對任意的非零實數 m 和 n ，且實數值 $\sigma \in [0,1]$ ，表達式

$$-m \log(\sigma) - n \log(1-\sigma) \quad (3)$$

在 $\frac{m}{m+n}$ 處得到最小值。因此，給定生成器 G 的情況下，目標函數(2)在

$$D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x) + p_g(x)} \quad (4)$$

處得到最小值，此即為鑑別器的最優解。由式(4)可知，生成式對抗網路估計的是兩個機率分佈密度的比值，這也是和其他基於最小最佳化或者馬可夫鏈(Markov chain)方法的關鍵不同之處。

另一方面， $D(x)$ 代表的是 x 來源於真實資料而非生成資料的概率。當輸入資料取樣自真實資料 x 時， D 的目標是使得輸出概率值 $D(x)$ 趨近於 1，而當輸入來自生成資料 $G(z)$ 時， D 的目標是正確判斷資料來源，使得 $D(G(z))$ 趨近於 0，同時 G 的目標是使得其趨近於 1。這實際上就是一個關於 G 和 D 的零和遊戲，那麼生成器 G 的成本函數為

$$Obj^D(\theta_D, \theta_G) = -Obj^G(\theta_G) \quad (5)$$

所以生成式對抗網路的最佳化問題是一個極小-極大化問題，生成式對抗網路的目標函數可以描述如下：

$$\begin{aligned} \min_G \max_D V(D, G) &= E_{x: p_{data}(x)} [\log D(x)] \\ &+ E_{z: p_z(z)} [\log(1-D(G(z)))] \end{aligned} \quad (6)$$

對於生成式對抗網路的學習過程，我們需要訓練模型 D 來最大化判別資料來源於真實資料或者偽數據分佈 $G(z)$ 的準確率，同時，我們需要訓練模型 G 來最小化 $\log(1-D(G(z)))$ 。這裡可以採用交替最佳化的方法：先固定生成器 G ，最佳化鑑別器 D ，使得 D 的判別準確率最大化；然後固定鑑別器 D ，使生成器 G 最佳化，使得 D 的判別準確率最小化。當且僅當 $p_{data} = p_g$ 時達到全域最佳解。訓練生成式對抗網路時，同一輪參數更新中，一般對 D 的參數更新 k 次再對 G 的參數更新 1 次。

總結綜合上述優點生成式對抗網路確實為無監督學習，提供了一個非常有潛力的演算法則方案。

(i)生成式對抗網路是最佳的生成器，避免了解馬爾科夫鏈 (Markov Chain) 的學習機制:能夠區別於傳統的機率生成器。傳統機率生成器一般都需要進行馬可夫鏈式的取樣和推斷，而生成式對抗網路避免了這個計算複雜度特別高的過程，直接進行取樣和推斷，從而提高了生成式對抗網路的應用效率，在實際應用上也就更為廣泛。

(ii)生成式對抗網路是一個非常靈活的設計框架，各種類型的成本函數(cost function)都可以整合到生成式對抗網路模型當中，這樣使得針對不同的任務，我們可以設計不同類型的損失函數，都會在生成式對抗網路的框架下進行學習和最佳化。

(iii)最重要的一點是，當機率密度不可計算的時候，傳統依賴於資料自然性解釋的一些生成器就不可以在上面進行學習和應用。但是生成式對抗網路在這種情況下依然可以使用，這是因為生成式對抗網路引入了一個非常聰明的內部對抗的訓練機制，可以逼近一些不是很容易計算的目標函數。

三. 半監督生成對抗網路

半監督生成對抗網路由一個生成器 G 和鑑別器 D 組成，且二者均為卷積神經網路。半監督生成對抗網路通過在大量無標籤資料的基礎上加入少量標籤資料，不僅實現了生成對抗網路鑑別器作為分類器以對輸入資料進行正確分類，而且在大量無標籤資料的輔助下，避免了傳統分類方法在少量標籤

資料上產生過擬合(overfitting)問題。原始生成對抗網路鑑別器 D 的最後一層是以 Sigmoid 單元作為輸出層，限制了網路只能輸出(0, 1)的概率值，即輸出的值接近於 1 表示輸入資料為真實資料，接近於 0 則表示輸入資料為生成器 G 生成的資料。這種概率值很重要，因為通過它，鑑別器 D 能夠將梯度回饋給生成器 G ，從而可以提高 G 生成逼真圖像的能力。但半監督生成對抗網路將 Softmax 函數作為鑑別器 D 的最後一層，讓鑑別器可以輸出 $k+1$ 個類別，其中 k 為原始資料集類別數目，而第 $k+1$ 類則代表生成器生成的資料為“假(Fake)”，即鑑別器 D 輸出的 $k+1$ 個單元對應為 $[class\ 1, class\ 2, \dots, class\ k, fake]$ 。此時，鑑別器 D 也充當了分類器的角色，並且具有三種不同的輸入資料來源用於訓練，即為真實無標籤資料和生成器生成的資料。一方面，鑑別器 D 通過學習區分真實無標籤資料和生成的資料，可以很好地捕獲到真實無標籤資料具有的特定特徵，並且能夠說明生成器學習生成逼真的圖像，這是通過半監督生成對抗網路的對抗訓練完成的;在另一方面，鑑別器 D 使用真實有標籤資料進行監督訓練，在能夠正確區分真實數據和生成資料的基礎上，進一步精確地對資料集類別進行分類。在訓練過程中，鑑別器 D 實際上同時進行了無監督學習和有監督學習，可以將其稱之為 D/C 網路。半監督生成對抗網路的流程圖如圖 2 所示。該演算法主要用於故障診斷領域，因此選取故障種類作為條件變數與生成器進行連接。為了增強條件變數的指導作用將故障標籤與生成器每一層的輸入連接後作為整體輸入到下一層網路中。由於鑑別器執行識別功能而不能引入條件變數干擾模型的識別能力。根據故障診斷的特點以及條件變數的影響對損失函數進行最佳化，在生成器損失函數中增加了條件變數 C 以達到更好的訓練效果。

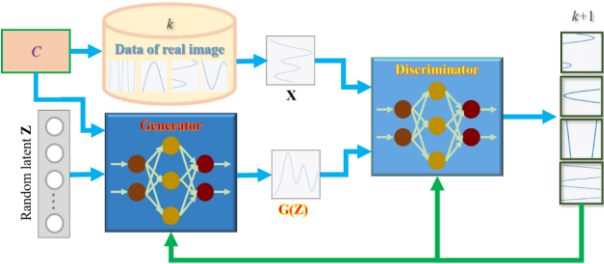


圖 2 改進生成式對抗網路方塊圖

有標籤資料、真實無生成器屬於生成式模型與其他模型不同之處在於生成器不需要預先建模，而是通過與鑑別器的對抗過程不斷生成逼近真實資料的生成資料，然而由於缺少限制條件導致生成資料過於自由。針對上述問題本文提出一類改進生成對抗網路演算法在半監督生成對抗網路演算法的基礎上添加條件變數(conditional variable)使用額外資訊對模型增加條件指導生成器和鑑別器的訓練過程。相比於原型生成對抗網路，主要區別在於鑑別器輸出一個 $k+1$ 的類別訊號（生成的樣本為第 $k+1$ 類）。對於鑑別器，損失函數(loss function) 包括兩部分，一個是監督學習損失（只需要判斷樣本真假），另一個是無監督學習損失（判斷樣本類別）。生成器則只需要儘量生成逼真的樣本即可。訓練完成後，鑑別器就可以作為一個分類模型去分類。圖 3 所示為半監督生成對抗網路方塊圖，圖中生成器 G 將隨機雜訊轉換為“偽(Fake)”示例。鑑別器接收帶有標籤的真實圖像 (x,y) ，沒有標籤的真實圖像 x 和由生成器生成的假圖像 x^* 。為了區分真假例子，鑑別器使用了 sigmoid 函數當輸出。為了區分真實類，鑑別器使用 softmax 函數當輸出。

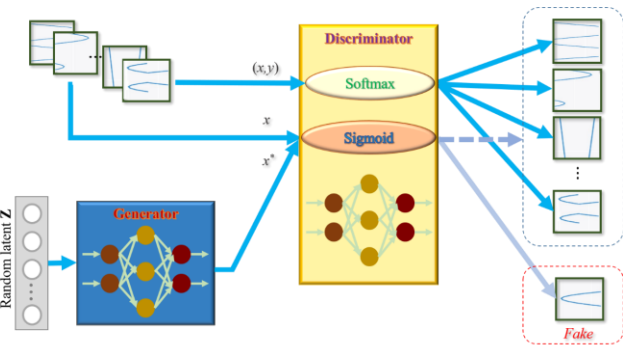


圖 3 半監督生成對抗網路方塊圖

表 1 半監督生成對抗網路整體結構描述

鑑別器	1.由 3 個網路結構組成，每個網路由 1 個多層感知器卷積層和 1 個最大池化層組成，多層感知器卷積層用於提取圖像特徵，最大池化層用於將圖像大小下採樣到輸入時的一半。 2.除輸出層使用 Softmax 啟動函數外，其他層均使用 Leaky Relu 啟動函數。 3.只在第二層和第三層上使用 Batch Normalization，因為不在第一層使用能夠增加模型的穩定性，不在最後一層使用是因為鑑別器的最後一層會作為生成器特徵匹配中的嵌入層。 4.第一層和第三層上使用 Dropout 可以有效減少網路參數，防止模型過擬合。 5.將其設置成一個最後一層使用 Softmax 啟動函數的 $k+1$ 類分類器，用於將真實資料正確地分類為 $k+1$ 類中的類別，將生成器生成的資料分類為第 $k+1$ 個代表“假”的類別。
生成器	1.使用 3 個卷積(反卷積)學習空間的上取樣。 2.除輸出層使用 Tanh 啟動函數外,其他層均使用 Relu 激勵函數(activation function)。 3.除最後一層外每一層均使用 Batch Normalization。

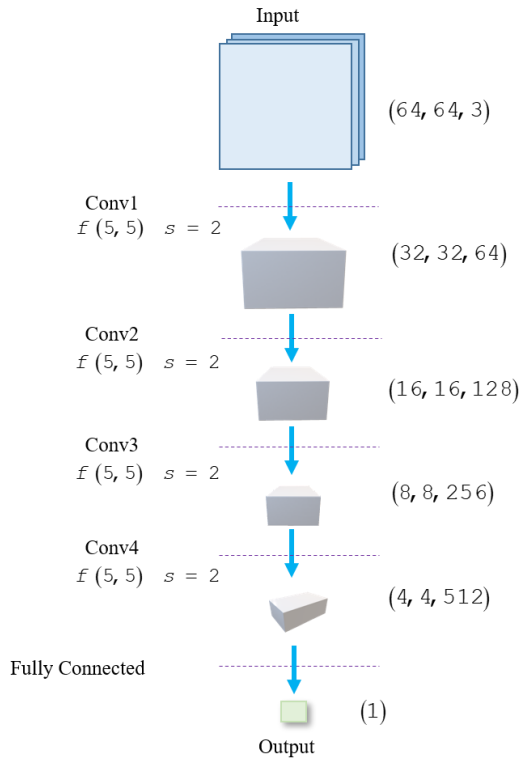


圖 4(a) 半監督生成對抗網路生成器方圖

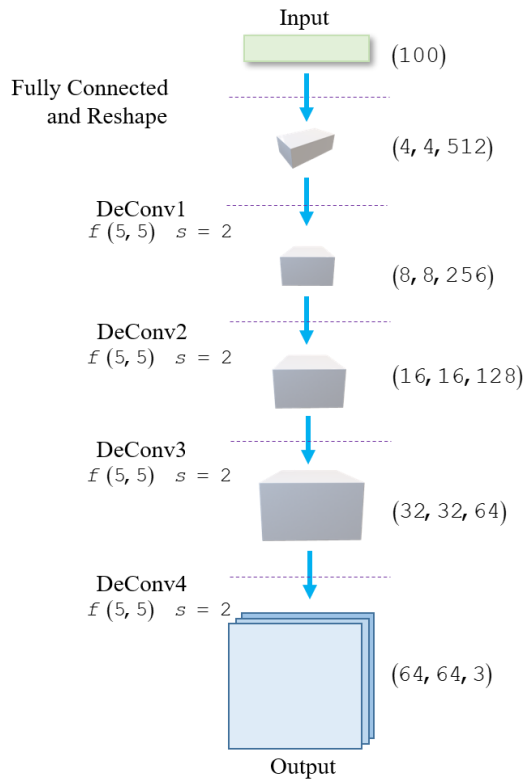


圖 4(b) 半監督生成對抗網路鑑別器方圖

四、故障模型小波變換

微衛星姿態故障診斷方法，深度卷積生成對抗網路的以故障信號作為主要依據，結合連續小波變換處理。小波轉換對於非平穩信號的能力和半監督生成對抗網路處理和識別圖像的功能，在半監督生成對抗網路的基礎上引入模型並對損失函數進行最佳化，最佳化過程中指導生成器和鑑別器的訓練。並利用其生成器和半監督學習能力分別解決了樣本資料不足和樣本標籤問題。模擬實驗表明，連續小波變換與深度卷積生成對抗網路結合的故障診斷方法與其他主流診斷方法相比能達到較高準確率。

● 連續小波變換(CWT)

對於任意的 $L^2(\bullet)$ 空間中的函數 $f(t)$ ，其CWT的定義為

$$\text{CWT}\{(a, \tau)\} = [f(t), \rho_{a,\tau}(t)]$$

$$= \frac{1}{\sqrt{a}} \int f(t) \rho^* \left(\frac{t-\tau}{a} \right) dt \quad (7)$$

$$\rho_{a,\tau}(t) = \frac{1}{\sqrt{a}} \phi \left(\frac{t-\tau}{a} \right) \quad a, \tau \in \mathbf{R}, a > 0 \quad (8)$$

式中： a 為尺度因數； τ 為平移因數； $\rho_{a,\tau}(t)$ 為小波基函數[9]。

通過對信號進行CWT，相當於將原來的一維信號投影到二維平面，CWT的時頻圖上同時反映了振動信號的頻率隨時間變化的關係。

Case 1:

$$f_{a_1} = \begin{cases} 0.5, & 40 \leq t < 70(\text{sec}) \\ 0 & \end{cases}$$

$$f_{a_2} = f_{a_3} = 0 \quad (9)$$

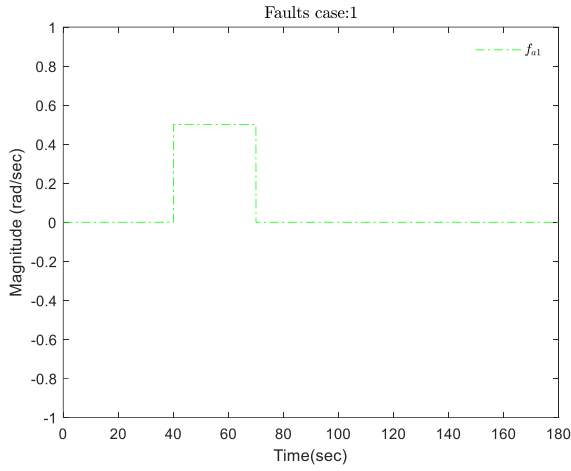


圖 5 故障樣本 case 1

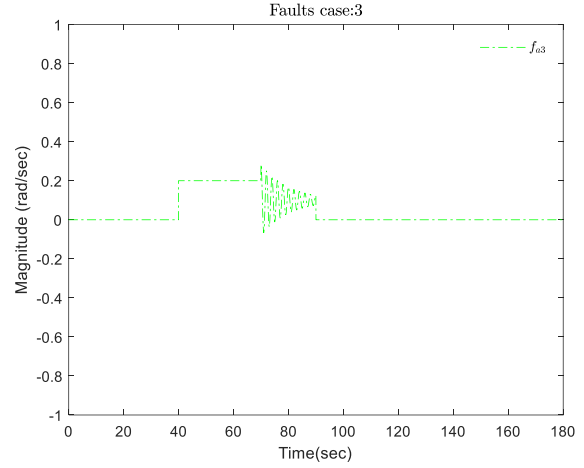


圖 7 故障樣本 case 3

Case 2:

$$f_{a_2} = \begin{cases} 0.1 + 0.3 \cos(\pi(t-40))e^{-0.1(t-40)}, & 40 \leq t < 70(\text{sec}) \\ 0 & \text{otherwise} \end{cases}$$

$$f_{a_1} = f_{a_3} = 0$$

(10)

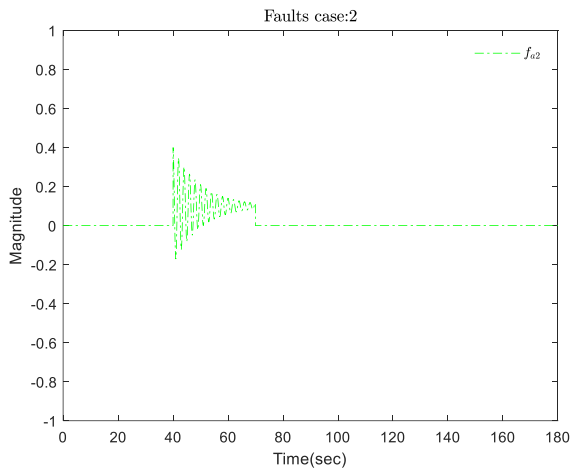


圖 6 故障樣本 case 2

Case 3:

$$f_{a_3} = \begin{cases} 0.2, & 40 \leq t < 70(\text{sec}) \\ 0.1 + 0.5 \cos(\pi(t-70))e^{-0.1(t-60)}, & 70 \leq t < 80(\text{sec}) \\ 0, & t \geq 80(\text{sec}) \end{cases}$$

$$f_{a_1} = f_{a_2} = 0$$

(11)

將故障信號轉化為頻率—時間—小波時頻圖，達到對振動信號的時頻分析，獲得更加全面細緻的資訊，利用深度生成對抗網路的特徵提取以及圖像識別能力，以壓縮後的時頻圖作為輸入，既能降低特徵圖各維的大小，又能提高訓練速度。相對應姿態故障 case 1~case 3 小波轉換時頻圖如圖 8、圖 9 和圖 10 所示。小波反轉換原信號與重構信號比較如圖 11、圖 12 和圖 13 所示。

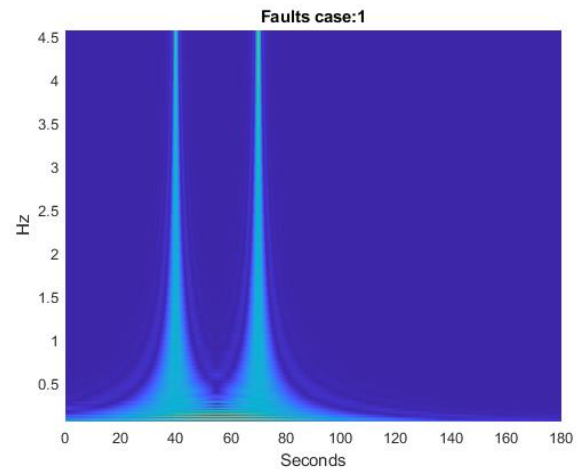


圖 8 姿態故障 case 1 小波轉換時頻圖

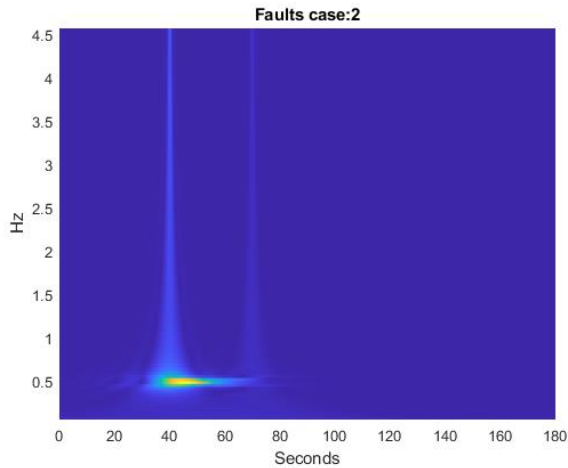


圖 9 姿態故障 case 2 小波轉換時頻圖

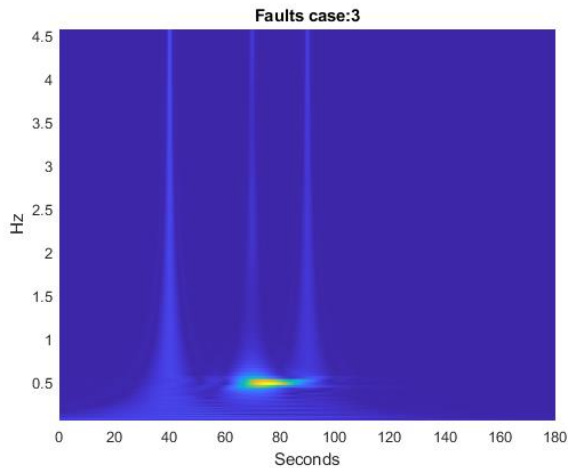


圖 10 姿態故障 case 3 小波轉換時頻圖

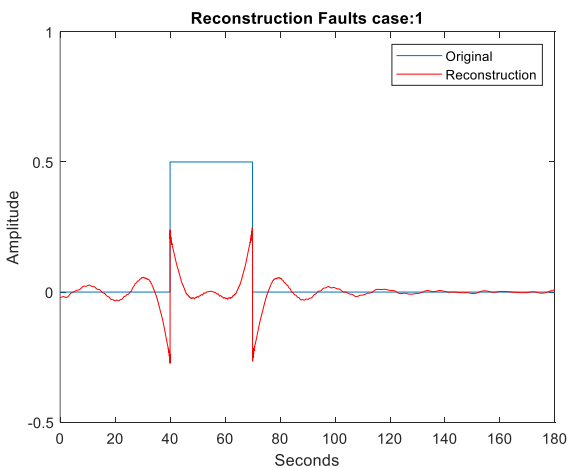


圖 11 姿態故障 case 1: 小波反轉換原信號與重構信號比較圖

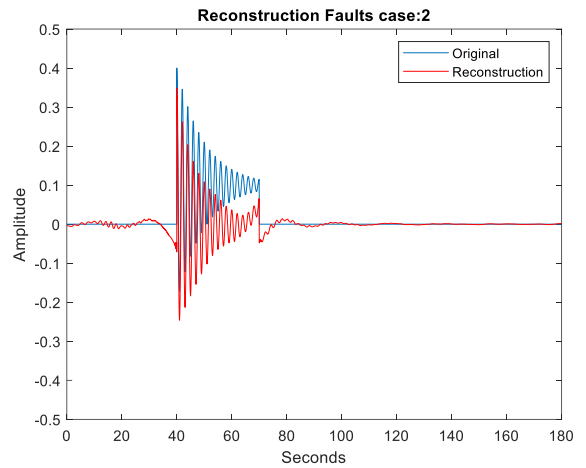


圖 12 姿態故障 case 2: 小波反轉換原信號與重構信號比較圖

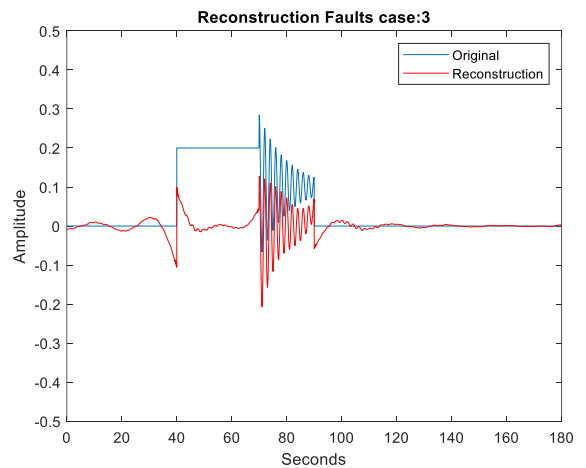


圖 13 姿態故障 case 3: 小波反轉換原信號與重構信號比較圖

經過資料預處理後，可以得到每種狀況下的二維時頻圖各 1800 個，選取其中的 75% 作為訓練樣本，其餘 25% 作為測試樣本。在讀取樣本圖片時，對不同的故障類型進行標籤，原始標籤為一位元數位，為了便於定義和使用。

五、損失函數計算

半監督生成對抗網路的鑑別器同時也充當了一個 $k+1$ 分類器的角色，它接收一個數據點作為輸入，輸出一個 $k+1$ 維的邏輯向量 $\{c_1, c_2, \dots, c_{k+1}\}$ 。這些邏輯向量可以轉換成類

概率

$$p_{\text{model}}(y = k+1 | x) = \frac{\exp(c_{k+1})}{\sum_{j=1}^{k+1} \exp(c_j)} \quad (12)$$

$$p_{\text{model}}(y = i | x, i < k+1) = \frac{\exp(c_i)}{\sum_{j=1}^{k+1} \exp(c_j)} \quad (13)$$

其中，式(12)表示 x 為“假”的概率，而公式(13)表示 x 為“真”且屬於類 i 的概率。

半監督生成對抗網路的損失函數就可以由「標記損失」和「無標記損失」兩部分組成，如式(14)所示。其中式(15)表示標記損失，即分類損失；式(16)表示無標記損失，即對抗損失，可以看出無標記損失是標準的生成對抗網路博弈值，與傳統生成對抗網路的鑑別器損失函數相同。

$$L_D = L_{\text{labeled}} + L_{\text{unlabeled}} \quad (14)$$

$$L_{\text{labeled}} = -E_{x, y \sim p_{\text{data}}} \log[p_{\text{model}}(y = i | x, i < k+1)] \quad (15)$$

$$L_{\text{unlabeled}} = E_{x \sim p_{\text{data}}} \log[1 - p_{\text{model}}(y = k+1 | x)] + E_{x \sim G} \log[p_{\text{model}}(y = k+1 | x)] \quad (16)$$

生成器 G 的損失函數由對抗損失和特徵匹配損失兩部分組成，如式(14)所示。公式(15)與傳統 GAN 相同，表示生成器在與鑑別器進行對抗訓練中學習到的對抗損失。特徵匹配由文獻[7]提出，其原理在於通過為生成器指定一個新的目標來防止它在當前鑑別器上過度訓練，從而解決 SGAN 訓練不穩定的問題。新的目標不是最大化鑑別器的輸出，而

是要求生成器生成與真實資料的統計資訊相匹配的數據，即訓練生成器匹配鑑別器中間層特徵的期望值。對於生成器來說，這是一個自然的統計資料選擇，因為通過訓練鑑別器，可以很自然地找到那些最能區別真實資料和當前模型生成資料的特徵。特徵匹配的損失如式(16)所示，其中 $f(x)$ 表示鑑別器 D 中間層上的啟動。

$$L_G = L_{\text{adversarial}} + L_{\text{feature-matching}} \quad (17)$$

$$L_{\text{adversarial}} = -E_{x \sim G} \log[1 - p_{\text{model}}(y = N+1 | x)] \quad (18)$$

$$L_{\text{feature-matching}} = \|E_{x \sim p_{\text{data}}} f(x) - E_{x \sim G} f(x)\|_2^2 \quad (19)$$

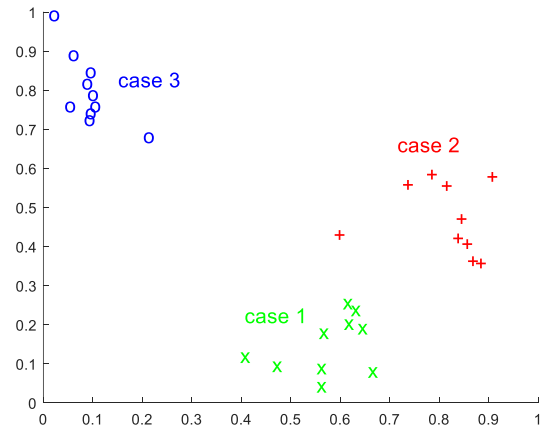


圖 14 經損失函數最佳化後生成對抗網路的鑑別判斷

六、小腦神經網路

小腦神經網路在控制領域又稱『小腦模型控制器』，是由一系列的映射所定義，敘述如下：

(1)小腦神經網路映射

小腦模型控制器： $Q \rightarrow Iv \rightarrow Hb \rightarrow Y$

其中 Q 為輸入向量， Iv 為內部變數的集合， Hb 為超立方體(Hypercube)的集合， Y 為輸出向量。

(i) $Q \rightarrow Iv$ ：輸入向量

$q = [q_1 \ q_2 \ \dots \ q_n]^T$ 是由 n 個變數所組成，

每一個變數 q_i 可以不同的方法分解為數個區段(Block)。每一個區段又包含許多個元素(Element)。不同的分區段分割方式利用每一個區段位移一個或多個元素，則可達到除了第一及最後一個區段外，其它區段的元素數目均相同。在圖 15 中的兩個狀態變數有 q_1 和 q_2 均被量化成 17 個間隔(元素)。每一個變數有 4 種不同的區段分割方法，每一種方法都將狀態變數的區域分成 5 個區段，每一個完整大小的區段由 4 個元素所構成，如圖 10 表示，若 $q_1=6$ 對映到集合

$I_{v1} = \{B, H, M, Q\}$ 及 $q_2=5$ 對映到集合

$I_{v2} = \{b, h, l, q\}$ 。

(ii). $Iv \rightarrow Hb$ ：內部變數是對映到一起超立方體(Hypercube)的集合，每一個超立方體都有一個指定的內部變數的組合(名稱)。Hb 定義為一起超立方體的集合。以圖 15 說明，

$I_{v1} = \{B, H, M, Q\}$ 及 $I_{v2} = \{b, h, l, q\}$ 都對映

到 $H_b^* = \{Bb, Hh, Ml, Qq\}$ ，其中 Bb ， Hh ， Ml

和 Qq 指定了 4 個超立方體。從這種分割的方式來看，我們可以想像有 N_e 層超立方體，其中 N_e 為一完整區段內之元素(Element)數目。超立方體的形成是由量子化的方法來決

定。每一輸入狀態被一層一個共有 N_e 個不同的超立方體所覆蓋，如圖 10 所示。

(iii). $Hb \rightarrow Y$ ：每一個原始小腦模型控制器的輸出會因超立方體選擇向量與記憶體內容值向量的不同造成不同的輸出，其輸出向量可表示成

$$y = \sum_{i=1}^{n_R} a_i w_i = \mathbf{a}^T \mathbf{w} \quad (20)$$

其中 $\mathbf{a} = [a_1 \ a_2 \ \dots \ a_{n_R}]^T$ 是超立方體選擇

向量， $\mathbf{w} = [w_1 \ w_2 \ \dots \ w_{n_R}]^T$ 是記憶體內容值向量，其中 n_R 是超立方體的數目。而且 a_i 只能是 1(被選) 或 0(不被選)。在圖 15 中，狀態(6,5)之輸出值等於超立方體 Bb ， Hh ， Ml

和 Qq 所對應到記憶體內容值的總和。

對於一個輸入向量而言，一個完整區段的元素個數決定了層數和一般泛化(Generalization)的程度。在小腦模型控制器中，一個超立方體相當於一個記憶體空間。一個離散狀態的資料是由儲存於涵蓋在此狀態的超立方體(也就是其對應記憶體)的內容值。

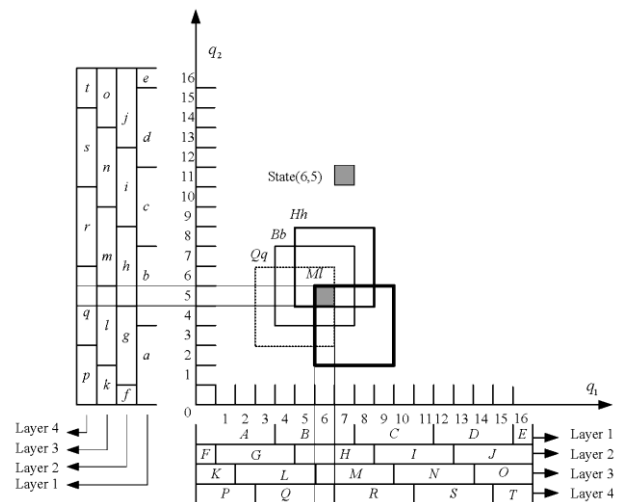


圖 15 雙輸入變數小腦模型控制器之區段分

割平面圖[14]

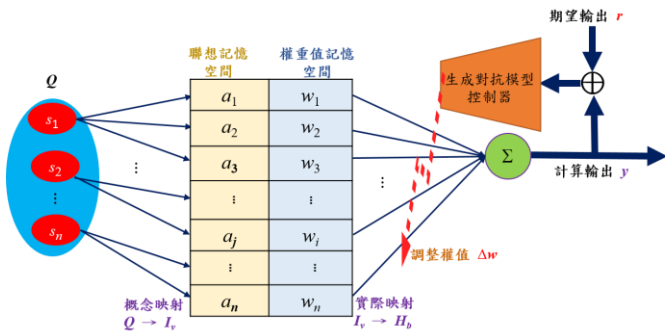


圖 16 基於小腦模型控制器生成式對抗網路基本架構

(2)小腦模型控制器的學習演算法

小腦模型控制器的學習演算法為：

若採用 δ 學習規則，權值調整指標

$$E = \frac{1}{2c} e^2(t) \quad (21)$$

式中， $e(t) = r(t) - y(t)$ 。

若採用在梯度下降法下，權值調整

$$\Delta w_j(t) = -\eta \frac{\partial E}{\partial w} = \eta \frac{e(t)}{c} \quad (22)$$

$$w_j(t) = w_j(t-1) + \Delta w_j(t) + \alpha(w_j(t-1) - w_j(t-2)) \quad (23)$$

上式中， η 為小腦模型控制器學習速率， α 為慣性係數。

七、微衛星姿態故障診測與容錯控制模擬
微衛星動力學方程以非線性方程表示

$$J\dot{\omega} = -\omega^\times J\omega + \tau_w + \tau_u \quad (24)$$

其中 ω 為衛星在體座標系中的角速度， J 是衛星的對稱慣性矩陣， τ_u 控制力矩向量和 τ_w 表示外部干擾力矩。

對於小姿態角，(16)式推導結果如下

$$\begin{cases} \dot{x}(t) = Ax(t) + g(t, x) + B_u u(t) + B_w w(t) + F_a f_a(t) \\ y(t) = Cx(t) + F_s f_s \end{cases} \quad (25)$$

其中狀態向量 $x = [\psi \ \theta \ \phi \ \dot{\psi} \ \dot{\theta} \ \dot{\phi}]^T$ ，

$$g(t, x) = \begin{bmatrix} 0 \\ \omega_0 \\ 0 \\ (J_z - J_y)\omega_y\omega_z/J_x \\ (J_x - J_z)\omega_x\omega_z/J_y \\ (J_y - J_x)\omega_x\omega_y/J_x \end{bmatrix},$$

控制輸入矩陣和外擾輸入矩陣

$$B_u = B_w = \begin{bmatrix} 0_{3 \times 3} \\ J^{-1} \end{bmatrix},$$

輸出矩陣 $C = I_6$

$$\text{制動器故障矩陣 } F_a = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T,$$

$$\text{感測器故障矩陣 } F_s = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T,$$

制動器故障向量 $f_a = [f_{a_1} \ f_{a_2} \ f_{a_3}]^T$ 和

感測器故障向量 $f_s = [f_{s_1} \ f_{s_2} \ f_{s_3}]^T$ 。

以前述資料與推導結果，數值模擬微衛星姿態故障診測與容錯控制。

表 2 微衛星模擬初始條件與制動器故障向量

	初始條件 $[\phi_0 \quad \theta_0 \quad \psi_0]$ $= [0.3491 \quad 0.1745 \quad -0.2618] rad.$ $[\dot{\phi}_0 \quad \dot{\theta}_0 \quad \dot{\psi}_0]$ $= [0.0316 \quad 0.0424 \quad 0.056] rad / sec$	數值模擬結果
Case 1	$f_{a_1} = \begin{cases} 0.5, & 40 \leq t < 70 (sec) \\ 0 & \end{cases}$ $f_{a_2} = f_{a_3} = 0$	圖 18(a)、18(b)、18(c)、18(d)
case 2	$f_{a_2} = \begin{cases} 0.1 + 0.3 \cos(\pi(t-40))e^{-0.1(t-40)}, & 40 \leq t < 70 (sec) \\ 0 & \end{cases}$ $f_{a_1} = f_{a_3} = 0$	圖 19 (a)、19(b)、19(c)、19(d)
case 3	$f_{a_3} = \begin{cases} 0.2, & 40 \leq t < 70 (sec) \\ 0.1 + 0.5 \cos(\pi(t-70))e^{-0.1(t-60)}, & 70 \leq t < 80 (sec) \\ 0, & t \geq 80 (sec) \end{cases}$ $f_{a_1} = f_{a_2} = 0$	圖 20(a)、20(b)、20(c)、20(d)

$$J = \text{diag}\{18, 21, 24\}$$

$$x_0 = [0.3491 \quad 0.1745 \quad -0.2618 \quad 0.0316 \quad 0.0424 \quad 0.056]$$

$$w(t) = \begin{bmatrix} A_0(3\cos(\omega_0 t) + 1) \\ A_0(1.5\cos(\omega_0 t) + 3\cos(\omega_0 t)) \\ A_0(3\cos(\omega_0 t) + 1) \end{bmatrix},$$

其中 $A_0 = 1.5 \times 10^{-5}$ ，軌道角速度

$$\omega_0 = 0.0012 \text{ rad/s}。$$

將以三種不同制動器故障向量討論深度生成對抗網路故障偵測與小腦神經網路演算法應用於故障容錯

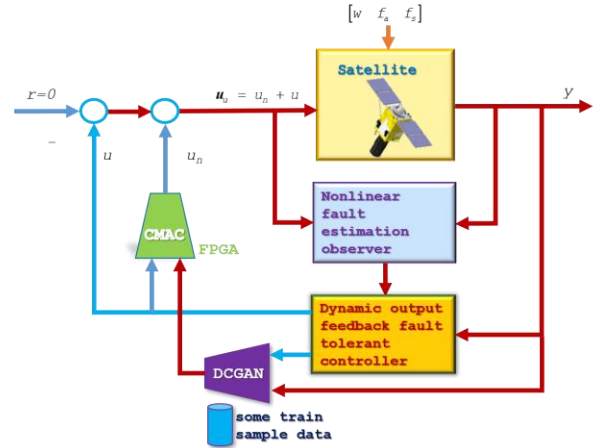


圖 17 微衛星姿態故障偵測與容錯控制基於生成式對抗網路與小腦模型控制器架構

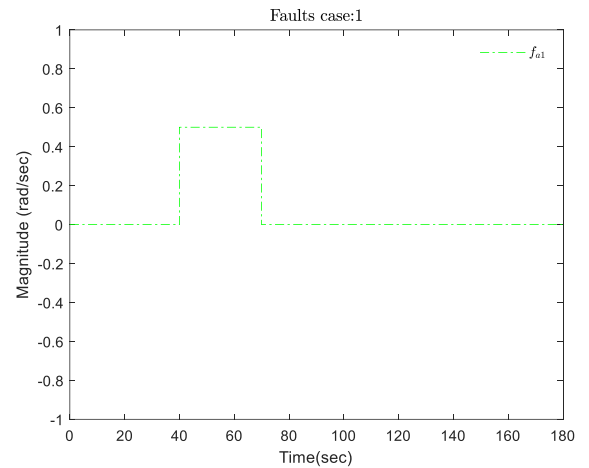


圖 18(a) 故障樣本 case 1

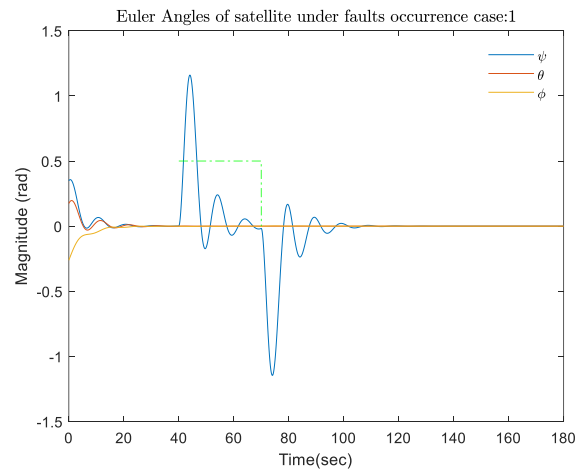


圖 18(b) 姿態角故障與復歸響應 case 1

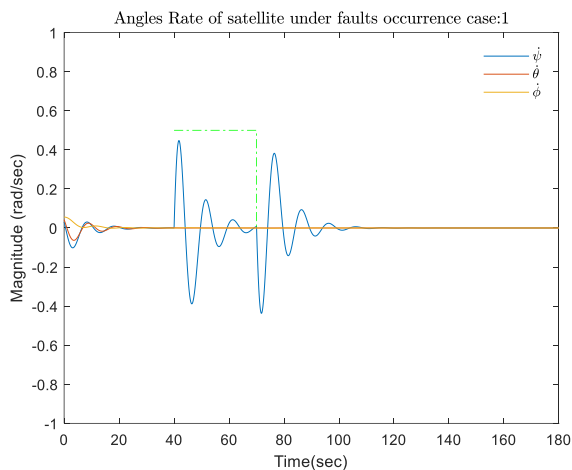


圖 18(c) 姿態角速率故障與復歸響應 case 1

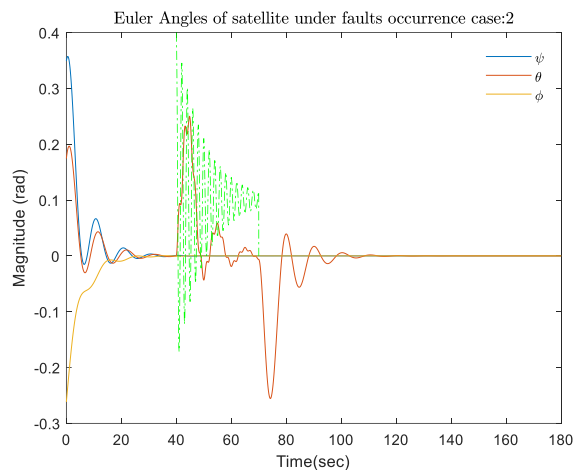


圖 19(b) 姿態角故障與復歸響應 case 2

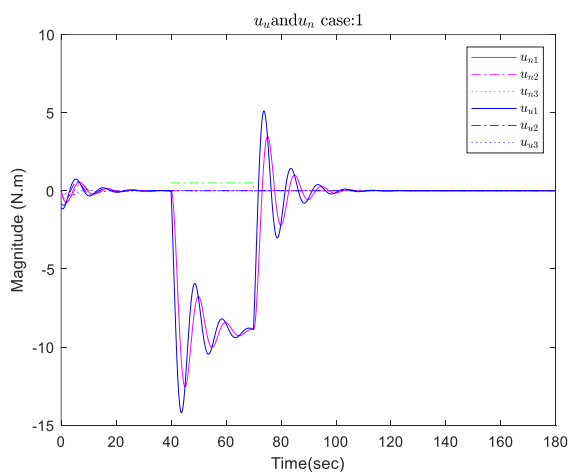


圖 18(d) 故障與復歸控制與容錯力矩響應 case 1

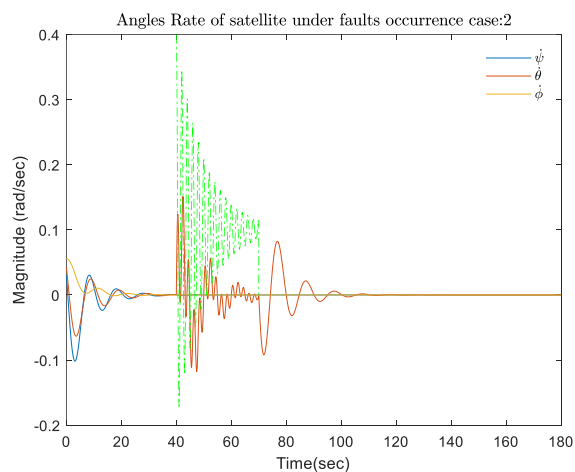


圖 19(c) 姿態角速率故障與復歸響應 case 2

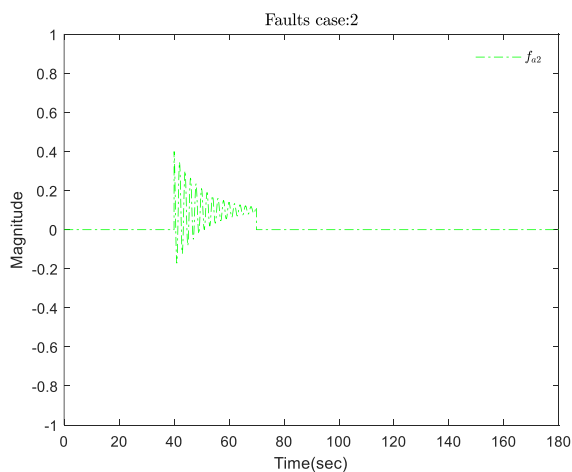


圖 19(a) 故障樣本 case 2

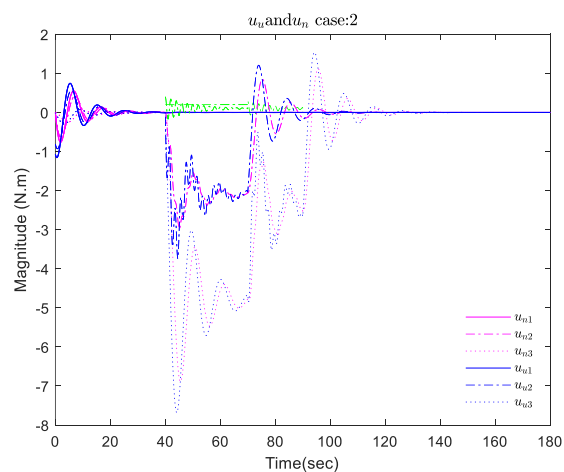


圖 19(d) 控制與容錯力矩故障與復歸響應

case 2

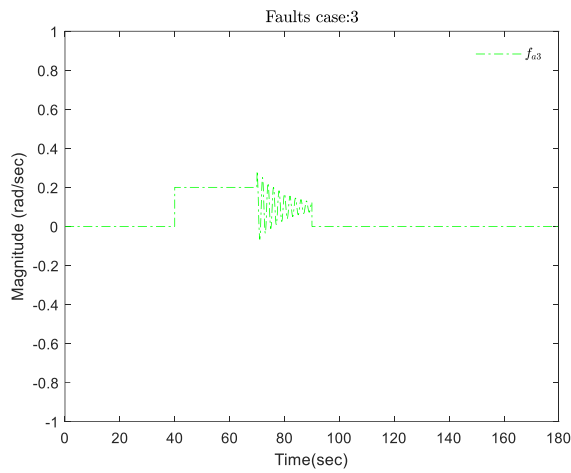


圖 20(a) 故障樣本 case 3

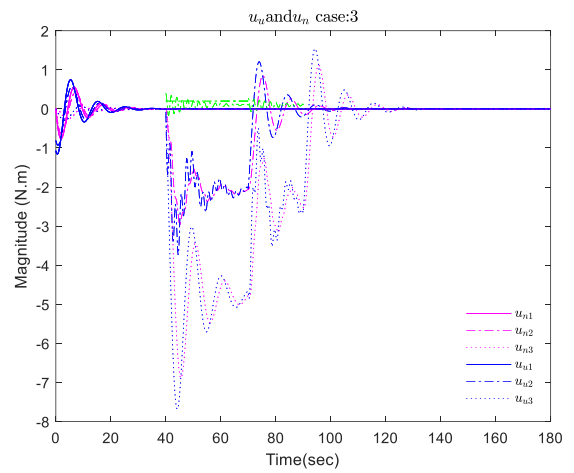


圖 20(d) 控制與容錯力矩故障與復歸響應 case 3

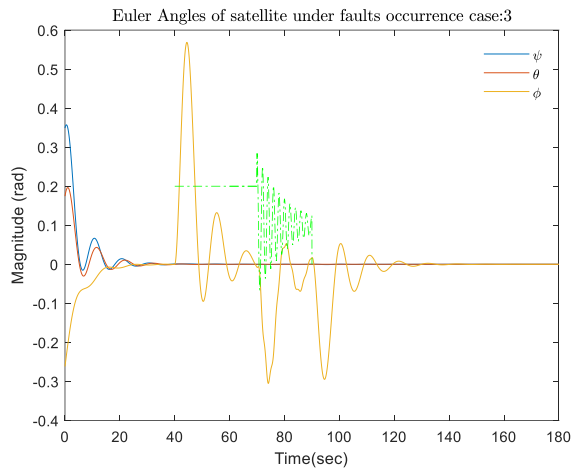


圖 20(b) 姿態角故障與復歸響應 case 3

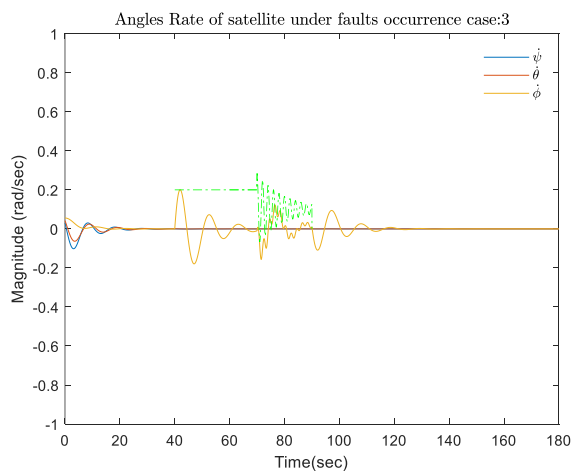


圖 20(c) 姿態角速率故障與復歸響應 case 3

八、 結果與討論

由上節數值模擬結果得知，半監督生成對抗網路運用於微衛星故障診斷的可行性。小腦網路透過半監督生成對抗網路樣本以及訓練參數，在制動器故障模型下可以發現有很好的泛化能力和穩定性，在樣本涉及不同的制動器故障模型，在取得不同模型仍達到良好的診斷結果模擬結果顯示平均準確率為 90%，可以得出半監督生成對抗網路演算法能達到良好的效果。

半監督生成對抗網路上的鑑別器是由線性卷積層組成，其抽象層次較低，不能提取圖像更深層次的特徵，從而不能完整學習到各個類別之間的差異性，因此分類精度不高。擬採用多層感知器卷積層代替半監督生成對抗網路鑑別器 D 上的線性卷積層對底層概念的每個局部特徵做一個更好的抽象，並接收大量無標記資料、少量標記資料，以及生成器生成的資料以提高正確分類率，解決了監督分類中因標記資料太少產生的模型泛化能力不強的問題；在生成器上使用特徵匹配，能

夠將生成資料分佈儘量與真實資料分佈拉近，進一步促進了鑑別器提取圖像不同類別特徵的能力。期望未來實驗結果，多層感知器卷積層的分類精度及其生成的圖像品質都有明顯提高。

誌謝

本文研究感謝感謝 108 年度國防部補助軍事院校教師（官）從事學術研究計畫編號國人培字第 1080017093 號之補助，使理論推導、實驗與數值模擬得以完成。同時並感謝共同主持人對實驗提供寶貴資料與協助。

參考文獻：

- [1] Yang Z Gao L. “Wavelet analysis and fault feature extraction of rolling bearing,” Engineering Asset Management 2011, Springer .
- [2] Jawade A Paraskar S Jadhav S., “Artificial neural network based induction motor fault classifier using continuous wavelet transform,” Systems Science & Control Engineering, vol. 2, no. 1, pp.684-690, 2014.
- [3] Yuan Jian-hu, Han Tao, Tang Jian, et al. “An approach to intelligent fault diagnosis of rolling bearing using wavelet time-frequency representations and CNN,” Machine Design and Research, vol.33, no.2, pp.93-97, 2017.
- [4] Xu Jiu-qiang Hong Li-ping Zhu Hong-bo et al., “Generative adversarial networks for the classification of lung nodules malignant,” Journal of Northeastern University Natural Science, Research, vol.33, no.2, pp.93-97, 2017.
- [5] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville and Yoshua Bengio, “Generative adversarial nets,” In Advances in Neural Information Processing Systems(NIPS), pp.2672-2680, 2016.
- [6] Salimans T., Goodfellow I., and Zaremba W., “Improved techniques for training GANs,” In NIPS, pp. 2234–2242. 2016.
- [7] Salimans T, Goodfellow I, Zaremba W, et al., “ Improved techniques for training GANs,” Advances in Neural Information Processing Systems, pp. 2234-2242, 2016.
- [8] Lantao Yu, Weinan Zhang, Jun Wang, Yong Yu, “SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient,” 31th AAAI Conference on Artificial Intelligence, pp.2852-2858, 2017.
- [9] U Ahsan, C Sun and I Essa, “Discrim Net: Semi-Supervised Action Recognition from Videos using Generative Adversarial Networks,” arXiv preprint, arXiv:1801.07230, 2018.
- [10] Levie R, Sochen N. “A wavelet Plancherel theory with application to sparse continuous wavelet transform,” <https://arxiv.org/abs/1712.02770>, 2018-08-17.
- [11] Odena A. “ Semi-supervised learning with generative adversarial networks,” arXiv preprint arXiv:1606.01583, 2016.
- [12] Dai Z, Yang Z, Yang F, et al., “Good semi-supervised learning that requires a bad GAN,” Advances in Neural Information Processing Systems, pp.6510-6520, 2017.
- [13] Qingbo He, Enhao Wu, Yuanyuan Pan, “Multi-Scale Stochastic Resonance Spectrogram for fault diagnosis of rolling element bearings,” Journal of Sound and Vibration, no.420, pp.174-184, 2018.
- [14] C. M. Lin, Y. F. Peng and C. F. Hsu, “Robust cerebellar model articulation controller design for unknown nonlinear systems,” IEEE Trans. on Circuit and systems: II: Express Briefs, vol. 51, no. 7, pp.354-358, 2004.