

基於深度學習分析手部動作腦波訊號

¹ 陳曹倫^{2*} 李建誠

¹ 元智大學 電機工程學系

a870128rlen@gmail.com

² 元智大學 電機工程學系

*通訊作者: cclee@saturn.yzu.edu.tw

摘要

本論文提出一種基於卷積類神經網路的腦波訊號動作分析方法；不同於傳統濾波後計算特定的特徵，而是用學習的方式尋找有效的特徵並分類。本論文利用非侵入式的腦波耳機擷取左右手運動時產生腦波訊號，經由經驗模態分解並重組數個頻段訊號，將其轉換成二維陣列交由卷積類神經網路擷取特徵，並分類。本論文嘗試使用不同網路與收集不同狀態下的資料來進行分析，最終在 VGG16 可以達到 72% 的準確度。

關鍵詞：卷積類神經網路、腦機介面、經驗模態分解

Abstract

In this paper, a method of brainwave signal motion analysis based on convolutional neural network is proposed. Different from the traditional analysis, the proposed method uses the deep learning to find effective features and builds the analysis model. This study uses non-invasive brain-machine interface (BCI) to collect brainwave signals when capturing left and right hand movements. It decomposes and recombines several frequency band signals through empirical mode decomposition (EMD), and converts them into two-dimensional arrays to be extracted and classified by convolutional neural networks. This study attempts to use different networks and collect data from different states for analysis, and finally achieves 72% accuracy in VGG16.

Keywords: Convolutional Neural Networks, Brain Machine Interface, Empirical Model Decomposition

一、簡介

大腦是掌管人類行為與感知的中樞，例如學習、運動、認知等。大腦有人體控制的最高權限，負責整個生命體發布命令，所以相較其他的器官而言格外重要。人類在運動前，大腦的皮質區會先產生執行的訊號，經由皮質脊髓束傳至脊髓運動神經元，再由脊髓運動神經元軸突所構成的周邊神經傳至肌肉，使肌肉縮。而這複雜的傳遞過程，很容易因為外部刺激、或生病，造成神經或肌肉功能受損，導致人們的行動能力受損障礙造

成生活的不便。

生理訊號相關裝置，近年來已經深入消費者市場中了，例如智慧手環，智慧椅墊等等。人體的生理訊號反映使用者的健康，漸漸地被大多數人開始重視了。在醫學方面，生理訊號常被應用在分析與紀錄，來了解許多細微的訊息，使醫生能更快速地得知病情。這些訊號包含了腦電圖(Electroencephalography, EEG)、肌電圖(Electromyography, EMG)、心電圖(Electrocardiography, ECG)。

對於那些肌肉控制能力喪失或脊髓束等相關傳遞神經出問題的病患，常常令病患家屬深感無能為力，加上醫院人力時有不足，病患出狀況時，醫院往往無法第一時間作處理，錯失最佳治療的時刻，如果此類病患能利用腦波控制的方式來與外界溝通，必能減少許多的遺憾發生，因此本研究目的在於探討大腦與外界溝通的橋樑—腦機介面(Brain-Computer Interface, BCI)的演算法與實作。

腦機介面近幾年來越來越受到重視，也有越來越多學術單位與公司開始著手研究，它的發展能讓那些，無法自主行動的人(例如中風患者)藉由不同動作觸發特定腦波的電位變化，進而辨識使用者的動作，透過這些研究使得人類不必藉肌肉運動，單憑腦波就能與外界溝通。

在過往的運動想像 BCI 系統中，已經有許多研究提出了各種演算法來提高 BCI 系統的辨識能力，但每個使用者對於相同的動作所產生的想像訊號皆不盡相同，導致誤差的出現，因此才難以實現在市場與醫院中。

本次研究我們將提出不同於以往取人工特徵的方式；而是利用能夠自動學習特徵的深度學習；在影像分類上，卷積類神經網路(Convolutional Neural Network, CNN)有相當好的成果，而影像為二維陣列，因此本計畫，我們擬將所得到的訊號轉換為二維陣列的方式，再交由 CNN 提取特徵並進行分類，最後，將建構出一套模型，讓受測者戴上腦波裝置，並開始右手活動與左手活動這些動作，將電腦判斷的結果呈現出來

二、文獻回顧

(一) 腦機介面

腦波訊號測量透過多個訊號量測電極來擷取不同位置的腦波以代表不同腦區的神經放電現象，但所擷取到的腦波訊號電位是相當的微弱，必須透過放大器進行放大紀錄，以作為分析所需的腦波訊號。

腦波訊號測量分成三種，侵入式、部分侵入式與非侵入式。侵入式訊號測量，是直接將電極植入大腦灰質，可以取得高品質的訊號，但須透過手術才能完成，而且容易引發免疫反應或癒傷組織活動，導致訊號會逐步衰減至消失；部分侵入式，是將電極植入頭顱腔內但於大腦灰質外，訊號品質次等於侵入式，也是需要透過手術才能完成，但是不會引發免疫反應與癒傷組織活動的發生；非侵入式，將電極放置於頭皮上，是最為方便的方法，但是所得到的訊號品質有限，比較適合實際的應用且更具便利性。

腦機介面讓使用者透過腦波發出命令，透過不同腦波特徵分析來達到分類的目的，例如：慢速皮質電位(Slow Cortical Potential, SCPs) [1]、事件相關電位(Event-Related Potential, ERP) [2]、 μ 節律與 β 波的事件相關同步化(Event-Related Synchronization, ERS)與事件相關反同步化(event-Related Desynchronization, ERD) [3,4]等，再結合深度學習，分類不同的腦波訊號。

(二) 類神經網路與深度學習

類神經網路(Artificial Neural Network, ANN)的研究可以追溯到二十世紀 40 年代的心理學家 Warren McCulloch 和數學家 Walter Pitts [5]對人類腦神經元的活動進行理論建模，構建了第一個類神經網路模型，隨後 ANN 經歷了感測模型、自我調整線性元件等神經網路的發展，在許多學科領域引起了廣泛的興趣和研究，在 60 年代末，Marvin Minsky 及 Seymour Papert [6]指出以感知器為代表的網路模型功能十分有限，直到 80 年代 Hopfield [7]網路模型以及誤差反向傳播演算法提出後，ANN 才在世界範圍內迅速的發展。如今 ANN 已經深入研究，已發展了上百種類神經網路模型以及演算法來解決實際問題。其中又以學習方式的不同分成兩種網路模型 [8]。

1. 監督式學習(Supervised Learning):在學習過程中，需給予類神經網路一個訓練範例，其中包含了輸入質和輸出值和目標輸出值。目標輸出值當作監督者的身分，當實際輸出結果與目標輸出結果之間出現誤差，模型透過誤差訊號不斷調整個層級的權重值，以使實際輸出不斷接近目標輸出結果。

2. 非監督式學習(Unsupervised Learning):在學習過程中，只提供輸入值而沒有提供目標輸出值，因此沒有監督者提供任何回饋資訊。網路透過分析輸入值的特徵不斷更新權重值，這個過程稱為自我組織，適用於數據類問題。

近年來，因卷積神經網路 CNN 在影像上的優異結果與 GPU 的平行處理，深度學習開始被多數人重視並開始研究。深度學習主要分成三種型態，深度神經網路(Deep Neural Networks, DNN) [9]、卷積神經網路[10]、與時間遞歸神經網路(Recurrent Neural Network, RNN)[11]，目前的成效相當優秀，如 AlphaGo 與自動駕駛等。卷積神經網路(CNN)在影像處理上有很大的突破，CNN 與傳統的方法不同，不用提取人工特徵，可直接輸入原始影像。CNN 最早的概念是 1962 年 Hubel 和 Wiesel 所提出 [12]，1989 年由 Yann Le Cun 所開發出 CNN [13]，並應用於光學字符識別和計算機視覺上；其中，CNN 有三個概念：局部感知域(local receptive fields)、權重共享(shared weights)和池化(pooling)，在 2012 年，Hinton 的學生 Alex Krizhevsky 提出了深度卷積神經網路模型 AlexNet [14]，它可以算是 LeNet 的一種更深更寬的版本。AlexNet 中包含了幾個比較新的技術點，也首次在 CNN 中成功應用了 ReLU、Dropout 等技術，並使用了 GPU 進行運算加速。

三、方法

(一) 經驗模態分解

本研究使用經驗模態分解(Empirical Model Decomposition, EMD)從腦波儀收集來的訊號，藉此得到各頻率之時域訊號。黃鵬博士(Dr. Norden E. Huang)於 1998 年提出希爾伯特-黃轉換(Hilbert-Huang Transform, HHT) [15]，將欲分析的資料分解成本質模態函數(Intrinsic Mode functions, IMF)，將得到的 IMF 作希爾伯特轉換(Hilbert Transform)後得到其瞬時頻率；而將資料分解為 IMF 的流程我們稱作 EMD。EMD 可以針對不同的複雜訊號，利用訊號自身的極值特性分解出具有不同尺度特性的 IMF 與一個趨勢函數。特別適用於非線性(non-linear)與非平穩(non-stationary)訊號的分析處理。

EMD 將訊號分解成 IMF 與趨勢函數的過程稱為篩選程序(Sifting Process)。EMD 藉由不斷重覆的篩選程序來逐步找出 IMF。以訊號 $x(t)$ 為例，篩選程序流程概述如下：

1. 找出原始訊號 $x(t)$ 中所有波峰波谷，並在波峰與波峰之間以及波谷與波谷利用多項式函數匹配，通常為三次樣條(cubic spline)，將所有波峰連接成上包絡線與所有波谷連接成下包絡線。
2. 求出上下包絡線之平均，得到包絡線的平均值 $m_1(t)$ 。
3. 原始訊號 $x(t)$ 減去包絡線的平均值，得到第一個分量 $h_1(t)$ 。

$$h_1(t) = x(t) - m_1(t). \quad (1)$$

4. 檢查 $h_1(t)$ 是否滿足 IMF 的兩個條件;如不滿足,則回到步驟 1 並且將 $h_1(t)$ 當作原始訊號,進行第二次的篩選,如(2)所示。

$$h_2(t)=h_1(t)-m_2(t). \quad (2)$$

重複篩選 k 次

$$h_k(t) = h_k - h_1(t) - m_k(t). \quad (3)$$

5. 直到符合 IMF 的條件或與前一個原始訊號的標準差小於設定的閾值即可停止,得到第一個本質模態函數 $imf_1(t)$,如(4)所示。

$$imf_1(t) = h_k(t). \quad (4)$$

6. 原始訊號 $x(t)$ 減去 $imf_1(t)$ 可得到殘餘訊號 $r_1(t)$,如(5)所示。

$$r_1(t) = x(t) - imf_1(t). \quad (5)$$

7. 將 $r_1(t)$ 當作新的資料,重新執行步驟 1 至步驟 5,得到新的殘餘訊號 $r_2(t)$ 。如此重複 n 次

$$r_2(t) = r_1(t) - imf_2(t). \quad (6)$$

$$r_3(t) = r_2(t) - imf_3(t). \quad (7)$$

...

$$r_n(t) = r_{n-1}(t) - imf_n(t). \quad (8)$$

8. 當 $r_n(t)$ 為單調函數(monotonic function)或其波形不超過 2 個極值,篩選過程停止, $r_n(t)$ 為殘餘訊號(或稱趨勢訊號),其可以看出原始訊號的趨勢變化。原始訊號 $x(t)$ 可以表示成 n 個 IMF 分量與一個殘餘訊號 $r_n(t)$ 的組合,如(9)所示。

$$x(t) = \sum_{k=1}^n imf_k(t) + r_n(t). \quad (9)$$

9. 最後,原始訊號便可以分解成 n 個 IMF 和一個殘餘訊號。圖 1 為 EMD 流程圖。

(二) 訊號轉二維陣列

以往的方法大多是設置一個截止頻率 (cutoff frequency) 得到一組高低頻;而本文採用的方法是捨棄殘餘訊號,其餘留下來。利用腦波儀收集到的資料經 EMD 分解後皆得到 6 個 IMFs,如圖 2 示意腦波訊號分解出的訊號。

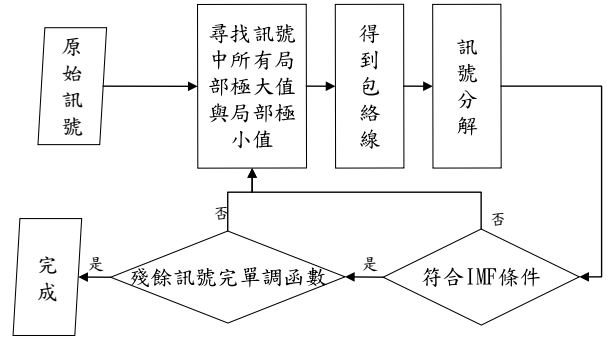


圖 1、EMD 流程圖

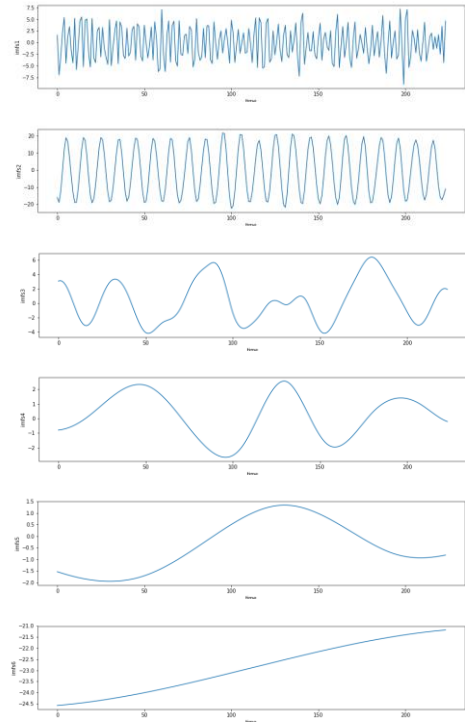


圖 2、分解腦波訊號示意圖

本研究的架構是交由 CNN 進行特徵提取,因此須將輸入轉為二維陣列;排列的方式我們採用重複性,重複性是指運用這 5 種不同的訊號進行排列組合,如 $(k*n)$ 的方式,當 $k=2$,則代表 5 種訊號間兩兩都有排列在一起。影像的長度與寬度則是取決於網路模型所需要的長寬(224*224)。

(三) VGGNet

VGGNet [16]是牛津大學計算機視覺組 (Visual Geometry Group) 研發的深度卷積神經網路。VGG-Net 探索了卷積神經網路的深度與其性能之間的關係,通過反覆堆疊 3x3 的小型卷積核取代 AlexNet 所用的大型卷積核, VGGNet 成功地構築了 16~19 層深的卷

積神經網路。VGGNet 的整個網路都使用了同樣大小的卷積核尺寸 (3*3)。

VGGNet 擁有 5 段卷積，每一段內有 2 至 3 個卷積層，同時每段尾部會連接一個最大池化層用來縮小圖片尺寸。其中多個完全一樣的 3*3 的卷積層堆疊在一起的情況，如圖 3 所示，兩個 3*3 的卷積層相當於 1 個 5*5 的卷積層，即一個像素會跟周圍 5*5 的像素產生關聯。而 3 個 3*3 的卷積層串聯的效果則相當於 1 個 7*7 的卷積層。除此之外，3 個串聯的 3*3 的卷積層，擁有比 1 個 7*7 的卷積層更少的參數，使得訓練時間更短。

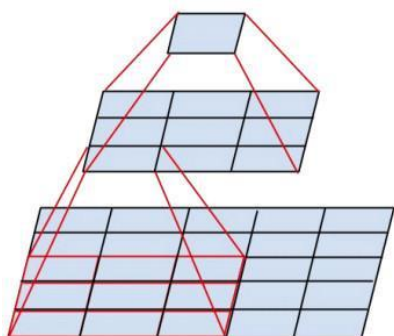


圖 3、兩個 3*3 的卷積層等於一個 5*5 的卷積層

(四) ResNet

ResNet (Residual Neural Network) [17]由微軟研究院的 Kaiming He 等 4 名華人提出，通過使用 Residual Unit 成功訓練 152 層深的神經網路，在 ILSVRC 2015 比賽中獲得了冠軍，同時參數量卻比 VGGNet 低，效果非常突出。ResNet 的結構可以極快地加速超深神經網路的訓練，模型的準確率也有非常大的提升。

ResNet 解決這個問題：在不斷加深的神經網路時，會出現一個 Degradation 的問題，即準確率會先上升然後達到飽和，再持續增加深度則會導致準確率下降。而論文中提到的使用全等映射直接將前一層輸出傳到後面的想法，就是 ResNet 的主要架構。

假定某段神經網路的輸入是 x ，期望輸出是，如果我們直接把輸入 x 傳到輸出作為初始結果。如圖 4 所示，這就是一個 ResNet 的殘差學習單元 (Residual Unit)，ResNet 相當於將學習目標改變了，不再是學習一個完整的輸出，只是輸出和輸入的差別，即殘差。

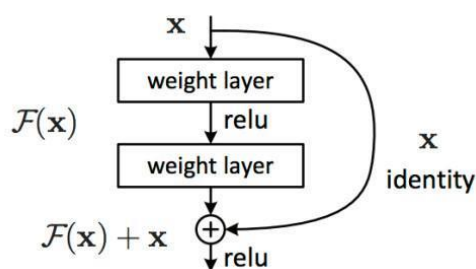


圖 4、ResNet 的殘差學習模塊

四、實驗結果

本實驗使用神念科技 NeuroSky Mindwave 單極腦波耳機來收集以 60Hz 輸出的 alpha、beta、gamma、delta、theta 資料，其訓練資料與測試資料為相關事件，並且觀察閉眼與睜眼上的差別，最後再以 ResNet18、VGG16 兩個網路模型來進行訓練。

(一) 閉眼收集資料

閉眼實驗收集了 8 組資料，分別以 224 筆為一組做分割，再經 EMD 後排列組合產生 224 筆資料，我們以所有資料與只取 alpha1 與 alpha2 的資料，分別使用 VGG16 與 ResNet18 兩種 CNN 網路模型進行比較，同時也比較是否採用 pre-train model，實驗結果準確率如表 1 所示，大致上而言，採用所有資料會比只用 alpha1 與 alpha2 的資料效果來得好，而使用 pre-train model 又會比沒有用 pre-train 來得好。

表 1、閉眼資料之準確率

	All data	alpha1+ alpha2
VGG16 pre-train model	60%	53%
ResNet18 pre-train model	65%	53%
VGG16 without pre-train	50%	50%
ResNet18 without pre-train	56%	60%

(二) 睜眼收集資料

睜眼實驗亦收集了 8 組資料，分別以 224 筆為一組做分割，再經 EMD 後排列組合產生 224 筆資料，我們以所有資料與只取 alpha1 與 alpha2 的資料，實驗安排如同上述的閉眼實驗一樣，分別使用 VGG16 與 ResNet18 兩種 CNN 網路模型進行比較，同時也比較是否採用 pre-train model，實驗結果準確率如表 2 所示，其中以 VGG16+pre-train model 效果最好，而睜眼實驗效果也比閉眼實驗效果好。

表 2、睜眼資料之準確率

	All data	alpha1+ alpha2
VGG16 pre-train model	72%	70%
ResNet18 pre-train model	69%	70%
VGG16 without pre-train	50%	50%
ResNet18 without pre-train	66%	71%

五、結論

本實驗嘗試使用不同的前處理方法、不同 CNN 架構的比較，自行收集的資料使用 VGG16 搭配 3 層全連接層準確率可達 72%，可以推論出以下問題，儀器導致的誤差非常巨大，其次可能是因資料量不夠或訓練過程遺失重要資訊導致準確率下降。未來，希望能使用精準度更高的儀器，並收集更多的資料與不同人的資料來進行分析、學習，以能達到更高的準確度。

四、致謝

感謝科技部 107 年度大專學生研究計畫，計畫編號 107-2813-C-155-024-E。

參考文獻

- [1] B. Rockstroh, T. Elbert, A. Canavan, W. Lutzenberger, and N. Birbaumer, Slow Cortical Potentials and Behavior, 2nd ed., Baltimore: Urban and Schwarzenberg, 1989.
- [2] E. E. Sutter and D. Tran, "The brain response interface: communication through visually induced electrical brain responses," Journal of Microcomputer Applications, vol. 15, pp. 31–45, 1992.
- [3] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," Clinical Neurophysiology, vol. 110, no. 11, pp. 1842–1857, 1999.
- [4] G. Pfurtscheller and F. H. Lopes da Silva, "Functional meaning of event-related desynchronization (ERD) and synchronization (ERS)," The Netherlands: Elsevier, vol. 6, pp. 51–65, 1999.
- [5] W.S.Mcculloch, W.Pitts. "A logical calculus of the ideas immanent in nervous activity," Bull Math Biol, 52(la)99-115.1990.
- [6] M.Minsky, S.Papert."Perceptron (expanded edition), "Cambridge, MA: MIT Press,1988.
- [7] D.Rumelhart, G.E. Himtont, R,J. Williams. "Learning representations by back-propagating errors," Nature,323(9):533-536.1986.
- [8] 叢爽，面向「MATLAB 工具箱的神經網路理論與應用」，第三版，合肥:中國科技技術大學出版社，1-7，2009。
- [9] D.Ciregan, U.Meier, and J.Schmidhuber, "Multi-column deep neural networks for image classification," in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, 2012, pp.3642-3649:IEEE.
- [10] Y.LeCun and Y.Bengio, "Convolutional networks for images, speech, and time series," The handbook of brain theory and nural networks, vol.3361, no.10, p.1995, 1995.
- [11] C.Lee Giles, G.M.Kuhn, and R.J.Williams, "Dynamic recurrent neural network:Theory and applications," IEEE Transactions on Neural Networks, vol.5, no.2, pp.153-156, 1994.
- [12] D.H.Hubel and T.N.Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," The Journal of physiology, vol.160, no.1, pp.106-154, 1962.
- [13] Y.LeCun et al., "Backpropagation applied to handwritten zip code recognition," Neural computation, vol.1, on.4, pp.541-551, 1989.Alex Krizhevsky.,
- [14] ImageNet Classification with Deep Convolutional Neural Networks" NIPS 2012.
- [15] N. E. Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," in Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 1998, vol. 454, no. 1971, pp. 903-995: The Royal Society.
- [16] Karen Simonyan and Andrew Zisserman, "VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION," ICLR, 2015.
- [17] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun., "Deep Residual Learning for Image Recognition," CVRP, Dec 25, 2015.