



# 人工智慧 技術在 海洋觀測 之應用

陳威成／國立成功大學水利及海洋工程學系 博士生

楊文榮／國家海洋研究院海洋產業及工程研究中心 助理研究員

傅仲偉／國立成功大學水利及海洋工程學系 碩士

董東璟／國立成功大學水利及海洋工程學系 教授

廖建明／國家海洋研究院海洋產業及工程研究中心 主任

台灣海域遊憩活動興盛，但海域也面臨著自然災害和人為污染等問題。有許多研究利用海洋觀測方法，使用衛星影像和光學影像並透過影像處理技術進行觀測和分析。近年來隨著電腦硬體和人工智慧技術的發展，應用人工智慧技術來研究自然現象問題也逐漸興起。本研究論述了與人工智慧應用在海洋觀測相關的研究，包含利用人工智慧技術偵測海岸或海面上的海洋廢棄物協助解決海洋垃圾問題、以及偵測海面下的石油或天然氣管線是否洩漏以防止海域污染擴散等。這些研究都展示了人工智慧在海洋觀測領域的潛在價值。最後，本研究以人工智慧應用於裂流觀測為例，使用更快速卷積神經網路（Faster R-CNN）建立裂流判識模型，並獲得良好的判識效果，正確率可達八成。顯示人工智慧技術可以有效地應用於海洋觀測。

關鍵詞：人工智慧、海洋觀測、裂流判識

## 前言

台灣四面環海，近年來海域遊憩活動更是逐漸興盛，許多民眾經常趁著空閒時候前往各處海岸進行遊憩活動，但在民眾親近海的同時，海域也常伴隨著自然災害、人為汙染等問題，例如突然在岸邊出現的瘋狗浪（freak wave）將民眾打入海中、暗藏在海面下將人帶離近灘的裂流（rip current）或是人類任意傾倒垃圾導致海岸堆積海洋廢棄物等問題。因此，許多研究透過海洋觀測方法來了解海域的情況，常使用衛星影像和光學影像並透過影像處理技術來觀測與分析海域的各種問題，如 Dankert *et al.* [1] 使用衛星雷達影像或微波雷達影像去分析海洋中的群波（wave groups），Goddijn and White [2] 提出利用數位相機拍攝海域來估計海域水質的方法，Pugliano *et al.* [3] 利用光學影像來觀察海岸線的變化等。

近年來因為電腦硬體與技術的增強，將人工智慧方法（Artificial Intelligence, AI）應用於自然現象問題的研究逐漸興起。人工智慧目的為讓機器或電腦等非人類的

人造物能透過各種不同的方法來擁有部分與人類相似的智慧，人類智慧的行為諸如模仿、學習、思考、判斷以及啟發等，因此人工智慧是一種理念、目的或是框架，可以囊括諸多概念與方法。在人工智慧的基礎下誕生了許多諸如資料探勘（data mining）、機器學習（machine learning）以及深度學習（deep learning）等概念及方法，這些人工智慧方法彼此相關聯，目標皆為讓機器擁有部分於人類智慧相似的能力。

與海洋領域相關的問題通常較為複雜，例如部分海洋現象尚未有明確的理論可以說明其機制，並且較常出現只能應用在特定區域，在其他區域應用需重新校正的問題。海洋的清澈度、天氣條件與地理位置等因素也容易影響海洋觀測的結果，並且海洋觀測常需耗費大量的人力與資源。因此，掌握人工智慧技術應用於海洋觀測，能讓人智慧學習如海洋現象或海洋污染等問題的特徵，透過學習大量的海洋觀測資料，讓人智慧能自動化且準確地觀測到相關問題，有利於後續的分析與研究。

## 人工智慧方法 深度學習方法

近年因電腦計算能力增強，對於利用機器學習以及深度學習的研究逐漸興起，包含能針對時序資料如股票走勢的預測、非時序資料如單張圖像的分析都能有不錯的效果。如圖 1 所示，機器學習是在已處理過的資料中透過學習已知的特徵來對新事物進行預測，機器學習在進行學習前必須給定一組包含經由人類知識萃取出特徵的資料。而深度學習基於神經網路（neural network）理論，不同於機器學習方法，深度學習捨去了需先由人類知識執行的知識萃取，而是多層結構的神經網路從大量的資料中學習特徵為何，對資料進行學習前不需透過人類知識的介入，因此，近年來在人工智慧領域中，深度學習方法相當熱門。

深度學習是機器學習分支下的一門學問，同樣是使機器或是電腦能透過經驗、試誤或模擬的方式進行學習。因此，深度學習可視為一個自適應的函數，通過結合多個簡單的概念來表達複雜的問題，深度學習透過過往的資料和經驗來學習資料中的規則或是特徵，尋找一個能從特定輸入與預期輸出找出關聯的功能函數，以回歸問題的例子來說，深度學習即是透過不斷地計算與修正函數中的參數，從資料分布中找出一個函數能夠擬合資料，進而能預測輸入新的資料時的輸出為何。深度學習旨在透過優化問題（optimization problem）來適應任何欲應用問題的函數，根據損失函數（loss function）來定義問題，例如與實際資料相差程度的均方根誤差（root mean square error），透過參數優化的技術使其適應和擬合所要解決的問題，而不同架構的深度學習目標都是為了幫助機器能有效適應問題的過程。

如圖 2 所示，機器學習方法包含隨機森林（Random Forest, RF）、支撐向量機（Support Vector Machine, SVM）以及淺層的類神經網路（Artificial Neural Network, ANN）等，而深度學習方法包含深層神經網路（Deep Neural Network, DNN）、遞迴神經網路（Recurrent Neural Network, RNN）以及卷積神經網路（Convolution Neural Network, CNN）等，遞迴神經網路處理需考慮時間屬性的時序列資料能有很好的效果，而卷積神經網路在影像處理領域有顯著的成果，主要能透過對二維的影像進行特徵擷取，從擷取的特徵中學習知識或是規律等，從而對新的資料能進行準確的辨識或

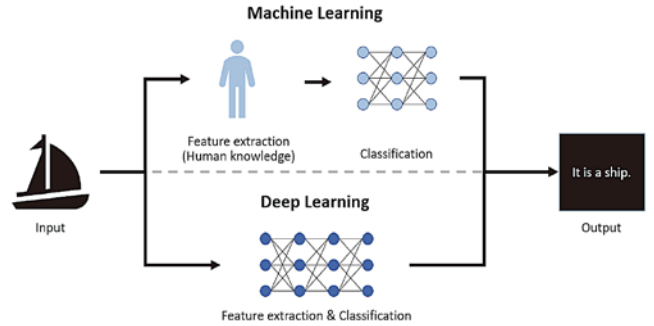


圖 1 機器學習與深度學習之差異

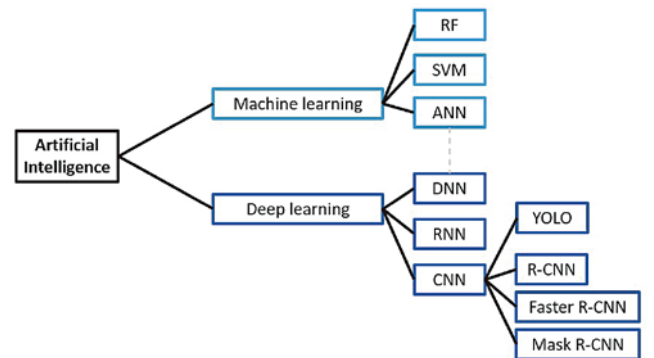


圖 2 人工智慧、機器學習與深度學習衍生之多種方法

是預測。因此，以卷積神經網路為基礎衍生了許多新的方法與架構來處理各式各樣的影像問題，如運算速度較高能快速偵測目標物的 YOLO（You Only Look Once）、針對影像中每一個區域仔細辨識的 R-CNN（Region Convolutional Neural Network）、以 R-CNN 為基礎改善辨識速度的 Faster R-CNN 以及以 Faster R-CNN 為基礎能偵測出目標物外型的 Mask R-CNN。

## 類神經網路

目前的深度學習方法皆以類神經網路（artificial neural network）為基礎衍生而來，類神經網路為模仿人類大腦組織及運作方式的資訊處理技術，人類大腦由密集相互連接的神經細胞所組成，每個神經細胞的構造簡單，但大量的神經細胞結合起來即能擁有可觀的計算與處理能力。神經元（neuron）為類神經網路最基本的組成單位，如圖 3 所示，神經元透過權重的連結將信號傳

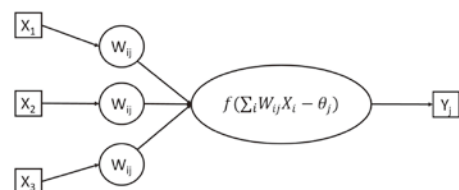


圖 3 神經元架構

送至其他神經元，輸入值  $X$  與權重值  $W$  輸入至神經元後，神經元內部即開始進行運算，運算完後的結果再經過啟動函數（activation function）轉換，計算公式如式 (1) 所示，而計算結果  $Y$  即為神經元之輸出值。

$$Y_j = f(\text{network}) = f\left(\sum_i W_{ij} X_i - \theta_j\right) \quad (1)$$

其中  $Y_j$  為類神經網路中神經元的輸出值， $f(\text{network})$  為啟動函數， $W_{ij}$  為類神經網路中各神經元間的連結權重值， $X_i$  輸入變數， $\theta_j$  類神經網路處理單元的門檻值。

類神經網路架構包含輸入層（input layer）、隱藏層（hidden layer）以及輸出層（output layer），若隱藏層的數目夠多且擁有大量的神經元進行運算，就會被稱為深度神經網路，如圖 4 所示，神經網路的輸入資料由輸入層輸入，傳遞至隱藏層的神經元進行運算，根據要解決的問題之複雜度，神經元的數量可以是一個或是多個，最後經由啟動函數（activation function）計算得到輸出值。而對於機制尚未明確或是較為複雜的問題，通常會使用如 S 型曲線函數（sigmoid function）和雙曲線正切函數（hyperbolic tangent function），使類神經網路具有非線性計算能力，建立複雜的函數關係。

近年來隨著硬體的計算能力和儲存空間的增加，能夠成功訓練隱藏層層數較多且神經元較多的神經網路，增加隱藏層與神經元數目來建構更為複雜的神經網路，比起淺層神經網路可以提供更準確的結果，使電腦能更進一步地模擬人腦複雜的神經系統，這種具有多層的類神經網路架構，被稱為深度神經網路（deep neural network），即為深度學習方法的基礎，透過多個隱藏層中的神經元去萃取資料中隱含的知識或是特徵，來處理較為複雜的問題。

而目前的深度學習方法皆以神經網路為基礎衍生而來，由 Rumelhart *et al.* [4] 研究出倒傳遞類神經網路（back propagation neural network）加上電腦能力日新月異，使得類神經網路開始突出於人工智慧領域，透過

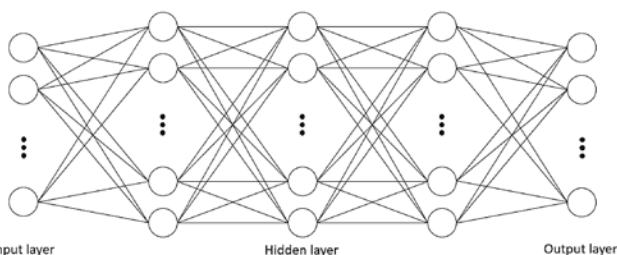


圖 4 深度神經網路架構概念圖

讓電腦學習與人腦結構相同的神經傳遞訊息的過程，從而讓電腦模擬神經傳遞訊息與反應動作的過程。倒傳遞神經網路即利用倒傳遞演算法來更新權重，透過不斷地計算與更新權重，最終得到可以適應或是擬合欲解決問題的函數。倒傳遞演算法是利用將網路中所有的權重，計算損失函數的梯度，根據梯度下降法（gradient descent）來決定如何修正網路中的權重，如圖 5 所示，梯度下降法會透過計算誤差函數的梯度，得到如何修正權重能使誤差降低的方向，訓練過程中不斷地計算梯度並修正權重，使誤差隨著訓練次數而逐漸降低，最終得到訓練效果良好的神經網路模型。

### 卷積神經網路

目前有許多種深度學習方法，例如深層神經網路（Deep Neural Network, DNN）、遞迴神經網路（Recurrent Neural Network, RNN）以及卷積神經網路（Convolution Neural Network, CNN），藉由能從大量資料中提取較為複雜的知識或是特徵等優勢，應用在各種領域上。深度學習方法在各種領域上的成功，使用深度學習方法應用各種問題上的研究逐漸興起，如影像分類、人類語音識別以及自然語言處理（natural language processing）等，在各項領域中皆能有顯著的成效。

卷積神經網路因其能有效地提取影像特徵並準確分析的優勢，至今仍在影像處理領域上蔚為主流。以 Le Cun *et al.* [5] 以神經網路進行手寫數字辨識技術為先驅，但當時受限於硬體設備的不足，神經網路在當時並無顯著優勢，而 Chellapilla *et al.* [6] 利用顯示卡

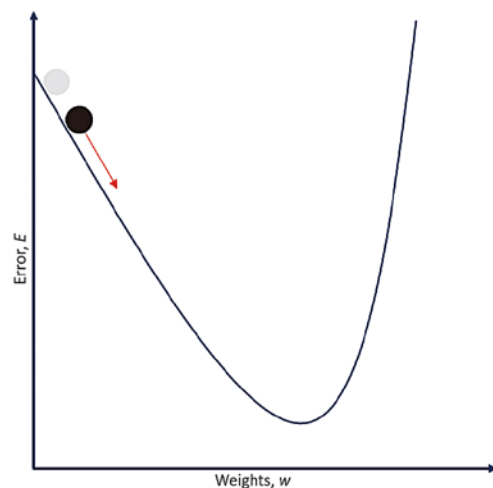


圖 5 梯度下降法概念圖

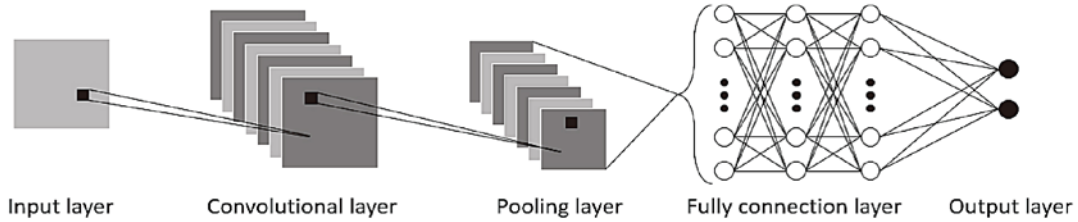


圖 6 卷積神經網路架構

(Graphics Processing Unit, GPU) 的平行運算能力，將原先的深度學習模型訓練的速度提升，也是第一個使用 GPU 進行平行運算來訓練的深度學習模型，Keizhevsky *et al.*<sup>[7]</sup> 在 2012 年參與 ILSVRC (ImageNet Large Scale Visual Recognition Competition) 競賽，利用卷積神經網路 (Convolutional Neural Network, CNN) 模型拿下影像分類比賽冠軍。

卷積神經網路架構與神經網路相同，可分為輸入層、隱藏層以及輸出層，輸入層與輸出層用於資料的輸入與運算結果的輸出，如圖 6 所示，卷積神經網路不同之處在於隱藏層中由卷積層 (convolutional layer)、池化層 (pooling layer)、全連接層 (fully connected layer) 所組成。如圖 7 所示，輸入資料在卷積層中通過卷積核 (convolutional kernel) 在影像上滑動進行卷積運算，尋找影像中的紋理或是邊緣等特徵，將找出的特徵結果結合成特徵圖 (feature map) 並傳入下一層中繼續尋找特徵，能保留目標的空間資訊並提取出合適的特徵。而卷積神經網路的訓練就是透過不斷地運算並調整卷積核參數以及神經元間的權重，使得輸入影像經過卷積層後能得到最有代表性的特徵。池化層目的為將卷積層輸出的特徵圖縮小，進行下採樣 (down sampling) 使得網路中的參數能減少，進而簡化網路的複雜程度以及降低運算成本，同時又能保留影像的特徵，常見的池化方式有平均池化 (average pooling) 和最大池化 (max pooling) 等，以平均池化為例，將輸入的影像分成多個子區域，計算每個子區域中的平均值，並將其作為新的特徵圖輸出至下一層中。全連接層將卷積層以及池化層提取出來的特徵整合，將原先二維的特徵圖轉換成一維的數據，並輸入至深度神經網路進行訓練，透過不斷地運算與更新權重，得出當特定的特徵圖輸入全連接層時，能正確進行影像分類或是偵測，因此在辨識新影像時，能透過影像中的空間位置與相關特性，將新影像經由多層的

運算與處理，最終判斷出新影像的特徵，進而進行影像分類與偵測。

卷積神經網路與一般的神經網路的區別在於卷積神經網路常用於影像特徵辨識 (pattern recognition)，卷積神經網路能透過卷積層可以從影像中捕捉到空間特徵 (spatial feature)，準確地判識影像中的目標和目標位置，相較於常用來處理一維數據的神經網路，一般的神經網路如欲處理二維影像資料，將二維影像的每個像素輸入網路中，會需要大量的神經元以及參數，並會失去原輸入資料的空間特性，而卷積神經網路的特點在於能處理二維數據，除了能特徵提取的同時保存資料的空間特性外，也能減少訓練模型所需要的參數，能判斷影像之特徵並分析，建構出層數多的網路，使網路有足夠多的參數可以進行訓練，具有較高的學習能力，能學習且處理較為複雜的問題。

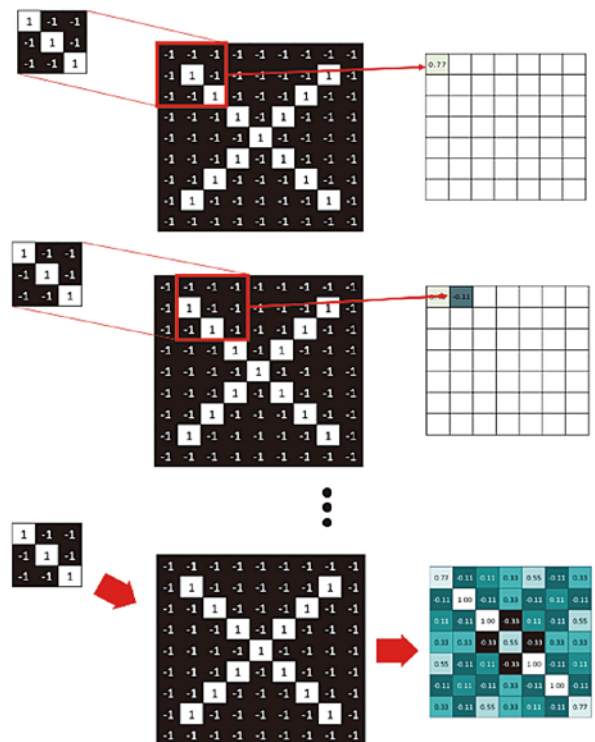


圖 7 卷積核對單張影像進行卷積並提取出特徵

## 導入 AI 的海洋觀測

### 應用 AI 於海洋觀測相關研究

人工智慧技術的演進與電腦硬體設備的增強誕生許多人工智慧方法，得利於深度學習方法在影像辨識上的發展，諸多利用深度學習方法來分析衛星或光學影像的研究逐漸興起，在海洋觀測領域上亦然。如 Dusek *et al.* [8] 探討美國國家大氣海洋局（National Oceanic and Atmospheric Administration, NOAA）對於裂流的觀測與分析模式，並展示使用岸邊攝影機以及時間平均（time-averaged）影像，讓救生員與模式兩者來辨識裂流並進行比較，此研究也介紹了使用基於影像流（video flow-based）和人工智慧方法來辨識裂流，結合這些方法可以提高裂流辨識的準確性。

除了可應用在觀測海洋現象外，也可透過人工智慧技術偵測海洋廢棄物或是環境汙染來解決海洋汙染問題。如 Fallati *et al.* [9] 指出人為海洋垃圾（anthropogenic marine debris）為海洋環境首要問題，其中以塑膠垃圾占所有人為海洋垃圾八成以上，此研究利用商用無人機（Unmanned Aerial Vehicle, UAV）搭配 RGB 高解析度攝影機空拍海岸影像，並且基於深度學習方法，偵測與分類出所有分散在海岸上的人為海洋垃圾，結果顯示透過此類方法能有效辨識海岸人為海洋垃圾。Chang *et al.* [10] 利用卷積神經網路以及遷移學習（transfer learning）的方法對濟州島的海灘圖像進行分類訓練，遷移學習能將已訓練好的模型轉移至新的模式，將某種問題已被訓練好的結果套用在另一個問題上，使用更少的運算成本來訓練。此研究以此種方法來訓練模型，學習對海灘影像是否乾淨或是遭受汙染的判斷，結果顯示此種演算法可以有效在光學監視影像中擔任海灘是否乾淨的偵測模式，幫助當地政府減少海灘垃圾的政策。De Laurentiis *et al.* [11] 指出使用合成孔径雷達影像（synthetic aperture radar imagery）辨識海洋浮油的研究中，面臨最大的挑戰之一是如何在海面存在多種光線散射機制下，準確辨識出漏油事件，此研究使用無人飛行器的數據，並且利用卷積神經網路提取非線性特徵的能力以及深度學習網路學習判斷漏油的能力來建置辨識海上漏油模型，研究結果顯示此種人工智慧技術能有效地判斷漏油事件。

而人工智慧技術不只能應用在海面上的觀測，也

可應用於海面下的觀測，如 Hu *et al.* [12] 指出海底石油和天然氣運輸管道若發生氣體洩漏可能導致海洋環境被嚴重破壞，甚至可能引起火災與爆炸。因此，此研究透過使用自主水下載具 - 機器魚（autonomous underwater vehicles-robotic fish）來偵測水下氣體洩漏事件，該機器魚除了裝備高解析度廣角攝影鏡頭拍攝水下畫面外，也裝載深度學習方法辨識模型，來辨識水下氣體洩漏事件的發生。透過不同尺度的水下物理實驗產生水下氣體洩漏資料集，提供給深度學習辨識模型進行訓練，此研究顯示結合深度學習與自主水下載具即時偵測水下氣體洩漏的可行性。

以上研究展示使用人工智慧技術可應用在各種海洋觀測的相關問題，包含海洋現象的判識、海洋廢棄物辨識、水下汙染偵測和船舶漏油追蹤等，顯示人工智慧技術在海洋觀測上的可行性與潛在價值。

### 透過 AI 觀測海洋裂流

海洋裂流是指在碎波帶（surf zone）內發生並具有時間短、尺度小且流速強等特點的海流。通常會發生在波浪與波浪之間或是波浪與地形交互作用導致波浪碎波後，引起以垂直海岸線方向離開碎波帶之海流。由於裂流的危險性相當高，許多學者致力於研究裂流的發生原因。根據現有的研究，裂流的出現、範圍以及強度主要取決於海氣象參數、潮汐特徵、海灘型態以及海底地形等因素。世界各國每年仍有許多人因為裂流而罹難，根據台灣交通部中央氣象局 [13] 統計國內近十年共計有 27 筆因為裂流所導致的意外事件發生，其中傷亡人數高達 118 人，並且實際上有發生裂流之事件會比僅統計傷亡事件的數據高出不少。因此，若能夠即時觀測到遊憩海域上裂流的出現，並盡早向民眾發出預警，請民眾遠離裂流發生之區域，進而減少傷亡事件的發生。

至今已有不少裂流觀測與機制探討的研究，Brander [14] 在澳洲的海灘進行現場實驗，透過近岸調查、佈放染料、航空測拍照片以及壓力計等方式來觀測與分析裂流，Dalrymple *et al.* [15] 透過現場觀測、實驗模擬與數值模式等方式分析，發現裂流會以不同的形態發生。除了可以使用影像處理技術來分析裂流的光學影像或航拍照片外，人工智慧中深度學習方法之卷積神經網路能透過卷積層擷取影像的特徵，為現今建置物件判識深度學習模型的常見方法，由 Grishick *et al.* [16] 提出的區域卷積

神經網路 (R-CNN) 奠定更有效且準確地進行物體識別的基礎，後續同樣由 Grishick *et al.* [17] 提出的快速區域卷積神經網路 (Fast R-CNN) 以及 Ren *et al.* [18] 提出的更快速區域卷積神經網路 (Faster R-CNN) 皆是以區域卷積神經網路 (R-CNN) 為基礎衍生而來。

因此，本研究利用深度學習方法中的更快速區域卷積神經網路 (Faster R-CNN) 建立裂流觀測模式，從 Google Map 與 de Silva *et al.* [19] 的研究共蒐集 320 張衛星影像，其中有 220 張為裂流影像，100 張則為無裂流之影像，藉由裂流影像建立資料庫並訓練深度學習模式，並進行模式超參數之率定，找到訓練效果最佳之裂流觀測模式。

Grishick *et al.* [16] 提出的區域卷積神經網路奠定準確辨識物件的基礎，如圖 8 所示，透過選擇演算法 (selective search) 的方法將影像切割成數個子區域，藉由子區域中影像的物體顏色、尺寸與紋理等關係合併成相似性較高的區域，找出具體有目標物的位置，因此選取了一定數量的候選區域 (region proposal)，再將每個候選區域輸入進卷積層中進行卷積運算，最後將卷積後提取出的特徵輸入支撐向量機 (support vector machine)，支撐向量機是一種善於處理分類問題的機器學習方法，支撐向量機會根據提取出的特徵對目標物進行分類並以邊界框 (bounding box) 表示其位置。區域卷積神經網路方法相較過往物件偵測之演算法，能大幅提升對物件辨識能力，並能解決傳統物體偵測方法須對影像中所有像素進行辨識，導致運算成本過高的問題。

然而區域卷積神經網路在訓練過程中，每一個候選區域均須經過卷積層進行特徵的提取，且卷積神經網路要求輸入的特徵圖尺寸必須為相同大小，區域卷積神經網路須對大小不同的候選區域進行裁減與縮放成相同大小後再進行卷積，因此運算過程較為緩慢且耗時，也容易導致影像的資訊失真或是形變，同時保存了許多卷積後提取出的特徵圖導致對於硬碟之儲存空間需求龐大。

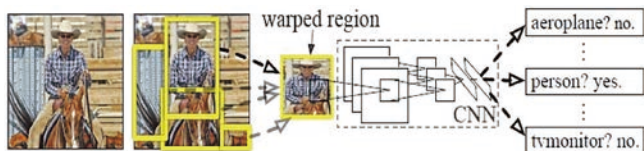


圖 8 R-CNN 架構圖 [16]

Grishick *et al.* [17] 以區域卷積神經網路為基礎，加入了 RoI 池化層 (region of interest pooling layer)，RoI 池化層先確定要輸出的影像大小為  $M \times N$ ，將輸入特徵圖的長與寬之像素分成  $M$  和  $N$  等份，每一份取最大值結合成  $M \times N$  的影像，即能輸出固定尺寸的特徵圖。因此，如圖 9 所示，快速區域卷積網路透過選擇演算法產生候選區域，並投影到輸入影像經過卷積後的特徵圖上，利用 RoI 池化層進行特徵的篩選並輸出新的特徵圖，連結全連接層找出目標物的分類與邊界框。快速區域卷積神經網路將整張影像輸入卷積層中，而非將各個候選區域進行卷積，並且將輸出分類與邊界框的任務一併結合在全連接層中的網路訓練，而非另外使用支撐向量機訓練，因此減少許多運算成本與儲存空間。

Ren *et al.* [18] 認為原先的快速區域卷積神經網路中選擇演算法所需耗費的時間依舊較長，因此提出以快速區域卷積神經網路為基礎，加入區域提取網路 (region proposal network, RPN) 取代原先的選擇演算法來提取候選區域，如圖 10 所示。因此，候選區域的提取、影像特徵的提取以及目標物分類與邊界框均由同一個深度學習網路架構來進行分析，減少運算時間與儲存空間。區域提取網路為一全卷積神經網路 (fully convolutional

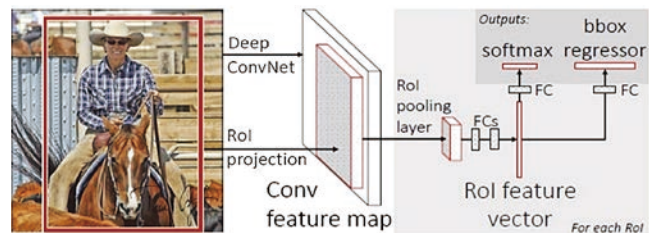


圖 9 Fast R-CNN 架構圖 [17]

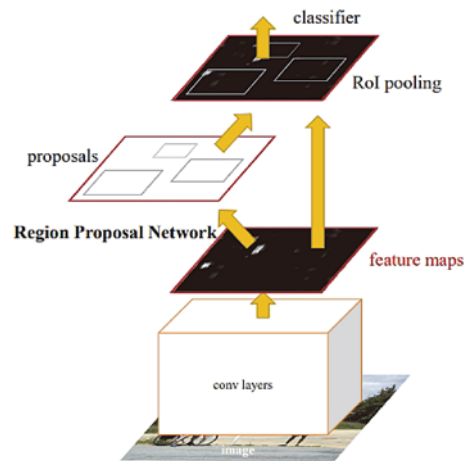


圖 10 Faster R-CNN 架構圖 [18]

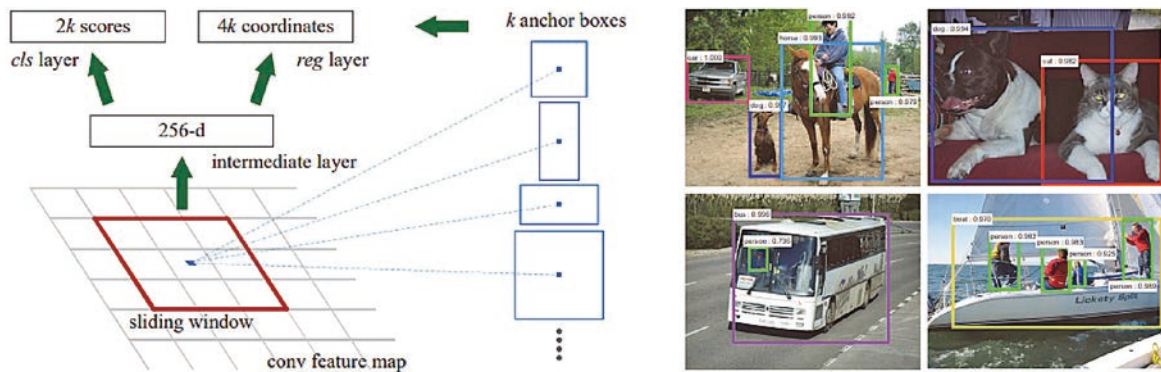


圖 11 Faster R-CNN 透過錨點選取候選區域並得到物體的分類與位置 [18]

network)，表示該神經網路中僅有卷積層，並無池化層與全連接層，用來產生候選區域。區域提取網路先經過一層  $3 \times 3$  之卷積層於特徵圖上進行卷積運算，再輸入兩層  $1 \times 1$  之卷積層，分別為進行物體或背景之二元分類的分類層（box-classification layer）與進行邊界框中心點與長寬大小回歸的回歸層（box-regression layer layer），因此可得多個候選區域。卷積層提取輸入影像的特徵輸出特徵圖，經過區域提取網路後卷積得到新的特徵圖，特徵圖中每個點結合了原特徵圖周圍的空間資訊。再對特徵圖上的每個點，生成多個大小不同的候選框，可由多種面積尺寸與長寬比例搭配而成，如圖 11 所示，這些候選框的中心點被稱為錨點（anchor），錨點包含了四個變數，分別為中心點座標、寬度以及高度，因此對於特徵圖上的每個點，得到多個候選區域。上述提到的分類層針對候選區域來判斷該區域為目標物或是背景，如果為背景即能篩選掉，減少後續的計算量；回歸層則用來確定目標物邊界框的中心點位置與範圍，未經過回歸層的邊界框部分框出物體的位置，透過與正確物體位置比較對邊界框進行調整，能回歸出能使原中心點與長寬更接近正確位置的函數，找出更準確的目標物位置。

## 裂流影像判識結果

### 評估指標

當一個人工智慧演算法模型建置完成後，需要以針對人工智慧的指標來評估訓練出來的模型成效如何。對於二元分類問題能以混淆矩陣（confusion matrix）來計算評估指標。以判識裂流影像的混淆矩陣為例，模型偵測該區域有發生裂流且實際亦為如此稱為真陽性（True Positive, TP），模型偵測該區域未發生

裂流且實際也未有裂流發生為真陰性（True Negative, TN），模型偵測該區域有發生裂流但實際卻未有裂流發生為（False Positive, FP），模型偵測該區域未有裂流發生但實際該區域有裂流發生為偽陰性（False Negative, FN）。根據此混淆矩陣可以計算出不同的評估指標來評估模型的好壞，本研究使用以下三種評估指標：

### 正確率 (accuracy)

正確率即模型在所有輸入的資料中，正確判識的比率，代表模型在實際有發生或是未發生裂流的情況下，均能正確判識的能力，可顯示出模型總體的判識效果。

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

### 召回率 (recall)

召回率即在實際有發生裂流的資料中，模型亦正確判識有發生的比率，又可稱為真陽性率（true positive rate），代表模型在海域實際有裂流發生時，能正確偵測該區域有發生裂流的能力，可顯示出模型準確判識裂流發生的效果。

$$\text{recall} = \frac{TP}{TP + FN} \quad (3)$$

### 精確率 (precision)

精確率即在模型判識有裂流發生的結果中，有多少比率實際亦有發生裂流的比率，又稱為陽性預測值（Positive Predictive Value, PPV），可用來評估模型是否常常將沒有裂流發生卻判識有發生的誤報程度，精確率越高代表誤報程度越低，若精確率低但召回率卻相對高時，代表模型在判識裂流發生時有高估的傾向。

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

### 模式訓練與測試

本研究使用的更快速區域卷積神經網路 (Faster R-CNN) 是以卷積神經網路為基礎架構，因此在建置模型的過程中神經網路透過一定數量的神經元以及神經元之間相連的權重來計算並傳遞資訊，擬合出一個複雜的函數去近似機制或特徵未明的問題，神經網路中的訓練意指透過不斷地進行試誤，並調整神經元之間的權重，直至神經網路能正確地輸出分類或預測等目標。權重的調整方式並非隨機調整，而是利用倒傳遞演算法將網路中的權重與誤差之關係 (即損失函數)，根據梯度下降法來決定如何修正網路中的權重，透過計算誤差函數中的梯度方向來決定如何修正權重能使誤差降低。因此，在訓練神經網路模型時，需先設定訓練時的超參數，超參數影響模型訓練時如權重的調整幅度與模型試誤的次數等，率定超參數能有效建置辨識效果最佳的模型本研究在模式訓練時進行訓練超參數 (hyperparameter) 之率定，目的為找出使訓練結果表現最佳之參數，超參數率定的好壞對模式辨識效果有顯著的影響。

本研究根據過往經驗與前人文獻先設定一組超參數，使用手動調整法 (manual tuning) 進行超參數之率定，選擇前人文獻已構建之參數能較有方向地設定超參數率定的範圍，找出正確率最佳之訓練模式。使用與訓練資料不同之 100 張影像進行超參數之率定，調整依序從影像模式進行一次權重調整過程時所訓練的

資料數量的批量大小 (batch size)、令權重更新時不會停滯的動量係數 (momentum)、抑制權重更新速度的權重衰減係數 (weight decay) 與影響模式權重調整次數的疊代次數 (iteration) 共四個超參數，根據率定好的超參數建置訓練正確率最佳的辨識模式，並進行後續模式測試。率定結果顯示，當超參數設置批量大小為 100、動量係數為 0.99、權重衰減係數為 0.0005 與疊代次數為 30000 時進行模式訓練，其訓練結果相較於其它超參數組合訓練之模式有較佳之正確率，可達 95%，代表設定此組率定好的超參數組合可以得到不錯的訓練效果。

因此，本研究設定率定好的超參數組合來建置裂流辨識模式，並且輸入與訓練資料不同的測試資料來評估模式的辨識能力。測試資料同樣蒐集於 Google Map 與前人研究所使用之影像，共計有 100 張衛星影像，其中 50 張為有裂流發生的影像、另外 50 張則為無裂流發生的影像。測試結果顯示模式建置的效果良好，正確率可達 86%，代表模式整體辨識效果不錯，部分的辨識結果圖如圖 12 所示。而模式的召回率可達 84%，代表模式能有效地辨識出裂流的發生，模式的精確率可達 88%，代表模式辨識出裂流的結果中，大多實際也有裂流發生，模式的召回率與精確率皆高，顯示模式在辨識裂流時不常有誤判的情形發生。以上三種指標皆達八成以上的辨識效果，顯示利用深度學習方法中的更快速卷積神經網路能建置良好的裂流辨識模式，使用此種人工智慧技術能有效地辨識裂流的發生。



圖 12 模式裂流辨識結果圖。圖中紅框處代表模式辨識裂流發生的位置



## 結論與建議

台灣四面環海，而海域也常伴隨著自然災害、人為汙染等問題，許多研究透過海洋觀測方法來分析海域相關問題，常使用衛星影像和光學影像並透過影像處理技術來觀測與分析海域的各種問題。因近年來電腦硬體與人工智慧（Artificial Intelligence, AI）技術的增強，將人工智慧方法應用於自然現象問題的研究逐漸興起。

人工智慧方法中的深度學習方法在影像辨識上的成功，許多研究使用深度學習方法應用於海洋觀測與相關問題上，包含可利用人工智慧技術來偵測海岸或海面上的海洋廢棄物來協助解決海洋垃圾問題、偵測海面下的石油或天然氣管線是否洩漏來阻止海域汙染擴散或是可追蹤海上船舶漏油區域來定位海上漏油事件並盡早防止漏油影響範圍擴大等。這些研究皆展示了人工智慧在海洋觀測領域的可行性與潛在價值。

本研究以深度學習方法中的更快速卷積神經網路（Faster R-CNN）建置裂流判識模式為例，從 Google Map 與前人研究共蒐集 320 張衛星影像，藉由裂流影像建立資料庫並訓練深度學習模式，並進行模式超參數之率定來找到訓練效果最佳之裂流觀測模式。結果顯示模式建置成果良好，正確率可達 86%，代表模式整體判識效果不錯，能準確地判識出裂流發生，並且在判識裂流發生時沒有顯著誤判的傾向，代表利用更快速卷積神經網路能建置良好的裂流判識模式。顯示使用人工智慧技術能有效地應用在海洋觀測問題上。

人工智慧方法應用在海洋觀測上往往需要大量且廣泛的資料來學習，未來針對所要應用的問題不同，建議可設置攝影機來固定拍攝海域影像，蒐集大量的影像資料提供給人工智慧模式學習，並且可搭配影像處理技術來減低因天氣條件等因素影響模式判斷的問題，提升模型判識的穩定性與準確性，增強人工智慧方法建置觀測模式時的判識效果。

## 參考文獻

- Ankert, H., Horstmann, J., Lehner, S., and Rosenthal, W. G. (2003), Detection of wave groups in SAR images and radar image sequences. *IEEE Transactions on Geoscience and Remote Sensing*, 41(6), 1437-1446.
- Goddijn, L. M. and White, M. (2006), Using a digital camera for water quality measurements in Galway Bay. *Estuarine, Coastal and Shelf Science*, 66(3-4), 429-436.
- Pugliano, G., Robustelli U., Benassai, G., Luccio, D., D., and Mucerino, L. (2019), Shoreline measurement obtained with direct and remote techniques on a sandy beach in Gulf of Pozzuoli (Campania). 2019 IMEKO TC-19 International Workshop on Metrology for the Sea.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986), Learning representations by back-propagating errors. *Nature* 323, 533-536.
- Le Cun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., and Jackel, L.D. (1989), Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation* 1 (4), 541-551.
- Chellapilla, K., Shilman, M., and Simard, P. (2006), Combining multiple classifiers for faster optical character recognition. *Document Analysis Systems VII, Proceedings 3872*, 358-367.
- Krizhevsky, A., Sutskever, I., and Hinton, G. (2012), ImageNet Classification with Deep Convolutional Neural Networks. *Advances in neural information processing systems*.
- Dusek, G., Churma, M., Im, J., Etienne, R., Pang, A., de Sliva, A., and Mori, I. (2020), From Lifeguards to AI: Advancing Rip Current Forecasting Through New Detection Methods. *Ocean sciences meeting. Computer Vision and Pattern Recognition.*, Cornell University.
- Fallati, L., Polidori, A., Salvatore, C., Saponari, L., Savini, A., and Galli, P. (2019), Anthropogenic Marine Debris assessment with Unmanned Aerial Vehicle imagery and deep learning: A case study along the beaches of the Republic of Maldives. *Science of the Total Environment*, 693.
- Chang, M., Yuan, Y., and Yue, Q. (2020), CNN Image Classification Analysis for 'Clean-Coast Detector' as Tourism Service Distribution. *Seed Science and Technology*, 18(1), 15-26.
- De Laurentiis, L., Jones, C. E., Holt, B., Schiavon, G., and Del Frate, F. (2021), Deep Learning for Mineral and Biogenic Oil Slick Classification With Airborne Synthetic Aperture Radar Data. *IEEE Transactions on Geoscience and Remote Sensing*, 59(10), 8455-8469.
- Hu, S. Y., Feng, A., Shi, J. H., Li, J. J., Khan, F., Zhu, H. W., Chen, J., and Chen, G. M. (2022), Underwater gas leak detection using an autonomous underwater vehicle (robotic fish). *Process Safety and Environmental Protection*, 167, 89-96.
- 交通部中央氣象局 (2021), 海岸裂流監測與預警技術研究 (1/3) 計畫報告, 財團法人成大研究發展基金會執行。
- Brander, R. W. (1999), Field observations on the morphodynamic evolution of a low-energy rip current system. *Marine Geology*, 157(3-4), 199-217.
- Dalrymple, R. A., MacMahan, J. H., Reniers, A. J.H.M., and Nelko, V. (2011), Rip Currents. *Annual Review of Fluid Mechanics*, 43, 551-581.
- Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014), Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 580-587.
- Girshick, R. (2015), Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, 1440-1448.
- Ren, S., He, K., Girshick, R.B., and Sun, J. (2015), Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 1137-1149.
- de Silva, A., Mori, I., Dusek, G., Davis, J., and Pang, A. (2021), Automated rip current detection with region based convolutional neural networks. *Coastal Engineering*, 166, 103859. 