

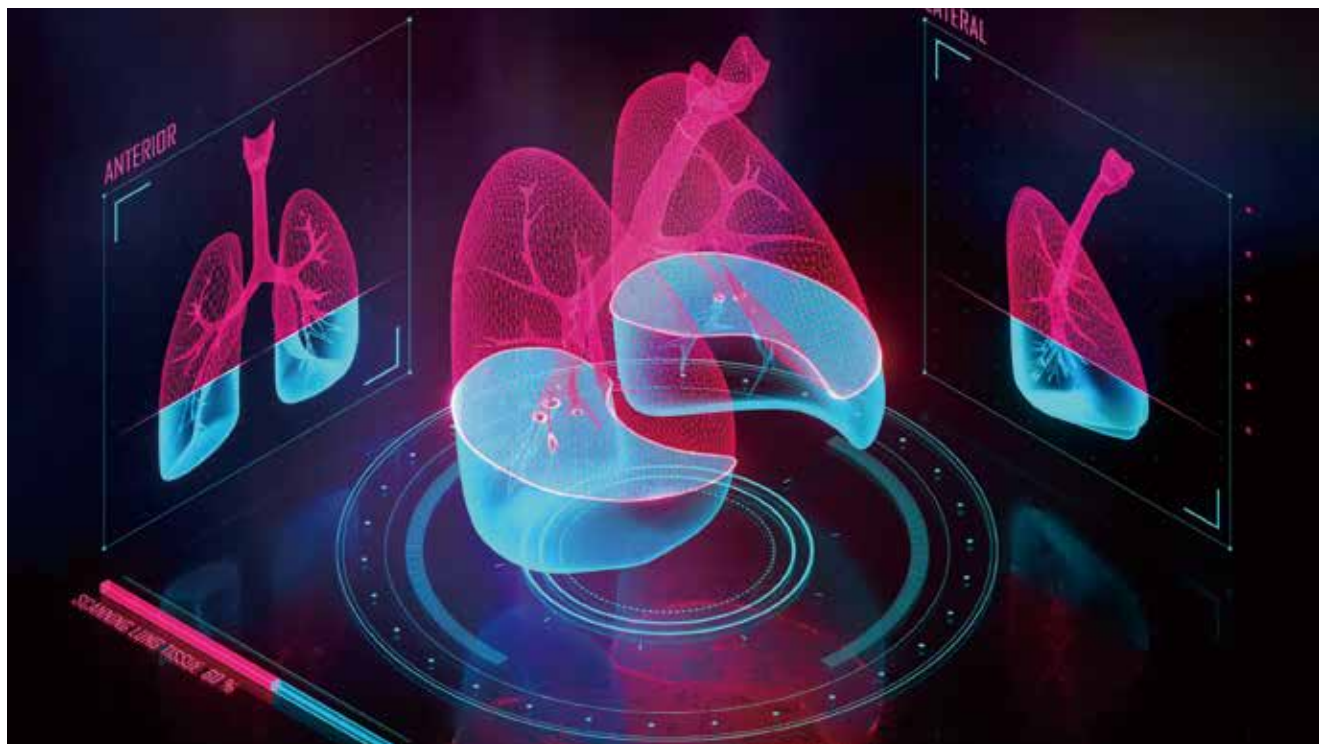
物理學點子幫人工智慧 發展出高維度視力

● 作者：約翰·帕夫勒斯（John Pavlus）

● 譯者：戴守煌

不論從什麼角度來看，物理定律都不會改變。這個構想使電腦能偵測出更高維度彎曲空間中的圖樣特徵。

作者簡介：帕夫勒斯為作家兼製片人，居住在奧勒岡州波特蘭市，作品刊載於《科學人》、《彭博商業周刊》等刊物。



新的深度學習技術在電腦斷層掃描影像的肺癌辨識上已展現出比以往更準確的前景，有朝一日將帶來更高品質的醫療診斷。（Olona Shmahalo 攝 / Quanta 雜誌）

如今電腦已經能夠開車、在西洋棋與圍棋等圖板遊戲中擊敗世界冠軍、甚至寫散文。人工智慧革命性進展主要有賴於一種特定人工類神經網路的強大能力，它的設計靈感來自哺乳類動物視覺皮質中層層連接的神經元。這種「卷積神經網路」（convolutional neural networks，簡稱 CNN）已證實十分善於學習 2 維資料中的型態，特別是在辨識諸如手寫筆跡與數位影像物體等電腦視覺任務方面。

不過一旦應用到不帶有平面幾何型態的資料集（例如 3D 電腦動畫中的不規則物件模型，或自駕車為了繪製周圍環境而生成的點雲）上時，這種能力強大的機器學習架構卻表現得不怎麼好。2016 年左右，一門稱為「幾何深度學習」（geometric deep learning）的學科出現了，目標是使 CNN 脫離 2 維平面的限制。

現在，研究人員已經建立新的理論架構，並打造



Quanta Magazine 是西蒙斯基金會（Simons Foundation）出版但編輯獨立之網路科普雜誌（<http://www.quantamagazine.org/>），希望能提高數學、物理與生命科學前沿研究進展的公眾能見度。本文譯自：

<https://www.quantamagazine.org/an-idea-from-physics-helps-ai-see-in-higher-dimensions-20200109/>

本刊感謝 Quanta magazine 與主編 Thomas Lin 同意翻譯轉載，翻譯之文責由本刊自負。



任職於高通與阿姆斯特丹大學的機器學習研究人員柯恩是規範等變卷積神經網路的主要推手之一。(Ork de Rooij 攝)

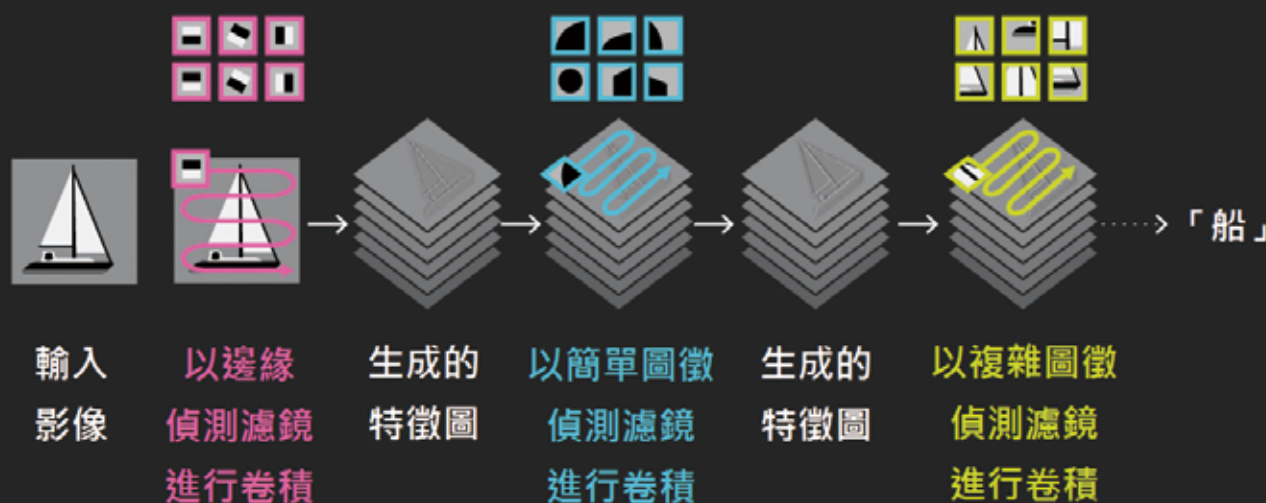
出可在任意幾何表面進行型態學習的類神經網路。這些由任職於阿姆斯特丹大學與高通（Qualcomm）人工智慧研究小組的柯恩（Taco Cohen）、威勒（Maurice Weiler）、基卡納格魯（Berkay Kicanaoglu）與魏林（Max Welling）等人所研發的「規範等變卷積神經網路」（gauge-equivariant convolutional neural network，簡稱「規範 CNN」）不僅能偵測出 2D 像素陣列中的型態，對球面與非對稱彎曲物件一樣有效。「這個架構是曲面上深度學習問題的最終解答。」魏林說。

規範 CNN 在全氣候模擬資料（可以自然地映射到球面上）的型態學習已經大幅超越其前輩的表現。這種演算法可能改善無人機與自駕車對 3D 物體的視覺，以及從心臟、大腦或其他器官不規則彎曲表面所讀取資料中辨識型態的能力。

研究人員讓深度學習跳脫 2 維平面的解決之道與物理學息息相關。用來描述這個世界的物理學，例如愛因斯坦廣義相對論或粒子物理標準模型等，都具有一種名為「規範對稱」的性質。這意謂世界中各種物理量與彼此之間的關係，並不會因為參考座標系（或稱「規範」）不同而改變；不論觀察者靜止或移動，不論數值差距多大，它們都保持一致。在不同規範下所做的量測結果必定能互相轉換，而且事物之間的基本關係在轉換過程中維持不變。

舉例來說，假如我們量測足球場長度，單位是碼。然後又量了一次，單位是公尺。長度數字會改變，但是改變方式是可預測的。類似的狀況是，兩位攝影師從兩個不同位置拍攝同一物體，會得到不同影像，但這些影像彼此相關。規範等價性確保物理學家各種描述實相的模型彼此一致，不論它們的

卷積神經網路如何看見物體



① 圖徵偵測濾鏡在輸入的影像上滑動，並記錄每個濾鏡與影像中每個位置的吻合度，產生一組特徵圖。

② 特徵圖經處理後，再換濾鏡逐次進行卷積，偵測越來越高階的圖徵。

③ 最後，類神經網路學會辨識並正確分類輸入的影像。

(Lucy Reading-Ikkanda 為 Quanta 雜誌繪製)

視角或量測單位為何。規範 CNN 對資料也抱持同樣的假設。

「（物理上）沒有任何一個方向是特殊的，這種概念他們想把它放進類神經網路中，」將機器學習應用到粒子物理資料的紐約大學物理學家克蘭莫（Kyle Cranmer）說：「而且他們想出實踐辦法。」

超脫平面

布隆斯坦（Michael Bronstein）是倫敦大學帝國理工學院（Imperial College London）電腦科學家，他在 2015 年提出「幾何深度學習」一詞，描述一項正在崛起的進展，大家開始嘗試跳脫平面並打造

在非平面資料進行型態學習的類神經網路。這個名詞以及相關研究活動很快就蓬勃發展起來。

布隆斯坦與他的合作者知道，要脫離歐幾里德平面，他們需要重新構思一種基礎運算程序，該程序曾經大幅提昇類神經網路 2D 影像辨識的效能。這種稱為「卷積」的程序，是先用一層類神經網路對輸入資料的許多小片段進行數學運算，再將結果傳送到下一層網路。

「大致上可以將卷積想像成一扇滑動的窗口，」布隆斯坦解釋道。一組卷積神經網路將許多這種「窗口」滑過整筆資料，每面窗口可偵測資料中的某種型態，功能有如濾鏡。以貓的照片為例，訓練過的 CNN 會用濾鏡偵測原始輸入資料的低階特徵，

例如圖案的邊。這些特徵接著傳送給別層網路，由該層進行卷積並找出高階特徵，例如眼睛、尾巴或三角形耳朵。訓練來辨識貓的 CNN，最後會利用這些分層迭次的卷積結果，為整張影像指定一個標籤，比如說「是貓」或是「非貓」。

但這些方法只在平面上管用。「一旦你想分析的表面變成彎的，麻煩就大了。」魏林說。

在曲面（幾何學中稱之為「流形」）上進行卷積，就好像在地球儀之外拿著一小方半透明方格紙，試圖準確描繪格陵蘭的海岸線一樣。你無法將方形平面完全貼合到格陵蘭上而不產生皺摺；這表示紙再度攤平後，你的作圖就變形了。但若將方格紙固定在相切於地球儀上某點的方向，透視描繪出來的格陵蘭邊界也會變形（此法稱為「麥卡托投影」）。另一種辦法是，你可以將方格紙疊在平面世界地圖上而非地球儀上，不過這樣也只是複製地圖本身的失真罷了，就像整個地圖的頂部邊緣其實只對應到地球儀上的一個點（北極）。而且如果流形不是像地球儀那樣單純勻稱的球體，而是更複雜或更不規則，例如瓶子的 3D 形狀或折疊起來的蛋白質，在上面進行卷積將更為困難。

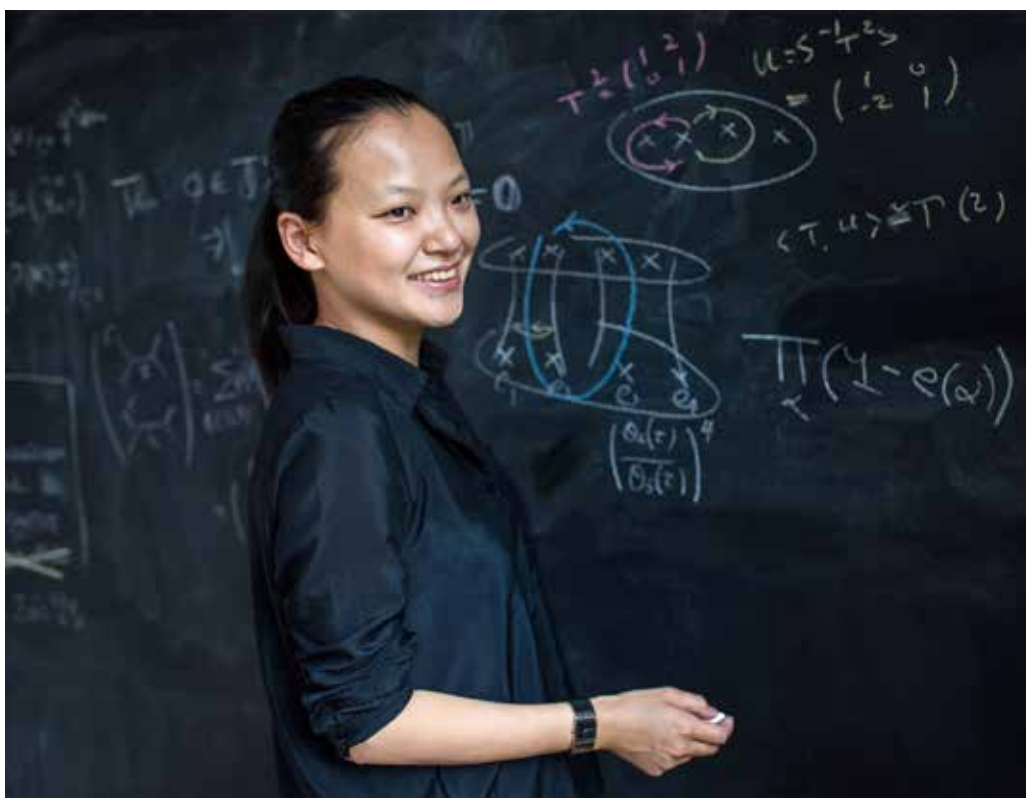
布隆斯坦與合作者在 2015 年發現一個方法，可以解決在非歐流形（non-Eulclidean manifolds）上卷積的問題，就是將滑動窗口的形狀改成更像一張圓形蜘蛛網而不是一張方格紙，因而可以貼合到球體（或任何曲面）上，卻不至於拉長、撕裂或弄皺它。

滑動濾鏡的性質這麼修改後，CNN 變得更善於「理解」某些幾何關係。舉例來說，類神經網路可以自動認得：從一個 3D 圖形彎折而成的兩種姿態

其實是同一物體的兩種不同狀態（比如站立的人像與抬起一腳的人像），而不是兩個截然不同的物體。這種修改也大幅提昇類神經網路的學習效率。標準 CNN「使用數百萬章圖形作為學習範例，（並且需要）經歷數週的訓練，」布隆斯坦說：「我們只用 100 個左右的圖形擺成不同姿態，而且只訓練了大概一個半小時。」

在此同時，柯恩與他在阿姆斯特丹的同事們也正開始從另一頭攻克這個問題。2015 年柯恩還是研究生，他當時研究的並不是如何讓深度學習脫離平面的限制。相反的，他感興趣的是當時自以為的實務工程問題——資料效率；或者如何用較少範例來訓練類神經網路，而非一般所需的數千筆或數百萬筆資料。「深度學習法就好比是非常緩慢的學習者。」柯恩說。如果你想訓練 CNN 辨識諸如貓的影像，這不會有什麼大問題（因為網際網路可以無止境供應貓的照片）。但如果你想要類神經網路偵測某些更重要的東西，例如肺部組織影像中的惡性腫瘤節結，那麼要找到夠多的訓練資料並不容易。這些資料必須診斷準確、標示得宜、並且沒有隱私疑慮。訓練類神經網路所需的範例越少越好。

柯恩知道，有個提昇類神經網路資料效率的方法是預先載入某些關於資料的假設——例如即使經過圖像內部的旋轉或鏡射，肺部腫瘤還是肺部腫瘤。卷積神經網路通常需要許多同型態但不同方向的圖案加以訓練，從頭學習這些資訊。2016 年，柯恩與魏林共同發表一篇論文，定義出將一些幾何對稱假設加入類神經網路的方法。這些方法十分有效；到了 2018 年，柯恩與共同作者溫克爾斯（Marysia Winkels）進一步拓展其應用，在電腦斷



阿姆斯特丹大學物理學家程之寧。(Ilvy Njokiktjen 為 Quanta 雜誌拍攝)

層掃描影像辨識肺癌的任務中展現出充滿前景的成果：他們的類神經網路只需訓練其他類神經網路十分之一的資料量，就能辨識出疾病的視覺證據。

這些阿姆斯特丹的學者們繼續開疆拓土。這就是他們最終發現規範等變性的由來。

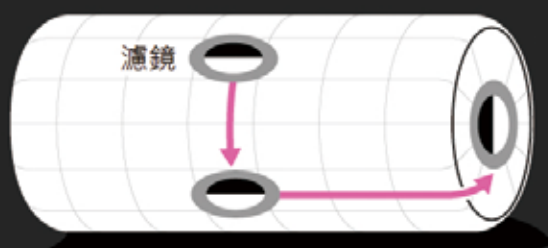
等變性的拓展

物理學與深度學習具有基礎上的相似性。就像柯恩說的：「兩個領域都關心實地測量，並建立模型以便預測未來觀測結果。」最重要的是，他注意到

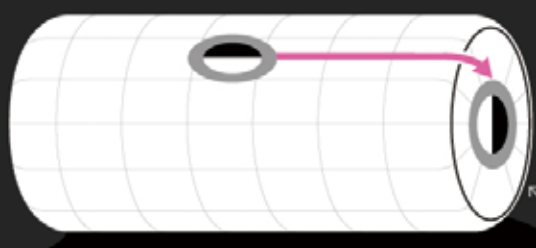
兩個領域尋找的都不是只能解釋個案的模型（氫原子有一套描述方式，上下顛倒的氫原子又有另一套描述方式，這樣是行不通的），而是普遍適用的模型。「當然，物理學在這方面做得相當成功。」

等變性（或稱協變性，物理學家比較喜歡這個詞）是一種假設，自愛因斯坦以降的物理學家都根據這個假設來延伸他們的模型。「它的意思不過是，如果你描述的物理是對的，那麼它應該與你用的『量尺』無關，或者更一般來說，跟你身為哪種觀察者無關。」阿姆斯特丹大學理論物理學家程之寧（Miranda Cheng）解釋道；她曾與柯恩以及其

在曲面上進行卷積



路徑 1



路徑 2

濾鏡最後的方向不一

標準卷積神經網路無法在曲面上運行。邊緣偵測濾鏡在曲面上滑動最後將朝向不同方向，依其滑動路徑而定。濾鏡以各種方式滑動會產生不同的特徵圖。

(Lucy Reading-Ikkanda 為 Quanta 雜誌繪製)

他研究者合撰論文，探討物理與 CNN 的關聯性。或者用愛因斯坦本人在 1916 年的話來說：「自然界的通則可用所有座標系中都成立的方程式來表示。」

透過此法則一個簡單範例的應用，也就是「位移等價性」，卷積網路成為深度學習最成功的方法之一。偵測圖像中某種特徵（例如垂直邊）的窗口濾鏡滑過（或者「位移通過」）整個像素點陣平面，並紀錄所有垂直邊的位置；然後它會建立一張「特徵圖」（feature map）標記這些位置，並傳遞給網路下一層。之所以能產生特徵圖，是因為位移等價性的緣故：類神經網路「假定」同樣特徵可以出現在 2D 平面上任何位置；不論垂直邊出現在右上角

或左下角，都能被成功辨識為垂直邊。

「等變類神經網路的重點在於，將這些顯而易見的對稱置入網路架構中，令它們成為不勞而獲的白吃午餐。」威勒說。

到了 2018 年，威勒、柯恩與他們的博士班指導教授魏林拓展這種「白吃午餐」，將其他等變性也包括進來。他們的「群等變」（group-equivariant）CNN 可以在平面影像中偵測出旋轉過或鏡射過的特徵，不需要先用該特徵朝向該方向的特定圖例進行訓練；球面 CNN 能從球面上的資料建立特徵圖，而不會像平面投影那樣變形失真。

這些方法的普適性仍不足以處理凹凸不平或不規則流形上的資料——這些流形幾乎涵蓋了一

切物體的幾何，從馬鈴薯到蛋白質、人體，以及彎曲時空。這些流形不具有整體對稱性（global symmetry）可以讓類神經網路事先預設：它們上面每個點都截然不同。

挑戰在於，在表面上滑動平面窗口有可能改變濾鏡的方向，依所走的個別路徑而定。設想有個濾鏡用來偵測左暗右亮的簡單型態。當它在平面網格上下左右滑動，可以一直維持正確方向朝上。但如果換成球面，狀況就不同了。當濾鏡沿著球的赤道滑動 180 度，濾鏡方向仍然一致，左暗右亮。然而，如果你改道經由球的北極滑動到對面同一位置，濾鏡就變上下顛倒了——右暗左亮。它將無法在資料中偵測出相同的型態，或製作相同的特徵圖。在更複雜的流形上移動濾鏡，可能導致它朝向與原來方向不同的任何方向。

幸運的是，愛因斯坦之後的物理學家已經處理過同樣的問題，並且找出解答：規範等變性（gauge equivariance）。

魏林解釋，關鍵是放棄掌握濾鏡沿不同路徑移動時方向如何改變。相反的，你可以只選擇一種濾鏡方向（或者「規範」），然後定義從其他方向轉換成這個方向的一致法則。

代價是，雖然初始方向可以採用任意規範，從其他規範轉換成該座標系時，原本的型態必須保持不變——就像光速從公尺 / 秒換算程英哩 / 小時，原本的物理量必須保持不變一樣。有了這個規範等變方法，魏林說：「真正的數字會改變，但改變方式都在掌握中。」

柯恩、威勒與魏林在 2019 年將規範等變性（也就是終極的「白吃午餐」）寫入他們的卷積神經網

路。他們的做法是在類神經網路透過卷積可從資料中「看到」的東西上加入數學約束條件；只有規範等變型態才可以在網路間層層傳遞。「基本上你可以給它任何形狀的表面」，從歐氏平面到任意彎曲物體，包括克萊因瓶（Klein bottle）之類的奇異流形，或者 4 維時空，「它都能在表面上進行深度學習。」魏林說。

管用的理論

規範等變 CNN 理論的適用範圍如此之廣，以至於它自動納入前一代幾何深度學習方法內建的假設——比如旋轉等變性以及球面上的濾鏡移動。甚至連布隆斯坦早先的方法也包含其中，也就是讓類神經網路辨識單一 3D 圖像彎曲而成的不同姿態。「規範等變性是一種廣泛適用的架構。我們 2015 年的工作成果也被涵蓋在內，成為某些特例。」

規範 CNN 理論上可以在任意維度下的任意曲面上運行；柯恩與同僚已經在全球氣象資料上測試過，這些資料天生具有 3D 球狀結構。他們用規範等變架構打造出一套 CNN，訓練它在模擬氣象資料中偵測極端天氣型態，例如熱帶氣旋。2017 年時，政府及學術研究人員用標準卷積網路從資料中找出氣旋的準確率為 74%；去年規範 CNN 偵測氣旋準確率高達 97.9%。（它的表現也勝過適用範圍沒那麼廣的幾何深度學習法，這是 2018 年時專為球面設計的，可達 94% 準確。）

穆迪岡達（Mayur Mudigonda）是使用深度學習的勞倫斯伯克萊國家實驗室（Lawrence Berkeley

National Laboratory) 氣候科學家，他表示將會持續關注規範 CNN。「這種人類視覺智慧的層面，」也就是不論方向為何，都能準確看出型態的能力，「是我們想應用在氣象領域的。」他說。最近聘用柯恩與魏林並收購他們新創公司（包含他們先前在等變類神經網路的研究成果）的晶片製造商高通打算將規範 CNN 理論用於發展更強大的電腦視覺應用，例如可以一口氣「看」360 度的無人機。（這種魚眼視野下的世界可以很自然地映射到球面上，就像全球氣象資料一樣。）

在此同時，規範 CNN 也吸引像克蘭莫這種物理學家的注意。他打算將它們應用在次原子粒子交互作用的模擬資料上。「我們正在分析有關於強（核）作用力的資料，試圖了解質子內部發生什麼事，」克蘭莫說。這些資料是 4 維的。他說：「這麼一來具有規範等變性的類神經網路就有了完美的應用場合。」

康多（Risi Kondor）是轉而研究等變類神經網路的前物理學家，他說規範 CNN 的潛在科學應用可能比它們在人工智慧上的運用重要得多。

「如果你的工作是在 YouTube 上辨識貓，而且發現你對於辨識上下顛倒的貓不是很在行，那就不太妙了，不過你也許可以睜隻眼閉隻眼。」他說。但是對於物理學家來說，確保類神經網路不會因為特定方向而誤判力場或粒子軌跡卻是至關重要的事。

「這不只是方便的問題而已，」康多說：「遵守與生俱來的對稱性是絕對必要的。」

不過儘管物理學家的數學對激發規範 CNN 的靈感有所助益，而且物理學家或許還能為它們找到眾多應用，柯恩注意到這些類神經網路本身並不會發

現任何新物理。「我們現在能設計出處理怪異資料的網路，但你得（事先）知道資料的結構是什麼」他說。換句話說，物理學家可以用規範 CNN 的理由是，愛因斯坦已經證明時空可以用四維彎曲流形來表示。柯恩的類神經網路無法自行「看出」這些結構。「我們還做不到學習對稱性這件事，」他說，雖然他希望這在將來有實現的可能。

柯恩忍不住為跨學科連結感到開心，這些連結曾經是他出於直覺想到，如今已用嚴謹數學表達出來的。「我一直有一種感覺，機器學習與物理都在做類似的事，」他說「這是我覺得妙不可言的事情之一：我們從這個工程問題出發，當我們開始改進我們的系統，卻有愈來愈多連結逐漸顯露出來。」

本文出處

Quanta Magazine January 9, 2020。本刊特別感謝阿姆斯特丹大學柯恩與程之寧授權刊登他們的照片。

譯者簡介

戴守煌畢業於台大物理系，英國德倫大學（Durham University）基本粒子理論博士，現為南臺科技大學助理教授。

延伸閱讀

- YouTube 關於規範 CNN 的介紹影片有 <https://www.youtube.com/watch?v=x1WRwq4tLlg> 和 <https://www.youtube.com/watch?v=wZWn7Hm8osA>；規範 CNN 的簡易簡介文章 <https://towardsdatascience.com/an-easy-guide-to-gauge-equivariant-convolutional-networks-9366fb600b70>
- 歐拉（Christopher Olah）是一位擅長解釋神經網絡的作家 <http://colah.github.io/archive.html> 這是他的部落格網址，其中有些神經網路與數學的科普文章，非常淺顯易懂。
- 史丹福大學「視覺辨識的卷積神經網路」（Convolutional Neural Networks for Visual Recognition）課程筆記網址 <https://cs231n.github.io/convolutional-networks/>