

# 基於 Visual Geometry Group Network 之交通號誌識別系統

## (一). 摘要

在自動駕駛技術蓬勃發展的今天，能夠輔助駕駛的各種系統也飛速成長，其中針對外界交通號誌進行定位並辨識的交通號誌辨識系統很早就進入人們對於智慧車輛的規劃藍圖中。其中針對交通號誌進行分類的系統也在深度學習當道的今天有長足的進步，隨著卷積神經網路接擁有更深的層數並針對結構進行改動，已逐步逼近人類的辨識率。然而越來越深的網路結構也對與訓練以及優化種下了隱憂。一個輕量且準確，並具備抗干擾能力的交通號誌辨識模型將增加自動駕駛技術安全性。為此本研究計畫提出了一個新模型的雛形，新的結構與舊結構相比更深，卻擁有更少量的參數。少量的參數較不容易過擬合，訓練時長縮短，更方便針對模型進行優化。本研究計畫也打算使用隨機關閉神經元及權重衰減等手段來防止模型的過度擬合，在初步測試中，新的結構相比舊的具備更優秀的辨識能力。新模型也在訓練過程中擁有抗干擾能力，透過在訓練過程中加上噪音，模型能分辨被錯誤輸入的資料並對系統作出反饋，以達到增進安全的目的。

## (二). 研究動機與研究問題

### 1. 研究動機

自動駕駛技術近年來引起越來越多研究人員和業界的關注。得益於先進駕駛輔助系統 (ADAS) 以及電子線控煞車等新興技術的發展，配備這些技術的車輛因其安全性和為駕駛員帶來的便利吸引了眾多新客戶。最近也有許多的行車記錄儀錄影帶顯示配備有自動駕駛技術的車輛確實具有檢測周圍環境危險及做出保護駕駛決定並相應採取行動的能力。例如，車輛在檢測到周圍有其他駕駛即將失控時，自動調整車速與行駛方向以維持車距，避免汽車和駕駛員本人遭受危險。而及時視覺訊息處理技術是自動駕駛系統的骨幹。借助高品質的傳感器（主要由攝像頭和激光雷達組成）以及強大的運算能力，車輛能夠從周圍環境中收集足夠的數據，經過處理之後輸出有意義的訊息。而交通號誌又是在行駛道路上能夠直接提示路況，對於行車資訊提供訊息的直接途徑。ADAS 必須使用交通標誌識別(TSR)等方法來處理獲得的交通號誌資訊，文獻[1]表明，可以將 TSR 提供的信息納入計算的 ADA 系統對於行車安全有重大改進。綜上所述，本次計劃的研究動機為改善現有交通標誌識系統，促進自動駕駛技術的進步與成熟。

### 2. 研究問題

一個成熟的交通號誌識別系統通常由兩個部分組成，交通號誌偵測系統(TSD)以及交通號誌分類系統(TSR)，流程如圖 1。前者對於攝像器所捕捉的影像進行運算，並標記出交通號誌所處在的區域，以便系統後續進行分類。傳統上交通號誌偵測系統多由工程師所撰寫的規則提取特徵，不但過度依賴程式設計師的經驗，在即時運用上遇到許多障礙而無法提高正確辨識率。近年來像 CNN 或 R-CNN 等深度學習方法[2-4]被廣泛用於子區域探測，功能更加強大，已經成為主流演算法，[5]提出了一種能夠快速標定小範圍目標的輕量級模型。在依據交通號誌偵測系統的輸出為依據切割圖片後，被判定含有交通號誌的圖片被送入交通號誌分類系統裡進行分類。交通標誌最初設計的

形狀有利駕駛將其與道路上其他自然物體區分，並且在惡劣環境下也能辨別標誌內容。但文獻[6]表明，即使交通號誌偵測系統提供了正確的區域，由於周圍環境依然會對圖片造成干擾（如由於長時間曝光或由於日光而導致的標誌色差）會弱化能夠進行分類的特徵，正確的分類圖片仍然是一件具有挑戰性的任務，詳見圖 2(引由[11])。再加上由交通號誌偵測系統錯誤判讀而產生的錯誤資料，分類系統應該具抗雜訊干擾能力並自動識別錯誤資料，以免輸出錯誤的訊息對於行車產生危險。交通號誌分類系統在將過多年的演進後取得長足的進展。許多新興的卷積神經網路在分類問題上都有著頂尖的判讀率。而隨著類神經網路的推陳出新，許多網路在加深的同時造成可訓練參數的指數增長，訓練模型盒優化都曠日費時。本計畫所研究問題為設計一個輕量同時擁有頂尖判讀率的模型，對於交通號誌分類系統將有莫大的幫助。

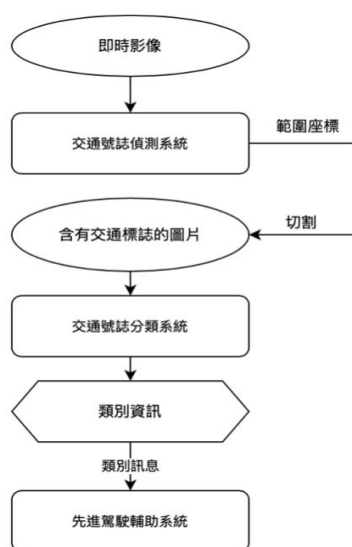


圖1 - 交通號誌識別系統

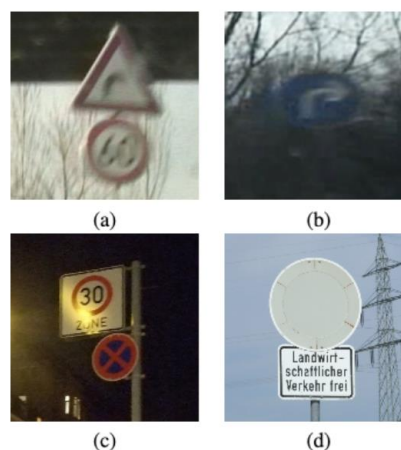


圖2 - 導致判讀困難的案例 [11]

(a)過久的曝光時間，(b)與背景的對比度過低，  
(c)環境燈光在交通號誌表面造成干擾(d)顏色在經過陽光照射的反光影像之後改變

### (三). 文獻回顧與探討

從大方向而言由，交通號誌分類系統由過去基於圖像或顏色的特徵分類方法逐步轉向由機器學習為主。近年來隨著類神經網路的發展，越來越多相關的技術被應用在特徵辨識與電腦視覺的處理。其中深度神經網路也早已被應用於許多自動駕駛領域，例如行人標示[7]或道路標示[8]，並取得了非常優越的成果。在早期的交通號誌分類領域，基於顏色的分類方法曾經當紅一時。使用顏色為標的之分類方法是對於圖片中每個像素的顏色分佈進行分析，進而將顏色分佈對應於一種交通號誌。將圖片分割為部分並分別標示上紅，藍，黃，綠，白和背景色等標籤。接著依據類別的特徵而建立出感興趣區。接著便能使用類似樹狀分類的方式對於多個類別進型分類。然而這些基於顏色的分類方法對於路牌的顏色變化過於敏感，一些陰影的變化或路牌上存在污漬都會大程度的影響輸出結果。[9]是基於 RGB 的顏色分類方法，而[10]是基於 HSV 的顏色分類方法。基於形狀的分類方法也曾經十分流行。這些方法基於將被分類圖片與模板進行比對，以捕捉交通號誌的形狀資訊。也有利用交通號誌經常出現圖案，如圓形，三角形和八角形的對成性進行偵測，藉而推測交通號誌[11]，或利用霍夫變換來偵測形狀[12]。這些方法的限制在於由於形狀的比對是和事前塑造好的模型進行對

比，對於實際運用中的路標旋轉造成角度偏轉和形狀缺陷將大大增加誤判機率。也有將以上兩者兼容並蓄而提出的模型。例如先利用顏色資訊進行分類，再利用圖片裡的形狀進行區分。但這些模型往往結合了上述兩者的缺點，對於照片裡的細節訊非常敏感而無法得到非常好的結果。例如圖片只要稍有色偏，分類器便無法分類。再者，這些模型通常非常複雜，實用的價值並不高。利用顏色或形狀資訊來進行分類的方法過度依賴程式設計師的經驗，已經逐漸被擁有更高正確率的機器學習方法所取代。第一代的機器學習方法以迭代最近點為分類演算法。他將分類工作分為三格步驟：針對原始圖片的特徵提取，利用基於迭代最近點的線性投影減少了維度再在進行分類。這類型方法的代表為[13]，當時在 GTSRB 獲得了 98.5%的正確率，遙遙領先其他以特徵提取為主的非機器學習方法。受益於更多機器學習方法的提出，例如利用支持向量機(SVM)，隨機森林(Random Forest)和利用 K-鄰近演算法(K-NN)都用有良好的分類能力。也有使用以上分類演算法配合早先被提出的圖像處理技術的交通號誌分類技術，例如由支持向量機配合圖像切割與感興趣區製成的分類器[14]。但相比起不斷在各項圖象分類領域刷新正確分辨率的卷積神經網路似乎還是略遜一籌，例如在著名的 GTSRB 暫居第一的神經網路即是由卷積神經網路進行改良而成，達成 99.71%的超高正確率。

卷積神經網路以其卷積層對特徵的萃取而聞名。現代的卷積神經網路往往透過多層卷基層互相疊加，上層的卷基層對於下層萃取出的特徵進一步萃取，進而達成比其他神經網路好上許多的正確率。這一切的開端由福島邦彥提出的神經感知機為源頭。透過模仿人類神經元的作用原理，感知機是含有輸入層和輸出層的神經網路，再加上激活函數使其能達到非線性分類功能。早期的卷積神經網路受制於當時計算能力得限制，與當時的其他機器學習方法例如支撐向量機相比並沒有多大的優勢。但在 Yann LeCun 等人的努力下，卷基層神經網路依然在各方面打穩根基。在 1985 年，[15]提出的後向傳播為神經網路提出了一個可靠高效的訓練方法。例如[16]便是採用後向傳播訓練多層神經網路以辨識手寫郵政編碼。以下提出現代卷積神經網路的著名模型：

### 1. LeNet

現代卷積神經網路的樣式在 LeNet[17]被大抵確定。卷積層，池化層，激活層以及全連接層的架構被後續的模型廣泛採用。LeNet 在當時沒有立刻大放異彩，歸因於沒有 GPU 輔助模型訓練，縱然使用後向傳播演算法訓練依然曠日費時，但其給之後的卷積神經網路提供了無數靈感而被稱作卷積神經網路的開山之作。

### 2. AlexNet

沉寂了許久的卷積神經網路終於在 AlexNet[18]迎來了爆發。其提出使用整流線性單位函數(ReLU)來代替由生物原理啟發的激活函數，例如 Sigmoid 函數。使用 ReLU 激活函數的模型不僅易於訓練，還減輕了梯度消失與誤差反向傳播演算法對於模型造成的影響，從而提高了模型準確性。AlexNet 也使用了 GPU 幫助訓練模型並一舉在 2012 年的 ImageNet 圖像分類競賽中拔得頭籌，而在此之後改良過的卷積神經網路也以其優異的性能不斷刷新各視覺辨識領域的辨識率紀錄。正式迎來了深度學習和卷積神經網路的時代。

### 3. VGG

如果說 AlexNet 是卷積神經網路的分水嶺，那 VGG[19]就是承上啟下的經典模型。在延續了前者的整流線性單位函數以及裡用 GPU 加速訓練的觀念後，一舉將模型加深並獲的巨大的成功。VGG 的一大特色便是採用較小的 3\*3 卷基層，這麼做能提高感受野並且在累積多個小卷積層後提高模型的非線性辨識能力。另外其採用的非重疊 2\*2 池化層也有類似的功能。上述所述的手段都對於加深網路深度有幫助，ReLU 激活函數加快了多層網路的訓練，較小的卷基層和池化層使網路能從下層提取更多資訊量，而使用 GPU 訓練使實際的訓練時間大幅縮短。可以說 VGG 模型的一切創新都圍繞著加深網路層數以達到更好效果。然而 VGG 也因為網路的深度而遇到了問題。在[19]中的表 3 明顯能看到此結構在到達 16 層以上時已經過於緊繃，在加深結構對於正確率已經沒有太多幫助了。縱使縮小卷積和池化層能延緩資料特徵縮減的問題，在層數加深的情況下依舊無法避免特徵的耗盡，在這個結構缺陷的限制下，VGG 結構的垂直發展已然遇到瓶頸。即便如此，VGG 依然將 Top 5 error rate 從原本 AlexNet 的 16.4% 下降到 7.3%，被視為既使用神經網路之後的另一句大突破。而隨著模型層數的不斷加深，訓練參數也隨之暴漲，對模型的訓練種下了許多隱患。

#### 4. Inception Net

Inception[20]透過同時加深與加寬網路來達到更好的效果。利用網路結構的稀疏性來避免過度擬合和梯度彌散等問題，並使用密集矩陣運算充分利用硬體資源。利用不同大小的池化層對上一層的網路輸出進行特徵提取，並利用 1\*1 的卷基層進行降維以保證在壓縮訓練參數的情況下達成基礎神經疊加的概念。Inception 在 2014 年提出後經過多次改良，例如改進卷基層的計算結構和改進分解甚至接和殘差連線。

#### 5. ResNet

後起之秀 ResNet[21]提出了殘差學習以修正 VGG 網路遇到的退化問題。透過段路連接，一舉將網路架構疊到了超過 100 層。也在 ImageNet 比賽上拔得頭籌，首次突破人類的分辨率。可以算是卷積神經網路的巔峰之作。

### (四). 研究方法及步驟

為了達成開發一個輕量同時擁有頂尖判讀率的交通號誌分類模型，本研究工作將建立新模型區分為三個部分：利用資料預處理來增進判讀率、建立新模型並測試其效能、使用優化方法優化模型並增進判讀率。利用以上三個部分來創建新的交通號誌分類模型，再使用各大交通號誌分類資料庫進行訓練。交通號誌雖然不是世界通用，但大多數相同含義交通號誌用有相似的外觀，善用這個原則能使訓練資料更加多元，數量也更多，能有效對抗過擬合問題，也能對於某些特定的資料類別使用資料增強來進一步拓展資料量。在訓練的過程中也需加入一些雜訊以及錯誤的資料進行訓練，以模擬上層交通號誌偵測系統出現錯誤的情況。交通號誌系統需要學會判斷輸入的資料是否是錯誤的，並將認為錯誤的輸入其歸類於一個類別中，這麼做對於交通號誌辨識系統的安全性有近一步保證。而被歸類為輸入錯誤的資料也能做為反饋用以改進交通號誌偵測系統。

#### 1. 利用資料預處理來增進判讀率

本研究工作將使用影像增強技術來強化模型的學習效果。有些原始照片太過灰暗，對於模型而言難以有效的提取特徵。而有些圖片的對比度太差，模型無法區分路標與背景物。應用影像增強技術可以對於原始圖片進行改進，解決對比度和清晰度問題，並調整整體亮度(見圖 3)。值得注意的是，原始照片（尤其是模糊照片）的亮度和對比度增強之間存在拮抗作用。增大對比度將使標誌本身與背景分離，從而將標誌調整為較深的色調。更高的亮度將使整體亮度更適合訓練，但會使交通標誌的圖案更加模糊。單通道圖像在訓練方面具有很多優勢，圖片所需參數比三通道圖像少，從而簡化了模型並減少了訓練參數，降低了訓練時間並減少了過度擬合的機會，與本研究計畫的目標不謀而合。灰階圖像有時甚至能凸顯照片的特徵，這使其在圖像分類 workflow 中很受歡迎。但是，在本研究計畫的模型中灰階圖片顯然不適用。這並不令人意外，畢竟顏色在交通標記中是一項傳遞大量訊息的因素，具有相同交通標誌圖案可以傳達完全相反的含義，只因他們的顏色不同。資料增強是一種防止過度擬合的常用方法，通過對於原始圖片的翻轉或調整角度，它可以增加數據種類並增加數據量。但在交通標誌分類系統使用全域的資料增強並不是一個好主意。很大一部分交通信號對於角度十分敏感。將右轉彎標誌翻轉將變為左轉彎標誌，使數據集混亂。該技術使用在大範圍下弊大於利。



圖3 – 照片預處理成效對比

## 2. 建立新模型並測試其效能

VGG 模型在許多應用中都達到了登峰造極的成效並取得了巨大的成功。但是 VGG 模型的缺點也是顯而易見。以更寬廣和更深層次的結構換取性能會導致需要訓練的參數暴增，其中 VGG-16 約有 1.3 億個參數，VGG-19 比前者多了 1000 萬個訓練參數。訓練如次臃腫的網路變成一項艱鉅的任務。不但有內存耗盡的風險，一次的訓練時間也曠日費時，優化大型模型更是一項困難重重的工作。為了解決上述問題，本研究工作對於固有的 VGG 模型進行改良。針對結構進行改良，以顯著減少需要訓練的參數，甚至獲得比傳統 VGG 模型更好的結果。[22] 通過實驗說明，在擁有相似學習參數的條件下，更深層的網絡往往比寬扁的網路結構擁有優勢，如果 VGG 有出色的性能但參數過於冗贅，為什麼不減少每層的參數，而使整體結構更深以彌補單層神經元數量不足的損失？

本研究工作構建了一個新結構，在完全連接的層之前，卷積層的數量略多，並且具有相同的 5 個池化層。傳統的 VGG 模型在每個大層結構中池化層之前的捲積層數量



是不一致的，經過驗證發現它並不能提高整個實驗的性能，因此本研究工作僅在結構中構造了相等的 5 層。另一個明顯的區別是圖層的厚度，VGG 提出了 [64, 128, 256, 512, 512] 的序列，這就是為什麼總參數非常大的原因，本研究工作將它們更改為相對較少的 [32, 64, 128, 256, 512] 序列，這項改變有助於節省需要訓練的參數。至於全連接層，AlexNet 和 VGG 模型均採用同一組的四個全連接層。一對具有 4096 個神經元的完全連接層，然後是 1000 個神經元層，最後以分類歸一化指數函式作為激活函數的輸出層。 [23] 得出結論，淺層網絡被迫增加每層神經元以補償數據集類型的重量損失，因此在訓練過程中會產生更多的總參數。本研究工作選擇了一個具有 [1024、512、256、128] 個節點的全連接層鏈，最後同樣接上分類歸一化指數函式作為激活函數的輸出層。本研究計畫結構除了比傳統 VGG 擁有更深的全連接層外，它以指數數列的形式減少了神經元的數量（Softmax 層大約有 50 個節點），而不是迅速地從數千個減少到半百個左右。VGG 與本研究計畫結構比較如圖 4:

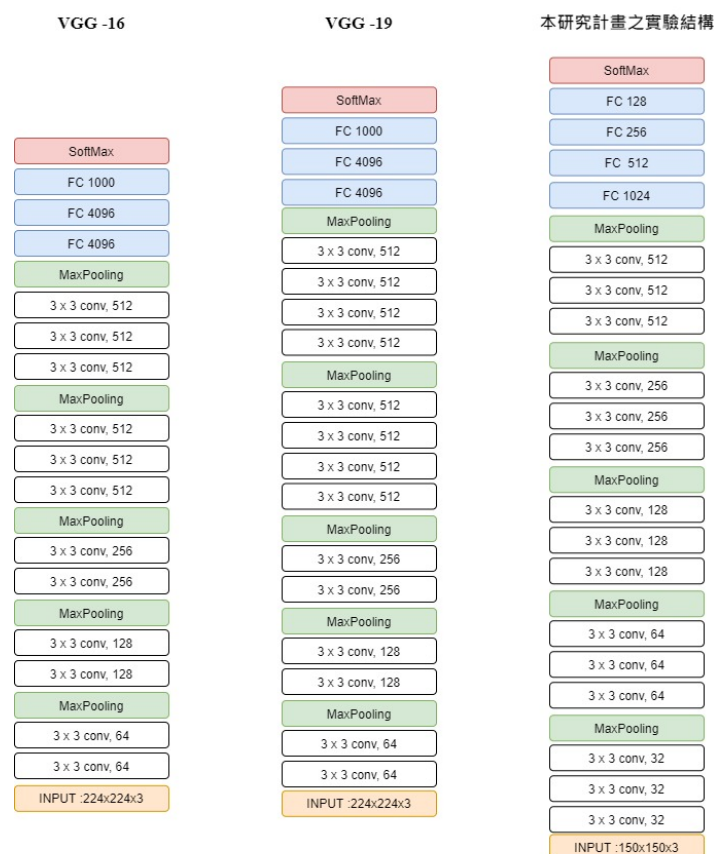


圖4 - 結構比較圖

本研究計畫實驗結構相比於 VGG 有兩個優勢，其中之一便是訓練參數。 假設將最後一個池化層的節點攤平為  $m$  個節點的一維數組，模型的任務是將數據分類為  $s$  類。在攤平層之間夾著標記為 0 到  $i-1$ ，總和為  $i$  個全連接層  $P$  的以及位於最尾端的輸出層，全連接層的參數是如式 1:

$$Parameters = ((currentlayer\ c) * (previouslayer\ p) + 1 * c) \quad \text{式 1}$$

整個模型的訓練參數是由各層之間的參數相加而得，由以上公式能推導出整個模型可訓練參數的總數為式 2:

$$P_0 * (m + 1) + \left\{ \sum_{j=0}^{i-1} P_{i+1} * (P_i + 1) \right\} + s * (P_{i-1} + 1) \quad \text{式 2}$$

由 VGG-19 原始模型的假設，最後一個池化層的輸出為  $7*7*512$ ，假設最後要將數據分成 50 個類別，本研究計畫實驗結構的可訓練參數和 VGG 結構的可訓練參數的比例表示為  $R$  如式 3：

$$\text{By assumption, } m = 7 * 7 * 512 = 25088, s = 50, \text{ thus}$$

式 3

$$R = \frac{1024m+129s+690,048}{4,096m+1001s+20,882,408} = \frac{26,385,586}{123,692,906} = 21.33\%$$

本研究計畫實驗結構在全連接層的可訓練參數相比 AlexNet 以及 VGG 模型要少得多。而全連接層的參數又是整體結構可訓練參數的主導因素，本研究工作的參數僅有約上述結構的五分之一。切換到新結構後，本研究計畫模型的訓練速度立刻有著飛躍性的成長。本研究也試圖在結構中添加更多完全連接的層，可惜的是，過擬合問題隨後開始加劇，上述結構似乎在神經元個數和層數的簡單和複雜性之間取得了平衡，詳細參數如表 1：

表1 – 可訓練參數比較表

| AlexNet & VGG models |       |                            | Proposed Structure |       |                       |
|----------------------|-------|----------------------------|--------------------|-------|-----------------------|
| Layers               | Nodes | Parameters                 | Layers             | Nodes | Parameters            |
| Flatten Layer        | m     | 0                          | Flatten Layer      | m     | 0                     |
| FC-4,096             | 4,096 | $4,096*(m+1)$              | FC-1,024           | 1,024 | $1,024*(m+1)$         |
| FC-4,096             | 4,096 | $4,096*(4,096+1)$          | FC-512             | 512   | $512*(1,024+1)$       |
| FC-1,000             | 1,000 | $1,000*(4,096+1)$          | FC-256             | 256   | $256*(512+1)$         |
| Softmax              | s     | $s*(1,000+1)$              | FC-128             | 128   | $128*(256+1)$         |
|                      |       |                            | Softmax            | s     | $s*(128+1)$           |
| Total                |       | $4,096m+1,001s+20,882,408$ | Total              |       | $1,024m+129s+690,048$ |

AlexNet 總共有 6200 萬個參數，VGG-16 有 1.38 億個參數，VGG-19 有超過 1.48 億個參數。本研究計畫實驗結構總共只有大約 1100 萬個可訓練參數，甚至比上述結構小了一個數量級(詳見圖 5)。本研究計畫實驗結構雖然參數比較少，卻擁有強勁的性能。在裡用 TSRD 的資料庫初步測試下，AlexNet 的測試準確值低於 55%，VGG-16 約 64% 和 VGG-19 的 66% 相差不遠，而本研究計畫結構在沒有優化的情況下就可以達到 71% 的正確率。所模型均未經優化，且與最終精度相比仍有 10-20% 的差距。

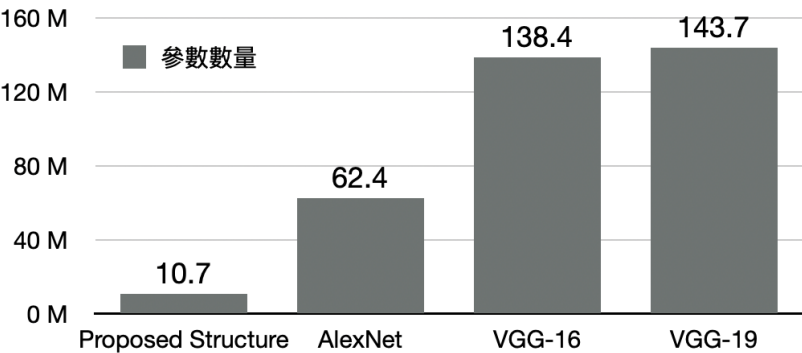


圖5 – 可訓練參數比較圖

### 3. 使用優化方法優化模型並增進判讀率

#### (1) 隨機關閉神經元

深度神經網絡受到過度擬合和梯度消失等問題的困擾，因此模型優化是使正確率升至頂尖的關鍵。文獻[24]顯示了在層數較淺前饋神經網絡中藉由隨機關閉神經元(Dropout)來增進訓練效果，並採用了不同的 dropout rate ( $p$ ) 進行了實驗。[25]在以 ImageNet 為範例顯示在更深層網絡中，隨機關閉神經元亦具有相似防止模型過度擬合的成效。由上述可知，隨機關閉神經元的效果已經得到驗證，且它們已存在於 AlexNet 和 VGG 結構中。

#### (2) 權重衰減

權重衰減亦是對抗過度擬合的武器之一。[26]推導出，這種簡單的方法可以抑制不相關的雜訊並消滅靜態噪聲，從而有助於模型的泛化。其中較著名的手法有 L1 正則化以及 L2 正則化。L1 正則化使用模型中的一些特徵而將某些捨棄，模型將會相對縮小，而 L2 正則化傾向保留著大部分的特徵，但每一項的權重係數都隨之縮小。

#### (3) 添加噪音

向類神經網絡添加噪音有助於提升未來模型的容錯能力，其工作原理與資料增強類似。當神經元記住特定數據的特徵時，神經元將學習局部特症而不是廣域的特徵，過度擬合就容易發生。增加噪音可以防止這種情況的發生。此外，訓練類神經網絡的降噪能力有助於模型進一步區分有用的資訊和瑣碎的細節，從而幫助類神經網絡在面對未知的資料時有更高的判讀力。

#### (4) 特徵縮放

特徵縮放也能幫助神經網絡學習。神經網絡使用梯度下降作為最佳化演算法，而該演算法對於輸入數據的縮放十分敏感，通過以某種形式統一饋入網絡的數據將對於學習有正面的效果。兩種常用的特徵縮放是正規化(Normalization)及標準化(Standardization)。前者在擁有不假設輸入數據為高斯分佈時能發揮較好的功效，而後者在輸入數據遵循常規分佈時將更有幫助。

特徵縮放甚至可能有助於處理內部協變量漂移的問題，在訓練過程中，為了應變輸入數據的分佈變化，必須降低學習率，並且初始化參數。內部協變量漂移會減慢訓練過程，尤其是在結構龐大時尤為明顯。在較大的類神經網絡中，[27]通過批標準化(Batch Normalization)解決了內部協變量漂移的問題，使模型達到了更好的性能。

### (五). 預期結果

考慮到許多數據量大的數據庫上正確率排行榜上時不時出現超過九成的高辨識率模型，又考慮到本研究實驗計畫的模型在初步測試階段表現優異，預期最終能擁有超過九成的正確率。但在考慮到有些數據庫的資料量過於偏頗，數據量不夠多與照片品質等多重因素之下，預計本計畫所提出的模型平均能繳出超過八成五的正確率。如果透過各個交通號誌資料庫的交叉訓練，模型的正確率理應大幅提升，這時便有機會挑戰人類辨識率(根據各資料庫略有不同，但普遍超過 96%)



## (六). 參考文獻

- [1] S. Krueger, J. Abele, C. Kerlen, H. Baum, T. Geißler, S. Grawenhoff, J. Schneider, and W. H. Schulz, "Exploratory study on the potential socio-economic impact of the introduction of intelligent safety systems in road vehicles," VDI/VDE Innovation + Technik GmbH, Institute for
- [2] Z. Liu, J. Du, F. Tian, and J. Wen, "MR-CNN: a multi-scale region-based convolutional neural network for small traffic sign recognition," *IEEE Access*, vol. 7, pp. 57120–57128, 2019.
- [3] Y. Wu, Y. Liu, J. Li, H. Liu, and X. Hu, "Traffic sign detection based on convolutional neural networks," in *The 2013 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, Dallas, TX, 2013.
- [4] Y. Zhu, C. Zhang, D. Zhou, X. Wang, X. Bai, and W. Liu, "Traffic sign detection and recognition using fully convolutional network guided proposals," *Neurocomputing*, vol. 214, pp. 758–766, 2016.
- [5] Hoang Nguyen, "Fast Traffic Sign Detection Approach Based on Lightweight Network and Multilayer Proposal Network", *Journal of Sensors*, vol. 2020, Article ID 8844348, 13 pages, 2020
- [6] Krizhevsky, Alex & Sutskever, Ilya & Hinton, Geoffrey. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems*. 25. 10.1145/3065386.
- [7] Tian, Y., Luo, P., Wang, X., & Tang, X. (2015). Pedestrian detection aided by deep learning semantic tasks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 5079–5087).
- [8] Li J., Mei X., Prokhorov D., Tao D. Deep neural network for structural prediction and lane detection in traffic scene, *IEEE Transactions on Neural Networks and Learning Systems*, 28 (3) (2017), pp. 690-703
- [9] Escalera A.D. L., Moreno L., Salichs M., Armingol J. Road traffic sign detection and classification, *IEEE Transactions on Industrial Electronics*, 44 (6) (1997), pp. 848-859
- [10] Shadeed, W., Abu-Al-Nadi, D., & Mismar, M. (2003). Road traffic sign detection in color images. In *10th IEEE international conference on electronics, circuits and systems*, 2003. *Proceedings of the 2003*, Vol. 2 (pp. 890–893).
- [11] Loy, Gareth & Barnes, Nick. (2004). Fast shape-based road sign detection for a driver assistance system. 1. 70 - 75 vol.1. 10.1109/IROS.2004.1389331.
- [12] Barnes, Nick, Gareth Loy, and David Shaw. "The regular polygon detector." *Pattern Recognition* 43.3 (2010): 592-602.
- [13] Mathias M., Timofte R., Benenson R., Van Gool L. Traffic sign recognition—How far are we from the solution?; *Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN)*; Dallas, TX, USA. 4–9 August 2013; pp. 1–8.
- [14] Chaiyakhon, Kedkarn & Hirunyanakul, Anusara & Chanklan, Ratiporn & Kerdprasop, Kittisak & Kerdprasop, Nittaya. (2015). Traffic Sign Classification using Support Vector Machine and Image Segmentation. 52-58. 10.12792/iciae2015.013.
- [15] DE Rumelhart, GE Hinton, RJ Williams, *Learning internal representations by error propagation*. 1985 – DTIC Document.

- [16] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [17] LeCun, Y.; Boser, B.; Denker, J. S.; Henderson, D.; Howard, R. E.; Hubbard, W. & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4):541-551.
- [18] Krizhevsky, Alex & Sutskever, Ilya & Hinton, Geoffrey. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems*. 25. 10.1145/3065386.
- [19] Simonyan, K. and Zisserman, A., “Very Deep Convolutional Networks for Large-Scale Image Recognition”, 2014.
- [20] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.
- [21] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [22] Lei Jimmy Ba and Rich Caruana. 2014. Do deep nets really need to be deep? In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2 (NIPS'14)*. MIT Press, Cambridge, MA, USA, 2654–2662.
- [23] Basha, S. H. Shabbeer et al. “Impact of Fully Connected Layers on Performance of Convolutional Neural Networks for Image Classification.” *ArXiv abs/1902.02771* (2020)
- [24] Srivastava, Nitish. "Improving neural networks with dropout." *University of Toronto* 182.566 (2013): 7.
- [25] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1 (January 2014), 1929–1958.
- [26] Krogh, Anders, and John A. Hertz. "A simple weight decay can improve generalization." *Advances in neural information processing systems*. 1992.
- [27] Loffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." *arXiv preprint arXiv:1502.03167* (2015).

## (七). 需要指導教授指導內容

1. 研究方法指導
2. 演算法建模訓練
3. 實驗設計與結果驗證分析