

OAC-112-023 (研究報告)

**基於深度學習計算機視覺技術之水下機器人
非法捕魚監控系統之研究與實現**

(正式報告)

海洋委員會補助研究

中華民國 112 年 8 月

「本研究報告僅供海洋委員會施政參考，並不代表該會政策，該會保留採用與否之權利。」

OAC-112-023 (研究報告)

基於深度學習計算機視覺技術之水下機器人 非法捕魚監控系統之研究與實現

(正式報告)

學校：國立澎湖科技大學

指導教授：陳良弼

學生：黃祥睿

研究期程：中華民國112年3月至112年10月

研究經費：新臺幣6萬元

海洋委員會補助研究

中華民國 112 年 8 月

「本研究報告僅供海洋委員會施政參考，並不代表該會政策，該會保留採用與否之權利。」

「本研究報告絕無侵害他人智慧財產權之情事，如有違背願自負民、刑事責任。」

目次

表次.....	iv
圖次.....	v
摘要.....	1
第一章 前言.....	3
1.1 研究緣起與問題背景.....	3
1.2 現況分析.....	3
1.3 研究目的及研究重點.....	6
1.4 預期目標.....	9
第二章 文獻回顧與探討.....	10
第三章 研究方法及過程.....	13
3.1 系統架構.....	13
3.2 非法捕魚行為分類的電腦視覺框架.....	14
3.2.1 資料預處理.....	14
3.2.3 整體系統流程.....	17
3.3 特徵注意力前置網路.....	18
第四章 實驗結果.....	20
4.1 Top-1 Accuracy 和 Top-5 Accuracy 評估分數.....	20
4.2 本計畫所使用的模型性能評估.....	20
4.3 模型效能優化.....	22
4.4 消融研究.....	24
4.5 AI 邊緣運算平台的可視化結果.....	25
第五章 結論.....	27
第六章 應用場域.....	29
第七章 計畫成果.....	30
參考文獻.....	34

表 次

表 1、進度表.....	9
表 2、Training and Test Original Datasets.....	17
表 3、Performance comparison between the trained models used in this paper on the illegal fishing image dataset.....	21
表 4、Training Setup Parameters.....	23
表 5、Training and Test Datasets.....	23
表 6、Performance Comparison of Different Methods in This Paper.....	24
表 7、Performance comparison results of ShuffleNet V2 among different modules.....	24
表 8、The actual test is on Jetson Nano.....	25

圖 次

圖 1、分數介於 1 到 5 之間，陰影較深的國家表現更差 [1].....	4
圖 2、本計畫系統架構圖.....	14
圖 3、Example of underwater fishing behavior samples. (a) Abnormal behavior.....	15
圖 4、Overall architecture of the proposed underwater fishing behavior classification framework. .	16
圖 5、Under water robot.	16
圖 6、The proposed FA-FENet architecture.	19

摘要

關鍵字：非法捕魚、計算機視覺系統、智慧物聯網 (AIoT)、遷移式學習

非法捕魚是一個全球性問題，為了解決非法捕魚問題，若能有一種能即時且自動的對捕魚行為進行分類的電腦視覺系統，用以自動化執行即時分類方法管理非法捕魚行為，便可以減少人力成本和效率。然而，至今這方面的實際應用仍然有限，因為水下非法捕魚行為並不受到重視，原因在於數據集不易收集且還有水下影像模糊不清問題。據我們所知，目前還沒有針對水下非法捕魚影像電腦視覺領域的相關研究，對此，本研究提出一種結合智慧物聯網(AIoT)的用於非法捕魚行為分類的電腦視覺框架，此框架由一個分類模型組成針對水下捕魚行為進行即時分類，並根據水下行為影像進行分析水下捕魚行為是否具有非法性，在分類模型中我們使用遷移式學習對預訓練模型進行參數微調提高模型性以建置更穩健的電腦視覺系統。目前，我們所收集的數據集為 20,500 張，分為正常水下行為和異常水下行為兩種類別，正常水下行為類別由 Scuba_Diver 和 VDD-C 兩種公開的水下潛水員數據組成，異常水下行為是蒐集水下持有非法捕魚設備人員(魚刺、漁網、魚槍、流刺網)。數據集中使用了數據增強方法補足數據集不足的情形，包括:旋轉、最大化和最小化過濾器、銳化、模糊、對比調整、像素打亂、邊緣增強、光滑度調整、亮度調整和顏色抖動調整。實驗結果顯示

出本研究提出的水下捕魚行為分類模型在 AI 邊緣運算平台上可在 20 毫秒內時間實現 100.00%的準確度。另外，本研究提出特徵注意力前置網路(FA-FENet)在實驗中有不錯的表現，即能提取水下捕魚模糊不清影像的特徵。

第一章 前言

1.1 研究緣起與問題背景

近年來隨著氣候劇烈變遷影響，非法捕魚問題成為了海洋永續發展的全球性重要議題，根據聯合國農糧署（Food and Agriculture Organization of the United Nations，簡稱為 FAO），全球已有超過 80% 魚種被人類開發利用，為了漁業能永續發展，全世界制定了相關的漁業收穫辦法以對非法捕魚行為進行的管理和監督。非法捕魚是指違反利用海洋和漁業資源的各種活動，稱為（illegal, unreported and unregulated fishing，簡稱 IUU）。

1.2 現況分析

如圖 1 所示，非法捕魚是一個全球性問題，發生在南中國海、非洲西海岸（據估計非法捕撈量為 40%）、南美洲兩岸、印度洋東部、整個大洋洲和南極洲周圍[1]。破壞性捕撈[2]在東南亞區域印度尼西亞、菲律賓和柬埔寨是普遍的，在 2018 年菲律賓薄荷島，漁民使用炸彈向海中發射，破壞性的捕撈海中漁獲使菲律賓大面積的珊瑚礁遭到破壞[3]。在 2020 年數百艘中國船隻在加拉帕戈斯群島沿岸非法捕魚[4]，非法的漁業入侵使智利每年造成 3.97 億美元的損失和漁業資源已逐漸崩潰[4]。

在牙買加南部加勒比海(southern coast of Jamaica)由於過度捕撈的情形，漁業資源造成嚴重的破壞[5]，如：使非法捕魚設備(魚叉和漁網)，其中在佩德羅島(Pedro Cay)在 2011 年有 10 艘外國船隻藏有大量潛水員進行非法捕

魚活動[5]。由於工業捕撈作業的影響，印度洋的魚種正處於崩潰的邊緣，以往日本在非洲和印度洋曾使用工業用的漁獲船進行捕獲大量的非法捕魚[6]。在 2018 年有多名漁業運販在意大利和馬爾他水域進行非法捕魚，非法的捕魚帶來的交易每年超過 1,200 萬歐元[7]，同一年在意大利也發現使用炸藥捕魚且在西非國家水域進行非法捕魚[7]。

非法捕魚的情形在臺灣也有許多案例，在 2019 年臺灣漁船以海豚為誘餌捕捉鯊魚[8]有 70 隻海豚被故意殺死。在 2022 年臺灣澎湖，因蒐集非法捕魚證據的觀察員在取締違法漁業者時，被漁業者傷害事件[9]。在 2021 年葡萄牙拖網漁船在加拿大水域進行違法捕魚，蒐集非法捕魚的證據的觀察員還遭到驅逐和恐嚇[10]。因此，全世界都有非法捕魚的案例，解決非法捕魚問題成為了本計畫主要的目的。

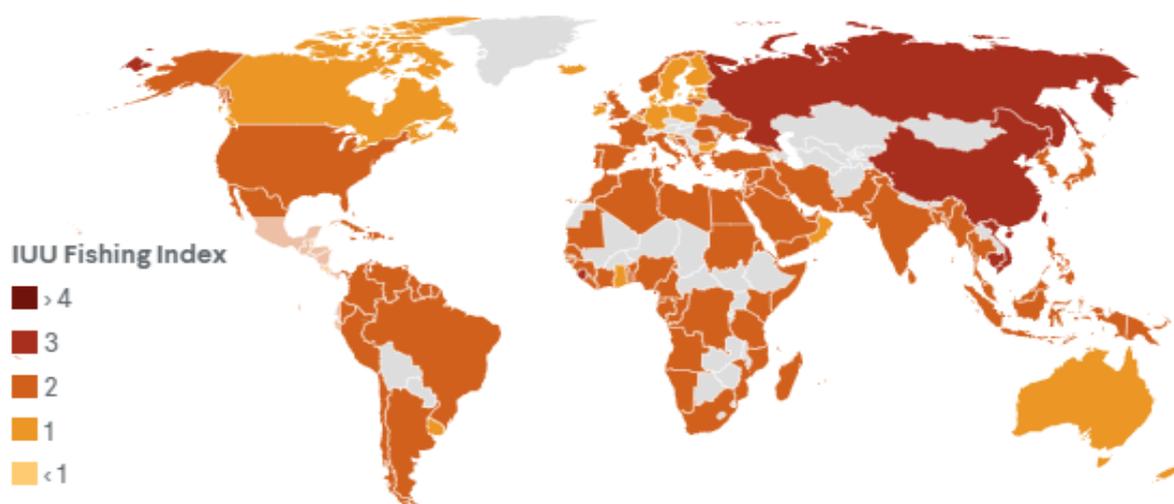


圖 1、分數介於 1 到 5 之間，陰影較深的國家表現更差 [1]

解決非法捕魚問題有許多困難需要克服，如：監督困難問題，在無人看管的地方進行捕魚，那裡沒有明確的管理者，屬於無人管理的區域，這種區域沒有明確的法規所以難以監督在此區域的非法捕魚的行為。改裝漁船並偽裝成其他船隻使海域的監督者難以辨識，通常改裝船舶外表會改裝成普通漁船、旅遊船，內部改裝非法漁具和使用非法潛水漁具進行捕魚。

因此，解決非法捕魚問題大多仰賴於巡邏海洋的監督者，隨著第四次工業革命，巡邏海洋的任務逐漸使用新興技術來替代傳統巡邏海洋方法，傳統大多是使用船隻完成，現今是由衛星、全球定位系統(GPS)和船載船舶監控系統(Automatic Identification System, AIS)[11]可幫助監管者輕鬆的跟踪漁船的位置，例如：使用監視器進行船舶監控和使用船隻標識(Unique Vessel Identifiers, UVI)進行船舶管理，使用物聯網技術和人工智慧技術幫助自動化遠程監控漁船，從而更容易監管非法捕魚活動。

在荷蘭已經支持在所有非法行為風險的歐盟漁船上安裝攝像頭[12]，除此之外，為了確保海產的來源是合法的，可追溯海產附有描述船隻、日期、地點、漁具和物種的文件，利用消費者對產品的需求以減少非法捕魚行為[13]。更嚴厲的處罰也是一個解決非法捕魚問題的方案，例如：阿塞拜疆修訂非法捕撈魚類和其他水生動物的罰款增加了 20 倍[14]。

物聯網技術在資通訊(ICT)領域具有重要潛力。報告預測，預計到 2026 年將達到 6,042.45 美元[15]，國際市場將在整個過程中以 24.7% 的驚人

復合年增長率增長估計的年份[15]。物聯網(Internet of Things 簡稱為 IoT)是一種透過網際網路的裝置互連本地及遠端設備與其他設備的新興技術，結合硬體層及軟體層的網路架構建置而成的技術，使網路設備有相當高的擴展性，同時不再侷限於傳統的網路媒體設備，即不同的設備(裝置)之間也能相互連通，透過與不同的設備及裝置間進行讀取並傳遞，達到智慧化控制、自動化及管理等技術及功能。如今，物聯網技術廣泛應用在智慧農業、智慧工業、智慧城市、智慧醫療及智慧家居等領域。近年來，人工智慧(Artificial Intelligence，簡稱為 AI)技術受到極大的關注，隨著軟體技術的日益成熟，AI 技術的應用趨勢持續性的成長[16]。

人工智慧為近年來主要發展趨勢，2026 年全球人工智慧市場預計將達到 45.1 億美元。人工智慧是用於代替人類決策的一種技術，此技術用於各個領域中促進人類社會型態的改變，如能將此技術導入漁業產業，有助於解決非法捕魚問題。在最近幾年，人工智慧結合物聯網的智慧物聯網(AIoT)技術，此技術將原本智慧化的設備(裝置)資料學習並自主做出決策，讓原本的物聯網技術更加的智慧化。

1.3 研究目的及研究重點

基於上述的問題，以下列出本研究計畫的主要貢獻，包括：本研究計畫所要解決技術問題和系統應用方法的創新上的立足點：

 在技術貢獻方面，針對水下捕魚行為數據不足，我們使用數據增

強方法補足了數據不足問題，包括：旋轉、最大化和最小化過濾器、銳化、模糊、對比調整、像素打亂、邊緣增強、光滑度調整、亮度調整和顏色抖動調整。

✚ 在技術貢獻方面，我們採用遷移式的深度學習方式進行模型效能優化並建置在 AI 邊緣運算平台上。實驗結果證明，在 AI 邊緣運算平台上在 20 毫秒內時間實現 100% 的準確度。

✚ 在技術貢獻方面，本計畫比較了現今新興的所有模型並選出最適用於水下捕魚行為分類的最優模型，並進行遷移式學習來優化模型提高準確度。

✚ 在系統應用創新方法方面，我們提出一種用於非法捕魚行為分類的電腦視覺框架，此框架由一個分類模型組成針對水下捕魚行為進行即時分類，並根據水下行為影像進行分析水下捕魚行為是否具有非法性，在分類模型中我們使用遷移式學習對預訓練模型進行參數微調提高模型性以建置更穩健的電腦視覺系統。

✚ 在系統應用創新方法方面，我們是第一個針對水下捕魚行為進行分類以驗證我們所提出的方法對於水下非法捕魚行為蒐證具有新穎性和創新性。

✚ 在系統應用創新方法方面，水下捕魚影像模糊問題在人力判讀過程是耗時且複雜的，我們提出一種基於深度學習的計算機視覺技

術之水下機器人非法捕魚監控系統以解決水下捕魚影像模糊問題
人力判讀過程是耗時且複雜的情形。

✚ 為了解決水下捕魚影像模糊問題，我們提出的特徵注意力前置網路 FA-FENet 表現了不錯的效能，能提取水下捕魚模糊不清影像的特徵。

✚ 在系統應用創新方法方面，我們使用智慧物聯網 (Artificial Intelligence over Internet of Things, 簡稱為 AIoT) 技術結合水下機器人 (Remotely operated underwater vehicle, 簡稱為 ROV)。

✚ 在系統應用創新方法方面，監督者可透過基於深度學習計算機視覺技術之水下機器人非法捕魚監控系統監控水下是否有非法捕魚行為，如有非法捕魚行為則通知監督者或執法人員進行水下非法捕魚的蒐證。

1.4 預期目標

以上已說明完本專題該如何運作，為了能讓本專題實現預期結果，故進度如表 1。

表 1、進度表

月次	112 年 3 月	112 年 4 月	112 年 5 月	112 年 6 月	112 年 7 月	112 年 8 月	112 年 9 月	112 年 10 月
工作項目								
收集資料								
尋找研究方法								
撰寫程式與改進								
彙整報告(期中及正式)								

本研究其餘部分安排如下。第二章的子章節 2.2 進行文獻回顧與探討，我們詳細討論關於現今非法捕魚應用以突顯本計畫使用電腦視覺技術解決非法捕魚問題的應用方法創新。第三章的子章節 3.1 我們會詳細描述了本計畫所提出的架構。在第三章的子章節 3.2 我們介紹了本計畫所提出的非法捕魚行為分類的電腦視覺框架，包括:數據集、資料預處理流程和訓練方法。在第三章的子章節 3.3 會介紹提出的特徵注意力前置網路。第四章我們會依照模型性能優缺點和問題並進行比較，然後選擇最適用的模型在水下捕魚行為分類任務上，並進行模型優化與改良和消融實驗。

第二章 文獻回顧與探討

為了解決非法捕魚問題，近年來有許多關於使用深度學習技術和物聯網技術解決非法捕魚問題的研究[17]-[23]，透過船載船舶監控系統(Automatic Identification System, 簡稱為 AIS)檢測非法捕魚行為[17], [20]透過合成孔徑雷達(synthetic aperture Radar, 簡稱為 SAR)影像檢測非法捕魚行為[19]透過船舶影像檢測非法捕魚行為[21]透過光學衛星影像檢測船舶影像[22]和使用物聯網技術的無人機進行船舶監控 (Vessel Monitoring System, 簡稱為 VMS) 檢測非法捕魚行為[23]。

Global Fishing Watch 平台[28]是一種結合 AI 和衛星數據可追蹤船隻活動雲端監控平台，使用的神經網路可辨識船隻大小和發動機功率，捕撈的類型和使用的漁具。Carlos 等人[17]並是使用 Global Fishing Watch 平台上的 IAS 數據集提出一種使用船載船舶監控系統(AIS)所分類的捕魚行為 IAS 數據集進行捕魚行為分類，通過監督自動編碼器(SA-DR)進行數據預處理以避免過度擬合情形，使用 Global Fishing Watch 平台上的 IAS 數據集，作者[17]所提出的分類模型針對 4 個類別(拖網捕魚行為、圍網行為、非法漁具釣魚行為、延吊繩行為)，模型提供 95%精準度。作者[17]所提出的方法數據來源需透過船載船舶監控系統(AIS)提供，船載船舶監控系統(AIS)可提供船舶 ID、船名、位置、航向和船速[18]。如果有漁船改裝成其他船舶，錯誤資訊傳入船載船舶監控系統(AIS)使作者[17]所提出的分類模型降低了分類

精準度，這樣無法有效的改善非法捕魚問題。另外，非法捕魚中有些漁船會利用潛水員進行非法捕魚，潛水員非法捕魚行為在東南亞時分常見，潛水員非法捕魚則無法使用分類模型[17]進行分類。

除了船載船舶監控系統(AIS)外也可以利用 SAR 進行船舶監控，Song 等人[19]提出一種 CNN 架構的血管辨識算法(Vessel Identification Algorithm) 增強合成孔徑雷達圖像海中不明確的船隻位置。SAR 的船舶監控在於不穩定氣候的海域，例如:暴風雨的大海和陽光不足的海域，可優於傳統的船載船舶監控系統(AIS)所檢測的船舶資訊。Young 等人[20]提出一種 R-CNN 架構的血管辨識算法(Vessel Identification Algorithm) 進行船舶監控，作者[20]所提出的方法可追蹤異常船載船舶監控系統(AIS)的船舶，改善 [17]方法的缺點，可辨識異常的船載船舶監控系統(AIS)的船舶，但潛水員非法捕魚問題仍無法解決，潛水員非法捕魚行為是無法透過船載船舶監控系統(AIS)檢測。

Syah 等人[21]提出一種使用卷積神經網路(Convolutional Neural Network, CNN)方法對船舶影像進行分類，類別分為兩種，漁船和軍用船舶，使用深度學習技術進行船舶檢測可對於船舶監控系統(Vessel Monitoring System, VMS)更加智慧化。實驗結果表明，Syah 等人[21]所提出的方法提供 90%的精準度。Kartal 和 Duman [22]使用卷積神經網路(Convolutional Neural Network, 簡稱為 CNN)方法對光學衛星的船舶影像(Satellite Imagery)進行

辨識。實驗結果顯示 Kartal 和 Duman [22]所提出的方法提供 99%的精準度。

另外，船舶監控系統(Vessel Monitoring System, VMS)解決非法捕魚問題存在挑戰，船舶監控系統(VMS)和船載船舶監控系統(AIS)存在同樣的問題，船舶關閉系統刻意傳入錯誤資訊。因此，Prayudi 等人[23]提出一種使用無人機系統檢測非法捕魚的船舶。實驗表明，此方法提供 75.03%的平均精度(AP)，96.44%的平均匹配精度(Amp)，使用 Faster-RCNN ResNet50 模型，整體系統運作為 SSD MobileNet V2 的輕量型架構。作者[23]使用無人機監控船舶的方法以解決錯誤資訊問題。在東南亞國家有許多小型漁船未安裝船舶監控系統(VMS)和船載船舶監控系統(AIS)，這種小型漁船則無法有效的監控。但是，潛水員非法捕魚問題仍無法解決，潛水員在海面下深處進行非法捕魚，在海面上的無人機無法及時有效的監控。

因此，上述文獻中並未針對水下非法捕魚行為問題的電腦視覺領域的研究，本計畫針對水下捕魚行為主要應用方法創新。我們所收集的水下捕魚行為數據集為正常水下行為和異常水下行為兩種類別，正常水下行為類別由 Scuba_Diver [50]和 VDD-C [51]兩種公開的水下潛水員數據組成，異常水下行為是蒐集水下持有非法捕魚設備人員(魚刺、漁網、魚槍、流刺網)。

第三章 研究方法及過程

本章節的子章節安排如下，首先我們會說明適用於本計畫的模型框架，然後介紹現今新興的模型。在實際上測試上，我們會在子章節 3.1 描述本計畫的系統架構，在子章節 3.2 呈現本計畫的非法捕魚行為分類的電腦視覺框架和訓練過程，子章節 3.3 為特徵注意力前置網路。

3.1 系統架構

探討完相關文獻後，我們查覺到現階段的研究未考慮到潛水員非法捕魚問題，潛水員非法捕魚無法透過船舶監控系統(VMS)和船載船舶監控系統(AIS)，潛水員非法捕魚定義在於使用非法捕魚設備(魚刺、漁網、魚槍、流刺網)。因此，本計畫提出一種使用計算機視覺技術之智慧物聯網水下機器人非法捕魚監控系統，使用水下機器人能蒐集水下的非法行為，圖 2 為本計畫系統框架圖，以下本計畫系統架構圖包含：子章節 3.1 非法捕魚行為分類的電腦視覺框架並詳細描述資料預處理和訓練平台。在子章節 3.2 詳細說明了特徵注意力前置網路細節。

如圖 2 所示，本計畫提一種基於深度學習的計算機視覺技術之水下機器人非法捕魚監控系統，以晶片系統技術和感測器技術來解決海洋生物多樣性及永續發展為主要研究議題。我們使用智慧物聯網(Artificial Intelligence over Internet of Things, 簡稱為 AIoT)技術結合水下機器人(Remotely operated underwater vehicle, 簡稱為 ROV)。整體操作流程是由監

督者可透過遙控器控制水下機器人進行水下影像蒐集，並將水下影像透過 4G LTE/5G 行動網路回傳影像至 AI 伺服器，當水下影像中有潛水員進行非法捕魚行為，AI 伺服器則即時辨識潛水員非法捕魚行為。另外，監督者可透過非法捕魚監控平台監控水下是否有非法捕魚行為，如有非法捕魚行為則通知監督者或執法人員進行水下非法捕魚的蒐證。

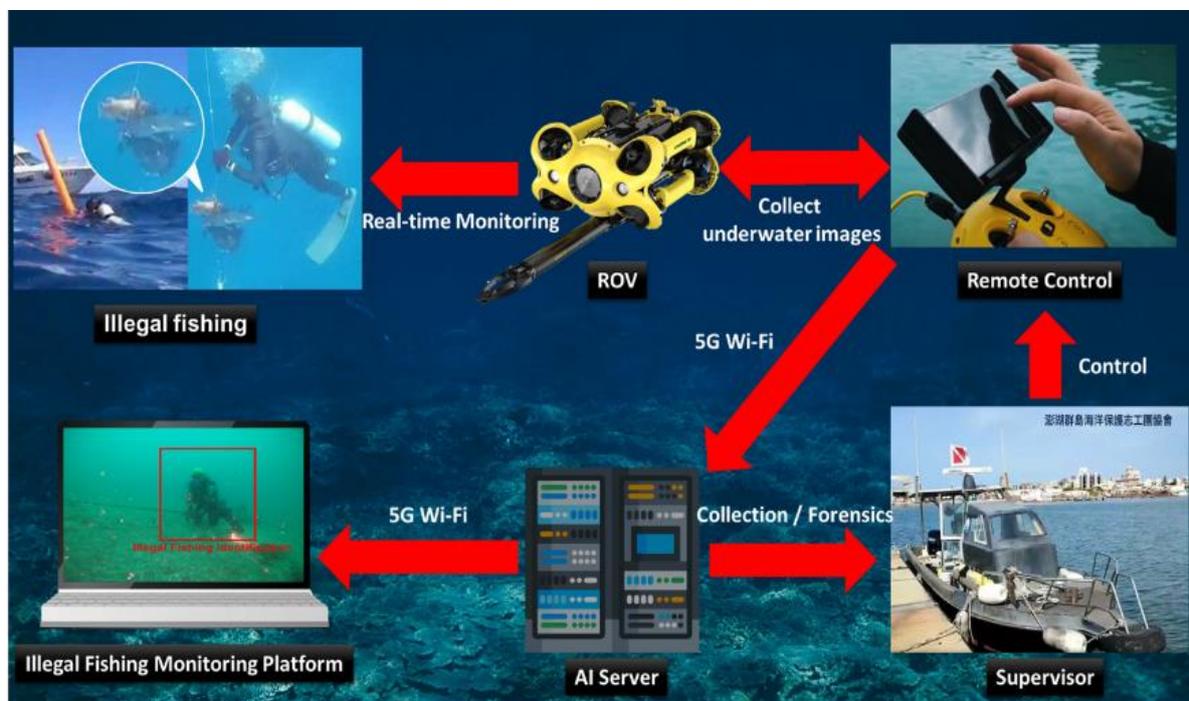


圖 2、本計畫系統架構圖

3.2 非法捕魚行為分類的電腦視覺框架

3.2.1 資料預處理

圖 3 說明了水下捕魚行為類別採樣範例。我們可以發現水下捕魚行為顏色、形狀及紋理特徵小、類別相似度高和特徵模糊問題。如圖 4 所示，非法捕魚行為分類的電腦視覺框架如下，首先我們將圖像餵入本計畫所提

出的訓練框架，我們總共分為三個部分 Meta、Train 和 Val，Meta 是用於部屬 Train 和 Val 的圖像位置。Train 為訓練集用於部屬訓練的圖像位置。Val 為測試集用於部屬測試的圖像位置。如表 2 所示，我們所收集的數據集為 15,000 張，分為正常水下行為和異常水下行為兩種類別，訓練集與測試集分為 9:1。

正常水下行為類別由 Scuba_Diver [50]和 VDD-C [51]兩種公開的水下潛水員數據組成，異常水下行為是由水下機器人(圖 5)蒐集水下持有非法捕魚設備人員(魚刺、漁網、魚槍、流刺網)。數據集中使用了數據增強方法補足數據集不足的情形，包括:旋轉、最大化和最小化過濾器、銳化、模糊、對比調整、像素打亂、邊緣增強、光滑度調整、亮度調整和顏色抖動調整。

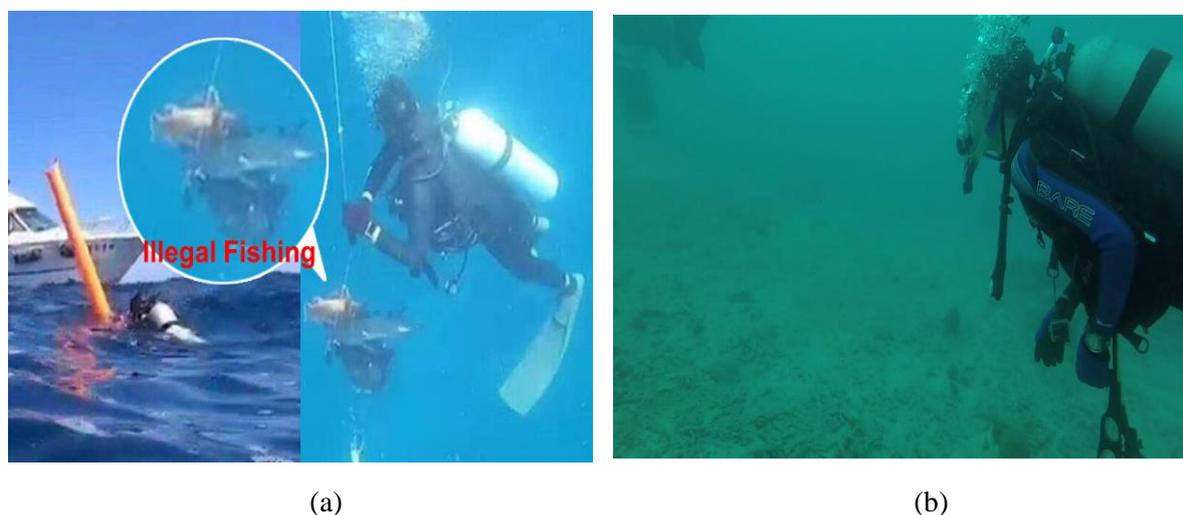


圖 3、Example of underwater fishing behavior samples. (a) Abnormal behavior. (b) Normal behavior.

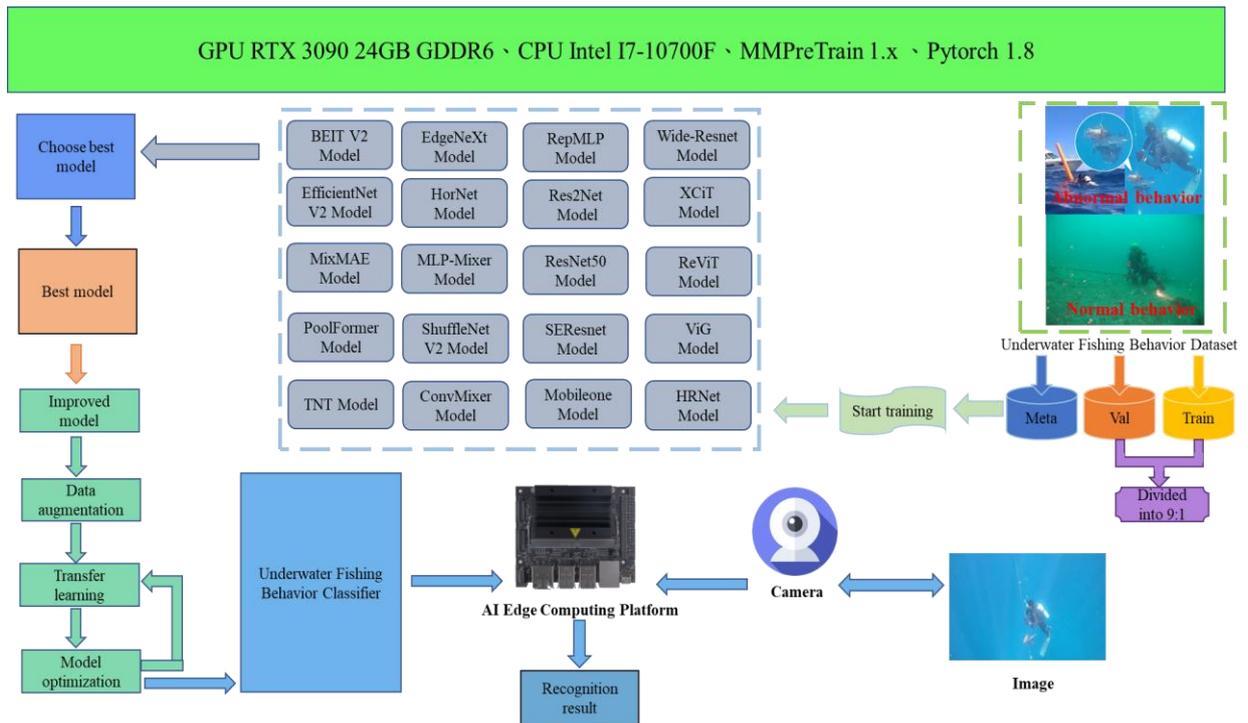


圖 4、Overall architecture of the proposed underwater fishing behavior classification framework.



圖 5、Under water robot.

表 2、Training and Test Original Datasets

Class	Training dataset	Test dataset
Normal behavior	6,750	750
Abnormal behavior	6,750	750
Total	13,500	1,500
All Images	15,000	

3.2.2 訓練平台

本計畫所使用的訓練平台是 GPU 為 RTX 3090 24GB GDDR6、CPU 為 Intel I7-10700F，並使用開源的 MMPreTrain 1.x [49] 作為分類工具，使用 PyTorch 1.8 作為訓練模組。我們會依序在以下子張節介紹我們的訓練框架細節。另外，我們會在第四章節介紹我們在訓練過程中的參數設置，並且比較模型性能選出最適用於水下非法捕魚分類的最優模型，並進行遷移式學習來優化模型提高準確度。

3.2.3 整體系統流程

基於以上貢獻點我們提出一種基於深度學習的計算機視覺技術之水下機器人非法捕魚監控系統，以下列出本系統操作流程：

- 監督者可透過遙控器控制水下機器人進行水下影像蒐集。
- 水下機器人將水下影像透過 4G LTE/5G 行動網路回傳影像至 AI 伺服器。
- 當水下影像中有潛水員進行非法捕魚行為，AI 伺服器則即時辨識潛水員非法捕魚行為。

- 監督者可透過非法捕魚監控平台監控水下是否有非法捕魚行為，如有非法捕魚行為則通知監督者或執法人員進行水下非法捕魚的蒐證。

3.3 特徵注意力前置網路

為了解決特徵模糊問題，我們提出一種特徵注意力前置網路全名為 Feature Attention Front-End Network 簡稱為 FA-FENet，如圖 6 所示。我們提出的前置網路可以提高對 CNN 主幹的特徵注意力。我們提出的 FA-FENet 有兩條通路，一條通路用藉由通道縮放來提升特徵注意力，網路佈局是 3×3 Conv、 3×3 Conv 和 2×2 Conv 的特徵壓縮注意力， 2×2 Conv、 3×3 Conv 和 3×3 Conv 的特徵放大注意力，另一條通路用以學習全局特徵的跳接設計。

另外，值得注意的是我們提出的 FA-FENet 為了不讓模型進行冗餘的卷積操作，我們的每一層的尺度規劃必須平衡通道設計，我們將 Front channel 設置為 3，整個網路佈局是分為兩部分，前期的特徵壓縮注意力，佈局是(3×3 Conv, Front channel = 3)、(3×3 Conv, Front channel = 6)和(2×2 Conv, Front channel = 9)，後期關注於壓縮的特徵進行重點學習(特徵放大注意力)，佈局是(2×2 Conv, Front channel = 6)、(3×3 , Conv Front channel = 3)和(3×3 Conv, Front channel = 6)。

這種佈局有別於 ResNet[43]的 bottleneck block(1×1 Conv、 3×3 Conv

和 1×1 Conv), bottleneck block 佈局是首先用 1×1 conv 降維, 其次藉由 3×3 conv 提取特徵, 再者是使用 1×1 Conv 升維。然而, 我們認為這種 bottleneck block 佈局不適用於模糊特徵, 模糊特徵導致網路前端學習成效低, 主要原因在於特徵搜索只依靠 3×3 Conv。

因此, 我們提出的網力主要改善模型的前端, 然後再與 CNN 主幹相連接, 有別於近年的注意力網路。一般而言, 注意力網路都會佈局在卷積層之後, 而我們提出的注意力網路是針對前端的注意力, 我們認為網路前端能主要關注重點特徵學習, 後期便不會學習到冗餘的特徵。請注意, 我們提出的 FA-FENet 是一種獨立的端到端的 CNN 架構可與不同的 CNN 主幹進行抽換, FA-FENet 可直接替換掉 CNN 主幹的前端 Conv1。我們將在第四章實驗結果報告加入特徵注意力前置網路後的 CNN 主幹的表現。

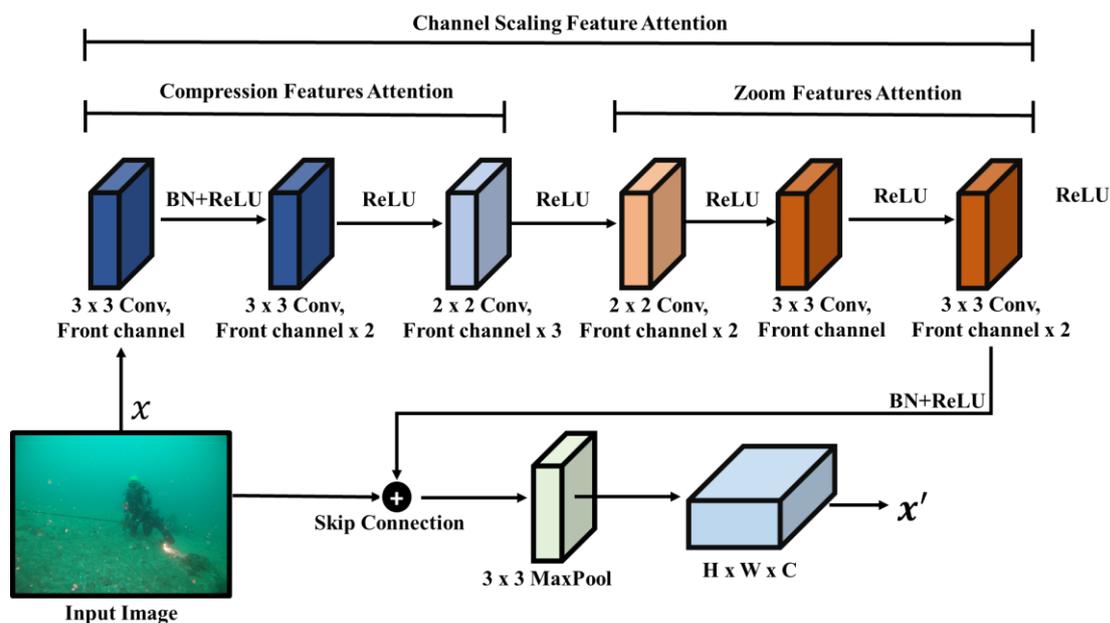


圖 6、The proposed FA-FENet architecture.

第四章 實驗結果

4.1 Top-1 Accuracy 和 Top-5 Accuracy 評估分數

本計畫使用的模型評估指標為多類別圖像分類任務常用的 Top-1 Accuracy 和 Top-5 Accuracy。Top-1 Accuracy 為預測概率分數最大的類別且預測正確的類別。Top-5 Accuracy 為實際結果相符且預測分數前五的類別。如公式 1 所示，Accuracy 計算是由 TP(True Positive)、FN、FP 和 TN。FN 為 False Negative，預測結果被判定為負樣本，但實際上是正樣本。FP 為 False Positive，預測結果被判定為正樣本，但實際上是負樣本。TN 為 True Negative，預測結果被判定為負樣本，實際也是負樣本。TP 為 True Positive，預測結果被判定為正樣本，實際也是正樣本。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

4.2 本計畫所使用的模型性能評估

表 2 說明了本計畫所使用的模型在水下捕魚行為分類任務上的性能評估。在 Top-1 Accuracy 和 Top-5 Accuracy 評估為 HorNet、MLP-Mixer、ShuffleNet V2、ConvMixer、RepMLP、Res2Net、ResNet50、SEResnet 和 Wide-Resnet 模型性能表現良好。在 Params(M)和 Flops(G) 評估為 ShuffleNet V2 模型性能表現良好。在 Weight Size 評估為 ShuffleNet V2 模型最小。在訓練

時間(Time)評估 ResNet50 模型時間最少。我們綜合以上表 2 模型之間比較的優缺點，我們選擇在水下捕魚行為分類任務上 Top-1 Accuracy 和 Top-5 Accuracy 表現良好的 ShuffleNet V2 模型進行模型最為本計畫的主要預訓練模型。

表 3、Performance comparison between the trained models used in this paper on the illegal fishing image dataset

Model	Top-1 Accuracy	Top-5 Accuracy	Params(M)	Flops(G)	Time	Epoch
BEIT V2 [29]	1.1236	100.0000	86.53	17.58	2 minutes	300
EdgeNeXt [30]	98.8764	100.0000	18.51	3.81	1 minutes	300
EfficientNet V2 [31]	53.9326	62.9213	14.36	3.50	12 minutes	300
HorNet [32]	100.0000	100.0000	88.42	15.42	7 minutes	300
MIXMIM [33]	1.1236	100.0000	114.67	16.35	12 minutes	300
MLP-Mixer [34]	100.0000	100.0000	59.88	1.61	4 minutes	300
PoolFormer [35]	1.1236	100.0000	56.17	8.96	10 minutes	300
ShuffleNet V2 [36]	95.2709	100.0000	2.28	0.15	5 minutes	300
TNT [37]	82.0225	100.0000	23.76	3.36	29 minutes	300
ConvMixer [38]	100.0000	100.0000	24.38	5.55	18 minutes	300
HRNet [39]	93.3820	100.0000	77.47	17.36	8 minutes	300
Mobileone [40]	1.1236	100.0000	14.84	2.98	16 minutes	300

Model	Top-1 Accuracy	Top-5 Accuracy	Params(M)	Flops(G)	Time	Epoch
RepMLP [41]	100.0000	100.0000	96.45	9.69	12 minutes	300
Res2Net [42]	100.0000	100.0000	45.21	8.12	7 minutes	300
ResNet50 [43]	100.0000	100.0000	25.56	4.12	2 minutes	300
ReViT [44]	17.9775	100.0000	87.34	17.49	17 minutes	300
SEResNet [45]	100.0000	100.0000	49.33	7.86	9 minutes	300
ViG [46]	7.8652	100.0000	95.21	16.86	20 minutes	300
Wide-Resnet [47]	100.0000	100.0000	68.88	11.44	8 minutes	300
XCiT [48]	97.7528	100.0000	6.72	1.24	31 minutes	300

4.3 模型效能優化

表 3 為本計畫所有原模型所設置的訓練參數。表 4 說明了原始數據和增強後數據的差別，原始數據總共為 15,000 張圖像增強後為 20,500 張圖像，訓練集與測試集維持在 9:1 的比例。如表 5 所示，我們使用數據增強方法增加數據量並使用遷移式學習的 fine-tuning 方法，learning rate 從 0.001 改為 0.01，Batch Size 從 32 改為 64，總共做了 100 步訓練，訓練時間約為 3 小時，Average Accuracy 從 95.27%改良到 100.00%，損失函數從 0.0713 降到 0.0683。

表 4、Training Setup Parameters

Model	learning rate	Channel	Batch size	Class	Weight decay	Num workers
BEIT V2 [29]	0.001	768	32	2	0.05	20
EdgeNeXt [30]	0.006	584	32	2	-	20
EfficientNet V2 [31]	0.001	1536	32	2	0.00001	20
HorNet [32]	0.004	1024	32	2	-	20
MIXMIM [33]	0.001	768	32	2	0.0001	20
MLP-Mixer [34]	0.003	768	32	2	0.3	20
PoolFormer [35]	0.0005	768	32	2	0.05	20
ShuffleNet V2 [36]	0.5	1024	32	2	0.00004	20
TNT [37]	0.003	3	32	2	0.05	20
ConvMixer [38]	0.0001	1024	32	2	1	20
HRNet [39]	0.1	2048	32	2	0.0001	20
Mobileone [40]	0.1	2048	32	2	0.0001	20
RepMLP [41]	0.0001	768	32	2	1	20
Res2Net [42]	0.1	2048	32	2	0.0001	20
ResNet50 [43]	0.001	2048	32	2	0.0001	20
ReViT [44]	0.0001	1536	32	2	0.05	20
SEResnet [45]	0.1	2048	32	2	0.0001	20
ViG [46]	0.1	1024	32	2	0.0001	20
Wide-Resnet [47]	0.1	1000	32	2	0.0001	20
XCiT [48]	0.0001	192	32	2	0.05	20

表 5、Training and Test Datasets

Original dataset		
Class	Training dataset	Test dataset
Normal behavior	6,750	750
Abnormal behavior	6,750	750
Total	13,500	1,500
All Images	15,000	
Data augmentation		
Class	Training dataset	Test dataset
Normal behavior	9,225	1,025
Abnormal behavior	9,225	1,025
Total	18,450	2,050
All Images	20,500	

表 6、Performance Comparison of Different Methods in This Paper

Method	Class	Average Accuracy rate	Loss	learning rate	Batch Size
This work proposed with Transfer learning based on pretrain model	2	96.61%	0.0683	0.01	64
ShuffleNet V2	2	95.27%	0.0713	0.001	32

4.4 消融研究

在上述實驗中，我們報告了 ShuffleNet V2 和 ResNet50 模型在水下捕魚行為分類任務中，除了 Top-1 Accuracy 和 Top-5 Accuracy 且有檢測速度快和推論時間少的優點。因此，我們選擇在水下捕魚行為分類任務上 Top-1 Accuracy 和 Top-5 Accuracy 表現良好的 CNN 主幹 ShuffleNet V2 模型進行模型最為本計畫的主要預訓練模型，AI 邊緣設備也可以進行水下捕魚行為的檢測效率，並且降低平台成本。表 6 我們報告了預訓練模型 ShuffleNet V2 在不同模組之間的消融實驗，模組的比較是基於 ShuffleNet V2 的原始 Bottleneck Block 和我們提出的 FA-FENet。實驗證明我們提出的方法，FA-FENet 使預訓練模型達到更好表現。

表 7、Performance comparison results of ShuffleNet V2 among different modules

Model	Top-1 Accuracy	Top-5 Accuracy	Params(M)	Flops(G)
ShuffleNet V2 [36] +FA-FENet (Ours)	97.04%	100%	2.286	0.15
+Original Bottleneck Block	96.61%	100%	2.283	0.08

4.5 AI 邊緣運算平台的可視化結果

為了實際測試了類別的準確性，我們將改良後的模型建置在 AI 邊緣運算平台，本計畫所選擇的 AI 邊緣運算平台為 Jetson Nano，表 7 表示了實際測試的可視化影像以異常捕魚行為為例和運算時間總體時間不超過 20 毫秒，我們將影像切割為數張圖像以利呈現模型類別的準確性。

表 8、The actual test is on Jetson Nano.

Image				
Seconds (ms)	19	14	17	16
Image				
Seconds (ms)	15	16	18	12
Image				
Seconds (ms)	14	18	12	18

Image	<p>Ground truth: 0 (abnormal behavior) Prediction: 0, 1.00 (abnormal behavior)</p> 	<p>Ground truth: 0 (abnormal behavior) Prediction: 0, 1.00 (abnormal behavior)</p> 	<p>Ground truth: 0 (abnormal behavior) Prediction: 0, 1.00 (abnormal behavior)</p> 	<p>Ground truth: 0 (abnormal behavior) Prediction: 0, 1.00 (abnormal behavior)</p> 
Seconds (ms)	15	17	16	17
Image	<p>Ground truth: 0 (abnormal behavior) Prediction: 0, 1.00 (abnormal behavior)</p> 	<p>Ground truth: 0 (abnormal behavior) Prediction: 0, 1.00 (abnormal behavior)</p> 	<p>Ground truth: 0 (abnormal behavior) Prediction: 0, 1.00 (abnormal behavior)</p> 	<p>Ground truth: 0 (abnormal behavior) Prediction: 0, 1.00 (abnormal behavior)</p> 
Seconds (ms)	15	17	16	17

第五章 結論

為了解決水下非法捕魚行為分類任務的人工判讀問題，我們選擇近幾年來優秀的模型進行蘭花圖像分類，包括: BEIT V2、EdgeNeXt、EfficientNet V2、HorNet、MIXMIM、MLP-Mixer、PoolFormer、ShuffleNet V2、TNT、ConvMixer、HRNet、Mobileone、RepMLP、Res2Net、ResNet50、ReViT、SEResnet、ViG、Wide-Resnet 和 XCiT。經過實驗結果證明，經過遷移式學習的分類模型 ShuffleNet V2 在水下捕魚行為分類在各方面得到不錯的成果。我們分析本計畫所訓練的模型在水下捕魚行為分類任務進行以下總結:

- ✚ 我們使用數據增強方法增加數據量並使用遷移式學習的 fine-tuning 方法，learning rate 從 0.001 改為 0.01，Batch Size 從 32 改為 64，總共做了 100 步訓練，訓練時間約為 3 小時，Average Accuracy 從 95.27% 改良到 96.61%，損失函數從 0.0713 降到 0.0683。
- ✚ 我們分析在 Top-1 Accuracy 和 Top-5 Accuracy 評估為 HorNet、MLP-Mixer、ShuffleNet V2、ConvMixer、RepMLP、Res2Net、ResNet50、SEResnet 和 Wide-Resnet 模型性能表現良好，其次為 TNT。
- ✚ 我們分析在 Params(M)和 Flops(G) 評估為 ShuffleNet V2 模型性能表現良好。
- ✚ 我們分析在 Weight Size 評估為 ShuffleNet V2 模型最小。
- ✚ 我們分析在訓練時間(Time)評估 ResNet50 模型時間最少。

- ✚ 我們綜合以上模型之間比較的優缺點，我們選擇在水下捕魚行為分類任務上 Top-1 Accuracy 和 Top-5 Accuracy 表現良好的 ShuffleNet V2 模型進行模型最為本計畫的主要預訓練模型。
- ✚ 我們分析的模型之中，在 Top-1 Accuracy 評估分數中，BEIT V2、EdgeNeXt、EfficientNet V2、MIXMIM、PoolFormer、HRNet、Mobileone、ReViT 和 ViG 表現較不佳，這些模型中對於特徵的處理各有不同方法，水下捕魚影像模糊成為了主要問題。
- ✚ 我們分析的模型之中，在 Top-1 Accuracy 評估分數中，HorNet、MLP-Mixer、ShuffleNet V2、ConvMixer、RepMLP、Res2Net、ResNet50、SEResnet、Wide-Resnet 和 TNT 表現都達到 80%以上。
- ✚ 我們分析的模型之中，ResNet 的變異模型，包括:Res2Net、ResNet50、SEResNet 和 Wide-Resnet 表現良好，代表 ResNet 架構適合提取水下捕魚模糊不清影像的特徵。
- ✚ 實驗證明，我們提出的特徵注意力前置網路 FA-FENet 表現了不錯的效能，能提取水下捕魚模糊不清影像的特徵。

第六章 應用場域

水下機器人屬於海洋科學與技術應用的範疇之一，不管是軍事任務的應用或是民用事業的應用，水下機器人發揮了重要的作用，而水下機器人替代了水下工作人員執行水下的危險任務，例如：水下機器人中的無人水下航行器（unmanned underwater vehicle, UUV）可長期在海中工作且高機動性，可執行各種水下任務，UUV 有許多子分類，例如：自主水下航行器（Autonomous Underwater Vehicle, AUV）、遙控潛水器（Remotely Operated Vehicle, ROV）、混合 ROV&AUV 的自主遠程操作車輛(Autonomous & Remotely Operated Vehicle, ARV)[24]，在軍事方面可用於海岸定位、海岸巡航[25]、港口監視船舶、跟蹤船舶，在一般民用方面可用於水底建築檢查、水底設備檢查、水底災害救助及海洋工程，在海洋領域的學術研究方面可用於海底環境 3D 製圖[26]、海洋資源的收集[27]和海洋資源探測。

因此，水域的監督者進行蒐集非法捕魚的證據時，便可透過水下機器人進行水下蒐證。使用水下機器人的優點在於安全性和便利性，使用者可安全且便利的進行水下蒐證。水下機器人可回傳即時的水中影像，並透過物聯網的無線區域網路(4G LTE/5G)上傳到非法捕魚監控平台。非法捕魚監控平台使用深度學習技術將回傳即時的水中影像進行辨識，若有非法捕魚行為即提醒監督者。

第七章 計畫成果

本計畫完成基於深度學習計算機視覺技術之水下機器人非法捕魚監控系統之研究與實現，同時本計畫相關研發成果投稿 2023 年 IEEE 全球消費研討會(2023 IEEE Global Conference on Consumer Electronics (GCCE 2023)) 將在今年 10 月於日本奈良英語口說發表論文，本計畫成果參加 2023 教育部智慧晶片系統應用創新專題實作競賽(Intelligent SoC Innovative Project Contest)智慧環境類組，入圍決賽入選獎。未來工作方面，預計將本計畫相關研究工作投稿於 IEEE Transactions on Instrumentation and Measurement 國際期刊。

【決賽通知1：請入圍決賽團隊留意事項】2023教育部智慧晶片系統應用



智慧晶片系統應用創新專題實作競賽 <isocipcontest@gmail.com>
寄給 智慧晶片系統應用創新專題實作競賽、陳春侯教授、密件副本：我

- (這是由主辦單位寄發決賽通知之第1封email。主辦單位預訂陸續於112年6月30位將「僅」透過email (iSoCIPContest@gmail.com) 及競賽網頁 (<https://www.isocipcontest.org/>)

您好！

恭喜您的團隊入圍決賽！

本競賽由教育部指導，國立高雄大學工學院主辦。協辦單位為國立成功大學電機系及教

- ★ 決賽日期：112年7月8日 (星期六)
- ★ 決賽時間：預定10:30~17:30 (10:30~11:40報到)
- ★ 決賽地點：國立高雄大學工學院大樓一樓
- ★ 決賽型式：實體決賽

1. 採Blind Review。參賽團隊不得以任何形式揭露可識別校名、系名、實驗室之獲獎資格。
2. 每一團隊之參賽作品報告 + 作品演示約10分鐘，評審問答約5分鐘。

★ 競賽網頁：<https://www.isocipcontest.org/>

另外，請您留意下列日期及事項：

1. 112年6月19日 (星期一) 或112年6月20日 (星期二)

請參賽團隊成員「務必」擇一場聆聽「智慧財產權法導論」演講
依教育部指示，主辦單位將線上播放「智慧財產權法導論」演講，為大家建立智慧財產
演講題：智慧財產權法導論—以專利法的重要概念為核心
演講人：謝國廉教授
國立高雄大學法學院教授

入圍 2023 教育部智慧晶片系統應用創新專題實作競賽證明



參加 2023 教育部智慧晶片系統應用創新專題實作競賽之實況

[IEEE GCCE 2023] Notification of your paper #1570920027 'A Deep Learning

edas=ieee-gc...@edas.info

寄給我、Guan-Zhi、Liang-Bi

英文

中文 (繁體)

[翻譯郵件](#)

Dear Mr. Xiang-Rui Huang:

Congratulations!

Your paper #1570920027 'A Deep Learning-Based ROV Surveillance System for Monitoring Illegal Fishing' has been accepted in C

The reviews are below.

Some papers submitted as the Oral category have been moved to Poster sessions against authors' preferences because of our co

To make presentations at the conference, the accepted authors must submit final camera-read manuscripts and IEEE copyright for At least an early-bird registration by one of the authors is also required.

More detailed information will be posted soon on the following page, so please visit and read the instruction carefully.

<https://www.ieee-gcce.org/2023/instructions.html>

Thank you very much for your patience, and we are looking forward to seeing you at the conference.

2023 年 IEEE 第 12 屆全球消費研討會(2023 IEEE 12th Global Conference on Consumer Electronics (GCCE 2023))論文接受函

A Deep Learning-based ROV Surveillance System for Monitoring Illegal Fishing

Xiang-Rui Huang, Guan-Zhi Huang, and Liang-Bi Chen*
Department of Computer Science and Information Engineering,
National Penghu University of Science and Technology, Penghu, Taiwan
E-mail: *liangbi.chen@gmail.com

Abstract—The issue of illegal fishing is a global concern that requires attention. This paper focuses on marine biodiversity and sustainable development and proposes a deep learning-based computer vision technology for monitoring illegal fishing using ROV surveillance systems. By combining deep learning-based technology with a remotely operated underwater vehicle (ROV), supervisors can remotely control the underwater robot to take images and send them to the AI server through a 4G/5G mobile network. The AI server quickly identifies any instances of illegal fishing by divers in underwater images. Additionally, supervisors can use the illegal fishing monitoring platform to monitor for any illicit activity fishing underwater.

Keywords—AIoT; illegal fishing; deep learning

I. INTRODUCTION

In recent years, with the impact of drastic changes in climate, the issue of illegal fishing has become an important global issue for sustainable marine development. According to the Food and Agriculture Organization (FAO) of the United Nations, more than 80% of fish species are developed and utilized by human beings. For the sustainable development of fisheries, the world has formulated relevant fishery harvesting methods to manage and supervise illegal fishing behaviors. Illegal fishing refers to a variety of activities that violate the use of marine and fishery resources, known as illegal, unreported, and unregulated (IUU).

The vessel monitoring system (VMS) is a solution to the problem of illegal fishing, but the same problem exists between the VMS and vessel automatic identification system (AIS). In 2011, there were 10 foreign vessels on Pedro Cay with a large number of divers using illegal fishing equipment (harpoons, fishing nets) to conduct illegal underwater fishing activities [1]. Therefore, the problem of illegal underwater fishing has not yet been resolved. To solve the problem of illegal fishing, there have been many studies on the use of deep learning technology and IoT technology in recent years [2]-[6]. For example, the Global Fishing Watch platform [2] is a cloud-based monitoring platform that combines AI and satellite data to track boat activity. The neural network used can identify boat size and engine power, fishing type and fishing gear used.

Carlos *et al.* [3] used the IAS dataset on the Global Fishing Watch platform to propose a fishing behavior classification using the fishing behavior IAS dataset classified by the AIS. Young *et al.* [4] proposed a vessel identification algorithm based on the R-CNN architecture for ship monitoring. The disadvantage of this method is that it can identify abnormal ships with an AIS, but the problem of illegal fishing by divers still

This work is supported in part by the Ocean Affairs Council (OAC), Taiwan, under Grant OAC-112-023. This work is also supported by the Higher Education Sprout Project, Ministry of Education (MoE), Taiwan, under Grant MOE-112G0004-1. *Corresponding author: Prof. Liang-Bi Chen.



Fig. 1. System construction of the proposed system.

cannot be solved. The illegal fishing behavior of divers cannot be detected by AIS.

Syah *et al.* [5] proposed a method to classify ship images using a CNN method. The categories are divided into two categories: fishing boats and military ships. Using deep learning technology for ship detection can be used for ship monitoring. The experimental results showed that the proposed method provided 90% accuracy. Kartal and Duman [6] used a CNN method to identify the satellite imagery of optical satellites. Experimental results showed that the proposed method provided 99% accuracy.

However, there is no research in the field of computer vision on the problem of illegal underwater fishing behavior in the abovementioned literature. The underwater fishing behavior dataset we collected is divided into two categories: normal underwater behavior and abnormal underwater behavior. The normal underwater behavior category consists of the normal behavior data of underwater divers. Fish equipment personnel (fishbone, fishing net, spear gun, gillnet) collected data and identified and analyzed underwater fishing behavior.

II. THE PROPOSED SYSTEM

Underwater illegal fishing is defined as the use of illegal fishing equipment (bone fishing, submersible fishing). Therefore, this paper proposes a remotely operated underwater vehicle (ROV)-based illegal fishing monitoring system using deep learning technology, which can use ROVs [7] to collect underwater illegal activities. Fig. 1 shows the system architecture of the proposed system.

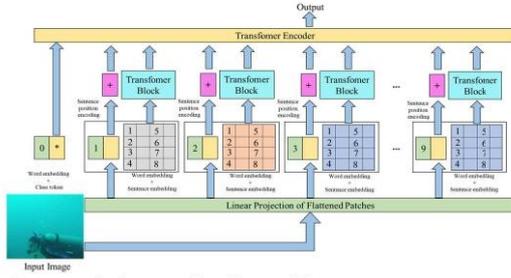


Fig. 2. Pretrained TNT model architecture [8].



Fig. 3. Recognition results. (a) Normal behavior. (b) Abnormal behavior.

As shown in Fig. 1, when the supervisors of the waters collect evidence of illegal fishing, they can use the ROV to search for evidence underwater. The advantages of using an ROV lie in safety and convenience, and users can conduct underwater evidence searches safely and conveniently. The ROV can return real-time underwater images and upload them to the illegal fishing monitoring platform through the 4G LTE/5G mobile networks. The illegal fishing monitoring platform uses deep learning technology to recognize real-time underwater images and alerts the supervisor if there is illegal fishing.

In this paper, we define illegal fishing as the use of illegal fishing equipment (fishbone, hook, submersible). The illegal fishing behavior images are fed into the illegal fishing recognition model for training. The dataset is approximately 5,000 images, the overall training and testing ratio is divided into 9:1, and the images in each category are averaged to allow the model to provide excellent accuracy. High-accuracy illegal fishing models can recognize highly variable blurred images in the sea, allowing supervisors to better detect illegal fishing divers. As shown in Fig. 2, the pretrained model TNT [8] is selected in this work. As shown in Table 1, the experimental results show that when the epoch is 264, the pretrained model reaches an accuracy of 81.0225, the loss function is 0.2018, the size of the TNT model is 280 MB, Params (M) is 23.76, and Flops (G) is 3.36. As shown in Table 2, the total training time is 42 minutes, and each step in the training process requires 11,363 memory. The following are the parameters of the TNT model set in this paper. The total category is 2, the channel is 3, the batch size is 32, the score is verified every 5 steps (num workers), the initial learning rate is set to 0.003, and the weight decay is 0.05.

Fig. 3 demonstrates the experimental results of the underwater fishing behavior recognition model. Fig. 3(a) shows

TABLE I. EVALUATION SCORE VALIDATION ON UNDERWATER FISHING BEHAVIOR DATASET

Model	Epoch	Accuracy	Loss function	Size	Params (M)	Flops (G)
This work	264	80.0225	0.2018	280	23.76	3.36

TABLE II. TRAINING PARAMETER SETTINGS

Model	Time (m)	Weight decay	Class	Channel	Batch size	learning rate
This work	42	0.05	2	33	23.76	0.003

that a normal diver does not carry illegal diving equipment. The underwater fishing behavior identification model identifies normal behavior. The net is illegal diving equipment, and the underwater fishing behavior identification model identifies it as an abnormal behavior, as shown in Fig. 3(b).

III. CONCLUSION

In this paper, an underwater robot illegal fishing monitoring system based on deep learning computer vision technology is proposed. When the supervisor collects evidence of illegal fishing, he can use the underwater robot to search for evidence underwater. The advantage of the underwater robot is safety, and convenience. The underwater image has a high-variable blur problem. The model proposed in this paper has an epoch of 264, and the pretrained model achieves an accuracy of 81.0225 and a loss function of 0.2018. Our proposed system uses an underwater robot to return real-time underwater images, which are uploaded to the illegal fishing monitoring platform via a 4G LTE/5G mobile network. The illegal fishing monitoring platform uses deep learning technology to identify the real-time underwater images returned. If there is illegal fishing, it will remind the supervisor to conduct underwater evidence search.

REFERENCES

- [1] Overharvesting and Poaching Devastate Jamaica's Fisheries. [Online]. Available: <https://insightcrime.org/investigations/overharvesting-poaching-devastate-jamaica-fisheries-nuu/>
- [2] Global Fishing Watch: Transparency for a Sustainable Ocean. [Online]. Available: <https://globalfishingwatch.org/>
- [3] H. Carlos, R. Aranda, M. R.-D. Velasco, A. Y. Rodríguez-González and M. E. Méndez-López, "Fishing Gear Pattern Recognition by Including Supervised Autoencoder Dimensional Reduction," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022, Art no. 8013105.
- [4] D. L. Young, "Deep nets spotlight illegal, unreported, unregulated (IUU) Fishing," *2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pp. 1-7, 2019.
- [5] A. Syah, M. Wulandari, and D. Gunawan, "Fishing and military ship recognition using parameters of convolutional neural network," in *Proc. 2020 3rd International Conference on Information and Communications Technology (ICOLACT)*, pp. 286-290, 2020.
- [6] M. Kartal and O. Duman, "Ship detection from optical satellite images with deep learning," in *Proc. 2019 9th International Conference on Recent Advances in Space Technologies (RAST)*, pp. 479-484, 2019.
- [7] X.-R. Huang and L.-B. Chen, "An underwater explorer remotely operated vehicle: Unraveling the secrets of the ocean," *IEEE Potentials*, vol. 42, no. 3, pp. 31-36, May-June 2023, doi: 10.1109/MPOT.2022.3233713.
- [8] K. Han, et al., "Transformer in transformer," *Neural Information Processing Systems*, vol. 34, pp. 15908-15919, 2021.

參考文獻

- [1] Illegal Fishing is a Global Threat. [Online]. Available: <https://www.cfr.org/article/illegal-fishing-global-threat-heres-how-combat-it>
- [2] The Horrors of Fishing with Dynamite. [Online]. Available: <https://www.nytimes.com/2015/02/05/business/energy-environment/the-horrors-of-fishing-with-dynamite.html>
- [3] In the Philippines, Dynamite Fishing Decimates Entire Ocean Food Chains. [Online]. Available: <https://www.nytimes.com/2018/06/15/world/asia/philippines-dynamite-fishing-coral.html>
- [4] Fishing in Dangerous Waters: A New Flashpoint for Latin America and China. [Online]. Available: <https://www.brinknews.com/fishing-in-dangerous-waters-a-new-flashpoint-for-latin-america-and-china/>
- [5] Overharvesting and Poaching Devastate Jamaica's Fisheries. [Online]. Available: <https://insightcrime.org/investigations/overharvesting-poaching-devastate-jamaica-fisheries-iuu/>
- [6] Satellites Reveal Japan's Mauritius Base for Africa Fishing Operations. [Online]. Available: <https://www.forbes.com/sites/nishandegnarain/2020/12/14/japan-vs-nature-satellites-reveal-japans-mauritius-base-for-africa-fishing-operations/?sh=18dc1f1b6be7>
- [7] Shocking cases of illegal fishing in EU highlight the need for stronger fisheries control. [Online]. Available: <https://www.cht.com.tw/home/campaign/5genterprise/CHT-Smart->

Ship.html

- [8] China and Taiwan named in us report against illegal fishing – greater transparency needed. [Online]. Available: <https://ejfoundation.org/news-media/china-and-taiwan-named-in-us-report-against-illegal-fishing-ejf-calls-for-greater-transparency-3>
- [9] <https://news.ltn.com.tw/news/life/breakingnews/4037376>
- [10] More illegal fishing 'notices' levelled against Portuguese draggers outside Canada's 200-mile limit. [Online]. Available: <https://www.sea-nl.ca/post/more-illegal-fishing-notices-levelled-against-portuguese-draggers-outside-canada-s-200-mile-limit>
- [11] Wikipedia, Automatic identification system. [Online]. Available: <https://www.credit.com.tw/NewCreditOnline/Epaper/ThemeContent.aspx?sn=78&unit=491>
- [12] NGOs call on EU to require electronic monitoring to stop illegal fishing. [Online]. Available: <https://www.seafoodsource.com/news/environment-sustainability/ngos-call-on-eu-to-require-electronic-monitoring-to-stop-illegal-fishing>
- [13] Five Ways to Fight Illegal Fishing. [Online]. Available: <https://sharkresearch.earth.miami.edu/five-ways-to-fight-illegal-fishing/>
- [14] Stronger penalties imposed to stop illegal fishing. [Online]. Available: <https://www.azernews.az/nation/114281.html>
- [15] Top 7 Trends in IoT to Look Out for in 2021. [Online]. Available: <https://dzone.com/articles/top-7-trends-in-iot-to-look-out-for-in-2021>
- [16] AI Trends: What can we expect from Artificial Intelligence in 2020. [Online]. Available: <https://www.xorlogics.com/2019/12/30/ai-trends-what-can-we-expect-from-artificial-intelligence-in-2020/>

- [17] H. Carlos, R. Aranda, M. R.-D. Velasco, A. Y. Rodríguez-González and M. E. Méndez-López, “Fishing Gear Pattern Recognition by Including Supervised Autoencoder Dimensional Reduction,” *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022, Art no. 8013105.
- [18] <https://transport-curation.nat.gov.tw/portAuthority/ais.html>
- [19] J. Song, D.-j. Kim, S. An, and J. Kim, “Restoration of Authentic Position of Unidentified Vessels in SAR Imagery: A Deep Learning Based Approach,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 1064-1078, 2022.
- [20] D. L. Young, “Deep Nets Spotlight Illegal, Unreported, Unregulated (IUU) Fishing,” in *Proc. 2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, 2019, pp. 1-7.
- [21] A. Syah, M. Wulandari, and D. Gunawan, “Fishing and Military Ship Recognition using Parameters of Convolutional Neural Network,” in *Proc. 2020 3rd International Conference on Information and Communications Technology (ICOIACT)*, 2020, pp. 286-290.
- [22] M. Kartal and O. Duman, “Ship Detection from Optical Satellite Images with Deep Learning,” in *Proc. 2019 9th International Conference on Recent Advances in Space Technologies (RAST)*, 2019, pp. 479-484.
- [23] A. Prayudi, I. A. Sulistijono, A. Risnumawan, and Z. Darojah, “Surveillance System for Illegal Fishing Prevention on UAV Imagery Using Computer Vision,” *2020 International Electronics Symposium (IES)*, 2020, pp. 385-391.
- [24] F. Kong, Y. Guo, and W. Lyu, “Dynamics Modeling and Motion Control of a New Unmanned Underwater Vehicle,” *IEEE Access*, vol. 8, pp. 30119-30126, 2020.

- [25] P. Yao, Z. Zhao, and Q. Zhu, “Path Planning for Autonomous Underwater Vehicles with Simultaneous Arrival in Ocean Environment,” *IEEE Systems Journal*, vol. 14, no. 3, pp. 3185-3193, Sep. 2020.
- [26] G. Salavasidis *et al.*, “Terrain-Aided Navigation with Coarse Maps—Toward an Arctic Crossing With an AUV,” *IEEE Journal of Oceanic Engineering*, vol. 46, no. 4, pp. 1192-1212, Oct. 2021.
- [27] C. Hu, Y. Pu, F. Yang, R. Zhao, A. Alrawais, and T. Xiang, “Secure and Efficient Data Collection and Storage of IoT in Smart Ocean,” *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9980-9994, Oct. 2020.
- [28] Global Fishing Watch: Transparency for a Sustainable Ocean. [Online]. Available: <https://globalfishingwatch.org/>
- [29] Z. Peng, *et al.*, “BEiT v2: Masked image modeling with vector-quantized visual tokenizers,” 2022, *arXiv:2208.06366*. [Online]. Available: <https://arxiv.org/abs/2208.06366>
- [30] M. Maaz, *et al.*, “EdgeNeXt: Efficiently amalgamated cnn-transformer architecture for mobile vision applications,” *Computer Vision—ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, Proceedings, Part VII. Cham: Springer Nature Switzerland*, 2023. pp. 3-20. [Online]. Available: <https://arxiv.org/abs/2206.10589>
- [31] M. Tan and Q. Le, “EfficientNetV2: Smaller models and faster training,” in *Proc. International conference on machine learning*, 2021, pp. 10096-10106. [Online]. Available: <https://arxiv.org/abs/2104.00298>
- [32] Y. Rao *et al.*, “Hornet: Efficient high-order spatial interactions with recursive gated convolutions,” *Neural Information Processing Systems*, vol. 35, pp. 10353-10366, 2022. [Online]. Available: <https://arxiv.org/abs/2207.14284>

- [33] J. Liu, *et al.*, “MixMAE: Mixed and Masked Autoencoder for Efficient Pretraining of Hierarchical Vision Transformers,” 2022, *arXiv: 2205.13137*. [Online]. Available: <https://arxiv.org/abs/2205.13137>
- [34] I. Tolstikhin, *et al.*, “Mlp-mixer: An all-mlp architecture for vision,” *Neural information processing systems*, vol. 34, pp. 24261-24272, 2021. [Online]. Available: <https://arxiv.org/abs/2105.01601>
- [35] W. Yu, *et al.*, “MetaFormer is actually what you need for vision,” in *Proc. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, 2022, pp. 10809-10819. [Online]. Available: <https://arxiv.org/abs/2111.11418>
- [36] N. Ma, *et al.*, “Shufflenet v2: Practical guidelines for efficient cnn architecture design,” in *Proc. European conference on computer vision (ECCV)*, 2018, pp. 116-131. [Online]. Available: <https://arxiv.org/abs/1807.11164>
- [37] K. Han, *et al.*, “Transformer in transformer,” *Neural Information Processing Systems*, vol. 34, pp. 15908-15919, 2021. [Online]. Available: <https://arxiv.org/abs/2103.00112>
- [38] A. Trockman and J. Z. Kolter, “Patches are all you need?,” 2022, *arXiv: 2201.09792*. [Online]. Available: <https://arxiv.org/abs/2201.09792>
- [39] J. Wang, *et al.*, “Deep high-resolution representation learning for visual recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 10, pp. 3349-3364, Oct. 2020.
- [40] P. Vasu, *et al.*, “An improved one millisecond mobile backbone,” 2022, *arXiv: 2206.04040*. [Online]. Available: <https://arxiv.org/abs/2206.04040>
- [41] X. Ding, *et al.*, “Repmlp: Re-parameterizing convolutions into fully-connected layers for image recognition,” 2021, *arXiv: 2105.01883*.

- [Online]. Available: <https://arxiv.org/abs/2105.01883>
- [42] S.-H. Gao, M.-M. Cheng, K. Zhao, X.-Y. Zhang, M.-H. Yang, and P. Torr, “Res2Net: A New Multi-Scale Backbone Architecture,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 2, pp. 652-662, 1 Feb. 2021.
- [43] K. He, *et al.*, “Deep residual learning for image recognition,” 2015, *arXiv: 1512.03385*. [Online]. Available: <https://arxiv.org/abs/1512.03385>
- [44] K. Mangalam, *et al.*, “Reversible Vision Transformers,” in *Proc 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, 2022, pp. 10820-10830.
- [45] J. Hu, L. Shen, and G. Sun, “Squeeze-and-Excitation Networks,” *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, pp. 7132-7141.
- [46] K. Han, *et al.*, “Vision GNN: An image is worth graph of nodes,” 2022, *arXiv: 2206.00272*. [Online]. Available: <https://arxiv.org/abs/2206.00272>
- [47] S. Zagoruyko and N. Komodakis, “Wide residual networks,” 2016, *arXiv: 1605.07146*. [Online]. Available: <https://arxiv.org/abs/1605.07146>
- [48] A. El-Nouby, *et al.*, “XCiT: Cross-Covariance Image Transformers,” 2021, *arXiv: 2106.09681*. [Online]. Available: <https://arxiv.org/abs/2106.09681>
- [49] MMPreTrain 1.x open source tool code. [Online]. Available: <https://github.com/open-mmlab/mmpretrain/tree/main>
- [50] https://images.cv/download/scuba_diver/1300/CALL_FROM_SEARCH/%22scuba_diver%22
- [51] <https://conservancy.umn.edu/handle/11299/219383>