

水下影像強化對水下魚隻偵測準確度之影響

¹ 黃晉晁, ¹ 何培魁, ¹ 鄭錫齊, ¹ 張欽圳

¹ 國立臺灣海洋大學 資訊工程學系

E-mail: {10957038, 11157010, csc, cvml}@ntou.edu.tw

摘要

智能化箱網養殖可有效提升漁業競爭力，但由於日夜與氣候的交替，光學相機面臨進光亮與模糊等挑戰。為解決影像模糊等問題，使用生成對抗網路強化影像做為各項偵測前處理。結果顯示，使用強化影像相較於模糊影像，在體長估測將誤差降至 10.9%，體重部分降至 10.3%，驗證了使用 GAN 強化影像用於偵測的可行性。

關鍵字: 深度學習; 生成對抗網路; 水下影像; 影像強化; 智能化箱網養殖。

1. 緒論

本文旨在於對水下影像強化前後的各項水下魚隻偵測結果進行比較，進而探討水下影像強化對於水下魚隻偵測準確度之影響。

目前的光學影像於陸地上可完整且精準分析出圖片中的物件，但若要进行水下分析時除了防水的難題外更會面臨影像模糊而產生分析成效不佳的困境。為解決水下影像模糊問題，使用生成對抗網路 (GAN) 進行水下影像的強化，GAN 可強化水下影像，令夜間光學影像轉換成日間，不但可使以水下影像為基礎的魚隻資訊分析更準確，更能讓光學影像於夜間也能進行分析。並與聲納結合，解決水下光學影像模糊問題。本文針對生成之強化影像與模糊影像進行偵測結果比較，希望可以驗證強化影像用於水下偵測應用的可靠性。

本實驗在臺灣的魚隻養殖場域架設了多種傳感器，並由不同時間拍攝的影像中找出清晰及混濁的影像進行 GAN 的訓練。接著在魚隻偵測神經網路中，分別以 GAN 網路生成之強化影像及混濁或夜間的水下影像為輸入進行偵測結果比較。在各方面皆顯示由生成對抗網路生成的強化影像在偵測方面皆有較高的準確性。

論文整體貢獻如下：

1. 修改並驗證了一個生成對抗網路模型強化水下影像，令夜間光學影像轉換成日間。
2. 同時透過實際的場域實驗，對比有無使用生成強化影像的偵測結果，驗證使用生成強化水下影像作為辨識前處理之影響。

第二節首先著重目前已發表之水下影像強化生成的相關研究與應用。第三節討論使用強化生成影像作為輸入的水下魚隻偵測系統，並將這些系統使用模糊影像及強化生成影像作為輸入。第四節為將強化影像作為偵測系統輸入後的影響與比較。第五節為結論。

2. 水下影像強化

在海上若要以光學影像對魚隻進行大範圍監控非易事，考量到水質與光線導致水下能見度降低(圖 1)等等的因素更是讓光學影像在水下的應用難上加難。過去的研究已經探討了多種方法來強化水下影像。其中一種主要方法是使用生成對抗網路 (GAN)。



圖 1: 不同水質與天氣下的光學影像(清澈(左); 混濁(中); 夜間(右))

2.1 透過生成對抗網路強化影像

生成對抗網路[1]屬於非監督式學習的神經網路，在影像處理、音訊處理、自然語言處理等領域中都得到了廣泛應用，他的學習方法為將網路分為兩部分，包含生成器與判別器，生成器負責產生圖片而判別器負責辨識圖片是否為生成器所生成亦或是原始圖片，兩者互相反覆訓練直到生成器的圖片無法被判別器所辨識為止，通過反覆訓練，GAN 能夠生成具有所需特性的影像。另外，一些研究則提出了使用 GAN 來增強水下影像的方法，以解決顏色偏差、曝光不足和模糊等問題。這些方法通常包括配合物理模型來進行影像顏色強化，並使用雙判別器來提高影像的真實性和視覺美感。這些相關方法與強化水下影像的需求具有高度相關性，因此我們以這些內容作為啟發，將 GAN 用為水下影像的前處理，用以解決在實際場遇上會遇到的影像模糊問題。而 GAN 又包含許多變體，比如 InfoGAN[2]、CGAN[3]、DCGAN[4]……等，皆為對原始的生成對抗網路進行優化。而此網路的目的為依據訓練樣

本生成出使用者所需要的結果，可藉由數據集的不同與各式模型架構建立出各種各樣的影像以應用於不同領域上。

在日本東京大學有人以 cGAN 為基礎設計提出一種將聲納與夜間光學影像作為輸入，並輸出日間清晰的光學影像的 GAN 演算法[5]，為了能更好的在夜間對魚隻進行監控，他們在水族館的大型水族箱進行實驗，將聲納與相機裝設在不同角度以便對齊兩者間的視圖。實驗中將影像加入噪聲與處理當作夜晚與水質混濁影像，再配合聲納輸入他們所設計的演算法模型進行實驗，期望重建出日間沙丁魚群的影像，最後也以色彩分布與影像的相似度證明該模型生成日間水下影像效果優異。

[6]中為了優化海事作業的效率以及促進海事作業與機器學習間的相關性，該論文提出一種 GAN 的水下影像增強模型，以解決水下影像中的顏色偏差、曝光度不足及模糊等問題。[7]同樣希望藉由 GAN 網路來強化水下影像來分析河川下游的生態問題，他們利用 GAN 配合物理模型來進行水下影像顏色強化並透過雙判別器來提高影像的真實性與視覺美感。

2.2 透過 CycleGAN 進行水下影像強化

鑒於 GAN 在水下影像強化的優秀表現以及其功能的多樣性，我們選擇透過 GAN 作為水下魚隻偵測的前處理，希望透過 GAN 強化影像，增強其準確性。其中採用 CycleGAN[9]模型進行訓練，CycleGAN 主要用於將兩種不同風格的模型進行轉換，該模型架構由兩組 GAN 組成，包含生成器(G,F)與判別器(D_x, D_y)，如圖 2 所示。當模糊影像(X)輸入時經過生成器(G)會生成影像(\hat{Y})，此時判別器 D_y 會結合 \hat{Y} 與真實清晰影像(Y)進行判別；而另一組 GAN 剛好相反由 Y 輸入生成器 F 時會生成出影像 \hat{X} ， D_x 便會將 \hat{X} 與真實模糊影像 X 結合並判別。

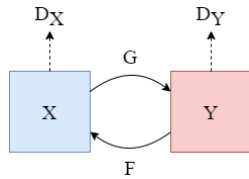


圖 2: CycleGAN 架構圖[9]

並且作者認為將 X 輸入至 G 時應該也要可以藉由 F 反向再映射回 X，反之亦然，因此提出一種循環一致性損失函數(cycle-consistency loss)，希望將原始影像映射至生成影像時也能再由另一個生成器映射回原始影像，如圖 3 所示。

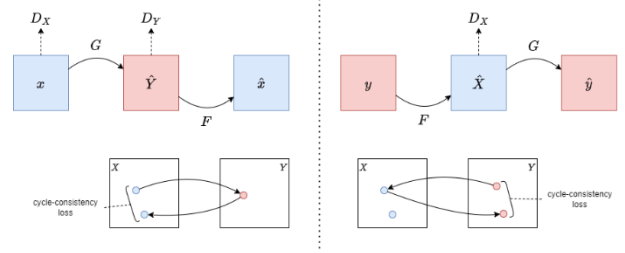


圖 3: forward cycle-consistency 與 backward cycle-consistency[9]

而在 CycleGAN 中相關的 loss 定義如下方式(1)~(4)。(1)與(2)表示兩組 GAN 的 loss 函數，包含生成器與判別器的 loss，其中 $P_{data(x)}$ 與 $P_{data(y)}$ 表示輸入模糊影像 X 與清晰影像 Y 的所有集合。式(3)中表示循環一致性 loss 函數，該函數表示 x 經由 G 進行映射再由 F 映射回 x 即 $F(G(x))$ 與實際 x 的差異再加上 y 經由 F 進行映射再由 G 映射回 y 即 $G(F(y))$ 與實際 y 的差異。整體的 CycleGAN loss 即為式(1)至式(3)結合為式(4)，其中 λ 為權重因子預設為 10。

$$L_{GAN}(G, D_Y, X, Y) = E_{y \sim P_{data}(y)} [\log D_Y(y)] + E_{x \sim P_{data}(x)} [\log(1 - D_Y(G(x)))] \quad (1)$$

$$L_{GAN}(F, D_X, Y, X) = E_{x \sim P_{data}(x)} [\log D_X(x)] + E_{y \sim P_{data}(y)} [\log(1 - D_X(F(y)))] \quad (2)$$

$$L_{cyc}(G, F) = E_{x \sim P_{data}(x)} [\|F(G(x)) - x\|_1] + E_{y \sim P_{data}(y)} [\|G(F(y)) - y\|_1] \quad (3)$$

$$L(G, F, D_X, D_Y) = L_{GAN}(G, D_Y, X, Y) + L_{GAN}(F, D_X, Y, X) + \lambda L_{cyc}(G, F) \quad (4)$$

我們將於 4.2 進一步討論 CycleGAN 的訓練及結果等其他內容。

3. 魚隻偵測

進行影像強化的最終目的，是為了提供養殖相關偵測較好的影像內容，取得並記錄正確的數據以進行精準養殖。分法中以水下影像與聲納以 KNN 回歸方法進行魚隻體長、體重、數量等分析，並將未強化前的影像與強化影像進行分析結果的精準度比較。在此討論了以下的偵測項目。

3.1 立體影像 3D 分析

立體影像 3D 分析流程分為四大部分，包含實例分割、立體物件偵測、視差微調與 3D 物件重建(圖 4)，首先以經過相機校正後的立體影像作為實驗影像，以 MaskRCNN 模型為每幀影像中的魚隻進行實例分割將每隻魚視為獨立個體。接著將偵測後的左右影像魚隻物件兩兩組成對應關係，利用物件中心

點座標可取得魚隻視差，但使用單一視差無法生成 3D 魚隻點雲，因此利用視差微調模型解決問題。該模型輸入左右影像對齊，接著以視頻插值 CNN (VICNN) [8] 合成中間影像，並計算像素殘差視差提供每個像素點獨立的視差值。最後將這些點蒐集起來集成 3D 點雲，以 PCA 方法取得最長軸即可獲得魚隻體長。

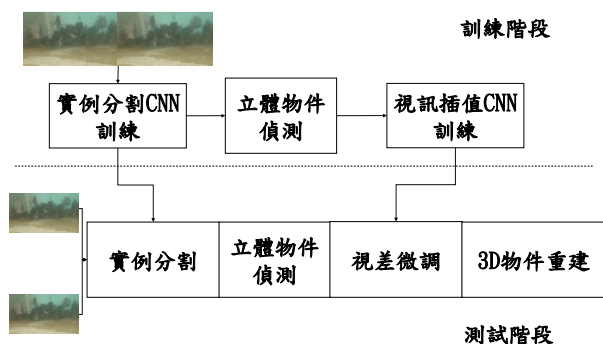


圖 4: 立體影像 3D 分析流程圖

3.2 立體影像體長、體重偵測

魚隻體長的量測以立體影像進行分析後再以 KNN 回歸。首先進行相機校正扭正影像計算魚隻 3D 座標，並使用 Mask-RCNN 進行魚隻標記與訓練。為確保左右影像可對應至同一隻魚，在每隻魚的 bounding box 加上 X 與 Y 軸的座標限制並以左右影像的魚隻像素長度作為判斷依據。若有正確對應則進行魚隻體長計算，使用 PCA 主成分分析法取得 3D 點雲的最長軸，並去除極端值計算魚隻體長。由於魚隻隨時在游動，體態不同會造成量測誤差，因此對魚隻做追蹤取得一秒內的體長做平均作為該隻魚體長(圖 5)。

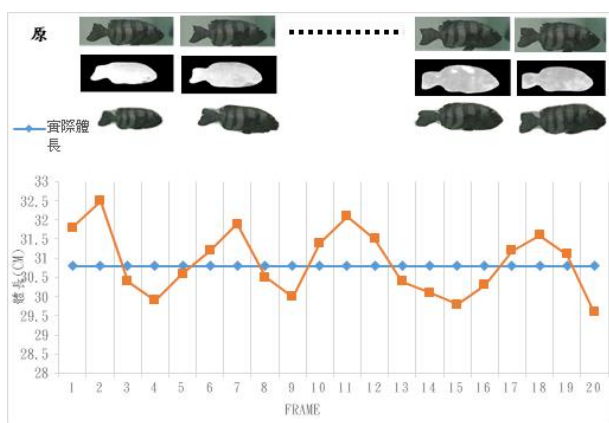


圖 5: 魚隻追蹤一秒內體長

在取得預測體長後以 KNN 配合實際量測體長進行回歸，分別以 K=3、K=5 與 K=7 時不同 K 值進行 KNN 回歸並計算其誤差。體重為透過體長的 KNN 回歸找到誤差最小 K 值的體長進行體重預測，同樣利

用 K=3、K=5 與 K=7 時不同 K 值建立體長與體重的回歸模型並以最小誤差 K 值作為預測結果。

3.3 聲納與影像結合魚隻數量估測

在立體影像上因水質與可視範圍小等問題無法取得全場域的魚隻數量，因此利用聲納擁有廣大偵測範圍的特性，配合兩者影像提高估測準確度。流程如下：

1. 利用物件偵測模型取得立體影像魚隻數量
2. 3D 融合演算法偵測聲納共同區域
3. 計算聲納上重疊區域魚隻數
4. 計算聲納與立體影像魚隻數比值
5. 推估至全場域取得場域魚隻數

首先利用 YOLOv4 取得立體影像上的魚隻數量，再來利用相機校正計算魚之 3D 座標，配合 ICP 變換矩陣對應至聲納後計算 convex hull 找到重疊區域，計算聲納影像上的重疊區域魚隻數量取得聲納與立體影像魚隻數比值，最後以該比值推估至全場域(圖 8)。

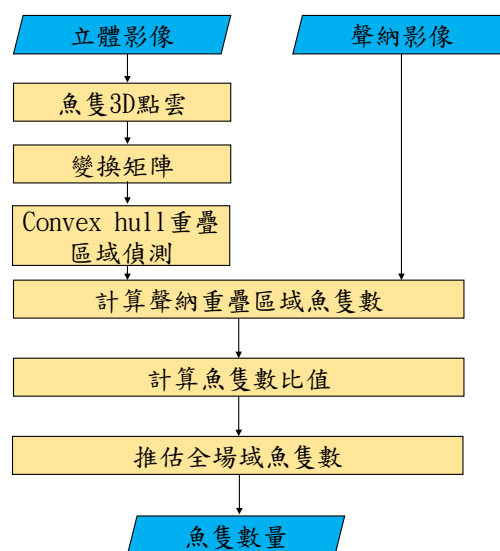


圖 8: 數量估計流程圖

4. 實驗結果

4.1 資料蒐集

實驗主要使用了三個場域做為實驗場地(圖 9)，分別為: A 海洋大學水生中心，B 屏東枋寮室內養殖池，C 屏東恆春海外養殖箱網。其中 A 場域魚種為斑石鯛，魚隻數量約 100 隻；B 場域魚種為紅九棘鱸，魚隻數量約 1100 隻；C 場域魚種為金鯧，魚隻數量約 17000 隻(表 2)。我們在這些場域中佈設了水下 RGB 立體攝影機、聲納及水質感測等系統，用於蒐集水下影像及各種相關資訊。

表 1: 實驗場域資料

場域	魚種	魚隻數量	影像數量
A 場域	斑石鯛	100	1800
B 場域	紅九棘鱸	1100	1800
C 場域	金鯧	17000	1800



圖 9: 實驗場域(左至右分別為 A 至 C 場域)

由於在 CycleGAN 中需要將 Domain A 的模糊影像生成為 Domain B 的清晰影像。將原始 RGB 立體攝影機取得的日間影像經過篩選後，取得各場域 1800 張作為 Domain B 的清晰影像，其中 1440 張做為訓練，360 張作為驗證。接著將 Domain B 的清晰影像降低亮度與對比度並加入高斯噪聲作為 Domain A 的模糊影像，同樣分為 1440 張訓練與 360 張驗證 (圖 10)。

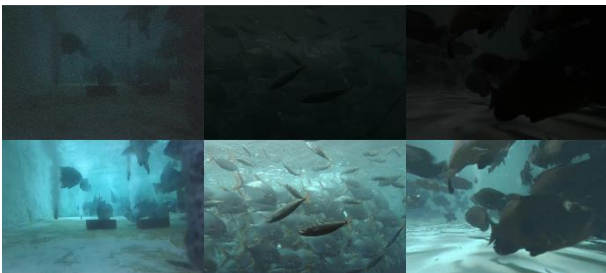


圖 10: 資料集內容(上為模糊影像，下為清晰影像。左右分別為 A 至 C 場域)

4.2 神經網路訓練

由於 CycleGAN 模型龐大，若輸入原始影像大小 1920*1080 會導致難以訓練，因此將影像裁剪成 256*256 大小的區塊進行訓練。CycleGAN 訓練包含兩組的生成器與判別器，分別為 Domain A to B 與 Domain B to A，loss 圖中 G_A 代表生成器由 Domain A to B 的 loss；G_B 代表生成器由 Domain B to A 的 loss；D_A 代表 G_A 生成的影像與真實 Domain B 影像間的 loss；D_B 代表 G_B 生成的影像與真實 Domain A 影像間的 loss。為了令 Domain A to B 還能再轉換為 A，cycle_A 代表 Domain A 遷移到 B 再遷移回 A 的 loss；cycle_B 代表 Domain B 遷移到 A 再遷移回 B 的 loss；idt_A 與 idt_B 則代表輸入輸出間顏色的一致性，為不讓色調改變過大所設置的 loss (圖 11)。影像強化結果如下 (圖 12)。

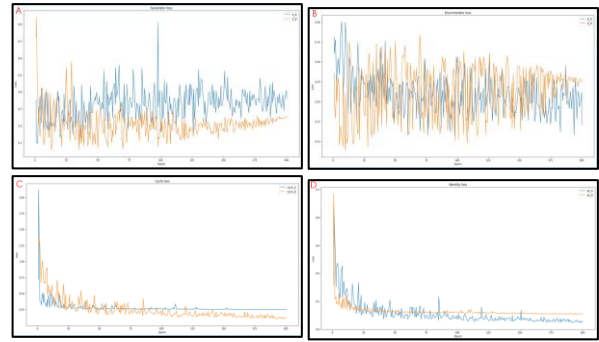


圖 11: CycleGAN loss 圖

(A)G_A 與 G_B (B)D_A 與 D_B
(C)cycle_A 與 cycle_B (D)idt_A 與 idt_B



圖 12: A 場域 CycleGAN 影像強化結果(上為模糊影像，下為清晰影像。左右分別為立體攝影機左右影像)

並且可以透過不同場域的影像強化結果 (圖 12、13、14) 觀察到 CycleGAN 影像強化在不同場域的強化表現都十分出色，這意味著這項技術可以有效地將模糊的水下影像轉換為清晰的影像，從而提高了在水下魚隻偵測方面的準確度和效果。並且在不同場域中的表現也體現出 CycleGAN 足以滿足不同場域的情境，在處理水下影像強化方面具有很大的潛力。



圖 13: B 場域 CycleGAN 影像強化結果(上為模糊影像，下為清晰影像。左右分別為立體攝影機左右影像)

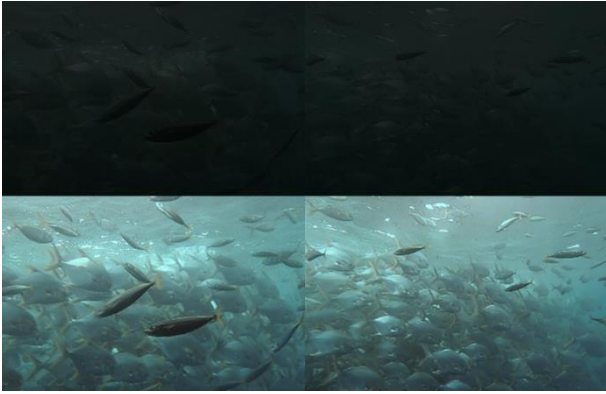


圖 14: C 場域 CycleGAN 影像強化結果(上為模糊影像，下為清晰影像。左右分別為立體攝影機左右影像)

4.3 分析比較

為證明影像強化後有助於提升水下魚隻分析結果，本節中使用水生中心、室內養殖池、恆春海上箱網三個場域，連續一分鐘共 1800 幀影像，利用第 3 節中提出的魚隻偵測項目，進行生成對抗網路的強化影像與原始影像的分析與比較。

4.3.1 立體影像 3D 分析

將影像經由內、外部參數與失真參數扭正，接著以 Mask RCNN 網路對影像進行魚隻實例分割找出左右影像中魚隻實例，再來找出左右影像中對應的魚隻物件進行匹配取得視差並做視差微調取得魚隻每個像素獨立視差，最後將 3D 座標進行 3D 建模分析魚隻資訊，結果如圖 12。



圖 12: 3D 點雲圖(模糊影像 3D 點雲(左)；清晰影像 3D 點雲(右))

由圖 12 中可以看到，右側由強化影像生成之 3D 分析結果點雲相較於左側由模糊影像生成之 3D 點雲較為清晰，可辨認出魚鰭、花色等樣貌。

4.3.2 實例切割

許多分析皆會使用實例切割作為初步定位及判定，透過 Mask RCNN 實例切割結果可以直觀的判斷結果好壞。並從圖 13 中看到各場域中模糊影像在切割時其精準度皆不如強化影像。

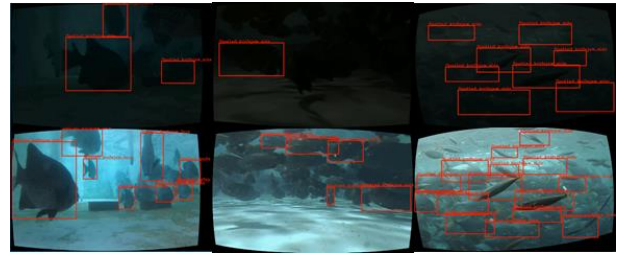


圖 13: MaskRCNN 實例分割(模糊影像(上)；強化影像(下))

4.3.3 體長估測

根據 3.2 使用立體影像分析出體長，以下為各場域分析結果。

表 2: 體長估測結果

影像種類	分析 cm	實際 cm	誤差 %
A 模糊影像	24.17	21.1	14.5
A 強化影像	23.41	21.1	10.9
B 模糊影像	16.39	15.5	5.7
B 強化影像	16.31	15.5	5.2
C 模糊影像	30.75	34.94	11.9
C 強化影像	31.8	34.94	8.9

透過上表的分析結果，清楚地發現強化影像的分析結果皆優於模糊影像，這顯示了經過影像強化後可以得到更精確的體長分析。在各場域中，強化影像的體長估測結果與模糊影像相比都顯示出更低的誤差率，分別為 10.9%、5.2% 和 8.9%，相比之下，模糊影像的誤差率分別為 14.5%、5.7% 和 11.9%。這說明了影像強化技術在提高魚隻體長分析準確性方面的有效性。

4.3.4 體重估測

在體重估計中同樣將模糊影像與清晰影像的分析結果進行比較，在此取 A 場域的分析結果作為代表參考，如表 3。

表 3: A 場域體重分析結果

影像種類	K	分析 g	實際 g	誤差 %
模糊影像	3	297.87	262.58	13.4
	5	299.18	262.58	13.9
	7	297.77	262.58	13.3
強化影像	3	290.54	262.58	10.6
	5	290.42	262.58	10.6
	7	289.64	262.58	10.3

由於體重的結果是由體長的分析結果藉由 KNN 方法回歸而來，透過選擇將同場域的不同 K 值回歸結果

列出比較。由此可以看到，在同個場域下，強化影像的體重分析結果較模糊影像接近真實資料。

4.3.5 魚隻數量分析

在水下魚隻數量分析方面，使用了人工神經網絡（ANN）建立了簡單的回歸模型。實驗中將強化影像取得每幀內的魚隻數量，每秒 30 幀所預測的魚隻加總作為一筆數據連續紀錄，5 分鐘共 300 筆。隨後透過 ANN 預測整體魚隻實際數量。各場域的預測魚隻數量與實際魚隻數如表 4。可以看到在各場域中，數量的預測誤差分別為 3.3%、2.2%、1.0%。

表 4: 魚隻數量估測結果

場域	分析隻	實際隻	誤差 %
A 場域	96.6	100	3.3
B 場域	1075	1100	2.2
C 場域	17184	17000	1.0

5. 結論

本論文透過對水生中心、室內養殖池和恆春海上箱網等場域拍攝的影像，實驗並建立了一個使用 CycleGAN 進行風格轉換的神經網路，該網路能夠成功地對水下影像進行強化。且針對強化影像進行了立體影像 3D 分析、實例切割、體長估測、體重估測以及魚隻數量分析的比較。

首先，在立體影像 3D 分析方面，通過對強化影像和模糊影像生成的 3D 點雲進行比較，能夠清楚觀察到強化影像生成的 3D 點雲更為清晰，能夠更準確地辨識出魚鰭和花色等魚隻特徵。其次，在實例切割方面，使用了 Mask RCNN 網路進行實例分割，結果顯示強化影像在實例切割方面表現優於模糊影像。在體長估測和體重估測方面，分別使用立體影像進行了魚隻體長和體重的估算，結果顯示強化影像的分析結果優於模糊影像，在實驗中誤差率分別降低了 3.6% 及 3.3%。最後，在魚隻數量分析方面使用了人工神經網路建立回歸模型，預測了各場域魚隻的實際數量，且與實際魚隻數量誤差值為 3.3% 以內。

綜合以上實驗結果可以得出以下結論：影像強化技術能夠顯著提升水下魚隻分析的準確性和效率。透過增強影像的清晰度和邊界，可以更準確地進行魚隻的立體影像分析、實例切割、體長估測、體重估測以及魚隻數量分析，提供更準確的水下魚隻分析結果供養殖戶參考。養殖戶透過更精準的養殖分析，便能夠更精準的掌控實際養殖狀況，達到精準養殖，避免不必要的環境資源浪費。

得益於影像處理技術的進步，許多相關應用都得以發展，而我們也將繼續探索和改進精準養殖的相關技術，以應對環境變遷等更多的挑戰，為相關領域的發展貢獻更多的力量。

致謝

在此由衷感謝所有參與研究和提供數據的合作者，他們的貢獻和努力為這項研究提供了寶貴的資源和支持，使得我們能夠進行深入的分析和探索。在此向所有人致以深深的謝意，是你們的幫助和支持使得這篇研討會論文得以成功完成。

國科會計畫編號: MOST111-2221-E-019-068-MY3

參考文獻

- [1] Creswell, Antonia, et al. "Generative adversarial networks: An overview." IEEE signal processing magazine 35.1 pp.53-65, 2018.
- [2] Chen, Xi, et al. "Infogan: Interpretable representation learning by information maximizing generative adversarial nets." Advances in neural information processing systems 29 (2016).
- [3] Deng, Jia, et al. "cGAN based facial expression recognition for human-robot interaction." IEEE Access 7 pp.9848-9859. (2019)
- [4] Fang, Wei, et al. "A Method for Improving CNN-Based Image Recognition Using DCGAN." Computers, Materials & Continua 57.1 (2018).
- [5] Terayama, Kei, et al. "Integration of sonar and optical camera images using deep neural network for fish monitoring." Aquacultural Engineering 86 (2019): 102000.
- [6] Zhang, Dehuan, et al. "Hierarchical attention aggregation with multi-resolution feature learning for GAN-based underwater image enhancement." Engineering Applications of Artificial Intelligence 125 (2023): 106743.
- [7] Cong, Runmin, et al. "PUGAN: Physical Model-Guided Underwater Image Enhancement Using GAN with Dual-Discriminators." IEEE Transactions on Image Processing (2023).
- [8] Yoon, Kuk-Jin, and In So Kweon. "Adaptive support-weight approach for correspondence search." IEEE transactions on pattern analysis and machine intelligence 28.4 (2006): 650-656.
- [9] Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." Proceedings of the IEEE international conference on computer vision. 2017.