

自適應河川異常檢測系統：結合運動學分析、影像處理與深度學習

徐英豈¹，彭康軒²，簡廷因³，陳伶志⁴

^{1, 2, 3} 元智大學資訊工程學系

{s1101444, s1103329}@mail.yzu.edu.tw

tinin@saturn.yzu.edu.tw

⁴ 中央研究院 資訊科學研究所

cclljj@iis.sinica.edu.tw

摘要

隨著臺灣都市化和工業化的加速，河流污染問題日益嚴重，對生態系統和公共健康構成重大威脅。本研究旨在應對現有河川監控系統的不足，開發一種自適應河川異常檢測算法，能在暴雨及平穩情況下有效識別河流異常事件。我們提出了一種結合運動學分析與異常檢測的創新方法，能夠在不依賴本地資料訓練的情況下，準確檢測河川漂流物及人員異常，從而更全面地保護環境和公共安全。

關鍵字：河流污染檢測、異常檢測、運動學分析、環境監控、自適應算法

1. 介紹

隨著臺灣的都市化和工業化進程的加速，都市河流的污染問題變得越來越嚴重。工業廢水、生活污水和農業徑流不斷將大量污染物帶入河流，迅速惡化水質，對生態系統和公共健康造成重大威脅。根據臺灣水利署的報告，自 2019 年（108 年）以來，臺灣的河川污染呈現逐年加劇的趨勢[1]。

在此背景下，對城市河流進行異常監測和分析成為環境管理的當務之急。尤其是在城市環境中，除了污染監測，檢測河流中潛在人員落水等異常事件也同樣重要。透過系統及時警報這些異常情況，可以迅速應對，減少對生命安全和環境的威脅。

然而，目前大多數河川監視器無法及時發現問題，通常是在大規模污染事件或異常發生後，才調取事發前的影像逐幀尋找，這不僅費時費力，且往往在問題發現時，環境已經遭受嚴重破壞。此外，現有的河流異常監測主要依賴人工巡查，例如：警

衛每隔幾個小時需實地檢查河流是否發生暴漲或出現異常，這種方式效率低下且浪費過多人力。

目前，大多數研究針對河川或海洋污染主要分為水質分析[2, 3]以及影像分析[4-6]。在水質分析上，主要檢查水中化學物質含量是否異常，但污染並不僅限於水質，肉眼可見之垃圾及漂流物皆是。而在影像分析方面，大多研究基於深度學習的方式，透過大量標註資料進行訓練，進而獲得極高的準確率，然而，由於所用數據均為本地收集，這些模型在面對未知環境條件或污染物特徵及分佈模式變化時，難以有效適應和應對。

基於先前的研究，我們認為在異常檢測中應引入運動學分析，以判斷異常物體與周圍環境的運動關係。當異常物體的移動方式與周圍環境相似時，將其判定為被動漂流物，如河水攜帶的物體；若運動方式不一致，則考慮其為主動物體，如人或鳥。雖然這可能會將部分異常錯誤分類，但已足以正確地歸類大部分異常情況。

本研究的貢獻包括以下幾點：

1. 開發了一種自適應異常檢測算法，無論是在暴雨還是正常情況下，都能有效偵測河川異常。
2. 提出了一種泛化且創新的方法，結合運動學分析與異常檢測，在不依賴本地資料訓練的條件下，能有效偵測河川漂流物。
3. 結合污染檢測與人員檢測，能夠及時發現河流異常事件，更全面地保護環境和公共安全。

2. 相關研究

在河川污染上，已有多個研究基於水質檢測

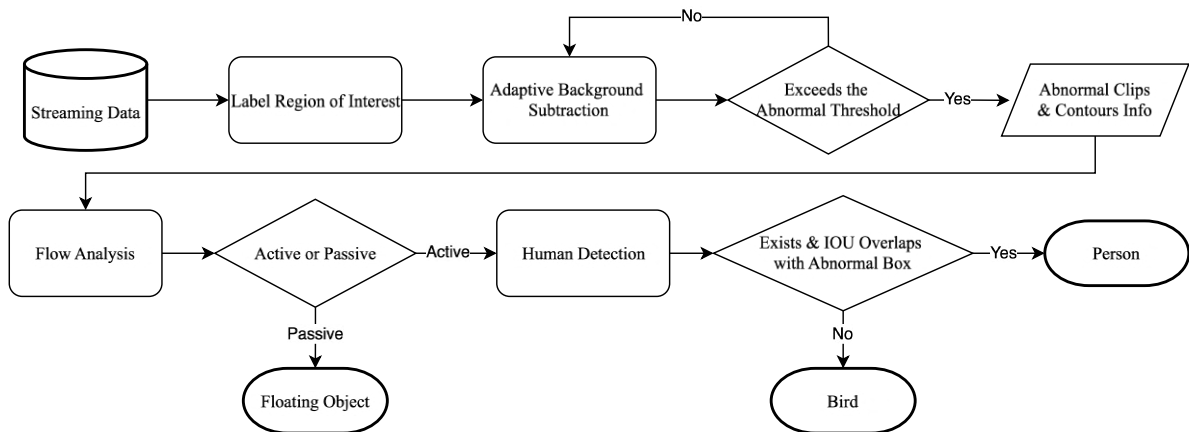


圖 1. 流程圖



圖 2. 前處理 (a)原圖(b)背景減除(c)模糊及開運算

來判別污染。例如，L. M. Guan[2]等人通過放置多個傳感器來監測水中的濁度及各種化學物質，並使用微控制器進行數據收集和處理，以實現即時的水質監測。J. Jiping[3]等人則分析水溫、含氧量、濁度、硝酸鹽、電導率...等等化學性質，並將其歸類為低頻（季節性）週期性波動以及高頻週期性波動來分析各種水質異常。

隨著機器學習的快速發展，許多分類以及分割模型已應用在污染檢測上，其中 M. J. H. Al-Battbooti [4]等人利用無人機拍攝的影像數據，透過 HSV 特徵進行油污特徵提取，並比較 SVM、KNN 以及 Random Forest 在油污、廢水污染、河流碎片上的分類結果，雖在本地資料集上取得不錯準確度，但他們也提到在不同環境中效果的不確定性，並以此為未來研究方向。在深度學習上，X. Yang[5]等人通過修改 YOLOv5 並添加注意力機制，證明其方法在河流漂浮垃圾物體檢測的公開資料集上，所有指標均優於原始 YOLOv5 和當時最先進的其他三個模型。此外，也有研究結合影像處理以及深度學習來追蹤污染異常，像是 C.-H. Wu[6]等人針對海洋油污污染使用全卷機神經網路分割污染區域並利用 SVM 分類有無油污。隨後，透過光流法[7]追蹤污染，幫助相關人員了解污染擴散的方向。

在過去的研究中，基於影像分析的方法主要依

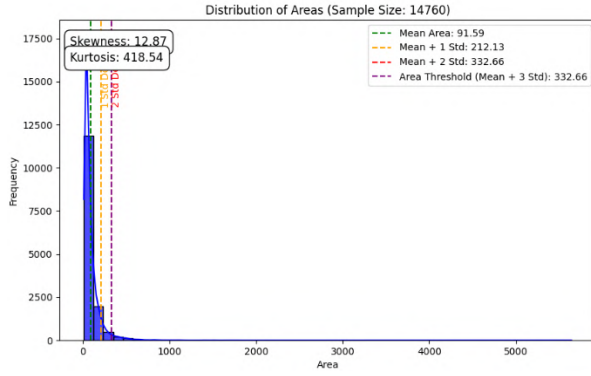
賴於物體的單幀靜態特徵進行分析，未充分考慮物體與環境之間的運動學關係。導致無法泛化至其他河川，因此，我們提出了一種結合運動學分析的新算法，旨在根據物體與周圍環境的運動相似性來提高異常檢測的準確性。

3. 方法

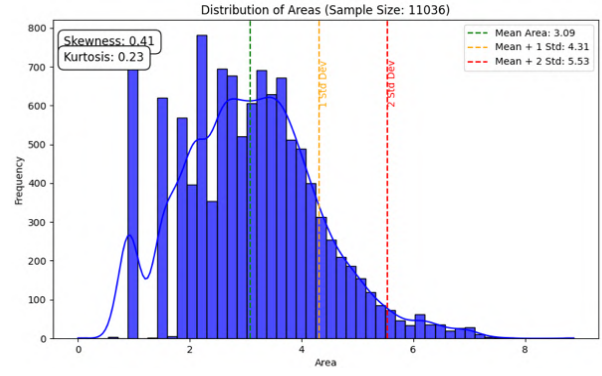
本研究設計了自適應背景減除法(Adaptive Background Subtraction)，該方法根據前景的混亂程度動態選擇合適的參數，以有效提取影像中的前景物體。並透過異常分數計算機制來識別真正的異常。隨後，分析前景異常物件與環境的運動模式，利用 3D 時空訊息和影像特徵來探討異常發生的原因，研究架構如圖 1。

3.1. 自適應背景減除法

本研究使用的影像為中研院所設置的四分溪河流監視器，解析度為 986*741，幀率為 30 fps，首先需標記出河道範圍，以降低干擾。我們採用 MOG2 (Mixture of Gaussians 2) [8] 來當作背景減除的核心方法，並通過高斯模糊和開運算處理前景遮罩(圖 2)，以去除噪點並增強物體的連續性，針對不同環境下前景的混亂程度，我們動態調整

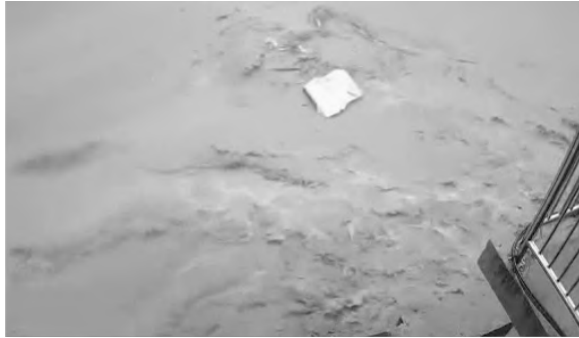


(a)原始面積分佈圖



(b)刪除 Q1 並對數轉換

圖 3. 對數轉換結果



(a)原圖



(b)應用 CLAHE 以及銳利化

圖 4. 影像增強

MOG2 算法中的關鍵參數方差閾值(`varThreshold`)，該參數基於馬哈拉諾比斯距離的平方值，用於判斷像素是否屬於背景模型。同時，我們設定了面積閾值，當異常面積大小超過特定值時即進行框選。

為了實現動態調整，我們建立了一個容量為 2,500 幀的佇列，用於存放每幀中前景物體的面積數據。在計算當前方差閾值和面積閾值時，我們首先考慮這些數據，然後刪除第一四分位數 (`Q1`)，因為這些通常來自河流的流動部分，面積較小因此導致面積分布正偏態過於嚴重。接著，對剩餘數據進行對數轉換，使其更接近常態分佈。最後，將轉換後的數據回推至原始尺度，並設定均值加上 3 個標準差作為面積閾值(圖 3)。方差閾值則根據面積閾值進行線性插值確定。

3.2. 異常加扣分機制

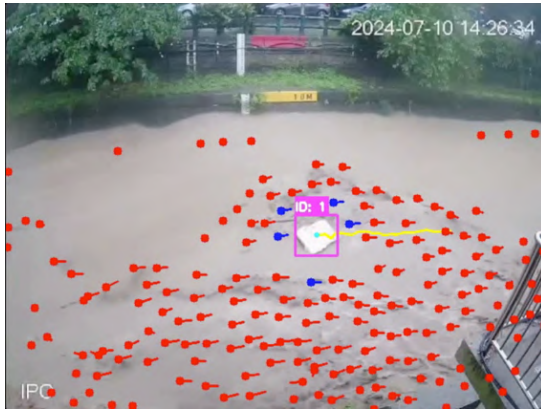
本研究設定當最大異常偵測數量超過 40 時進行後續分析。為避免異常計數器過度累積並確保在偵測後能迅速衰減，同時避免對非連續的異常事件生成多個影片，設計了一個動態的加扣分機制。當

每幀發生異常時，計數器 (`counter`) 會加 1，因此僅分析持續超過 1 秒的異常。若無異常發生，則進行扣分，扣分方式如公式 1，`step` 為衰減步長，`counter` 為當前檢測到的異常總數，`init` 為容忍的異常誤差數量。

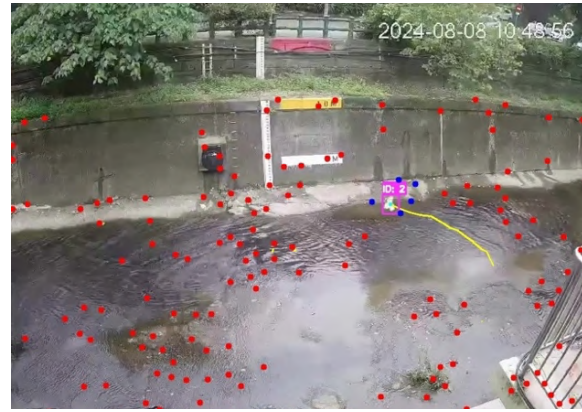
$$step = \begin{cases} 1, & \text{if } counter < init, \\ 2^{\lfloor \frac{counter - init}{5} \rfloor}, & \text{if } counter \geq init, \end{cases} \quad (1)$$

3.3. 運動模式分析

在測試河流異常物體時，我們發現分類模型難以有效區分鳥類和垃圾。我們嘗試了多個 YOLO 系列的物件偵測模型，但由於鳥類在影片中面積較小且模糊，導致偵測信心度不高，甚至錯誤地將許多河流垃圾和其他無關背景框選出來，且這些錯誤的信心度往往很高。因此，僅依靠像素訊息不足以提供有效資訊。我們進一步分析異常影片中物體的運動軌跡，以判斷其與周圍環境的關係是主動還是被動的。首先，我們對所有幀應用了直方圖均衡化 (`CLAHE`) 和圖像銳化等圖像增強方法(圖 4)，以



(a) 漂流物追蹤範例



(b) 鳥類追蹤範例

圖 5. 光流法及 SORT 結果 (藍色為最近五點外部向量)

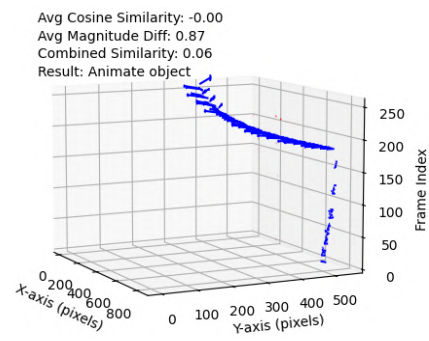
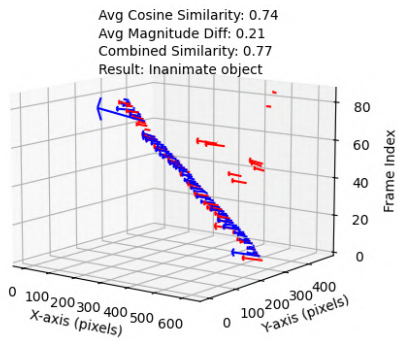


圖 6. 漂流物 3D 運動向量分析

(紅色向量為周圍外部相向，藍色為內部向量)

強化 Shi-Tomasi 角點檢測[9]的特徵點提取效果。並以每 5 幀更新一次角點，接著使用 L-K 金字塔光流法[7]來追蹤這 5 幀連續時間序列中的特徵點。

我們基於 SORT (Simple Online and Realtime Tracking) [10]算法引入目標追蹤技術，針對每個邊界框進行追蹤，以識別物體的軌跡。在圖 5 中，粉色框和標註的 ID 表示 SORT 的追蹤結果，黃色線段顯示了追蹤路徑。我們選取框中心外最近的 5 個光流追蹤向量，計算其平均向量作為周圍運動向量，框內向量則取自上一幀的向量差(圖 5)。接著，我們統計每幀的框內向量與框外平均向量，並比較它們的運動模式相似度(圖 6，為圖 5 對應的 3D 運動向量分析)。通過計算每幀的餘弦相似度和強度差異，判斷物體的運動模式是否與環境相似(向量相似度 ≥ 0.5)。若運動模式與環境非常接近，則定義為被動關係，反之則為主動關係。

3.4. 異常識別

若物體為被動關係，我們定義為此異常是跟著

水流流動的，由於異常有各種可能，我們將其定義為「其他異常」，而主動關係若非人則判定為鳥。

在人類識別上，本研究選擇使用異常影片進行物件偵測，而不是僅使用框選出的異常框進行分類。因為分類通常會喪失物體與周圍環境的關聯性，導致上下文信息的丟失。而使用物件偵測方法，可以同時考慮物體的特徵及其在場景中的位置和周圍環境。我們使用 YOLOv9[11]的預訓練模型來檢測人類，該模型僅用於人類的檢測(conf=0.5)。針對 YOLOv9 框選出來的人類會再與異常框計算 IoU (Intersection over Union)，以降低錯誤檢測的可能性(圖 7)。

影像損壞是監視器常見的問題，經測試該亂碼異常面積大部分都很大，且皆會被視為主動關係，因此定義超過當前 30 倍面積閾值且非人類的都視為其他異常，因為其大小不可能為鳥(圖 8)。



圖 7. 物件偵測結果

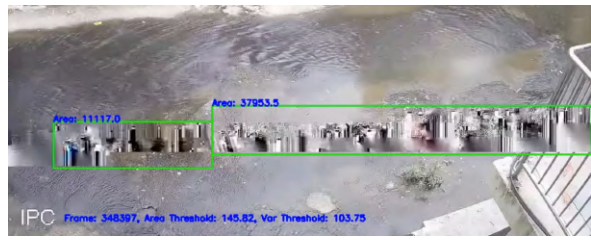


圖 8. 影像損壞範例

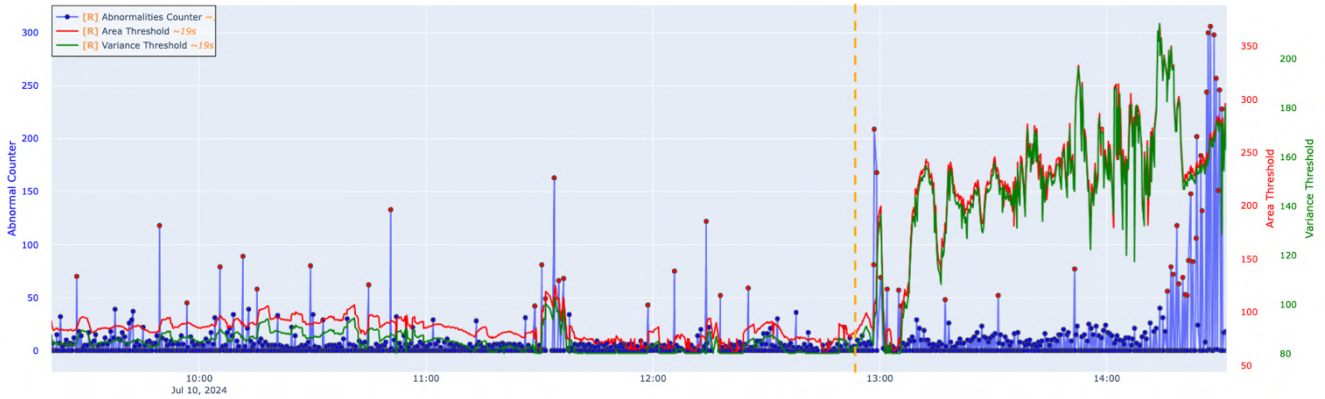


圖 9. 某段時間動態參數及異常分數結果

4. 結果

4.1. 自適應背景減除法結果

在圖 9 中展示了該動態算法的有效性，紅點表示被識別為異常的事件。藍色曲線代表異常檢測分數，紅色曲線表示面積閾值，綠色曲線表示方差閾值。在黃色虛線處，面積閾值和方差閾值迅速上升，這是由於開始下暴雨所引起的變化。

數據顯示，該天四分溪在早上並沒有太多異常，異常主要來自鳥類活動和影像傳輸的亂碼。在 13:00 時開始下暴雨，導致動態閾值快速上升，而動態異常分數機制在此表現出色，並不會因為大量瞬間雨滴異常而導致分數一直上升導致後續分析錯誤，反而是快速地增加面積、方差閾值。而因為下午的大暴雨導致 14:00 後大量異常被檢測出來，主要由暴雨沖刷下來的垃圾和樹木殘段引起。

4.2. 運動模式分析及識別結果

我們測試了 8 天白天到下午的資料，其中有兩天發生了暴雨，分類結果如圖 10 所示。在其他異常事件中，共有 158 個事件（以 90 個影像損壞 49 個暴雨垃圾為多數），其中 29 個被錯誤地預測為鳥

類。這 29 個錯誤中，有 21 個是由於影像損壞引起的，因為損壞範圍較小；3 個是魚類激起的水花；1 個是暴雨導致的垃圾，由於出現時間過短，導致運動分析出錯；2 個是光影變化過大；1 個是樹葉掉落；還有 1 個是小型垃圾引起的誤判。

在實驗過程中，我們發現此算法在環境快速變換時具有良好的排錯能力。以大雨為例，由於環境變化迅速，SORT 無法追蹤到有效框，因此這些情況會被分類為其他異常。此外，前景提取會捕捉到許多小型異常，但透過 SORT，可以有效地追蹤固定 ID 的物體不受干擾。本研究還嘗試了使用結合 CNN 資訊的 Deep SORT，但測試結果不如 SORT。

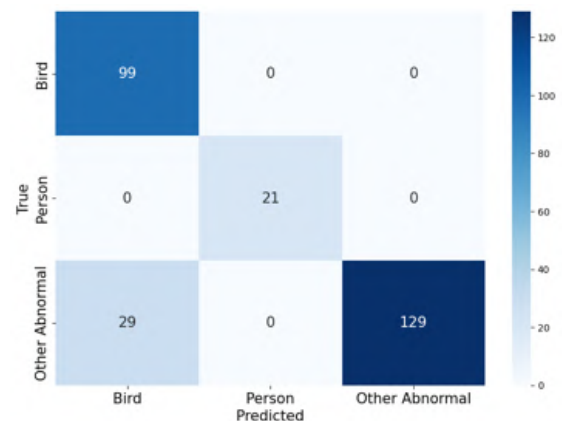


圖 10. 異常識別混淆矩陣

4.3. 討論

本研究目前僅適用於白天，因為夜間監視器上方的燈光吸引了過多昆蟲遮擋鏡頭，導致畫面過於混亂而無法識別。

為了降低影像損壞的誤判率，之後可以將 30 倍面積閾值降低至 10 倍來測試，因為觀測顯示鳥類的最大尺寸並未超過該參數的 10 倍。在鳥類檢測方面，該方法的優勢在於即使在極低解析度下也能進行追蹤，且不需耗費大量算力。然而，現實中的異常情況繁多，僅憑框選影像難以準確辨別鳥類，即使人類也難以在缺乏前後幀資訊的情況下作出正確判斷。我們曾嘗試對影像進行 4 倍超解析處理，期望在分類器上獲得更好的效果。然而，由於動態算法能捕捉到僅 10×10 像素的鳥類異常，這樣的小區域在超解析後的分類效果反而不如原圖。因此，我們放棄了像素級處理，轉而利用異常影片中的 3D 資訊進行分析，這導致了在水流平緩（看不出水流）時容易誤判為鳥類，為了減少誤判，未來可以考慮使用 ConvLSTM 技術，通過對整段影片進行鳥類追蹤來提高準確性。

5. 結論

本研究結果證明，所提出的自適應 MOG2 方法能在不同環境下有效提取真實異常，並成功結合運動學分析和 SORT 技術，實現對漂流物的檢測與追蹤，並且能精確地偵測人類。

目前，我們將異常分為三類（人、鳥、其他異常）。未來，我們可以通過結合其他特徵數據進一步細分更多類別，例如結合水流高度資訊，將暴雨引起的垃圾單獨分離出來。我們將繼續測試更多資料，並驗證此方法在不同環境中的監視器上是否同樣有效，以及是否能在邊緣運算裝置上運行。此外，我們計劃建立一個網站，提供每日的異常分析結果，並自動生成異常報告供相關人員查看和分析。目前，雛型網站已經開發。

6. 參考文獻

- [1] 行政院環境保護署. "105 年至 110 年重要河川污染指標概況." https://epaper.wra.gov.tw/Article_Detail.aspx?s=8239&n=30177 (accessed).
- [2] L. M. Guan, F. Yakub, A. R. M. Hanapiah, M. S. A. Rahim, A. Azizan, and M. Z. B. A. Rashid, "Water Quality Detection System using Real Time Vision Computing with Trash Collector," 2023.
- [3] J. Jiping, Z. Yi, P. Tianrui, W. Baoyu, C. Ritik, and T. Yu, "A comprehensive study on spectral analysis and anomaly detection of river water quality dynamics with high time resolution measurements," *Journal of Hydrology*, vol. 589, p. 125175, 2020, doi: <https://doi.org/10.1016/j.jhydrol.2020.125175>.
- [4] M. J. H. Al-Battbooti *et al.*, "Designing and Developing an Advanced Drone-Based Pollution Surveillance System for River Waterways, Streams, and Canals Using Machine Learning Algorithms: Case Study in Shatt al-Arab, South East Iraq," *Applied Sciences*, vol. 14, no. 6, p. 2382, 2024.
- [5] X. Yang *et al.*, "Detection of river floating garbage based on improved YOLOv5," *Mathematics*, vol. 10, no. 22, p. 4366, 2022.
- [6] C.-H. Wu, J.-W. Hsieh, C.-Y. Wang, and C.-H. Ho, "Marine Pollution Detection based on Deep Learning and Optical Flow," 2020.
- [7] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," presented at the Proceedings of the 7th international joint conference on Artificial intelligence - Volume 2, Vancouver, BC, Canada, 1981.
- [8] P. Kaewtrakulpong and R. Bowden, "An Improved Adaptive Background Mixture Model for Realtime Tracking with Shadow Detection," *Proceedings of 2nd European Workshop on Advanced Video-Based Surveillance Systems; September 4, 2001; London, U.K.*, 05/23 2002, doi: 10.1007/978-1-4615-0913-4_11.
- [9] S. Jianbo and Tomasi, "Good features to track," in *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 21-23 June 1994 1994, pp. 593-600, doi: 10.1109/CVPR.1994.323794.
- [10] A. Bewley, Z. Ge, L. Ott, F. T. Ramos, and B. Upcroft, "Simple online and realtime tracking," *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 3464-3468, 2016.
- [11] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "Yolov9: Learning what you want to learn using programmable gradient information," *arXiv preprint arXiv:2402.13616*, 2024.