

# 建構無感測器條件下的都市水位推論模型：以生成式影像為訓練基礎之研究

## Towards Low-Cost Urban Flood Monitoring: A Generative Image-Based Approach to Water Level Recognition

逢甲大學水利工程與資源保育學系

預研究生  
蔡杰儒  
Jie-Ru Tsai

碩士班研究生  
王健明  
Jian-Ming Wang

助理教授  
陳柏蒼  
Boris P.-T.  
Chen

### 摘要

極端氣候與都市化加劇使得都市地區面臨日益嚴峻的淹水風險，然而傳統水位監測方法如感測器與人工巡查，因建置與維運成本高昂、覆蓋範圍有限，難以即時應對災害。本研究提出一套結合卷積神經網路(Convolutional Neural Network, CNN)與影像生成技術(Stable Diffusion)的水位辨識系統，透過 AI 模擬不同水位影像進行訓練，無需歷史淹水資料即可快速建模，有效補足現行監測系統的盲點。

系統分為固定視角與非固定視角兩類模型設計，前者利用閉路電視(closed-circuit television, CCTV)影像空景生成訓練影像，進行水位分類辨識；後者整合社群媒體與新聞等動態影像，建立五級水位分級模型以處理視覺變異。研究亦參考國際標準制定符合台灣環境的分級制度。初步結果顯示，本系統能有效辨識多種來源的淹水影像，具備應用於不同地點與災害場景的潛力。然而，模型在面對極端光影變化、拍攝角度嚴重偏差或缺乏明顯參照物的影像時，辨識準確度仍可能受到影響，未來尚需進一步強化資料多樣性與模型適應性，以提升整體實務應用效能。

關鍵詞：水位辨識、影像生成、卷積神經網路、都市淹水監測、Stable Diffusion

### Abstract

The intensifying impacts of extreme climate events and rapid urbanization have significantly increased the risk of flooding in metropolitan areas. Traditional water level monitoring approaches, such as sensor-based systems and manual inspections, are hindered by high implementation and maintenance costs, as well as limited spatial coverage, making them inadequate for timely disaster response. This study presents a novel water level recognition framework that integrates Convolutional Neural Networks (CNNs) with generative image modeling using Stable Diffusion. By simulating diverse flood scenarios through AI-generated imagery, the proposed system enables rapid model training without the need for historical flood data, effectively addressing the limitations of

existing monitoring infrastructures.

The system is designed under two model architectures: a fixed-view model that employs synthesized CCTV background scenes for categorical classification, and a variable-view model that incorporates dynamic visual content from social media and news footage to construct a five-level water classification scheme capable of handling visual variance. The classification system references international standards while adapting to Taiwan's urban hydrological environment. Preliminary results indicate that the proposed models can accurately classify flood levels from heterogeneous visual sources, demonstrating strong potential for deployment across different locations and disaster contexts. However, model performance may decline under extreme lighting conditions, significant camera angle distortions, or in the absence of clear reference objects. Future research will focus on enhancing training data diversity and improving model generalizability to strengthen the overall applicability and robustness of the system in real-world flood monitoring scenarios.

Keywords: Water Level Recognition, Image Generation, Convolutional Neural Networks, Urban Flood Monitoring, Stable Diffusion

## 一、前言

近年臺灣與全球多處城市受極端降雨衝擊，短時強降雨使排水系統瞬間超載，造成道路與低窪地區積淹水。以 2025 年度為例，丹娜絲颱風於 7 月 6 日登陸，期間帶來強風豪雨，多處地區發生淹水；其後 8 月的楊柳颱風接續影響臺灣，再度造成強降雨與風浪威脅。此類短延時、高強度且階段相繼的風雨型態，使都市積淹水的發生頻率與影響程度上升，並對供電、交通與排水等關鍵設施的運作造成更大壓力。為回應此風險，政府與地方推動與自然共生的韌性城市，強調在災害衝擊下維持基本功能並迅速恢復，而對決策者而言，即時掌握災區現況並在有限資源下做出有效因應與分配，是核心關切。

基於上述背景，本研究利用社群媒體既有淹水影像資料與閉路電視(closed-circuit television, CCTV)影像資料，採用卷積神經網路(Convolutional Neural Network, CNN)並結合生成式合成影像建立自動水位辨識模型，使模型能從影像中學習水位特徵，於不同影像來源情況下提供即時判讀，並結合相關水情資訊以促進城市防災基礎設施的智慧化。

## 二、研究方法

本研究擬建構一套無須依賴長期歷史水位影像且可取代淹水感測器之人工智慧水位辨識流程。既有方法多仰賴長期累積之監視畫面與同步水位紀錄下，以人工標記訓練分類或分級模型，此途徑在人力與時程上成本高，且易受災例稀疏、鏡頭飄移與資料品質不一等因素限制，難以支援災中快速建模與部署的需求。

為回應上述限制，本研究採取「先以生成式影像構築訓練集，後以真實影像檢核」之設計，於定點視角，以 CCTV 影像作為目標區域並使用 Stable Diffusion 產生多種水位，取得分布均衡之樣本；於多來源視角(分固定視角)，彙整社群與新聞影像並依參照物規則進行有無淹水進行標註，用作模式作用標的。基於文章篇幅限制，本文將主以固定視角模式建置進行說明，非固定視角模式建置因相對簡單，故僅於討論與結論進行陳述。

### 三、研究結果 3.1 影像樣例與用途說明

淹水水位高度間距設定，本研究之水位分級系統係整合各國家機構標準與本土實務需求所建立。透過分析經濟部水利署、美國國家氣象局之洪水警示標準，以及日本國土交通省之都市浸水基準，並考量臺灣地區之道路特性與防災實務，研擬出五級制水位分類框架。國際標準與本土化調整之對照如表 1 所示：

分級	水深範圍	適用情境	本土化調整依據與換算說明
層級一	<15cm	<ul style="list-style-type: none"> <li>•機車排氣管進水風險</li> <li>•人行道積水影響通行</li> <li>•市區低窪地區開始積水</li> </ul>	<ul style="list-style-type: none"> <li>• 依據臺灣機車排氣管平均高度(12.3cm±2.1cm)</li> <li>• 參考臺北市人行道最低路緣石高度(13cm)</li> </ul>
層級二	15~30cm	<ul style="list-style-type: none"> <li>• 機車可能熄火</li> <li>• 小型車底盤進水</li> <li>• 地下室開始進水</li> </ul>	<ul style="list-style-type: none"> <li>• 依據臺灣小型車最低離地高度調查(18.5cm±3.2cm)</li> <li>• 考量臺灣常見水溝蓋高度(25cm)</li> </ul>
層級三	30~50cm	<ul style="list-style-type: none"> <li>• 多數房車無法通行</li> <li>• 一樓商店進水</li> <li>• 地下停車場積水</li> <li>• 變電箱可能短路</li> </ul>	<ul style="list-style-type: none"> <li>• 依據臺灣建築技術規則一樓門檻最低高度(45cm)</li> <li>• 參考經濟部標準局小型車最大涉水深度(40cm)</li> </ul>
層級四	50~100cm	<ul style="list-style-type: none"> <li>• SUV 車輛進水風險</li> <li>• 一樓住宅家具泡水</li> <li>• 道路中斷</li> <li>• 變電所可能受損</li> </ul>	<ul style="list-style-type: none"> <li>• 依據臺灣災害史統計(87%建築進水案例≥55cm)</li> <li>• 考量 SUV 平均進氣口高度(78cm)</li> </ul>
層級五	>100cm	<ul style="list-style-type: none"> <li>• 成人行走困難</li> <li>• 車輛可能漂浮</li> <li>• 建築結構受損</li> <li>• 維生系統中斷</li> </ul>	<ul style="list-style-type: none"> <li>• 依據臺灣成人平均身高比例(100cm≈女性腰部高度)</li> </ul>

表 1 國際淹水分級及適用情境

圖 1 為真實世界 CCTV 之影像，呈現同一監視器於事件期間的兩種狀態：有淹水與無淹水。該組影像作為本研究之最終測試資料，未參與任何訓練與驗證流程；圖 2 為以 Stable Diffusion 於同一空景條件下生成之模擬淹水影像，涵蓋不同水位對應之 L1~L5 分級訓練樣本，本組影像作為判斷水位等級模型之訓練資料。生成過程同時對濁度、波紋、反射、陰影及模糊等因素進行隨機化，以擴大資料分布並縮小合成與真實之差異。生成樣本近百張，模型訓練以圖 2 例生成影像為主，並搭配少量、不與圖 1 重複之真實影像作為微調與校準樣本，完成訓練後，再以真實影像進行模式驗證，以檢視學習穩定度。



(a) 實際淹水影像

(b) 無淹水

圖 1 台 9 線 421K+500 即時影像



(1) 影像生成底稿(無淹水)

(2) L1 影像生成(0~15cm)

(3) L2 影像生成(15~30cm)



(4) L3 影像生成(30~50cm)

(5) L4 影像生成(50~100cm)

(6) L5 影像生成(100cm 以上)

圖 2 由 AI 生成模擬 LV1~LV5 淹水影像

### 3.2 訓練過程與學習穩定度

本研究以生成影像資料訓練模型，並以真實影像驗證；圖 3 與圖 4 為訓練過程之損失與準確率曲線，以檢視學習是否穩定且無明顯過擬合。

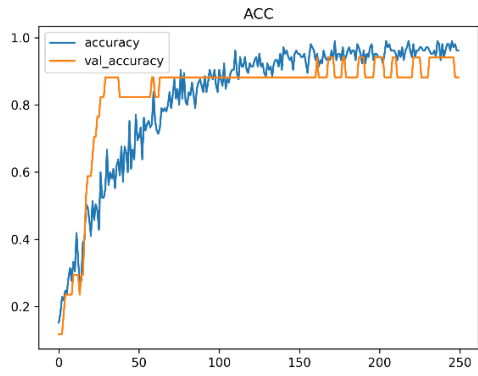


圖 3 學習準確率

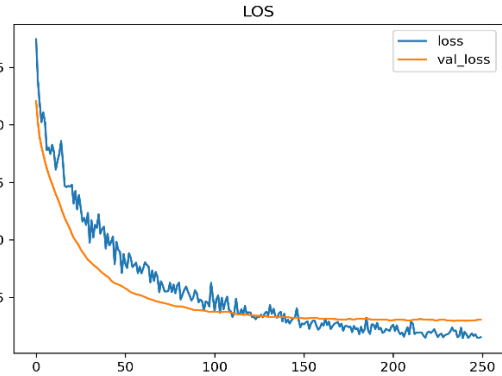


圖 4 學習損失

圖 3 顯示，準確率曲線在初期呈現快速上升，隨後逐步趨緩並於中後段趨於穩定；圖 4 顯示學習歷程穩定且無顯著過擬合。

### 3.3 真實與推論等級差異（混淆矩陣分析）

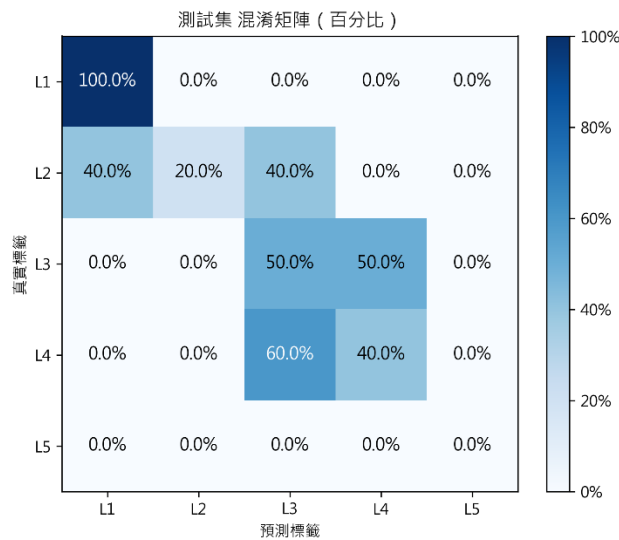


圖 5 固定視角之混淆矩陣

圖 5 呈現固定視角情境下的淹水等級辨識混淆矩陣。矩陣以斜對角為主，顯示模型整體能維持等級次序的一致性；錯誤多集中於相鄰等級，例如 L2 誤判成 L3、L3 誤判成 L4，而遠距離錯分完全沒有。另就資料組成而論：訓練資料已涵蓋等級五之合成樣本；然而本研究真實測試影像集因資料稀缺不含 L5 樣本，故混淆矩陣中 L5 列與欄為 0 屬於資料設計所致，並非模型無法識別。亦即，在測試所呈現的 L1~L4 級範圍內，模型未出現將其他等級誤判為 L5 之情形，顯示高等級門檻設定具嚴謹性；未出現 L5 亦符合測試集真值分佈。

### 3.4 結果討論



圖 6 L3 判斷成 L4



圖 7 L2 判斷成 L1



圖 8 接近清晨

圖 6 該畫面中可見大型車輛與水面相對接近。由於訓練階段以車輛部位（如輪胎高度、車門下緣）作為相對水位的參照，模型在含車輛的畫面會傾向強化此類線索；然而本例車輛並未行駛於較低窪的積水段，車體與路面之相對高度造成視覺上的「等級抬升」，導致 L3 被高估為 L4。

圖 7 本例中模型注意力主要落在靠近路側、未積水之行車動線，遠側的淺層積水區因對比不足、紋理不明顯而被忽略，致使判定傾向低估。此情境顯示當可用參照物（路緣、車輪等）主要出現在未積水區時，模型會將濕潤路面誤視為無積水。

圖 8 光照色偏導致的誤判在清晨場景中，道路表面出現明顯的藍色光源反射。首次測試時，模型將此視為水面鏡射與高反射度的證據，因而輸出 L4。當納入與該場景相似之光照條件真實樣本後，模型不再將色偏視為積水特徵，改能判為 L1，顯示其對照明變化的辨識較為穩健。

整體而言，訓練與測試表現相互貼合，顯示學習歷程穩定；其次，本研究固定視角的混淆矩陣圖 5 以斜對角為主要目標，錯誤主要落在相鄰等級，符合有序分類的行為特性，模式推論結果支持本研究方法在等級任務辨識上的可行性與穩健性；此外，就資料組成而言，訓練資料涵蓋 L5，而真實測試集不含 L5 真值，而測試過程亦未將其他等級誤判為 L5 且在本次測試範圍(L1-L4)內的等級辨識皆在 1 個級序內；而錯誤案例分析圖 6~8 則說明模型主要倚賴參照物位置與反射、紋理線索，在車輛不位於積水區或光照色偏時易出現高估或低估；當補入相似條件的樣本後，模型對照明與紋理變化之辨識更為穩健。

## 四、結論與建議

本研究提出一套以生成式影像為核心的水位等級辨識流程；於定點視角以空景條件生成多種水位與環境組合，縮小合成與真實世界之差異。整體流程可在近百張規模的合成樣本，搭配少量真實影像微調的情況下，快速得到可運轉的模型，顯示此方法具免水尺、低門檻之特性。

此外，針對非固定視角之應用，本研究亦以前述模型對社群、新聞等異質來源的淹水畫面進行等級辨識，但實際上，因拍攝距離、俯仰角、相機高度與焦距差異所造成的尺度與透視變化，使參照物在畫面中的相對位置與尺寸產生偏移，易導致高估或低估，相較於固定視角，誤差值會高於 1 個級序範圍，為降低此類誤差，需輸入更多相關畫面作為模型學習的樣本，盡可能達到誤差等級在正負 1 個級序。

綜上所述，生成式影像驅動的水位等級辨識可在短時程、低成本下於新道路建置，即使未設置水尺亦可運作，補強傳統監測佈建的盲點，並可與既有 CCTV 系統以及社群媒體畫面結合，作為災時即時判讀與通報之輔助。未來可沿著本研究框架，依作業需求擴充至連續水位推估與道路客製化校正，並建立跨機關的影像共享與去識別治理流程，以加速實務落地。