

基於遮罩引導之失真影像修復系統

Mask-Guided Distortion Image Restoration System

指導教授：韓欽銓

學生：李政勳 車延信 吳佳霖

國立聯合大學 資訊工程學系

苗栗市南勢里聯大2號

{cchan, U1124049, U1124030, U1124031}@o365.nuu.edu.tw

摘要

本專題建構一套基於遮罩引導之失真影像修復系統，以深度學習技術自動辨識雨滴、水紋與光暈等失真區域，並透過遮罩預測引導修復流程，提升影像重建精確度。系統採用 U-Net 結合注意力機制，搭配改良版 BRISQUE 指標進行品質評估，最終整合為具互動操作的修復功能之平台，兼具研究與實務應用價值。

關鍵詞：遮罩引導、影像修復、深度學習、注意力機制、即時修復。

Abstract

This project develops a mask-guided-distortion image restoration system based on deep learning techniques. The system automatically identifies distortion regions such as raindrops, water ripples, and glare, and utilizes mask prediction to guide the restoration process, thereby improving reconstruction accuracy. A U-Net architecture combined with an attention mechanism is adopted, with BRISQUE metrics employed for quality evaluation. Ultimately, the system is integrated into an interactive restoration platform, providing both research and practical application value.

Keywords: Mask-guided, Image Restoration, Deep Learning, Attention Mechanism, Real-time repair.

一、前言

隨著行車記錄器、無人載具、智慧監控與自動駕駛技術發展，影像品質已成為影響系統判斷的重要關鍵。然而實際環境中常出現雨滴、雨條紋與光暈等失真，導致畫面模糊、邊緣不清與辨識錯誤，且多種失真常同時疊加，使傳統濾波與一般修復方法難以有效處理。

為解決此問題，本專題建構一套基於遮罩預測之影像修復系統，運用深度學習自動標示受損區域，並以預測遮罩引導局部修復[1]，以提升重建精度與穩定性。

二、研究方法及進行步驟

(一) 訓練資料製作

資料來源包含 RainDS [2] 與 Flare7K++ [3] 兩公開資料集，涵蓋雨滴、雨條紋與光暈等失真。針對光暈部分，另以素材合成方式套疊至乾淨影像建立失真資料。資料前處理中，所有影像統一縮放為 640×360 並轉換為 PNG，以確保格式一致。之後根據乾淨與失真影像的亮度差異生成灰階遮罩(公式 1 與公式 2)，標示受損區域，作為後續模型的監督訊號。

$$f(I) = 0.299 \times R + 0.587 \times G + 0.114 \times B \quad (1)$$

$$M = f(I_r) - f(I_c) \quad (2)$$



圖 1：乾淨影像(左圖)、失真影像(中圖)與遮罩影像(右圖)

由於訓練資料數量有限，同時為增強模型泛化能力，對訓練影像進行水平與垂直翻轉等資料增強操作，以模擬多樣化場景並降低過擬合風險，如圖 2 所示。

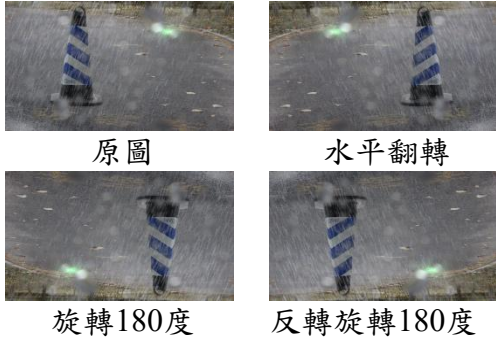


圖 2：資料增量示意圖

(二) 無參考影像品質指標驗證實驗

為驗證其在多樣失真條件下的適用性與穩定性，本專題以 TID2013 [4] 標準資料集為基準，該資料集含 25 張原始影像，經 24 種失真與五種強度處理，共計 3000 張測試影像，可用以評估指標對失真強度變化的敏感度與區辨能力。

比較三種主流指標：NIQE、PIQE 與 BRISQUE [5]，並依據五大失真類別（雜訊、模糊與銳化、壓縮與傳輸、幾何與結構變形、複合與其他）分析其合理變化趨勢。如圖 3 所示，BRISQUE 表現最穩定，合理趨勢比例達約 93%，能準確反映品質隨失真強度變化的單調關係；PIQE 約為 85%，於幾何與混合失真下略有波動；NIQE

約為 70%，對低對比與結構性失真敏感度不足。針對自然場景中的雨滴、雨條紋與光暈失真，分別屬於幾何變形、雜訊及模糊類。

綜合分析可知，BRISQUE 在這些自然型失真下具最佳穩定性，因此本專題最終採 BRISQUE 作為影像修復品質之主要評估指標。

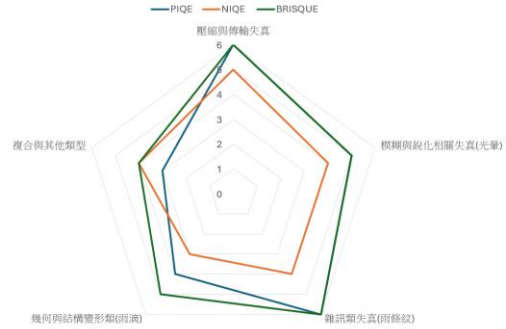


圖 3：雷達圖：各影像品質評估指標於五大失真類別之合理情況數量

(三) BRISQUE 特徵差異分析

BRISQUE 原始分數對雨滴遮蔽並不敏感，使得加入雨滴後的分數變化不明顯，但其 36 維特徵向量仍可能包含可辨識的統計差異（公式 3）。

為驗證此點，選取 RainDS 中 218 組原始與雨滴合成影像，計算特徵差並進行成對 t 檢定（公式 4）。結果顯示多數特徵在兩類影像間呈極顯著差異（ $p < 10^{-20}$ ），說明雨滴會強烈影響自然場景統計，僅少數特徵未達顯著性。特徵分布亦呈一致方向偏移，使原始與失真影像在特徵空間產生可分性。

設計改良後的雨滴量化模型，能特別適應雨滴失真並提升對遮蔽區域的敏感度。然而此調整僅對雨滴有效。最終使用 SVM 驗證，分類準確率約 70%，顯示 BRISQUE 特徵可作為雨滴相關分析的輔助資訊。

$$\Delta f_i = f_{i,\text{rain}} - f_{i,\text{original}} \quad (3)$$

$$H_0: \Delta f_i = 0 \quad (4)$$

(四) 遮罩預測模型 (如圖 4)

採 U-Net [6] 架構結合 CBAM [7] 注意力機制,使模型能同時關注通道與空間特徵,以精確定位受失真影響的區域。輸入 RGB 影像,輸出灰階遮罩圖。為生成訓練遮罩,將失真影像與對應乾淨影像轉為灰階後取差異以突顯失真區域。損失函數採綜合策略:以 Dice Loss、IoU Loss 與 Perceptual Loss 提升區域一致性,再輔以 Perceptual Loss 提升結構細節的辨識能力。

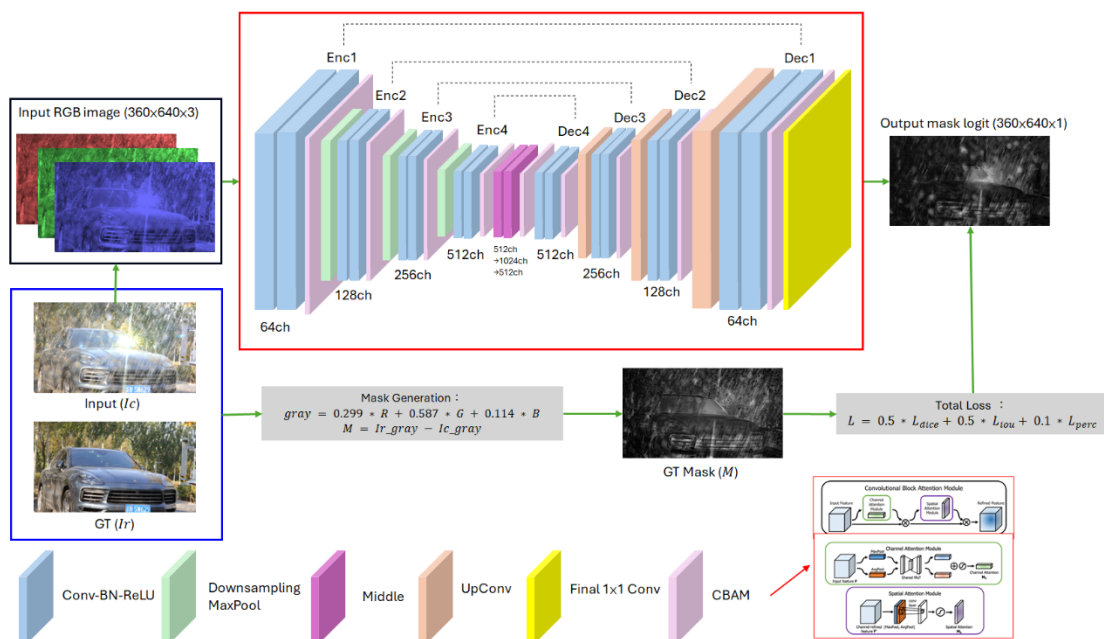


圖 4：遮罩預測模型示意圖

(五) 失真修復模型 (如圖 5)

輸入為遮罩與原始失真影像所組成的四通道影像。模型主體採用結合 RCBAM [8] 模組的 U-Net 生成器,強化特徵提取與區域重建能力。引入 PatchGAN 判別器 [9],以局部影像區塊為單位進行真實與生成影像的對抗判別,使模型能專注於局部細節修復與紋理還原。訓練過程整合 SSIM Loss、Perceptual Loss 與 GAN Loss。

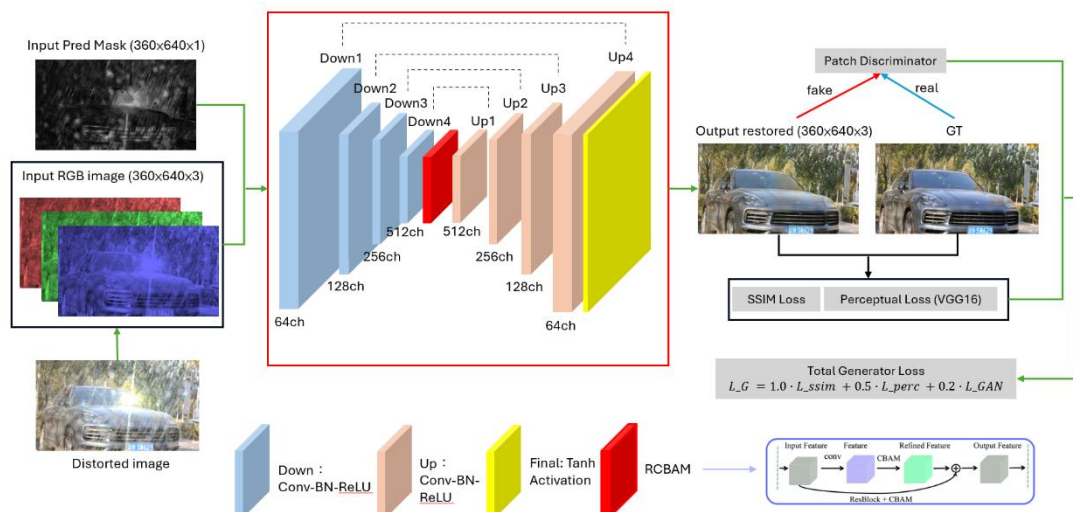


圖 5：修復失真模型示意圖

三、模型評估與分析

(一) 模型評估結果

由表 1 可見，在光暈、雨滴與雨條紋同時存在的複合失真環境下，使用人工真實標記的精準遮罩（GT Mask）引導之模型在合成與真實資料中皆明顯優於使用預測遮罩（Pred Mask）結果，顯示遮罩準確度對修復品質具有關鍵影響。

精確遮罩可有效引導生成器聚焦於失真區域，減少背景干擾並提升結構與紋理一致性；反之，當遮罩精度降低時，模型注意力易受誤導，導致細節與亮度還原不足。

如圖 6 所示，預測遮罩模型仍能重建主要物體輪廓與場景結構，整體影像具良好一致性與可辨識性。

綜合而言，所提出之方法於複合失真場景下展現穩定修復效能，未來若能進一步提升遮罩預測精度，預期將能更接近理論修復上限。

(二) 與 RainDS 文獻[10]評估比較

根據表 2 與表 3，在雨滴與雨條紋單一雨況下，本專題方法於 GT Mask 條件多數皆優於 CCN，顯示當遮罩能準確標示失真位置時，遮罩引導確實能有效提升修復品質。

然而在 Pred Mask 下，特別是雨滴和雨條紋同時出現的多重雨況中，本方法整體仍略低於 CCN，反映現階段遮罩模型尚難完整捕捉細碎雨滴、細長雨條紋及其交疊區域，造成局部漏標並影響生成器的還原判斷。

進一步分析可知，雨條紋的形狀明確，即使遮罩略有偏差仍具足夠引導效果；但在複雜雨滴和雨條紋同時出現場景中，失真透明度變化大且常覆蓋高頻細節，一旦遮罩資訊不足，生成器即難以區分失真與真實細節，導致 SSIM 與 PSNR 下降。整體結果指出，遮罩品質仍是系統效能的主要瓶頸，也是未來改進的核心方向。

表 1：光暈、雨滴與雨條紋同時存在之影像修復評估結果

遮罩類型	合成 (Synthesis) 資料		合成 (Real) 資料	
	SSIM	PSNR	SSIM	PSNR
GT Mask	0.9170	28.68	0.7831	24.46
Pred Mask	0.8635	26.77	0.6800	22.14

表 2：各方法於真實(real) RainDS 資料集上的定量比較

Methods	雨滴		雨條紋		雨滴+雨條紋	
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
CCN (文獻報告)	0.820	26.83	0.757	24.81	0.704	23.09
Ours (Pred Mask)	0.745	23.66	0.797	25.85	0.683	22.47
Ours (GT Mask)	0.778	24.71	0.847	27.43	0.766	24.18

表 3：各方法於合成(syn) RainDS 資料集上的定量比較

Methods	雨滴		雨條紋		雨滴+雨條紋	
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
CCN (文獻報告)	0.969	35.12	0.974	33.29	0.950	32.16
Ours (Pred Mask)	0.952	31.30	0.955	32.02	0.884	28.02
Ours (GT Mask)	0.971	32.48	0.976	33.25	0.932	29.75

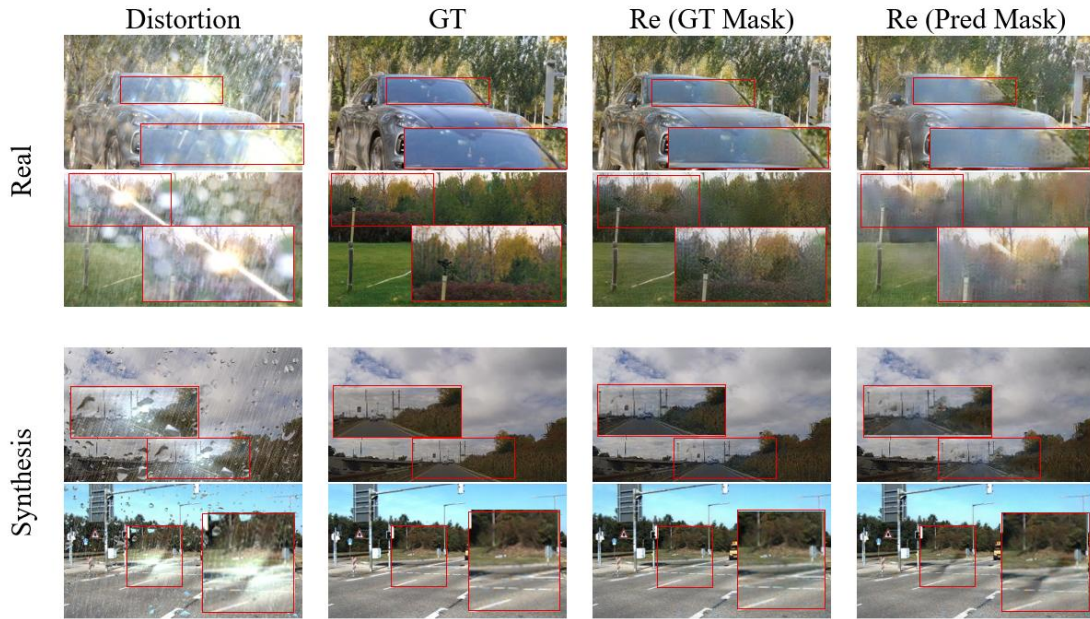


圖 6：光暈、雨滴與雨條紋同時存在時之影像修復結果示意圖

(三) BRISQUE 指標評估

如表 7 所示，本研究以 BRISQUE 所採用的 NSS（自然影像統計）特徵與 MSCN/AGGD 參數為基礎，重新設計一個能量化雨滴程度的無參考指標，結果顯示該指標在雨滴（RD）類型上最能凸顯影像中局部遮蔽的強度。

在原始 BRISQUE 評分中，雨滴影像的 Input 分數甚至低於 GT

（BRISQUE 分數越低代表越自然），顯示標準 BRISQUE 對雨滴造成的局部模糊與高亮遮蔽極度不敏感，常將受損影像誤判為更自然。

本研究提出的雨滴量化指標能清楚放大這些「偏離自然影像統計的局部失真」。其結果使得 Input 分數明顯高於 GT，成功呈現雨滴失真的自然度下降。值得注意的是，對於移除雨滴後的 Re(GT)，其得分偶爾會低於原 GT，但這並非意味修復影像比乾淨影像更自然，而是雨滴移除後，局部紋理的 NSS 分佈更趨一致，進而提高指標對修復品質差異的敏感度。

表 8 用於觀察不同失真類型在修復前後，Input (I)、GT (G) 與 Restore (R) 的評分排序是否符合合理品質關係 ($R < I$)。表中「張數」代表落在各排序類別的影像數量，反映指標是否能正確辨識修復品質。

在雨滴（RD）類型中，原始 BRISQUE 的排序相當混亂，許多影像落入 $I < G$ 或 $I < R$ 等明顯不合理的分類，顯示 BRISQUE 本身對局部高亮遮蔽極不敏感。

引入本研究的雨滴量化指標後，RD 類別的張數大量集中於合理的 $R < I < G$ ，錯誤排序幾乎完全消失，證明所設計的雨滴特徵能讓模型更有效識別遮蔽程度並準確反映修復效果。少數 $R < G$ 的情況，則多為移除雨滴後使局部自然統計更一致，並不代表修復超越 GT。

其他失真類型（RS、RDS、FRDS）也呈現改善，但變化幅度較 RD 小，顯示本指標主要針對雨滴失真的辨識能力。若未來能針對其餘失真類型分別設計對應的量化方式，將可進一步提升多失真場景的整體品質評估準確度。

表 7：各失真類型 BRISQUE 平均分數（分數越低越佳）

	BRISQUE				雨滴數量指標			
	GT	Input	Re (GT)	Re (Pred)	GT	Input	Re (GT)	Re (Pred)
RD	15.53	11.22	14.16	14.75	20.56	48.73	17.61	19.91
RS	15.53	22.98	15.31	14.51	20.56	13.39	17.61	19.91
RDS	15.53	16.79	16.43	16.64	20.56	36.47	17.11	19.29
FRDS	15.51	16.89	14.00	14.84	20.45	37.18	13.13	17.84

表 8：各失真類型修復時指標分數關係分類統計

RD：雨滴，RS：雨條紋，RDS：雨滴+雨條紋，FRDS：光暈+雨滴+雨條紋

R：Restore，I：Input，G：Gt

			理想情況 (R > I)			不理想情況 (I > R)			總共
			R<I<G	R<G<I	G<R<I	I<R<G	I<G<R	G<I<R	
BRISQUE	GT Mask	RD	21張	5張	3張	29張	37張	3張	98張
		RS	7張	33張	36張	8張	6張	8張	98張
		RDS	6張	13張	29張	11張	31張	8張	98張
		FRDS	15張	28張	18張	21張	13張	5張	100張
	Pred Mask	RD	12張	5張	2張	38張	37張	4張	98張
		RS	11張	32張	40張	7張	3張	5張	98張
		RDS	5張	6張	33張	15張	28張	11張	98張
		FRDS	16張	16張	29張	22張	11張	6張	100張
雨滴量化指標	GT Mask	RD	1張	88張	8張	1張	0張	0張	98張
		RS	9張	20張	8張	53張	4張	4張	98張
		RDS	3張	69張	14張	7張	3張	2張	98張
		FRDS	6張	84張	4張	5張	1張	0張	100張
	Pred Mask	RD	1張	73張	23張	1張	0張	0張	98張
		RS	5張	14張	12張	39張	22張	6張	98張
		RDS	7張	43張	41張	4張	2張	1張	98張
		FRDS	8張	54張	32張	3張	1張	2張	100張

四、系統實作與驗證

(一) 影像修復平台系統

本專題建置一套整合靜態、動態與即時影像修復的平台，使用者上傳影像或影片後即可快速獲得修復結果。系統以遮罩預測與深度學習模型自動辨識並重建多重失真區域，形成「失真辨識、區域修復、結果回饋」的一站式流程，如圖 7 所示。

後端採 Flask 負責接收資料與模型推論，前端以 React.js 提供影像／影片上傳、模式切換與結果展示等互動介面。系統亦支援影片逐幀修復，確保連續畫面的一致性。平台同時提供 WEB 與桌面 GUI 版本。如圖 8 和圖 9 所示，介面設計直觀，三大功能模組（靜態、動態、即時修復）清楚呈現並可快速啟動修復流程。

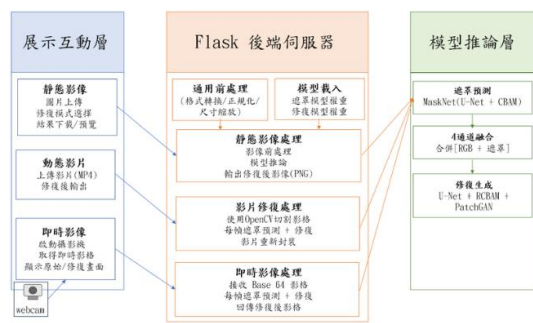


圖 7：系統整體系統架構示意圖



圖 8：系統 WEB 首頁介面



圖 9：系統 GUI 首頁介面

(二) 靜態影像修復功能

本功能面向一般使用者，提供上傳影像、檔案檢查、模式選擇、遮罩預測、區域性修補、評分顯示、前後對照與 PNG 下載等流程。圖 10 為 Web 介面，顯示輸入檢查畫面，若格式或尺寸不符，系統會即時跳出警示並阻擋後續操作。整體流程由檔案檢查、防呆、修復與結果輸出串接而成，確保操作順暢並提升使用體驗。

圖 11 為 GUI 介面。使用者上傳圖片後，系統會自動驗證格式與尺寸，通過後依選擇模式載入對應模型並執行靜態影像修復，使受損區域得以自然重建。

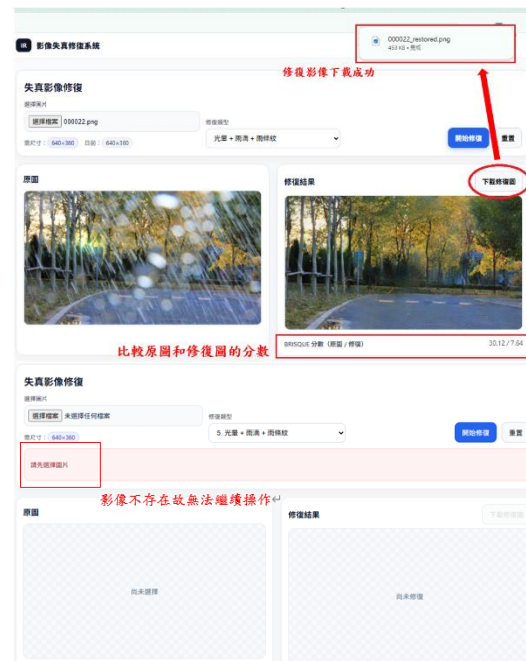


圖 10：靜態修復功能 WEB 成果展示

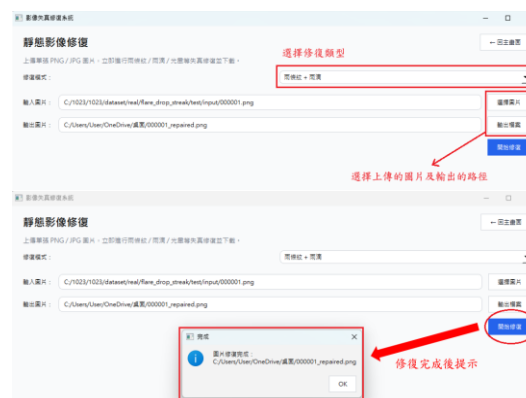


圖 11：靜態修復功能 GUI 成果展示

(三) 動態影像修復功能

本功能提供五項操作：上傳或錄製影片並預覽；自動切割影格作為逐幀輸入；按下「開始修復」後依序進行推論；呈現修復前後對比；最後下載完整修復影片。

如圖 12、圖 13 所示，影片載入後左側顯示原始畫面，右側為修復結果區域。系統自動完成影格切割與模型推論。完成後可查看還原後的影片品質提升並提供下載。整體操作簡潔流暢，使用者能快速完成影片修復並立即預覽成果。

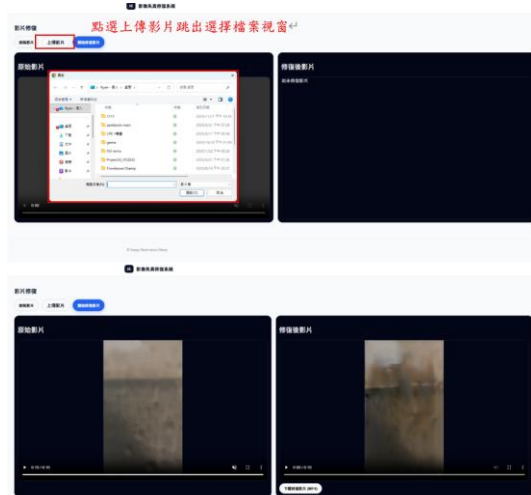


圖 12：動態修復功能 WEB 成果展示

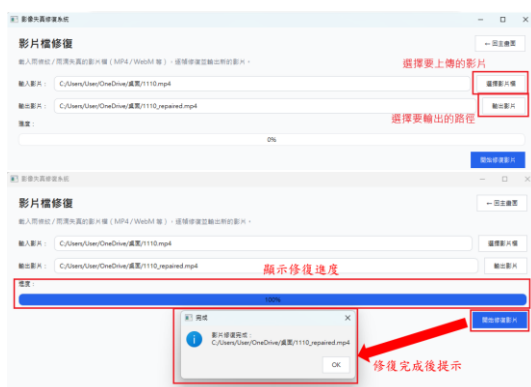


圖 13：動態修復功能 GUI 成果展示

(四) 即時影像修復功能

本功能提供五項操作：啟動鏡頭並取得權限；持續擷取影格作為模型輸入；以四通道影格（RGB+遮罩）進行即時修復；同步顯示原始與修復畫面；關閉鏡頭以終止串流並釋放資源。

如圖 14、圖 15 所示，使用者按下即時修復後授權攝影機，系統便將擷取的影格加入修復隊列並立即進行推論。模型針對受損區域進行填補與細節還原，使模糊或缺損位置自然融入整體畫面。使用者可在介面上直接預覽即時修復結果，整體流程順暢直觀。

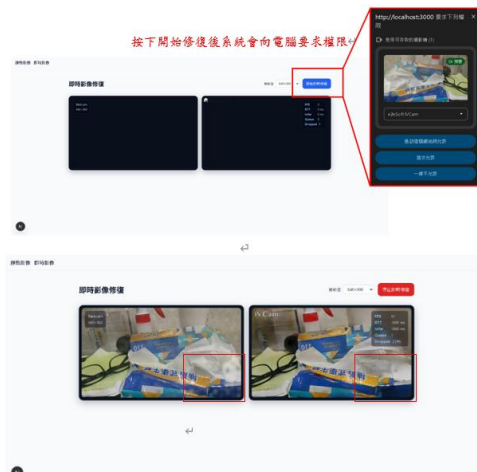


圖 14：即時修復功能 WEB 成果展示

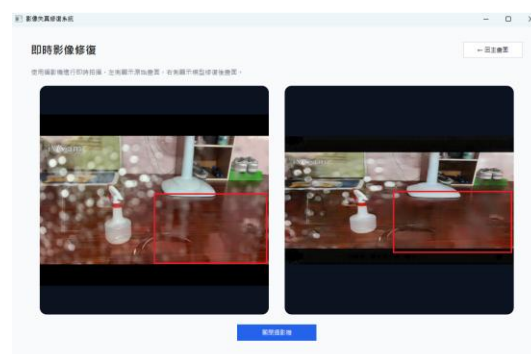


圖 15：即時修復功能 GUI 成果展示

五、結論

本專題旨在針對多重影像失真問題進行探討與系統設計，透過模型訓練與平台整合，嘗試提升影像在惡劣環境下的可辨識性與穩定性。

整體成果顯示，所提出的方法能在多數情境下有效改善畫面品質，並維持一定的穩定表現，驗證專題研究方向的可行性。雖仍受資料多樣性與遮罩準確度等因素影響，使部分結果未達理想，但整體流程已具雛形，能作為後續改良的基礎。

平台系統的建置亦展現出初步應用潛力，使使用者能以直觀介面操作影像修復流程。

本專題雖仍有進步空間，但已成功建立一套具實作價值的影像修復原型，為後續研究提供參考方向。

六、專題作品影音平台網址

<https://www.youtube.com/watch?v=V7Gs7deEzdA>



七、參考文獻

[1] 何笠 (He, Li), 《基於深度學習使用 U-Net 進行雨天影像多重失真之同時修復》，碩士論文，國立中山大學資訊工程學系，台灣，高雄市，2024 年

[2] Quan, Y., Ma, Y., Yang, H., and Ji, H., "Deep Raindrop Removal from a Single Image," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13873–13882, 2019.

[3] Wei, Y., Liu, S., Wang, W., et al., "Flare7K: A Dataset for Lens Flare Removal," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14456–14466, 2021.

[4] Ponomarenko, N., Ieremeiev, O., Lukin, V., et al., "Color Image Database TID2013: Peculiarities and Preliminary Results," *Proceedings of the 4th European Workshop on Visual Information Processing (EUVIP)*, pp. 106–111, 2013.

[5] Anish Mittal, Anush Krishna Moorthy, and Alan C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," *IEEE Transactions on Image Processing*, Vol. 21, No. 12, pp. 4695–4708, 2012.

[6] Ronneberger O., Fischer P., Brox T., "U-Net: Convolutional Networks for Biomedical Image Segmentation," *MICCAI*, pp. 234–241, 2015.

[7] Woo S., Park J., Lee J. Y., Kweon I. S., "CBAM: Convolutional Block Attention Module," *ECCV*, pp. 3–19, 2018.

[8] Si J., Kim S., "Single Image Rain Streak Removal Using Harris Corner Loss and R-CBAM Network," *arXiv:2507.23185*, 2025.

[9] Isola P., Zhu J.-Y., Zhou T., Efros A. A., "Image-to-Image Translation with Conditional Adversarial Networks," *CVPR*, pp. 1125–1134, 2017.

[10] Qian R., Tan R. T., Yang W., Su J., Liu J., "Attentive Generative Adversarial Network for Raindrop Removal from a Single Image," *CVPR*, 2018.