國立臺灣師範大學科技與工程學院電機工程學系
碩士論文

Department of Electrical Engineering
College of Technology and Engineering
National Taiwan Normal University
Master's Thesis

基於雨嵌入一致性和注意力機制之單張影像去雨
Single Image Deraining Using Rain Embedding Consistency and Attention Mechanism

黃冠樺

Kuan-Hua Huang

指導教授：康立威 博士

Advisor: Li-Wei Kang, Ph.D.

中華民國 112 年 8 月

August 2023

誌　　謝

快樂的時光總是過得如此的快，想當初興高采烈地來就讀國立臺灣師範大學電機研究所，轉眼即逝在這裡的三年生活也即將結束了。回首這些日子裡，回憶都還猶如昨日一樣，心中還是有很多的不捨。非常感謝在這段時間以來，受到許多人的陪伴、教導與支持。首先，我要感謝我的碩士班指導教授康立威博士，教授在任何時候都能給予我很大的幫助，不管是論文還是研究上，都會耐心地聆聽我的疑問，並給予適當的意見，使我能夠學習到許多影像處理的相關知識和技術，並且在研究中成長學習，而且在獨立思考、報告能力、論文研究…等專業能力上都有大幅提升。此外，特別感謝國立臺北科技大學電機工程系助理教授李曉祺博士和國立成功大學統計系助理教授許志仲博士在百忙之中特別撥空來擔任我的口試委員，並對於我的論文研究給予專業上的建議，使得本論文可以有更完整性的幫助。

感謝屈軒宇學長、高漢棋學長，在我剛進實驗室時充滿熱情的給我幫助，不管是在研究上還是課業上。感謝實驗室夥伴蔡洪弦、謝政霖、張政期、李政霖、劉晟岳、倪至謙、李沃晏、李汪翰，讓實驗室除了研究外還增添了許多的歡樂。最後，還想感謝我的家人以及親友們，心中除了感謝還是感謝，並期許未來自己能夠不負大家的期待，成為你們的驕傲。

冠樺 謹誌於

中華民國一一二年八月

基於雨嵌入一致性和注意力機制之單張影像去雨

學生：黃冠樺　　指導教授：康立威

國立臺灣師範大學電機工程學系碩士班

中文摘要

由於數位媒體的快速發展，影像處理的技術越來越受到人們的重視。不過由於影像資料之來源非常廣泛且品質難以控制，往往會有不同種因素的干擾，包括障礙物、光源、天氣等等，造成影像品質過低，可能會使其相關應用之效能大打折扣，甚至毫無用途。因此，為了解決這些難題，人們投入數位影像品質回復或強化的研究，在近些年來取得明顯的提高影像判讀性及可視性，還能幫助提高物件偵測的準確率。而在我們日常生活中，下雨是最常出現的情況，造成不管是拍攝影像或影片，都會因為雨水而造成影像不清晰。在目前現有的研究方法裡，有使用深度學習、多尺度、Transformer模型等影像去雨方法。其中在使用編碼器解碼器的去雨方式裡，通常是根據輸入的有雨影像來預測雨層。因此，編碼器解碼器的網路架構引起了廣泛的關注。但由於在編碼器階段需要提取影像裡有雨的特徵，而在提取的效果及精確度就很重要。為了解決這個問題，許多論文會加上各種模塊來提升提取的效果。

為了解決上述問題，本篇論文提出一個編碼器解碼器網路架構， 並且加上注意力模塊，使其在編碼器階段可以提取更多更準確的有雨特徵，且在編碼器解碼器裡常用的跳躍連接也改成注意力機制的模塊，以讓編碼器提取的特徵可以加強傳遞，使得解碼器可以更為準確預測雨層。在實驗階段，我們使用了多個知名影像資料集，包括Rain100H、Rain100L以及Rain800來訓練及測試所提出的網路架構效果。

關鍵字：影像去雨；編碼器解碼器；注意力機制。

Single Image Deraining Using Rain Embedding Consistency and Attention Mechanism

Student：Kuan-Hua Huang Advisor：Dr. Li-Wei Kang

Department of Electrical Engineering
National Taiwan Normal University

Abstract

Because of the rapid development of digital media, people pay more and more attention to image processing technology. However, due to the wide range of sources of image data and the quality is difficult to control, there are often interferences from various factors, including obstacles, light sources, weather, etc., resulting in low image quality, which may greatly reduce the performance of related applications, or even Useless. Therefore, in order to solve these problems, people have invested in the research of digital image quality restoration or enhancement, which has significantly improved image interpretation and visibility in recent years, and can also help improve the accuracy of object detection. In our daily life, rain is the most common occurrence, and whether it is shooting images or videos, the images will be unclear due to rain. Among the existing research methods, there are image deraining methods using deep learning, multi-scale, and Transformer models. Among them, in the rain removal method using an encoder-decoder, the rain layer is usually predicted based on the input rainy image. Therefore, the network architecture of encoder-decoder has attracted extensive attention. However, since the feature of rain in the image needs to be extracted in the encoder stage, the effect and accuracy of the extraction are very important. In order to solve this problem, many papers will add various modules to improve the extraction effect.

In order to solve the above problems, this paper proposes an encoder-decoder

network architecture, and adds an attention module, so that it can extract more and more accurate rainy features in the encoder stage, and is commonly used in encoder-decoder The skip connection is also changed to an attention mechanism module, so that the features extracted by the encoder can be strengthened, so that the decoder can predict the rain layer more accurately. In the experimental phase, we used several well-known image datasets, including Rain100H, Rain100L and Rain800, to train and test the effect of the proposed network architecture.

Keywords: Image deraining; Encoder-Decoder; Attention mechanism.

目　錄

圖 目 錄

表　目　錄

第一章　緒論

1.1　研究動機與背景

單影像去雨(Image deraining)是一種典型的 low-level 影像處理(Image processing)，目標從一張充滿雨水或雨痕的影像上，重建還原成乾淨的清晰影像，這是一個有挑戰性的不適定問題。而隨著現在自動駕駛汽車以及無人機已成為全球趨勢，兩者都需要依靠攝影機來實現對周圍環境的觀察。然而，一些環境的因素，比如雨、霧霾和其他障礙，會影響到攝影機的可見度，因此影像去雨等相關等可以提升影像品質的方法也顯得更為重要。再加上人工智慧(Artificial Intelligence，AI) 中的深度學習(Deep Learning)在影像處理快速蓬勃發展，各種影像恢復的技術已經取得非常出色的成果。

在過去十年裡，單影像去雨出現了各種方法，早期的方法通常是根據雨痕和清晰影像的統計特性施加各種先驗，到之後基於卷基神經網路(CNN)的學習法，再到近期的 Transformers 模型，使用編碼器解碼器的架構再加上注意力機制(Attention mechanism)。

為了測試影像恢復的表現，我們使用幾種不同的公開資料集來驗證我們的方法，像是合成影像數據集，包含合成小雨影像的 Rain100L，大雨影像的 Rain100H，還有同時合成大雨和小雨影像的 Rain800。

1.2 研究目的與方法概述

為了能夠將單張有雨影像做影像還原，我們使用編碼器解碼器來作為我們方法的主要架構。在網路架構中，主要分為兩個部分，去雨主架構以及自動編碼器。在我們的方法中，使用了注意力機制來加強雨特徵的提取，使重建的雨層可以與輸入的 ground-truth 雨層一樣接近。

我們在使用合成有雨影像資料集的結果發現，我們使用注意力機制實現了最先進的去雨性能，這表明注意力機制可能是可以提高編碼器解碼器的去雨整體性能的可行因素。

1.3 論文架構

本篇論文如下排序：總共由五個章節所組成，第一章，第二章，第三章，第四章，第五章。第一章提出了研究這個主題的動機，並講述本論文提出對應此議題，所提出的方法。第二章針對本研究相關之文獻作探討與介紹，我們會詳細介紹本篇論文透過影像處理上的重要文獻與研究，逕而得到靈感與啟發，第三章詳細說明本論文所提出的網路架構，第四章說明實驗結果，在該章節介紹我們所使用的資料集，以及實驗所使用的軟硬體架構、訓練細節與資料集等，我們會詳細的介紹與展示所提出架構的效能與面對各式不同資料集的結果，並針對實驗結果及消融測試做充分的分析，第五章對本研究之方法與實驗結果總結，並對未來研究提出改善方向。

第二章　文獻探討

近年來影像去雨已經取得相當亮眼的研究成果並且實際應用於實際場域，然而單幅影像去雨還是非常具有挑戰性，因為將觀察到的下雨影像分解為雨層和無雨背景影像是一個非常不適定的問題。本章將節介紹不同的去噪方法與相關研究文獻，2.1 傳統影像去雨相關方法，2.2 基於深度學習影像去雨相關方法，2.3 使用殘差學習相關方法，2.4 Transformer 相關方法，2.5 使用多尺度相關方法。

2.1　傳統影像去雨相關方法

影像去雨的問題普遍存在，影像去雨的主要目的是為移除影像中的雨水部分，使得影像可以看似清晰且乾淨。而且影像去雨技術可以提高影像品質，帶給觀賞者有更良好的觀看體驗，並且在視訊影像方面具有相當的潛力，如何消除雨水也就成為了我們主要研究議題。

在現今的生活中，人手一支智慧型手機已成為常態，也間接造成影像去雨或是美顏等透過後製的影像重建技術被廣為利用，而影像去雨在影像處理上仍然佔有相當重要的部分，目前在學術上已經有許多種不同的去雨方法被提出並解決此類問題。而傳統的方法有像稀疏編碼和字典學習法、低秩矩陣法以及基於梯度先驗等相關方法。

2.1.1 稀疏編碼和字典學習法

字典學習的目標，是提取事物最本質的特徵（類似於字典當中的字或詞語）。如果我們能獲取最本質的特徵，那我們就能掌握這個事物的最本質的內涵。換句話說，字典學習將我們得到的物體的訊息降維，以減少該物體一些無關緊要的訊息干擾我們對該物體的定義。而稀疏編碼是將大量的多餘變量去除，只保留與相應變量最相關的解釋變量，在簡化模型的同時又保留了數據中最重要的訊息。由此可見，稀疏編碼所要達到的目標與字典學習是一致的，都是要將多餘的訊息去除，並將最重要的本質訊息保留下來。

Kang 等人[1]提出一種基於單圖像的雨痕去除框架，首先使用雙邊濾波將輸入的雨圖像分解為高頻(HF)和低頻(LF)的部分。接著通過基於MCA[2]-[6]的字典學習法和稀疏編碼將 HF 部分分解為「雨分量」和「非雨分量」。然後把提取部分的字典分為兩個子字典。接著將兩個子字典進行稀疏編碼，可以得到 HF 部分的幾何分量，然後與圖像的 LF 部分結合得到去除雨水的圖象。

2.1.2 低秩矩陣法

低秩矩陣，表示行與列都具有相關性。簡單來講，就是可以用一部分的行與列，來表示另一部分的行與列。而剛好矩陣是稀疏性的，所以可以利用這種相關性來填充矩陣。

Chen 和 Hsu[18]在二維影像中提出一種低秩去雨模型，將輸入圖像分解為雨條紋分量、無雨分量。通常雨影像在不同局部斑塊中包含相似的雨痕模式，所以使用低秩模型來表徵雨痕出現。而該模型不需要字典學習，

也不需要其他可能會造成問題的預處理(例如,減少低頻內容、雨水檢測),因為低秩模型可以抓取由雨紋模式跨越的子空間。

2.1.3 基於梯度先驗的相關方法

Zhu 等人[8]提出一種通過將單個輸入影像分解為無雨背景 B 和雨紋層 R 的去除雨痕方法。過程中是由三個影像先驗輔助,首先分析雨影像的局部梯度統計,以分辨由雨痕主導的影像區域。從這些區域估計了主要的雨條紋方向。接下來,在背景層 B 上定義兩個先驗,一個基於估計雨的方向,另一個則是基於集中稀疏表示。接著基於補丁與提取的雨補丁的相似性,在雨紋層 R 上定義了第三個先驗。

2.2 基於深度學習影像去雨相關方法

從 2017 年以來,單影像去雨方法進入使用深度學習的時代,利用各種類型的網路,像是卷基神經網路(Convolutional Neural Network,CNN)、生成對抗網路(Generative Adversarial Network,GAN)、遞歸神經網路(Recurrent Neural Networks,RNN)等。這些利用深度卷基網路可以自動提取特徵,使它們能夠模擬從有雨圖像到乾淨圖像的更複雜映射。

Yang 等人[9]提出一種聯合雨水檢測和去除的深度網路,如圖 2-1 所示。自動檢測雨紋區可以約束雨水去除,使得網路能夠對雨區及非雨區執行自適應操作,保留更多豐富細節。接著為了檢索更多上下文訊息,提出一個擴張網路來擴大感受野,通過合成具有不同擴張因子的多個並行卷積。

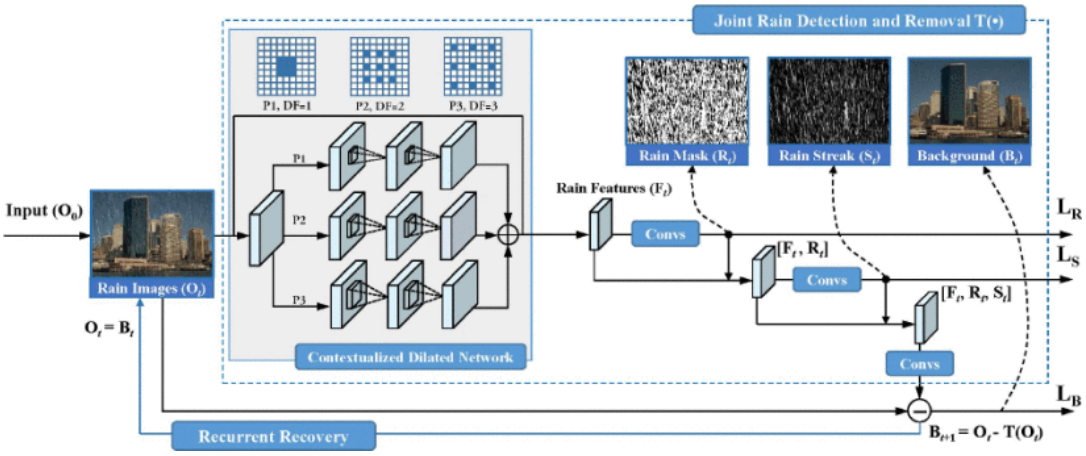


圖2-1　雨水檢測和去雨網路架構

（資料來源：[6]）

Fu 等人[10]提出一個基於深度 CNN 的 DerainNet 架構，如圖 2-2 所示。用於從單幅影像中去除雨水。DerainNet 直接從數據集中自動學習乾淨和雨水細節層之間的非線性映射函數。架構中並無使用堆疊隱藏層或增加神經元等方法來達到所需的映射函數，而是使用影像處理來修改目標函數並提高去雨品質。

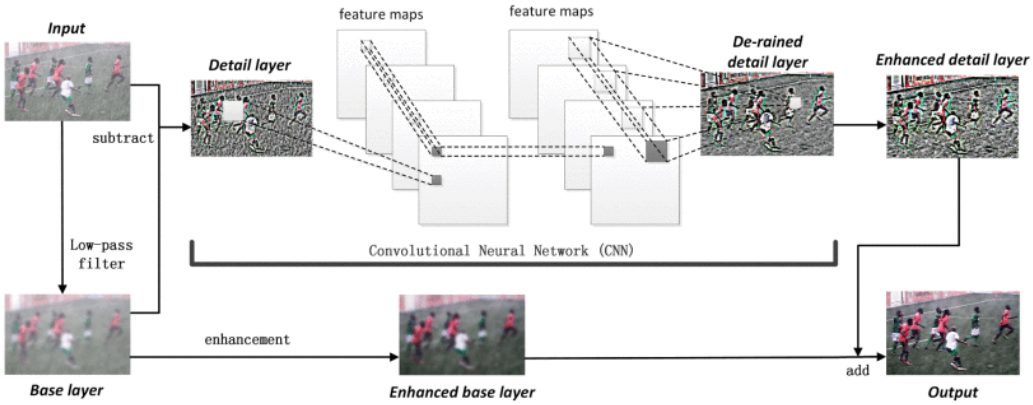


圖2-2　DerainNet 架構

（資料來源：[6]）

Li 等人[11]提出一個叫 RESCAN 模型，是一種基於深度卷積和遞歸神經網路的深度網路架構，如圖 2-3 所示。在每個階段使用具有多個全卷積的上下文聚合網路來去除雨痕。由於雨痕有不同的形狀和方向，在網路中的每個通道都有相對應的雨痕。Squeeze-and-Excitation(SE)模塊用於依據每個卷基層中的屬性，為各種通道分配不同的 alpha 值。受益於呈指數增長的卷基擴張，該網路具有較大的深度接受場，可以幫助獲取更多的上下文訊息。為了更好的去除雨水，使用循環神經網路(RNN)架構將三種循環單元結合起來，來幫助後面的去雨效果。

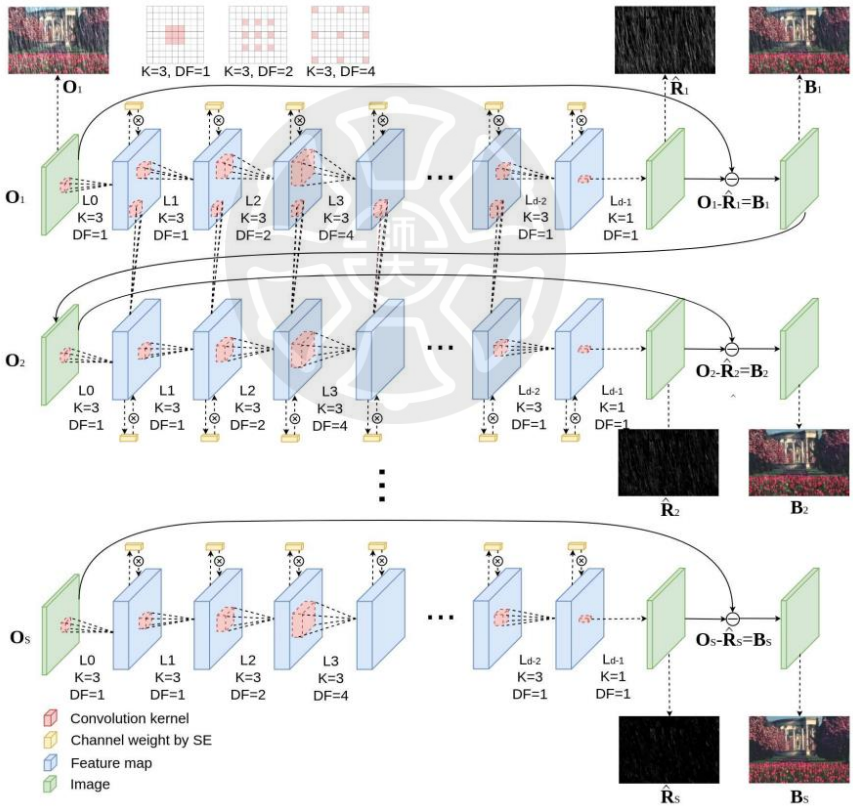


圖2-3 RESCAN 模型架構

（資料來源：[11]）

2.3 使用殘差學習相關方法

He 等人[12]在 2015 年提出深度殘差網路(ResNet)，並獲得 CVPR2016 最佳論文獎，而 ResNet 解決了深度 CNN 模型難訓練的問題。深度網絡的退化問題說明深度網絡不容易訓練，作者提出了殘差學習來解決退化問題，如圖 2-4 所示。

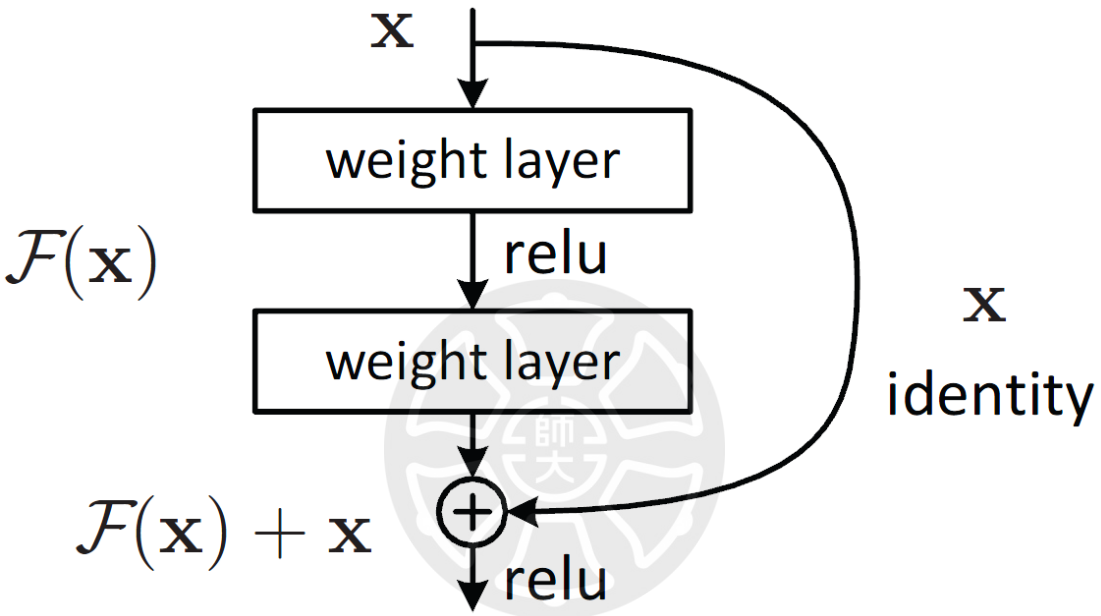


圖2-4 殘差學習圖

（資料來源 ：[12]）

ResNet 是參考了 VGG19[13]網路，在其基礎上進行了修改，並加入了殘差學習，如圖 2-5 所示。

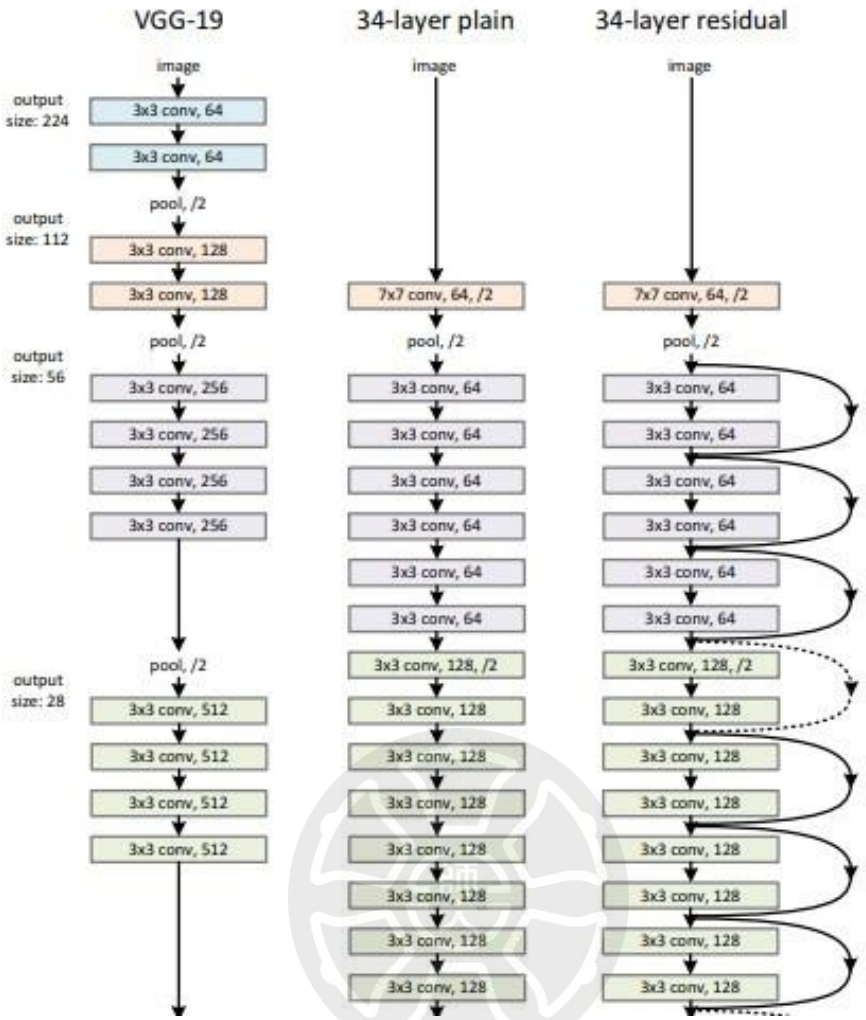


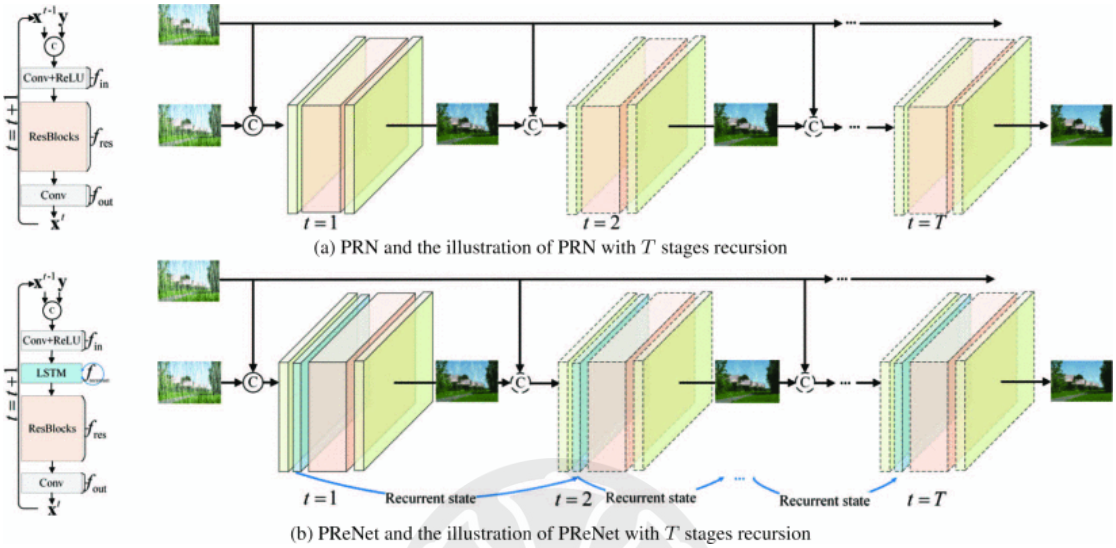
圖2-5 ResNet 網路架構圖

（資料來源 ：[13])

Fu 等人[14]提出一種深度細節網路架構，用於基於深度卷基神經網路從單個影像中去除雨痕。受到 ResNet 啟發，作者提出了一個深度細節網路來直接減少從輸入到輸出的映射範圍，使得學習過程更加容易。作者採用 ResNet 結構作為參數層來深入探索圖像特徵，不是直接在圖像上應用 ResNet，而是利用先驗知識並使用圖像細節層作為輸入。

Ren 等人[15]在考慮了網路架構、輸入輸出以及損失函數，提出一種基線網路。作者使用具有五個殘差塊(ResBlocks)的 ResNet，在不增加模型

參數的情況下，通過 ResNet 遞歸展開多個階段來引入漸進式 ResNet(PRN)，如圖 2-6 裡的(a)。此外，還引入循環層以利用跨遞歸階段的深度特徵依賴性形成 PReNet，如圖 2-6 裡的(b)。



(a) PRN and the illustration of PRN with T stages recursion

(b) PReNet and the illustration of PReNet with T stages recursion

圖2-6 漸進式網路架構

(資料來源：[15])

2.4 Transformer 相關方法

Transformer 模型[16]是 Google 團隊在 2017 提出的一種自然語言處理(NLP)的模型。此模型是一種基於 Seq2Seq[17] 和 self-attention mechanism 合成的深度學習模型，減少了 CNN 的缺點(即有限的感受野和對輸入內容的不適應性)，而該模型主要是應用在自 NLP 與電腦視覺(CV)領域。

在 2020 年，Google Brain Team[18]提出一個叫 Vision Transformer(ViT)的網路架構，如圖 2-7(左)所示。作者受到 NLP 中的 Transformer 成功的啟發，直接將一個標準的 Transformer 模型應用在圖像上，完全捨棄了 CNN。

而相對於 NLP 領域的每個輸入單位使用 Word Embedding，本篇論文提出了 Patch Embedding 作法，將影像切分為圖塊像量。在把影像資料送入 Transformer Encoder 前，需先將其轉換為類似 NLP 模型中類似 Sequence 的形式，如圖 2-8(1)所示。而 Transformer Encoder 包含兩個重要模組，多頭注意力(MSA)和全連接層(MLP)，如圖 2-7(右)所示。

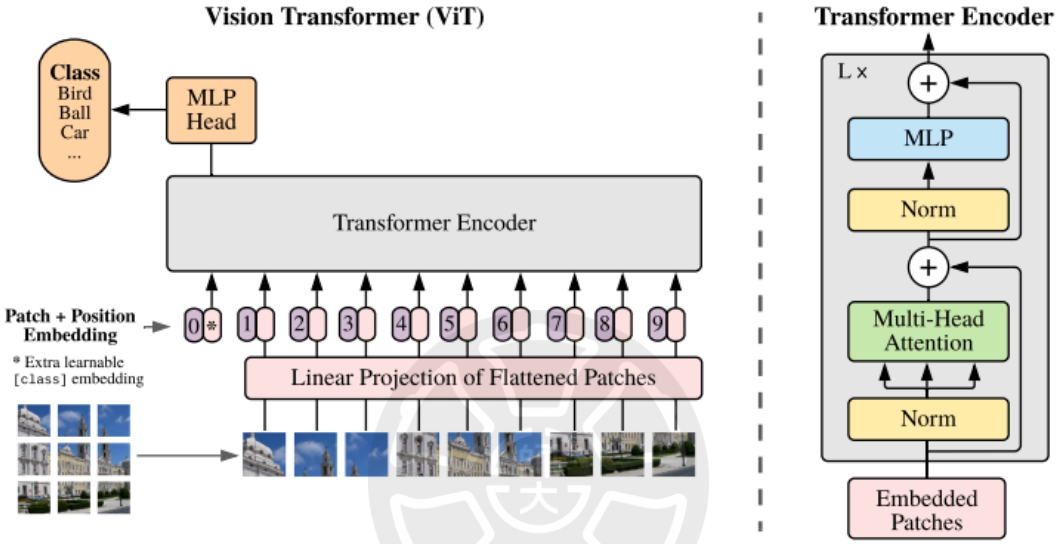


圖2-7 ViT 網路架構
(資料來源：[13])

$$\mathbf{z}_0 = [\mathbf{x}_{\text{class}}; \mathbf{x}_p^1\mathbf{E}; \mathbf{x}_p^2\mathbf{E}; \cdots; \mathbf{x}_p^N\mathbf{E}] + \mathbf{E}_{pos}, \qquad \mathbf{E} \in \mathbb{R}^{(P^2 \cdot C) \times D}, \mathbf{E}_{pos} \in \mathbb{R}^{(N+1) \times D} \tag{1}$$

$$\mathbf{z}'_\ell = \text{MSA}(\text{LN}(\mathbf{z}_{\ell-1})) + \mathbf{z}_{\ell-1}, \qquad \ell = 1 \ldots L \tag{2}$$

$$\mathbf{z}_\ell = \text{MLP}(\text{LN}(\mathbf{z}'_\ell)) + \mathbf{z}'_\ell, \qquad \ell = 1 \ldots L \tag{3}$$

$$\mathbf{y} = \text{LN}(\mathbf{z}_L^0) \tag{4}$$

圖2-8 ViT 的演算法
(資料來源 ：[18])

Chen 等人[19]使用 Transformer 架構開發了一個預訓練的圖像處理模型，即 Image Processing Transformer(IPT)的網路架構，如圖 2-9 所示。由於預訓練模型需要兼容不同的圖像處理任務，包含去雨、去噪、超分率等，

需要用大規模數據集來開發 Transformer 的潛力。而該網路架構是由四個部分組成，第一部分是用於從輸入損壞圖像(例如，有雨的影象或帶有噪聲的影象)提取特徵，中間部分是恢復數據中缺失訊息的編碼器解碼器架構，最後一部分是將特徵映射到恢復的圖象中。

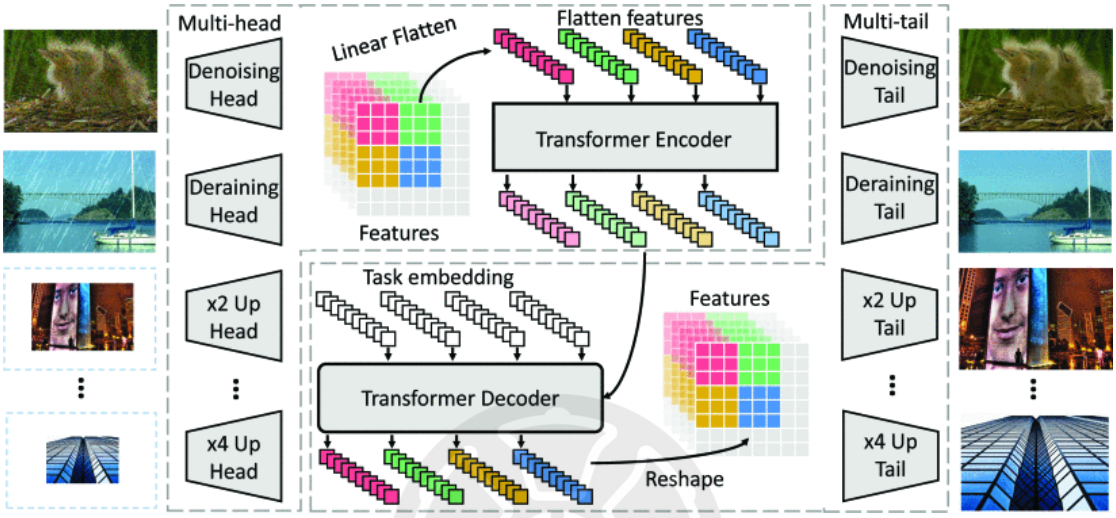


圖2-9 IPT 網路架構
(資料來源：[15])

Zamir 等人[20]提出一種叫 Restormer 的網路架構，是一種編碼器解碼器的 Transformer，如圖 2-10 所示。用於在高分辨率圖像上進行多尺度局部與全局的表示學習，而不將圖像分解成局部窗口，從而利用遠端圖象的上下文訊息。

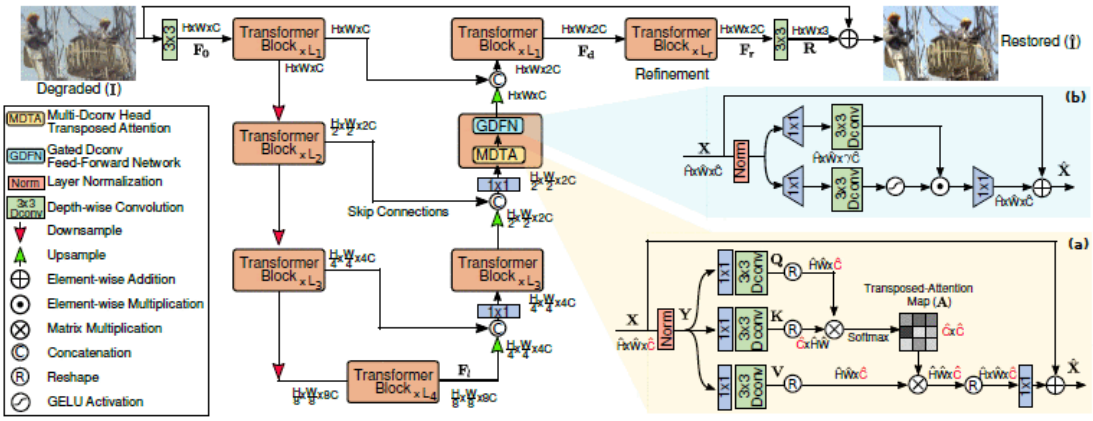


圖2-10 Restormer 網路架構
(資料來源：[15])

Wang 等人[21]提出了 Uformer 網路架構，是一種基於 Transformer 的圖像恢復結構，如圖 2-11 所示。Uformer 建立在一個 UNet 架構[22]，作者將卷基層修改為 Transformer 塊，同時保持相同的整體分層編碼器解碼器和跳躍連接。

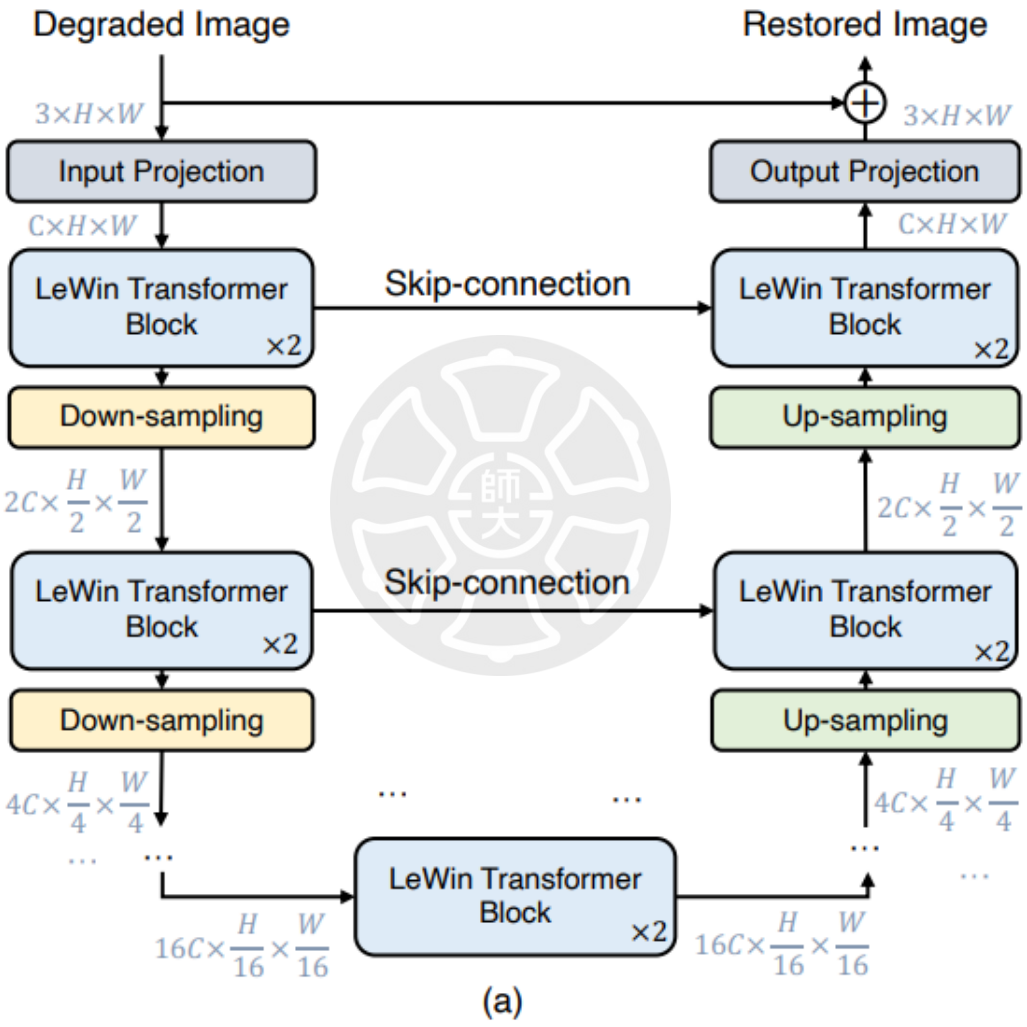


圖2-11 Uformer 網路架構
(資料來源：[21])

2-5 使用多尺度相關方法

Jinag 等人[23]提出 Multi-Scale Progressive Fusion Network(MSPFN)網路，如圖 2-12 所示。作者設計了三個基本模塊，粗融合模塊(CFM)、精細融合模塊(FFM)和重建模塊(RM)，以更有效的提取和整合多尺度訊息。在這些模塊中，具有相同尺度或不同尺度(金字塔層)的雨紋相似模式的互補訊息來逐漸整合，以協同的方式表徵雨紋分佈。

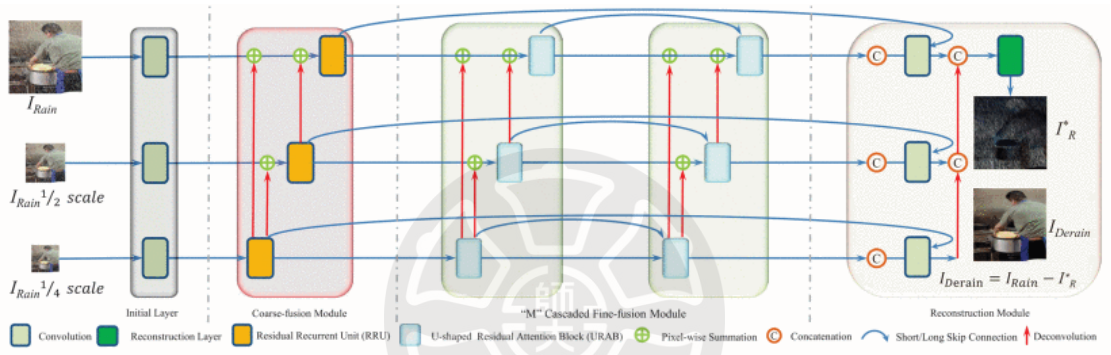


圖2-12 MSPFN 網路
(資料來源：[23])

Zamir 等人[24]提出 Multi-Stage Progressive Image Restoration(MPRNet)架構，如圖 2-13 所示。一開始的部分使用編碼器解碼器來學習多尺度上下文訊息，在最後的部分對原始圖像分辨率進行操作以保留精細的空間細節。作者還添加了跨接段特徵融合(CSFF)機制，有助於將多尺度上下文特徵從一開始傳到最後的部分。

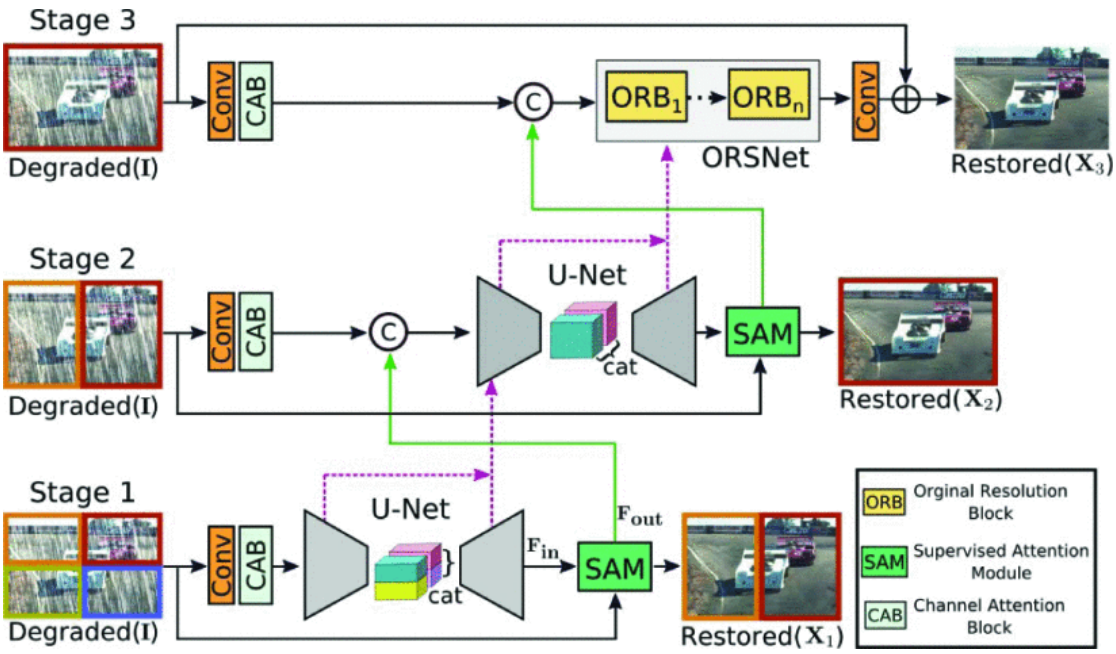


圖2-13 MPRNet 架構
(資料來源：[16])

2.6 使用編碼器解碼器相關方法

Li 等人[25]提出了一個編碼器解碼器架構，並提出“Rain Embedding Consistency”的概念，將自動編碼器的編碼嵌入視為理想的雨嵌入。作者還提出了一種具有長短期記憶（LSTM）的循環神經網絡（RNN）框架，將 LSTM 應用於編碼器的每個尺度，並讓每個尺度的編碼特徵傳遞到後續階段。

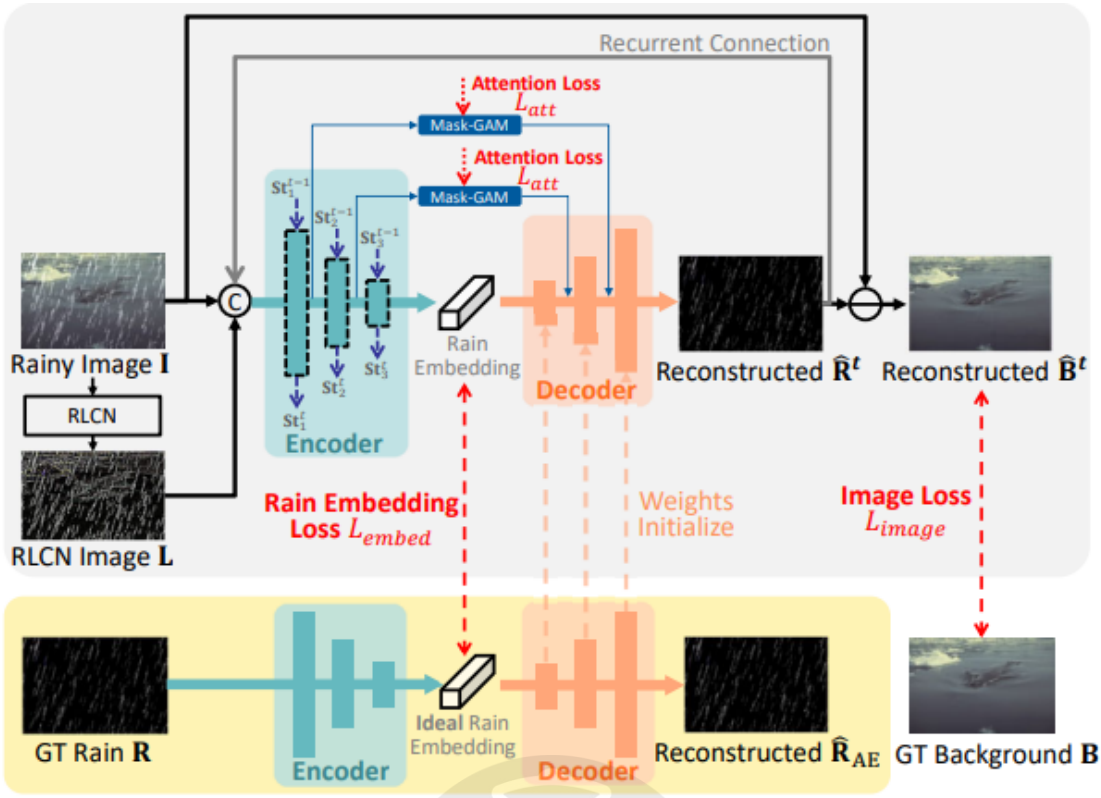


圖2-14 編碼器解碼器架構
(資料來源：[25])

Li 等人[34]提出了一個非局部網路增強的編碼器解碼器(NLEDN)網路架構，如圖 2-15 所示。在編碼器階段能夠學習特徵以獲得準確的雨紋建模，並在解碼器階段能夠去除所有級別的雨紋，同時保留雨紋的細節。作者還加入了非局部增強密集塊(NEDB)，不僅可以充分利用所有嵌入的卷積層特徵，還可以很好地捕獲空間上下文。

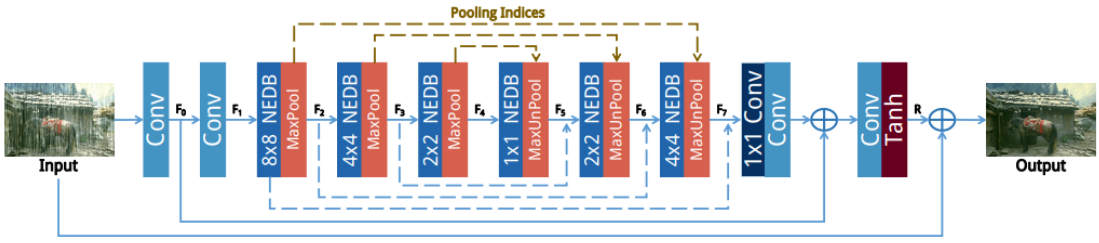


圖2-15 NLEDN 架構
(資料來源：[16])

Chen 等人[35]提出了 Half Instance Normalization Network(HINet)的網路架構，如圖 2-15 所示。架構中包含了兩個編碼器解碼器網路。在編碼器的部分，作者設計了 HIN 模塊來提取每個尺度的特徵，並採用文獻[24]的

CSFF(Cross-Stage Feature Fusion)與 SAM(Supervised Attention Module)來對兩個子網路進行連接。

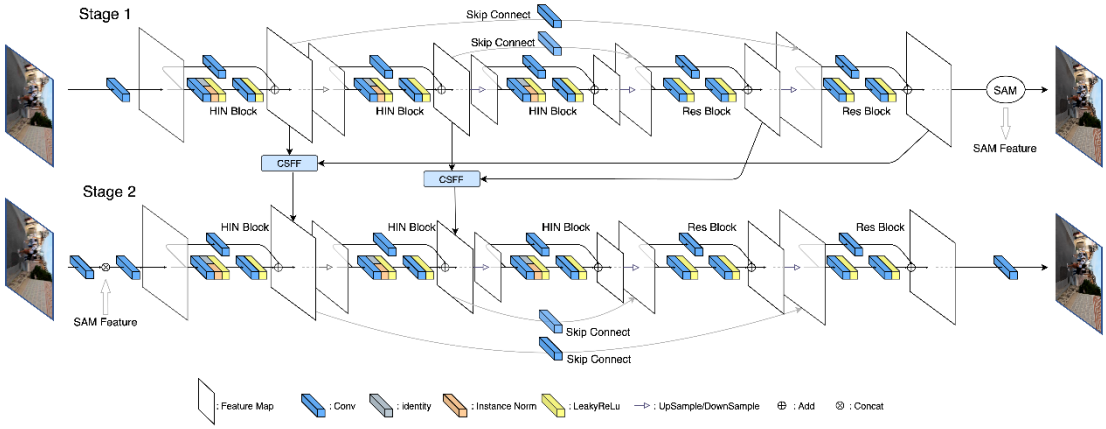


圖2-16 HINet 網路
(資料來源：[35])

第三章　研究方法

本章節將介紹本論文提出的去雨研究方法，為了重建出乾淨的影像，我們提出的研究方法架構由兩個部分組成：去雨架構與自動編碼器，透過以上兩部分可以將有雨的影像重建成乾淨的影像。本章各節將提及：3.1 自動編碼器；3.2 去雨的網路架構；3.3 介紹損失函數。

3.1　自動編碼器

本篇論文提出一個編碼器解碼器的去雨網路架構，是由一個編碼器解碼器架構以及自動編碼器組成。其中編碼器解碼器的典型目標是預測輸入有雨影像之間的殘差。由於殘差對應於雨層，所以預測無雨影像是通過輸入的有雨影像減去預測殘差所得到的。在這種情況下，編碼器需要有一個能很好地表示雨層訊息的高質量雨嵌入。

我們整體的做法與[25]相似，但在提取與傳遞影像特徵的方法與原論文不同，我們在其中加入了通道注意力模塊和多尺度像素注意力模塊。去雨架構如圖 3-1 所示，自動編碼器如圖 3-2 所示。

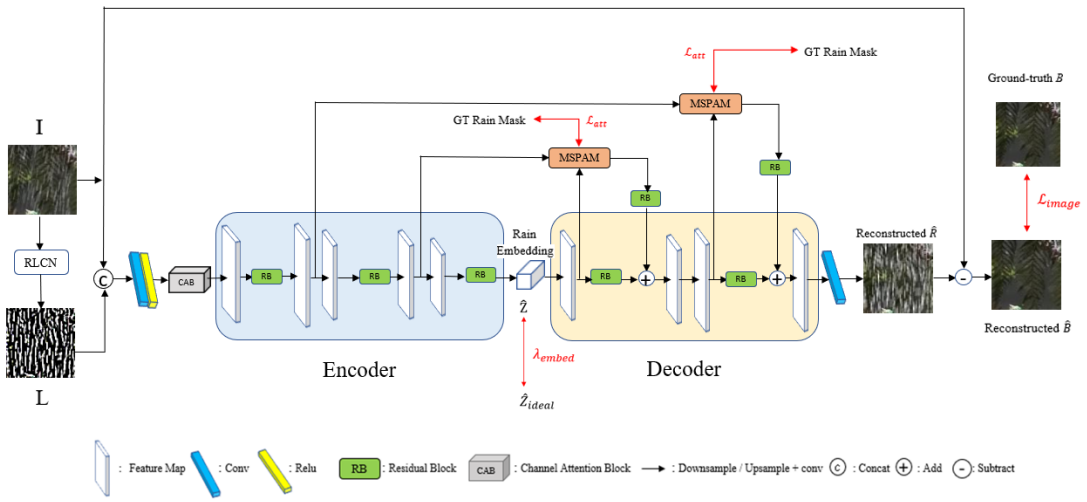


圖3-1 提出的去雨網路架構(因版面不足只畫出三個尺度，實際有四個尺度)

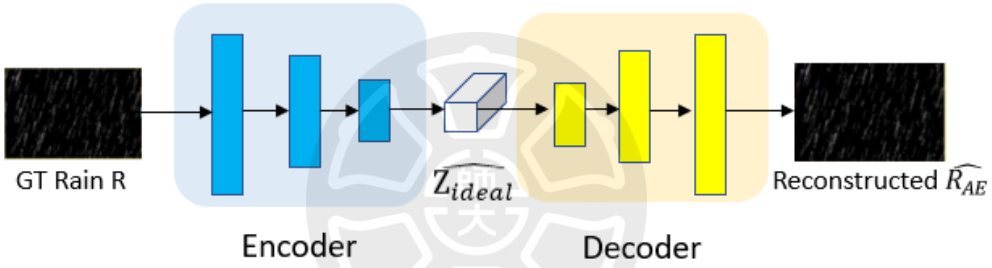


圖3-2 自動編碼器架構

我們使用自動編碼器來進行雨嵌入監督，採用和去雨架構相同的編碼器解碼器。為了訓練自動編碼器，ground-truth 的雨層生成方式為 R = I – B，其中 I 和 B 是從訓練數據集中採樣的一對有雨影像 I 和 ground-truth 背景影像 B。自動編碼器是利用 ground-truth 雨層作為輸入，生成理想的雨嵌入編碼$\hat{Z}_{ideal}$。接下來，解碼器使用$\hat{Z}_{ideal}$生成重建雨層$\hat{R}_{AE}$。

3.2 去雨網路主架構

為了進一步規範雨嵌入學習，我們使用校正局部對比度歸一化(Rectified Local Contrast Normalization，RLCN)來作為編碼器的輸入。首先將有雨影像輸入到 RLCN 中，如(式 3-1)所示。

$$L(i,j,c) = \frac{\phi_{rec}(I_c(i,j) - \mu I_c(i,j))}{\sigma I_c(i,j) + \epsilon} \quad \text{(式 3-1)}$$

其中，L 是輸出的 RLCN 影像，(i,J)表示像素座標，$c \in (R,G,B)$表示顏色通道，I_c表示輸入影像的第 c 個通道，$\mu I(i,j)$表示在像素(i,j)中以像素為中心的局部方形窗口內的平均像素強度，$\sigma I(i,j)$表示局部窗口內像素強度的標準差，ϵ表示一個穩定性的小值，$\phi_{rec}(\cdot) = \max(\cdot, 0)$表示輸出與零相比的最大值的整流函數。

在式 3-1 中，如果像素(i,j)的強度顯著高於或低於局部窗口內其他像素的強度，則$I_c(i,j) - \mu I(i,j)$的值取絕對值。因為通常我們專注在顯示較高像素的雨，所以使用 max 函數來執行校正並過濾負值，校正後的正值通過對應於局部對比度的局部標準偏差$\sigma I(i,j)$進一步歸一化，以能夠提取弱雨區和強雨區的像素。

圖3-3 (左)輸入的有雨影像 (右)輸出的 RLCN 影像
(資料來源：[25])

如圖 3-3 所示，生成的 RLCN 影像可以覆蓋大部分雨區，因此可以作為雨嵌入編碼的指導，從而提高編碼器的雨嵌入一致性的性能。

在提取影像特徵的方法使用了殘差模塊(Residual Block)，如圖 3-4 所示。

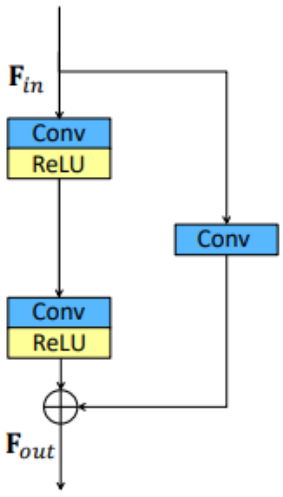


圖3-4 殘差模塊

為了加強影像特徵的提取，我們使用了通道注意力模塊(Channel Attention block，CAB)[26]，如圖 3-5 所示。

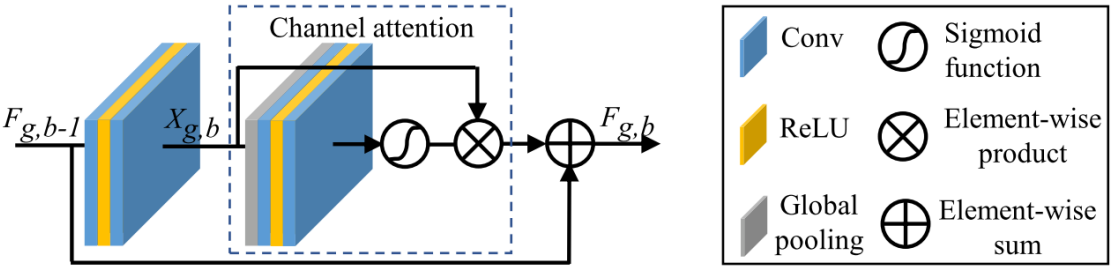


圖3-5 通道注意力模塊(虛線處部分)

通道注意力模塊的作法是，輸入一個是 H x W x C 的特徵，先進行一個空間的全局平均池化(Global Average Pooling)得到一個 1 x 1 x C 的通道特徵。接著再經過一個下採樣層和上採樣層得到每一個通道的權重係數，再將權重係數和原來的特徵相乘即可得到縮放後的新特徵。如圖 3-6 所示。

其中，下採樣層和上採樣層都是使用 1 x 1 的卷積，下採樣層的通道數所縮小 r 倍，r 取 16。接著，下採樣使用 Relu 激活函數，而上採樣用的是 Sigmoid 函數。

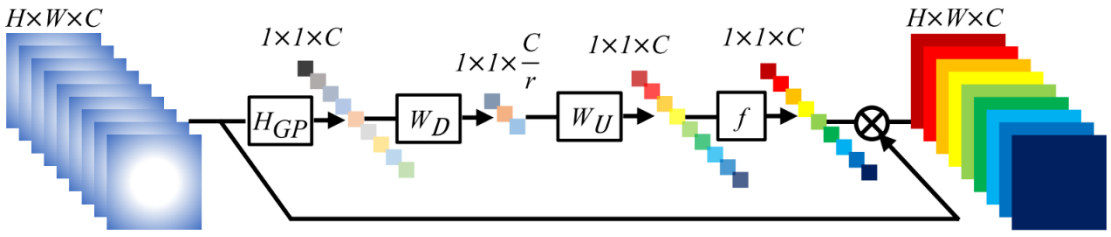


圖3-6 通道注意力機制

$$g_c = H_p(F_c) = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W} X_c(i,j) \qquad \text{(式 3-2)}$$

其中$X_c(i,j)$代表第 c 個通道X_c在位置(i,j)的值，H_p代表全局池化函數。

$$CA_c = \sigma(Conv\left(\delta\left(Conv(g_c)\right)\right)) \quad \text{(式 3-3)}$$

其中σ表示 Sigmoid 函數，δ表示 Relu 函數。

跳躍連接(Skip connection)廣泛應用於編碼器解碼器之中，用來重用編碼器的特徵以防止解碼時過度平滑。但由於 RLCN 影像含有一些關於非雨區的附加邊緣，會讓編碼器的特徵中存在和雨無關的特徵。如果直接將特徵從編碼器傳遞到解碼器，可能會影響到解碼階段。

為了解決這個問題，我們提出了多尺度像素注意力模塊(Multi Scale Pixel Attention Module，MSPAM)，如圖 3-7 所示。多尺度注意力模塊包含了兩個 3×3 卷積層、三個不同擴張率的擴張卷積、ReLU 函數、一個 1×1 卷基層，還有 Sigmoid 激活函數。

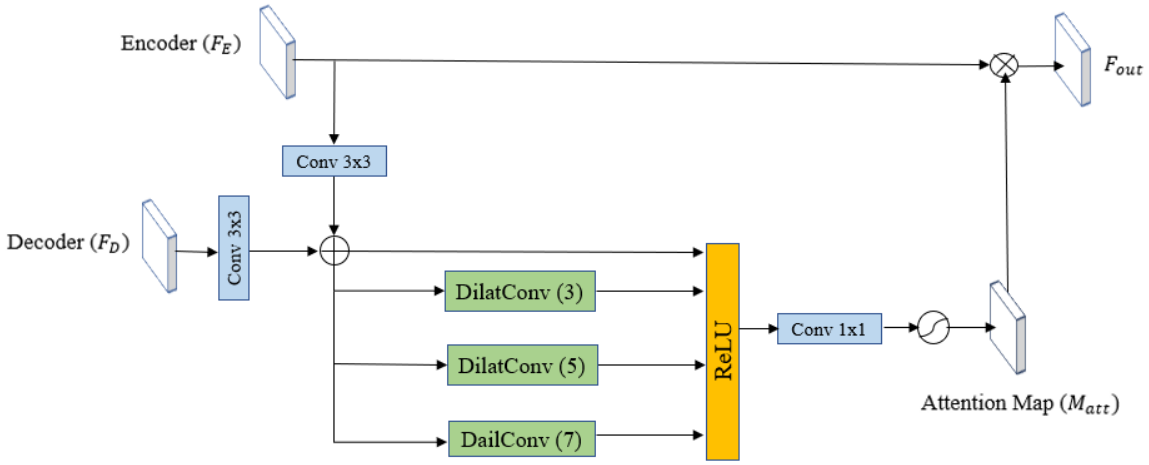


圖3-7 多尺度像素注意力模塊

$$\hat{F}_m = W_E * F_E + W_D * F_D \quad \text{(式 3-4)}$$

其中F_E與F_D分別表示編碼器的特徵和解碼器的特徵，W_E與W_D表示3×3 卷積層的權重。

$$\hat{F} = [\hat{F}_m, DConv_3(\hat{F}_m), DConv_5(\hat{F}_m), DConv_7(\hat{F}_m)] \quad \text{(式 3-5)}$$

其中$\hat{F}$表示$\hat{F}_m$與三個經過$DConv$的特徵圖進行合併。

$$M_{att} = \phi_2(W_1 * (\phi_1(\hat{F}))) \quad \text{(式 3-6)}$$

其中M_{att}表示由多尺度注意力模塊生成的注意力圖，W_1表示 1×1 卷積層的權重，ϕ_1表示 ReLU 函數。ϕ_2表示 Sigmoid 激活函數，用來生成注意力圖。接著使用獲得的注意力圖來重新校準局部特徵$\hat{F}_E$，最後將 MSPAM 生成的注意力增強特徵F_{out}向後傳遞。

3.3 網路訓練目標及損失函數

我們使用了三個損失函數來實現，總體損失函數如(式 3-7)所示：

$$\mathcal{L} = \lambda_{embed}\mathcal{L}_{embed} + \lambda_{att}\mathcal{L}_{att} + \lambda_{image}\mathcal{L}_{image} \quad \text{(式 3-7)}$$

第一個損失$\mathcal{L}_{embed}$，在從自動編碼器對高雨層重建性能的觀察表明，自動編碼器可以通過其推導出一個理想的雨嵌入$\hat{Z}_{ideal}$，再通過解碼器推導

出的雨層$\hat{R}_{AE}$，能夠非常接近輸入的$\hat{R}$。這使我們能夠專注於訓練編碼器 E，以提高自動編碼器的理想雨嵌入$\hat{Z}_{ideal}$，與提出的去雨網路編碼器預測的雨嵌入$\hat{Z}$之間的雨嵌入一致性。而$\mathcal{L}_{embed}$使用的是 L1 loss，如(式 3-8)所示。

$$\mathcal{L}_{embed} = \| \hat{Z}_{ideal} - \hat{Z} \|_1 \quad \text{(式 3-8)}$$

第二個損失$\mathcal{L}_{att}$，為了監督從 MSPAM 所產生的注意力圖的學習，我們使用通過對 ground-truth 雨層 R 做閾值化而導出的 ground-truth 圖來做損失函數，並對三種不同尺度的注意力圖做平均。而$\mathcal{L}_{att}$使用的是 L2 loss，如(式 3-9)所示。

$$\mathcal{L}_{att} = \frac{\sum_1^3 \|\hat{M}_{att} - M\|_2}{3} \quad \text{(式 3-9)}$$

其中，$\hat{M}_{att}$代表不同尺度的注意力圖，M 代表 ground-truth 圖。

最後一個損失L_{image}是圖像損失，我們使用負 SSIM 損失[33]來評估 ground-truth 與重建的背景影像差異，如(式 3-10)所示。

$$\mathcal{L}_{image} = -SSIM(B, \hat{B}) \quad \text{(式 3-10)}$$

其中B代表的是 ground-truth，$\hat{B}$代表重建的背景影像。

第四章　實驗結果分析

此章節將說明本研究嘗試之所有實驗，首先於 4.1 介紹實驗相關配置，4.2 影像去雨結果，4.3 消融測試。

4.1　軟硬體架構

硬體部分，我們在實驗使用的整體硬體配置如表 4-1 所示如下：

表4-1　電腦硬體設備

CPU	Intel® Xeon® Gold6154@ 3.00GHz
GPU	NVIDIA Tesla V100 32G
RAM	90GB

隨著深度學習的發展，目前仍然以使用 Python 程式語言為主，以 Python 為主衍伸了許多不同的開發環境，我們主要是使用 PyTorch 來做為我們的開發環境，使用 PyTorch 的好處，主要是可以開啟半精度運算和加快運算速度，並且減少 GPU 資源佔用，我們使用的作業系統為 Linux，整體環境配置如表 4-2。

表4-2　開發環境規格

Linux 版本	Ubuntu 20.04
CUDA 版本	11.7
cuDNN 版本	8.4
Python 版本	3.8
Pytorch 版本	2.0.0

4.1.1　訓練細節

以下我們的實驗設定。

1. 本實驗總共訓練 100 個時期(Epoch)。
2. 所有資料集都裁剪成 96 x 96 的補丁大小用於訓練。
3. RLCN 影像計算的局部窗口設置為 9 x 9 的大小，網路的循環時間 N 設為 6。
4. 模型訓練都不載入預訓練權重，皆從頭開始訓練模型。
5. 模型訓練批量(Batch size)大小皆為 4。
6. 我們優化器(Optimizer)使用 Adam(Adaptive Moment Estimation)
7. 模型學習率(Learning Rate)設置為 0.001，我們在同時達到第 25、50、75 個 Epoch 時會乘以 0.2 來降低學習率。
8. 損失權重的部分，$\lambda_{embed} = 0.02$, $\lambda_{att} = 0.1$, $\lambda_{image} = 1$ 。

4.1.2　資料集

在資料集的選擇上主要採用合成的有雨影像資料集，我們選擇有發表過論文的資料集做為測試模型架構的依據，讓實驗較有客觀性，並具有說服力。

資料集我們使用的是 Rain800[31]、Rain100H[30]、Rain100L[30]、

Rain200H[30]資料集。其中，Rain800 包含 700 個訓練對與 98 個測試對（由於影像太大，從原論文的 100 個測試影像中刪除 2 個影像），Rain100H 包含 1800 個訓練對與 100 個測試對，Rain100L 包含 200 個訓練對與 100 個測試對，Rain200H 包含 1800 個訓練對與 200 個測試對。如圖 4-1 與 4-2 所示。

圖4-1 Rain100H 資料集
(資料來源:[30])

其中 Rain100H 包含五種不同方向的雨紋，Rain100L 只有一種方向的雨紋，Rain800 是不同強度、方向的雨紋都有，Rain200H 則是與 Rain100H 同樣的訓練對。

圖4-2 Rain100L 資料集

(資料來源:[30])

4.2 影像去雨結果

在本節中，我們將說明所提出方法的實現細節。主要分成三個小節，第一部分我們會介紹兩種常用的性能指標用於比較和評估，像是峰值訊號與雜訊比 PSNR(Peak Signal-to-Noise Ratio)和結構相似性指數測度 SSIM(Structural Similarity Index Measure)[19]。第二部分是我們使用合成的有雨影像進行實驗。第三部分是我們進行了消融研究以證明所提出方法的有效性。

兩種常用的性能指標用於比較和評估：PSNR 和 SSIM。在三個通道的 RGB 顏色空間上，使用 PSNR 和 SSIM 的標準來評估影像去雨的結果。PSNR 的標準是基於每個對應影像的像素之間誤差。因為這個標準只考慮了數值誤差，所以有時 PSNR 的值比較高時，並不一定有很好的視覺效果。因此，通常會搭配 SSIM 的準則，SSIM 是通過比較影像結構訊息的變化來觀察影像的失真，從中獲得客觀的質量評估。PSNR 和 SSIM 的標準都是基於亮度，這些標準的值越高，影像重建的品質就越好，PSNR 指標公式

如(式 4-1)。

$$PSNR(X,Y) = 10log_{10}\frac{255^2}{MSE} \quad \text{(式 4-1)}$$

其中，255 代表 8bit 內最大的值，MSE 代表統計學內的均方誤差，其公式如(式 4-2)。

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\ \sum_{j=1}^{N}\ (X(i,j) - Y(i,j))^{\,2} \quad \text{(式 4-2)}$$

其中，$X(i,j)$和$Y(i,j)$分別代表兩張圖片位於相同位置的 pixel，MSE 代表平均一個 pixel 的均方誤差，因此較專注於局部的差異。

乾淨影像 X 和去雨影像 Y 在同一位置的 X 影像塊和 Y 影像塊之間的 SSIM 標准定義如(式 4-3)所示：

$$SSIM(X,Y) = \frac{(2\mu X\mu Y + c_1)2\sigma X\sigma Y + c_2)}{(\mu_X^2 + \mu_Y^2 + c_1)(\sigma_X^2\sigma_Y^2 + c_2)} \quad \text{(式 4-3)}$$

μX 和 μY 分別代表 X 與 Y 的平均值。同時，σX和σY是 X 與 Y 的偏差。c_1和c_2是小常數。然後，SSIM(X, Y)的標準是影像上基於塊分布的 SSIM 的平均值。

4.2.1 與現有方法比較

為了確認合成影像的去雨效果，我們選擇了以前的三種最先進方法做比較，包括:PReNet[15]、MPRNet[24]、ECNet[25]。測試於 Rain800 資

料集的效果如圖 4-3 所示:

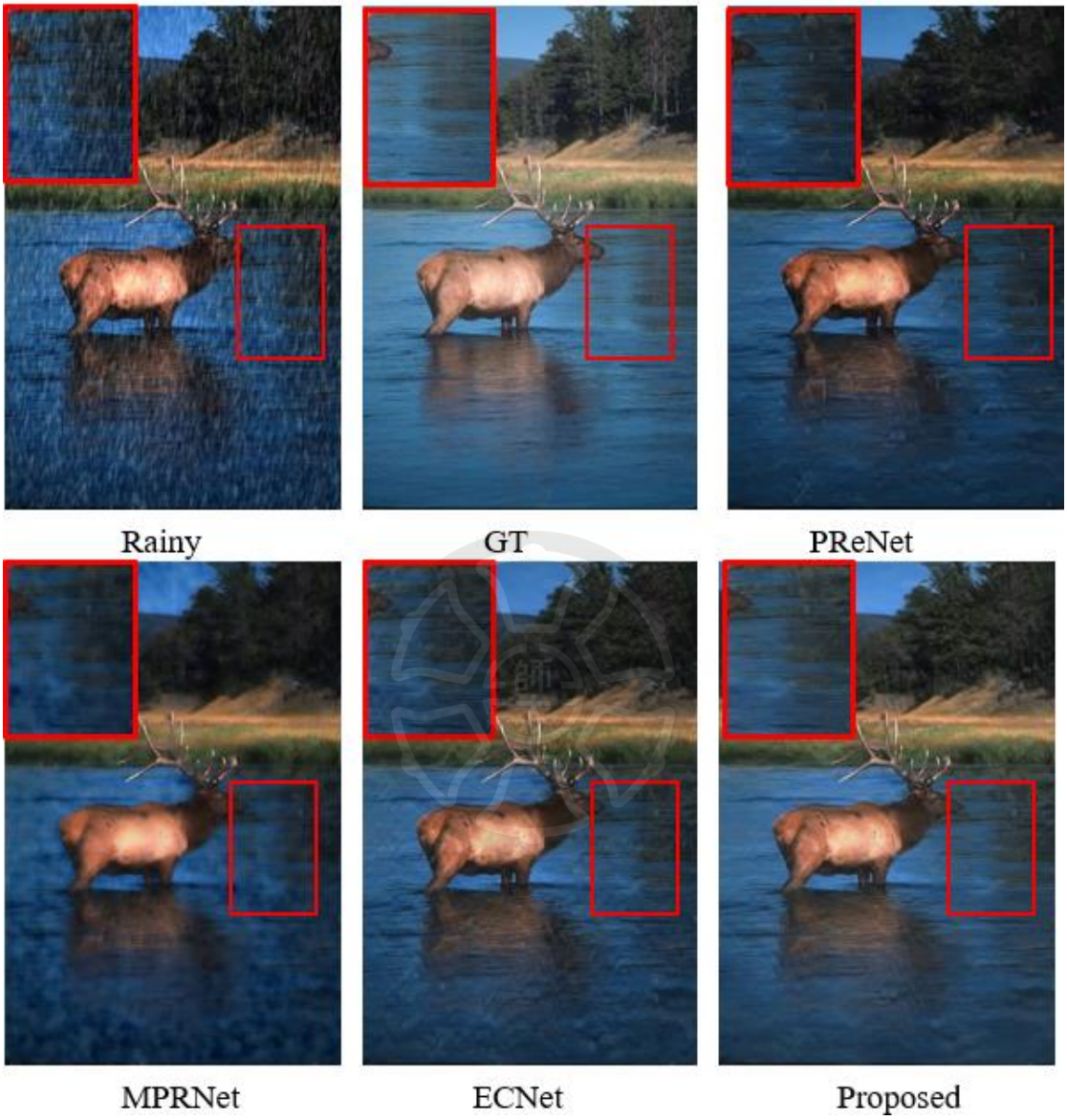


圖4-3 Rain800 資料測試集結果

從測試結果來看，我們的影像顏色比較接近 ground-truth，且去雨的效果也比較好。MPRNet 的影像過於模糊，ECNet 影像比 PreNet 效果來的好一些。我們也計算在兩個指標(PSNR/SSIM)的表現，其結果如表 4-3。與 PReNet 相比，PSNR 提升了 1.45dB，SSIM 提升了 0.006。與 MPRNet 相

比，PSNR 提升 2.86dB，SSIM 提升了 0.093。與 ECNet 相比，PSNR 提升了 1.11dB，SSIM 提升了 0.018。

表4-3　Rain800 去雨重建後結果(PSNR/SSIM)

Method	PSNR(dB)	SSIM
PReNet	26.33	0.891
MPRNet	23.84	0.804
ECNet	26.67	0.879
Proposed	**27.78**	**0.897**

我們的方法在 Rain100H 資料集中的結果，PSNR 和 SSIM 分別取得了不錯的性能，如表 4-4 與圖 4-4 所示。我們的影像看起來比較柔順，且髮絲也比較明顯。ECNet 的頭髮中段會有些模糊，而 PreNet 在頭髮中段部分有比較好，MPRNet 還是過於模糊。與 PReNet 相比，PSNR 提升了 0.04dB，SSIM 提升了 0.004。與 MPRNet 相比，PSNR 提升了 2.52dB，SSIM 提升 0.068。與 ECNet 相比，PSNR 提升了 0.58dB，SSIM 提升 0.008。

表4-4　Rain100H 去雨重建後結果(PSNR/SSIM)

Method	PSNR(dB)	SSIM
PReNet	30.07	0.904
MPRNet	27.59	0.834
ECNet	29.53	0.900
Proposed	**30.11**	**0.908**

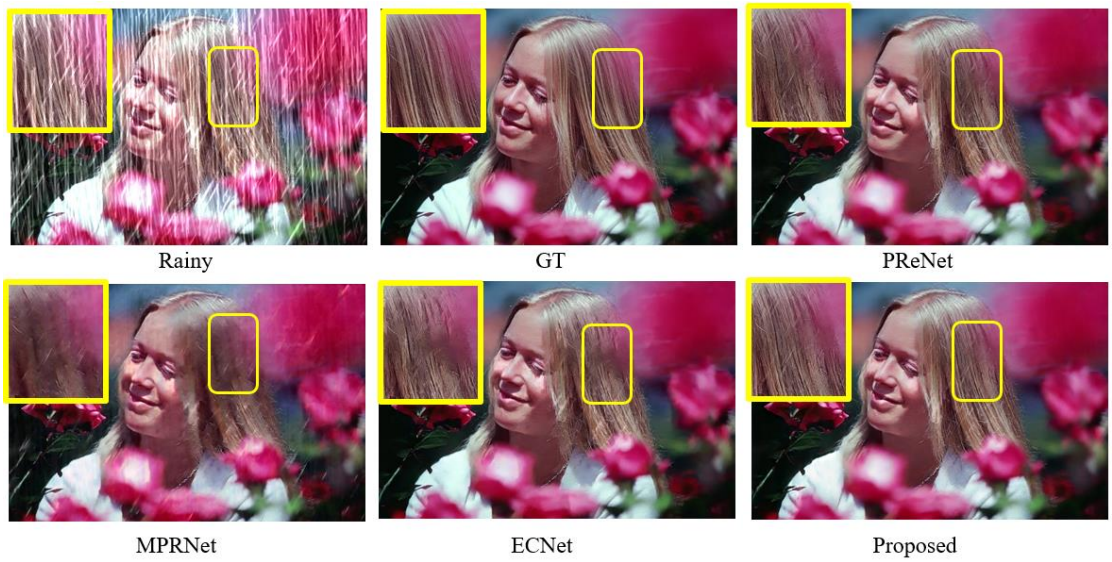


圖4-4 Rain100H 資料測試集結果

我們的方法在 Rain100L 資料集中，如表 4-5 與圖 4-5 所示。整體來說，PSNR 和 SSIM 與 MPRNet 差異較大，但在影像的效果上差異不大。與 PReNet 相比，PSNR 提升了 0.04dB，SSIM 則是持平。與 MPRNet 相比，PSNR 提升 1.24dB，SSIM 提升 0.006。與 ECNet 相比，PSNR 提升 0.3dB，SSIM 提升 0.001。

表4-5 Rain100L 去雨重建後結果(PSNR/SSIM)

Method	PSNR(dB)	SSIM
PReNet	37.29	**0.978**
MPRNet	36.09	0.972
ECNet	37.03	0.977
Proposed	**37.33**	**0.978**

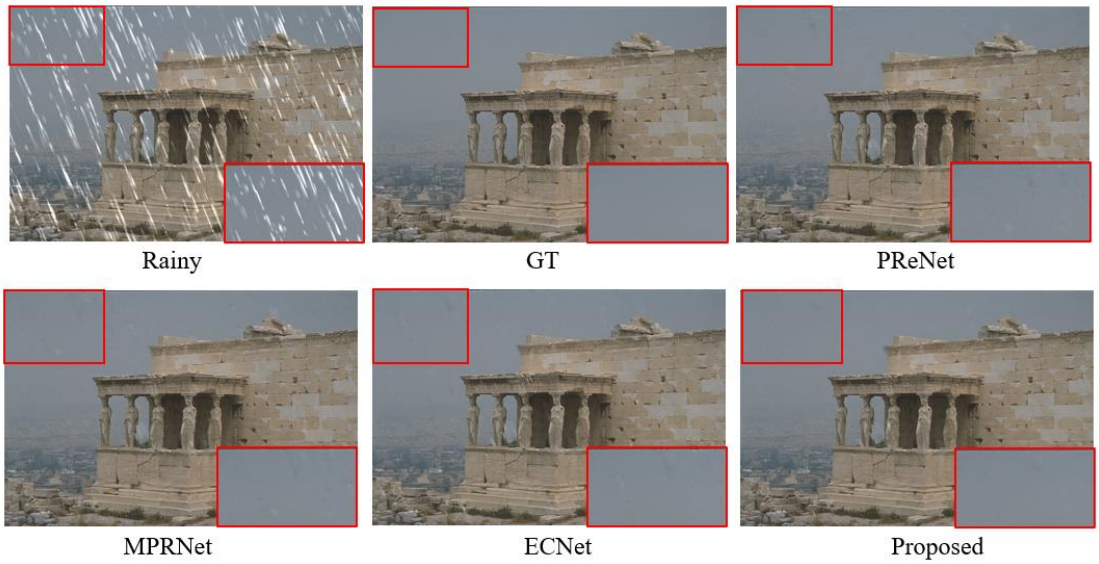


圖4-5 Rain100L 資料測試集結果

我們的方法在 Rain200H 資料集中，PSNR 和 SSIM 分別取得了不錯的性能，如表 4-6 與圖 4-6 所示。與成 PReNet 相比，PSNR 提升 0.19dB，SSIM 提升了 0. 01。與 ECNet 相比，PSNR 提升 0.75dB，SSIM 提升了 0.009。

表4-6 Rain200H 去雨重建後結果(PSNR/SSIM)

Method	PSNR(dB)	SSIM
PReNet	28.94	0.900
MPRNet	26.98	0.882
ECNet	28.38	0.892
Proposed	**29.13**	**0.901**

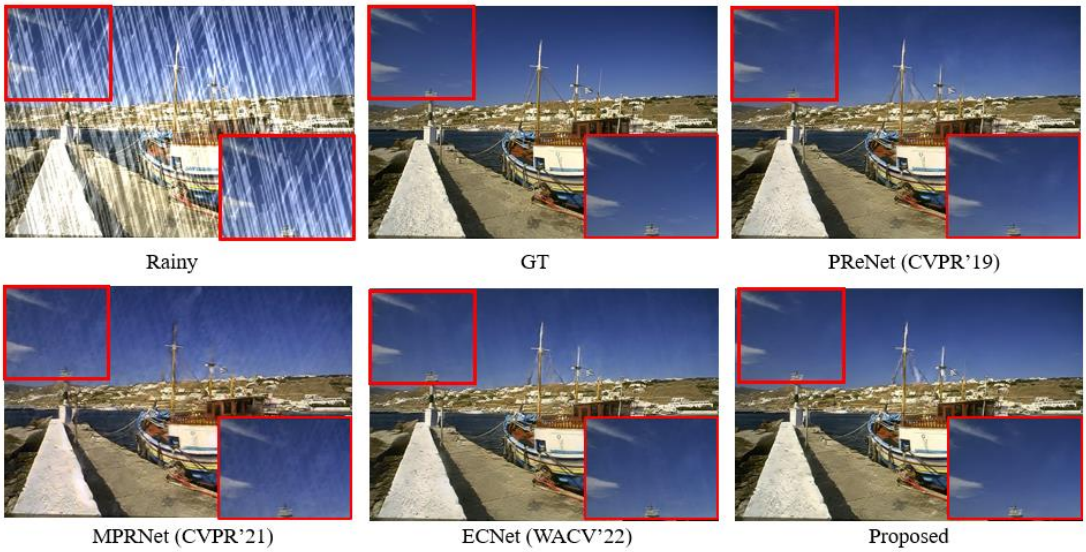


圖4-6 Rain200H 資料測試集結果

4.3 消融測試

在消融測試的部分，我們試圖消去與融合我們提出的方法來證明，透過添加的方法可以有效的提高去雨後的效果，分別將通道注意力模塊(CAB)與多尺度像素注意力模塊(MPSAM)來做消融測試。

首先，我們在 Rain100H 資料集做測試，結果如表 4-7 所示。加入 CAB，PSNR 提升了 0.12dB，SSIM 提升了 0.003。說明通道注意力模塊可以加強特徵的提取。接著加入 MSPAM，兩項指標有明顯進步，PSNR 提升了 0.56dB，SSIM 提升了 0.005。說明我們提出的 MSPAM 可以增強有用特徵的提取，減少不相關特徵的向後傳遞。

表4-7　Rain100H 消融測試結果(PSNR/SSIM)

CAB	MSPAM	PSNR(dB)	SSIM
×	×	29.53	0.900
✓	×	29.65	0.903
✓	✓	**30.11**	**0.908**

為了增加說服力，我們以相同方法，改用 Rain200H 資料集做試，結果如表 4-8。首先加入 CAB，PSNR 上升 0.09dB，SSIM 上升 0.004。提升了一定程度的效果。接著加入 MSPAM，兩項指標也有明顯進步，PSNR 上升 0.45dB，SSIM 上升 0.004。再度證明 MSPAM 可以有效的將特徵提取出來。

表4-8　Rain200H 消融測試結果(PSNR/SSIM)

CAB	MSPAM	PSNR(dB)	SSIM
×	×	28.38	0.892
✓	×	28.49	0.896
✓	✓	**28.94**	**0.900**

4.4　測試時間及模型大小

最後，我們針對所有方法，做測試時間及模型大小的比較，結果如表 4-9。測試時間採用的是 Rain100H 數據集，計算出產生一張圖片所需要的時間(s/image)。模型大小則是針對測試時需要用到的網路架構，來計算其網路架構大小。從表 4-9 可以看到 MPRNet 的模型大小最大，測試的時間也最長。PReNet 的模型雖然最小，但測試的時間卻不是最快，我們的方法也

比其方法快。而與 ECNet 相比雖然模型大小與測試時間都較大一點，但實驗結果優於其方法。

表4-9　測試時間及模型大小比較

Method	測試時間 (s/image)	模型大小(MB)
PReNet	0.08	0.169M
MPRNet	0.168	20.1M
ECNet	0.026	2.461M
Proposed	0.032	2.883M

第五章　結論與未來展望

在我們的研究中，我們提出了一種單張影像去雨架構，來有效處理去雨的議題。我們使用注意力模塊來提升特徵的提取能力及精確度，並且使特徵能有效地傳遞到解碼器，使解碼時不因無關的特徵而降低重建雨層的效果。

在合成影像去雨實驗表明，我們的方法優於其他去雨方法。在數值及影像品質上，生成視覺上更令人愉悅的影像，經過實驗我們所提出的方法可以有效的提升 PSNR 與 SSIM。

在消融測試部分，我們印證本論文中的注意力模塊的重要程度。雖然各個模塊影響程度不同，但都對於去雨效果都是有所提升，因此所有注意力模塊都是缺一不可。

本論文目前只應用於影像的去雨上，希望未來有機會透過改善網路模型架構來實現即時的影像去雨，用來應用在其他諸如影像超解析或是影像去模糊等影像處理上的問題，來提升我們方法在應用上的多元性。

參 考 文 獻

[1] L. W. Kang, C. W. Lin, and Y. H. Fu. Automatic singleimage-based rain streaks removal via image decomposition. IEEE Transactions on Image Processing, 21(4):1742–1755, 2012..

[2] J. M. Fadili, J. L. Starck, J. Bobin, and Y. Moudden, “Image decomposition and separation using sparse representations: an overview,” Proc. IEEE, vol. 98, no. 6, pp. 983–994, June 2010.

[3] J. M. Fadili, J. L. Starck, M. Elad, and D. L. Donoho, “MCALab: reproducible research in signal and image decomposition and inpainting,” IEEE Computing in Science & Engineering, vol. 12, no. 1, pp. 44–63, 2010.

[4] J. Bobin, J. L. Starck, J. M. Fadili, Y. Moudden, and D. L. Donoho, “Morphological component analysis: an adaptive thresholding strategy,” IEEE Trans. Image Process., vol. 16, no. 11, pp. 2675–2681, Nov. 2007.

[5] G. Peyré, J. Fadili, and J. L. Starck, “Learning adapted dictionaries for geometry and texture separation,” in Proc. SPIE, vol. 6701, 2007.

[6] J. L. Starck, M. Elad, and D. L. Donoho, “Image decomposition via the combination of sparse representations and a variational approach,” IEEE Trans. Image Process., vol. 14, no. 10, pp. 1570–1582, Oct. 2005.

[7] Y. L. Chen and C. T. Hsu. A generalized low-rank appearance model for spatio-temporally correlated rain streaks. In Proceedings of the IEEE International Conference on Computer Vision, pages 1968–1975, 2013.

[8] L. Zhu, C.-W. Fu, D. Lischinski, and P.-A. Heng. Joint bilayer optimization for single-image rain streak removal. In Proceedings of the IEEE International Conference on Computer Vision, pages 2526–2534, 2017.

[9] W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo, and S. Yan. Deep joint rain detection and removal from a single image. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1357–1366, 2017.

[10] X. Fu, J. Huang, X. Ding, Y. Liao, and J. Paisley. Clearing the skies: A deep network architecture for single-image rain removal. IEEE Transactions on Image Processing, 26(6):2944–2956, 2017.

[11] X. Li, J. Wu, Z. Lin, H. Liu, and H. Zha. Recurrent squeezeand-excitation context aggregation net for single image deraining. In European Conference on Computer Vision, pages 262–277, 2018.

[12] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition.

In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 770–778, 2016.

[13] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations, May 2015.

[14] X. Fu, J. Huang, D. Zeng, Y. Huang, X. Ding, and J. Paisley. Removing rain from single images via a deep detail network. In CVPR, 2017..

[15] D. Ren, W. Zuo, Q. Hu, P. Zhu, and D. Meng. Progressive image deraining networks: a better and simpler baseline. In IEEE Conference on CVPR, pages 3937–3946, 2019.

[16] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need. In NeurIPS, 2017.

[17] I. Sutskever, O. Vinyals, and Q. V. Le. Sequence to sequence learning with neural networks. In Advances in neural information processing systems, pages 3104–3112, 2014.

[18] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In ICLR, 2021.

[19] H. Chen, Y. Wang, T. Guo, C. Xu, Y. Deng, Z. Liu, S. Ma, C. Xu, C. Xu, and W. Gao. Pre-trained image processing transformer. In IEEE Conference on Computer Vision and Pattern Recognition, pages 12299–12310, 2021..

[20] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M. H. Yang. Restormer: Efficient transformer for high-resolution image restoration. In CVPR, 2022.

[21] Z. Wang, X. Cun, J. Bao, and J. Liu. Uformer: A general u-shaped transformer for image restoration. arXiv preprint arXiv:2106.03106, 2021.

[22] O. Ronneberger, P.Fischer, and T. Brox. U-Net: Convolutional networks for biomedical image segmentation. In MICCAI, 2015.

[23] K. Jiang, Z. Wang, P. Yi, B. Huang, Y. Luo, J. Ma, and J. Jiang. Multi-scale progressive fusion network for single image deraining. In CVPR, 2020.

[24] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, M.-H. Yang, and L. Shao. Multi-stage progressive image restoration. In IEEE Conf. Comput. Vis. Pattern Recog., pages 14821–14831, 2021.

[25] Y. Li, Y. Monno, and M. Okutomi, “Single image deraining network with rain embedding consistency and layered LSTM,” in Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis., 2022, pp. 3957–3966.

[26] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu. Image super-resolution using very deep residual channel attention networks. In ECCV, 2018.

[27] T. Wang, X. Yang, K. Xu, S. Chen, Q. Zhang, and R. WH Lau. Spatial Attentive Single-Image Deraining with a High Quality Real Rain Dataset. In CVPR, 2019.

[28] Z. Wang, A. C Bovik, H. R Sheikh, and E. P Simoncelli. Image Quality Assessment:From Error Visibility to Structural Similarity. TIP, 13(4):600–612, 2004.

[29] H. Wang, Q. Xie, Q. Zhao, and D. Meng. A modeldriven deep neural network for single image rain removal. In Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pages 3103–3112, 2020.

[30] W. Yang, R. T Tan, J. Feng, J. Liu, Z. Guo, and S. Yan. Deep joint rain detection and removal from a single image. In Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pages 1357–1366, 2017.

[31] H. Zhang, V. Sindagi, and V. M Patel. Image de-raining using a conditional generative adversarial network. IEEE Trans. on Circuits and Systems for Video Technology, 30(11):3943–3956, 2019.

[32] W. Wei, D. Meng, Q. Zhao, Z. Xu, and Y. Wu. Semi-supervised transfer learning for image rain removal. In Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pages 3877–3886, 2019.

[33] D. Ren, W. Shang, P. Zhu, Q. Hu, D. Meng, and W. Zuo. Single image deraining using bilateral recurrent network. IEEE Trans. on Image Processing, 29:6852–6863, 2020.

[34] G. Li, X. He, W. Zhang, H. Chang, L. Dong, and L. Lin. Non-locally enhanced encoder-decoder network for single image de-raining. In ACM International Conference on Multimedia, pages 1056–1064, 2018.

[35] L.Chen, X.Lu, J.Zhang, X.Chu, and C.Chen. Hinet: Half instance normalization network for image restoration. In CVPRW, pages 182–192, 2021.

自　　傳

我是黃冠樺，出生在彰化縣的小家庭，家裡有父母親以及一個妹妹。平常喜歡打籃球與健身，來維持一定的體態，並也藉此磨練心智。從小最喜歡的科目就是數學，因此高中時選擇了二類組。加上高三下的物理課有許多電的課程，使我對電的部分很感興趣，因此在大學時選擇了就讀國立彰化師範大學電機工程系。而在大學期間除了修習必修的工程數學、電子學、電路學、電磁學等，剛好在學習程式的微處理機與程式設計等課程上讓我十分感興趣。因此決定在就讀研究所時選擇跟程式相關的組別。

後來研究所考上國立臺灣師範大學電機研究所，在選擇研究上就以影像處理的教授為主，有幸加入了康立威博士的多媒體實驗室。在研究所期間除了學習到如何透過傳統影像處理的方式來提升影像品質，也修習了其他領域的課程，以充實自己各個方面的能力。

在研究領域方面，以深度學習及影像處理為主。我的論文題目是使用 Pytorch 的開發環境來研究編碼器解碼器的架構以去除單張影像上的雨水，並使用各種注意力模塊，來完成影像品質回復的任務。在結果方面，能夠優於近幾年的方法，且在視覺上能有更好的效果。期許自己未來也能繼續在深度學習及影像處理的領域發展，並能夠使用研究所所學為這個領域做出貢獻。

學 術 成 就

1. K.-H. Huang and L.-W. Kang, “Learning Single Image Rain Streak Removal Based on Deep Attention Mechanism,” Asisa Pacific Signal and Information Processing Association Annual Summit and Conference, Taipei, Taiwan, October 2023.