1. 隊名及隊員
隊名：NTU_b05901005_QQ
隊員：b05901005賴沂謙
b05901006高珮瑋
b05901068蕭如芸
b05901158林緯瑋

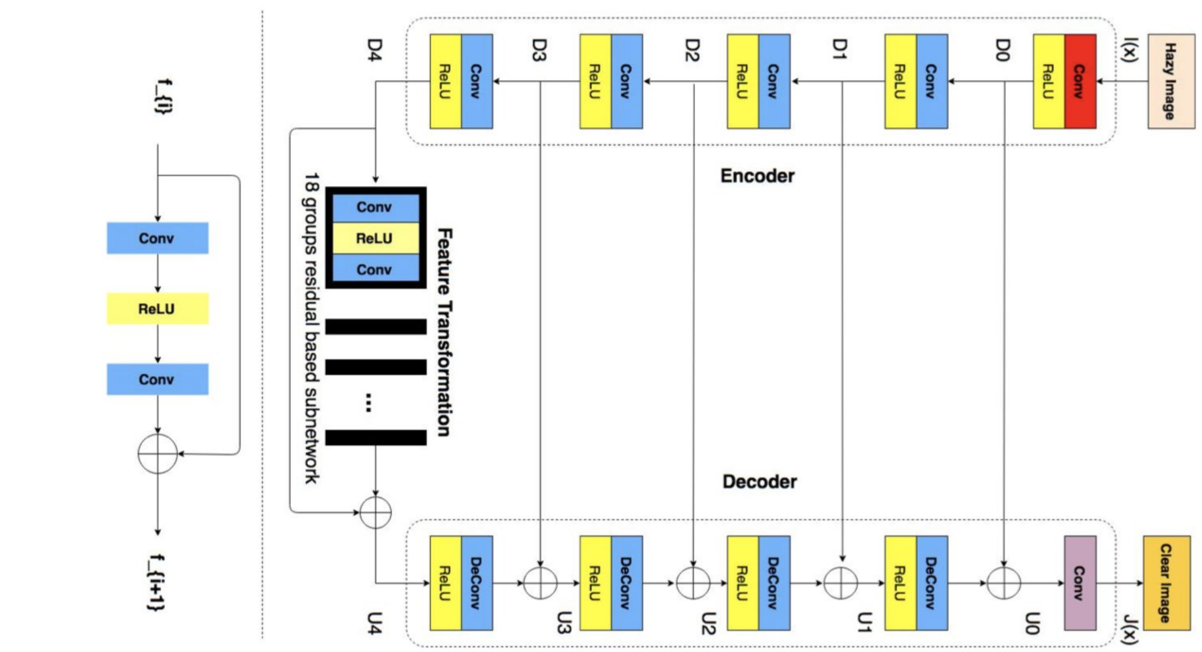
2. 所選擇的題目: Image Dehazing

3. Problem study: 與此次作業相關的 paper 閱讀，並寫出 paper 所提出的方法。（model架構、資料處理或是 training tips......相關的 paper 皆可) 最後面請務必附上 reference。

主要參考Paper: Progressive Feature Fusion Network for Realistic Image Dehazing (PFFNet)

特點：Performs feature fusion on spatial pyramid mappings between encoder and decoder, which enables maximally <u>preserved structural details from inputs for deconvolution layers</u>, and further makes the dehazing network more input-adaptive.

模型架構：



references:

(1) Image Dehazing by Joint Estimation of Transmittance and Airlight using Bi-Directional Consistency Loss Minimized FCN
(2) Progressive Feature Fusion Network for Realistic Image Dehazing (PFFNet)
(3) Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing
(4) Multi-scale Single Image Dehazing using Perceptual Pyramid Deep Network

4. Proposed method: 簡要的寫出現在的模型架構及訓練方法，再提出之後預計使用的改進方法。

(1)目前做法：

(a)模型架構

使用*Progressive Feature Fusion Network for Realistic Image Dehazing*提出的PFFNet，架構如上題所示。

詳細架構如下：

[Encoder]

```
(sub_mean): MeanShift(3, 3, kernel_size=(1, 1), stride=(1, 1))
(add_mean): MeanShift(3, 3, kernel_size=(1, 1), stride=(1, 1))
(conv_input): ConvLayer(
  (reflection_pad): ReflectionPad2d((5, 5, 5, 5))
  (conv2d): Conv2d(3, 16, kernel_size=(11, 11), stride=(1, 1))
)
(conv2x): ConvLayer(
  (reflection_pad): ReflectionPad2d((1, 1, 1, 1))
  (conv2d): Conv2d(16, 32, kernel_size=(3, 3), stride=(2, 2))
)
(conv4x): ConvLayer(
  (reflection_pad): ReflectionPad2d((1, 1, 1, 1))
  (conv2d): Conv2d(32, 64, kernel_size=(3, 3), stride=(2, 2))
)
(conv8x): ConvLayer(
  (reflection_pad): ReflectionPad2d((1, 1, 1, 1))
  (conv2d): Conv2d(64, 128, kernel_size=(3, 3), stride=(2, 2))
)
(conv16x): ConvLayer(
  (reflection_pad): ReflectionPad2d((1, 1, 1, 1))
  (conv2d): Conv2d(128, 256, kernel_size=(3, 3), stride=(2, 2))
)
```

[Residual Block] (共疊18個)

```
(res1): ResidualBlock(
  (conv1): ConvLayer(
    (reflection_pad): ReflectionPad2d((1, 1, 1, 1))
    (conv2d): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1))
  )
  (conv2): ConvLayer(
    (reflection_pad): ReflectionPad2d((1, 1, 1, 1))
    (conv2d): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1))
  )
  (relu): PReLU(num_parameters=1)
)
```

[Decoder]

```
(convd16x): UpsampleConvLayer(
  (reflection_pad): ReflectionPad2d((1, 1, 1, 1))
  (conv2d): ConvTranspose2d(256, 128, kernel_size=(3, 3), stride=(2, 2))
)
(convd8x): UpsampleConvLayer(
  (reflection_pad): ReflectionPad2d((1, 1, 1, 1))
  (conv2d): ConvTranspose2d(128, 64, kernel_size=(3, 3), stride=(2, 2))
)
(convd4x): UpsampleConvLayer(
  (reflection_pad): ReflectionPad2d((1, 1, 1, 1))
  (conv2d): ConvTranspose2d(64, 32, kernel_size=(3, 3), stride=(2, 2))
)
(convd2x): UpsampleConvLayer(
  (reflection_pad): ReflectionPad2d((1, 1, 1, 1))
  (conv2d): ConvTranspose2d(32, 16, kernel_size=(3, 3), stride=(2, 2))
)
(conv_output): ConvLayer(
  (reflection_pad): ReflectionPad2d((1, 1, 1, 1))
  (conv2d): Conv2d(16, 3, kernel_size=(3, 3), stride=(1, 1))
)
(relu): LeakyReLU(negative_slope=0.2)
```

(b) Data augmentation

對Indoor/Outdoor的Hazy/GT圖片做相同的data augmentaion，將每張圖片crop出多張512*512的圖片。從左上方開始，由左至右，由上而下，每次平移256，一張圖片約可以crop出170張圖片。Indoor的training data（不含validation）原本只有25張，經過data augmentaion後，約有4200張；Outdoor的training data（不含validation）原本只有35張，經過data augmentaion後，約有4900張。

(c) 訓練方法

Indoor和Outdoor的圖片分開訓練，使用相同的模型架構和訓練方法。訓練參數為batch size = 16，number of epochs = 300，optimizer為Adam(lr = 1e-4)，loss function為生成的圖片和ground truth之間的mean square error。

(d) 目前結果

目前一共有三個模型，分別是依照上述模型架構和訓練方法得到的indoor和outdoor model，以及github上的pretrained model。我們訓練的model和pretrained model的差別是，我們疊了18個Residual Blocks，pretrained model則是疊13個。嘗試不同組合得到的結果如下：

	使用模型		PSNR x SSIM
	Indoor	Outdoor	
1	pretrained model	pretrained model	16.802760
2	our indoor model	our outdoor model	16.764459
3	our indoor model	pretrained model	16.737262
4	pretrained model	our outdoor model	16.830134

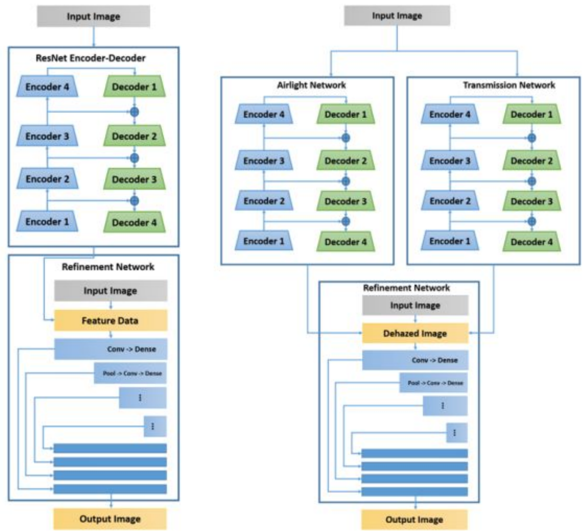
(2)後續改良作法：

希望再試試看不同論文所提出的方法,

例如：*Feature Forwarding for Efficient Single Image Dehazing*

https://arxiv.org/pdf/1904.09059.pdf

所提出的改良LinkNet的架構，將pretrained Resnet18 model所提出來的features，直接接到pyramid pooling network中去做refinement



以及**Multi-scale Single Image Dehazing using Perceptual Pyramid Deep Network**

http://openaccess.thecvf.com/content_cvpr_2018_workshops/papers/w13/Zhang_Multi-Scale_Single_Image_CVPR_2018_paper.pdf

所提出的encoder-decoder模型：利用dense block+residual block所建構的encoder，加上dense-residual block(two-layer dense block + an upsampling transition block)+pyramid pooling module所建構的decoder，得到最後的dehazed image。

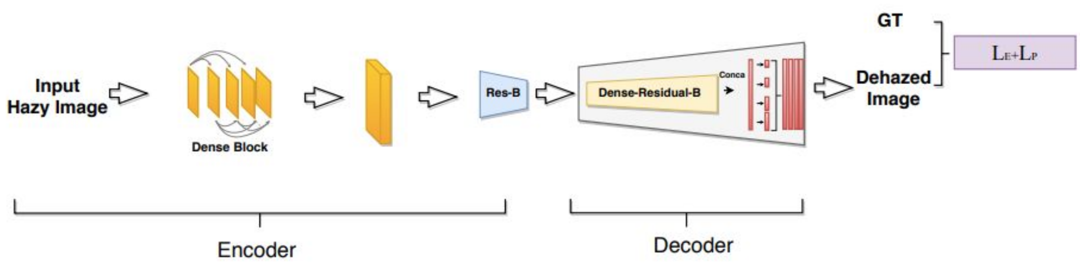


Figure 2. Overview of the proposed Multi-scale Single Image Dehazing using Perceptual Pyramid Deep Network.

之後我們會想辦法利用pretrained model所得到的features，加上例如Densely Connected Pyramid Dehazing Network的方法，得到更好的結果。