**AI CUP 2024 春季賽**

**以生成式AI**

**建構無人機於自然環境偵察時所需之導航資訊**

**競賽 II － 導航資料生成競賽**

**競賽報告**

隊伍：TEAM\_5137

隊員：吳家萱（隊長）、許中萁、吳建中。

Private leaderboard： 0.722013 / Rank 17

**壹、環境**

作業系統：Windows11。

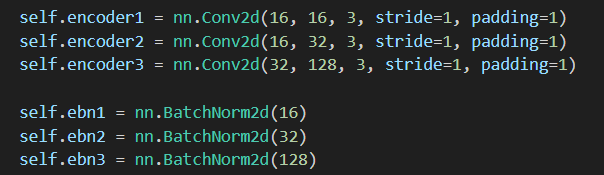
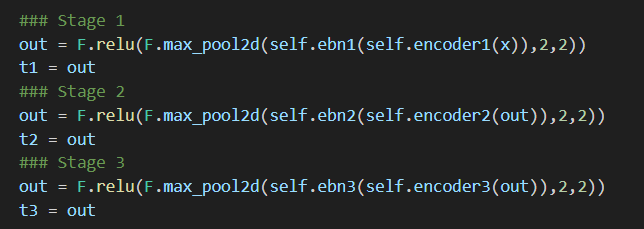
程式語言：Python。

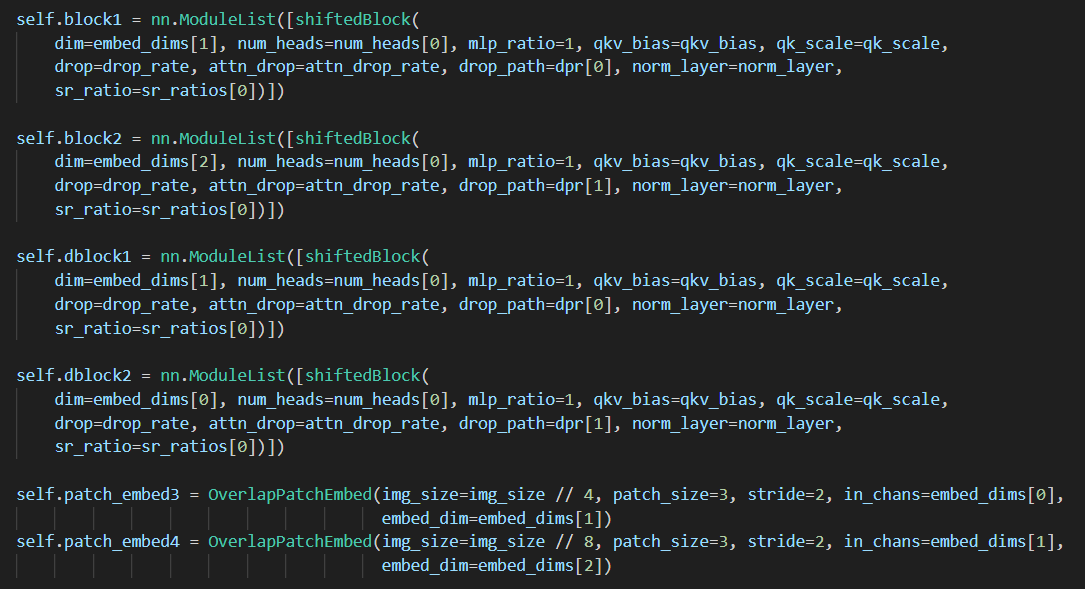
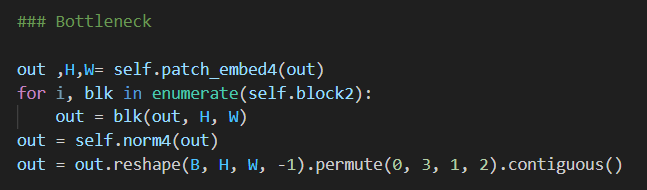
套件（函式庫）：PyTorch、numpy、cv2、os、torchvision、timm、matplotlib、math。

程式透過CUDA使用GPU訓練模型，無使用預訓練模型及額外資料集。

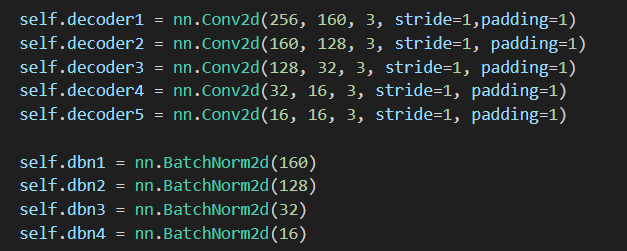
**貳、演算法與模型架構**

我們設計了一個改良U-NeXt的模型架構，該模型結構包含：Encoder、Bottleneck、Decoder、Output Layer。

1. Encoder:  
   Encoder由三個捲積層組成，結構如下:  
   encoder1: in channel=16, out channel=16, kernel size=3, stride=1, padding=1  
   encoder2: in channel=16, out channel=32, kernel size=3, stride=1, padding=1  
   encoder3: in channel=32, out channel=128, kernel size=3, stride=1, padding=1   
   每次forward都會對每一層捲積層做BatchNormalization，透過max pooling減少特徵圖尺寸，使用ReLU激活函數。  
   
2. Bottleneck  
   使用Overlap Patch Embed將圖像特徵分割成重疊的patch，通過捲積進行嵌入（從而保留圖片邊緣的資訊，提高分割的精度），然後提供給shifted MLP結合卷積操作和MLP，為了更有效地捕捉特徵。  
   Overlap Patch Embed結構如下：

  
Bottleneck寫法如下，使用上述的self.patch\_embed4來做Overlap Patch Embed：

1. Decoder:

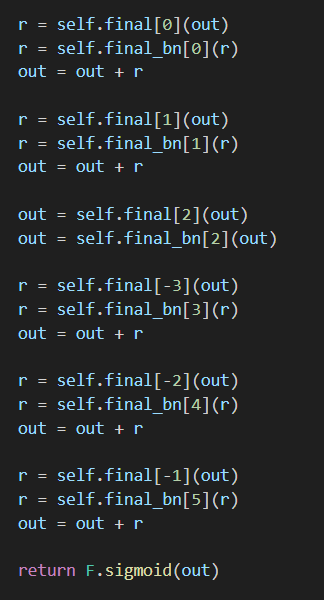
Decoder由多層卷積組成，逐層進行Up-sampling以恢復原特徵圖的尺寸，每個解碼層之間加入了Residual Connection，使得梯度能更好的反向傳播，結構如下：  


每次forward都會對每一層捲積層做BatchNormalization，透過F.interpolate做上採樣，改變數據的大小，使用的是bilinear採樣法。





1. Output Layer  
   輸出層由多層卷積組成，最後使用sigmoid函數進行激活，輸出分割後的圖像，結構如下：



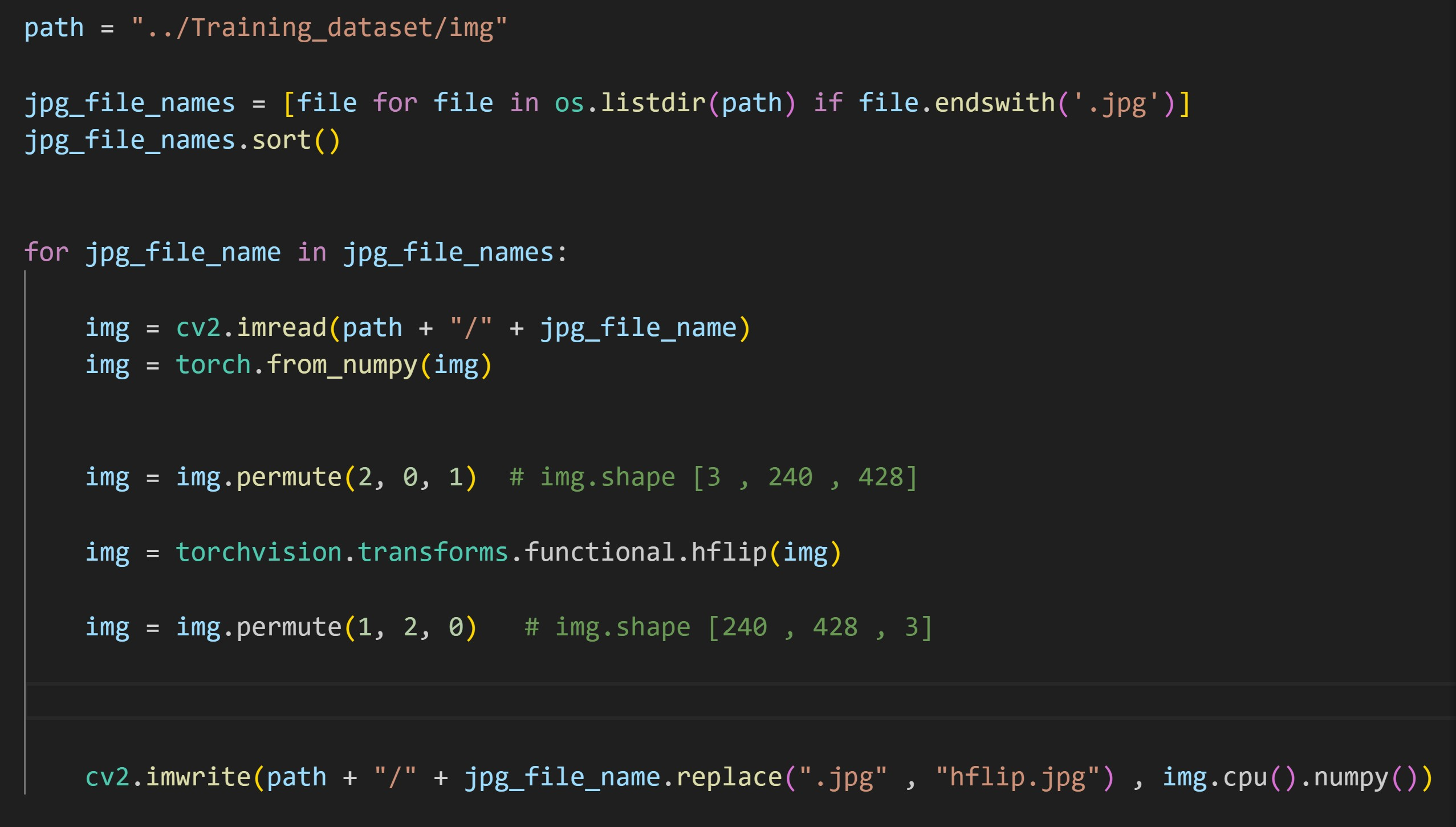
**參、技術模型原創性或改良成效**

**肆、資料分析與處理過程**

我們有對訓練資料集做水平翻轉擴增，方法如下：  
img.permute()為對tensor進行維度換位，轉換後的img.shape為[3, 240, 428]

對img做水平翻轉，使用torchvision.transforms.functional.hflip(img)

水平翻轉完再將其用img.permute()轉換為原始圖片維度，即[240, 428, 3]。



在Training dataset讀取時，轉換成numpy型態，接著用unsqueeze做擴維，



**伍、訓練方式**

**陸、結果分析與結論**

**柒、程式碼**

Github連結：<https://github.com/lavinia0724/AI-CUP-Navigational-Data-Generation-with-Generative-AI>

**捌、使用的外部資源與參考文獻**

**UNeXt: MLP-based Rapid Medical Image**

**Segmentation Network**

**報告作者聯絡資料表**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 隊伍名稱 | TEAM\_5137 | Private Leaderboard 成績 | 0.722013 | Private Leaderboard 名次 | Rank 17 |
| 身分  （隊長/隊員） | 姓名 (中英皆需填寫)  (英文寫法為名,姓，例：Xiao－Ming, Wu，名須加連字號，姓前須加逗號) | 學校＋系所中文全稱  (請填寫完整全名，勿縮寫) | 學校＋系所英文中文全稱  (請填寫完整全名，勿縮寫) | 電話 | E-mail |
| 隊長 | 吳家萱  Chia－Hsuan,  Wu | 天主教輔仁大學  資訊工程學系 | Fu Jen Catholic University Department of Computer Science and Information Engineering | 0916921480 | lavi＠nisra.net |
| 隊員1 | 許中萁  Chung－Chi,  Hsu | 天主教輔仁大學  資訊工程學系 | Fu Jen Catholic University Department of Computer Science and Information Engineering | 0966702318 | f0978473376@gmail.com |
| 隊員2 | 吳建中  Jian－Zhong,  Wu | 天主教輔仁大學  資訊工程學系 | Fu Jen Catholic University Department of Computer Science and Information Engineering | 0932101121 | andywu10523@gmail.com |
| 隊員3 |  |  |  |  |  |
| 隊員4 |  |  |  |  |  |
| 指導教授資料 | | | | | |
| 每隊伍至多可填寫兩名 | 指導教授  中文姓名 | 指導教授  英文姓名  (英文寫法為名,姓，例：Xiao－Ming, Wu，名須加連字號，姓前須加逗號) | 任職學校＋系所  中文全稱  (請填寫完整全名，勿縮寫) | 任職學校＋系所  英文全稱  (請填寫完整全名，勿縮寫) | E-mail |
| 教授 1 | 黃貞瑛 | Jen－Ing, Hwang | 天主教輔仁大學  資訊工程學系 | Fu Jen Catholic University Department of Computer Science and Information Engineering | Jihwang  @csie.fju.edu.tw |
| 教授 2 |  |  |  |  |  |