

DRCT(Dense Residual Connected Transformer)

論文ソース

- [DRCT: Saving Image Super-Resolution away from Information Bottleneck](#)

概要

- SOTAであるHATを上回る(2年ぶりに更新)
- 既存のモデルではnetworkの後半の層にいくにつれて空間情報が減少しているという問題を提起
- これを解決するためにDense residual blockを導入
 - RRDB-Net, RDNから着想を得たらしい

前提知識

Leaky ReLU

- ReLUと異なり負のときに0ではなく小さい負の値をとる

$$f(x) = \begin{cases} x, & x \geq 0 \\ \alpha x, & x \leq 0 \end{cases}$$

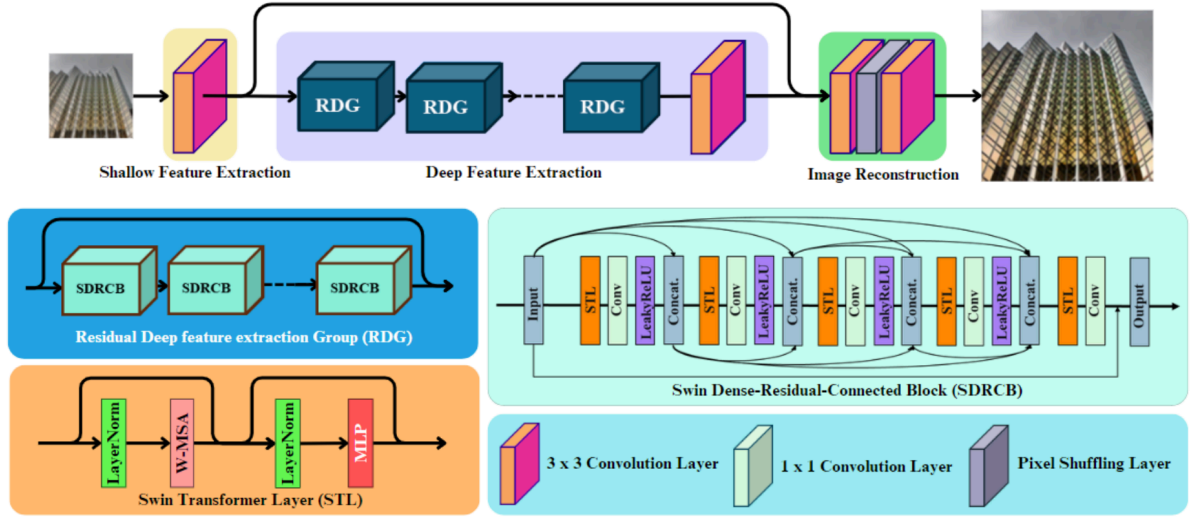
- α は0.01など

DRCT

- 構造はHATと同じく Shallow Feature Extraction, Deep Feature Extraction, Image Reconstruction の3つからなる
- Shallow Feature ExtractionとImage Reconstructionは同じ
- Deep Feature ExtractionのHABが代わりにSDRCB(Swin Dense-Residual Connected Block)になっている
 - これは(STL->Conv->LeakyReLU)を1単位として前層のすべての出力をconcatenateしている
 - これが5個ある

- 。この時、ソースコードをみると次元が一致するように調整されている

- https://github.com/ming0531/DRCT/blob/main/drct/archs/DRCT_arch.py
- 292行目以降のadjust1,adjust2など



- 。以下、数式で表現

Shallow Feature Extraction

$$F_0 = \text{Conv} (I_{LQ}) \quad I_{LQ} \in \mathbb{R}^{H \times W \times C_{in}}$$

Deep Feature Extraction

- 。全体は以下

$$F_{DF} = H_{DF} (F_0)$$

- 。これはRDGを繰り返すことで得られる

$$\begin{aligned} F_i &= \text{RDG}_i (F_{i-1}) \quad i = 1, 2, \dots, K, \\ F_{DF} &= \text{Conv} (F_K) \end{aligned} \quad (1)$$

- 。またRDGは複数のSDRCBから成る
- 。SDRCBはdense residual blockを導入した5つのSTLから成る
- 。SDRCB間の入出力を Z とすると以下のようにかける

$$\begin{aligned} Z_j &= H_{trans} (\text{STL} ([Z, \dots, Z_{j-1}])) \quad j = 1, 2, 3, 4, 5, \\ \text{SDRCB} (Z) &= \alpha \cdot Z_5 + Z \end{aligned} \quad (2)$$

- 。 H_{trans} は 1×1 Convolution と LeakyReLU

Image Reconstruction

$$I_{SR} = H_{rec}(F_0 + F_{DF})$$

- H_{rec} は function of the reconstruction

論文中の英単語メモ

- meticulous(adj) : 几帳面な
- commendable(adj) : 称賛に値する
- mitigate(v) : 軽減する
- unleash(v) : 解き放つ
- oscillation(n) : 振動、発振
- efficacy(n) : 効能
- consecutive(adj) : 連続した