You Only Need 90K Parameters to Adapt Light: a Light Weight Transformer for Image Enhancement and Exposure Correction

Github Link

Ziteng Cui¹, Kunchang Li², Lin Gu¹,², Shenghan Su², Peng Gao², Zhengkai Jiang³, Yu Qiao², Tatsuya Harada¹,³

1. The University of Tokyo 2. Shanghai AI Laboratory 3. RIKEN AIP

4. Shanghai Jiao Tong University 5. TenCent YouTu Lab

Illumination-Adaptive-Transformer

Only 90K Parameters !
0.004s inference speed per image !
State-Of-The-Art !

Several tasks:
(1) Low-light enhancement
(2) Exposure correction
(3) Low-light object detection & Low-light semantic segmentation
(4) Various-light condition object detection

 Enhancement Result

Exposure Correction Result

<table>
<thead>
<tr>
<th>Method</th>
<th>PSNR</th>
<th>SSIM</th>
<th>FLIKR</th>
<th>POCS</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIME (我们的)</td>
<td>32.8</td>
<td>0.91</td>
<td>30.5</td>
<td>0.79</td>
<td>22.0</td>
</tr>
<tr>
<td>Our Model</td>
<td>33.2</td>
<td>0.92</td>
<td>30.8</td>
<td>0.80</td>
<td>21.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>PSNR</th>
<th>SSIM</th>
<th>FLIKR</th>
<th>POCS</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIME (我们的)</td>
<td>32.8</td>
<td>0.91</td>
<td>30.5</td>
<td>0.79</td>
<td>22.0</td>
</tr>
<tr>
<td>Our Model</td>
<td>33.2</td>
<td>0.92</td>
<td>30.8</td>
<td>0.80</td>
<td>21.8</td>
</tr>
</tbody>
</table>

Model Structure:

Our local branch is adopt depth-wise convolution for light weight design, consists of two branch to learn global and local map.

Our global branch is inspired by Detection Transformer (DETR) Using attention queries to dynamic control ISP-related parameters

Paper Link

Pixel-wise Enhancement Module (PEM)

Depth-wise -Conv Transformer

Global Prediction Module (GPM)

Dynamic learnable attention queries

Several tasks:
(1) Low-light enhancement
(2) Exposure correction
(3) Low-light object detection & Low-light semantic segmentation
(4) Various-light condition object detection

Table 1: Experimental results on LR (Y1 & Y2) [104] dataset, best and second best results are marked in red and blue respectively, meanwhile [21] is a non-supervised training method and [21] is self-supervised training method.