Progressive Splitting and Upscaling Structure for Super-Resolution

Qiang Li, Tao Dai, Shu-Tao Xia
liqiang19@mails.tsinghua.edu.cn, daitao.edu@gmail.com, xiaist@sz.tsinghua.edu.cn

Problem
Most super-resolution (SR) methods focus on the design of network architecture and adopt a sub-pixel convolution layer at the end of network, but few have paid attention to exploring potential representation ability of upsampling layers.

Method
Sub-pixel convolution layer aggregates several low resolution (LR) feature maps and builds super-resolution (SR) images in a single step. However, those LR feature maps share similar patterns as they are extracted from a single trunk network. In this paper, we propose a progressive splitting and upsampling structure (PSUS) for SR task. It works with a certain basic block and aims at generating decoupled SR features progressively. It uses fewer parameters and lower computational cost, whose details are shown in the paper.

Overview
Flexible Structure We propose a progressive splitting and upsampling structure (PSUS) for image SR to explore the potential representation ability of upsampling layer.

Novel Splitting Strategy We propose a progressive splitting module (PSM) which can produce decoupled deep features using approximately the same computational cost.

Efficient Upscale Module We propose a multi-branch upsacle module (MUM) which aggregates LR features. Besides, we propose a transition strategy to further reduce computational cost and parameters.

Proposed Architecture
Our proposed PSUS consists of four parts: shallow feature extraction, global feature extraction, progressive splitting module (PSM) and multipath upscale module (MUM). We use $L_{LR}$ and $L_{SR}$ to denote input and output of network. One convolutional layer is used to extract the shallow feature $F_0$ from LR input image. Assuming that corresponding chain-like model stacks $B$ basic blocks, we use $A = (0 < \lambda < 1)$ of them to extract global feature. Further, we adopt a PSM to progressively decouple features and generate $F_{LRB}$, $F_{LRB} \cdot \cdot \cdot F_{LRB,2}$. Then, we have got $r^2$ groups of features which correspond to each position of $r \times r$ patch respectively. What MUM does is to aggregate these features and generate $C$ SR features. The architecture of PSM and MUM have been shown in the figure. More details could be found in the paper.

Experiments
We compare the performance of our method to three widely-used models: a small model EDSR-baseline from [1], a large model RCAN[2] and an unsupervised SR model ZSSR[3]. We conduct experiments on a small model EDSR-baseline firstly. It is a single-scale model and only contains 16 ResBlocks. For PSUS with ResBlock, we study the effects of different $\lambda$ for $\times 2$ model and then choose proper values of $\mu_1$, $\mu_2$ for $\times 4$ model. Quantitative metrics are presented. $\lambda$ denotes the ratio of basic blocks using for extracting global feature. For $\times 2$ model, its network architecture is determined by the single hyperparameter $\lambda$. We retrain EDSR-baseline model in our environment for fair comparison and train $\times 2$ PSUS with $\lambda \in \{0.875, 0.75, 0.5\}$. PSNR and SSIM are shown in Table I. As for $\times 4$, results are shown in Table III and Table IV.

Table I: Quantitative metrics of model complexity and computational cost for different $\times 4$ models.

Table III: Quantitative metrics of model complexity and computational cost for different $\times 4$ models.

Table IV: PSNR (dB) and SSIM results (scale $\times 4$) of baseline and our proposed PSUS. Best results are highlighted.

References

Acknowledgements
This work is supported in part by the National Natural Science Foundation of China under Grant 61771273, Guangdong Basic and Applied Basic Research Foundation on 2019A1515110344, the China Postdoctoral Science Foundation under Grant 2019M660645, the R&D Program of Shenzhen under Grant JCYJ2018050815220444, and the project “PCL Future Greater-Bay Area Network Facilities for Large-scale Experiments and Applications (LZC0019).”

Fig. 5: PSNR on validation set of $\times 2$ models during first $2 \times 10^3$ iterations of training.

Fig. 7: Visual comparison for $\times 4$ SR. Best results are highlighted.