Abstract

Most existing methos recover high-resolution representations from low-resolution representations, out proposed network maintains high-resolution representations through the whole process.

1. start from a high-resolution subnetwork
2. gradually add high-to-low resolutio subnetworks to form more stages
3. connect the multi-resolution sbnetworks in parallel.
4. conduct repreated multi-scale fusions such that each of the high-to-low resolution representations receives information from other parallel representations over and over, leading to rich highresolution representations

1. Introduction

High-Resolution Net,Two benifits:

1. connects subnetowrks in parallel than in series.（所以能保持高分辨率的信息，而不需要从低到高的恢复过程）
2. repeated multiscale fusions with the help of the low-resolution representations of the same depth and similar level, resulting in that high-resolution respresentations are also rich

2. Related Work

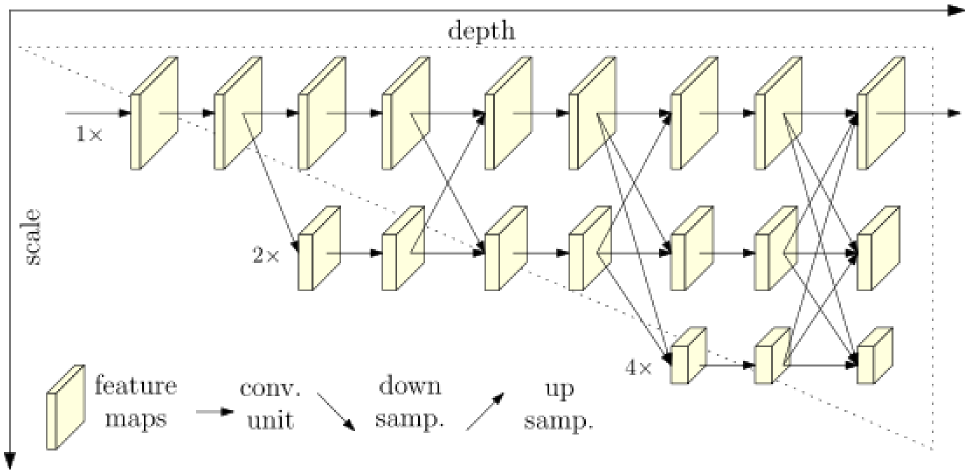
1. Representative network design patterns:
   - Symmetric high-to-low and low-to-high processes(Hourglass)
   - Heavy high-to-low and light low-to-high(Cascaded pyramid networks, SimpleBaseline)
   - Combination with dilated convolutions.
2. Multi-scale fusion
3. INtermediate supervision
4. Our approach:
   - maintain high-resolution representations through the whole process
   - generate reliable high-resolution representations through repeatedly fusing the representations produced by the high-to-low subnetworks.
   - **without** using intermediate heatmap supervision

3. Approach

to detect the locations of K keypoints or parts, generally transform the problem to estimating K heatmaps of size WxH,$\{H_1, H_2, \ldots, H_k\}$, where each heatmap H_k indicate the feature map of the kth keypoint.

3.1 Structure

1. a stem consisting of two strided convolutions decreasing the resolution
2. a main body outputing the feature maps with the same resolution as its input feature maps
3. A regressor estimating the heatmaps
这篇论文重点是将HRNet引入第二步中,模型结构如下：



3.2 Repeated multi-scale fusion

如上图，在第二个stage， 有两个并行的branch， 在stage的最后会相互融合，然后生成两个和之前channels数、shape一样的feature map, 作为下一层的输入。 *值得注意的是，若下一层的branch不同于上一层的branch时，参考代码中的实现，新增的branch用上一层的最后一个branch生成。*

3.3 Heatmap estimation

3.4 Network instantiation

contains four stages with four parallel subnetworks, The first stage contains 4 residual units where each unit, the same to the ResNet-50, is formed by a bottleneck with the width 64, and is followed by one 3×3 convolution reducing the width of feature maps to C.
The 2nd, 3rd, 4th stages contain 1, 4, 3 exchange blocks, respectively. One exchange block contains 4 residual units where each unit contains two 3 × 3 convolutions in each resolution and an exchange unit across resolutions.