

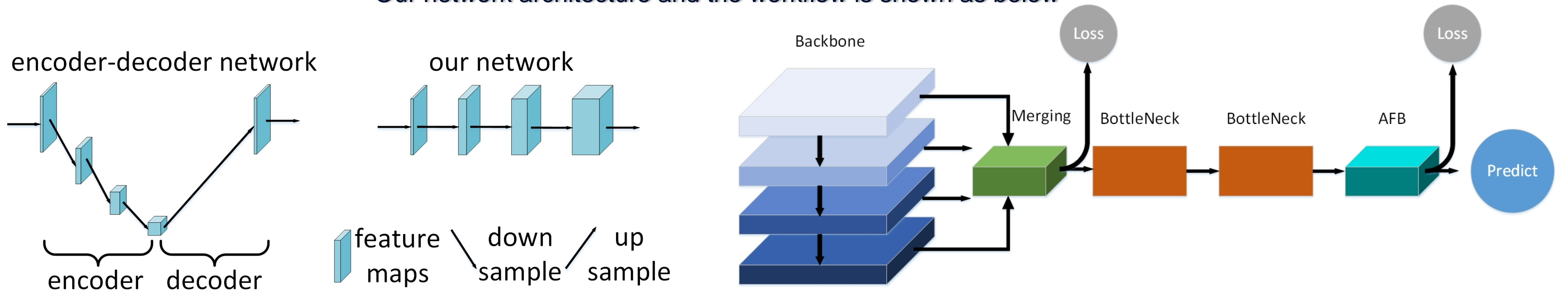
# The Same Size Dilated Attention Network for Keypoint Detection

Yuan Chang, Zixuan Huang et al.

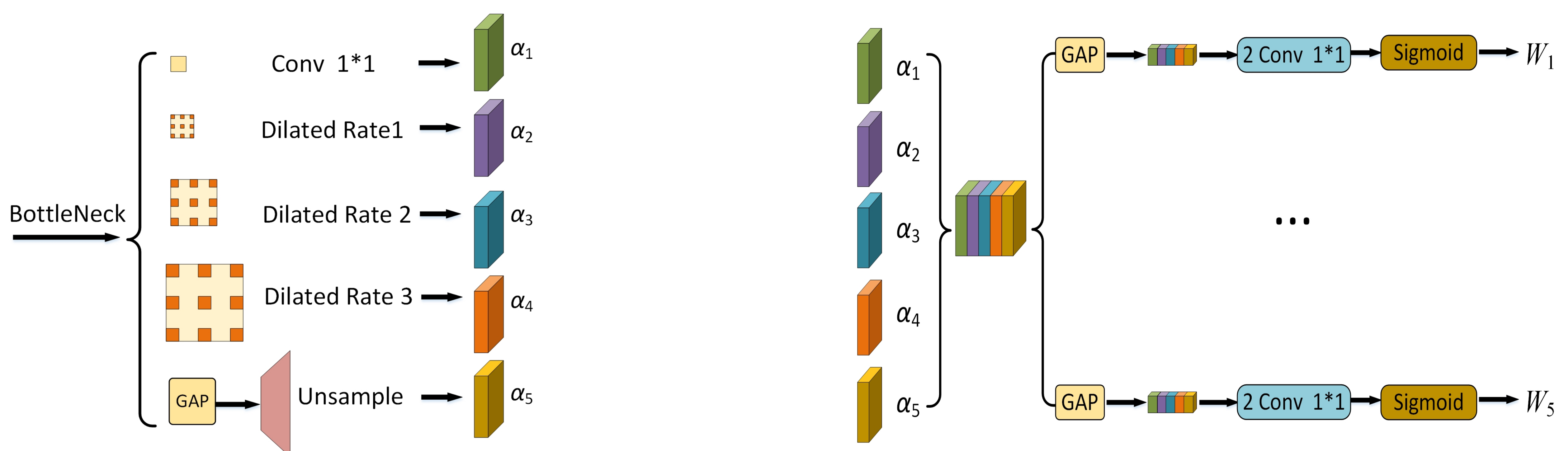
changyuan@bupt.edu.cn  
hangzixuan0508@bupt.edu.cn



Our network architecture and the workflow is shown as below



The AFB module is shown as below



$$W_i = \sigma(g_i(\text{GAP}(\text{Concat}(\alpha_{1..5}))))$$

$$\text{Output} = \sum_{i=1}^5 W_i \cdot \alpha_i$$

The main contributions of this work are twofold.

- We design a new network which is different from the traditional encoder decoder architecture. It eliminates the upsampling operation of the network and avoids the information loss caused by bilinear interpolation.
- We propose a new feature fusion block named AFB which first resamples the input at multiple scales and then uses the attention mechanism to allow network to learn the weights of the components to be fused.

COCO2017 Val

Models	Backbone	Inputsize	AP
8-stage hourglass	—	256*256	67.1
CPN	ResNet50	384*288	71.6
Simplebaseline	ResNet50	384*288	72.2
Ours	ResNet50	384*288	73.4
Ours*	ResNet50	384*288	75.3