Appendix

A. Rep-Net incorporates with different network architectures

In the experiments, we incorporate Rep-Net with four popular network architectures (backbone models), including ResNet50 [2], MobileNetv2 [4] and MobileNetv3 [3] and proxylessNAS-Mobile [1]. Rep-Net consists of 6 modules for all the architectures. To reduce the activation memory consumption, we apply a 4×4 average pooling layer at the beginning of the Rep-Net to downsample the input. In addition, we don't add the module to connect with the first layer of the backbone model to capture general low-level features, but add a module to the last layer before classifier to improve the transfer capacity by enriching the high-level features. More specifically, following the ***Strategy*** mentioned in the Section 3.3, 1) for the ResNet50, we add the modules to connect with the first Bottleneck block of each stage; 2) for the MobileNetV2, we add the modules to connect with the $2^{ed}, 4^{th}, 7^{th}, 11^{th}, 14^{th}, 17^{th}$ InvertedResidual Block; 3) for the MobileNetV3, we add the modules to connect with the $2^{ed}, 4^{th}, 7^{th}, 11^{th}, 13^{th}, 15^{th}$ InvertedResidual Block; 4) for the proxylessNAS-Mobile, we add the modules to connect with the $1^{st}, 5^{th}, 9^{th}, 13^{th}, 17^{th}, 21^{st}$ InvertedResidual Block.