BKDSNN: Enhancing the Performance of Learning-based Spiking Neural Networks Training with Blurred Knowledge Distillation

Zekai Xu[1], Kang You[1], Qinghai Guo[2], Xiang Wang[2], Zhezhi He[1]

[1]Shanghai Jiao Tong University, Shanghai, China, [2]Huawei Technologies, Shenzhen, China

Problems

ANN Propagation   SNN Propagation

ANN and SNN Propagation.

Method	Learning-based	Conversion-based
Description	Replace activation function with spike-based neuron and learning from scratch.	Convert pretrained ANN to SNN with mathematically equivalent spike-based neuron
Lower Time-Step	✓	✗
Higher Accuracy	✗	✓

Motivation

Feature distribution of SNN and ANN

Pretrained ANN → serve as teacher model → Learning-based SNN
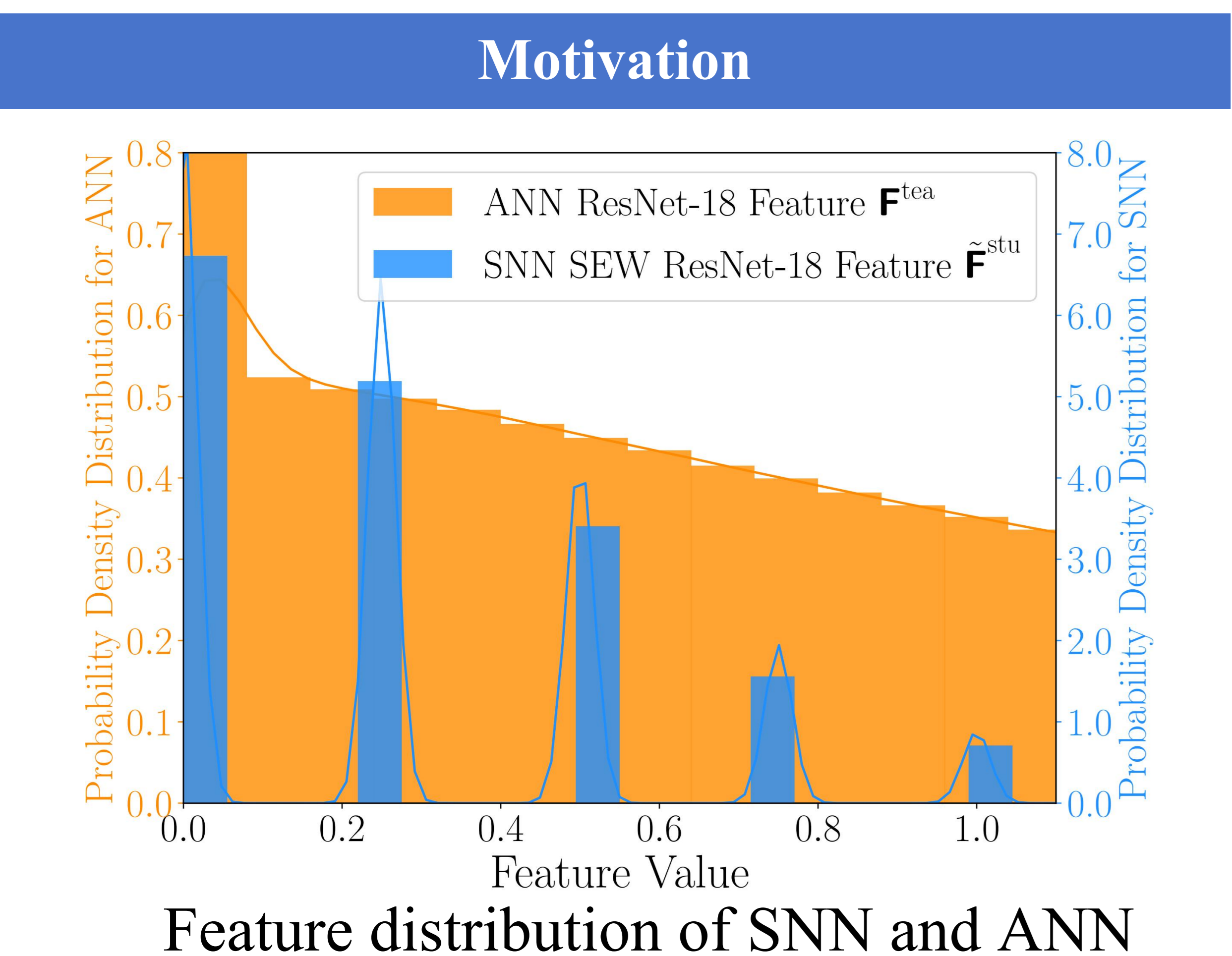
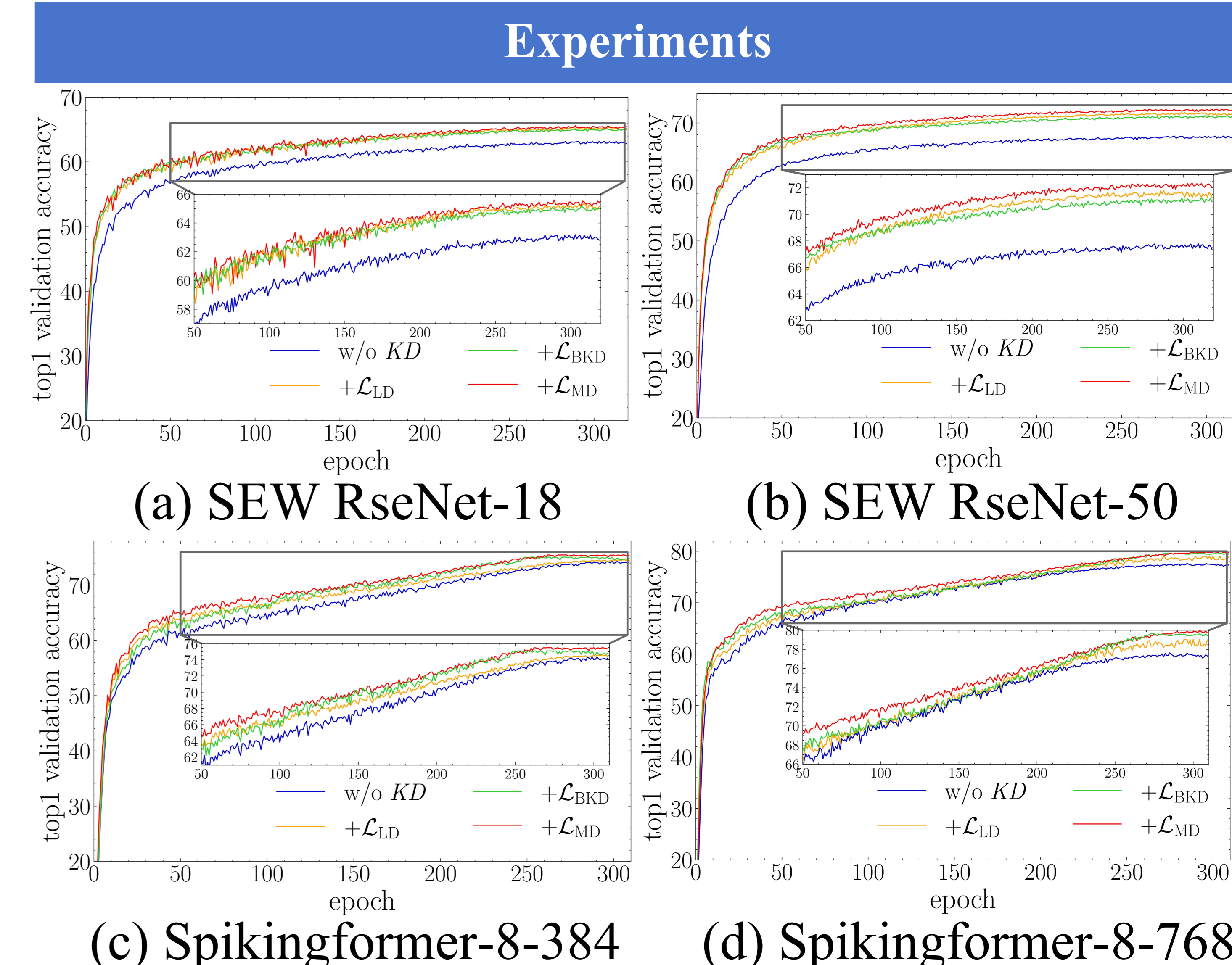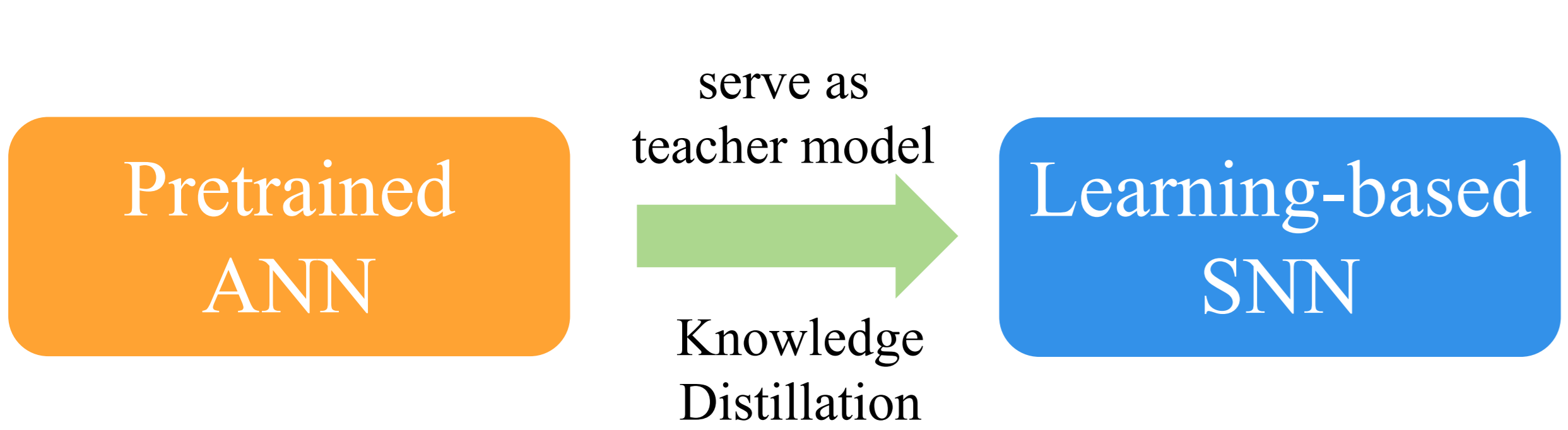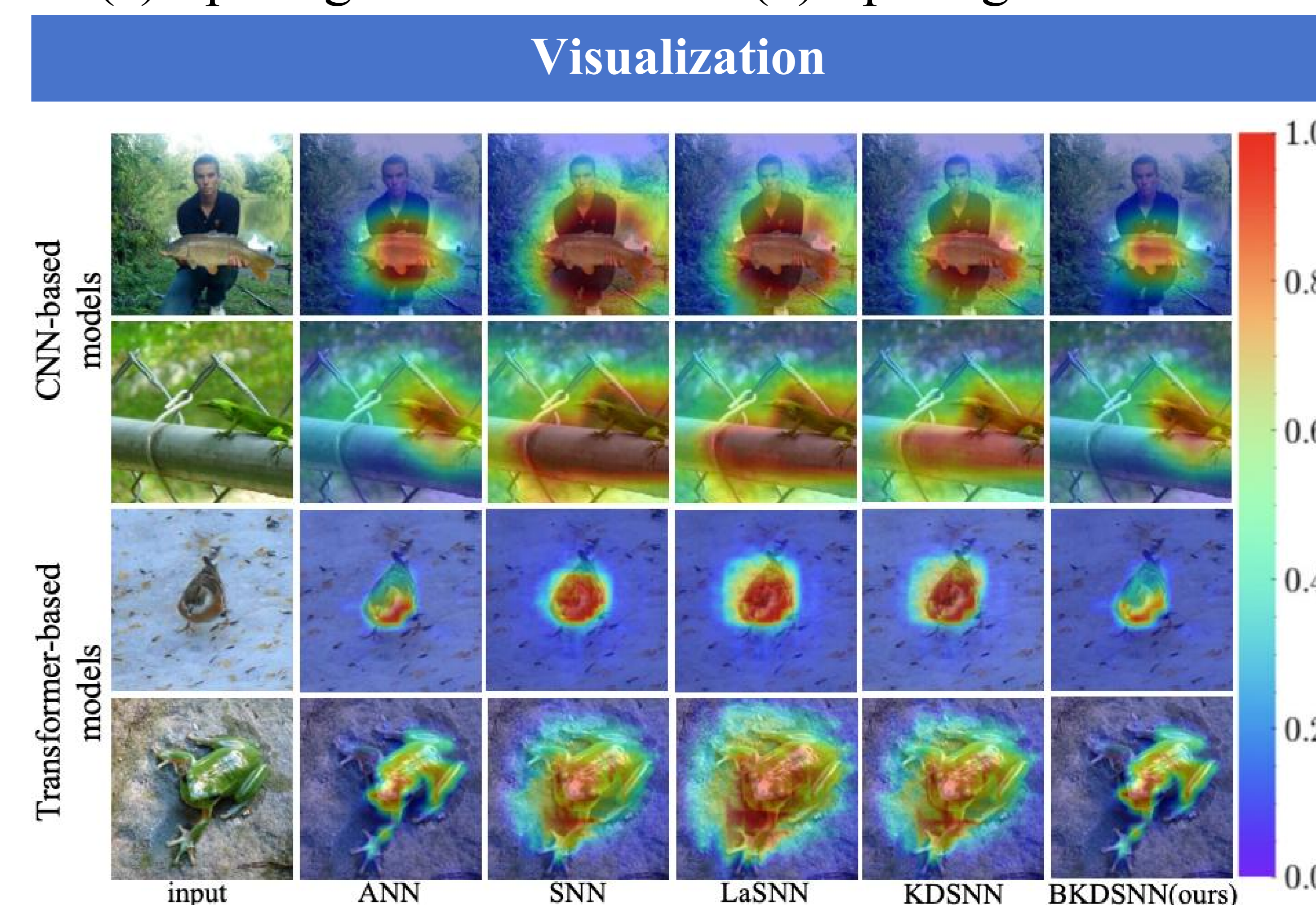Knowledge Distillation

Architecture

Overview of training SNN with blurred knowledge distillation (BKD). The SNN is directly trained with back-propagation through time (BPTT), where we utilize a blurred variant of KD to achieve higher accuracy. BKD block is highlighted in yellow shaded region.

Blurred Knowledge Distillation (BKD)
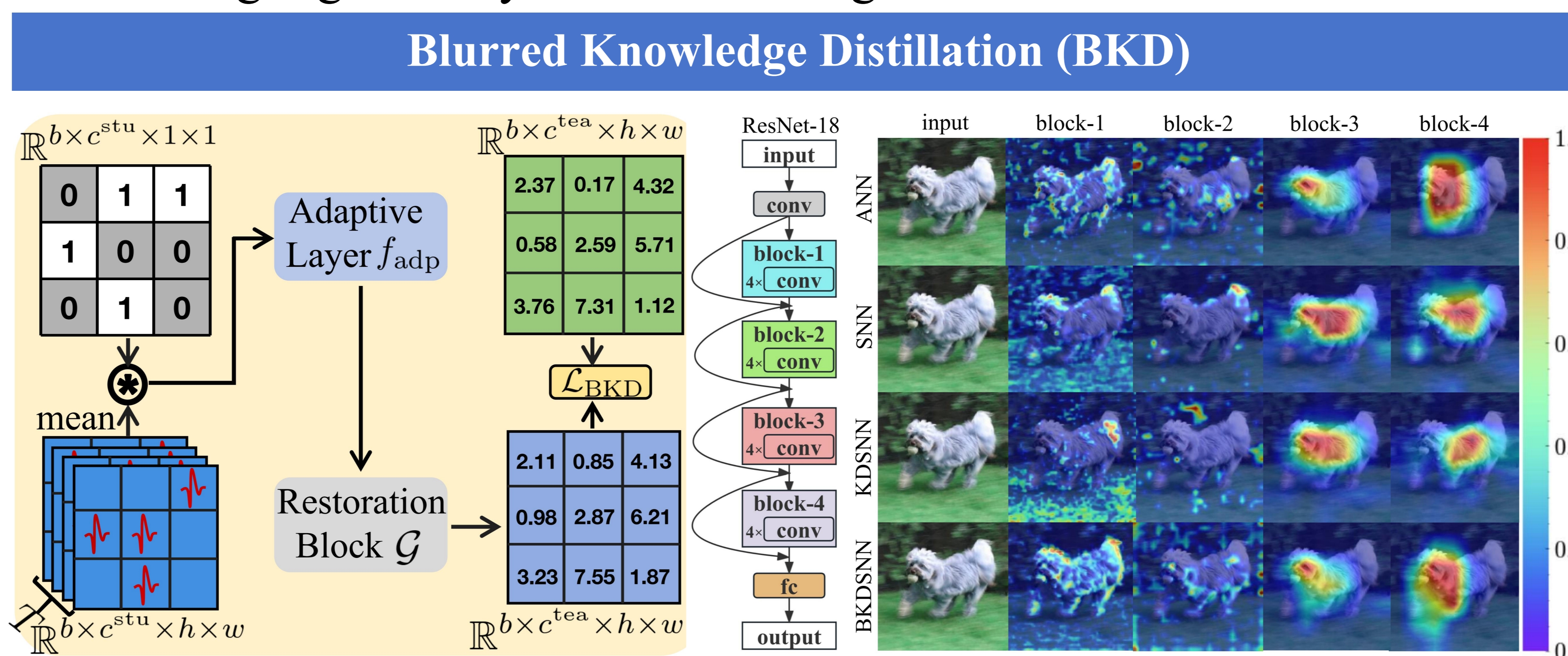
BKD highlighted in yellow shaded region differs from prior SNN KD in three perspectives: 1) A blurred matrix is **randomly sampled on the fly** (per input image) to mask out the feature of student SNN; 2) A restoration block consisting of two convolutional layers connected by ReLU layer is applied on **blurred SNN features** to **restore and mimic ANN features**; 3) BKD is applied only to the **intermediate features before the last layer**.

Experiments

(a) SEW RseNet-18   (b) SEW RseNet-50

(c) Spikingformer-8-384   (d) Spikingformer-8-768

Visualization

input   ANN   SNN   LaSNN   KDSNN   BKDSNN(ours)

Feature map visualization of different methods on SEW ResNet-18 and Spikingformer-8-384. Shaded colors from blue to red indicate the impacts of the regions on the classification scores from low to high.