

基于自注意力机制的自知识蒸馏研究

■ 导师: 何琨 教授

■ 答辩人: 高也

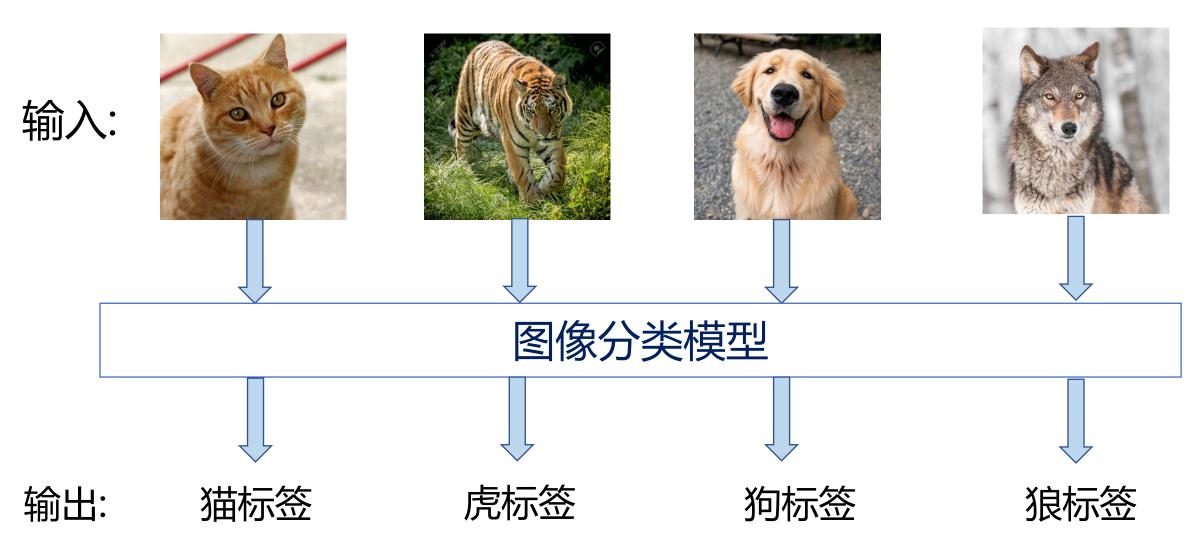
■ 答辩日期: 2022年5月21日



- 1. 研究背景:知识蒸馏和自知识蒸馏。
- 2. 论文方法:基于自注意力机制的自知识蒸馏(SKDSAM)。
- 3. 实验结果: SKDSAM与其他自知识蒸馏模型的对比。
- 4. 总结: SKDSAM显著提升了自蒸馏的分类准确率。

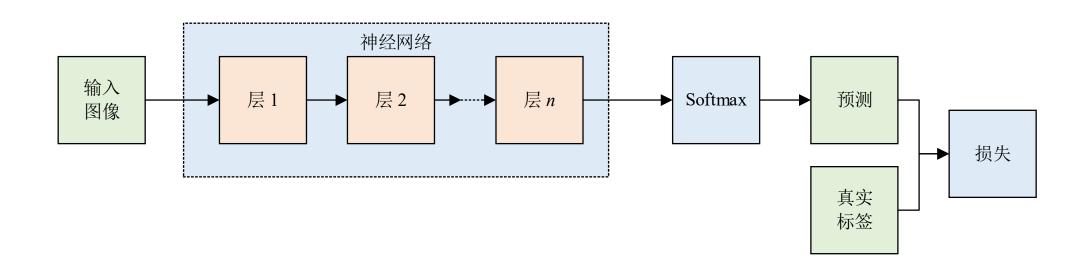


1.1 任务: 图像分类





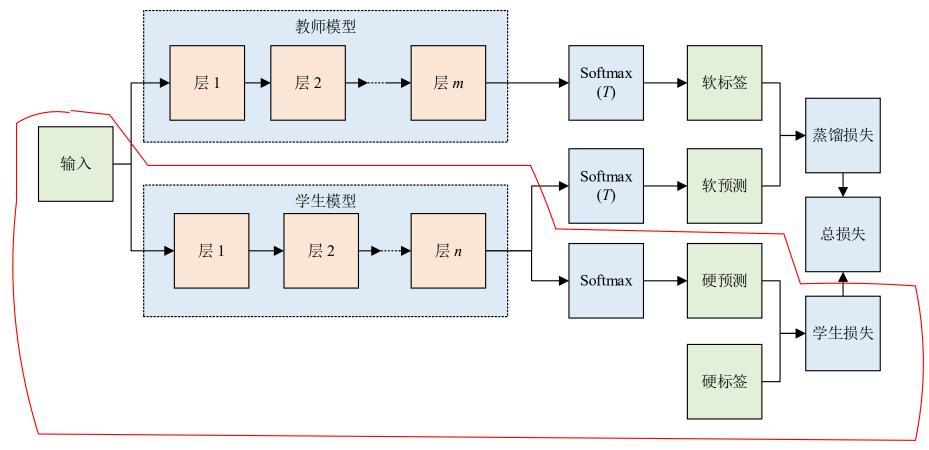
1.2 图像分类模型



$$L = CE(P_S, y)$$



1.3 知识蒸馏:一种神经网络压缩的技术



$$Softmax(T) = \frac{exp(z_i/T)}{\sum_{j} exp(z_j/T)}$$

$$L = CE(P_s, y) + \lambda T^2 \cdot KL(P_s(T), P_t(T))$$

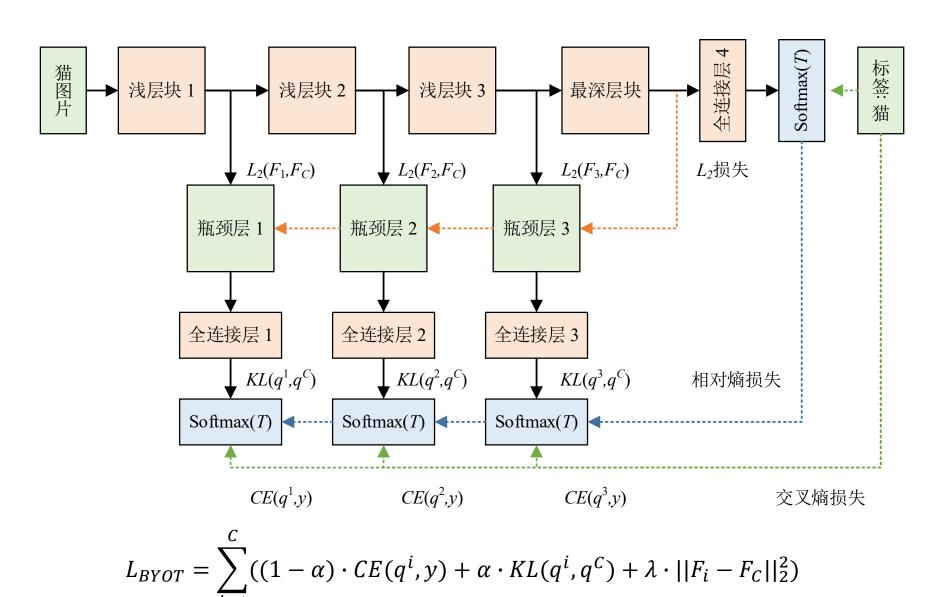


1.4 传统知识蒸馏存在的不足

• 高容量的教师模型的训练过程需要大量的计算和存储资源。



1.5 自知识蒸馏: 不需要外部的教师模型





1.6 BYOT模型存在的不足

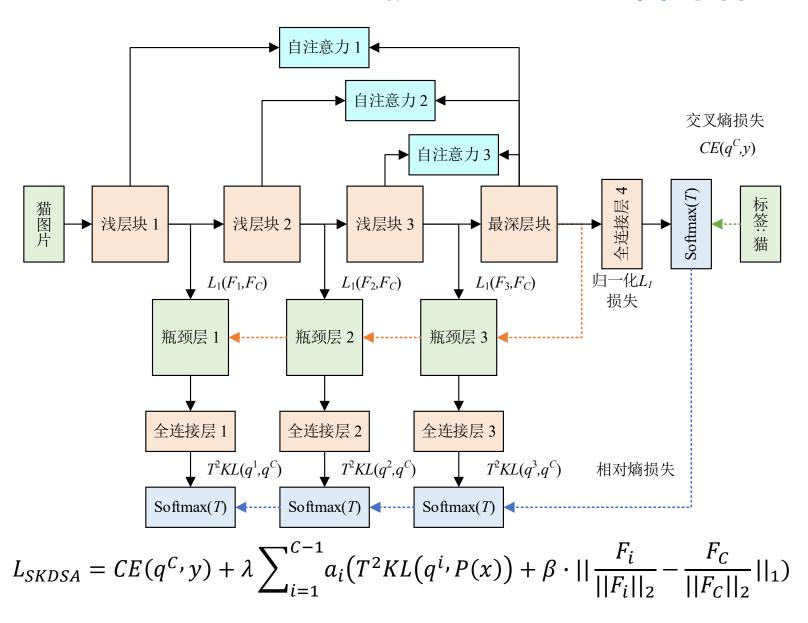
・现有的自知识蒸馏模型(如BYOT模型)将作为学生模型的各个浅层块一视同仁,忽略了各个浅层块对最深层块的不同影响。



- 1. 研究背景:知识蒸馏和自知识蒸馏。
- 2. 论文方法:基于自注意力机制的自知识蒸馏(SKDSAM)。
- 3. 实验评估: SKDSAM与其他自知识蒸馏模型的对比。
- 4. 总结: SKDSAM显著提升了自蒸馏的分类准确率。

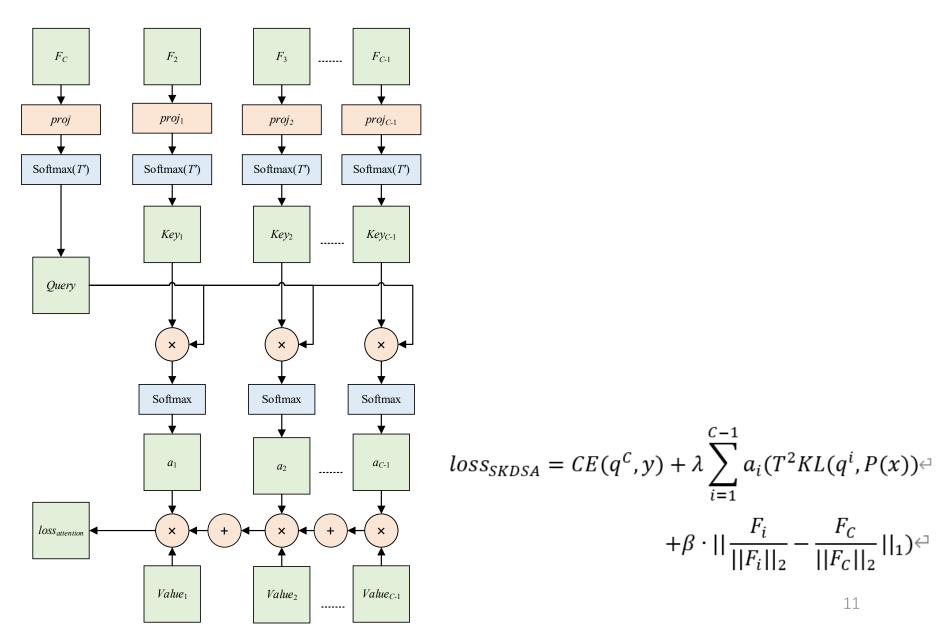


2.1 SKDSAM模型的网络结构



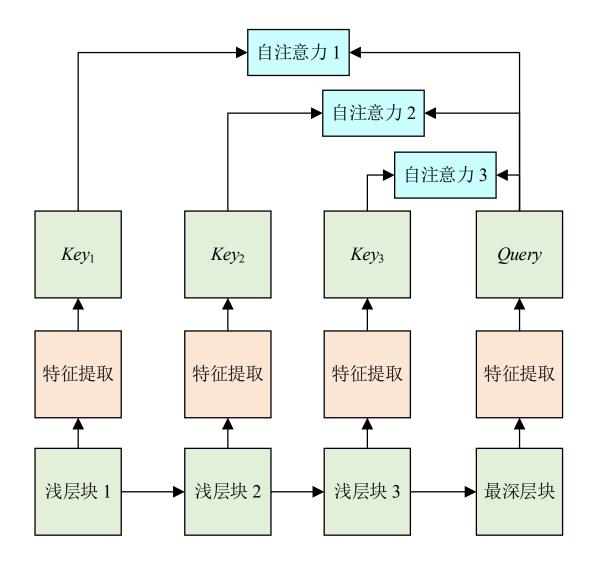


2.2 SKDSAM模型中的自注意力机制





2.3 SKDSAM模型中自注意力机制的直观理解





2.4 结合数据增强技术的SKDSAM模型

- 数据增强是一种在不收集新数据的前提下增加原始数据多样性的技术。
- 方式包含对于原图像进行几何变换、颜色转换、随机擦除、对抗训练和神经风格迁移等等。





图 3.6 原图像(左)和 Cutout 后的图像(右) ←



- 1. 研究背景:知识蒸馏和自知识蒸馏。
- 2. 论文方法:基于自注意力机制的自知识蒸馏(SKDSAM)。
- 3. 实验评估: SKDSAM与其他自知识蒸馏模型的对比。
- 4. 总结: SKDSAM显著提升了自蒸馏的分类准确率。



实验结果

表 4.6 基于 ResNet 框架的各模型在通用图像数据集上的分类准确率(%)

模型	CIFAR-100	Tiny ImageNet
交叉熵	74.80	54.60
DDGSD	76.68	57.76
BYOT	76.87	56.76
CS-KD	78.01	57.72
SLA-SD	77.88	58.67
FRSKD	78.61	59.61
SKDSAM	80.51	60.42
SKDSAM+Cutout	81.77	62.61

- 1. 研究背景:知识蒸馏和自知识蒸馏。
- 2. 论文方法:基于自注意力机制的自知识蒸馏(SKDSAM)。
- 3. 实验评估: SKDSAM与其他自知识蒸馏模型的对比。
- 4. 总结: SKDSAM显著提升了自蒸馏的分类准确率。



结论

- SKDSAM模型取得了当前最高(SOTA)的分类准确率。
- · 结合数据增强技术可以进一步提升SKDSAM模型的性能。



Q&A



致谢

