

算子转化方案

1. 目前实现情况

对目前比较常见的网络框架，例如Tensorflow，Pytorch，Mxnet，针对给出的样例均能实现从前端往后的算子转化，并生成对应的数据流图/网络结构图

2. 算子转化优化

2.1. 明确信息含义

对于算子中的某些信息含义不甚明了，难以确定其参数对应的数据值。

参数	存在算子
input_log2scale	conv & upsample
weight_log2scale	conv
bias_log2scale	conv
output_log2scale	conv & upsample
output_shift	conv
bias_shift	conv
pl_name	add
add_name	add

2.2. 详细的参数说明

对于算子中的一些参数想要一些更详细的描述，例如某些参数的可能值有哪些

算子	参数	可能值
conv	activation_type	Relu / None / ...
upsample	mode	nearest / ...

2.3. 更多的算子样例

为了更好的针对各类算子进行前端往后的转化，可能需要更多的样例进行尝试。

2.4. 问题

- 算子中没有layout相关的信息，不清楚是否会对后续产生有影响
- 算子内部的量化是否需要在外部添加额外的操作进行配合