

Input: (batch, vocab\_size + 1, seg\_len)

output: (batch, vocab\_size, seg\_len)

