

# Depthwise Separable Convolutions for Neural Machine Translation

Depthwise separable convolutions reduce the number of parameters and computation used in convolutional operations while increasing representational efficiency. They have been shown to be successful in image classification models, both in obtaining better models than previously possible for a given parameter count (the Xception architecture) and considerably reducing the number of parameters required to perform at a given level (the MobileNets family of architectures). Recently, convolutional sequence-to-sequence networks have been applied to machine translation tasks with good results. In this work, we study how depthwise separable convolutions can be applied to neural machine translation.

**Author:** ŁUKASZ , Kaiser (Google Brain)

**Presenter:** ŁUKASZ , Kaiser (Google Brain)