

Preprocessed images in
batches of shape (B, C, H, W)

Backbone

CBAM Block

ConvMixer Block

Depth-Wise Convolution

GELU

Batch Normalization

residual connection



$\times depth$

Point-Wise Convolution

GELU

Batch Normalization

Global Average Pooling

Flatten

Fully-connected Layer

Classes