SPA

Prompt Encoder

$[\text{SP}]_2^t$ $[\text{PT}]_2^t$

Transformer Block 2 🔥 ↑

$[\text{SP}]_1^t$ $[\text{PT}]_1^t$

Transformer Block 1 🔥 ↑

Prompt Pool in Task 1, 2, ..., t-1 ❄️   Prompt Pool in Task t 🔥 $[\text{PT}]_0^t$ 🔥

$+$

Segmented Positional Encoding

NKA

$[\text{PT}]_2^{t-1}$

$\times \alpha$

$\oplus$
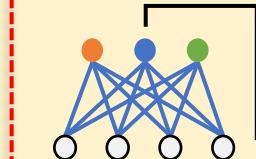
$\times (1-\alpha)$

$[\text{PT}]_1^{t-1}$

$\times \alpha$

$\oplus$

$\times (1-\alpha)$

$g(\cdot)$

l^{t-1} Class ID

$\alpha \updownarrow$

l^t Class ID

Vision Transformer $\mathcal{V}(\cdot)$ ↑

Transformer Block 3 ❄️ ↑
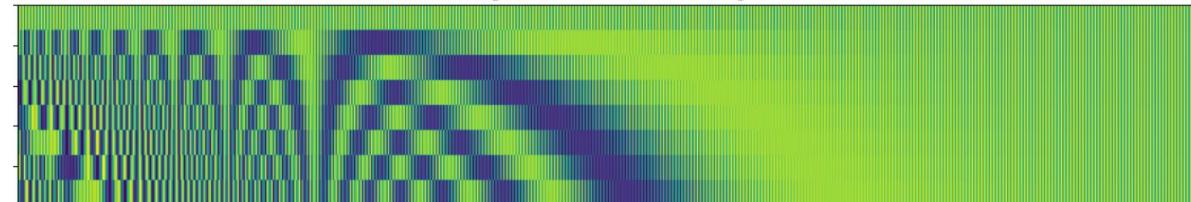
Transformer Block 2 ❄️ ↑

Transformer Block 1 ❄️ ↑

$[\text{CLS}]$

Image   Image Patches