BrAIn Lab.
Brain and Artificial Intelligence Lab.

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

QKFormer: Hierarchical Spiking Transformer usingQ-K Attention

Chenlin Zhou et. al

2026.04.30

BrainLAB Journal Club
인공지능응용학과 이규빈

Overview

01 Author & Journal

Chenlin Zhou

팔로우

내 프로필 만들기

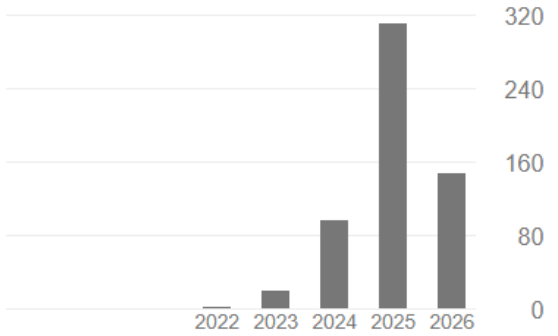
Peking University | Pengcheng Laboratory
stu.pku.edu.cn의 이메일 확인됨

Efficient Artificial Intelligence Brain-inspired Computing

제목	인용	연도
Spikingformer: A Key Foundation Model for Spiking Neural Networks C Zhou, L Yu, Z Zhou, H Zhang, W Jiaqi, H Zhou, Z Ma, Y Tian AAAI 2026 (oral)	187 *	2023
QKFormer: Hierarchical Spiking Transformer using QK Attention C Zhou, H Zhang, Z Zhou, L Yu, L Huang, X Fan, L Yuan, Z Ma, H Zhou, ... NeurIPS 2024 (spotlight)	129	2024
Direct Training High-Performance Deep Spiking Neural Networks: A Review of Theories and Methods C Zhou, H Zhang, L Yu, Y Ye, Z Zhou, L Huang, Z Ma, X Fan, H Zhou, ... Frontiers in Neuroscience	85	2024

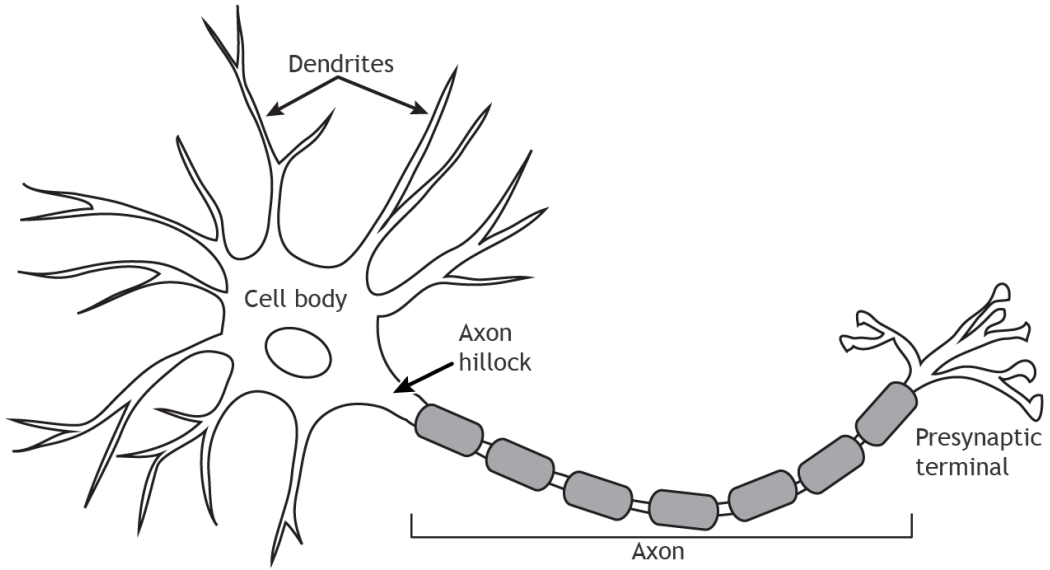
인용

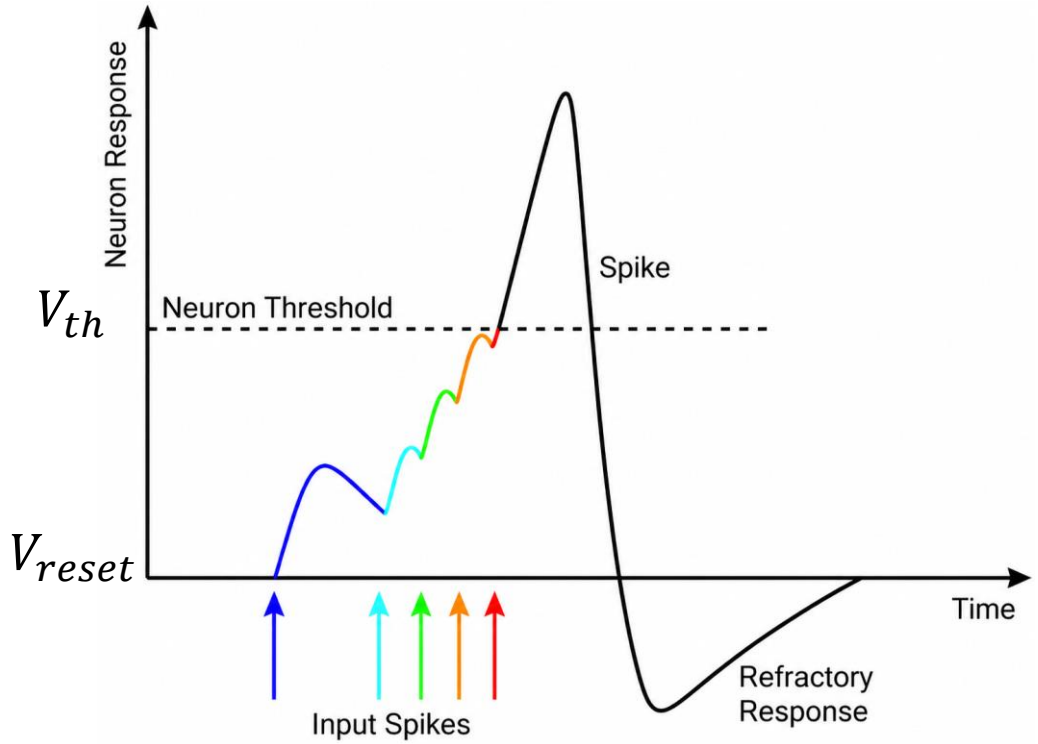
	전체	2021년 이후
서지정보	577	577
h-index	9	9
i10-index	7	7



BrAIn Lab.
Brain and Artificial Intelligence Lab.

1) LIF model



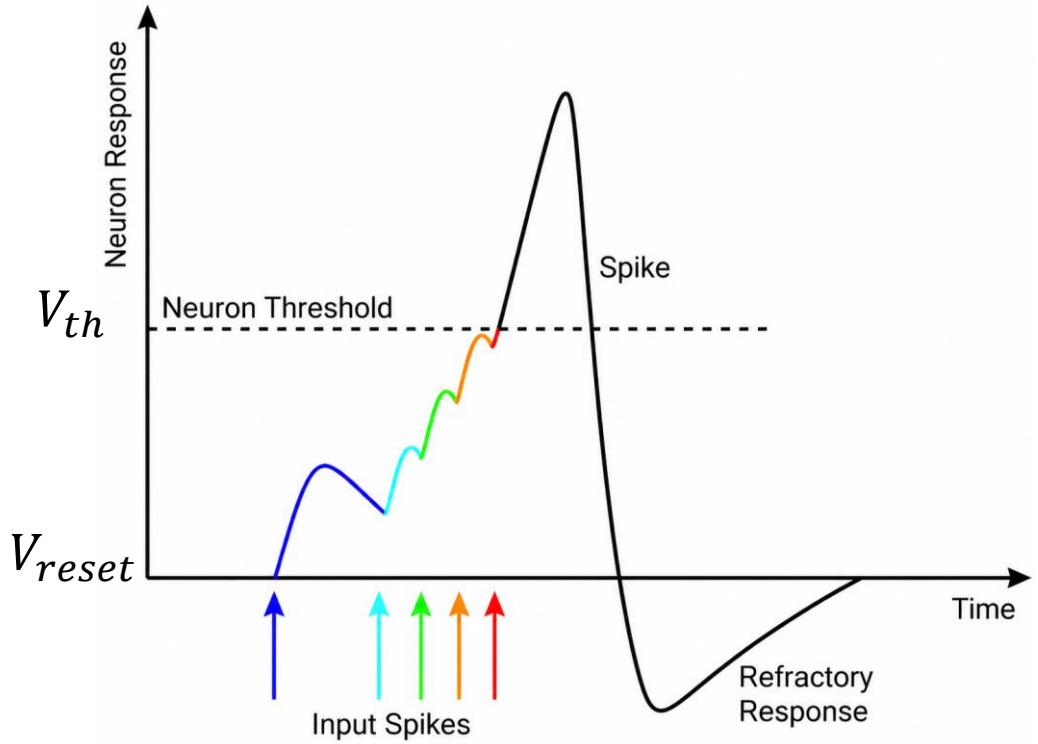


1) LIF model

$$U[t+1] = H[t] + X[t+1]$$

$$S[t+1] = \Theta(U[t+1] - V_{th})$$

$$H[t+1] = \begin{cases} V_{reset} & (\mathrm{S}[t+1] = 1) \\ \tau U[t+1] & (\mathrm{S}[t+1] = 0) \end{cases}$$

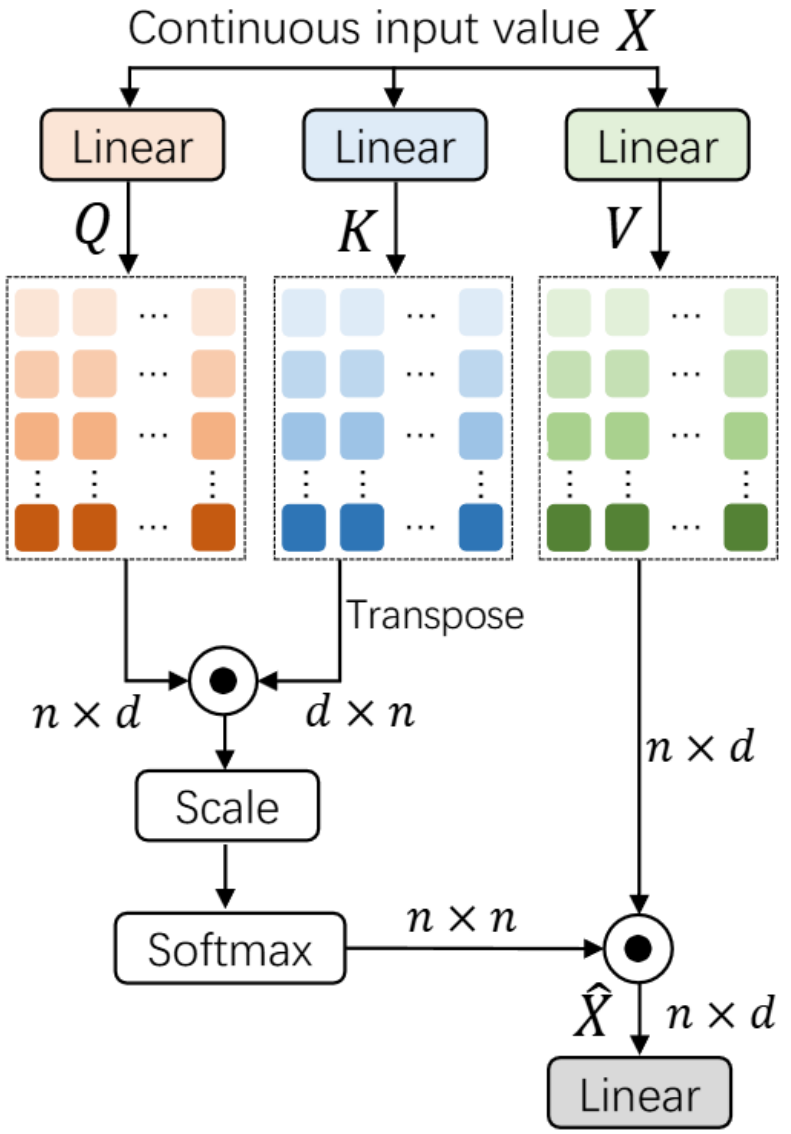


BrAIn Lab.
Brain and Artificial Intelligence Lab.

1) Vanilla Self Attention

$$Q_F, K_F, V_F = X(W_Q, W_K, W_V)$$

$$VSA(Q_F, K_F, V_F) = \text{Softmax}\left(\frac{Q_F K_F^T}{\sqrt{d_k}}\right) V_F$$

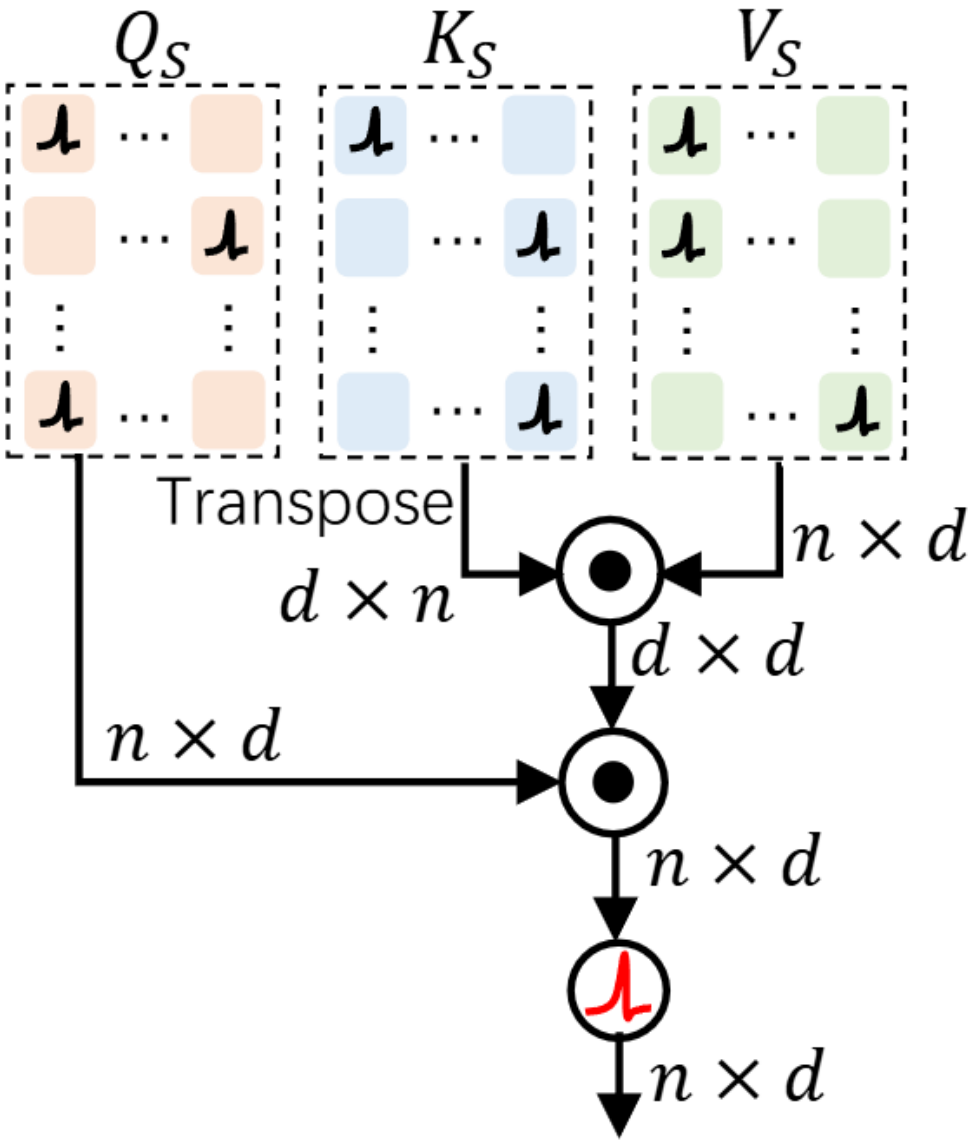


BrAIn Lab.
Brain and Artificial Intelligence Lab.

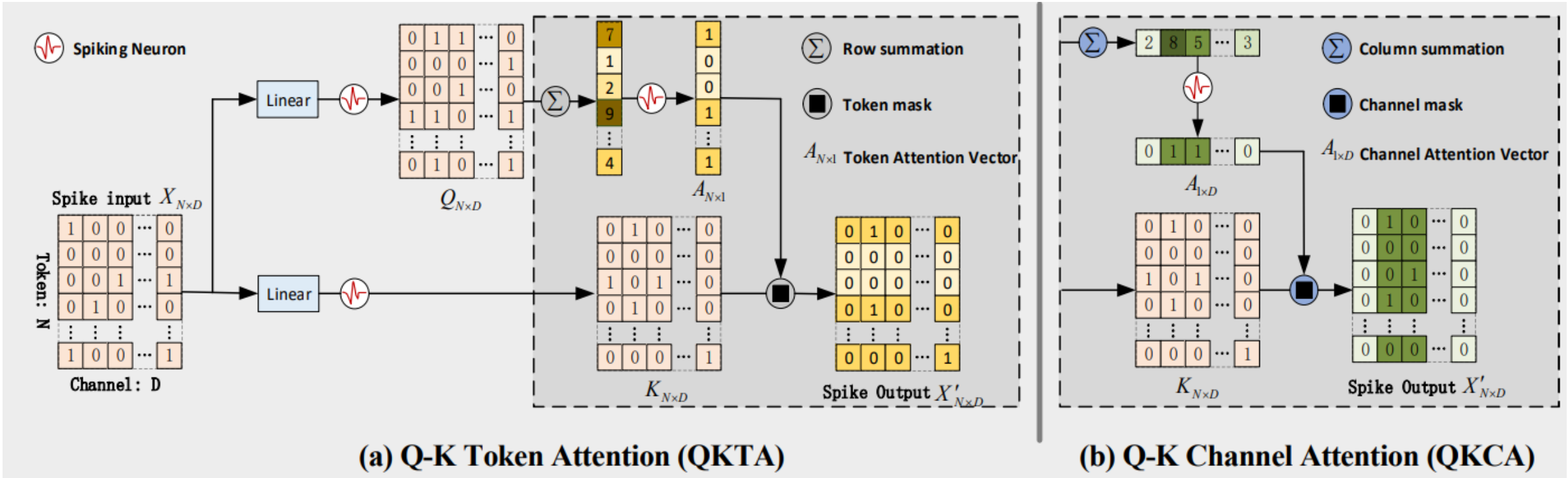
2) Spiking Self Attention

$$I = \mathrm{SN}_I(\mathrm{BN}_I(X(W_I))), \quad I \in \{Q, K, V\}$$

$$\mathrm{SSA}'(Q, K, V) = \mathrm{SN}(QK^T * s)V$$



3) Query-Key Attention



3) Query-Key Attention

$$Q = \mathrm{SN}_Q(\mathrm{BN}(XW_Q)), \quad K = \mathrm{SN}_K(\mathrm{BN}(XW_K))$$

QK Token Attention(QKTA)

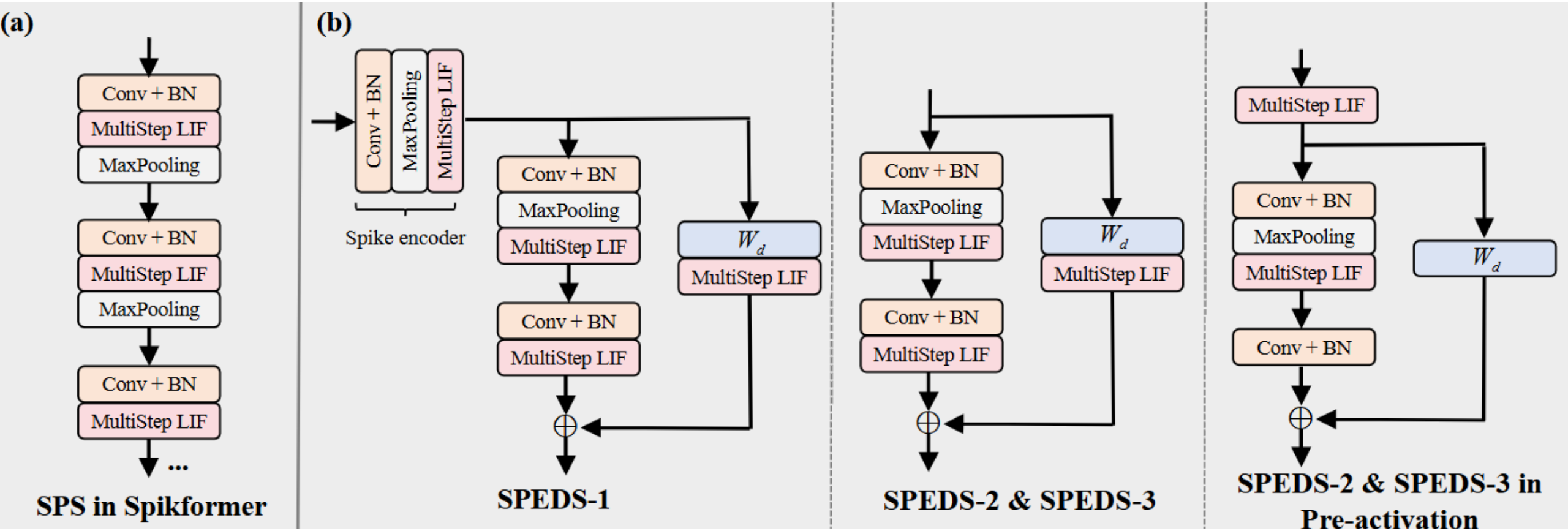
$$A_t = \mathrm{SN}\left(\sum_{i=0}^{D} Q_{i,j}\right), \quad X' = A_t \otimes K$$

QK Channel Attention(QKTA)

$$A_c = \mathrm{SN}\left(\sum_{j=0}^{N} Q_{i,j}\right), \quad X' = A_c \otimes K$$

03 **Key Idea** : Spiking Patch Embedding with Deformed Shortcut

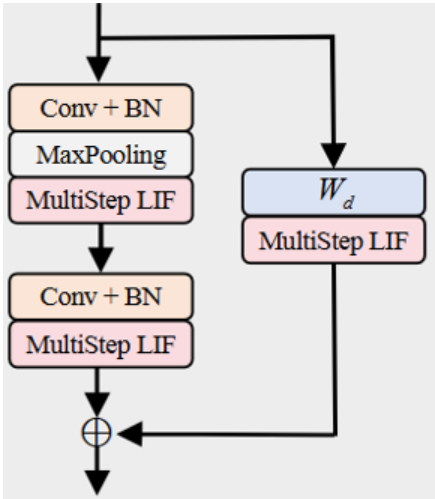
BrAIn Lab.
Brain and Artificial Intelligence Lab.



BrAIn Lab.
Brain and Artificial Intelligence Lab.

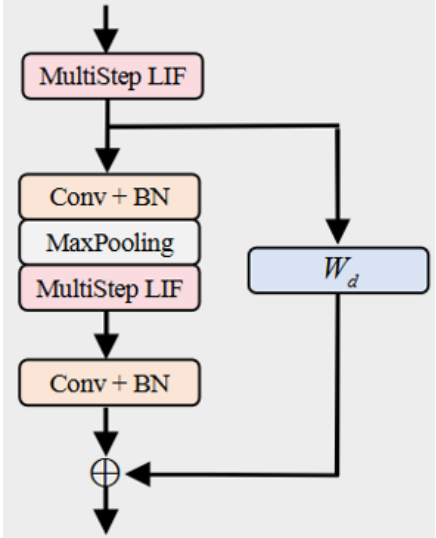
1) Activation Before Addition

$$Y = F(X, \{W_i\}) + \mathrm{SN}(W_d X)$$



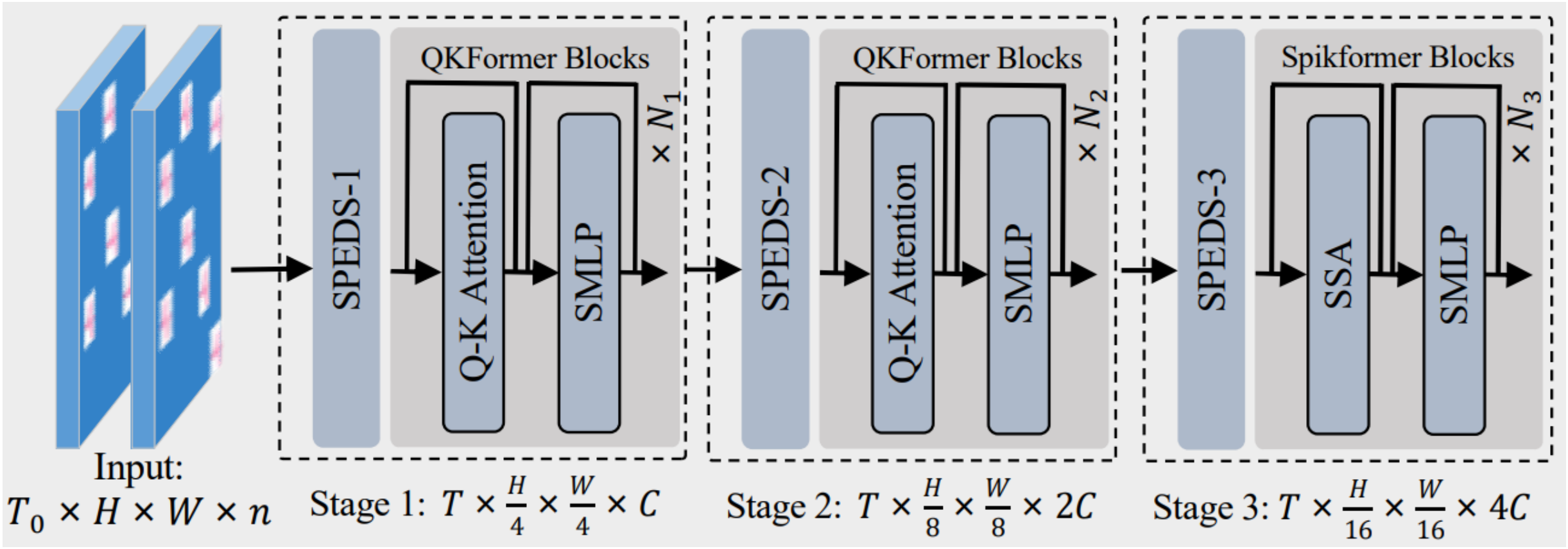
1) Pre-Activation

$$Y = \mathrm{SN}(G(X, \{W_j\}) + W_d X)$$



QKFormer (baseline)	81.15%
QKFormer (ABA->PA)	81.18%

03 **Key Idea** : Hierarchical Architecture



1) ImageNet-1k Classification

Methods	Type	Architecture	Input Size	Param (M)	Power (mJ)	Time Step	Top-1 Acc (%)
RMP[21]	A2S	VGG-16	224^2	39.90	-	2048	73.09
QCFS[22]	A2S	ResNet-18	224^2	11.70	-	1024	74.32
MST[23]	A2S	Swin Transformer-T	224^2	28.50	-	512	78.51
SEW ResNet[28]	SNN	SEW-ResNet-34	224^2	21.79	4.89	4	67.04
	SNN	SEW-ResNet-101	224^2	44.55	8.91	4	68.76
	SNN	SEW-ResNet-152	224^2	60.19	12.89	4	69.26
Spikformer[11]	SNN	Spikformer-8-384	224^2	16.81	7.73	4	70.24
	SNN	Spikformer-8-512	224^2	29.68	11.58	4	73.38
	SNN	Spikformer-8-768	224^2	66.34	21.48	4	74.81
Spikingformer[12]	SNN	Spikingformer-8-384	224^2	16.81	4.69	4	72.45
	SNN	Spikingformer-8-512	224^2	29.68	7.46	4	74.79
	SNN	Spikingformer-8-768	224^2	66.34	13.68	4	75.85
S-Transformer[13]	SNN	S-Transformer-8-384	224^2	16.81	3.90	4	72.28
	SNN	S-Transformer-8-512	224^2	29.68	1.13	1	71.68
	SNN	S-Transformer-8-512	224^2	29.68	4.50	4	74.57
	SNN	S-Transformer-8-768*	288^2	66.34	6.09	4	77.07
ViT[4]	ANN	ViT-B/16	384^2	86.59	254.84	1	77.90
DeiT[32]	ANN	DeiT-B	224^2	86.59	80.50	1	81.80
	ANN	DeiT-B	384^2	86.59	254.84	1	83.10
Swin[8]	ANN	Swin Transformer-B	224^2	87.77	70.84	1	83.50
	ANN	Swin Transformer-B	384^2	87.77	216.20	1	84.50
QKFormer	SNN	HST-10-384	224^2	16.47	15.13	4	78.80
	SNN	HST-10-512	224^2	29.08	21.99	4	82.04
	SNN	HST-10-768	224^2	64.96	8.52	1	81.69
	SNN	HST-10-768	224^2	64.96	38.91	4	84.22
	SNN	HST-10-768*	288^2	64.96	64.27	4	85.25
	SNN	HST-10-768**	384^2	64.96	113.64	4	**85.65**

2) CIFAR, DVS Classification

Method	CIFAR10			CIFAR100			DVS128			CIFAR10-DVS		
	Param	T	Acc	Param	T	Acc	Param	T	Acc	Param	T	Acc
Spikformer [11]	9.32	4	95.51	9.32	4	78.21	2.57	16	98.3	2.57	16	80.9
Spikingformer [12]	9.32	4	95.81	9.32	4	78.21	2.57	16	98.3	2.57	16	81.3
CML [14]	9.32	4	96.04	9.32	4	80.02	2.57	16	98.6	2.57	16	80.9
S-Transformer[13]	10.28	4	95.60	10.28	4	78.4	2.57	16	**99.3**	2.57	16	80.0
STSA[15]	–	–	–	–	–	–	1.99	16	98.7	1.99	16	79.93
ResNet-19 (ANN)	12.63	1	94.97	12.63	1	75.35	–	–	–	–	–	–
Trasnformer (ANN)	9.32	1	96.73	9.32	1	81.02	–	–	–	–	–	–
QKFormer	6.74	4	**96.18**	6.74	4	**81.15**	1.50	16	98.6	1.50	16	**84.0**

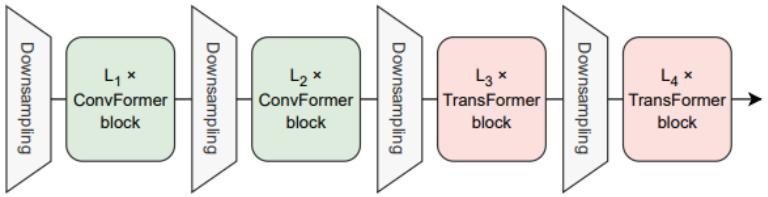
3) Ablation Study

Model	CIFAR100 (Acc)	CIFAR10-DVS (Acc)
QKFormer (QKTA + SSA, baseline)	81.15%	84.00%
QKFormer (QKTA + SSA, w/o SPEDS)	80.08%	83.40%
Spikformer (SSA, w/o scaling)	76.95%	79.30%
Spikformer (SSA)	78.21%	80.90%
Spikformer (SSA) + SPEDS	80.26%	82.20%

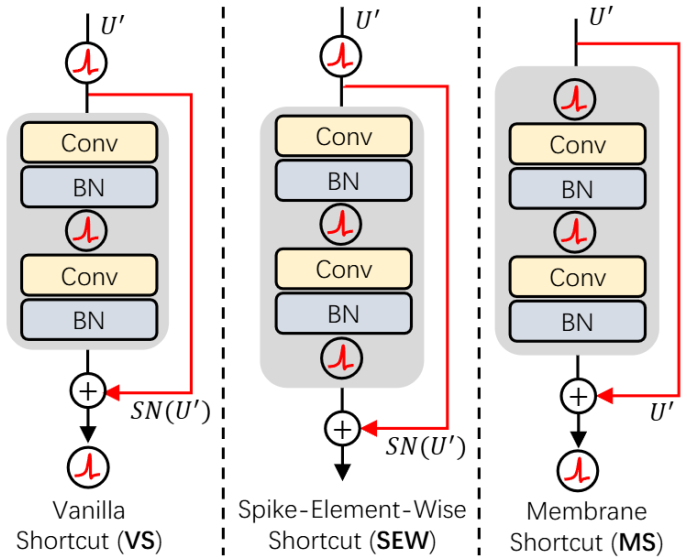
Model	CIFAR100 (Acc, Param)	CIFAR10-DVS (Acc, Param)
QKFormer (QKTA + SSA, baseline)	81.15%, 6.74M	84.00%, 1.50M
QKFormer (QKCA + SSA)	81.07%, 6.74M	84.30%, 1.50M
QKFormer (QKTA + QKCA)	81.04%, 6.44M	83.10%, 1.44M
QKFormer (SSA)	81.23%, 6.79M	84.10%, 1.52M
QKFormer (QKCA)	81.00%, 6.44M	80.70%, 1.44M
QKFormer (QKTA)	79.09%, 6.44M	80.70%, 1.44M

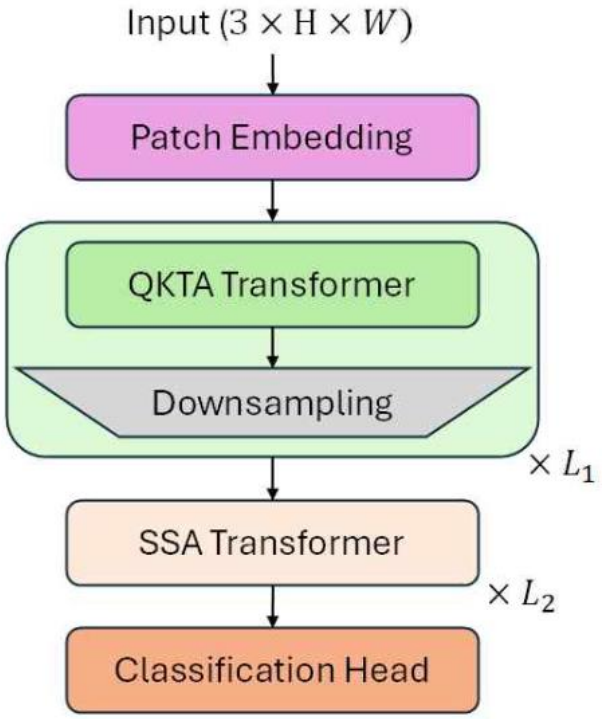
BrAIn Lab.
Brain and Artificial Intelligence Lab.

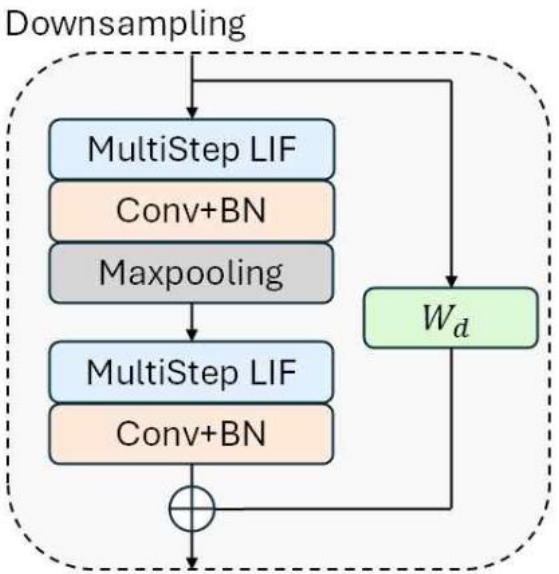
1) Meta-Former

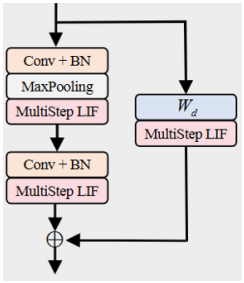


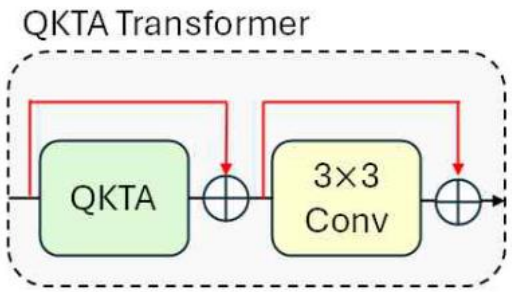
2) Membrane Shortcut

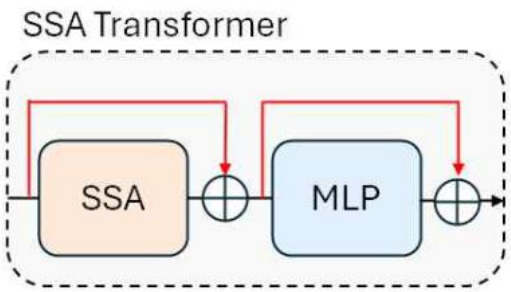












Thank You

2026.04.30

BrainLAB Journal Club
인공지능응용학과 이규빈