Dynamic Aggregated Network for Gait Recognition

Kang Ma[1], Ying Fu[1], Dezhi Zheng[1], Chunshui Cao[2],
Xuecai Hu[2], Yongzhen Huang[2,3]

[1] Beijing Institute of Technology, [2] WATRIX.AI, [3] Beijing Normal University

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

北京理工大學 BEIJING INSTITUTE OF TECHNOLOGY

北京師範大學 BEIJING NORMAL UNIVERSITY

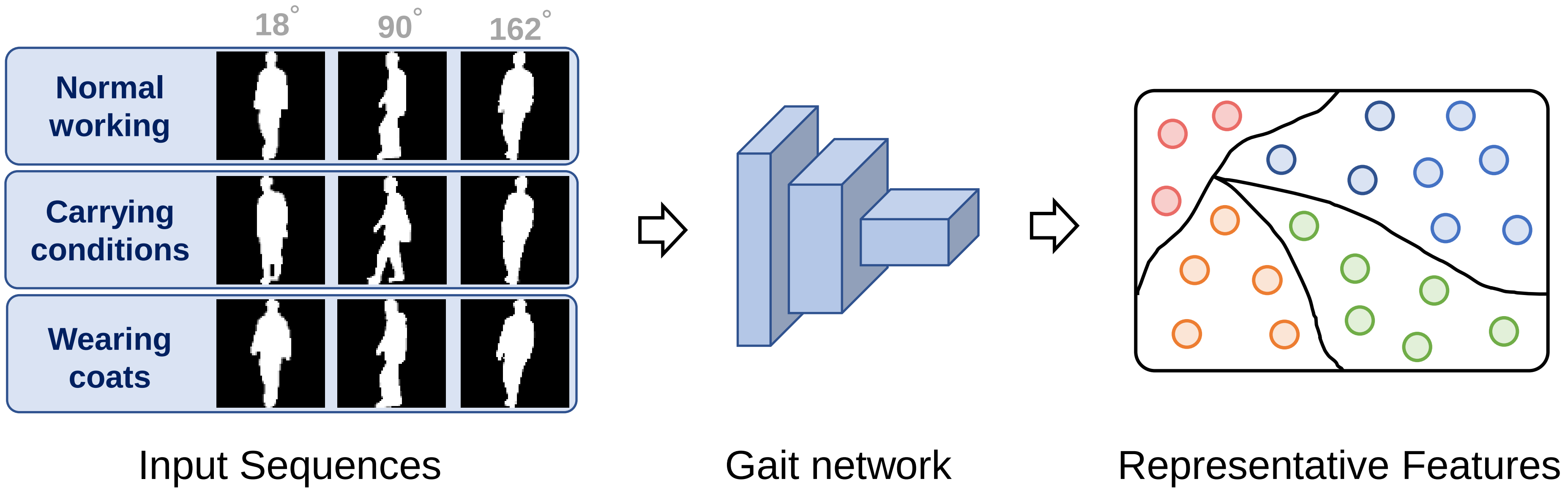
银河水滴 WATRIX. AI

Overview

北京理工大學
BEIJING INSTITUTE OF TECHNOLOGY

Brief introduction to Gait Recognition

Gait recognition is beneficial for **a variety of applications**, including **video surveillance**[1], **crime scene investigation**[2], and **social security**[3], to mention a few.



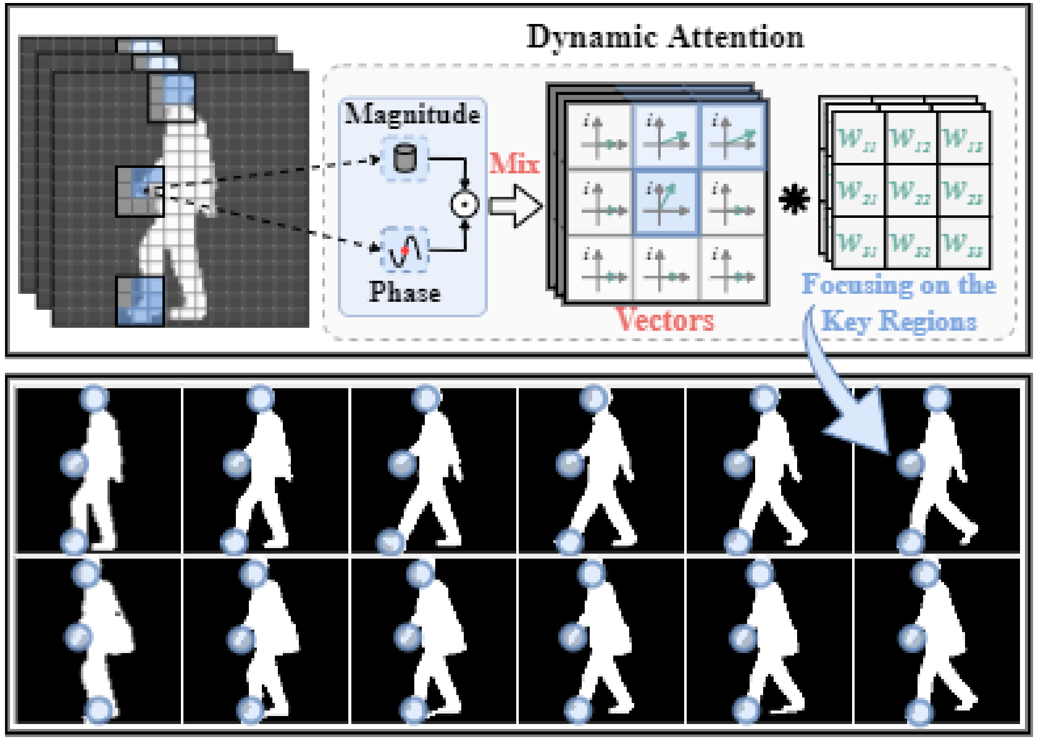
[1] Imed Bouchrika, Michaela Goffredo, John Carter, and Mark Nixon. On using gait in forensic biometrics. Journal of forensic sciences, 2011.
[2] Haruyuki Iwama, Daigo Muramatsu, Yasushi Makihara, and Yasushi Yagi. Gait verification system for criminal investigation. Information and Media Technologies, 2013.
[3] Peter K Larsen, Erik B Simonsen, and Niels Lynnerup. Gait analysis in forensic medicine. Journal of forensic sciences, 2008.

Background

Motivation

The features of each pixel are mapped as a vector with both **magnitude** and **phase** components.

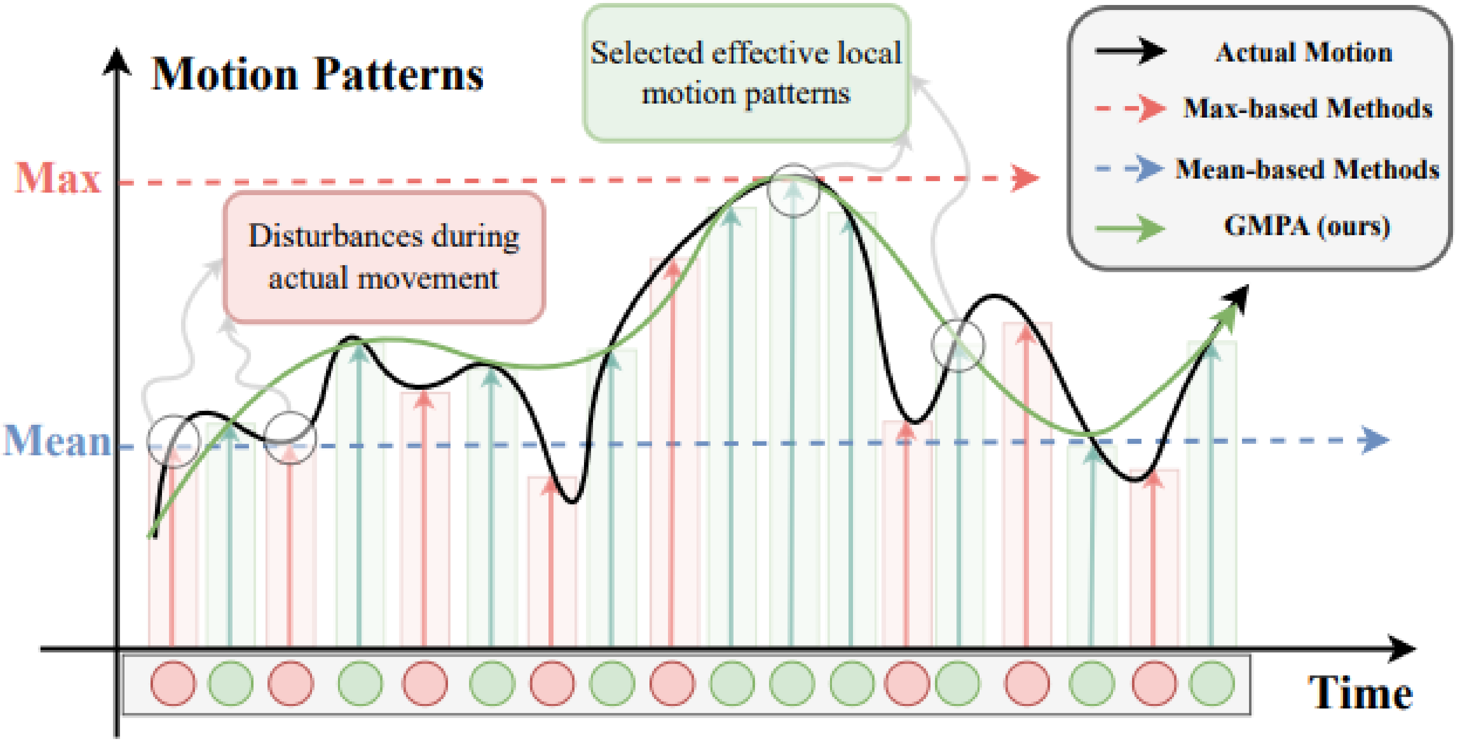
- **Magnitude,** which represents contextual information;
- **Phase,** which is used to construct dynamic attention models for the key regions;



Background

Motivation

In addition, we develop a **self-attention mechanism** to select representative **local motion patterns** and further learn **robust global motion patterns**.

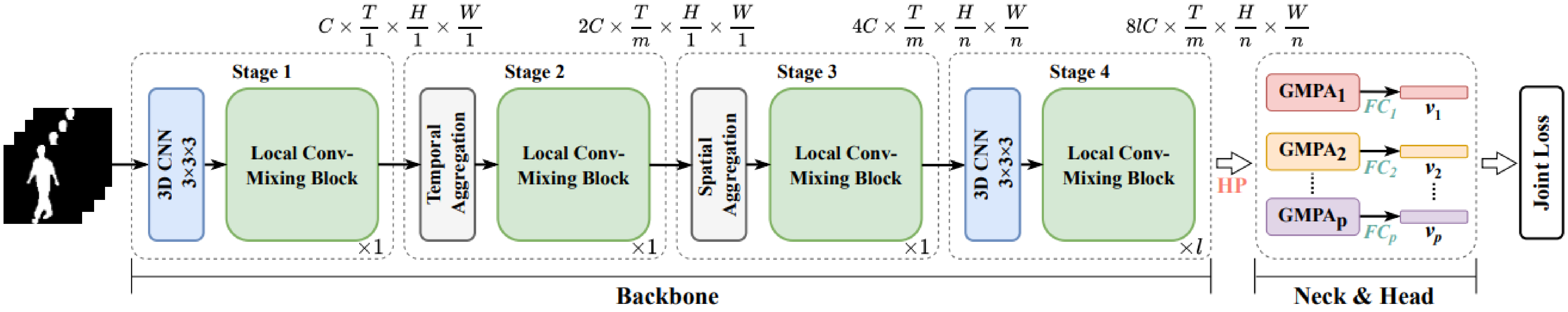


Method

DANet

The overview of the proposed Dynamic Aggregated Network for gait recognition.

- **Local Conv-Mixing Block**, which utilizes a **dynamic attention model** to establish relationships among neighboring pixels of focus.
- **Global Motion Pattern Aggregator**, which is responsible for aggregating the local motion patterns of each part and mapping them separately to produce the global motion patterns.

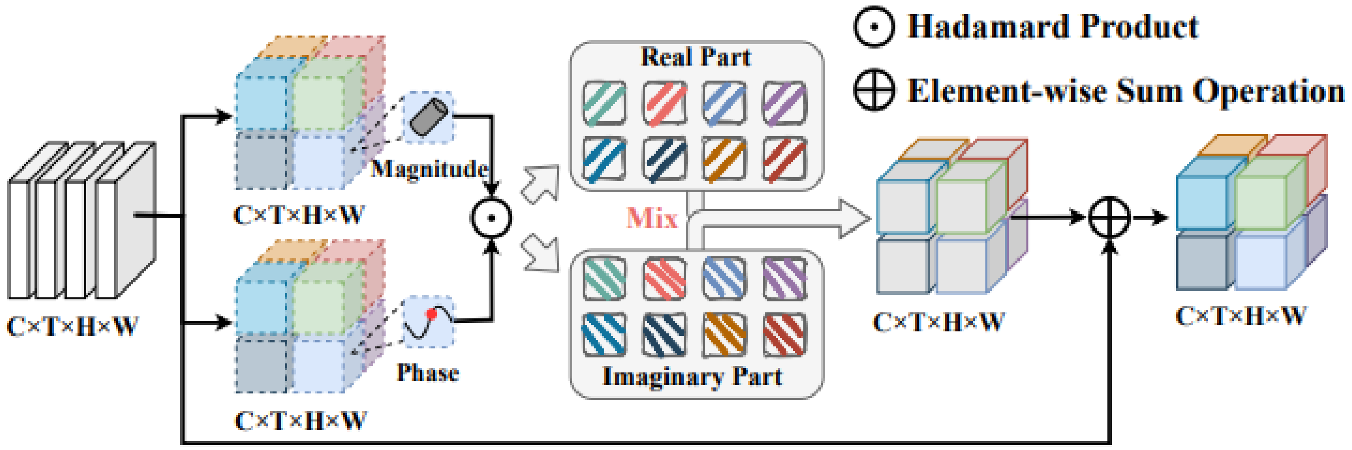


Method

LCMB

The architecture of the proposed Local Conv-Mixing Block.

- **Vector Representation**, which mapping the features of each pixel to the complex domain.
- **Vector Aggregation**, which aggregate the local spatio-temporal domain of each features according to receptive field size.

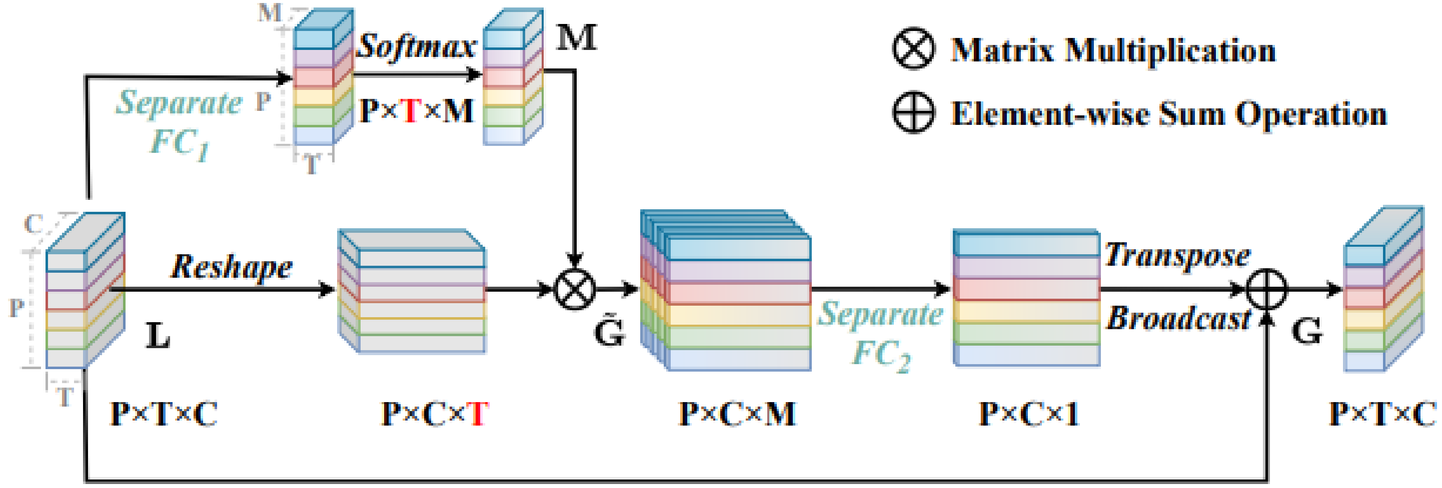


Method

GMPA

The architecture of the proposed Global Motion Pattern Aggregator.

- **Lower-order Global Motion Patterns**, which first propose to squeeze variable local motion patterns into a preset number of channel descriptors.
- **Higher-order Global Motion Patterns**, which perform a further mapping aiming at fully capturing the high-order global motion patterns.



Results

Performance Comparison on CASIA-B

	Method	Probe View											Mean
		0°	18°	36°	54°	72°	90°	108°	126°	134°	162°	180°	
NM	CNN-LB [52]	82.6	90.3	96.1	94.3	90.1	87.4	89.9	94.0	94.7	91.3	78.5	89.9
	GaitSet [8]	90.8	97.9	99.4	96.9	93.6	91.7	95.0	97.8	98.9	96.8	85.8	95.0
	GaitPart [14]	94.1	98.6	99.3	98.5	94.0	92.3	95.9	98.4	99.2	97.8	90.4	96.2
	GLN(*) [20]	93.2	**99.3**	**99.5**	**98.7**	96.1	**95.6**	97.2	98.1	99.3	98.6	90.1	96.9
	MT3D [32]	95.7	98.2	99.0	97.5	95.1	93.9	96.1	98.6	99.2	98.2	92.0	96.7
	GaitGL [33]	96.0	98.3	99.0	97.9	**96.9**	95.4	97.0	98.9	99.3	98.8	94.0	97.4
	LagrangeGait [7]	95.7	98.1	99.1	98.3	96.4	95.2	97.5	99.0	99.3	98.9	94.9	97.5
	DANet(ours)	**96.4**	99.1	99.2	98.2	96.6	95.5	**97.6**	**99.4**	**99.5**	**99.3**	**96.9**	**98.0**
BG	CNN-LB [52]	64.2	80.6	82.7	76.9	64.8	63.1	68.0	76.9	82.2	75.4	61.3	72.4
	GaitSet [8]	83.8	91.2	91.8	88.8	83.3	81.0	84.1	90.0	92.2	94.4	79.0	87.2
	GaitPart [14]	89.1	94.8	96.7	95.1	88.3	**94.9**	89.0	93.5	96.1	93.8	85.8	91.5
	GLN(*) [20]	91.1	**97.7**	97.8	95.2	92.5	91.2	92.4	96.0	97.5	95.0	88.1	94.0
	MT3D [32]	91.0	95.4	97.5	94.2	92.3	86.9	91.2	95.6	97.3	96.4	86.6	93.0
	GaitGL [33]	92.6	96.6	96.8	95.5	93.5	89.3	92.2	96.5	**98.2**	96.9	91.5	94.5
	LagrangeGait [7]	94.2	96.2	96.8	95.8	94.3	89.5	91.7	96.8	98.0	97.0	90.9	94.6
	DANet(ours)	**95.0**	97.3	**98.3**	**97.4**	**94.7**	91.0	**93.9**	**97.4**	**98.2**	**97.6**	**94.2**	**95.9**
CL	CNN-LB [52]	37.7	57.2	66.6	61.1	55.2	54.6	55.2	59.1	58.9	48.8	39.4	54.0
	GaitSet [8]	61.4	75.4	80.7	77.3	72.1	70.1	71.5	73.5	73.5	68.4	50.0	70.4
	GaitPart [14]	70.7	85.5	86.9	83.3	77.1	72.5	76.9	82.2	83.8	80.2	66.5	78.7
	GLN(*) [20]	70.6	82.4	85.2	82.7	79.2	76.4	76.2	78.9	77.9	78.7	64.3	77.5
	MT3D [32]	76.0	87.6	89.8	85.0	81.2	75.7	81.0	84.5	85.4	82.2	68.1	81.5
	GaitGL [33]	76.6	90.0	90.3	87.1	84.5	79.0	84.1	87.0	87.3	84.4	69.5	83.6
	LagrangeGait [7]	77.4	90.6	93.2	90.2	84.7	80.3	85.2	87.7	89.3	86.6	71.0	85.1
	DANet(ours)	**82.8**	**94.8**	**96.9**	**94.3**	**89.0**	**83.9**	**87.9**	**92.3**	**95.1**	**92.0**	**80.3**	**89.9**

Results

北京理工大學
BEIJING INSTITUTE OF TECHNOLOGY

Performance Comparison on OUMVLP

Method	Probe View														Mean
	0°	15°	30°	45°	60°	75°	90°	180°	195°	210°	225°	240°	255°	270°	
GaitSet [8]	79.5	87.9	89.9	90.2	88.1	88.7	87.8	81.7	86.7	89.0	89.3	87.2	87.8	86.2	87.1
GaitPart [14]	82.6	88.9	90.8	91.0	89.7	89.9	89.5	85.2	88.1	90.0	90.1	89.0	89.1	88.2	88.7
GLN [20]	83.8	90.0	91.0	91.2	90.3	90.0	89.4	85.3	89.1	90.5	90.6	89.6	89.3	88.5	89.2
GaitGL [33]	84.9	90.2	91.1	91.5	91.1	90.8	90.3	88.5	88.6	90.3	90.4	89.6	89.5	88.8	89.7
LagrangeGait [7]	85.9	90.6	91.3	91.5	91.2	91.0	90.6	88.9	89.2	90.5	90.6	89.9	89.8	89.2	90.0
CSTL [23]	87.1	91.0	91.5	**91.8**	90.6	90.8	90.6	89.4	90.2	90.5	90.7	89.8	90.0	89.4	90.2
DANet(ours)	**87.7**	**91.3**	**91.6**	**91.8**	**91.7**	**91.4**	**91.1**	**90.4**	**90.3**	**90.7**	**90.9**	**90.5**	**90.3**	**89.9**	**90.7**

Conclusions

1. In this paper, we propose a novel **Dynamic Aggregated Network** (**DANet**) for gait recognition, which consists of a serial of **Local Conv-Mixing Block** (**LCMB**) and **Global Motion Pattern Aggregator** (**GMPA**) to adaptively aggregate the robust discriminative global motion patterns.
2. The proposed method can dynamically **locate the key regions and extract the local motion patterns**, and then adaptively select the distinguishing local motion patterns to further **construct robust global motion patterns**.
3. The experimental results on **three popular gait datasets**, ***i.e.***, **CASIA-B**, **OUMVLP**, and **Gait3D**, verify the effectiveness of the proposed method and show great potential for practical applications.
4. In the future, we will further investigate adaptive learning of the local and global motion patterns in the complex-valued domain to aggregate more representative gait features.

北京理工大學
BEIJING INSTITUTE OF TECHNOLOGY

Thank you!