



中国科学技术大学
University of science and technology of China

The presentation for Gait recognition competition

Reported by Hao Sun



CONTENT

- 01 Research background and purpose
- 02 Experimental materials
- 03 Model analysis and data processing methods
- 04 Conclusion analysis and recommendations



中国科学技术大学
University of science and technology of China

1

Research background and purpose

Research object description



01

The farthest recognition distance of gait recognition using high-definition camera can reach 100 meters, and the city camera can also reach 50 meters

02

Gait recognition in the general sense does not require the cooperation of the recognition object, and the user experience is better

03

Compared with most biological characteristics, gait features are quite high in security

Current research development



GaitSet

Powerful cross-view gait recognition method



GaitPart

Targeted capture of the gait pattern of different parts of the human body



GaitNet

Construct encoder and decoder to learn features

Research background and purpose

Improve the cross-view capability of the recognition model

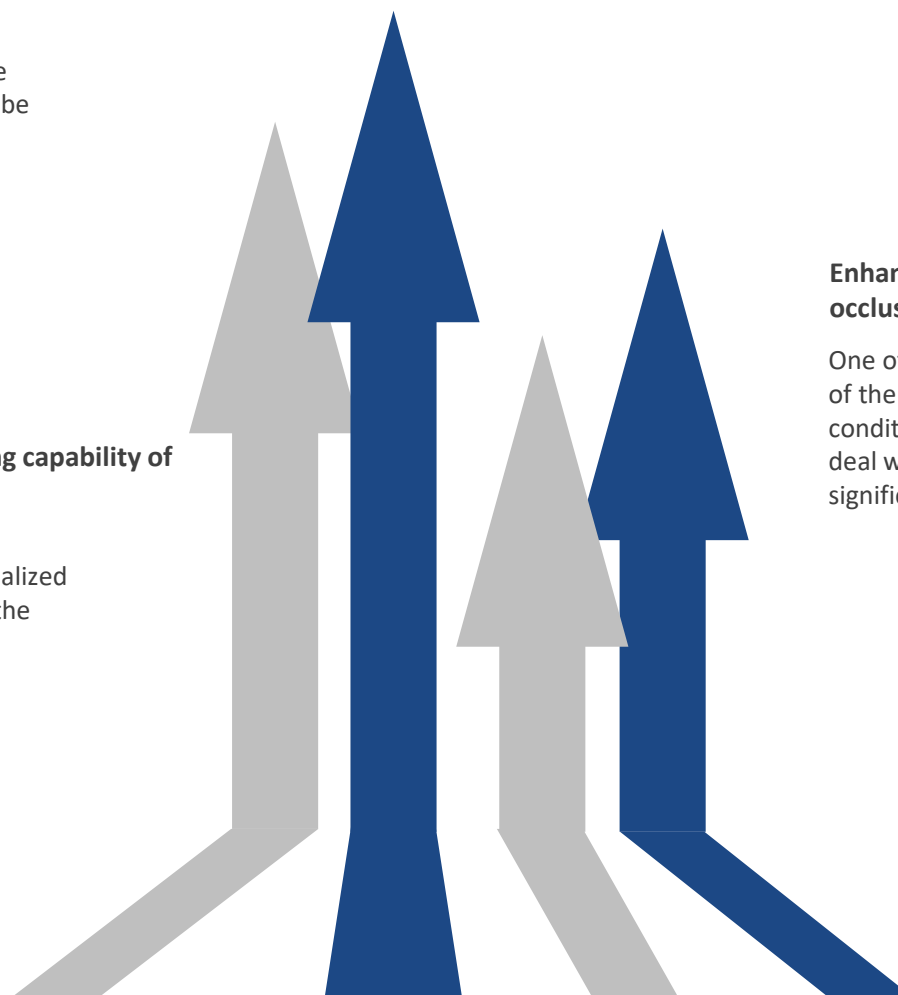
The performance of cross-view recognition is the key to whether the gait recognition method can be applied to the actual scene, and it is also the key factor of the evaluation model

Improve the sequence information processing capability of the recognition model

Effective processing of spatial and temporal serialized information in the contour map is beneficial to the improvement of prediction accuracy

Enhance the performance of the recognition model under occlusion

One of the important factors affecting the accuracy of the gait recognition model is the occlusion condition of the environment, which can effectively deal with the impact of occlusion and can significantly improve the robustness of the model





中国科学技术大学
University of science and technology of China

2

E x p e r i m e n t a l m a t e r i a l s

Data processing method



RGB images use human body detection and human body segmentation to generate contour maps to focus the model's attention on the learning of gait patterns

Experimental Materials



CASIA-E dataset

Our training data set is a background-free contour map obtained by human body detection and human body segmentation processing on the captured RGB image set, which helps us focus on the processing of walking patterns

This data set is obtained by processing 500 different pedestrian walking videos



中国科学技术大学
University of science and technology of China

3

Model analysis and data processing methods

Model analysis

the frame-level CNN Structure of GaitSet

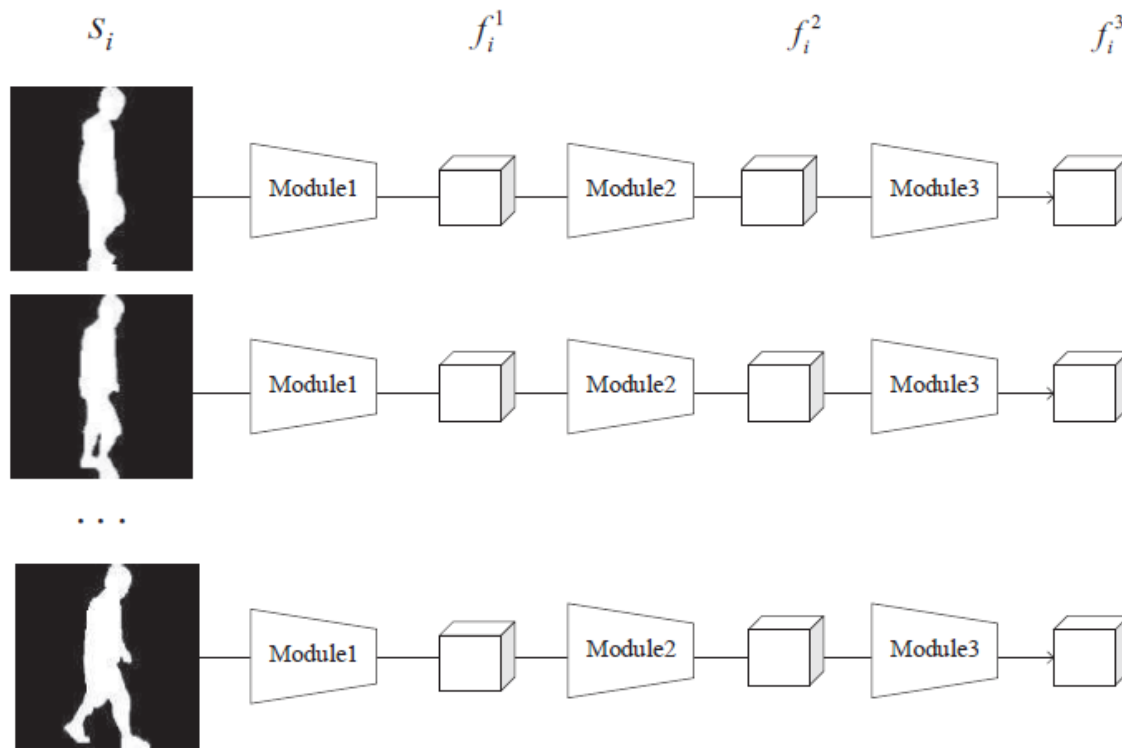


Table 1. The convolution network settings

module	layer	kernel size	stride
module1	conv1-1	5×5	1
	conv1-2	3×3	1
	pooling1	2×2	2
module2	conv2-1	3×3	1
	conv2-2	3×3	1
	pooling2	2×2	2
module3	conv3-1	3×3	1
	conv3-2	3×3	1

Fig. 2. The convolution network of GaitSet.



Model analysis

Set Pooling (SP) operation

$$F^l[i, j, k] = SP(\{f_1^l[i, j, k], f_2^l[i, j, k], \dots, f_m^l[i, j, k]\})$$

$$SP \in \{Max, Median, Mean, JointFunc, etc\}$$

$$JointFunc(*) = 1_1Conv(concat[Max(*), Median(*), Mean(*)])$$

Model analysis

Horizontal Pyramid Mapping (HPM)

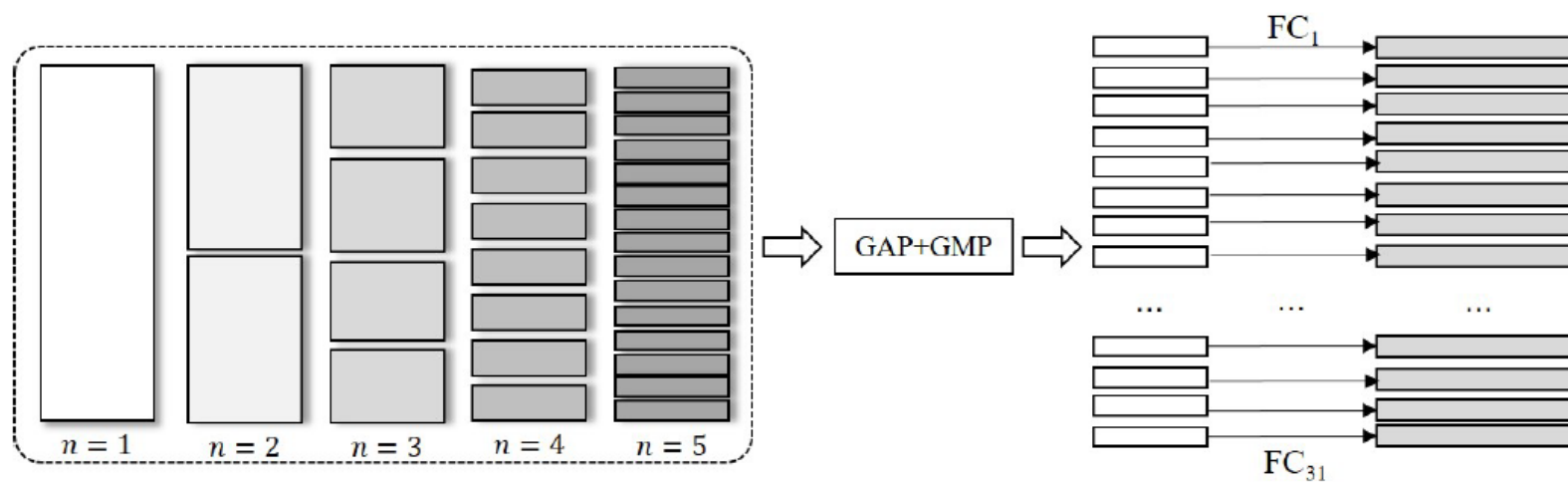


Fig. 4. The Horizontal Pyramid Mapping module.

Model analysis

Gaitset is adopted based on the hypothesis that the frame sequence of a walking video can be recovered from the unordered set of frames.



When the walking subject moves with high-frequency turning and waving actions, the unordered set may lose some temporal information, consequently, the accuracy of GaitSet decreases.



Model analysis

Micro-motion Capture Module (MCM)*

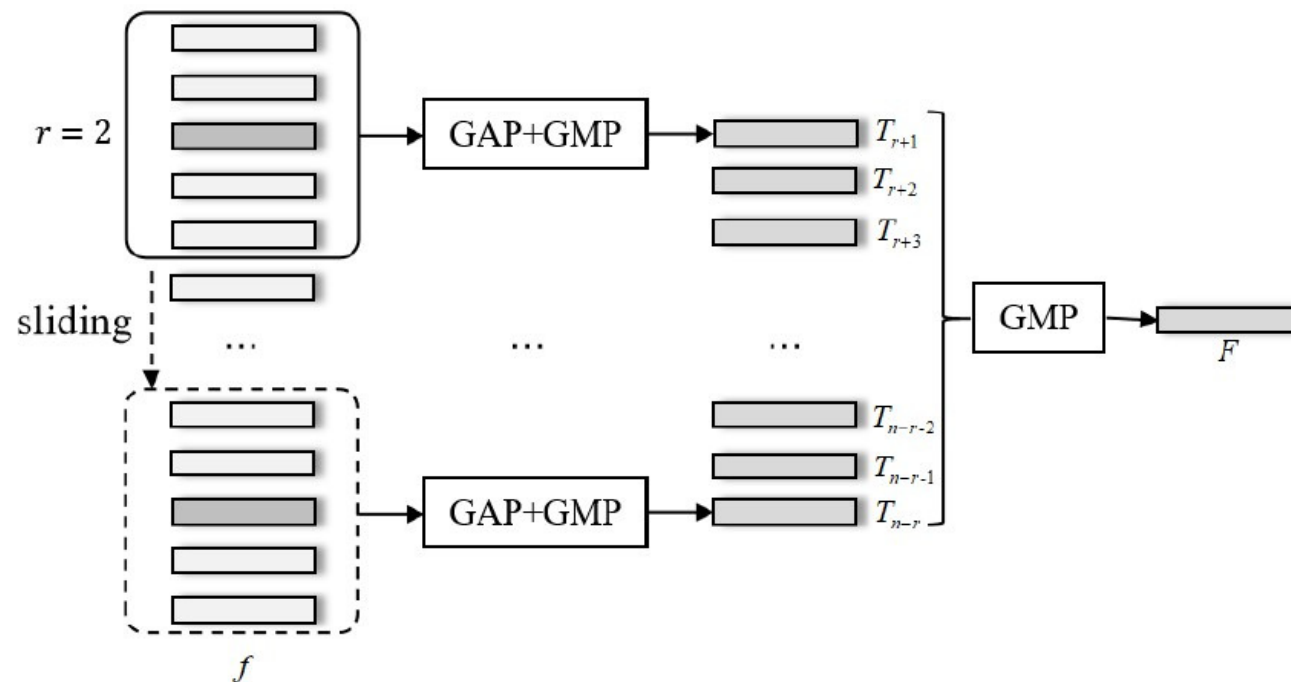


Fig. 5. The Micro-motion Capture Module.

*Fan C, Peng Y, Cao C, et al. GaitPart: Temporal Part-Based Model for Gait Recognition[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 14225-14233.

Data processing method

01

Picture flip

The flip of the picture can effectively enhance the recognition accuracy of the model in multiple perspectives

02

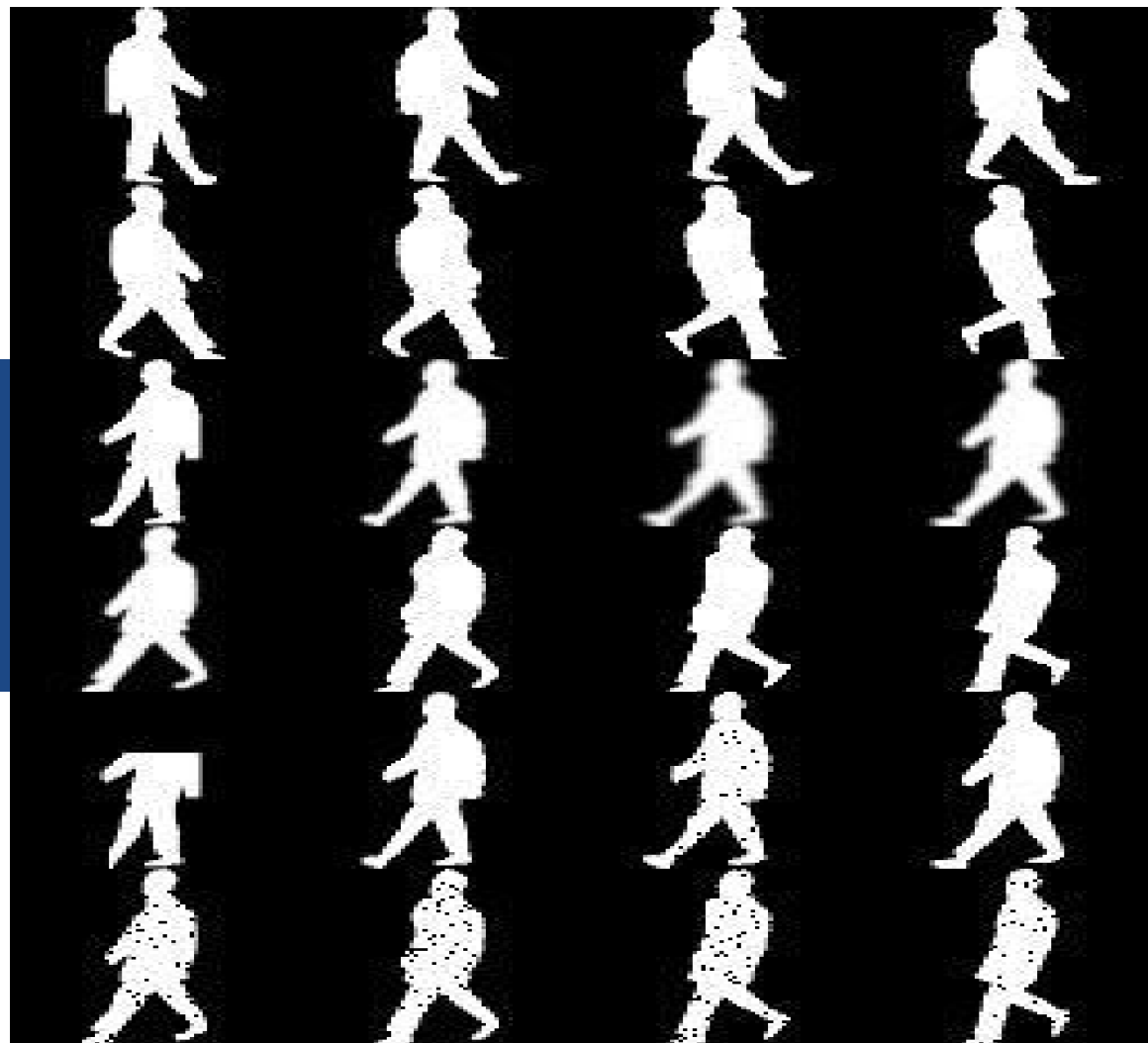
Gaussian blur

Gaussian blur further eliminates high-frequency features and focuses the model's attention on the learning of low-frequency features

03

Partial occlusion processing

Partial occlusion processing enables the training data set to handle the partial occlusion of the human body by the model in a complex environment



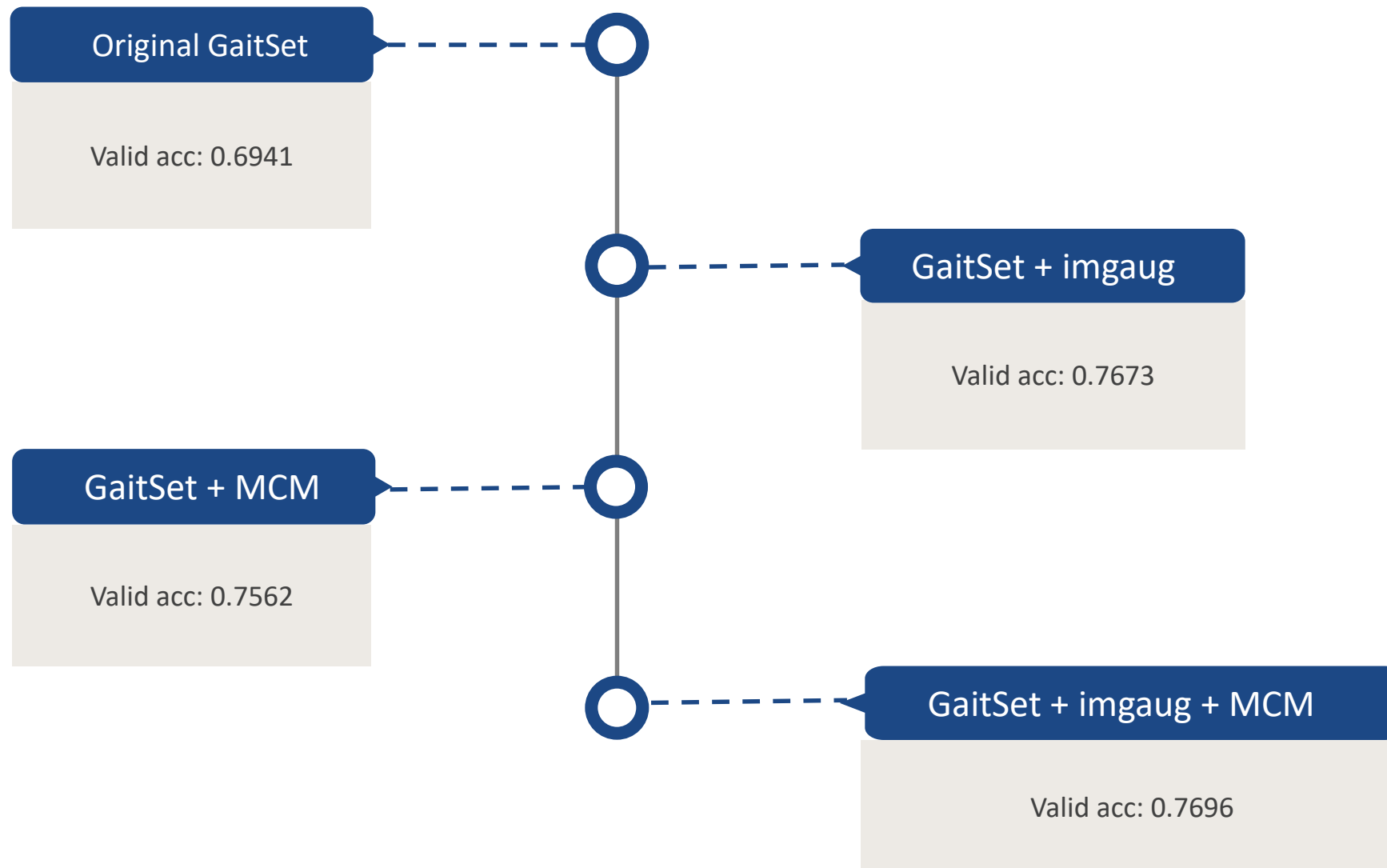


中国科学技术大学
University of science and technology of China

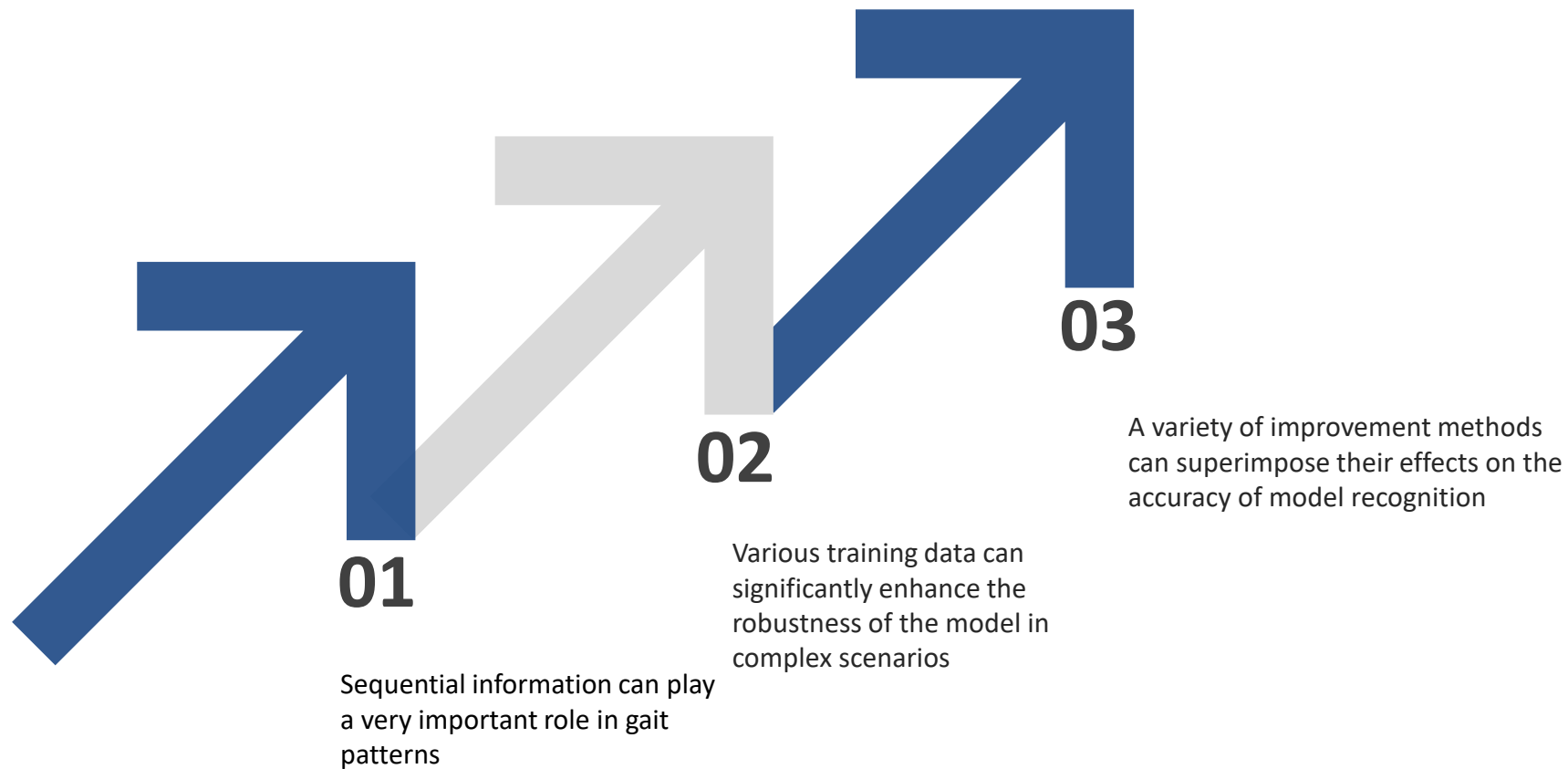
4

Conclusion analysis and summary

Accuracy comparison



Research summary





中国科学技术大学
University of science and technology of China

感谢评委的指导
