



南京大學

NANJING UNIVERSITY

RINC



基于度量学习和注意力机制的步态识别研究

- 答辩人：李雪健 MG1833044
- 导师：申富饶 教授
宋方敏 教授

目录

CONTENTS

1 研究背景

2 提出方法

- ◆ 结合识别与分类的步态识别损失函数设计
- ◆ 基于注意力机制的步态识别网络的设计

3 实际应用

- ◆ 步态识别系统

4 研究生期间工作成果



第一部分

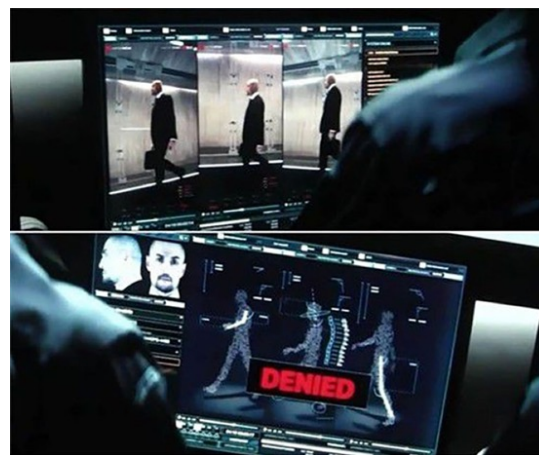
Research Background 研究背景



- 步态识别：依据人走路的姿态对人进行识别
- 优势：远距离识别、难以伪装和模仿、应用成本低
- 应用场景



监控视频查看



高精度身份验证



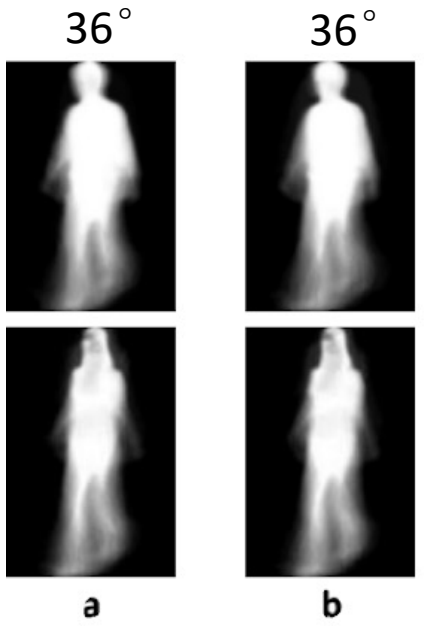
疫情下的身份验证



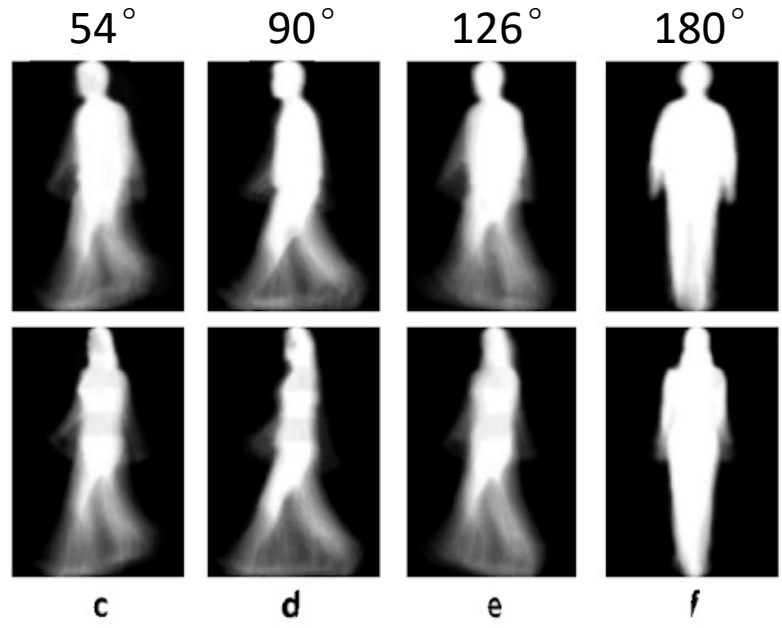
目前所面临的难点

两个不同的人

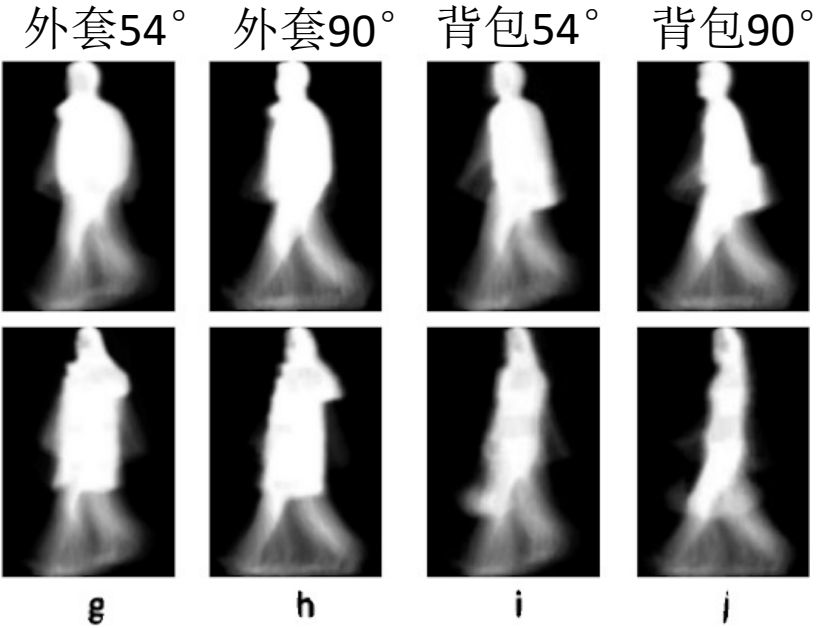
库中的GEI 同视角的GEI



跨视角的GEI



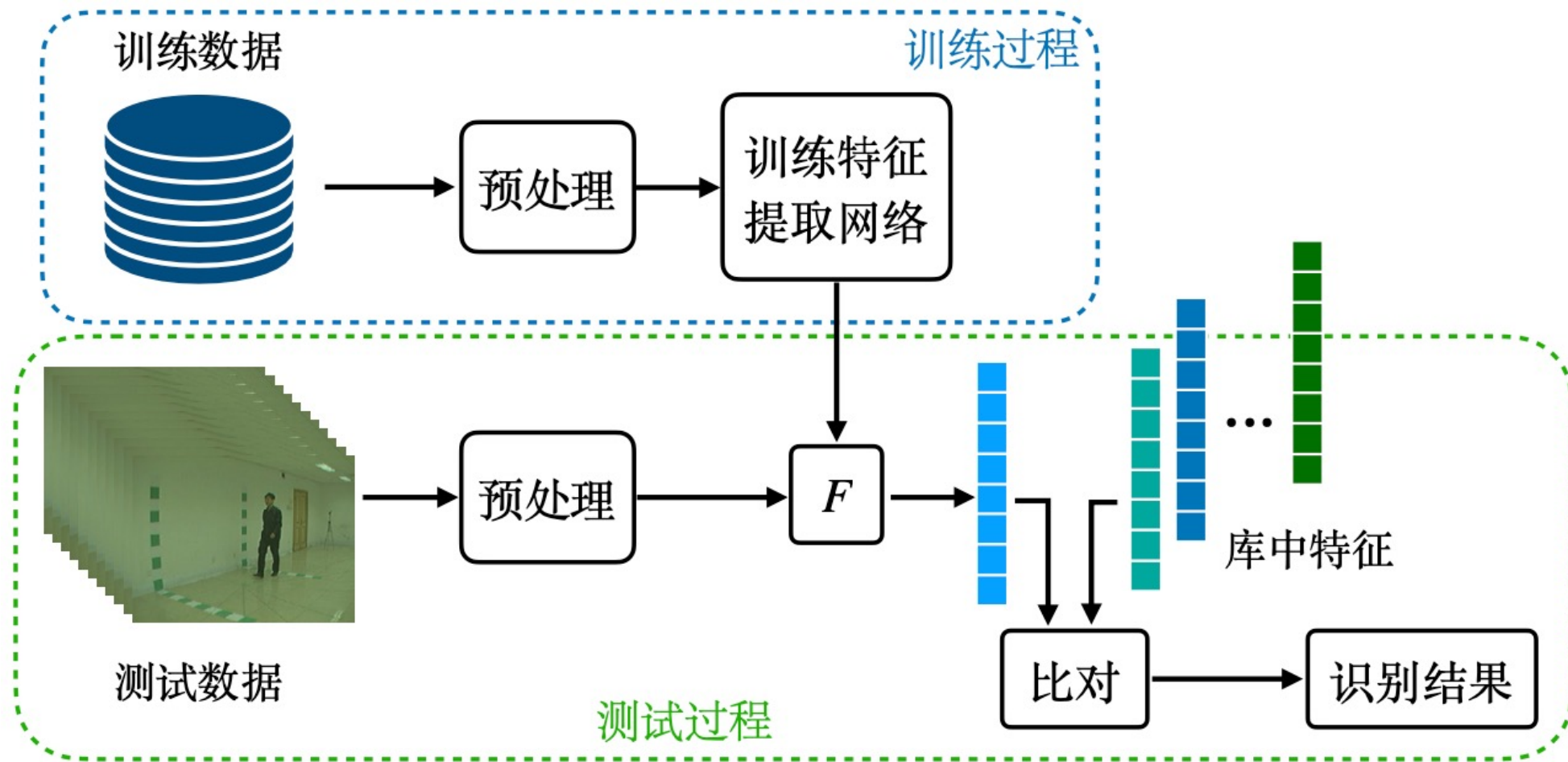
跨视角和走路情况的GEI



摄像头视角、穿衣情况、携带物品情况变化时都会很大程度上对走路姿态造成影响。



整体流程 | Pipeline





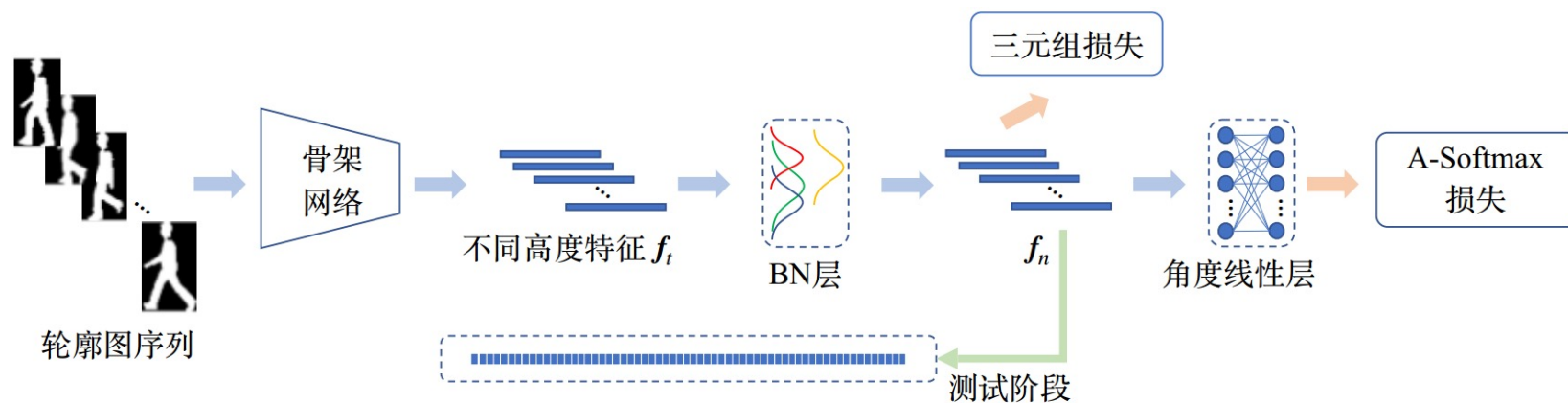
第二部分

Proposed Methods 提出方法

- ◆ 结合识别与分类的步态识别损失函数设计
- ◆ 基于注意力机制的步态识别网络的设计

结合识别与分类的步态识别损失函数设计

- 损失函数设计
- 协同优化





研究动机

■ 基于距离的损失函数

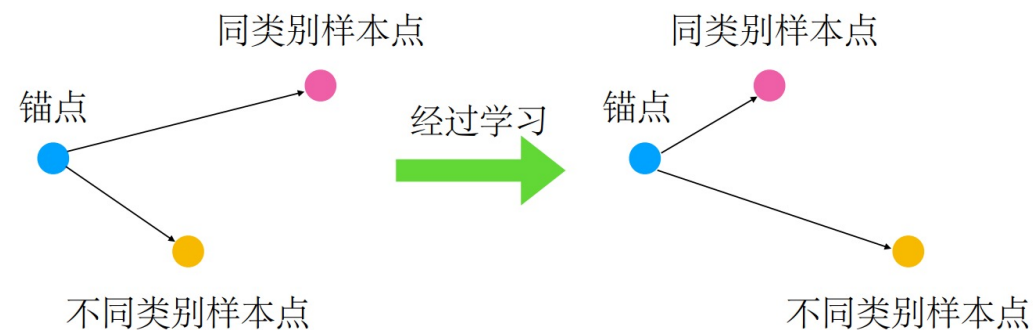
- 三元组损失

■ 基于分类的损失函数

- Softmax 损失

- A-Softmax 损失：将 W_i 归一化到长度为 1，偏置

置为 0，同时增加角度间隔

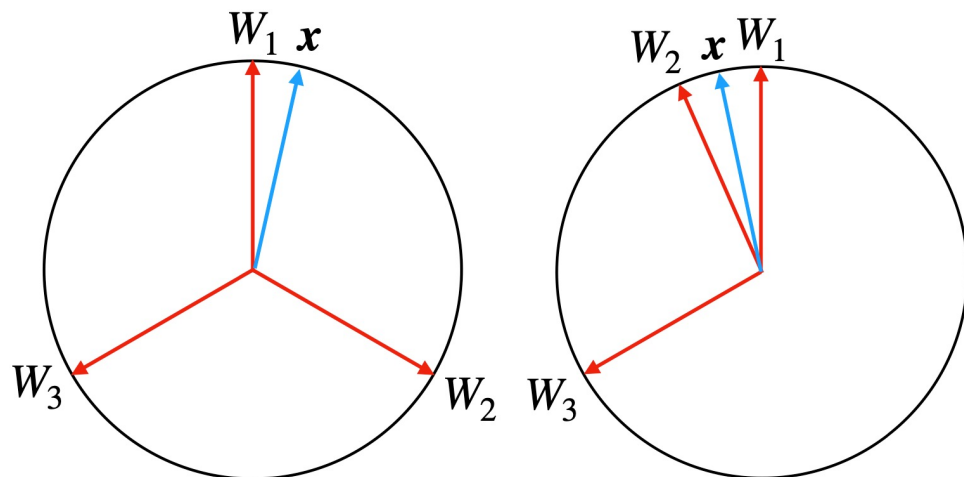


$$\mathcal{L}_{\text{Softmax}} = \frac{1}{N} \sum_{i=1}^N -\log \left(\frac{e^{\mathbf{W}_{y_i}^T \mathbf{x}_i + b_{y_i}}}{\sum_{j=1}^K e^{\mathbf{W}_{y_j}^T \mathbf{x}_i + b_{y_j}}} \right)$$

$$\mathcal{L}_A = \frac{1}{N} \sum_{i=1}^N -\log \left(\frac{e^{\|\mathbf{x}_i\| \psi(\theta_{y_i, i})}}{e^{\|\mathbf{x}_i\| \psi(\theta_{y_i, i})} + \sum_{j \neq y_i} e^{\|\mathbf{x}_i\| \cos(\theta_{j, i})}} \right)$$

研究动机

- 三元组损失
 - 优势：直接优化度量学习的目标
 - 缺陷：训练非常困难，三元组难度难以把握
- A-Softmax 损失
 - 优势：能够学习到很好的特征表达
 - 缺陷：用于分类的参数向量在特征空间中分布不均衡；在未知类别上泛化能力不强



(a) 理想情况下的特征空间

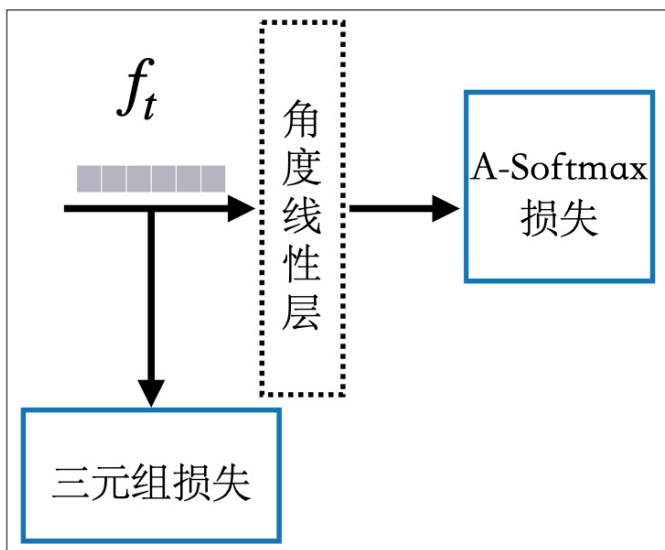
(b) 实际可能的特征空间

图 3-4: Softmax 损失函数不能直接优化特征距离

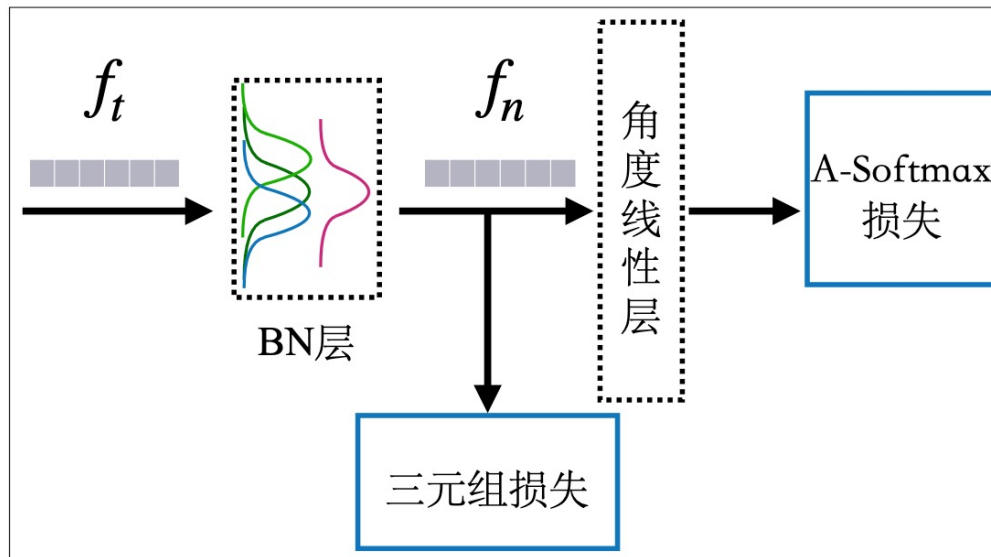


协同优化

- 两种损失很可能需要网络参数朝着不同方向更新
- 增加 BN 层，使特征始终位于同一分布下
- 降低了损失优化难度



(a) 直接优化两种损失



(b) 增加 BN 层



实验验证

损失函数	NM	BG	CL	平均识别精度
三元组损失	95.42	88.54	68.88	84.28
A-Softmax	95.37	89.08	61.15	81.87
A-Softmax + Triplet	96.10	91.88	74.36	87.45

1. 损失函数有效性

实验	NM	BG	CL	平均识别精度
e1	94.96	89.65	69.89	84.83
e2	94.62	89.84	67.55	84.00
e3	95.84	91.66	70.64	86.05
e4	95.66	91.28	72.72	86.55
e5	96.10	91.88	74.36	87.45

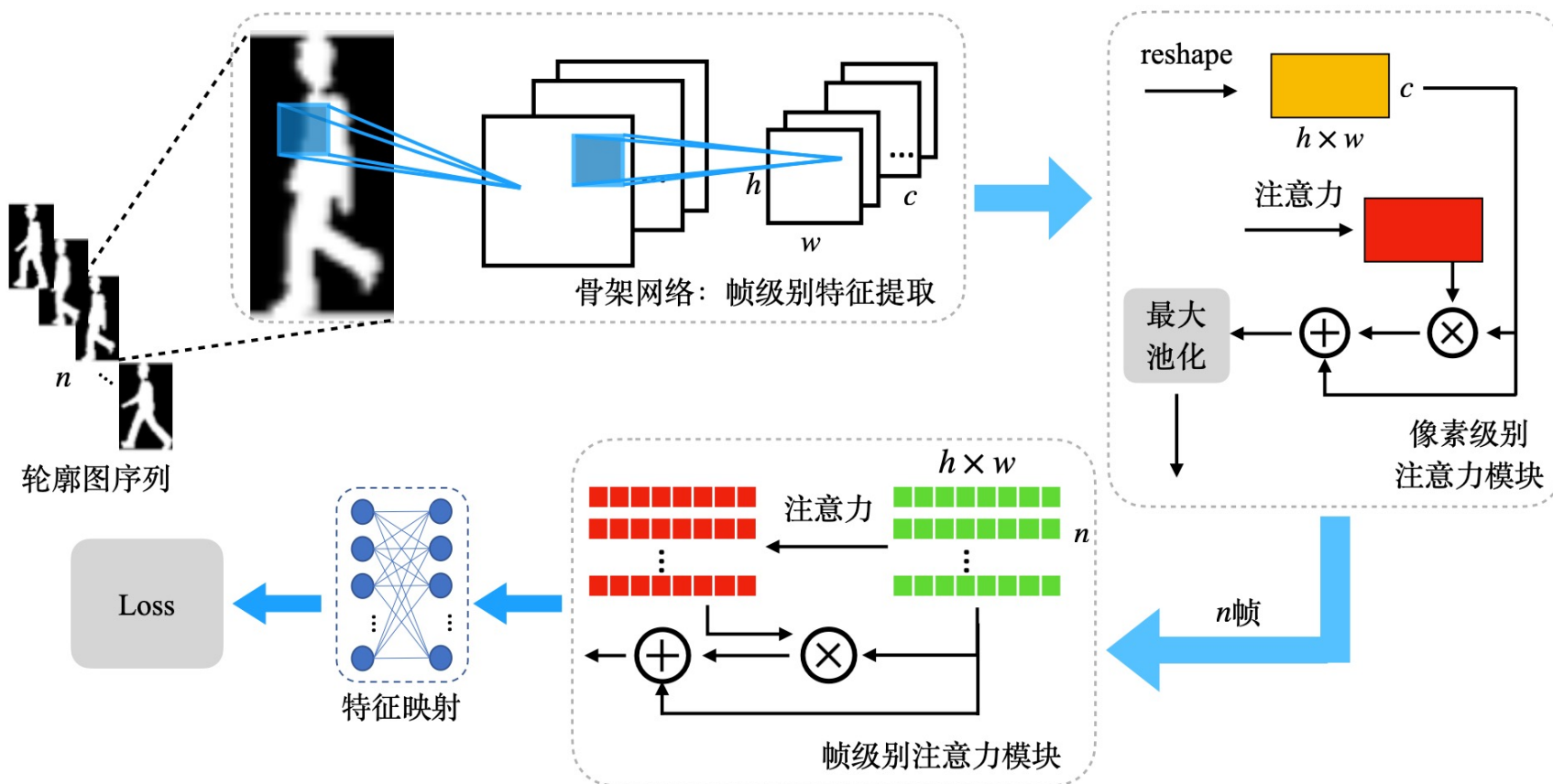
2. 协同优化

Method	NM	BG	CL	平均识别精度
MGAN ^[24]	68.1	54.7	31.5	51.4
CNN-Ensemble ^[21]	94.1	-	-	-
CNN-LB ^[21]	-	72.4	54.0	-
DisGait ^[6]	93.9	82.6	63.2	79.9
GaitSet ^[9]	95.0	87.2	70.4	84.2
ACL+local+temporal ^[8]	96.0	-	-	-
Ours	96.1	91.9	74.4	87.4

3. 方法对比



基于注意力机制的步态识别网络的设计





帧级别特征提取

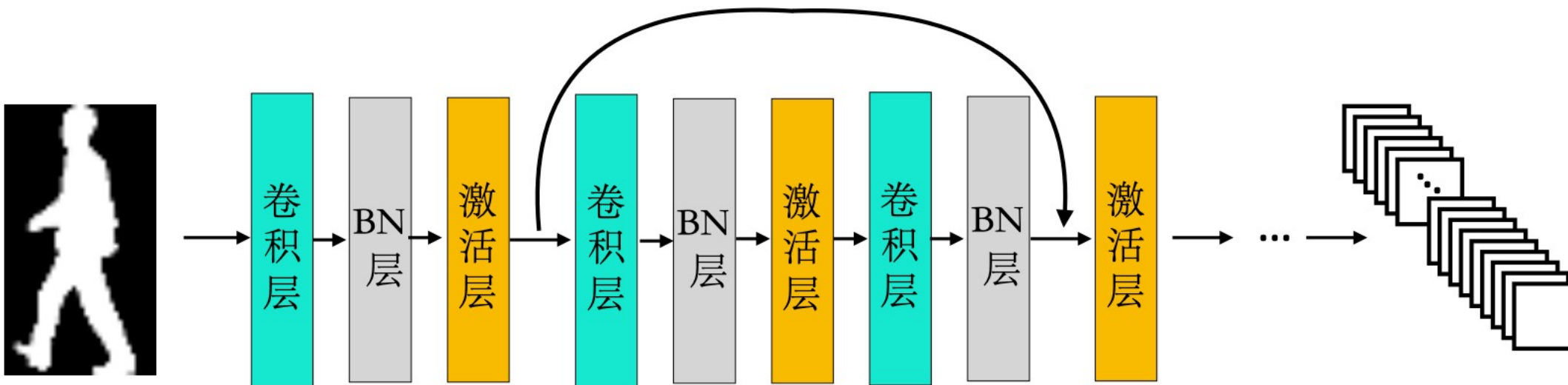


图 4-4: 特征提取网络设计



像素级别注意力机制

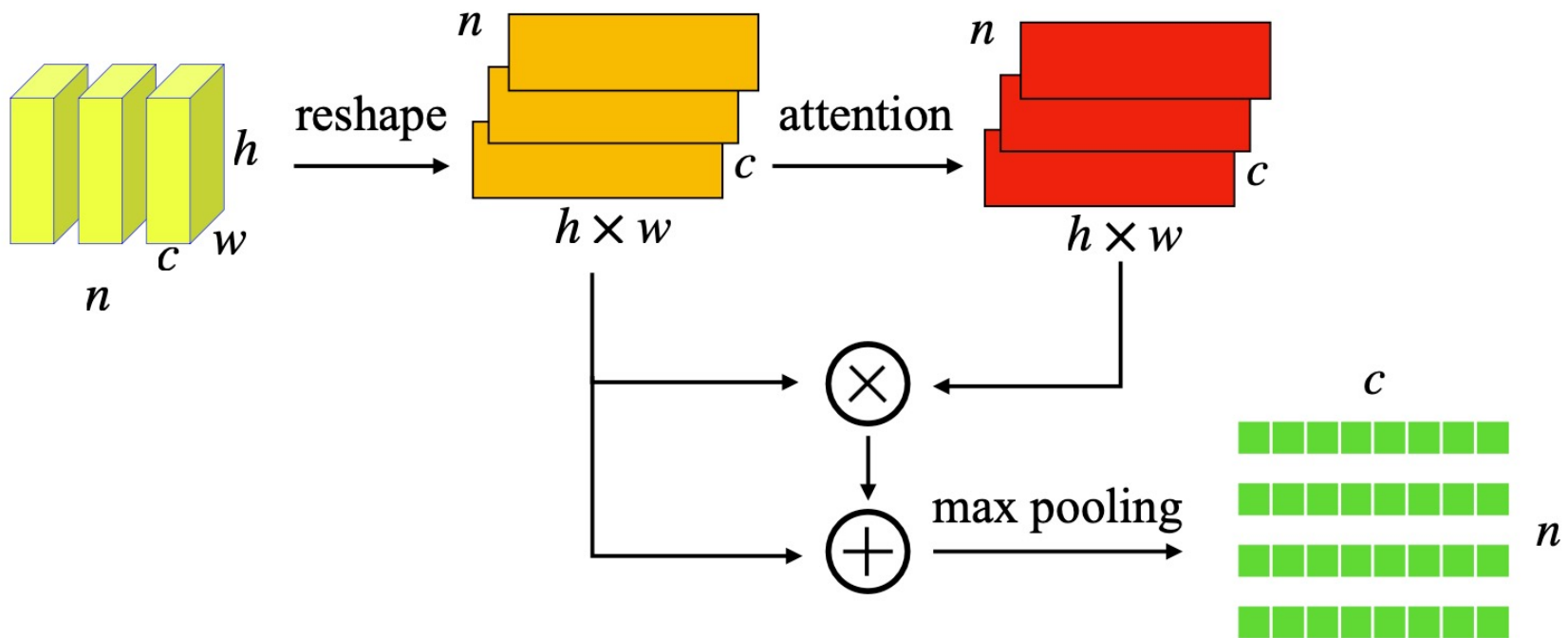


图 4-5: 像素级别注意力机制



帧级别注意力机制

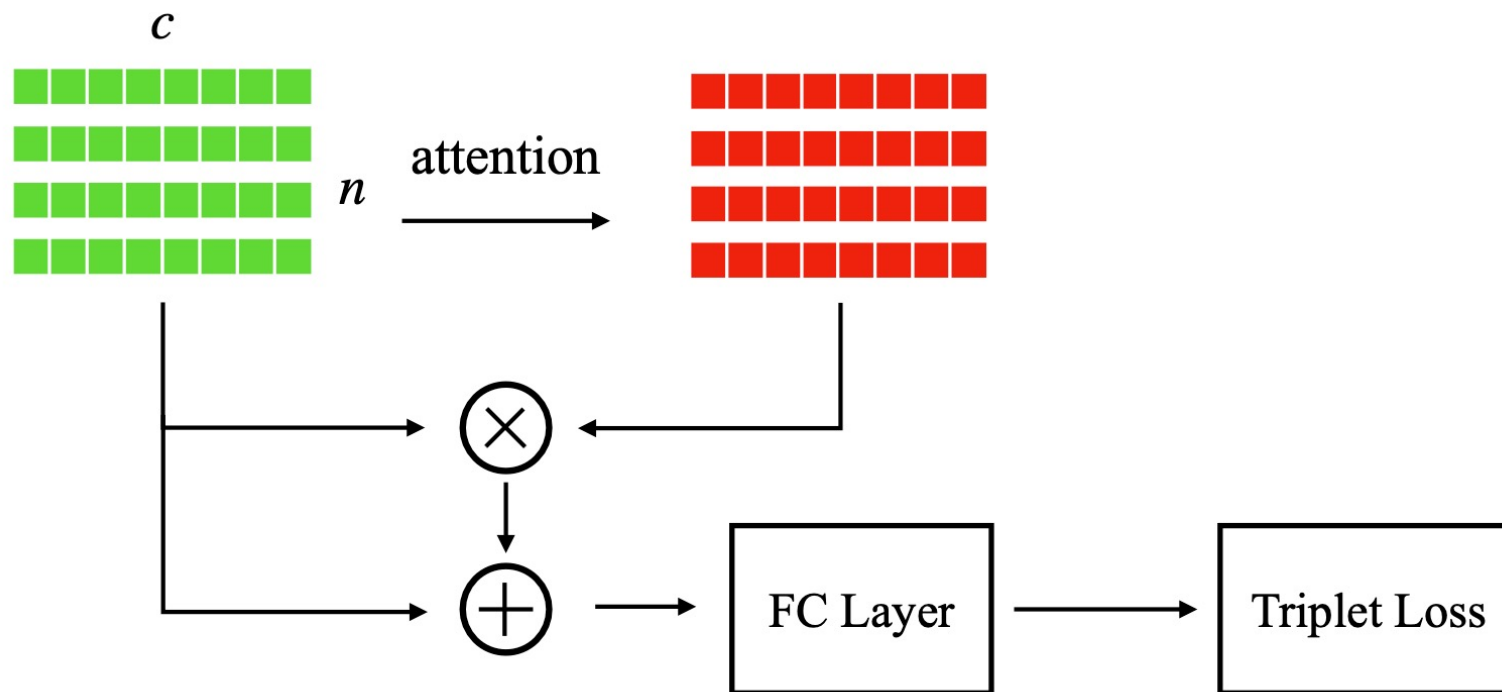


图 4-6: 帧级别注意力机制



实验验证

表 4-1: 不同注意力机制实验结果

注意力作用方式	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	平均识别精度
像素级别注意力	99.18	98.36	99.18	96.72	100.0	99.18	99.18	98.36	98.36	96.72	98.36	98.51
帧级别注意力	96.72	98.36	97.54	95.90	98.36	97.54	95.90	97.54	95.90	97.54	95.90	97.02
连续注意力	98.36	96.72	97.54	97.54	100.0	99.18	96.72	96.72	100.0	98.36	96.72	97.99
拼接注意力	98.36	95.90	99.18	97.54	95.90	97.54	97.54	96.72	98.36	98.36	97.54	97.54

表 4-2: 损失函数实验结果

损失函数	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	平均识别精度
三元组损失	98.36	97.54	99.18	97.54	98.36	97.54	99.18	94.26	97.54	100.0	100.0	98.14
A-Softmax 损失	98.36	99.18	98.36	98.36	100.0	100.0	100.0	98.36	98.36	98.36	97.54	98.81
第三章所提损失	100.0	98.36	99.18	99.18	100.0	99.18	99.18	98.36	99.18	99.18	99.18	99.18

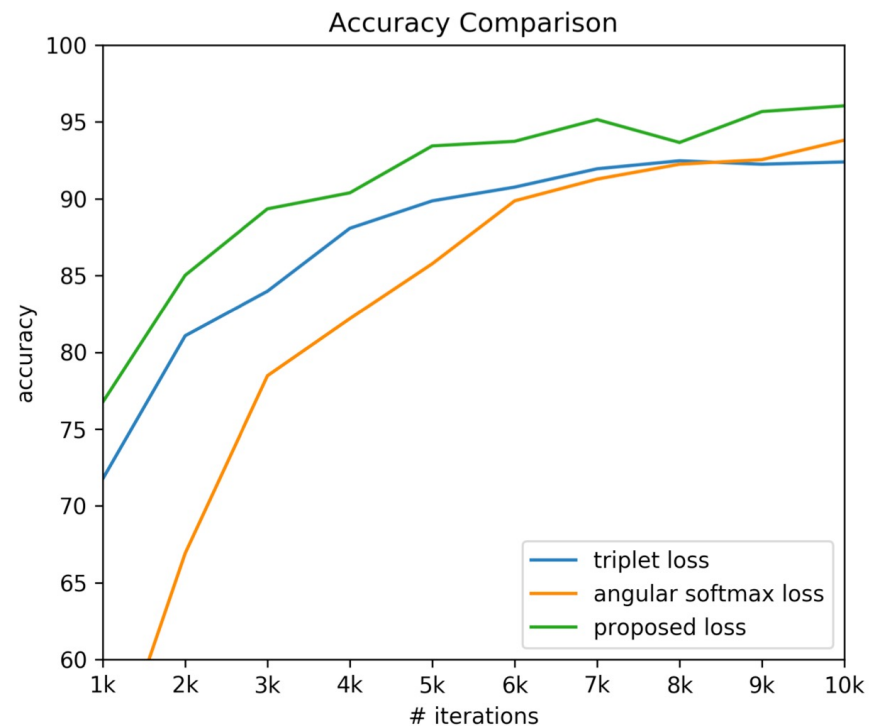


图 4-7: 识别精度变化曲线



第三部分

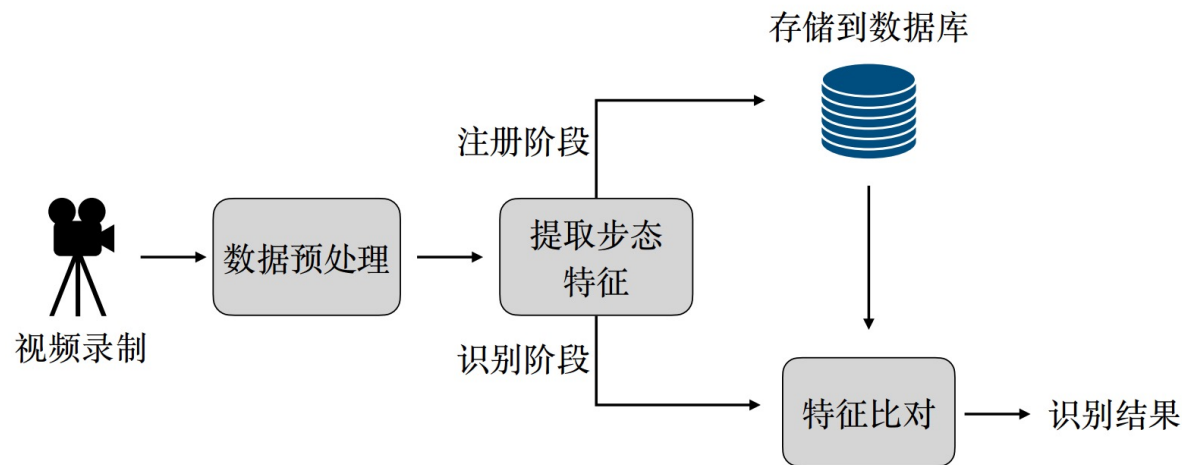
Applications

实际应用

◆ 步态识别系统

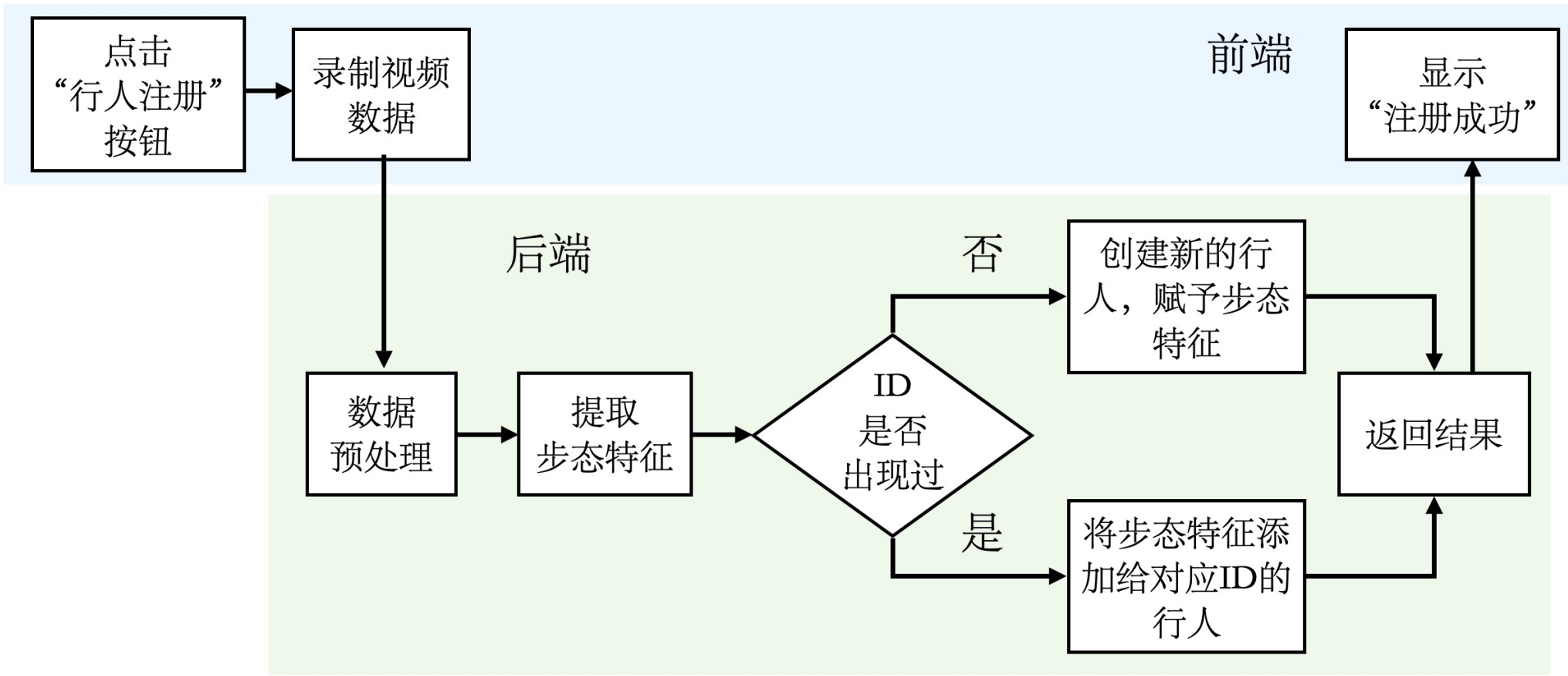
步态识别系统

- 利用走路姿态对人进行识别
- 身份验证系统的一部分
- 目的
 - 作为人脸识别、声纹识别的补充
 - 通过多种识别方式共同作用，达到更高的精度



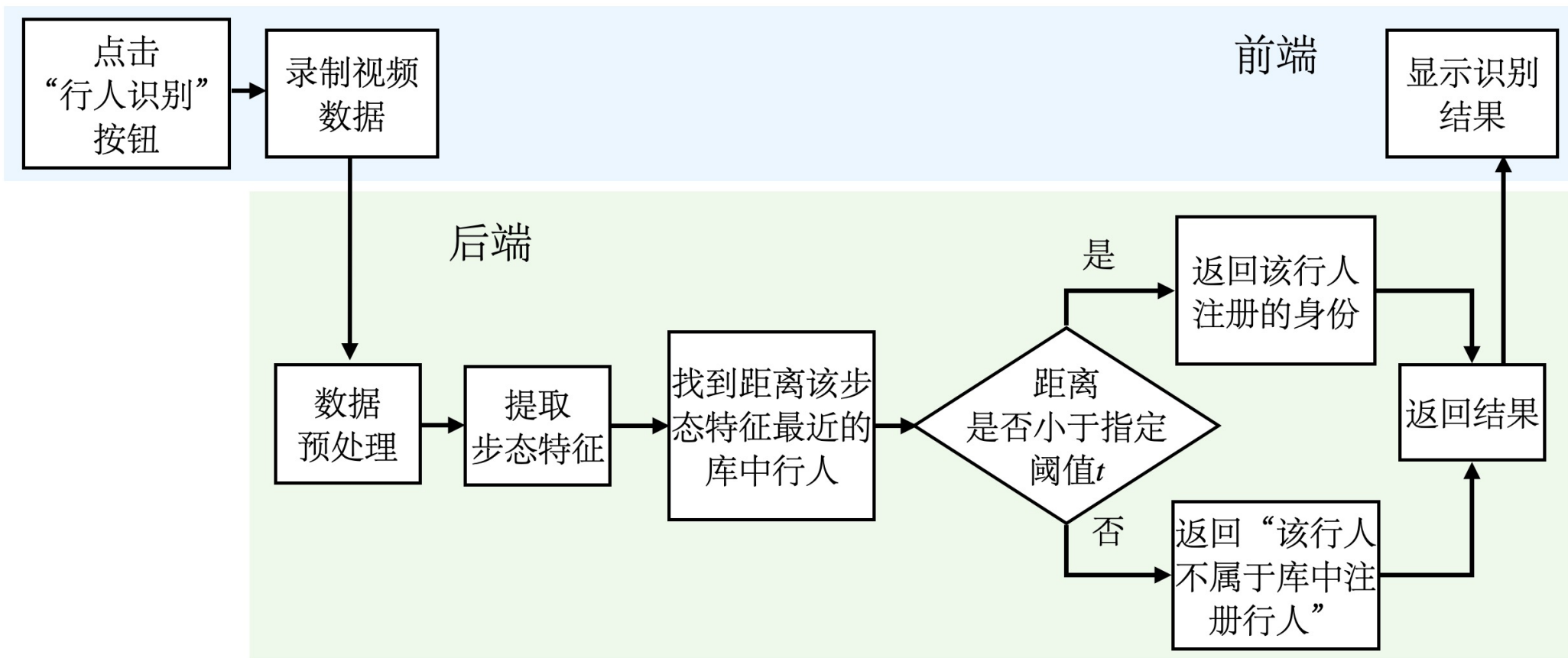


注册过程



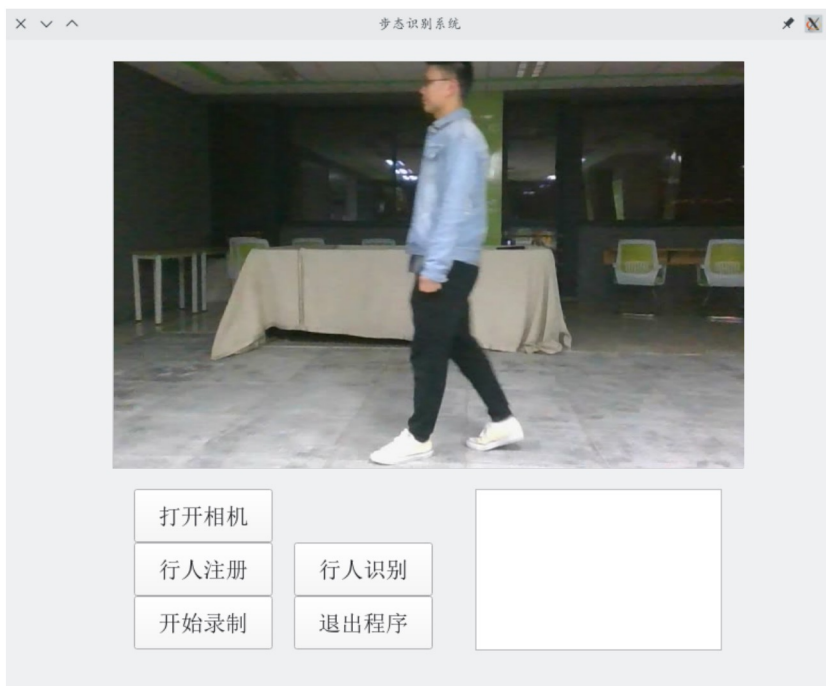


识别过程





程序界面



1. 程序主界面



2. 识别结果显示



- 从度量学习的角度，提出了一种全新的损失函数，该损失函数能够有效结合基于距离的损失与基于分类的损失，直接优化度量学习目标的同时，提取得到有代表性的特征；
- 提出了两种注意力机制的作用方式，分别是像素级别注意力机制与帧级别注意力机制。其中，像素级别注意力机制能够关注到特征图内部对于识别有帮助的信息，而帧级别注意力机制能够关注到整段序列内有区分力的信息。
- 在这两个算法的基础之上，实现了用于身份验证的步态识别系统，通过前后端分离的方式优化识别效率，该系统已经被成功应用于现实环境中。



第三部分

Summary

研究生期间工作总结



学术论文

- **Xuejian Li**, Feng Han, Jian Zhao, Furao Shen, “A More Robust Gait Recognition System Based on Metric Learning”, in Pattern Recognition(PR), under-review.

发明专利

- 申富饶, **李雪健**, 韩峰, 赵健, 一种结合两个向量嵌入空间的高精度步态识别方法。专利申请号：202110109320.0
- 陈力军, 梁雨, **李雪健**, 申富饶, 一种基于深度图像的图书馆机器人障碍识别方法。专利申请号：201810644120.3

比赛奖项

- 韩峰, **李雪健**, 赵加成, 姜少魁。2019 首届 IKCEST “一带一路” 国际大数据 竞赛, 国际二等奖



南京大學
NANJING UNIVERSITY



谢谢！

