

蘭州資源環境職業技術大學

LANZHOU RESOURCES & ENVIRONMENT VOC-TECH UNIVERSITY

本科毕业设计



题 目： 基于 VGG 的交通手势识别系统设计

学 院： 信息工程

专 业： 人工智能工程技术专业

年 级、班： 2023级人工智能工程技术(专升本)11班

学生姓名： 赵浩浩

指导教师： 尉雅晨

二〇二五年五月三十日

兰州资源环境职业技术大学学位论文原创性声明

本人郑重声明:所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。除文中已经注明引用的内容外,本论文不含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体,均已在文中以明确方式标明。因本学位论文引起的法律后果完全由本人承担。

学位论文作者签名: 赵浩浩
签字日期: 2025 年 5 月 20 日

基于 VGG 的交通手势识别系统设计

赵浩浩

(兰州资源环境职业技术大学信息工程学院人工智能工程技术专业 2023 级人工智能工程技术
(专升本) 11 班, 甘肃 兰州 730021)

指导教师: 尉雅晨

[摘要]随着人工智能技术的不断进步, 无人驾驶技术逐渐进入公众视野。在实际交通行驶环境中, 无人驾驶仍无法避免交通意外事件的发生。传统手势识别技术包括模板匹配、特征提取以及基于传感器的方法, 这些方法受到手势变化多端和环境因素的干扰, 往往无法达到实际应用的要求。鉴于此, 本文采用视觉几何组网络 (Visual Geometry Group, VGG) 在图像空间特征提取方面的卓越性能并结合卷积姿势机在人体关键点定位上的优势, 以及长短期记忆网络对图像时间序列信息的补充设计并实现了基于 VGG19 网络的交通手势识别系统。本文通过模型设计、数据选取、模型训练, 选择最佳模型构建基于 VGG19 的交通手势识别系统。系统测试结果表明杰卡德 (Jaccard) 的微平均分数总体接近 82.81%, 并将模型与时空图卷积网络和长短期记忆网络、骨骼长度角度网络和长短期记忆网络作对比实验, 实验结果表示本文模型在测试集中的识别性能具有一定优势, 能够有效融合空间、姿态和时序特征, 从而在复杂交通环境下实现准确的手势识别。

[关键词]VGG19, 交通手势识别, 卷积姿势机, 长短期记忆网络

Design of traffic gesture recognition system based on VGG

Haohao Zhao

(Grade 2023, Class 11, Major of Artificial Intelligence Engineering Technology , School of Information Engineering, Lanzhou Resources & Environment Voc-Tech University, Lanzhou 730021, Gansu Province)
Tutor: Yachen Wei

Abstract: With the continuous progress of artificial intelligence technology, driverless technology is gradually entering the public eye. In the actual traffic driving environment, driverless still cannot avoid traffic accidents. Traditional gesture recognition techniques include template matching, feature extraction, and sensor-based methods, which are often unable to meet the requirements of practical applications due to the many variations of gestures and the interference of environmental factors. In view of this, this paper adopts the excellent performance of Visual Geometry Group (VGG) network in image spatial feature extraction and combines the advantages of convolutional poser in human key point localization, as well as the long and short-term memory network to supplement the image time series information to design and implement a traffic gesture recognition system based on VGG19 network. In this paper, we select the best model to construct the VGG19-based traffic gesture recognition system through model design, data selection, and model training. The system test results show that Jaccard's (Jaccard) micro-average score is close to 82.81% overall, and the model with spatio-temporal map convolutional network and long and short-term memory network, bone length angle network and long and short-term memory network for comparison experiments, the experimental results indicate that this paper's model in the test set of the recognition performance has a certain advantage, and can effectively fuse spatial, gesture and temporal features, so as to achieve accurate hand gesture recognition under complex gesture recognition in complex traffic environments.

Key words:VGG19, traffic gesture recognition, CPM, LSTM

目录

摘要	I
Abstract	II
1 绪论	1
1.1 研究背景及意义	1
1.2 国内外研究现状	2
1.3 研究内容及安排	2
2 相关理论及技术基础	4
2.1 VGG19 网络模型结构	4
2.2 卷积姿势机(CPM)	4
2.3 长短期记忆网络(LSTM)	5
3 基于 VGG19 的交通手势识别模型设计	6
3.1 需求分析	6
3.2 数据集	6
3.3 图像特征提取	7
3.4 手势识别	7
3.5 姿态估计	8
3.6 模型设计	8
3.7 模型评估	9
4. 系统架构设计与功能实现	11
4.1 系统总体架构	11
4.2 实验环境	11
4.3 系统界面与功能演示	12
5 系统测试	15
6 总结与展望	16
6.1 总结	16
6.2 展望	16
参考文献	17
致谢	20

1 绪论

随着人工智能（Artificial Intelligence, AI）技术的飞速发展，无人驾驶技术正逐步从理论研究走向实际应用，但在复杂的交通环境中，仍需要依靠交警的手势来应对特殊情况。然而，由于手势的多样性以及环境因素的干扰，传统手势识别技术难以满足实际需求。因此本文提出了一种基于深度学习的模型设计融合 VGG19 的空间特征提取、卷积姿势机（Convolutional Pose Machines, CPM）的人体关键点定位以及长短期记忆网络（Long Short Term Memory, LSTM）的时序建模技术，构建了一个“空间-姿态-时序”的三级特征融合模型，有效解决了在复杂交通环境下手势识别问题。

1.1 研究背景及意义

1.1.1 研究背景

交通安全是城市可持续发展和保障人民生命财产安全的重要组成部分。随着汽车的持续增长和交通环境的日益复杂，传统的交通手势识别面临着诸多挑战。此外，手势识别本身极易受到观察者主观性的影响，准确性难以保证。人工智能技术的飞速发展为解决上述问题提供了新的思路和方法。在众多的深度学习模型中，卷积神经网络在图像识别领域取得了巨大的成功。其中，由牛津大学视觉几何组提出的 VGG19 网络以其简洁的网络结构和优秀的性能，成为了图像分类任务的基准模型之一。VGG19 网络通过堆叠小卷积核的卷积层和池化层，能够有效地捕获图像中的层次化特征，并在大规模图像数据集上展现出卓越的性能。现有的交通管理主要依赖于交通信号灯、交通标志、交警手势指挥等方式进行信息传递，对于需要即时、灵活沟通的场景，存在一定的局限性。交通场景中的手势可能受到执行者姿态、遮挡、视角变化等多种因素的影响，使得手势的识别更具挑战性。因此，本文设计并实现基于 VGG19 的交通手势识别系统，为复杂情况下交通指挥提供更多的可能性。

1.1.2 研究意义

随着人工智能和自动驾驶技术的发展，智能交通系统已成为重要的研究方向。在此背景下，设计并实现准确可靠的交通手势识别系统显得尤为重要，其对于无人驾驶车辆在复杂交通环境中的安全性和效率具有显著意义。卷积神经网络在图像识别领域展现出卓越性能，VGG19 网络作为其中的典型代表，在图像特征提取方面具备优势。因此，结合深度学习模型构建高性能的交通手势识别系统对于智能交通的建设与发展具有显著的支撑价值。

传统的手势识别技术，如模板匹配和特征提取等，在复杂环境和手势变化的情况下表现出局限性。本文融合 VGG19 网络与卷积姿势机以及长短期记忆网络，对手势的空间特征和时序动态信息捕捉，能够在多变环境下保持较高的识别准确率。相较于传统方法，该模型在关键点定位、动作连续性方面具有显著优势。

1.2 国内外研究现状

1.2.1 国内研究现状

在国内，基于视觉的手势识别方法借助摄像头采集图像信息，随后将这些信息传输至计算机。计算机对图像的颜色、像素分布、形状轮廓进行分析处理。2020年，张丞等人^[1]设计了融合空间上下文和时序特征的交警指挥手势识别机(Chinese Traffic Police Gesture Recognizer, CTPGR)该实验可以快速准确地识别交警指挥手势。2021年，彭雪萍等人^[2]设计并开发一款关于交通安全教育的体感游戏，能够让儿童在进行肢体游戏的同时进行安全知识的学习，提升学习效果，减少安全事故发生。2022年，阳嘉佳等人^[3]采用 OpenCV 图像视频处理、目标检测算法检测车内人员行为，降低了交通事故的发生率。2022年，张丞等人^[4]提出高度分层分区的图卷积交警手势识别方法，该方法的识别准确率显著提高。2023年，程贝芝等人^[5]提出了一种结合全局上下文信息的交警手势识别方法具有更强的可移植性。2023年，梁钊铭等人^[6]提出了基于手部特征点空间位姿信息的动态手势数据集构建 方法验证了该场景下特定数据采集方式与算法模型组合在装配动作智能识别上的有效性。2023年，李泰国等人^[7]提出一种改进的 YOLOv5s 列车驾驶员手势识别算法该算法更好地提升预测框精准度、加强模型的识别性能。2024年，鲁杰伟等人^[8]提出了一种改进 YOLOv5s 模型的手势识别算法。提高了手势识别的准确率和效率。2024年，孔皓^[9]对人机交互技术进行概述，分析人机交互技术在汽车中的应用以提升汽车性能，改善汽车驾驶体验。2024年，杜兵等人^[10]提出了一个 GCPMAGRU 模型进行交通警察手势识，有效改进了人体关键点定位的问题。

1.2.2 国外研究现状

在国外，手势识别方法主要集中于多模态传感器融合与三维姿态分析，技术路径涵盖数据手套、多点触控交互及肌电信号，通过获取手部三维姿态信息实现。2014年，S.Swainali 和 PG Chilveri^[11]使用加速度计传感器进行交通信号灯控制系统的手势识别。2018年，G.Yasmeen 等人^[12]使用 ti 传感器标签实现交通控制系统的高效手势识别为控制设备提供了虚拟环境。2019年，P.Sharma 和 N.Sharma^[13]设计了“手势识别系统”，2020年，MD Artemov 等人^[14]使用神经网络方法设计手势语言识别软硬件综合体。2020年，J.Wiederer 等人^[15]设计了自动驾驶汽车的交通控制手势识别。在探讨基于 VGG 网络的交通手势识别系统的研究现状时，可以观察到学术界和工业界均在努力克服一系列挑战。

1.3 研究内容及安排

1.3.1 研究内容

本文设计并实现基于 VGG19 网络引入卷积姿势机和长短期记忆网络的交通手势识别系统，该系统在复杂交通场景中对手势识别的精确度和响应速度显著提高。本文详细阐述了 VGG19 网络的架构，以及 CPM 与 LSTM 的功能和工作原理。采用专门的交通手势数据集进行模型训练，并选用适当的性能评估指标进行模型性能的评估，并通过与其他模型的对比分析说明模型的整体性能稳定。最后对交通手势识别系统进行设

计实现及功能展示。

1.3.2 章节安排

本文共由 6 章组成，组织结构与安排如下

第 1 章 绪论。介绍了研究背景和意义，同时介绍了手势识别方法的国内外研究现状，最后给出了本文的主要研究内容和章节安排。

第 2 章 相关理论及技术基础。简要介绍了相关概念以及理论基础。详细介绍了本设计模型涉及的 VGG19 网络、卷积姿势机及长短期记忆网络的工作机制，并对基础理论知识进行了阐述。

第 3 章 基于 VGG19 的交通手势识别模型设计。本设计采用 VGG19 提取 AI Challenger 数据集上图像的特征，然后用 CPM 和部分亲和场（Part Affinity Fields，PAF）在数据集上训练精确的人体关键点检测模型。随后，把 Police gesture dataset 数据集输入到训练好的关键点模型中预测空间特征序列。最终，用 LSTM 对预测的空间特征序列进行分类，构建完整的交通手势识别模型。

第 4 章 系统架构设计与功能实现。将交通手势识别模型用于系统实现，对系统总体架构进行设计，并对系统的应用模块（用户交互）、数据处理模块（数据预处理与特征提取）、模型模块和可视化模块四部分进行详细的设计与功能实现进行描述。

第 5 章 系统测试。对系统注册登录功能测试和交互功能测试，系统功能完备。

第 6 章 总结与展望。对研究的成果进行总结，并提出下一步研究。

2 相关理论及技术基础

2.1 VGG19 网络模型结构

VGG19 是卷积神经网络（Convolutional Neural Network, CNN）中的一个标志性模型，其设计原理基于“深度堆叠”和“小卷积核高效特征提取”。该模型通过多个 3×3 的小卷积核堆叠来取代大尺寸卷积核，利用“感受野等效性”原理，即两个 3×3 卷积核的堆叠相当于一个 5×5 卷积核的感受野，三个 3×3 卷积核等同于一个 7×7 卷积核的感受野，这样既减少了参数数量，又增强了通过增加非线性变换的特征学习能力。VGG19 网络结构由输入层、卷积层组、池化层和全连接层组成。输入层负责接收 $224 \times 224 \times 3$ 的图像数据，然后通过多组“卷积层 + ReLU 激活 + 池化层”的模块逐层提取图像特征。前两组卷积层组各自包含 2 个 3×3 卷积层，输出通道数从 64 增加到 128；后三组卷积层组则包含 3 至 4 个卷积层，通道数进一步增加至 256、512，从而实现从边缘、纹理等基础特征到物体部件、整体语义的层次化提取。池化层采用 2×2 的最大池化操作，以降低特征图的尺寸并减少计算量。最终，全连接层整合高层特征，并通过 Softmax 函数完成分类任务，构建起从“数据输入—特征层次提取—分类输出”的完整流程，网络结构图如图 2.1 VGG19 结构所示。

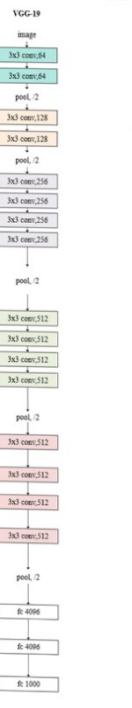


图 2.1 VGG19 结构

2.2 卷积姿势机

卷积姿势机将深度学习应用于人体姿态分析。通过多阶段全卷积网络，从输入图像中逐级细化地预测人体关节点热力图。支持单人及多人姿态估计，具有较大感受野以处理遮挡问题。以多阶段迭代优化架构为核心。首先，对输入图像进行初始特征提取，利用基础卷积网络（如 VGG）获取包含边缘、纹理等底层信息的特征图。随后进入多阶段处理过程，每个阶段以前一阶段输出的关键点置信图和初始特征图为输入，

通过卷积操作生成更精确的置信图。置信图的每个通道对应一个关键点，像素值反映该位置作为关键点的置信度。此外，CPM 引入中间监督机制，在每个阶段设置损失函数，计算预测置信图与真实关键点的误差，总损失由各阶段损失累加，既避免误差累积，又加速模型训练收敛。多阶段迭代使模型逐步修正关键点预测，尤其在复杂姿态、遮挡场景中，通过多轮特征融合优化，提升定位精度；各阶段结合图像全局与局部特征，即使关键点被遮挡，也能依托周围关联信息推断位置；独立阶段监督信号解决梯度消失问题，让模型早期接收反馈，加快收敛，提升训练效率。

2.3 长短期记忆网络

长短期记忆网络作为 RNN 的改进，通过“细胞状态”（信息传递主干道）与遗忘门、输入门、输出门三大门控机制，遗忘门控制细胞状态信息的保留或遗忘，输入门处理当前输入的候选状态及更新量，输出门基于更新后的细胞状态生成隐藏状态。LSTM 解决传统 RNN 长序列依赖难题（缓解梯度消失 / 爆炸），利用门控动态筛选关键信息，且广泛适用于自然语言处理、时间序列预测等时序任务，有效提升序列建模效果。

3 基于 VGG19 的交通手势识别模型设计

3.1 需求分析

系统通过 VGG19 实现对交通手势的识别，有效提升在交通指挥方面的效率及安全性，为交通管理部门、智能交通系统提供自动化的手势识别的解决途径。系统的功能性需求包括用户的注册与登录，并设置权限划分；能够支持交通场景视频上传、处理与分析；将分析结果进行可视化展示，并支持历史记录查询功能。非功能性需求要保障系统质量与运行性能，识别准确率不低于 70%，具备安全机制，保障用户数据、识别结果的保密性与完整性；系统开发语言选用 Python，深度学习框架采用 PyTorch，后端框架使用 Flask，数据库选用 MySQL 实现用户数据管理；系统架构采用前后端分离，使系统具有一定的兼容性和可维护性。

3.2 数据集

在基于 VGG19 的交通手势识别系统中，数据集收集方法的科学性和多样性对于训练出准确、鲁棒的模型至关重要。

(1) Police gesture dataset。数据是一个由 21 个视频组成的公开数据集，其中，包括 11 个训练集和 10 个测试集，数据集的总视频时长约 2 小时，视频文件是一些交通警察的混淆手势。这是目前最符合我国交通法则中国交警的 8 种指挥手势，它是由几名不同的受试者身穿交通警察服装，在不同场景下采集的交通警察指挥手势的视频片段。

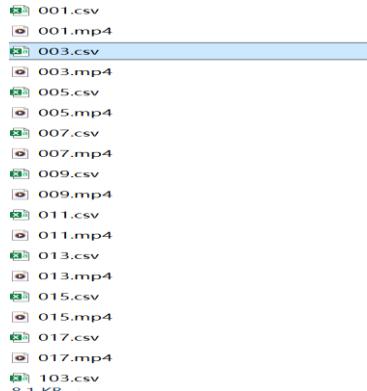


图 3.1 Police gesture dataset

(2) AI Challenger。数据集中包含人体骨骼关键点数据集，此数据集包含 30 万张图片，标注 70 万量级的人体关键点，远超 MSCOCO 和 MPII 数据集，覆盖超过 100 种复杂生活场景，包含密集遮挡、多姿态变化等挑战。本设计选取 1000 张图片进行应用，其中 700 张用于训练，300 张用于测试。如图 3.2 AI Challenger 数据集所示，图片中关键点数据通过热图显示，标注了 14 个关键点，包括头部、肩部、肘部、手腕、臀部、膝盖等。因此，数据集训练的模型在处理高相似性人体结构和环境干扰问题时能表现出较好的稳定性。

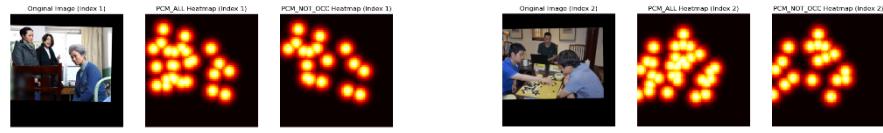


图 3.2 AI Challenger 数据集

3.3 图像特征提取

图像特征提取阶段采用 VGG19 网络作为主干特征提取器，对输入的 AI Challenger 图像进行多层次空间特征编码。在处理过程中，图像首先被统一调整至 $224 \times 224 \times 3$ 的尺寸，以符合 VGG19 网络的输入规范。随后通过若干个由 3×3 卷积核、ReLU 激活函数及最大池化组成的卷积模块，逐步提取图像的特征。VGG19 通过其深层结构实现了从低层感知到高层语义的有效转换，有助于后续模块捕捉与手势相关的显著视觉特征。网络中卷积层的堆叠结构不仅增强了非线性表达能力，还在保证参数可控的前提下提升了模型对图像局部细节的感知能力。最终得到的高维特征图作为交通手势识别系统中的核心表示，为后续动作识别模块提供稳定且富含判别信息的基础输入。如图 3.3 特征提取所示

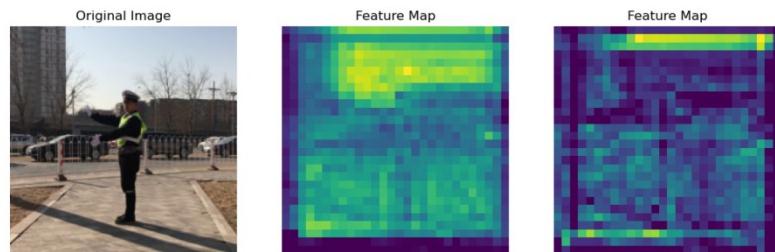


图 3.3 特征提取

3.4 手势识别

本文模型融合 LSTM 进行手势识别，能够有效地处理序列数据，并捕捉手势的动态变化。在手势识别过程中，首先需要将提取的人体关键点信息转换为骨骼长度和角度特征。通过计算关键点之间的距离和角度，得到骨骼长度和角度特征，并将其作为模型的输入。具体来说，通过计算关键点之间的欧几里得距离得到骨骼长度，通过计算向量之间的夹角得到骨骼角度。具体计算公式如下：

欧几里得距离公式如式(3.1)所示：

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (3.1)$$

点积公式如式(3.2)所示：

$$\vec{a} \cdot \vec{b} = a_x \times b_x + a_y \times b_y \quad (3.2)$$

差积公式如式(3.3)所示:

$$c_z = a_x \times b_y - a_y \times b_x \quad (3.3)$$

夹角的余弦值公式如式(3.4)所示:

$$\cos\theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \times \|\vec{b}\|} \quad (3.4)$$

夹角的正弦值公式如式(3.5)所示:

$$\sin\theta = \frac{c_z}{\|\vec{a}\| \times \|\vec{b}\|} \quad (3.5)$$

最终将转换后的骨骼长度和角度特征输入到 LSTM 模型中进行处理。在模型训练过程中，采用交叉熵损失函数来优化模型的参数，以提高模型的分类准确性。同时，为了防止过拟合，还在模型中添加了 Dropout 层，随机丢弃部分神经元，以增加模型的泛化能力。

3.5 姿态估计

本文通过分析图像中的人体特征，确定人体各个部位的位置和姿态。姿态估计主要基于人体关键点检测技术，通过检测人体的关键点位置来推断人体的姿态。姿态估计通过采用 CPM 与 PAF 进行人体关键点检测，通过输出人体关键点的置信图和部分亲和场，从而确定人体各个关键点的位置。

在姿态估计过程中，首先需要对输入的图像进行预处理，以确保图像数据的一致性和可用性。然后，通过 CPM 对预处理的图像进行处理，得到人体关键点的位置。在得到人体关键点位置后，为了增强关键点之间的结构建模能力，引入 PAF，用于编码关键点对之间的空间连接和方向信息，并对关键点之间的关系进行分析，推断出人体的姿态如图 3.4 人体关键点及部分亲和场所示。

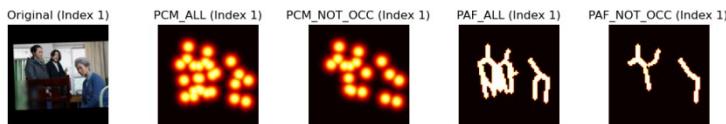


图 3.4 人体关键点及部分亲和场

3.6 模型设计

交通手势识别系统的模型设计如图 3.5 模型设计所示，在结构上融合了空间特征提取、姿态估计与时序建模三大模块，构建了一个端到端的深度学习识别框架。整体架构以 VGG19 卷积神经网络作为基础特征提取单元。VGG19 通过多层 3×3 卷积核与 ReLU 非线性激活函数，配合最大池化操作，从输入的 AI Challenger 数据集的原始图像中逐步提取低级边缘特征至高级语义特征，形成多尺度、多层次的图像表征。随后，提取得到的特征图输入至 CPM 模块中进行姿态估计。CPM 采用多阶段全卷积网络结构，在每一阶段预测关键点热力图并不断优化定位精度，有效缓解了遮挡与姿态模糊所带来的识别困难。在 CPM 输出的关键点后，引入 PAF 以确定关键点间的关系，通过对骨

架连通性的建模，增强模型对肢体结构的整体感知能力。Police Gesture Dataset 被用于训练基于 PAF 的关键点模型。对 PAF 模块处理后得到的人体骨架关键点数据，进一步提取出骨骼间的欧氏距离、角度信息和连接向量特征，构建成反映动作动态结构的时序特征序列。该特征序列作为输入传递至 LSTM 中，LSTM 通过其门控单元捕捉手势动作在时间维度上的演化规律，实现对序列信息的长期依赖建模，进而完成交通手势的准确分类。

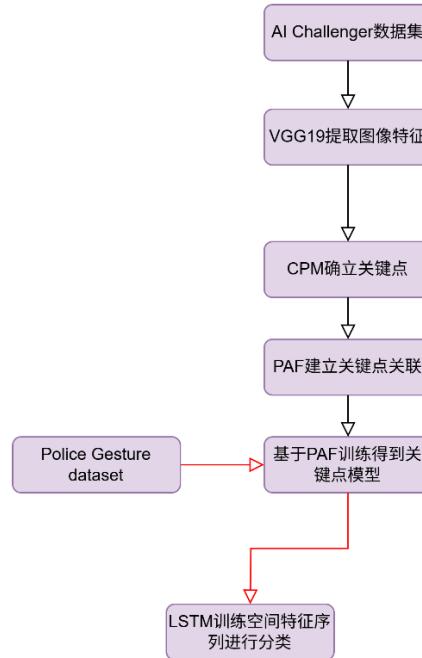


图 3.5 模型设计

3.7 模型评估

本文采用了 Jaccard 分数来评估手势识别模型的性能，评价公式如式(3.6)所示。Jaccard 分数是一种常用的评估指标，用于衡量两个集合之间的相似度。在手势识别任务中，Jaccard 分数可以用来衡量模型预测的手势类别与真实手势类别之间的相似度。在模型评估过程中，将测试集输入到训练好的模型中，得到模型的预测结果。接着将模型的预测结果与真实标签进行比较，计算 Jaccard 分数。从表 3.1 可以看出所有视频 Jaccard 的微平均分数总体接近 82.81%。

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (3.6)$$

表 3. 1 Jaccard 得分

视频文件名	Jaccard 得分
002.mp4	88.49%
004.mp4	84.68%
008.mp4	84.25%
010.mp4	80.92%
012.mp4	84.39%
014.mp4	72.26%
016.mp4	79.38%
018.mp4	79.51%
102.mp4	88.69%
104.mp4	84.89%
所有视频	82.81%

本文将具有时空图卷积网络的 STGCN+LSTM、有骨骼长度与角度信息的 BLA+LSTM、与 VGG+CPM+LSTM 三种模型进行对比分析，实验对比结果如图 3. 6 所示。图中展示了 STGCN+LSTM、BLA+LSTM 和 VGG+CPM+LSTM 三种模型在视频中 Jaccard 指数情况，其中横坐标表示视频片段编号，纵坐标表示对应的 Jaccard 指数，用于衡量模型在手势识别任务中的性能表现。

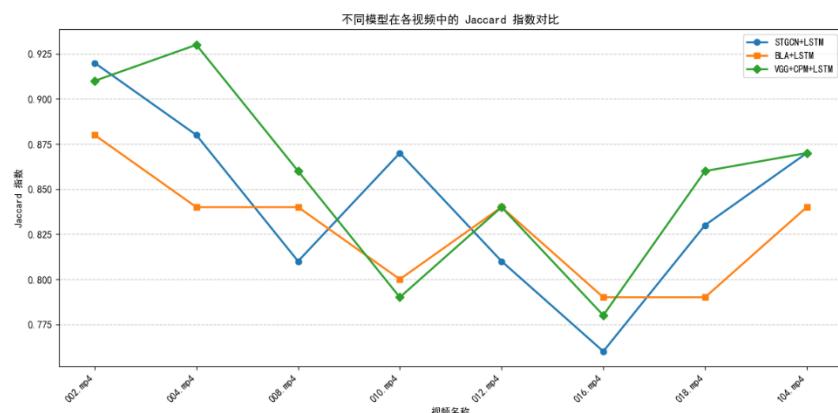


图 3. 6 模型对比

观察对比结果可知，VGG+CPM+LSTM 在大多数视频中取得了较高的 Jaccard 指数，特别是在遮挡物多的 "002_mp4" 和 "004_mp4" 视频中，VGG+CPM+LSTM 的识别性能显著优于 STGCN+LSTM 和 BLA+LSTM，这表明 VGG+CPM+LSTM 在处理纯序列数据，尤其是具有较稳定动作模式的手势识别任务中，依然具备较强的建模能力。STGCN+LSTM 在部分视频片段中表现波动较大，在光照比较强的 "010_mp4" 和 "016_mp4" 中，Jaccard 指数出现明显下降。这种现象可以归因于 STGCN+LSTM 对时空特征的敏感性，当骨骼动作变化复杂或数据噪声较大时，时空卷积可能未能充分提取关键特征，导致识别准确率下降而 VGG+CPM+LSTM 比较稳定。然而，在没有噪声的 "012_mp4" 和 "018_mp4" 视频中，STGCN+LSTM 的表现逐渐回升，表明其在处理具有一定时空模式但变化不剧烈的动作序列时能够发挥出优势。BLA+LSTM 整体表现相对稳定，但其 Jaccard 指数普遍低于 VGG+CPM+LSTM。这种现象说明基于骨骼长度与

角度的特征虽然简洁且具有物理解释性，但在复杂手势识别任务中存在信息损失的风险，难以完全捕捉细粒度的动作动态变化。因此，尽管 BLA+LSTM 能作为良好的模型进行对比分析，但在实际应用中识别精度受到一定限制。综合分析，VGG+CPM+LSTM 凭借对时间依赖关系的直接建模，在本任务中展现出整体优越的性能，而 STGCN+LSTM 虽然在某些片段中波动较大，但具有进一步挖掘时空特征潜力，适合处理具有复杂动态结构的数据。

4.系统架构设计与功能实现

4.1 系统总体架构

交通手势识别系统总体架构图如图 4.1 所示，系统由用户操作模块、数据处理模块、模型模块、可视化模块四个模块组成：

(1) 用户模块作为用户与系统的交互接口，集成了注册 / 登录、视频检测、停止检测以及注销登录功能，为用户提供了便捷的操作入口，使用户能够顺利发起视频检测请求或进行其他系统管理操作。系统识别的准确率毕竟不是 100%，安全员可以通过用户模块对无人驾驶汽车进行控制从而保证乘客的安全。

(2) 数据处理模块，首先对图像或视频进行读取，获取原始数据，随后通过关键点检测技术精准定位人体的关键部位，再进行特征提取，将数据转换为适合模型分析的形式，为后续的模型处理提供有价值的信息基础，确保输入数据能够被系统有效地理解与利用。

(3) 模型模块包含 CPM，PAF，LSTM。PAF 专注于人体姿态估计，通过对关键点的分析和处理，描绘出人体的姿态特征；LSTM 侧重于对时序特征的捕捉，能够处理数据中的时间维度信息，为手势识别等任务提供支持。训练模块通过数据加载获取训练所需的数据，然后进行模型训练，在训练过程中不断调整模型参数，以提升模型的准确性和泛化能力，模型验证环节对训练后的模型进行评估，检查模型的性能是否达到预期，确保模型在实际应用中的可靠性。

(4) 可视化模块将系统的处理结果以直观的方式呈现给用户，检测结果可视化则直接呈现对图像或视频的检测结果，使用户能够快速获取系统的分析结论。整个系统架构通过各模块的紧密配合，实现了从用户操作触发、数据处理准备、模型分析推理到结果展示的流程。

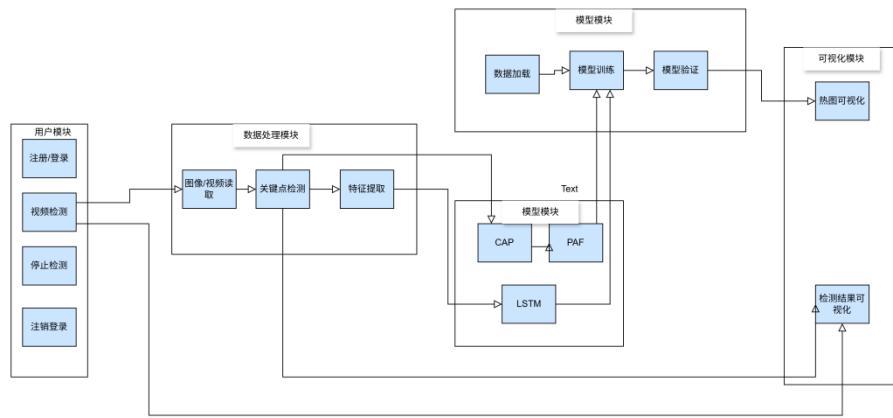


图 4.1 系统架构

4.2 实验环境

本文通过 VGG19 网络为基础引入卷积姿势机（CPM）以及长短期记忆网络（LSTM）设计并实现的交通手势识别系统是在 PyCharm 开发工具结合 PyTorch 框架实现的。实验的主要配置如表 4.1 实验配置所示。

表 4.1 实验配置

配置名称	详情
操作系统	Windows11
处理器	11th Gen Intel(R) Core(TM) i5-1135G7 @ 2.40GHz 2.42 GHz
显卡	Intel(R) Iris(R) Xe Graphics
内存	16.0 GB
硬盘	NVMe BC711 NVMe SK hynix 512GB
基础语言	Python
深度学习框架	PyTorch
CUDA	10.2.89
cuDNN	7.6.5

还安装了 opencv_python、torchvision 等学习库，主要开发环境如表 4.2 开发环境配置所示。

表 4.2 开发环境配置

库	版本
Flask	3.1.0
imgaug	0.4.0
numpy	1.22.0
numpy	1.24.3
opencv_python	4.11.0.86
PyMySQL	1.1.1
torch	1.10.0+cu102
torchvision	0.11.0+cu102
ujson	5.10.0

4.3 系统界面与功能演示

登录与注册界面包含用户名、密码输入框及相应按钮，注册功能对新用户开放，支持表单验证以确保用户名唯一性与密码复杂度要求，所有用户信息加密存储至 MySQL 数据库以保障数据安全。登录功能结合会话管理机制，在用户验证成功后跳转至主操作界面并保持用户状态。如图 4.2 所示。



图 4.2 注册登录界面

主操作界面采用上下分栏设计，上册侧为功能栏，下侧为主要区，集成了视频上传、检测控制、检测进度展示与结果展示等功能模块。视频上传模块通过 HTML5 文件上传组件实现，检测控制模块包括开始检测与停止检测按钮，检测进度模块实时反馈视频解析与分析状态，检测结果模块则以可视化方式展示人体关键点识别与手势分类结果。系统后端完成视频解码与逐帧提取，并对每帧图像执行人体关键点检测与手势分类，结果返回前端并进行动态渲染。如图 4.3 所示。

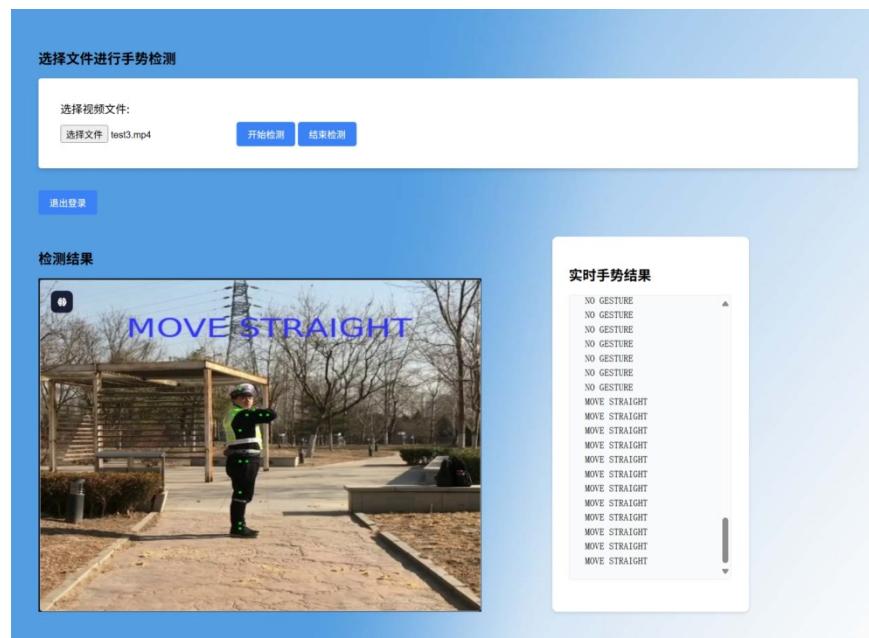


图 4.3 主界面

用户界面具备用户管理功能，支持查看注册用户列表、新增或删除用户账号、查询用户操作日志及管理模型训练与更新任务。界面设计遵循模块化和功能分离原则，为系统未来扩展复杂权限控制与审计机制提供了良好基础。通过系统界面与功能模块

的有机集成，交通手势识别系统实现了高效可靠的人机交互体验，能够有效支持交通场景下的人体姿态识别与手势分类任务。如图 4.4 所示。

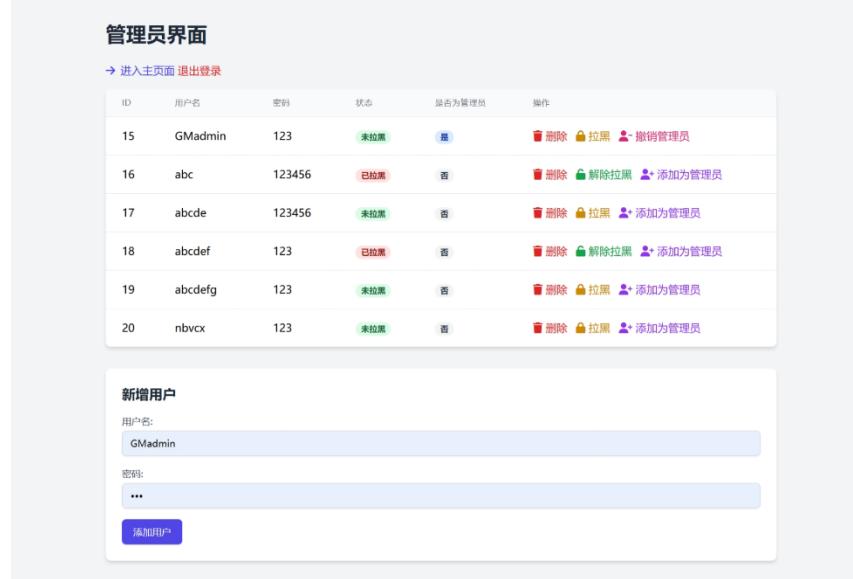


图 4.4 管理员界面

5 系统测试

为了验证交通手势识别系统的稳定性、完整性，本文围绕用户注册登录、模型调用响应、交互可视化、以及管理员权限管理等方面进行了系统测试。如表 5.1、表 5.2、表 5.3 所示。

表 5.1 注册登录测试

测试说明	预期结果	输出	判定结果
未输入账号密码点击登录	提示账号信息不能为空	提示账号信息不能为空	通过
账号空密码不为空点击登录	提示账号不可为空	提示账号信息不能为空	通过
密码空账号不为空点击登录	提示密码不可为空	提示密码不可为空	通过
当输入错误账号密码时	提示账号或密码错误	提示账号或密码错误	通过
当输入正确账号密码时	登录成功进入后台页面	登录成功进入后台页面	通过
输入管理员账号及密码时	进入管理员界面	进入管理员界面	通过
输入普通用户账号及密码时	进入用户界面	进入用户界面	通过

表 5.2 手势识别功能测试

测试说明	预期结果	输出	判定结果
上传视频格式文件	前端页面展示文件名	前端页面展示文件名	通过
上传其他格式文件	提示文件格式错误	提示文件格式错误	通过
文件上传点击开始检测按钮	显示检测结果	显示检测结果	通过
点击结束检测按钮	系统停止检测	系统停止检测	通过

表 5.3 用户功能测试

测试说明	预期结果	输出	判定结果
提交空用户名或密码	表单校验失败，提示字段不能为空	表单校验失败，提示不能为 空	通过
填写用户名和密码后提交	用户添加成功	用户添加成功	通过
点击“删除”按钮	用户删除成功	用户删除成功	通过
点击“拉黑”按钮	用户拉黑成功	用户拉黑成功	通过
点击“解除拉黑”按钮	用户恢复登录权限	用户恢复登录权限	通过

通过对系统用户注册登录、手势识别、视频处理、的测试，验证了系统功能完备、识别准确、操作流畅、安全稳定。系统能够在多种交通场景视频中保持准确的识别精度，对应手势反应迅速，具备实际部署与使用价值。

6 总结与展望

6.1 总结

本研究基于 VGG19 网络架构，融合 CPM 与长短期记忆网络 LSTM，构建了面向复杂交通场景的手势识别系统。实验采用 Police Gesture Dataset 与 AI Challenger 数据集，模型的 Jaccard 微平均分数达 82.81%。为了验证本文模型的有效性，将 VGG+CPM+LSTM 与 STGCN+LSTM、BLA+LSTM 模型作对比实验，实验结果表明 VGG+CPM+LSTM 模型在遮挡、视角变化和光照干扰性场景中表现更为稳定。在光照强烈的视频中，本模型的 Jaccard 指数较 STGCN+LSTM 提升显著，且整体波动幅度更小，能够有效捕捉手势的动态变化与空间关联性，显著提升复杂交通场景下的识别精度。本文采用 VGG+CPM+LSTM 模型设计并实现了手势识别系统，系统采用模块化架构设计，集成用户交互、特征提取（VGG19）、关键点检测（CPM+PAF）及分类（LSTM）模块，系统融合空间、姿态与时序特征的深度学习框架在遮挡、视角变化下对处理图像的能力稳定，具有一定的实际应用价值。

6.2 展望

未来进一步研究模型，将模型部署于嵌入式设备或移动平台，以满足手势识别的应用需求。手势识别技术与自动驾驶系统、交通信号控制系统、交通系统的整合，可以进一步实现智能化的交通管理。探索更高效的特征提取方法和更精确的时序建模技术，以提升手势识别的准确率。此外，当前系统处于测试阶段，后续研究考虑融入用户管理模块，并赋予安全员对车辆的控制权限，从而增强乘客的安全保障。

参考文献

- [1] 张丞,何坚,王伟东.空间上下文与时序特征融合的交警指挥手势识别技术[J].电子学报,2020,48(05):966-974.
- [2] 彭雪萍.基于Unity与Kinect的儿童交通安全教育体感游戏的构思与设计[J].信息与电脑(理论版),2021,33(19):114-117.
- [3] 阳嘉佳,刘治国.基于计算机视觉下的安全驾驶系统[J].数字技术与应用,2022,40(06):204-206.DOI:10.19695/j.cnki.cn12-1369.2022.06.62.
- [4] 张丞,侯义斌,何坚.高度分层分区的图卷积交警手势识别技术[J].计算机辅助设计与图形学学报,2022,34(07):1037-1046.
- [5] 程贝芝,伍鹏,寇静雯,等.结合全局上下文信息的交警手势识别方法[J].中南民族大学学报(自然科学版),2023,42(03):349-356.DOI:10.20056/j.cnki.ZNMDZK.20230309.
- [6] 梁钊铭,段旭洋,王皓.视觉导引装配场景动态手势识别方法[J].机械设计与研究,2023,39(02):12-18.DOI:10.13952/j.cnki.jofmdr.2023.0049.
- [7] 李泰国,张英志,张天策,等.基于改进YOLOv5s算法的列车驾驶员手势识别[J].铁道学报,2023,45(01):75-83.
- [8] 鲁杰伟,盘轩,彭雯蝶,等.一种基于改进YOLOv5s的手势识别算法[J].电脑知识与技术,2024,20(12):1-3.DOI:10.14004/j.cnki.ckt.2024.0593.
- [9] 孔皓.人机交互技术在汽车中的应用及对驾驶体验的影响[J].汽车测试报告,2024,(04):149-151.
- [10] 杜兵,赵骥.基于时空特征融合的交通警察手势识别[J].计算机工程与应用,2024,60(08):250-257.
- [11] S.Swapanali and P.G.Chilveri, "Hand Gesture Recognition using accelerometer sensor for traffic light control system," 2014 International Conference on Electronics and Communication Systems (ICECS), Coimbatore, India, 2014, pp. 1-5, doi: 10.1109/ECS.2014.6892648.
- [12] G.Yasmeen, S.Arun, J.N.Swaminathan, S.A.K.Jilani and S.Asif, "Efficient Hand Gesture Recognition for Traffic Control System Using ti Sensor Tag," 2018 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2018, pp. 1-3, doi: 10.1109/ICCCI.2018.8441483.
- [13] P.Sharma and N.Sharma, "Gesture Recognition System," 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), Ghaziabad, India, 2019, pp. 1-3, doi: 10.1109/IoT-SIU.2019.8777487.
- [14] M.D.Artemov, L.I.Voronova and A.G.Vovik, "Designing Soft-Hardware Complex for Gesture Language Recognition using Neural Network Methods," 2020 International Conference on Engineering Management of Communication and Technology (EMCTECH), Vienna,Austria, 2020, pp. 1-6, doi: 10.1109/EMCTECH49634.2020.9261564.
- [15] J.Wiederer, A.Bouazizi, U.Kressel and V.Belagiannis, "Traffic Control Gesture Recognition for Autonomous Vehicles," 2020 IEEE/RSJ International Conference on Intelligent Robots and

d Systems (IROS), Las Vegas, NV, USA, 2020, pp. 10676-10683, doi: 10.1109/IROS45743.2020.9341214.

致 谢

首先,我由衷感谢我的指导老师。在整个毕业设计的过程中,指导老师给我专业的指导和耐心的帮助。当我遇到困难时,老师都会耐心地解答我的疑问,给我提供的建议和思路。在课题选取、资料整理环节,老师都给予了细致的指导和鼓励,使我对研究有了更深入的认识。同时,我也要感谢家人。在毕业设计期间,他们在精神上给了我无微不至的关怀和鼓励,在生活上提供了无私的支持,为我分担压力。他们的理解和鼓励成为我前进的动力,让我专注于学业和毕业设计。

在老师和家人的支持下,我得以完成毕业设计项目,同时也在实践中巩固并拓展了我的知识和技能。我通过这次毕业设计,坚定了对学术追求的信心,也深刻体会到自己的成长。面对困难时,学会了冷静和自信。同时我也认识到还需努力。

关于学位论文使用授权说明

本人完全了解兰州资源环境职业技术大学有关保留、使用学位论文的规定，即：学校有权保留学位论文的复印件，允许该论文被查阅和借阅；学校可以公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存该论文。

(涉密的学位论文在解密后应遵守此规定)

签 名： 赵浩浩 导师签名： 陈海波 日 期： 2025年5月30日