**基于yolo的手语识别系统设计与实现 （任务书）**

**一**、**设计目的**

手语作为听障群体的视觉语言是帮助听障人士融入社会的重要交流方式。但是由于普通大众一般缺乏系统性的手语学习，难以准确地理解手语的意思，因此，健听人士能够理解听障人士的手语表达是解决听障群体无障碍沟通的主要需求。目前大多数通信技术都是为了支持口语或书面语而开发的，支持手语的通信技术平台少之又少。虽然语音聊天和视频通话等通信技术和工具已经成为我们生活的重要组成部分，但听障人士在使用这些技术方面仍存在许多问题。为解决健全人士与听障人士沟通交流困难的问题，同时为了弥补传统识别网络中存在的缺陷，本文提出了一种改进 YOLO网络模型的手语识别网络，对手语识别任务进行了相应研究，以求能更好的帮助听障人士解决日常交流。

**二、内容与要求**

图像采集：手语识别通常从使用摄像头捕捉手部动作和手势的视频图像开始。这可以是通过普通摄像头、深度摄像头或者其他传感器来实现。

预处理：采集到的图像可能需要进行一些预处理步骤，以提高后续分析的准确性。这可能包括去噪、图像增强、手部检测等。

手部检测和跟踪：为了识别手语，系统需要能够检测和跟踪手部在图像中的位置。这可能涉及到计算手部的轮廓、关键点或手部的运动轨迹。

特征提取：从手部的图像中提取关键的特征，例如手指的位置、手掌的形状等。这些特征有助于定义手语中的不同手势。

手势识别：利用机器学习或深度学习技术，训练一个模型来识别不同的手势。这可能包括静态手势（手的形状）和动态手势（手的运动轨迹）。

模型训练：使用标注好的手语数据集，训练模型以学习不同手势的表示和分类。深度学习方法，如卷积神经网络（CNN）和循环神经网络（RNN），在手语识别中常被使用。

语义解释：一旦识别了手势，系统需要将其转化为可理解的文字或语音信息。这可能涉及到将手势与事先定义的手语词汇或短语进行匹配，以生成相应的文本或语音输出。

实时性和交互性：对于实时的手语识别系统，确保系统能够快速响应和适应变化的手势是关键的。这对于与聋哑人士的实际交流非常重要。

**三、成果要求**

1、创新性模型设计：

首先搭建了 YOLO 手语识别任务实验环境。其次将注意力机制、聚类算法和损失函数这三个方面的改进方法进行组合，通过消融实验验证了三个方面的改进方法组合使用的有效性，并与不同网络模型进行了对比实验。

2、详细的模型实现和代码：

提供详细的模型实现代码，包括两阶段和单阶段目标检测算法、CBAM 注意力机制的原理的添加和使用 K-means++聚类算法 代替原始的 K-means 聚类算法，提高了先验锚框的尺寸匹配度，实现先验锚框与目标手 语图像的精确匹配和模型训练的全部过程。

3、实验结果和性能评估：

提交在合适的手语采集数据集上进行的实验结果。包括模型在训练集和测试集上的性能评估，使用准确性、召回率等指标。与传统方法进行比较，以验证创新模型的有效性。

4、性能优化和改进策略：

相较于基于传感器的手语识别方法仍然存在一定的提升空间, 目前本文所使用的数据集只涵盖了26个英文字母的手语 手势，应该采集更多的常用词汇来丰富和扩大数据集，以扩展模型的应用范围，同时提高识别准确度。

**四、进度计划**

2023.12.1-2023.12.25 选题、课题调研、收集资料、外文翻译、编写开题报告；

2023.12.26-2023.12.29 开题答辩；

2024.2.26-2024.3.3   阅读有关书籍或文献、初步构想、论文的初稿撰写；

2024.2.26-2024.3.24  毕业实习；

2024.3.25-2024.5.8   详细论证、中期工作汇报；

2024.5.9-.2024.5.31  论文撰写与完善；

2024.6.1-2024.6.4   资料提交上传；

2024.6.5-2024.6.11   指导教师评阅、评阅教师评阅；

2024.6.12-2024.6.16  毕业论文答辩；

2024.6.17-2024.6.20  毕业设计总结、评优；

**五、参考文献**

[1]王琪.基于深度学习的手语识别系统设计与实现[D].西安电子科技大学,2021.DOI:10.27389/d.cnki.gxadu.2021.000324

[2]邢晋超.基于改进YOLOv5s的手语识别算法研究[D].中北大学,2023.DOI:10.27470/d.cnki.ghbgc.2023.001085

[3]邢晋超,潘广贞.改进YOLOv5s的手语识别算法研究[J].计算机工程与应用,2022,58(16):194-203.

[4]蒲俊福,基于深度学习的视频手语识别研究[D]. 中国科学技术大 学,2020.DOI:10.27517/d.cnki.gzkju.2020.000530.

[5]韩凯. 基于深度学习的目标检测研究[D].西南科技大学,2018.

[6]吕情深.基于目标检测的中国手语识别研究[D]. 新疆大 学,2020.DOI:10.27429/d.cnki.gxjdu.2020.000068.

[7]刘建伟,刘俊文,罗雄麟.深度学习中注意力机制研究进展[J].工程科学报,2021,43(11):1499-1511.DOI:10.13374/j.issn2095-9389.2021.01.30.005.

[8]王丽光,张根源,刘子龙.基于单目视觉的实时手语识别系统设计与实现[J].电子科技,2017,30(03):130-133+137.DOI:10.16180/j.cnki.issn1007-7820.2017.03.036

[9]罗文杰.基于深度学习的手语识别研究与应用[D].南昌大学,2023.DOI:10.27232/d.cnki.gnchu.2023.003074

[10]高辉.基于深度学习的手语识别系统设计[D].哈尔滨理工大学,2023.DOI:10.27063/d.cnki.ghlgu.2023.000303

[11]王婧瑶,范飞,刘豪宇等.基于机器视觉的聋哑人手语识别——语音交互系统[J].物联网技术,2021,11(12):3-5.