[](http://news.cau.edu.cn/upload/2005/caunews_20050711181514.jpg)

**本科生毕业论文（设计）**

微型肠道机器人机电系统设计

及肠道息肉识别

**Electromechanical system design of miniature intestinal robot and intestinal polyp recognition**

学生姓名： 陶胜

指导教师： 胡标 副教授

专业名称： 机械电子工程

所在学院： 工学院

2023年 5 月

**摘 要**

随着人类社会的发展，人们的生活节奏越来越快，受到饮食结构改变、精神压力、食品安全等因素的影响，肠道疾病也逐渐成为影响人民健康的一个显著因素。在肠道疾病的诊断与治疗中，结肠镜检查是一个辅助医生治疗的重要手段，但是传统的电子肠镜只有前部的导向段具有自主弯曲的能力，后端主要靠操作者将镜体向肠腔内推进过程中受到肠壁挤压后被迫弯曲。在这个过程中，患者容易感到不适，而且严重时会导致出血穿孔，这些都不利于疾病的恢复。同时，结肠镜诊断对操作者的专业技能和经验积累的要求较高。

针对以上问题，本文设计了一种履带式微型肠道机器人。本文介绍了微型肠道机器人的工作原理，对机器人的机械结构进行设计，并对关键结构进行了力学性能的仿真与验证。该机器人内置直流微电机，以同步履带作为行走机构，操作者可在外部的通过控制器来控制微型肠道机器人在肠腔内部的运动，并依靠搭载的摄像头获得肠道内部图像。与传统的电子肠镜相比，主动运动肠镜机器人可以帮助操作者更好地控制镜体的姿态，在一定程度上减少了患者的不适程度，降低操作者的操作难度，缩短诊断时间，对于疾病的诊断与恢复和提升诊断效率都具有重要意义。

在临床结肠镜检查的过程中，息肉的检查是一个重要的检查项目。结肠癌大部分都是由结肠息肉发展而来，早期发现并切除结肠息肉可以大大地提升患者的生还率，然而在临床上，由于肠道内部环境复杂、息肉形态各异以及医生长时间易疲劳等因素的影响，息肉漏检很难避免。针对此问题，本文基于YOLOv5神经网络训练了用于结肠息肉识别和语义分割的算法。与传统靠医生肉眼识别结肠息肉相比，用YOLOv5神经网络实时辅助医师诊断，可以一定程度降低医师的工作强度，极大地缩短诊断时间，并有效提高息肉诊断效率。

最后，本文搭建了一套人机交互系统。首先，根据机器人运动控制的要求，搭建了一套遥控式机器人运动控制电路。然后，为进一步简化操作，开发了一套基于PyQt开发的上位机软件，其中集成了运动控制和肠镜图像显示与识别模块。通过人机交互系统，操作者可以通过简单地控制机器人的主动运动，实时获得肠镜的图像和息肉检测的结果。

**关键词：**微型肠道机器人；息肉识别；语义分割；YOLOv5

**Abstract**

Along with the development of human society, people 's pace of life is getting faster and faster. Affected by factors such as changes in dietary structure, mental stress, and food safety, intestinal diseases have gradually become a significant factor affecting people 's health. Colonoscopy is an important medical method in the diagnosis and treatment of intestinal diseases. However, the traditional electronic colonoscopy only has the ability to curve autonomously in the front part, and the back part is mainly forced to curve by the intestinal wall during the process of pushing the mirror into the intestinal cavity. During this process, patients are prone to feel discomfort, and severe cases can lead to bleeding and perforation, these are not conducive to the recovery of the disease. At the same time, traditional colonoscopy diagnosis requires a high level of professional skills and experience of the operator.

To solve the above problems, this paper proposes a tracked micro intestinal robot. This paper introduces the working principle of the micro intestinal robot, designs the mechanical structure of the robot，simulates and verifies the mechanical properties of the key structure. The robot applies crawlers as walking mechanism, which are drove by built-in DC micro-motors. The operator can control the movement of the micro-intestinal robot inside the intestinal cavity through the external controller, and get the intestinal internal images which is collected by the mounted camera on the robot. Compared with the traditional electronic colonoscopy, the operator can control the posture of the micro intestinal robot easier, which can also reduce the discomfort of the patient and the difficulty of the operator 's operation, and shorten the diagnosis time. It is of great significance for the diagnosis and recovery of the disease and the improvement of the diagnosis efficiency.

In the process of clinical colonoscopy, polyp detection is an important diagnosis item.

Colon cancer is mostly developed from colon polyps. Early detection and resection of colon polyps can greatly improve the survival rate of patients. However, in clinical practice, due to the complex intestine environment, the different appearance of polyps, and the fatigue of doctors caused by the long-time diagnosis, it is difficult to avoid the missed detection of polyps. Aiming at this problem, this paper trains an algorithm for colon polyp recognition and semantic segmentation based on YOLOv5 neural network. Compared with identifying colonic polyps by the eyes of doctors, using YOLOv5 neural network to assist doctors in real-time diagnosis can reduce the work intensity of doctors, greatly shorten the diagnosis time, and improve the efficiency of polyp diagnosis.

Finally, this paper builds a human-machine interaction system. Firstly, according to the requirements of robot motion control, a robot motion control circuit is built. Then, in order to further simplify the operation, a set of PC software based on PyQt is developed, which integrates motion control and colonoscopy display and recognition modules. Through the human-machine interaction system, the operator can simply control the active motion of the robot to obtain real-time colonoscopy images and polyp detection results.

**Key words:** micro intestinal endoscope robot; polyp recognition; semantic segmentation; YOLOv5

**目 录**

[**摘 要** I](#_Toc136257085)

[**Abstract** II](#_Toc136257086)

[第一章 绪论 1](#_Toc136257087)

[1.1选题背景和意义 1](#_Toc136257088)

[1.2内镜机器人的研究现状 3](#_Toc136257089)

[1.3基于深度学习的肠镜图像处理的研究现状 7](#_Toc136257090)

[1.4主要研究内容和章节安排 10](#_Toc136257091)

[第二章 机械结构设计 11](#_Toc136257092)

[2.1机械结构总体构思 11](#_Toc136257093)

[2.2传动机构设计与选型 12](#_Toc136257094)

[2.3机器人的材料选择及有限元仿真 15](#_Toc136257095)

[2.4本章小结 18](#_Toc136257096)

[第三章 结肠镜息肉的识别和语义分割 19](#_Toc136257097)

[3.1YOLO算法介绍 19](#_Toc136257098)

[3.2网络模型结构 19](#_Toc136257099)

[3.2 实验与分析 20](#_Toc136257100)

[3.3本章小结 27](#_Toc136257101)

[第四章 机器人的运动控制及肠镜图像显示 28](#_Toc136257102)

[4.1微型肠道机器人的上位机软件 28](#_Toc136257103)

[4.2微型肠道机器人的主动运动控制硬件电路 30](#_Toc136257104)

[4.3本章小结 31](#_Toc136257105)

[第五章 总结 32](#_Toc136257106)

[参考文献 34](#_Toc136257107)

[致 谢 37](#_Toc136257108)

[作者简介 39](#_Toc136257109)

**插图和附表清单**

[图1-1传统电子肠镜 1](file:///C:\Users\T\Desktop\23231.docx#_Toc135315509)

[图1-2 PillCamCOLON2系统 3](file:///C:\Users\T\Desktop\23231.docx#_Toc135315510)

[图1-3尺蠖 4](file:///C:\Users\T\Desktop\23231.docx#_Toc135315511)

[图1-4尺蠖式内镜机器人 4](file:///C:\Users\T\Desktop\23231.docx#_Toc135315512)

[图1-5动压旋进式内镜机器人原理图 5](file:///C:\Users\T\Desktop\23231.docx#_Toc135315513)

[图1-6足式内镜机器人 5](file:///C:\Users\T\Desktop\23231.docx#_Toc135315514)

[图1-7轮式内镜机器人 6](file:///C:\Users\T\Desktop\23231.docx#_Toc135315515)

[图2-1微型履带式肠道机器人驱动结构示意图 11](file:///C:\Users\T\Desktop\23231.docx#_Toc135317158)

[图2-2微型履带式肠道机器人整体外观正面示意图 11](file:///C:\Users\T\Desktop\23231.docx#_Toc135317159)

[图2-3微型履带式肠道机器人上从动同步履带结构示意图 12](file:///C:\Users\T\Desktop\23231.docx#_Toc135317160)

[图2-4直齿锥齿轮安全系数分析结果图 16](file:///C:\Users\T\Desktop\23231.docx#_Toc135317161)

[图2-5直齿锥齿轮的米赛斯等效应力分析结果图 16](file:///C:\Users\T\Desktop\23231.docx#_Toc135317162)

[图2-6直齿锥齿轮的位移分析结果图 16](file:///C:\Users\T\Desktop\23231.docx#_Toc135317163)

[图2-7零件35安全系数分析结果图 17](file:///C:\Users\T\Desktop\23231.docx#_Toc135317164)

[图2-8零件35的米赛斯等效应力分析结果图 17](file:///C:\Users\T\Desktop\23231.docx#_Toc135317165)

[图2-9零件35的位移分析结果图 18](file:///C:\Users\T\Desktop\23231.docx#_Toc135317166)

[图3-1 Mosaic数据增强的训练图像 19](file:///C:\Users\T\Desktop\23231.docx#_Toc135250010)

[图3-2 Kvasir-Seg内息肉图像和mask图像示例[47] 21](file:///C:\Users\T\Desktop\23231.docx#_Toc135250011)

[图3-3 Labelme对Kvasir-SEG数据集手动分割息肉像素 22](file:///C:\Users\T\Desktop\23231.docx#_Toc135250012)

[图3-4精确率与置信度的关系曲线 23](file:///C:\Users\T\Desktop\23231.docx#_Toc135250013)

[图3-5召回率与置信度的关系曲线 23](file:///C:\Users\T\Desktop\23231.docx#_Toc135250014)

[图3-6精确率与召回率的关系曲线 24](file:///C:\Users\T\Desktop\23231.docx#_Toc135250015)

[图3-7损失曲线 25](file:///C:\Users\T\Desktop\23231.docx#_Toc135250016)

[图3-8各个轮次训练的精确率曲线 25](file:///C:\Users\T\Desktop\23231.docx#_Toc135250017)

[图3-9各个轮次的召回率曲线 26](file:///C:\Users\T\Desktop\23231.docx#_Toc135250018)

[图3-10 mAP曲线 26](file:///C:\Users\T\Desktop\23231.docx#_Toc135250019)

表2- 1福尔哈贝传动0615N4.5S+06/1 64:1直流减速微电机参数…………….…13

表2- 2直齿锥齿轮设计参数………………………………………………….…….14

表2- 3圆柱直齿轮设计参数………………………………………………………..14

表2- 4同步带及带轮设计参数……………………………………………………..15

第一章 绪论

本章主要介绍了本文的选题背景和意义，论证了肠道机器人及肠镜图像处理的必要性和可行性，然后阐述了肠道机器人的研究现状和基于深度学习肠镜图像处理的研究现状，最后介绍了本文主要研究内容和章节安排。

## 1.1选题背景和意义

结直肠癌是严重威胁人类健康的常见肿瘤之一，根据国际癌症研究机构编制的GLOBOCAN2020癌症发病率和死亡率统计，结直肠癌在男性最常诊断的恶性肿瘤中排行第三，在女性最常诊断的恶性肿瘤中排行第二[1]，在所有癌症的死亡率中排行第二[2]。在过去30年中，中国结直肠癌的发病率稳步上升，并且在未来有进一步增加的趋势。如何降低结直肠癌的发病率和死亡率是一个急需解决的公共卫生问题[3, 4]。

图1-1传统电子肠镜

### 1.1.1肠道机器人的研究意义

目前医学临床上，仅通过患者的外在体征和既往病史很难对结直肠疾病做出准确的诊断，主要通过结肠镜观察并辅以其他医疗手段对结直肠疾病进行诊断和治疗[5]。

结肠镜检查是一种进入患者体腔来获得图像的侵入式医疗诊断手段，医生将柔性电子肠镜通过肛门进入患者肠腔，获取实时肠镜图像，进行分析诊断。整个过程可以分成两个主要阶段。第一个阶段是进镜，医生将结肠镜插入患者结肠，直到肠镜到达盲肠或者小肠的末端，在这个过程中，医生会对肠道进行一些清洗工作，对肠腔内部进行大致的观察，主要目的是将肠镜伸到结肠末端。第二阶段是退镜，在这个过程中，医生缓缓将镜体从结肠末端向后回退，这时，医生会仔细观察肠腔，有时为了便于肠镜的移动会向肠腔充气[6, 7]。目前临床结肠镜检查主要存在两个问题：（1）仅导向段可控，结肠镜检查时间长。由于柔性电子肠镜只有前段较短的导向段可以弯曲，后端的镜体通过与肠壁接触而被动弯曲，所以调整镜体的时间较长，整个检查过程会持续10-30分钟，患者可能会感到不适，如腹部痉挛，也有少数患者会出现并发症，比较严重的如出血或穿孔[8]，这无疑增加了患者的痛苦和风险，不利于疾病的治疗与恢复。同时，为保证获取较好的肠镜图像，医生需要长时间保持相对固定的姿势，医生容易感到疲劳、注意力分散，导致医疗手段难以达到最佳效果。（2）对医生专业技能要求高，往往需要多人操作。在肠镜检查的过程中，需要助手负责插镜、压腹，而医生需要长时间单手控制拨轮，来调节内窥镜前段的导向段，进而引导肠镜在肠道中的移动，另一只手通过图像信息和手感来调节和旋转镜体来观察病灶。这些操作需要医生和助手有一定的专业知识和经验积累[7]。针对传统结肠镜检查的问题，可以利用机器人技术，设计一款可控程度更高的结肠镜机器人，降低结肠镜检查的操作难度和门槛，减少对患者的入侵程度，减轻医师的工作强度，提升结肠镜检查的诊断效率。

### 1.1.2肠道息肉识别的研究意义

结直肠癌的一个显著特征就是90%以上由腺瘤-腺癌序列发展而来，而结肠腺瘤多表现为息肉样病变，如果在无症状早期通过结肠镜筛查发现结肠息肉并将其切除，就可以有效避免它发展成恶性肿瘤，早期发现的患者5年内的生存率可达90%，这也使结直肠癌成为少数可以完全预防的癌症[9]。传统的结肠镜筛查是医师通过肉眼观察电子肠镜传回的图像来诊断肠道病变情况。但是在长期的临床实践中发现，即使是经验丰富的医师也无法保证一次就发现所有病灶位置，肠镜筛查的效果高度依赖于医师的经验，漏诊率在6%-27%[10]，由此可见，传统的肠镜筛查的效率仍然存在一定的局限性。

随着深度学习和计算机视觉的迅速发展，相关技术在医疗图像领域的应用也在不断深入，早期受限于算法和计算机的运算能力，用计算机辅助诊断一直处于基础应用研究的阶段，但是，随着深度学习研究的深入和计算机运算能力的发展，计算机辅助结肠诊断技术正在不断走向现实[11, 12]，深度学习利用计算机模拟人类学习的过程，通过大量学习相关的医疗图像，提取环境中病灶的特征，模拟医师的诊疗思维和方式，对临床医疗视频进行检测，可以减轻医师的工作强度，缩短检测时间，这对于疾病的防控具有重要意义。

## 1.2内镜机器人的研究现状

经过多年的研究，世界上已经有多种内镜机器人问世，有的已经在医学临床上实现应用。根据机器人能否在肠道内主动运动，可将其分成主动式内镜机器人和被动式内镜机器人。

### 1.2.1被动式内镜机器人研究现状

被动式内镜机器人一般是指胶囊机器人，由于形似胶囊而得名。患者吞服机器人后，机器人从口腔进入消化道，机器人通过肠道的蠕动在消化腔内被动运动，在此过程这种机器人记录消化道的图像，胶囊机器人的续航时间一般为6-8小时，经过6-72小时后，胶囊机器人在经过整个消化系统后自然排出体外。一般胶囊机器人内包含图像传感器、照明装置、电池、图像传输模块，通过这些模块的协调运作来获取消化道的图像信息，为医生诊断提供依据。

2000年Given Image公司研制了世界上第一款胶囊内镜机器人，同年获得了FDA的批准进入临床[13]，2002年8月，Given Image推出了M2A一次性塑料胶囊式内窥镜[14]，这款胶囊内镜机器人由胶囊机器人本体、附带便携硬盘的外部接受天线和图像处理的计算机工作站组成。它可以提供140°的视角，以每秒两帧的频率输出图像，续航时间为7±1小时，在实验中，约有一半的胶囊式内窥镜在进入盲肠之前停止工作。在进行诊断时，患者需要佩戴肩带式腰包，其中放置有电池和一个用于存档接受图像305G硬盘，在口服胶囊之前，需要在患者的腹部贴附指定图案的天线阵列。这一款胶囊内窥镜机器人虽然存在续航时间较短，输出图像频率低的问题，但是它的出现对后续胶囊内窥镜的发展和改进具有重要意义。之后，Given Image公司推出了PillCam系列的胶囊内窥镜系统，其中PillCamCOLON2系统[15, 16]主要用于结肠镜的检查，它前后各搭载了一个CMOS图像模块和4个LED照明管，图像模块可以提供172°的视角，图像输出频率为4-35帧每秒，续航时间最少10小时，重量为2.9±0.1g，相比与初代的胶囊内镜，PillCamCOLON2可以直接显示直肠，并应用了帧率自适应调整技术，在胶囊运动或者图像变化时，系统会以高帧率来捕捉图像。

图1-2 PillCamCOLON2系统

上图中从左到右依次为胶囊机器人本体、传感腰带、记录仪、8导联数据接收阵列。

2004年，重庆金山科技集团推出了我国首个自主研发的胶囊内镜[17, 18]，它与能在消化道内随肠道蠕动实现从胃部到直肠的移动，同时对消化腔道进行图像采集和输出，金山科技集团的胶囊内镜系统已经在临床上得到了应用。

与传统的内镜相比，胶囊内镜机器人有无需麻醉、无创、无痛、操作简单的优点，同时由于采用无线传输数据，对患者的入侵程度低，而且胶囊机器人一般不会重复使用，避免了交叉感染，这些都使它成为了少数应用于临床的内镜机器人。但是，胶囊内镜机器人不能通过充气喷液的方式对肠道进行清洗，所以在诊断的过程中容易受肠道清洁程度的影响，不能全面的观察、发现病灶，容易出现漏诊。同时，胶囊机器人采用无线数据传输，这就限制了图像的分辨率，影响了图像的清晰度，而且，传统肠镜可以搭载活检钳等其他辅助设备来进行采样和部分息肉的切除治疗，而胶囊内窥镜机器人只能进行诊断观察的任务。所以，胶囊内窥镜系统目前主要用于胃部内镜检查，还不能完全取代传统电子内窥镜。

### 1.2.2主动式内镜机器人研究现状

与被动式机器人相比，主动式内镜机器人不需依靠肠道的蠕动等外界力量，可以实现自主运动，它也比被动式内镜机器人出现的更早。主动式内镜机器人按照其行进的方式又可以分为仿生蠕动式、旋进式、足式、轮\履带式。

在仿生蠕动式内镜机器人系统中，仿尺蠖式内镜机器人的研究较多。1995年，A. B. Slatkin等研制了一款尺蠖式小肠内窥镜机器人系统[19]，该机器人的主体为3节可伸缩的橡胶气囊组成，3节气囊协通过调运作来模仿尺蠖爬行的步态实现机器人在小肠内的运动，由于机器人主要依靠气囊与肠壁的摩擦进行运动，而肠壁往往因附有消化液而比较光滑，所以此机器人的运行效果不佳。

图1-4尺蠖式内镜机器人

图1-3尺蠖式

针对尺蠖式机器人与肠道接触摩擦力不足的问题，P.Dario等研制了利用气压吸附肠壁的内窥镜机器人[20]，通过向机器人的吸附孔内通低压，使机器人对肠道实现吸附定位，需要脱离肠壁时再通微弱高压，该机器人的最大行进速度为5mm/s。在2017年，Calderon等人为了进一步提高尺蠖式机器人的推进效率，在机器人的首尾添加了基于液体电路的电子皮肤传感器来感知机器人与肠道的接触力[21]。尺蠖式内镜机器人通过负压吸附肠壁，可以实现机器人在肠道内定位，但是尺蠖式机器人的行进速度较慢，不利于诊断效率的提升。

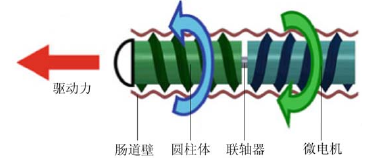
1996年，Ikeuchi等研制了一款利用动压效应旋进式机器人，它利用两个旋向相反的螺旋柱与肠道粘液摩擦产生的合牵引力驱动机器人移动，在动压效应的作用下，机器人与肠壁之间充满了粘液，与肠壁没有直接接触[22]。

图1-3动压旋进式内镜机器人原理图

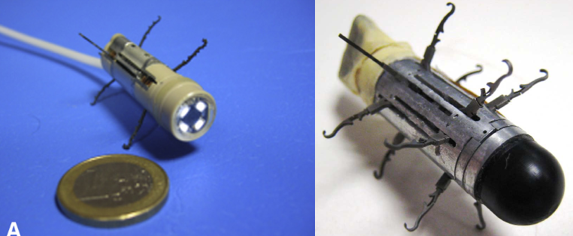
 也有研究人员研制了足式内镜机器人，它通过在机器人上布置足式结构，利用电机驱动其在肠道内行走[23, 24]。

图1-4足式内镜机器人

面对胶囊式机器人运动不受控的问题，也有主动胶囊内镜机器人的研究。2009年，Tortora等提出了一种通过螺旋桨驱动的可吞咽胶囊内镜机器人[25]，内镜医师通过操纵杆和专门开发的用户界面可以远程控制胶囊机器人，它主要用于胃部内镜，它可以1.5m/s的速度在充满液体的胃部持续运动和获取图像30min，但是目前仅在离体猪胃中实验。目前，更为可行并实现临床应用的是磁控式胶囊内镜机器人，磁控胶囊的理念最早是在2006年被提出[26]，磁控式胶囊内镜按控制磁场的类型可以分为三类：MRI式、手柄式、机械臂式。2010年，Rey等研发了一种MRI式的磁控胶囊内镜系统[27]，MRI式磁控胶囊内镜系统用核磁共振系统来对胶囊内镜进行控制，它主要用于胃部诊断，在诊断时胶囊需要借助液体浮力才能进行有效控制，所以在诊断前需要患者饮用大量的水，由于这套系统需要依靠核磁共振系统，所以成本较高、操作复杂。2010年，Swain等基于Given Imaging的结肠无线胶囊内镜系统[28]，在胶囊机器人内置入稀土磁性材料，在体外用一个桨状永磁体手柄对机器人进行控制。手持式磁控胶囊内窥镜系统的磁场范围较小，磁场强度较弱。2012年，Liao等研发了一款机械臂式胶囊内镜系统[29]，机械臂式胶囊内镜系统通过机械臂上的永磁体对胶囊内镜进行控制，它比MRI和手柄式胶囊内镜系统的磁场强度大很多，可实现对胶囊内镜毫米级的移动，与MRI相同，它也需要体积庞大的磁场发生和控制装置。

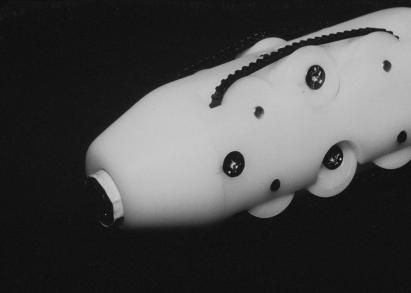
轮/履带式内镜机器人是一种基于传统的轮/履带的自主推进内镜机器人。1999年，Goh等提出了一款轮式内镜机器人[30]，轮式履带机器人利用布置的轮子运动，但是轮子在肠道内湿滑环境中容易打滑，采用接触面更大的履带则可以获得更好的通过性和可控性，这也是本文所研究的内镜机器人所采用的结构。

图1-5轮式内镜机器人

## 1.3基于深度学习的肠镜图像处理的研究现状

### 1.3.1结肠息肉目标检测研究现状

基于深度学习的目标检测算法主要有两种方式：基于锚框识别（Anchor base）和无锚框识别（Anchor free）。在基于锚框识别的目标识别算法中又分为单阶段（one-stage）和两阶段（two-stage）算法，两阶段算法的识别过程可分为两步，首先生成一系列的样本候选框，再从候选框中生成最终的物体边框；而单阶段算法则是直接通过主干网络得到类别和位置信息，没有使用区域生成网络。目前主要的两阶段算法有R-CNN（Region with CNN feature）[31]、SPPNet（Spatial Pyramid Pooling Networks）[32]、Fast R-CNN[33]、Faster R-CNN [34]、FPN（Feature Pyramid Networks）[35]。R-CNN是第一个基于深度学习的目标检测和语义分割算法，它可以将高容量的卷积神经网络用于自下而上的区域建议，以便于定位和目标的分割，而且当训练数据较少时，可以进行监督预训练，对指定的区域进行微调，进一步提升网络检测的性能[31]。但是，R-CNN算法中对于重叠候选框的冗余运算使检测的速度很慢。SPPNet为了改变先前的卷积神经网络需要输入固定大小图像的情况，采用了一种新的池化策略——空间金字塔池化层（Spatial Pyramid Pooling Layer，SPP），改进了传统基于卷积神经网络的图像分类的方法，使网络只需要一次计算就能生成整个图像的特征图，然后将指定区域的特征集中起来，该网络提高了卷积神经网络架构的准确性和鲁棒性[32]。之后，Fast R-CNN提出了一种快速基于区域的卷积网络的对象检测方法，该方法在之前RCNN的基础上，使用深度卷积神经网络进行分类，进一步提高了训练和测试时的速度和检测精度[33]。Faster R-CNN采用了区域建议网络，这是一个完全卷积网络，它可以预测每个位置的对象边界和对象分数，经过端对端的训练，生成高质量的区域建议，通过与检测网络共享全部图像的卷积特征，到达了以较低的计算成本获得识别区域建议的目的，Faster R-CNN进一步提高了检测的精度 [34]。FPN为了解决之前目标检测网络中尺度变化大的问题，在特征金字塔网络的基础上，提出了一种自适应特征金字塔网络来缓解此问题，该网络采用了自适应上采样和自适应特征融合的策略，通过模型预测每一个像素的一组采样点，并用采样点的特征组合来构成像素的特征表示，并用注意力机制来构建融合特征之间的像素级融合权重，进一步提高了算法对识别目标的定位准确性[35]。单阶段算法中比较有代表性就是YOLO系列目标检测算法[36,37,38,39]、SDD（Single Shot MultiBox Detector）。YOLO系列算法的一大优势就是检测速度快，能实现实时检测，具备不错对迁移和泛化的功能，而且在之后迭代的算法中，YOLO对于设备运算能力的要求也在不断减低，进一步降低了算法的应用门槛。YOLOv1在2015年提出，它将目标检测问题转化成一个回归问题，在一次的评估中网络直接从整个图像预测边界框和目标的类别概率，由于整个网络是一个单一的网络，它可以被直接优化成一个端对端的目标检测网络，它可以实现实时检测并具有不错的准确率[36]；之后YOLOv2进一步改进了检测方法，联合训练对象的检查与分类，进一步提高了实时目标检测的类别数限制和准确度[37]，YOLOv3采用了一个由YOLOv2、Darknet-19和残差组件混合的新网络——Darknet53来进行特征提取,提高了网络每秒最高的测量浮点运算，使网络可以更加充分利用计算机的算力，提高了检测的效率和速度[38]，YOLOv4充分吸取了之前的目标识别算法的优势，进行了大量的调试，尝试了大量的算法调整技巧，进一步提升了算法的综合性能[39]。SDD提出了一种使用单个深度神经网络目标检测算法，在检测时，网络为每个默认框中的每个对象的类别存在的概率生成一个指标分数，并将调整框来匹配对象的形状，网络结合了不同分辨率下的特征图对目标进行预测，提升了网络对不同尺度的目标的检测精准度。SDD训练的过程较为简单，可以直接集成到需要检测组件的系统中，与需要额外对象建议的方法相比，具有更高的准确度和高快的速度[40]。

目标检测算法在结肠息肉诊断上也有许多应用。结肠息肉目标检测主要是指在结肠镜检查的过程中实时检测视频流，在这个过程中图像检测算法主要是对输入的图像特征进行提取和分类，在其中找出疑似的结肠息肉区域并进行标注，为医生的诊断提供辅助，在一定程度上降低医生的工作量。

Rahim等人将一种卷积神经网络模型应用于息肉检测。卷积神经网络是一个典型的两阶段网络，及它在检测的过程中对于图像的提取和分类是分为两个阶段进行，这就使它的检测的准确率较高，但是检测速率较慢， Rahim等人在NVIDIA Titan RTX GPU上进行实验，平均处理时间为0.6秒/帧，这样的检测速率对于实时检测则稍显吃力[41]。

Zhang等提出了一种基于回归的卷积神经网络，用于结肠息肉的检查。它用从视频中提取的息肉图像调整快速检测算法ResYOLO，通过高效卷积算法的跟踪器并入视频的时间信息，进一步细化ResYOLO的检测结果。在AsuMayoDB的内窥镜视频中提取图像帧进行检测结果评估，算法能以每秒6.5帧的速度处理图像，准确率为88.6%[42]。

Shin等提出了一个基于R-CNN并结合最近深度CNN模型的息肉检测算法，该算法是直接将整个图像作为对象分析，而不是像传统目标检测算法对划分后的部分图像帧进行分析，但是由于卷积神经网络的深度较深，导致运算步骤较多，检测速度较慢，每帧处理时间为0.39s[43]，这就使该算法无法满足实时结肠息肉检测的功能。

Yu等提出了一种在线和离线3D深度学习算法，通过3D-FCN从结肠镜视频流中提取息肉，3D网络可以有效的学习视频流图像帧的时空特征，集成在线和离线学习的融合模型可以有效减少检测结果中的假正例数量，提高了检测精度[35]。

### 1.3.2结肠息肉语义分割研究现状

语义分割不同于目标识别，对图像的处理更加精细，需要分割图像中属于目标的像素。基于深度学习的语义分割主要有像素块分类和全卷积分类（FCN）两个方法。由于早期的分类网络中全连接层需要输入指定尺寸的图像，像素块分类采用同像素与其周围的像素组成的像素块来判断像素的类别。之后，Long等人提出了FCN，该网络不采用全连接层处理像素分类，输入图像的尺寸也不受限制，而且相对于像素块分类速度得到了很大的提升[44]。之后的大部分语义分割的研究都是基于FCN进行的。FCN中的池化层虽然能增加感受野的范围从而能更加广泛地获取图像信息，但是，会导致丢失移不动位置信息，这对于语义分割是不利的。针对此问题主要有两种解决思路，其一是采用编码-解码结构，先通过池化来弱化空间维度的信息，再利用解码层来恢复目标的细节和空间维度信息，这样既兼顾了池化层的优点，又恢复了像素的位置信息。U-net就是采用此结构的典例[45]。另一种就是采用扩张卷积和空洞卷积，它不采用池化层，利用扩张卷积和空洞卷积来在保留图像空间信息的同时使网络能更将广泛地获取图像。经过多年发展，语义分割在医学图像中有着许多的研究。

[Tashk](https://ieeexplore.ieee.org/author/37541481400)等人提出了一种新U-Net架构全自动息肉检测神经网络，其中包含一个全3D层，这使网络可以处理多光谱、高光谱和视频流的图像数据，网络采用了一个骰子预测输出层，采用概率方法预测，进一步提高了网络的预测能力[46]。

Feng等人提出了一种新的梯形架构实时结肠镜图像息肉分割网络，该网络在编码阶段对图像进行空间图像信息提取，然后用全注意力模块和多尺度融合模块来融合图像的不同尺度的特征，该网络的速度和息肉图像分割性能均优于U-Net[47]。

S.Sornapudi等人提出了一种改进的基于区域卷积神经网络的息肉分割网络(SSN)，通过特征提取和微调技术，并用预训练模型提取息肉的图像特征，在息肉数据集的实验中获得了不错的息肉定位效果[48]。

Fan等人提出了一种并行反向注意力的息肉图像分割网络（PraNet）。该网络并行部分解码器来融合高层次层中的特征，并基于融合特征为底层进初始引导。同时，用反向注意的方法提取边界的特征，建立区域与边界线索的关系。在实验中，该网络有着不错的分割精度和效率，并有着良好的迁移泛化能力[49]。

[Ji](https://arxiv.org/search/cs?searchtype=author&query=Ji%2C+G)等人提出了一种新的渐进式归一化自注意力息肉分割网络（PNS-Net），该网络基于归一化自注意力模块，并配备了递归和卷积神经网络模块，增加了网络的感受野，使之能充分利用息肉视频中的时间和空间信息。在实验中，PNS-Net网络可无需后处理地实时学习视频流中的特征[50]。

## 1.4主要研究内容和章节安排

针对结肠内窥镜对操作者专业技能和经验积累要求较高、操作者工作强度较大的问题，本文提出了一种履带式微型肠道机器人，此机器人能够在肠道内主动运动，使操作者能够更简单地控制机器人在肠道的位置和内窥镜的朝向。同时，为了提高结肠息肉的诊断效率、降低结肠息肉漏诊率、进一步降低医师的工作强度，本文将基于深度学习的机器视觉技术应用于结肠息肉，基于YOLOv5训练了一个结肠息肉的目标识别和语义分割的网络，并开发了集成微型肠道机器人控制和结肠镜图像显示与息肉诊断的人机交互软件，使操作者通过上位机软件就能实现对微型肠道机器人的运动控制、肠镜的显示和息肉的检查。本文章节的安排如下：

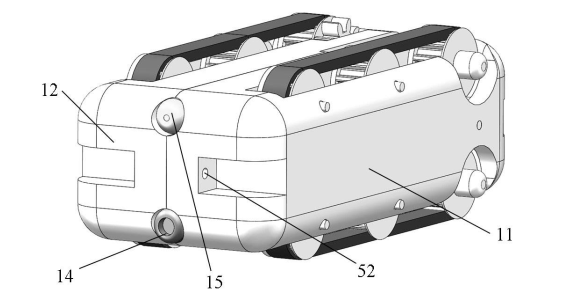
第一章主要介绍课题的研究背景和研究意义，阐述当前肠道机器人的研究现状和基于深度学习的目标识别和语义分割算法在医学诊断领域的应用现状，分析结肠镜检查的问题，明确本文的研究内容。

第二章主要介绍履带式微型肠道机器人的机械结构设计和力学性能分析。首先介绍履带式微型肠道机器人的机械结构和工作原理，然后根据工作环境对驱动电机和元件材料进行了选型，设计其中主要传动环节的相关参数，最后对机械结构进行建模并用软件模拟工作时的载荷，对关键部件的力学情况进行有限元分析，验证设计的可行性。

第三章主要介绍基于YOLOv5结肠息肉识别和语义分割功能的实现。首先介绍YOLOv5网络结构，之后阐述结肠息肉图像数据集制作、实验环境配置、训练参数设置的过程，然后对训练的结果进行分析，最后将结肠息肉视频数据输入检测网络进行检测，验证模型实时检测结肠镜图像的可行性。

第四章主要介绍上位机软件的开发和微型肠道机器人的运动控制电路。对于上位机软件开发部分，首先介绍上位机软件的主要功能，然后阐述了各个功能的实现方式和实现效果。对于运动控制电路部分，首先介绍微型肠道机器人运动控制的需求，然后根据工况和需求对控制硬件进行选型，最后介绍了控制的实现原理。

# 第二章 机械结构设计

针对结肠内部空间狭小，肠壁凹凸不平且附着黏膜消化液的特点，提出了一种微型履带式肠道机器人，其主要特征包括壳体，以及壳体内部的驱动电机、传动部件和运动结构。本章主要内容为机器人的机械结构设计和关键部件的工况仿真以及力学性能分析。

## 2.1机械结构总体构思

图2-2微型履带式肠道机器人驱动结构示意图

图2-1微型履带式肠道机器人整体外观正面示意图

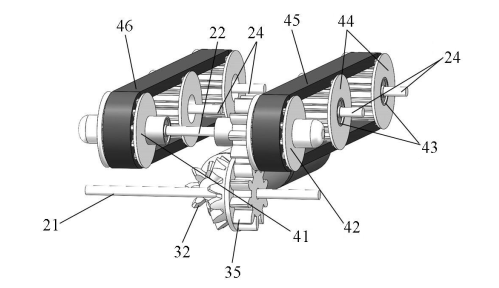
如图2-1所示，机器人外部包覆壳体11、12，肠镜摄像头安装在14处，15是辅助工具预留孔，可以搭载活检钳、气液输入管等，主动运动通过上下两组从动同步履带实现，使上下段均能沿肠壁运动，减少了机器人翻车后无法运动的风险，增加了运动的稳定性。如图2-2所示，在运动的过程中上下从动同步履带的履带轮的转动方向相反，所以机器人内部的传动系统设计为中心对称的形式。上下两组从动同步履带的结构完全相同，以上从动同步履带为例（图2-3），控制器通过信号控制驱动电机31正/反转，从而带动锥齿轮32正/反转，进而使齿轮35反/正转，最终使同步履带轮轴22正反转，从而实现上从同步动履带正/反向运动。下从动同步履带上下两组从动同步履带的结构完全相同组传动方式对称，运动方向相反，实现下方从动同步履带的反/正向运动，由此带动机器人的运动。

图2-1微型履带式肠道机器人上从动同步履带结构示意图

## 2.2传动机构设计与选型

为尽可能缩小机器人的体积，机器人采用德国福尔哈贝传动公司的0615N4.5S直流微电机做驱动器，配备06/1 64:1行星减速箱，该电机为贵金属换向直流微电机。具体参数见表2-1。

### 2.2.1直齿锥齿轮传动设计

驱动电机的动力通过锥齿轮32输出，然后通过两个完全的镜像的锥齿轮实现交叉轴运动传递、减速、将动力平等地分配到上下两组同步履带的三项功能，直齿锥齿轮设计参数见表2-2。

电机的额定转速为2500r/min，减速箱减速比为64:1，则锥齿轮的转速为39r/min，两个从动齿轮的转速为26r/min。

根据表2-1可知电机连续转矩为MN=25mNm，峰值转矩为MH=35mNm，根据圆周力计算公式[51]，可知

（公式2-1）

（公式2-2）

（公式2-3）

其中T1为主动齿轮传递的名义转矩，dm1是齿宽中点节线处的直径，K为修正系数，Ftc为计算圆周力。可得输入转矩为连续转矩时计算载荷为9.64N，输入转矩为峰值转矩时的计算载荷为13.50N。由于直齿锥齿轮的动力传输到两个相同的从动直齿锥齿轮，单一接触面的峰值计算载荷为6.75N。

### 2.2.2直齿圆柱齿轮传动设计

查阅资料得知圆锥直齿轮的效率为0.92-0.98[52]，则圆柱齿轮的输入功率范围是0.094W-0.1W,由于圆锥直齿轮输出到两组镜像圆柱直齿轮，故一组圆柱直齿轮的功率为0.047W-0.05W，根据齿轮工作转矩计算公式[51]，

（公式2-4）

可知，圆柱直齿轮工作时最大载荷为24.5mNm，主从齿轮的转速均为26r/min。

直齿圆柱齿轮设计参数见表2-3。

### 2.2.3同步履带设计

采用T型同步带作为机器人的运动机构可以有效减少带和带轮之间的打滑现象，可以大大提升传动效率，同时也继承了履带接比压小，通过性强的有点。同步带使用标准的MXL带，同步履带及带轮的设计参数见表2-4。

同步带带轮的转速为26r/min，根据同步带尺寸，可知，机器人在电机额定转速下的移动速度为16.31mm/s

表2- 5福尔哈贝传动0615N4.5S+06/1 64:1直流减速微电机参数

|  |  |  |
| --- | --- | --- |
| 参数名称 | 数值 | 符号 |
| 名义电压 |  |  |
| 额定电流 |  |  |
| 额定转速 |  |  |
| 空载转速 | *20000r/min* |  |
| 减速箱减速比 |  |  |
| 减速箱的连续转矩 |  |  |
| 减速箱的峰值转矩 |  |  |
| 总长度 |  |  |

表注：定值基于名义电压和环境温度22℃条件下。

表2- 6直齿锥齿轮设计参数

|  |  |  |
| --- | --- | --- |
| 参数名称 | 主动齿轮 | 从动齿轮 |
| 齿数 |  |  |
| 模数 | 1 | |
| 大端分度圆直径 |  |  |
| 锥顶角 |  |  |
| 压力角 |  | |
| 锥顶距 |  | |
| 齿宽 |  | |
| 齿宽系数 | 0.35 | |
| 齿顶高 |  | |
| 齿根高 |  | |
| 全齿高 |  | |

表2- 7圆柱直齿轮设计参数

|  |  |  |
| --- | --- | --- |
| 参数名称 | 主动齿轮 | 从动齿轮 |
| 齿数 |  |  |
| 模数 | 1 | |
| 分度圆直径 |  | |
| 压力角 | 20° | |
| 齿顶高 |  | |
| 齿根高 |  | |
| 全齿高 |  | |
| 齿厚 |  | |
| 标准中心距 |  | |

表2- 8同步带及带轮设计参数

|  |  |  |  |
| --- | --- | --- | --- |
| 参数名称 | 同步带轮 | 同步带（MXL） | |
| 齿数 |  |  | |
| 齿距 |  | | |
| 节顶距 |  | | - |
| 齿高 | 0.69mm | |  |
| 带轮厚度 |  | | - |
| 同步带宽 | - | |  |
| 同步带厚 | - | |  |

## 2.3机器人的材料选择及有限元仿真

微型肠道机器人样机的壳体、齿轮和内部的同步带轮均由ABS材料3D打印得到，同步履带由聚二甲基硅氧烷制成。ABS是一种强度高、韧性好的高分子材料，得益于其耐热性、易于热塑加工成型、制品尺寸稳定的特点，ABS是3D打印的主要耗材之一。聚二甲基硅氧烷是一种无毒、具有良好的化学稳定性的高分子材料，在药品、日化用品等领域广泛应用。

本文采用Altair公司开发的Inspire软件对微型肠道机器人的关键结构进行力学性能分析，Inspire是一款面向增材制造，以拓扑优化算法为基础的轻量化软件，能够实现几何建模、施加载荷和约束、有限元分析、轻量化设计等功能。本文主要利用Inspire软件对微型肠道机器人中锥齿轮32和零件35进行载荷约束仿真和力学有限元分析，两个零件的建模由SolidWorks完成。

### 2.3.1直齿锥齿轮32的有限元分析

设置零件的材料为Plastic（ABS），根据计算结果施加约束与载荷，此处的约束设置在直齿锥齿轮32与电机输出轴的连接处，载荷均取峰值载荷，分析单元尺寸设置为0.1mm。

分析结果：

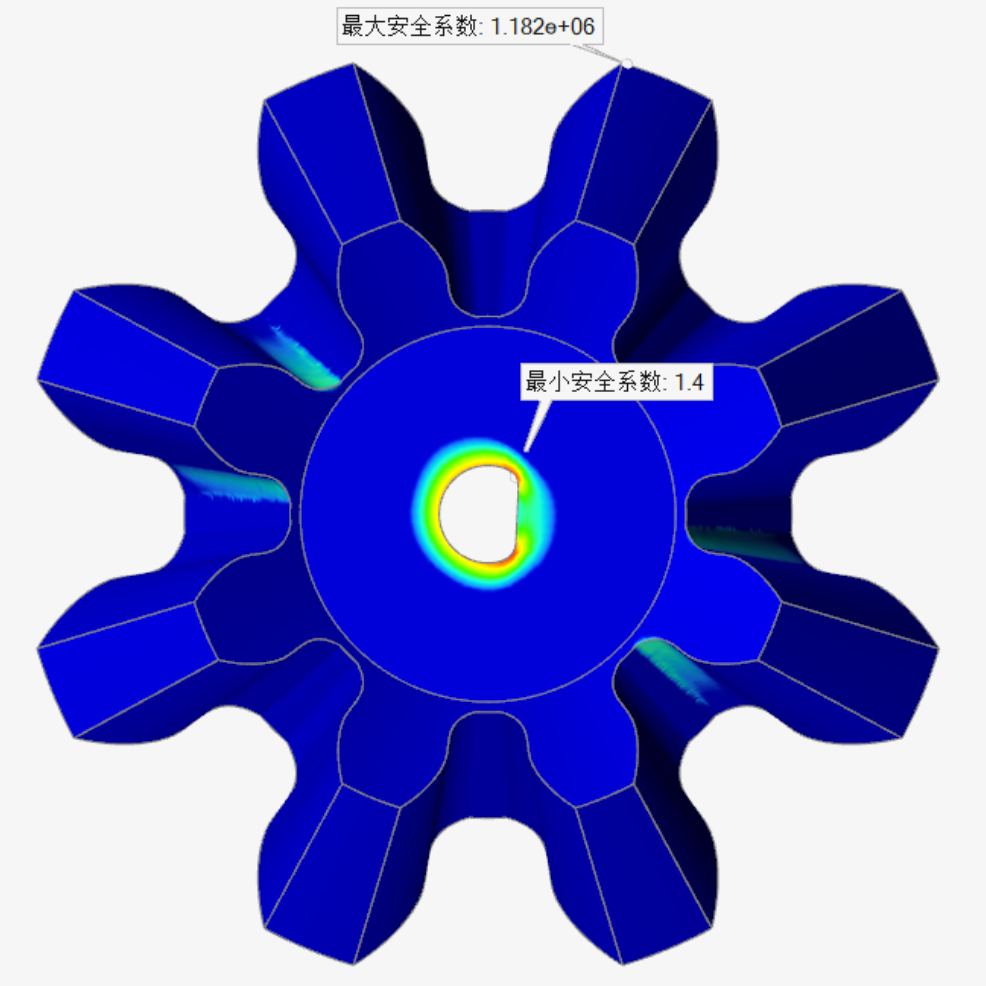
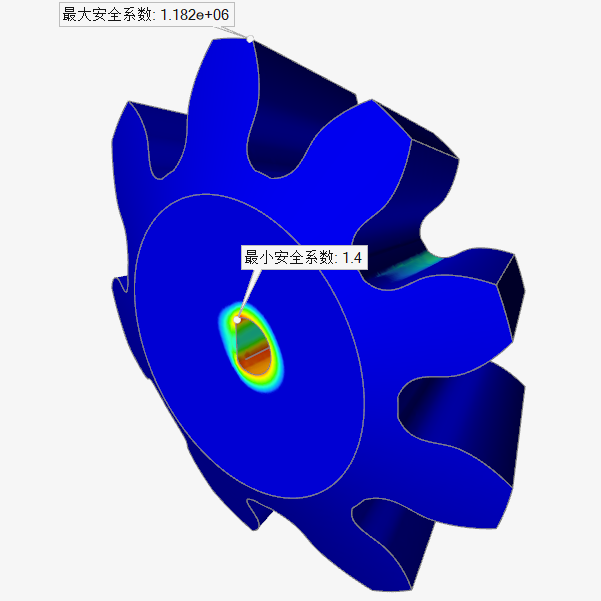
1. 安全系数：最小安全系数为1.4，最大安全系数为1.182×106；1

图2-2直齿锥齿轮安全系数分析结果图

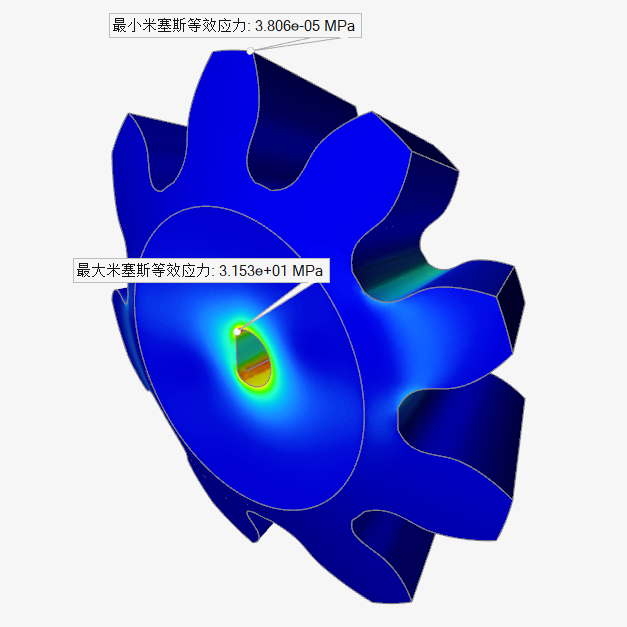
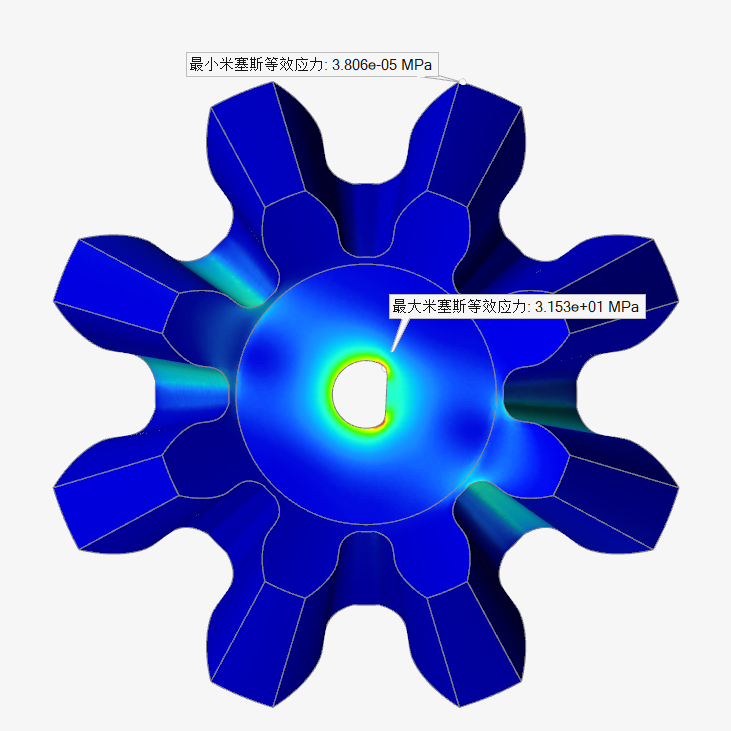
1. 米塞斯等效应力：最大米塞斯等效应力为31.53MPa，最小米塞斯应力为3.806×10-5MPa；

图2-3直齿锥齿轮的米赛斯等效应力分析结果图

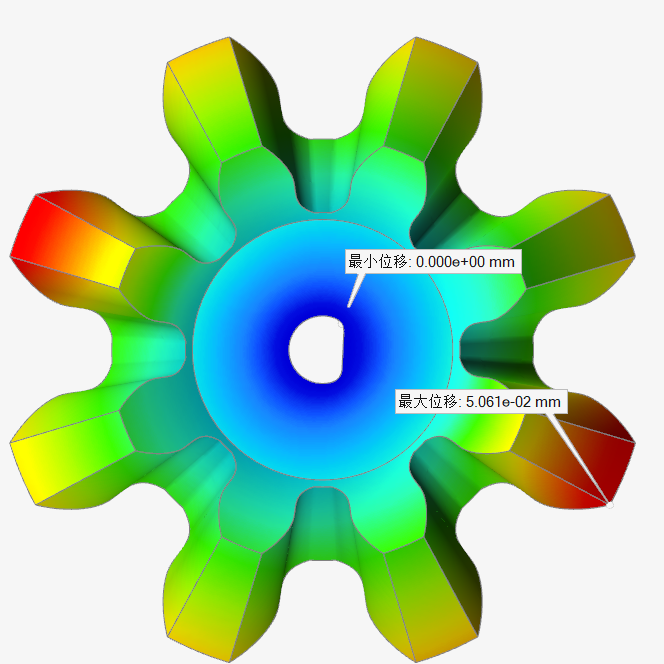
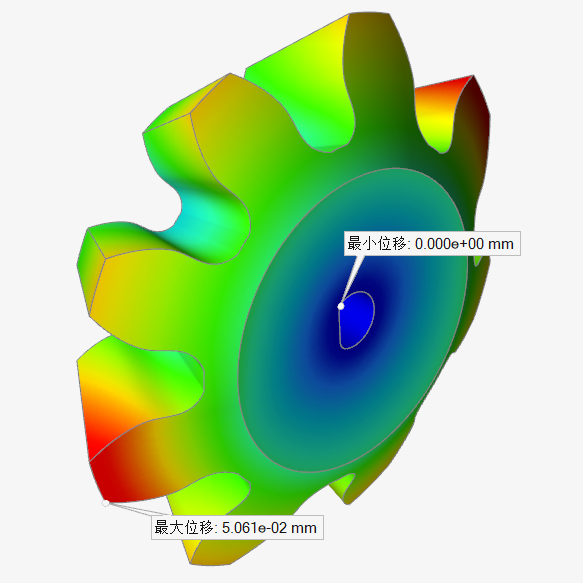
1. 位移：最大位移为5.375×10-2mm，最小位移为0mm。

图2-4直齿锥齿轮的位移分析结果图

### 2.3.2 零件35的有限元分析

设置零件的材料为Plastic（ABS），根据计算结果施加约束与载荷，此处的约束设置在零件35与轴承连接处，分析单元尺寸设置为0.2mm。

分析结果：

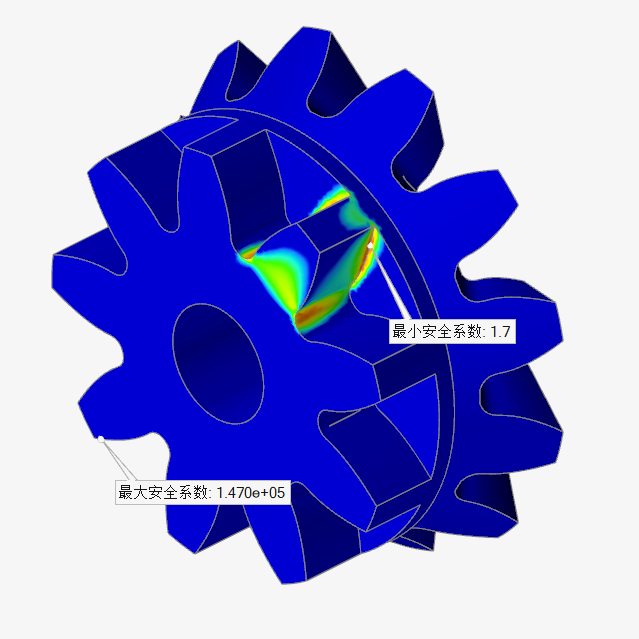
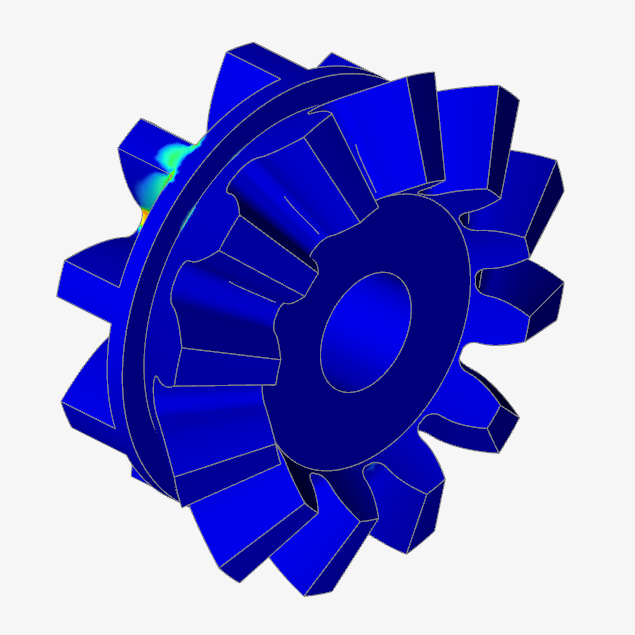
1. 安全系数：最小安全系数为1.7，最大安全系数为1.47×105；

图2-5零件35安全系数分析结果图

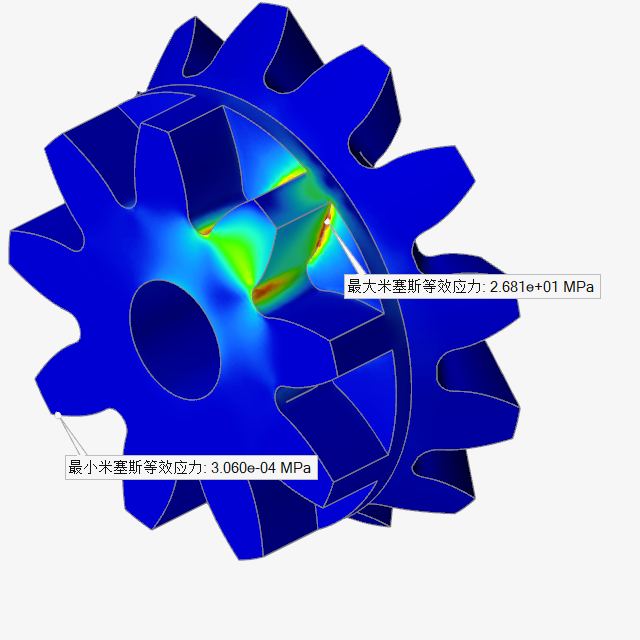
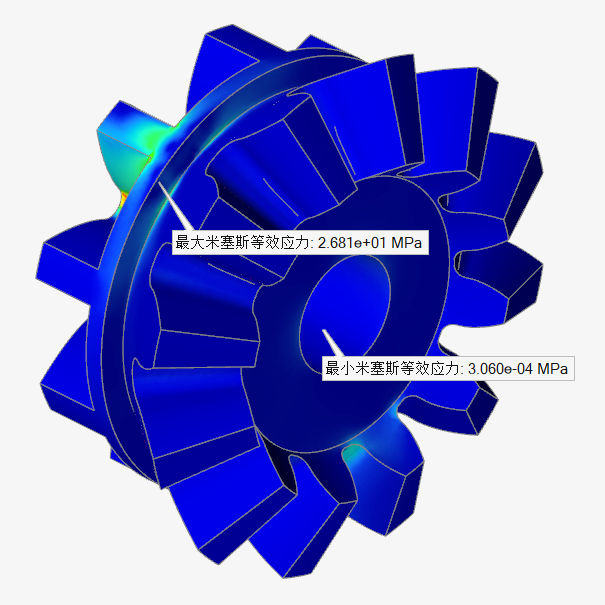
1. 米塞斯等效应力：最大米塞斯等效应力为26.81MPa，最小为3.06×10-4MPa；

图2-6零件35的米赛斯等效应力分析结果图

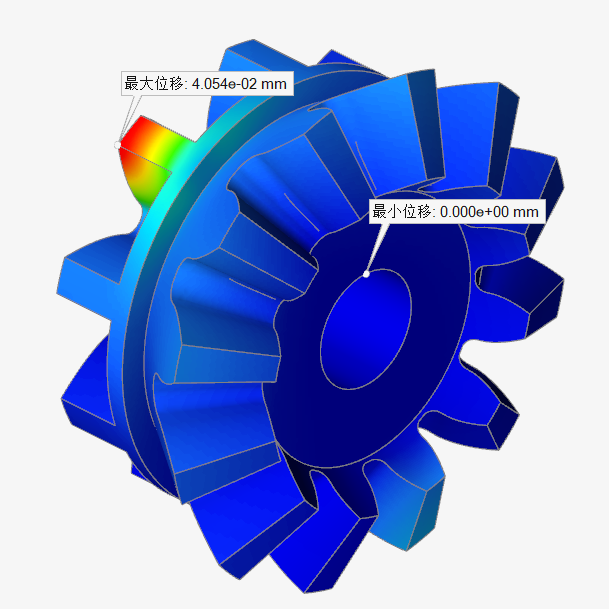
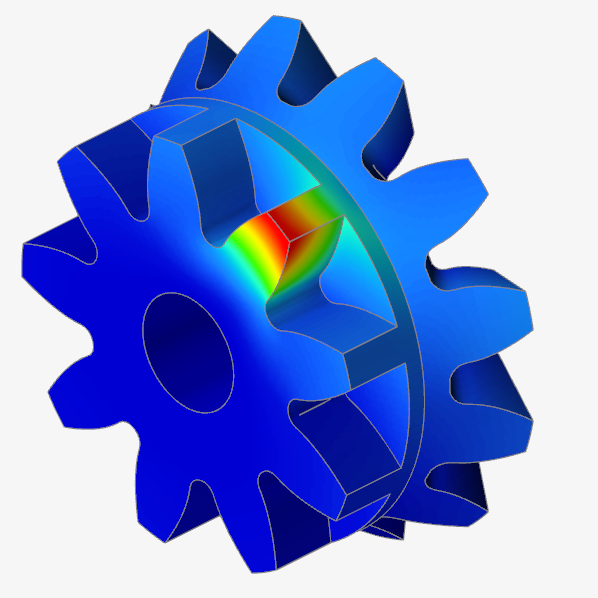
1. 位移：最大位移为4.054×10-2mm,最小位移为0mm。

图2-7零件35的位移分析结果图

综上，两个零件的材料为ABS，在工况为峰值载荷时，零件32的最大米塞斯应力31.53MPa，最小安全系数为1.4，最大位移为5.375×10-2mm；零件35的最大米赛斯应力为26.81MPa,最小安全系数为1.7，最大位移为4.054×10-2mm，强度不超过材料的屈服应力，满足实际工况的强度要求。

## 2.4本章小结

本章主要介绍了微型肠道机器人内部的机械结构以及传动原理，然后对其中的直齿齿轮、直齿圆柱齿轮和同步履带的参数进行了设计并计算了这些结构所受的载荷，最后在Inspire软件中对关键零件进行力学性能的分析，根据分析结果可知，在微型肠道机器人的峰值载荷工况下，零件的设计强度能够满足要求。

# 第三章 结肠镜息肉的识别和语义分割

结肠镜息肉的识别和语义分割主要目的有两个，一个就是在结肠镜图像中找到疑似的息肉组织，并用矩形框进行指示，其二就是对识别矩形框内属于息肉像素的图像进行分割。本文采用的是YOLOv5网络对结肠镜息肉进行识别和语义分割，在Kvasir-SEG数据集[47]中达到了不错的检测性能。

## 3.1YOLO算法介绍

在本文中采用的是YOLOv5网络7.0版本，该算法[基于YOLOv5](https://github.com/ultralytics/yolov5/tree/v7.0），YOLOv5-7.0基于YOLOv5)网络集成支持图像分割，并且提供了5种预训练模型，最小模型YOLOv5n-seg在CPU端的处理速度最快能达到62.7ms，[最大模型YOLOv5x-seg的mAPbox@.5:.95=50.7%、mAPmask@.5:.95=41.4%](mailto:最大模型YOLOv5x-seg的mAPbox@.5:.95=50.7%25、mAPmask@.5:.95=41.4%25)。

## 3.2网络模型结构

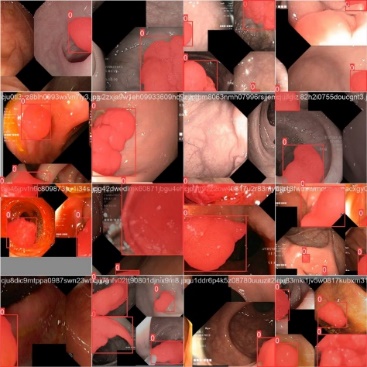
本文采用的网络为YOLOv5s-seg，它是基于YOLOv5s网络模型迭代改进后应用于图像分割的轻量化网络。网络主要由主干网络、颈部网络和头部预测网络三部分组成。

图3-1 Mosaic数据增强的训练图像

YOLOv5会在图像输入端进行处理，主要是Mosaic数据增强、Anchor框的计算和图片的自适应缩放。Mosaic数据增强通过将随机的4张图片进行随机缩放、裁剪、排布后进行拼接，经过这些操作后，图像中的小目标显著增加，丰富了检测物体的背景并且扩展了数据集，这将有利于网络的鲁棒性的提升，同时减少了数据计算量，从而降低了部署设备的要求，有利于算法的推广[39]。如图3-1所示，这是一次训练选取16个样本时，经过Mosaic数据增强后的图片，其中的每一个样本都是由4张随机处理的图片随机拼接而成。在模型训练时，网络会在初始的Anchor框格的基础上输出预测框，然后将预测框和真实的物体框进行比较，量化两者的差距，迭代网络参数，而初始的Anchor框就是在输入端中根据训练集来进行计算生成的。在图像检测中，原始的图像数据的长宽往往是不同的，所以常常在输入检测网络前需要将图片统一成固定的尺寸，在这个过程中，图片被按照一定比例缩放，然后在多余的地方进行填充，为避免过多填充造成信息冗余而影响推理速度，在输入端会根据原始图片的尺寸进行自适应缩放并填充最小面积的黑边，。

主干网络采用CSPDarknet53作为主干，主要功能是对输入的图像进行特征提取。由CBS模块、CSP模块、SPPF模块等组成。CBS模块中包含一个卷积层、一个Batch归一化层和激活函数，主要用于不同层次的信息特征提取。CSP模块主要借鉴了CSPNet，包括3个卷积层和多个残差组件，主要用于优化反向传播的路径，提升网络的学习能力的同时减少计算量和内存占用。SPPF（Spatial Pyramid Pooling-Fast）模块借鉴了空间金字塔池化的思想，由四个并联支路组成，采用多种尺寸的池化层以最大池化的方式进行多尺度融合，通过最大池化可以有效避免图像在剪裁缩放过程中图像失真等问题，增加了感受野，通过改进池化连接的方式，提高了检测的速度。

由于深层的特征图具有较强的语义特征和较弱的位置信息，浅层的特征图有较强的位置信息和较弱的语义特征，为了使网络充分提取数据各个层次的特征，颈部特征金字塔网络采用了特征金字塔网络（FPN）和路径聚合网络（PAN）[48]结合的结构，FPN能从深层将语义特征向浅层传递，从而增加多尺度的语义特征。PAN则从浅层像深层传达位置信息，增强了网络多尺度的定位能力。两个网络的融合可以使网络实现多尺度特征获取与融合，提高了网络的特征融合能力。

## 3.2 实验与分析

### 3.2.1实验环境及训练参数

本章的所有实验均在配置GeForce RTX 3060 Laptop GPU，i7-12700H [CPU@2.3GHz，安装了cuda11.3](mailto:CPU@2.3GHz的计算机上进行，安装了cuda11.3)、cudnn8.2.1的Pytorch1.12.1框架的计算机上进行。

### 3.2.2数据集

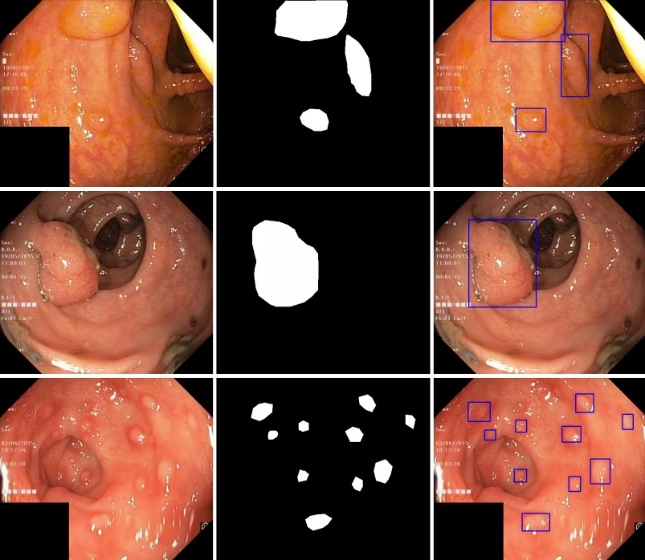
训练所用的结肠息肉的数据集是Kvasir-SEG[47]，数据集中包含分辨率从332×487到1920×1072的1000张息肉肠镜图片，结肠内窥镜镜图像由挪威Vestre Viken Health Trust的肠胃病专家收集并用矩形框标记，Kvasir-SEG数据集将数据上传到Labelbox程序[49]，由一位工程师和一位医生手动进行息肉像素分割。Kvasir-SEG数据集中包含了图像的原始图像和分割后的mask图像，以及包含息肉边框信息的JSON文件。

图3-2 Kvasir-Seg内息肉图像和mask图像示例[47]

注：3列图像由左至右依次为：肠镜原始图像、mask图像、JSON文件转化后的息肉图像

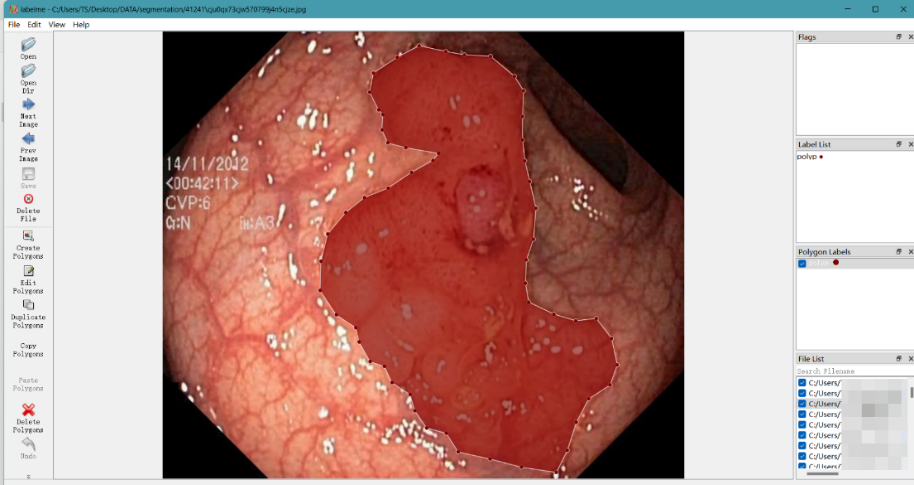
由于Kvasir-SEG公开的数据中并未包含语义分割后的数据集，本文将数据集中的原始图像输入到Labelme程序[50]中，对照数据集中的mask图像手动进行息肉像素的分割，得到包含像素分割节点信息的JSON文件后，通过Python脚本实现JSON文件向txt文件的转换，并将1000张图片按照训练集4:1的比例随机划分，为数据集输入YOLOv5网络训练做准备。

图3-3 Labelme对Kvasir-SEG数据集手动分割息肉像素

### 3.2.3评判标准

对于训练结果的评判标准主要有精确率、准确率、召回率、F1和均值平均精度。

对于训练的结果可以分为四类：真正例（TP）：正例正确预测为正例；真负例（TN）负例正确预测为负例；假正例（FP）：正例错误预测为负例；假负例（FN）：负例错误预测为正例。

精确率（Precision）是所有正例的预测结果中正确预测正例的占比：

准确率（Accuracy）是指所有预测结果中正确预测的占比：

召回率（Recall）是指实际正例中预测为正例的占比：

均值平均精度（mean Average Precision, mAP）是各个类别平均精度的均值：

### 3.2.4实验结果

将包含Kvasir-SEG训练集信息的txt文件输入YOLOv5网络，训练时，采用YOLOv5s-seg来初始化网络权重，训练轮次设置为300，每次训练输入神经网络的图片数为16，其他均采用默认值。训练的初始学习率和周期学习率为0.01。

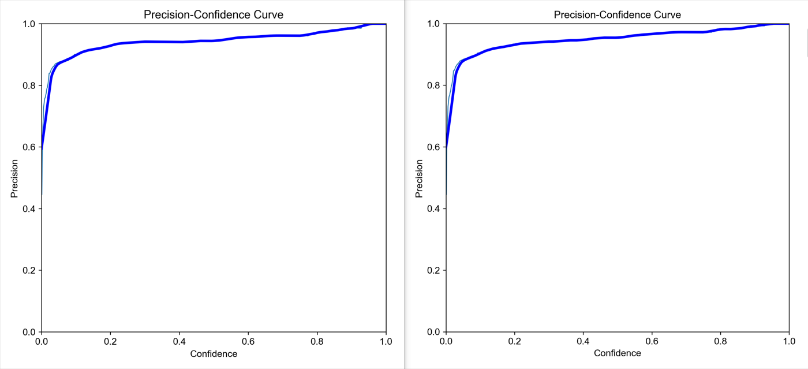
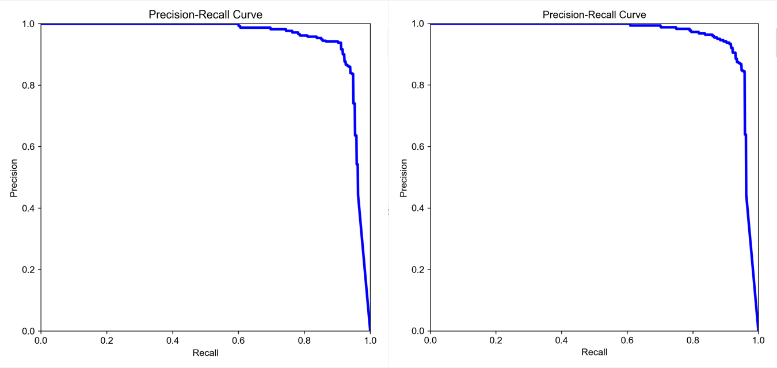
1. P曲线-精确率与置信度的关系曲线。

图3-4精确率与置信度的关系曲线

注：左图为box的P曲线 右图为mask的P曲线

1. R曲线-召回率与置信度的关系曲线

图3-5召回率与置信度的关系曲线

注：左图为box的R曲线 右图为mask的R曲线

召回率与精确率是衡量模型识别精准程度的重要评价标准，从（1）（2）中可以看出，召回率和精确率是相互矛盾的衡量指标，召回率越低，精确率越高，反之亦然。提高IoU阈值，相当于提高了网络预测识别对象为正例的门槛，那么，网络所识别的正例是真正例的概率就更高，从而提高了精确率，与此同时，过高的IoU阈值，也增加了网络漏掉一些特征稍弱的真正例，从而降低了网络的召回率。

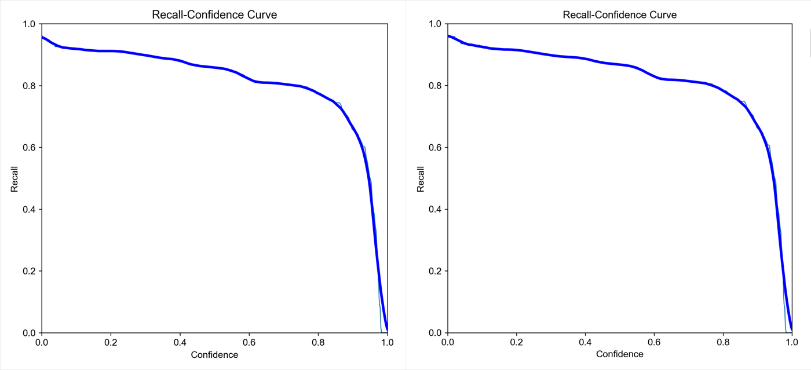
1. PR曲线-精确率与召回率的关系曲线。PR曲线与坐标轴所围成的面积即为平均精度（Average Precision，AP）。

图3-6精确率与召回率的关系曲线

注：左图为box的PR曲线 右图为mask的PR曲线

由PR曲线可知，YOLOv5网络在IoU阈值为0.5时对息肉目标识别的平均精度为95.1%,语义分割的平均精度为95.7%。

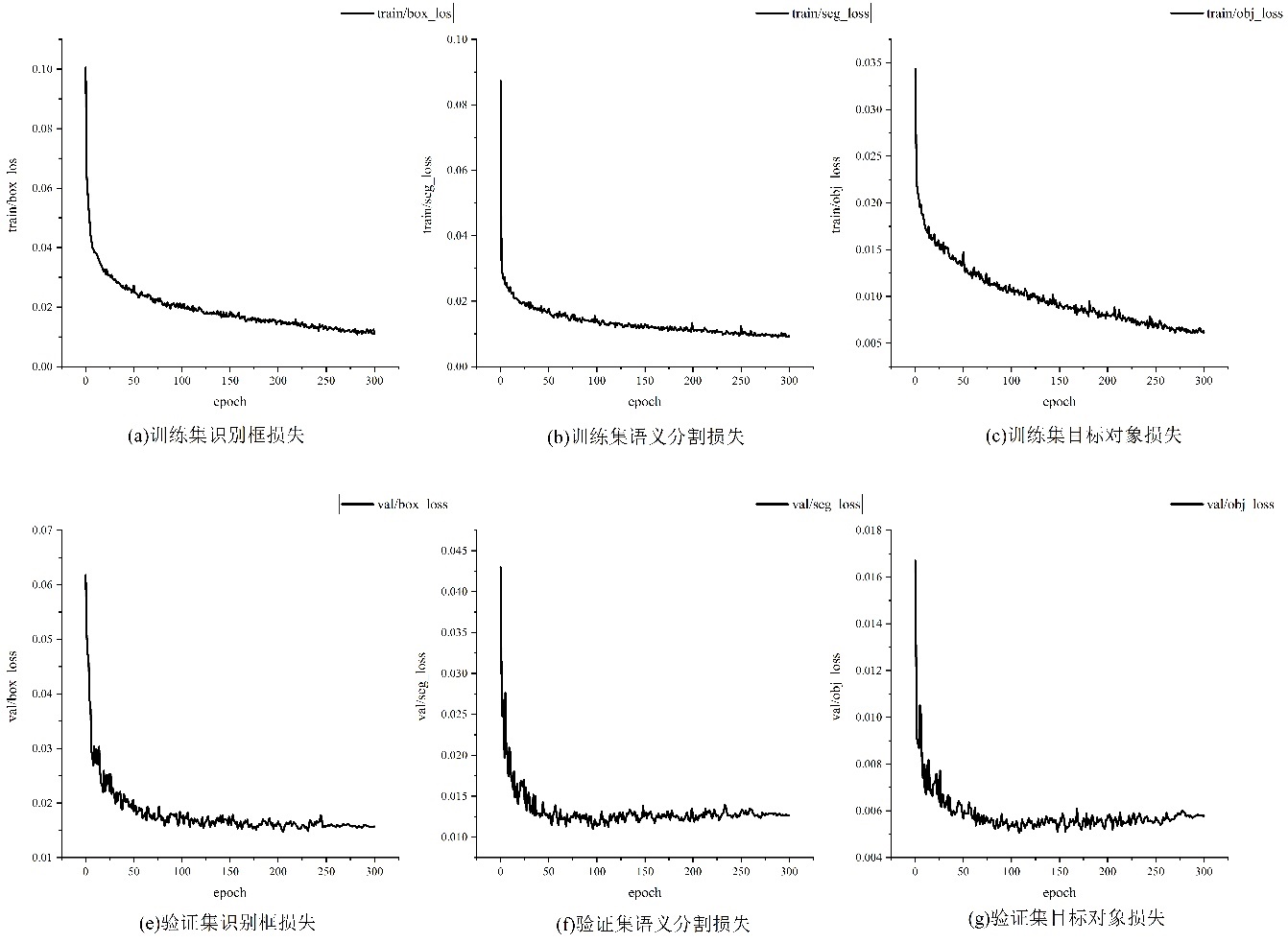
1. 损失曲线

图3-7损失曲线

根据损失曲线可知，训练集和验证集的各种损失总体上随着训练的进行先下降然后保持平稳，仅（f）和（g）中有轻微的上升趋势，这就说明在训练的过程中，模型对识别目标的特征总结逐渐准确，网络预测结果与真实的目标之间的偏差逐渐减小，由此可判断训练模型拟合适度。

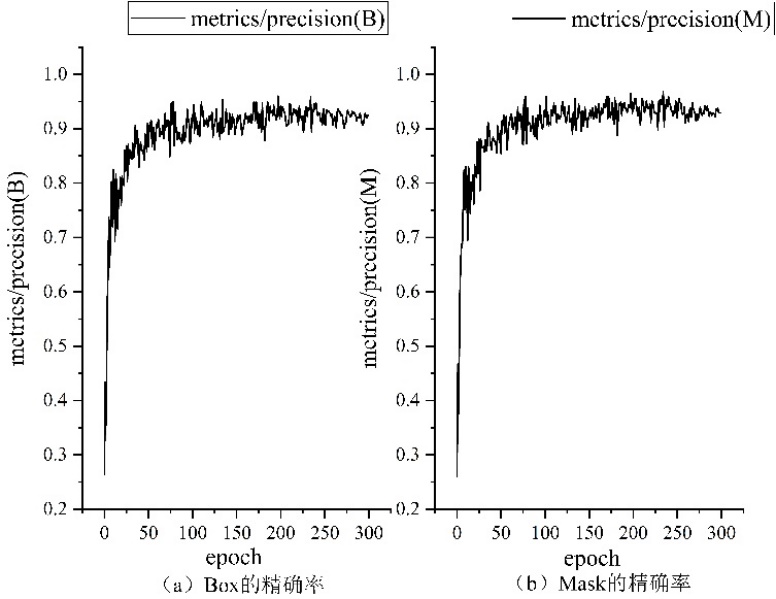
1. Precision曲线-各个轮次训练的精确率曲线

图3-8各个轮次训练的精确率曲线

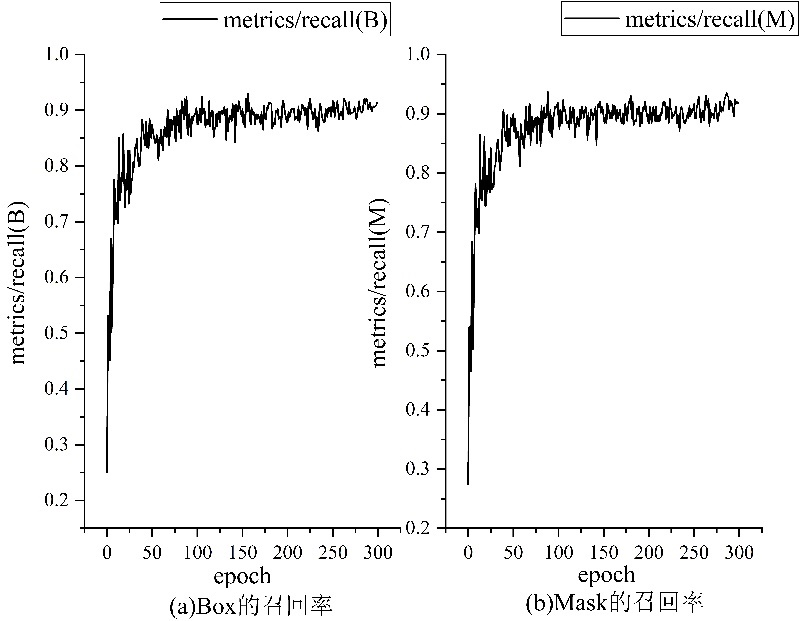
1. Recall曲线-各个轮次的召回率曲线

图3-9各个轮次的召回率曲线

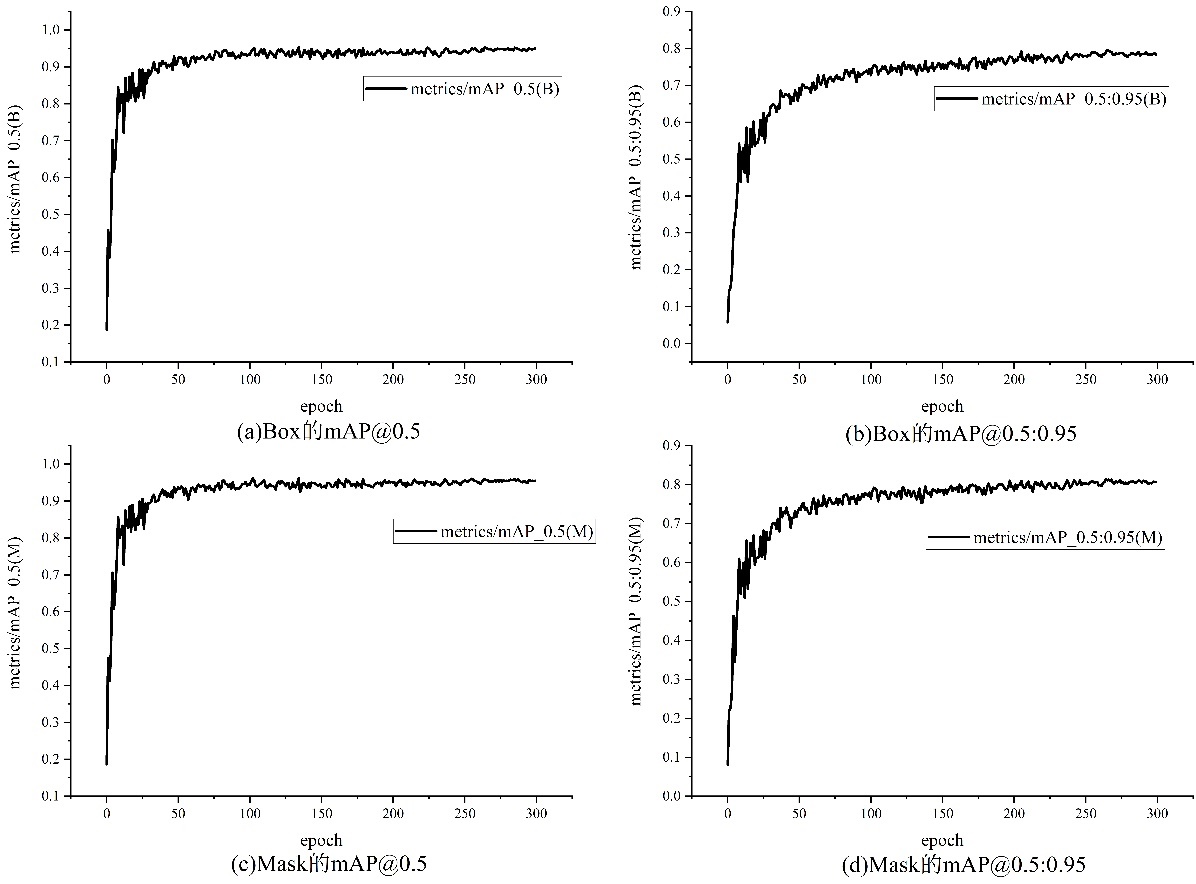
1. [mAP曲线-mAP@0.5是IoU阈值为0.5](mailto:mAP曲线-mAP@0.5是IoU阈值为0.5)时的均值平均精度，[mAP@0.5-0.95是IoU阈值在0.5-0.95之间，步长为0.05时mAP](mailto:mAP@0.5-0.95是IoU阈值在0.5-0.95之间，步长为0.05时mAP)的均值。

图3-10 mAP曲线

在300轮次训练后，在验证集中表现最好的一轮所得到的模型的息肉识别的精确率为93.752%，召回率为90.734%，[mAP@0.5为95.1](mailto:mAP@0.5为95.1)%；息肉语义分割的精确率为93.772%，召回率为91.033%，[mAP@0.5](mailto:mAP@0.5)为95.7% 。

为进一步验证所训练的模型在视频流检测时的时效性，本文采用Pablo Mesejo等人[51]分享的结肠镜息肉视频数据集来进行实验，视频流的分辨率为768×576，帧速率为25帧每秒，输入到模型后，视频流的分辨率压缩为640×480，在实验中，对1031帧视频数据的处理时间进行统计，平均的预处理时间为0.4ms，平均预测时间为7.9ms，平均后处理时间为0.9ms，单帧图像的处理速率远小于视频流原始的帧速率，所以训练所得的模型可以很好地实现实时检测的功能。

## 3.3本章小结

本章介绍了YOLOv5算法和YOLOv5-seg的网络结构，阐述了根据Kvasir-SEG数据集训练结肠息肉识别和语义分割模型的过程和实验结果，最后用息肉肠镜视频数据对模型检测的时效性进行了验证。息肉识别的均值平均为95.1%，语义识别的均值平均精度为95.7%，对视频流的帧检测速率平均为9.2ms。综上基于YOLOv5的息肉识别和语义分割网络有着很好的准确度和实时性。

# 第四章 机器人的运动控制及肠镜图像显示

这章主要介绍实现微型履带式结肠机器人主动运动的硬件电路，以及在Windows环境下开发的集成了运动控制和结肠镜息肉图像检测和语义分割的上位机软件。

## 4.1微型肠道机器人的上位机软件

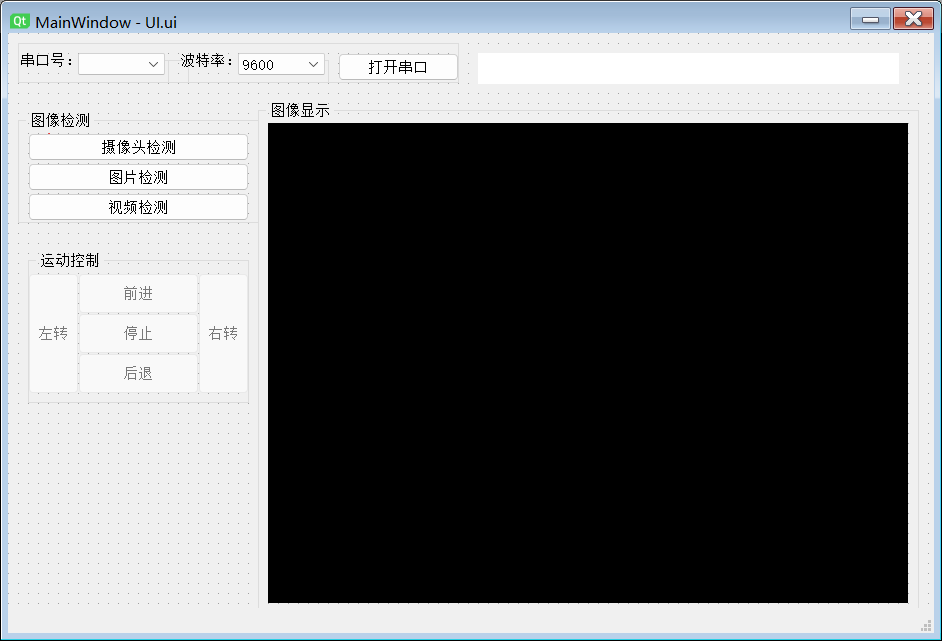
微型肠道机器人的上位机软件主要实现两个功能，其一就是作为PC端上位机，通过与下位机串口通信实现对微型肠道机器人的主动运动控制；其二就是接收微型肠道机器人所搭载的图像模块所传回的实时图像信息，并通过息肉识别和语义分割算法来实时显示肠镜图像。

图4-1微型肠道机器人上位机界面

本文的上位机软件是基于PyQt5.15开发。Qt是一个基于C++的跨平台完全面向对象的图形用户界面应用程序框架，可以用于Windows、Linux和macOS的图形界面开发，由于Qt易于扩展、可应用于组件编程的特点，目前已经广泛应用于智能汽车、智能制造、移动和嵌入式设备的多个领域。PyQt是Qt与Python结合的产物，它的主要功能是实现用Python开发Qt框架的用户界面程序。

上位机界面中微型肠道机器人的主动运动控制模块主要通过串口通信实现，上位机软件会自行扫描系统中可用的串口，使用者也可以根据需求和实际情况选择出口和设定串口波特率的数值，在开启串口后，运动控制部分按钮就被激活，点击按钮后，上位机软件就会通过串口向下位机的控制模块发送指定信号，控制直流微电机以一定的速度转动，实现微型肠道机器人完成指定的主动运动动作。

上位机的息肉检测部分根据输入的数据源类分成了三个类型：图像检测、视频检测、摄像头实时检测。上位机软件中集成了息肉检测网络，当操作者在上位机软件中输入了需要检测的数据源时，上位机软件会将数据喂入基于YOLOv5的息肉检测网络，然后，网络会将检测结果反馈到上位机软件的显示框，向操作者展示疑似息肉的目标检测和语义分割的结果。对于肠镜图片的检测，由于检测速度很快，上位机软件在网络检测完毕后再在上位机的显示框显示检测结果，并且会将检测结果自动保存到指定路径；对于视频流的检测，上位机会在网络检测完每一帧图像后将结果显示出来，实现了一边检测一边显示的功能，在视频流数据播放完毕或者暂停时，上位机软件会将检测结果以.mp4格式保存到指定路径，方便操作者后续的检查和诊断。

用肠镜视频流数据对上位机软件进行测试，在检测刚开始时，由于息肉检测网络初始化的原因，会有短时间的等待，在初始化完成之后，上位机可实现对视频数据检测结果的同步实时显示。

图4-2上位机息肉图像检测效果

## 4.2微型肠道机器人的主动运动控制硬件电路

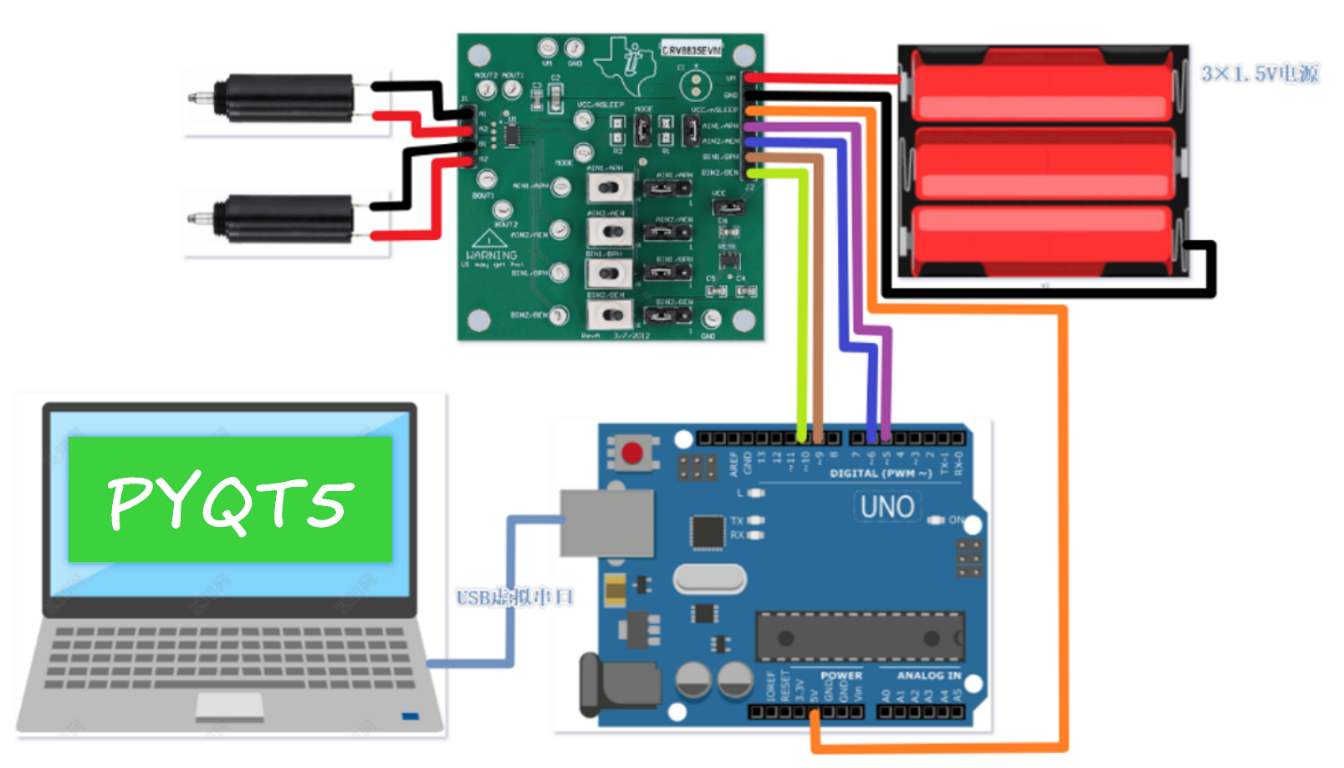
微型肠道机器人的运动控制的目的是使机器人能按照操作者的指令在肠道内主动运动，由于机器人的移动速度较慢，而且控制精度要求较低，所以对微型肠道机器人运动的控制采用的是开环控制，PC端上位机通过串口通信对机器人的运动实现遥控控制。

图4-3机器人主动运动控制硬件连接示意图

在硬件电路中，选用基于ATmega328P芯片的Arduino UNO作为控制器，该开发板上有14个数字输入\出引脚，其中有6个引脚可以实现PWM的输出，数字输出的高电平为5V，它可以在计算机上模拟生成一个USB串口，使ATmega328和计算机之间可以进行串口通信。电机的驱动器采用TEXAS INSTRUMENTS生产的DRV8835双H桥电机驱动器，其内部有两个H全桥，可以支持电源电压范围2-7V，可以通过PWM对该驱动器控制，实现对直流微电机的转速与转向的控制。

在控制过程中，Arduino接收到控制信号后，会根据控制信号在指定的输出引脚产生指定占空比的PWM信号输出到驱动器，DVR8835驱动根据接受到的PWM信号对直流微电机两端的电压大小和方向进行控制，从而实现对电机转向和转速的控制，最终实现对机器人运动的控制。

## 4.3本章小结

本章主要介绍了基于PyQt开发的上位机软件界面、功能和微型肠道机器人的主动运动控制硬件电路的构成。PC端上位机软件与微型肠道机器人控制器进行串口通信，实现对机器人主动运动的遥控控制。上位机软件的图像检测模块集成了YOLOv5的息肉检测网络，可以对肠镜图片、视频和肠道机器人传回的实时图像进行检测，对视频流可以一边检测一边显示检测结果，保证了肠道机器人息肉肠镜图像检测的实时性。

# 第五章 总结

本文针对传统电子肠镜仅导向段可控和对操作者专业技能要求高的特点，设计了一款履带式微型肠道式机器人，其精巧的结构能实现结肠内的主动运动，然后，根据设计的相关参数对驱动电机和制作材料进行选择，最后，利用Inspire对机器人内主要的机械结构进行力学性能分析，最后验证在峰值载荷的工况下，机器人相关结构的强度能够满足要求。

然后，针对临床上结肠息肉检测容易漏检的现状，本文将基于深度学习的机器视觉技术应用到肠道机器人的视觉模块，用YOLOv5算法训练了一个结肠息肉识别和语义分割的网络，在实验中，对于息肉识别的识别精度可达95.1%，语义分割的精度可达95.7%，同时，对于息肉肠镜视频的处理速度9.2ms/帧，能够实现实时的息肉识别和语义分割。

最后，为了简化操作步骤，便于人机交互，本文基于PyQt搭建了一套上位机系统，在PC端的上位机通过虚拟串口与机器人主动运动的控制器之间串口通信，使操作者可以在上位机上实现机器人运动的控制。上位机软件还集成了结肠息肉识别和语义分割网络，可以将从上位机返回的图像进行检测并实时反馈检测结果。

在以上的基础上，作者有以下展望：

（1）肠镜图像中，肠壁上由于附着有消化液，在灯光照射下，容易出现反光、光斑的情况，在图像的远端存在暗区和阴影，这些对于息肉识别的准确度都有干扰，还需要进一步的研究。

（2）由于互联网上公开的结肠息肉的数据集极为稀少，本文训练的数据集仅有1000张图片，这些图片并不能涵盖所有情况的息肉，因此需要构建一个更全面的息肉图像数据库。

（3）作者受限于相关专业知识的缺失，本文未对结肠息肉的类型、大小、病变程度等进行细分，在临床实践中，往往需要从息肉的形态对病情信息进行判断，所以可以结合医学知识对息肉进行更加细致的划分，使之更加具有使用意义。

**参考文献**

[1] Arnold M, Sierra M S, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality[J]. Gut, 2017,66(4):683-691.

[2] Sung H, Ferlay J, Siegel R L, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries[J]. CA: a cancer journal for clinicians, 2021,71(3):209-249.

[3] Zhang L, Cao F, Zhang G, Shi L, Chen S, Zhang Z, Zhi W, Ma T. Trends in and Predictions of Colorectal Cancer Incidence and Mortality in China From 1990 to 2025[J]. Frontiers in Oncology, 2019,9.

[4] Xie Y, Shi L, He X, Luo Y. Gastrointestinal cancers in China, the USA, and Europe[J]. Gastroenterology report, 2021,9(2):91-104.

[5] 中国结直肠癌诊疗规范（2020年版）[J]. 中国实用外科杂志, 2020,40(06):601-625.

[6] 中国早期结直肠癌筛查及内镜诊治指南(2014年,北京)\_柏愚[J].

[7] 徐智颖. 结肠镜检查单人操作法有效性及安全性的临床评价[D]. 中南大学, 2011.

[8] 毛伟芳, 顾秀珍. 结肠镜并发肠穿孔的原因分析及其防治[J]. 中华消化杂志, 2006,26(1):61.

[9] Joseph D A, DeGroff A S, Hayes N S, Wong F L, Plescia M. The Colorectal Cancer Control Program: partnering to increase population level screening[J]. Gastrointestinal endoscopy, 2011,73(3):429-434.

[10] 刘揆亮, 林香春, 吴静, 魏南, 王沧海, 刘红. 结肠镜下结直肠息肉漏诊危险因素分析[J]. 中国内镜杂志, 2013,19(09):914-918.

[11] 李一哲, 刘卓, 陈树银, 张雯倩, 任帅辉. 人工智能技术在医疗领域的应用及存在的问题[J]. 数字技术与应用, 2022,40(6):133-135.

[12] Mori Y, Kudo S, Berzin T M, Misawa M, Takeda K. Computer-aided diagnosis for colonoscopy[J]. Endoscopy, 2017,49(2017):813-819.

[13] Iddan G, Meron G, Glukhovsky A, Swain P. Wireless capsule endoscopy[J]. 2000,405(417).

[14] Ginsberg G G, Barkun A N, Bosco J J, Isenberg G A, Nguyen C C, Petersen B T, Silverman W B, Slivka A, Taitelbaum G. Wireless capsule endoscopyAugust 2002[J]. Gastrointestinal Endoscopy, 2002,56(5):621-624.

[15] Cristiano Spada C H M M, Amininejad L, Dekker E, Galmiche J, Frederic M, Johansson G W, Cesaro P, Costamagna G. Second-generation colon capsule endoscopy compared with colonoscopy [J]. Gastrointest Endosc, 2011,3(74):581-589.

[16]  PillCam™ COLON 2 System[EB/OL]. https://www.medtronic.com/covidien/en-us/products/capsule-endoscopy/pillcam-colon-2-system.html.

[17] 金山科技集团[EB/OL]. www.jinshangroup.net.

[18] 王雷, 李宜辉, 达四平, 赵晓晏. 国产胶囊内镜OMOM临床应用的进一步研究[J]. 中华医学杂志, 2006,86(6):421-423.

[19] Slatkin A B, Burdick J, Grundfest W, Authors A. The Development of a Robotic Endoscope[J]. Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots, 1995(2):162-171.

[20] P D, M C C, B A A E, Guglielmelli. Micromechatronics in Medicine[J]. 1996,2(1):137-148.

[21] Calderon A A, Ugalde J C, Zagal J C, Perez-Arancibia N O. Design, fabrication and control of a multi-material-multi-actuator soft robot inspired by burrowing worms[C]//: IEEE, 2016: 31-38

[22] Ikeuchi K, Yoshinaka K, Hashimoto S, Tomita N. Locomotion of medical micro robot with spiral ribs using mucus[C]//: IEEE, 1996: 217-222

[23] Quirini M, Menciassi A, Scapellato S, Dario P, Rieber F, Ho C, Schostek S, Schurr M O. Feasibility proof of a legged locomotion capsule for the GI tract[J]. Gastrointestinal Endoscopy, 2008,67(7):1153-1158.

[24] Valdastri P, Webster R J, Quaglia C, Quirini M, Menciassi A, Dario P. A New Mechanism for Mesoscale Legged Locomotion in Compliant Tubular Environments[J]. IEEE Transactions on Robotics, 2009,25(5):1047-1057.

[25] Tortora G, Valdastri P, Susilo E. Propeller-based wireless device for active capsular endoscopy in the gastric district[J]. Minimally Invasive Therapy & Allied Technologies, 2009,5(18):280-290.

[26] Carpi F, Galbiati S, Carpi A. Magnetic shells for gastrointestinal endoscopic capsules as a means to control their motion[J]. Biomedicine & Pharmacotherapy, 2006,60(8):370-374.

[27] Rey J, Ogata H, Hosoe N, Ohtsuka K, Ogata N, Ikeda K, Aihara H, Pangtay I, Hibi T, Kudo S, Tajiri H. Feasibility of stomach exploration with a guided capsule endoscope.[J]. Endoscopy, 2010,7(42):541-545.

[28] Swain P, Toor A, Volke F, Keller J, Gerber J, Rabinovitz E, Rothstein R I. Remote magnetic manipulation of a wireless capsule endoscope in the esophagus and stomach of humans (with videos)[J]. Gastrointestinal Endoscopy, 2010,71(7):1290-1293.

[29] Liao Z, Duan X, Xin L, Bo L, Wang X, Xiao G, Hu L, Zhuang S, Li Z. Feasibility and safety of magnetic-controlled capsule endoscopy system in examination of human stomach: A pilot study in healthy volunteers[J]. Journal of Interventional Gastroenterology, 2012,2(4):155-160.

[30] Goh P, Krishnan S M. Micromachines in endoscopy[J]. Best Practice & Research Clinical Gastroenterology, 1999,1(13):49-58.

[31] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation: CVPR.2014.81, 2014[C].

[32] He K, Zhang X, Ren S, Sun J. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015,37(9):1904-1916.

[33] Wang C, Zhong C. Adaptive Feature Pyramid Networks for Object Detection[J]. IEEE Access, 2021,9:107024-107032.

[34] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017,39(6):1137-1149.

[35] Girshick R, Research M. Fast R-CNN[J]. 2015 IEEE International Conference on Computer Vision (ICCV), 2015:1440-1448.

[36] J R, S D, R G. You onlylook once: unified，realtime object detection[J]. Proc.of theIEEE Conference on Computer Vision and Pattern Recognition, 2016:779-788.

[37] J R, A F. YOLO9000:better，faster，stronger[J]. 2017:6517-6525.

[38] J R, A F. YOLOv3: an incremental im-provement[J]. 2022.

[39] A B, CY W, H L. YOLOv4:optimalspeed and accuracy of object [J]. 2022.

[40] Rahim T, Hassan S A, Shin S Y. A deep convolutional neural network for the detection of polyps in colonoscopy images[J]. Biomedical Signal Processing and Control, 2021,68:102654.

[41] Zhang R, Zheng Y, Poon C C Y, Shen D, Lau J Y W. Polyp detection during colonoscopy using a regression-based convolutional neural network with a tracker[J]. Pattern Recognition, 2018,83:209-219.

[42] Shin Y, Qadir H A, Aabakken L, Bergsland J, Balasingham I. Automatic Colon Polyp Detection Using Region Based Deep CNN and Post Learning Approaches[J]. IEEE Access, 2018,6:40950-40962.

[43] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation: CVPR (2015), 2015[C].

[44] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical

Image Segmentation[J]. 2015,136.

[45] Tashk A, Herp J, Nadimi E. Fully Automatic Polyp Detection Based on a Novel U-Net Architecture and Morphological Post-Process[C]//: IEEE, 2019: 37-41

[46] Feng R, Lei B, Wang W, Chen T, Chen J, Chen D Z, Wu J. SSN: A Stair-Shape Network for Real-Time Polyp Segmentation in Colonoscopy Images[J]. 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), 2020:225-229.

[47] Sornapudi S, Meng F, Yi S. Region-Based Automated Localization of Colonoscopy and Wireless Capsule Endoscopy Polyps[J]. Applied Sciences, 2019,9(12):2404.

[48] Fan D, Ji G, Zhou T, Chen G, Fu H, Shen J, Shao L. PraNet: Parallel Reverse Attention Network for Polyp Segmentation: 23rd International Conference on Medical Image Computing and Computer-Assisted Intervention, Lima, Peru, 2020[C].

[49] Ge-Peng Ji Y C D F. Progressively normalized self-attention network for video polyp segmentation[J]. 2015.

[50] 刘莹, 吴宗泽. 机械设计教程[M]. 第三版. 北京: 机械工业出版社, 2019.

[51] 张淑敏. 机械设计课程设计指导书[M]. 北京: 中国农业大学出版社, 2014.

[52] Jha D, Smedsrud P H, Riegler M A, Halvorsen P A, de Lange T, Dag J A H A. Kvasir-SEG: A Segmented Polyp Dataset[J]. 2019.

[53] S L, L Q, H Q. Path aggregation network for instance segmentation Salt Lake City[J]. Proceedings of the 31st IEEE/CVF conference on Computer Vision and Pattern Recognition(CVPR)., 2018:8759-8768.

[54] Sharma, Rieger M R D. The best way to create and manage training data. software, LabelBox, Inc,[EB/OL]. [2019-05-21

]. https://www.labelbox.com/.

[55] Welcome to LabelMe, the open annotation tool.[EB/OL]. http://labelme.csail.mit.edu/Release3.0/.

[56] Mesejo P, Pizarro D, Abergel A, Rouquette O, Beorchia S, Poincloux L, Bartoli A. Computer-Aided Classification of Gastrointestinal Lesions in Regular Colonoscopy[J]. IEEE Trans Med Imaging, 2016,35(9):2051-2063.

# 致 谢

论文写到这里，也意味着我的本科生活接近尾声。四年大学，三年疫情，其中纵有许多遗憾和无奈，但是四年的回忆与收获对于我来说依旧弥足珍贵。

首先，感谢学校。作为一个来自农村的学生，学校为我提供了一个良好的学习环境，为我接触新事物、新知识提供了丰富多样的渠道和资源，让我有继续学习和不断进步的机会。

感谢在四年大学生活中遇到的每一位师长。在课堂上老师们为我们传业授道，课下为我们答疑解惑。感谢班主任张小栓老师四年来对我们学习生活负责认真的照顾指导。感谢导师老师胡标老师在我毕业设计过程中对我的指导和支持，让我能够顺利有序地推进毕设。

感谢在四年大学生活中遇到的每一位同学。在学习中，我们相互鼓励，一起参加比赛，一起熬夜，一起度过了一段充实的时光。在生活中，我们一起分享快乐与烦恼，相互帮助与照顾。感谢四年来，每一位朋友对我的包容与帮助，让我拥有一段珍贵的大学时光。

最后，感谢一直默默支持我的家人。感谢你们对我学习生活的无私投入和关心，感谢你们对我每一个选择的无条件支持，让我在千里之外也感到家的温暖。

四年前的夏天，我带着期待和未知来到农大，四年后的今天，我带着满满的收获和不舍即将离开，感谢四年的所遇所识，我将永远铭记这闪闪发光的四年。

陶胜

2023年5月

# 作者简介

基本介绍（姓名、性别、出生日期、籍贯等）：

陶胜，男，2000年10月26日出生，湖北省黄冈市黄梅县人。

教育经历：

2019年9月-2023年6月，在中国农业大学工学院机械电子工程系进行本科阶段学习。

本科期间主持/参与的科研项目：

2020-2022年参与国家级大学生创新创业项目—《基于 ADRC 迭代学习控制的太阳能无人机姿态控制》

本科期间获得的奖励和荣誉：

2020年，第三十七届全国部分地区大学物理竞赛二等奖

2020年，中国农业大学 2019-2021 学习优秀二等奖学金

2020年，中国农业大学 2019-2020 学年优秀学生奖学金

2021年，中国农业大学 2020-2021 学年学业三等奖学

2022年，第七届北京市大学生工程设计表达竞赛一等奖

2022年，中国农业大学 2021-2022 学年学业二等奖学

2022年，第15届“高教杯”全国大学生先进成图技术与产品信息建模创新大赛机械类 3D打印一等奖、Inspire优化创新设计三等奖、个人全能三等奖。