目录

[**第一章 需求分析** 1](#_Toc160813420)

[**1.1 项目背景与目的** 1](#_Toc160813421)

[**1.2 设计目标** 2](#_Toc160813422)

[**第二章 概要设计** 2](#_Toc160813423)

[**2.1 模块和框架设计** 2](#_Toc160813424)

[**2.2 整体框架结构图和整体数据流图** 3](#_Toc160813425)

[**第三章 详细设计** 3](#_Toc160813426)

[**3.1 用户交互设计** 3](#_Toc160813427)

[**3.1.1使用流程** 4](#_Toc160813428)

[**3.1.2 streamlit** 6](#_Toc160813429)

[**3.2 综合智能问答模块（数据流）** 6](#_Toc160813430)

[**3.2.1 LLM问答** 7](#_Toc160813431)

[**3.2.2 基于知识库的问答** 7](#_Toc160813432)

[**3.2.3 基于浏览器问答** 7](#_Toc160813433)

[**3.2.4 实时知识库创建** 7](#_Toc160813434)

[**3.3 接Langchain和大语言模型和词向量切割模型** 7](#_Toc160813435)

[**3.4 鸟类图片科普模块（数据流）** 7](#_Toc160813436)

[**Yolov8** 7](#_Toc160813437)

[**第四章 测试报告** 11](#_Toc160813438)

[**第五章 安装及使用** 11](#_Toc160813439)

[**第六章 项目总结** 11](#_Toc160813440)

[**参考文献** 11](#_Toc160813441)

**第一章 需求分析**

**1.1 项目背景与目的**

随着人类活动的不断扩张和生态环境的恶化，许多鸟类面临着濒危和灭绝的威胁。这些濒危鸟类在生态系统中扮演着重要角色，对维护生态平衡和生物多样性具有至关重要的作用。然而，传统的保护方法往往面临着监测和识别方面的挑战，导致保护工作的效果不佳。

本项目旨在利用图像识别技术与对话语言模型，开发一个基于YOLOv8模型的濒危鸟类图像识别系统与关于濒危鸟类保护的智能对话平台，提高实施濒危鸟类保护工作与宣传保护濒危鸟类相关知识的效率和准确性。通过应用图像识别算法，实现对濒危鸟类的监测、识别，为保护者提供及时的数据和情报，从而更好地制定保护策略和采取行动；当提出关于濒危鸟类的问题时，语言模型可以提供详细的回答，以普及保护濒危鸟类的重要性。

**1.2 设计目标**

|  |  |
| --- | --- |
| **设计目标** | **描述** |
| 用户友好性 | 确保系统的用户界面友好、直观，使用户能够轻松地访问和使用系统的各项功能，无论是问答科普还是图片科普。 |
| 功能完善 | 系统应包括问答科普和图片科普两大模块，覆盖鸟类科普和保护的各个方面。问答科普应提供多种问答模式，并允许用户管理知识库；图片科普应支持上传图片并进行自动识别和解释。 |
| 性能优化 | 保证系统的性能和稳定性，特别是针对后台处理鸟类图片识别和基于语言模型的问答生成，需要考虑优化算法和资源利用。 |
| 安全性 | 确保用户上传的图片和信息的安全性，采取必要的安全措施，如数据加密、访问控制等，防止数据泄露和恶意攻击。 |
| 可扩展性 | 系统应设计为可扩展的，方便日后添加新的功能或模块，以应对不断变化的需求和技术发展。 |
| 易维护性 | 代码应具有良好的结构和注释，方便团队成员理解和维护，同时应有完善的文档和测试，确保系统的稳定性和可靠性。 |
| 用户参与性 | 鼓励用户参与系统的建设和完善，例如允许用户上传自己的鸟类观察数据或科普资料，为系统的发展贡献力量。 |

**第二章 概要设计**

**2.1 模块和框架设计**

B/S架构

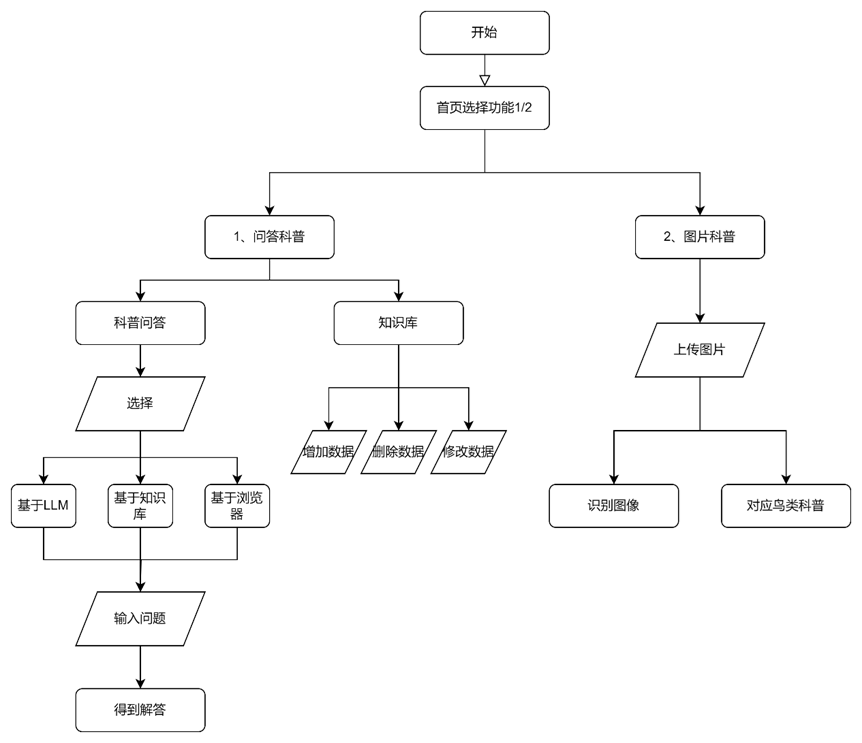
|  |  |  |  |
| --- | --- | --- | --- |
| Server端 | 模块名称 | | 功能描述 |
| 鸟类图像科普 | 图像识别 | 自动识别并展示濒危鸟类的名称、特征等信息 |
| 智能科普 | 基于最新人工智能技术提供动态的濒危鸟类相关知识、保护措施等科普内容 |
| 鸟类综合智能问答 | 基于LLM问答 | 利用LLM处理自然语言查询，提供灵活的问答能力。 |
| 基于知识库问答 | 依托专门构建的鸟类知识库来快速检索信息，对特定、结构化的查询提供精确答案。 |
| 基于浏览器问答 | 维护和更新鸟类知识库，包括添加新的知识点、更新现有信息和优化知识结构。 |
| 知识库管理 | 结合LLM和知识库技术，提供准确、及时的鸟类信息问答服务，支持多语言互动。 |
| Browser端 | 首页 | | 展示项目介绍、团队介绍和用户互动入口。 |
| 关键技术介绍页面 | | 详细介绍项目所采用的关键技术，如图像识别、LLM等。 |
| 开发团队、流程介绍页面 | | 介绍开发团队背景、项目开发流程和未来规划。 |

**2.2 整体框架结构图和整体数据流图**

**第三章 详细设计**

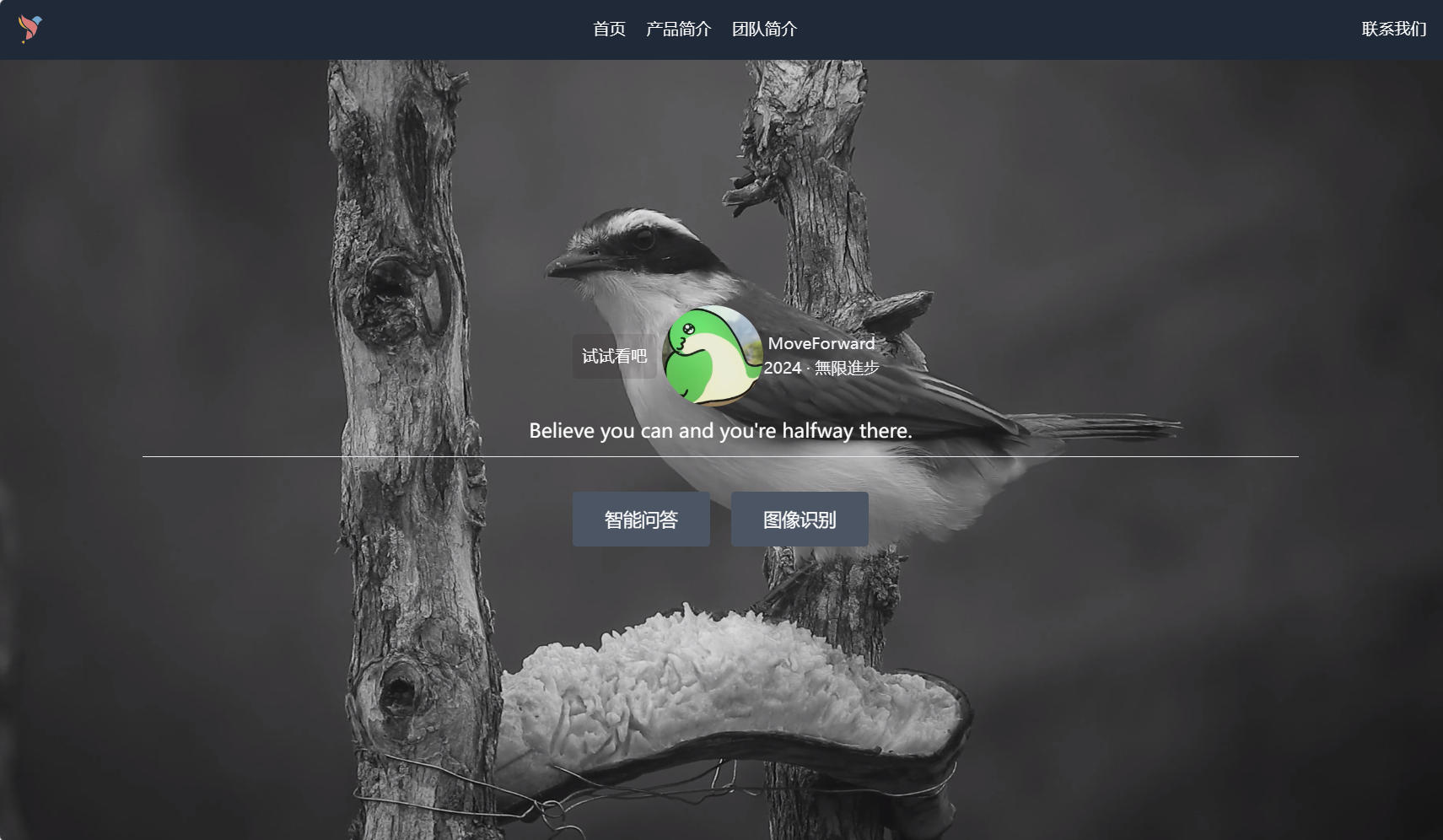
**3.1 用户交互设计**

这个流程不仅简化了用户操作，还确保了信息的准确传递，提高了用户的参与度和满意度。通过集成先进的技术，我们的系统为鸟类保护和科普教育提供了一个强有力的工具，使公众能够更容易地获取有关鸟类的知识，并激发他们对自然保护的兴趣和意识。



**3.1.1使用流程**

在我们的系统中，鸟类问答科普模块提供了一个易于使用且信息丰富的平台，旨在增强公众对鸟类及其保护的认识。用户首先通过浏览器访问系统首页，这是一个利用传统HTML、CSS、JavaScript以及Tailwind CSS和FontAwesome精心设计的响应式界面。



在选择问答科普模块后，用户将进入一个由streamlit构建的简洁前端界面，在此他们可以键入具体的鸟类相关问题。系统提供了三种问答模式以满足不同的用户需求：基于LLM模型的问答、基于知识库的问答以及基于浏览器的问答。

针对用户的问题，系统采用LangChain框架集成的ChatGLM大预言模型和m3e-base词向量切割模型进行处理。在知识库模式下，问题会被转化为向量并与本地向量数据库中的文档进行匹配，以找到最相关的上下文信息，随后利用LLM模型生成准确的答案。对于浏览器模式，系统通过网络搜索来寻找和展示最相关的信息。不论哪种模式，系统均能在用户界面上直接展示详细的答案或提供链接至答案源。

除了提供深入的问答科普模块外，我们还特别设计了一个鸟类图片科普模块，旨在通过视觉识别技术加深用户对鸟类多样性和特征的了解。用户可通过上传鸟类照片来启动这一模块。系统后台使用经过训练的yolov8模型识别图片中的鸟类种类和位置，并在图片上直接标注。此过程不仅展示了先进的图像识别技术，而且还能即时向用户提供关于所识别鸟类的详细科普信息。

一旦识别出鸟类种类，系统将结合预设的提示词和经过微调的ChatGLM模型，自动生成该鸟类的介绍。这包括鸟类的习性、栖息地、保护状况等有价值的信息，旨在教育用户并提高他们对生物多样性的保护意识。此外，为了优化模型的性能并适应不同的运行环境，我们对yolov8模型进行了减枝操作，包括约束训练、剪枝和回调训练（finetune），以确保即使在资源有限的设备上也能保持高效的识别准确率。

**3.1.2 streamlit**

Streamlit是用于快速构建数据应用和可视化界面的Python库。它集成了LangChain框架与多个大型语言模型，用户可轻松进行自然语言处理鸟类知识查询，前端友好。上传鸟类图片，后台YOLOv8模型迅速识别鸟类种类，结合ChatGLM模型生成科普内容，确保从图像识别到文本解释的全过程高效、互动性强。

|  |  |  |  |
| --- | --- | --- | --- |
| 特征 / 技术 | Streamlit | Dash | Bokeh |
| **易用性** | 非常易用，简单直观的API，适合快速开发 | 较易用，需要一定学习成本，但提供了丰富的文档和示例 | 较易用，提供了交互式绘图和可视化的功能，但需要更多配置 |
| **实时更新** | 支持实时更新，自动响应代码变化 | 需要手动刷新页面或重新加载应用 | 需要手动刷新页面或重新加载应用 |
| **Python生态系统兼容性** | 完全兼容，与Python库和工具集成无缝 | 与Python生态系统兼容，但需要额外的配置 | 与Python生态系统兼容，但需要额外的配置 |
| **部署便捷性** | 部署简单，可轻松部署到各种平台 | 部署相对简单，但需要一些配置 | 部署相对简单，但需要一些配置 |
| **交互式组件** | 提供丰富的交互式组件，如滑动条、选择框等 | 提供了一些交互式组件，但不如Streamlit丰富 | 提供了一些交互式组件，但不如Streamlit丰富 |
| **图形化功能** | 提供了简单易用的图形化功能 | 提供了一些图形化功能，但不如Streamlit灵活 | 提供了一些图形化功能，但不如Streamlit灵活 |

Streamlit易用，实时更新，与Python生态无缝，部署简单，提供丰富交互组件和图形化功能。Dash和Bokeh相对较易用，部署和交互组件方面不如Streamlit。

**3.2 综合智能问答模块（数据流）**

问答科普模块集成了LangChain框架和streamlit前端，结合ChatGLM模型和m3e-base词向量模型，支持多种问答模式。用户可以通过LLM模型、知识库或浏览器问答方式获取关于鸟类的科普知识，系统还支持用户自定义知识库，提供了灵活的问答体验。

**3.2.1 LLM问答**

LLM（Large Language Model）模式是一种基于大型语言模型的问答方式，通过训练的模型能够理解自然语言问题，并生成相应的回答。在鸟类科普系统中，LLM模式利用了ChatGLM模型，结合LangChain框架和streamlit前端，提供了多种问答方式，包括基于LLM的问答。用户可以通过输入问题，系统会利用LLM模型从已有的语料库中找到最相关的信息，并生成专业的科普回答。

**3.2.2 基于知识库的问答**

基于知识库的问答是指系统利用预先构建的知识库来回答用户的问题。在鸟类科普系统中，用户可以通过这种方式获取关于鸟类的科普知识。系统会提前构建一个包含丰富鸟类信息的知识库，用户可以输入问题，系统会在知识库中匹配相关信息，并给出相应的答案。这种方式能够快速、准确地响应用户的问题，特别适用于对于固定知识范围的查询。

**3.2.3 基于浏览器问答**

基于浏览器的问答是指系统利用浏览器的本地资源来回答用户的问题，而不需要向服务器发送请求。在鸟类科普系统中，这种问答方式可以提高用户体验和响应速度。系统会将部分知识库或者语言模型加载到浏览器中，当用户输入问题时，系统会在本地进行处理并生成答案，而不需要等待服务器的响应。这种方式通常适用于对计算资源要求不高的场景，能够在一定程度上提高系统的性能和响应速度。

**3.2.4 实时知识库创建**

实时知识库创建是指用户可以即时地创建和更新知识库，以适应不断变化的需求和知识。在鸟类科普系统中，用户可以通过系统提供的管理界面实时添加、编辑和删除鸟类相关的信息，包括物种介绍、习性特点、栖息地等。这样的功能使得用户可以根据最新的科学研究或者观察到的现象及时更新知识库，保持系统的信息更新和准确性，提高系统的实用性和可信度。窗体顶端

**3.3 Langchain和大语言模型和词向量切割模型**

**3.4 鸟类图片科普模块（数据流）**

**Yolov8**

**3.1 模型选择**

**3.1.1 YOLOv8**

图像识别技术在鸟类识别领域的应用面临着多种挑战：鸟类种类繁多，不同种类的鸟在外观上可能仅有细微差别，这要求模型能够捕捉到并识别这些微小的特征差异；鸟类图片通常包含复杂的背景，如树木、天空或水面，这些背景可能会干扰模型对鸟类的识别；同一种鸟在不同的姿态和视角下拍摄的图片可能差异较大，给准确识别带来了额外的挑战；不同的光照条件会影响图片的质量和鸟类的颜色表现，增加识别的难度；在一些图片中，鸟类可能只占很小的一部分，这要求模型能够准确检测到小尺寸目标。

在当前的人工智能领域，YOLO模型因其在处理复杂图像识别任务中表现出的显著优势而受到广泛关注。特别是在鸟类图像识别这一具有挑战性的任务中，YOLO通过其高精度、快速处理能力和优秀的泛化性能，显著提升了识别效果。这些优势使YOLO成为研究和应用中的优选模型，尤其是在需要从复杂背景中准确识别众多相似鸟类种类的场景里。如表1所示。

**表1 YOLO优势**

|  |  |
| --- | --- |
| 难点 | YOLO的优势 |
| 多样性和相似性 | YOLO模型凭借其深层的神经网络可以区分细微的特征，即使是外表相似的鸟类也能被准确识别。 |
| 复杂背景 | 采用先进算法优化背景抑制，减少背景噪声干扰，提高对鸟类目标的识别准确性。 |
| 姿态和视角变化 | 模型能够学习鸟类在不同姿态和视角下的表现，提高识别的灵活性和准确性。 |
| 光照条件变化 | YOLO模型能够适应不同光照条件下的图像识别，保持稳定的性能。 |
| 小目标检测 | YOLO模型在提升了速度的同时，也针对小目标检测进行了优化，有助于识别图像中的小型鸟类。 |

与传统的卷积神经网络（CNN）模型相比，YOLO在鸟类图像识别上的表现同样值得关注。传统CNN模型虽然在图像分类等任务中取得了显著成果，但在进行复杂的目标检测和分类任务时，往往需要更多步骤和复杂的处理流程。相比之下，YOLO作为一种端到端的目标检测模型，不仅提高了处理速度，而且通过一次性推理即可完成目标的检测和分类，极大提高了效率和准确性。如表2所示。

然而，随着计算机视觉领域的快速发展，YOLO系列模型也在不断进化。每一次迭代，都带来了新的改进和特性，以适应日益复杂的应用需求。YOLOv8作为该系列中的最新成员，继承了YOLO模型高速和高精度的核心特点，同时引入了创新的技术，以提升模型性能和效率。

|  |  |  |  |
| --- | --- | --- | --- |
| 特征 | YOLOv8 | YOLOv5 | CNN模型 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**3.1.2 ChatGLM**

ChatGLM-6B是一个开放源代码的双语对话语言模型，基于General Language Model(GLM)架构开发，拥有62亿个参数。该模型采用了与ChatGPT类似的技术，并专门针对中文问答与对话场景进行了优化。通过接受大约1T中英文标识符的训练，并结合了监督微调、自我反馈以及基于人类反馈的强化学习技术，ChatGLM-6B展现出了生成符合人类偏好回答的强大能力。

**表4 ChatGLM优势**

|  |  |
| --- | --- |
| 优势 | 描述 |
| 对话优化 | 特别为中文问答和对话场景进行了优化，确保ChatGLM可以理解并生成符合语境和文化相关性的回复 |
| 大规模训练数据 | 经过大约1T标识符的中英文训练，ChatGLM展现了对语言细微差别的广泛理解，这对于生成连贯多样的对话至关重要 |
| 技术先进 | 集成了监督微调、人类反馈强化学习等技术，使ChatGLM能够提供与人类偏好更为一致的回应 |
| 可扩展性和集成性 | 凭借先进的语言模型能力，ChatGLM可以扩展到各种应用程序，从客户服务自动化到互动教育平台 |
| 类人交互 | 采用了基于人类反馈的强化学习，意味着该模型经过训练能生成更具吸引力和类似人类的回复，提升用户体验 |

**3.1.3 m3e-base**

m3e-base主要通过其先进的文本向量化技术支持问答科普和知识库管理功能。m3e-base能够将自然语言文档和用户的查询问题转换成数学向量，这一过程对于理解和处理大量自然语言信息至关重要。特别是在进行问答时，m3e-base通过文本和问句的向量化，以及后续的向量匹配，确保了系统能够快速而准确地找到与用户查询最相关的信息，为大型语言模型（LLM）生成回答提供了准确的上下文。m3e-base的加入，极大提升了问答系统的效率和准确性，使得用户能够获得更为精准和丰富的鸟类保护知识。

|  |  |
| --- | --- |
| m3e-base优势 | 说明 |
| 高效率 | 快速处理文本向量化，确保实时响应用户查询 |
| 精确匹配 | 精准的向量匹配技术提高问答系统的准确性 |
| 适应性强 | 支持多种本地向量数据库，灵活适应不同的存储和检索需求 |
| 增强问答质量 | 为LLM提供准确上下文，生成更贴近用户查询的回答 |

**3.5 数据集制作及模型训练过程**

**3.5.1 鸟类图片数据集制作**

**3.5.2 LLM微调数据集制作**

**3.7 创新点**

**3.7.1 版本控制**

**3.7.2 Langchain**

**3.7.3**

**第四章 测试报告**

**第五章 安装及使用**

**第六章 项目总结**

**参考文献**