

视频转 Word 智能图文生成工具使用说明书

1. 项目简介

本工具是一个基于 Python 开发的自动化脚本，旨在将视频内容(本地文件或 Bilibili/YouTube 网络链接)高效转换为图文并茂的 Word 文档。它集成了 **OpenAI Whisper** 进行语音转文字、**DeepSeek LLM** 进行智能总结与文本润色，以及 **OpenCV** 进行关键帧截图。

核心功能：

- 多源支持: 支持本地视频文件及网络视频链接(B站/YouTube)自动下载。
- 智能听写: 利用 Whisper 模型实现高精度语音转文字。
- AI 总结与润色: 自动生成会议/视频纪要，并提供“去除口癖”的润色版全文。
- 图文对照: 生成带有时间戳、字幕和对应视频画面的表格，方便回顾。
- 格式优化: 内置字体强制修正功能(强制使用微软雅黑)，解决生成的 Word 文档中文字体显示异常的问题。

2. 配置指南

在运行代码之前，需要根据您的实际情况修改脚本顶部的 === 配置区域 ===。

变量名	说明	示例/建议
COOKIE_PATH	B站 Cookie 文件路径	用于下载高清或会员限制视频。建议使用 Netscape 格式导出的 txt 文件。若无需登录可忽略。
DEFAULT_OUTPUT_BASE	默认输出目录	D:/VideoWork/Output 这是所有生成文件的根目录。
LLM_API_KEY	大模型 API 密钥	必须替换为您自己的 DeepSeek 或兼容 OpenAI 格式的 Key。
LLM_BASE_URL	模型接口地址	默认为 DeepSeek 地址： https://api.deepseek.com
MAX_SUBTITLE_LENGTH	字幕分段长度	默认为 20。决定了 Word 表

		格中每一行文字的长度和截图的密度。
--	--	-------------------

4. 使用步骤

第一步：启动程序

在终端或 IDE 中运行脚本：

```
python your_script_name.py
```

第二步：输入视频源

程序启动后会提示：

请输入视频链接 或 本地文件路径：

```
>>>
```

- 网络视频：直接粘贴 Bilibili 或 YouTube 的 URL。
- 本地视频：直接粘贴电脑上的视频文件路径（如 D:\Videos\meeting.mp4）。

第三步：确认存储路径

程序会显示默认存储路径，按 回车(Enter) 键确认，或输入新的路径进行修改。

第四步：等待处理

程序将自动执行以下流程（视视频长度，耗时几分钟至数十分钟不等）：

1. 获取/下载视频：如果是链接，会自动下载最高画质。
2. 加载模型：加载 Whisper 语音模型（首次运行会自动下载模型权重）。
3. 语音识别：将音频转换为带有时间戳的文字。
4. AI 智能处理：调用 DeepSeek 进行全文总结和文本清洗。
5. 生成文档：截取视频画面，合成 Word 文档。

第五步：查看结果

完成后，程序会显示：

 完成！

 文档：[路径]/video_name.docx

 图片：[路径]/images

您可以直接打开生成的 .docx 文件查看成果。

5. 输出文档结构说明

生成的 Word 文档包含以下三个主要部分：

AI 智能总结

- 包含一句话摘要。
- 核心观点列表 (Key Points)。
- 详细的内容脉络梳理。

第一部分：字幕全文对比

- 1.1 原始全文：忠实还原音频的逐字稿。
- 1.2 AI 优化全文：去除了“呃、那个”等口语废话，修正了标点和段落，适合阅读。

第二部分 & 第三部分：图文对照表

- 采用双栏表格形式。
- 左栏：精确的时间戳(蓝色加粗)和对应的字幕文本。
- 右栏：对应时间点的视频截图。
- 注：第二部分使用 AI 润色后的文本，第三部分使用原始文本。