**YouTube 쇼츠 자동화 프로젝트: 파이썬 기반 '쇼츠 메이커' 개발 개요 및 분석**

**Executive Summary**

본 문서는 유튜브 채널 '잠코딩'에서 공개한 파이썬 기반 '쇼츠 메이커' 프로젝트의 개발 동기, 기술 구성, 초기 버전의 한계점, 그리고 향후 개선 계획을 종합적으로 분석한다. 프로젝트의 핵심 목표는 Make, n8n과 같은 유료 자동화 플랫폼을 사용하지 않고, 파이썬 코드와 여러 API를 연동하여 키워드 입력만으로 대본 작성부터 영상 제작, 자막 동기화까지 전 과정을 자동화하는 것이다.

버전 1.0은 300줄 미만의 코드로 키워드 기반 대본, 음성, 이미지 생성을 통해 쇼츠 영상을 성공적으로 제작했으나, 결과물의 품질은 저조했다. 주요 문제점으로는 ▲부실한 GPT 프롬프트로 인한 낮은 대본 퀄리티 ▲AI 이미지 생성 비용 부담으로 인한 시각 자료 부족 ▲기본 폰트와 템플릿 부재로 인한 조악한 디자인이 지적되었다.

향후 버전 2.0에서는 이러한 문제 해결을 위해 ▲GPT 프롬프트 정교화 ▲비용 절감을 위한 Pixabay 무료 이미지/영상 API 연동 ▲디자인 개선을 위한 폰트 변경 및 전문 템플릿 활용 ▲YouTube API를 통한 업로드 자동화 등을 계획하고 있다. 이는 단순한 기능 구현을 넘어, 비용 효율적이면서도 높은 품질의 콘텐츠를 지속적으로 생산할 수 있는 실용적인 자동화 솔루션으로의 발전을 목표로 하고 있음을 시사한다.

--------------------------------------------------------------------------------

**1. 프로젝트 개요 및 개발 동기**

최근 유튜브에서 '쇼츠 자동화'가 높은 조회수를 기록하며 큰 관심을 끌고 있다. 이러한 트렌드에 따라, 개발자는 직접 쇼츠 영상 제작 자동화 프로젝트에 착수했다.

* **개발 동기**:
  + **비용 문제**: 대부분의 쇼츠 자동화 솔루션은 Make, n8n과 같은 자동화 플랫폼을 사용한다. 이러한 플랫폼은 월 구독료가 발생하며, 이는 음성 및 이미지 생성 API 비용에 더해 추가적인 재정적 부담이 된다.
    - n8n: 무료 플랜이 없고 월 20유로부터 시작.
    - Make: 무료 플랜이 존재하지만, 작업 간 15분의 대기 시간, 월 사용 횟수 제한 등 기능적 제약이 커 결국 유료 결제가 필요.
  + **비용 효율성 추구**: 초기 단계에서는 시간 효율성보다 비용을 절감하는 것이 더 중요하다고 판단하여, 플랫폼 구독료 없이 직접 개발하는 방식을 선택했다.
* **프로젝트 목표**:
  + 키워드와 생성할 이미지 개수만 입력하면 **원클릭**으로 모든 프로세스가 자동 처리되는 프로그램을 개발한다.
  + 자동화 범위는 대본 작성, AI 음성 생성, 이미지 생성, 영상 합성, 그리고 자막 싱크 맞추기를 포함한다.

**2. '쇼츠 메이커' 기술 스택 및 핵심 기능**

'쇼츠 메이커' 버전 1.0은 파이썬을 기반으로 다양한 외부 도구와 API를 연동하여 구축되었다. 전체 코드는 단일 파일에 300줄 미만으로 구성되어 있다.

|  |  |  |
| --- | --- | --- |
| 구분 | 도구 및 라이브러리 | 역할 |
| **개발 환경** | Python 3.1 이상 | 프로그램의 주 개발 언어 |
| **핵심 도구** | FFMPEG | 영상과 음성의 편집 및 합성을 위한 필수 도구 |
| **API** | OpenAI API (GPT, DALL-E) | GPT 모델로 대본을 생성하고 DALL-E 모델로 관련 이미지를 생성 |
|  | Typecast API | 생성된 텍스트를 자연스러운 AI 음성으로 변환 |
| **라이브러리** | Whisper | 음성 파일을 분석하여 텍스트와 자막 싱크를 자동으로 맞춤 |

**핵심 자동화 프로세스**

1. **대본 생성**: 사용자가 입력한 키워드를 기반으로 GPT API에 대본 생성을 요청하여 자연스러운 텍스트를 얻는다.
2. **음성 생성**: 생성된 대본을 Typecast API로 전송하여 AI 음성을 생성하고 MP3 파일로 다운로드한다.
   * **Typecast 선택 이유**: 무료 TTS(Text-to-Speech) 옵션보다 훨씬 자연스러운 음성 품질을 제공하며, 실제 유튜브 영상에서 많이 사용된다.
   * **주요 특징**: 다양한 캐릭터 음성(사투리 포함) 선택 가능. 프로젝트에서는 '사투리 쓰는 부산 여자' 캐릭터를 선택했다.
   * **무료 체험**: 14일간 250쿼리(API 1회 호출 = 1쿼리)를 무료로 제공하여, 최대 250개의 쇼츠 음성 생성이 가능하다.
3. **이미지 생성**: 대본 내용에 맞춰 DALL-E API를 통해 원하는 수량의 이미지를 생성한다.
4. **자막 싱크**: Whisper 라이브러리를 사용해 생성된 음성(MP3)과 대본 텍스트의 싱크를 정밀하게 맞춘다.
5. **영상 합성**: FFMPEG을 이용하여 준비된 이미지, 자막, 음성을 최종 쇼츠 영상 파일로 합성한다.

**3. 버전 1.0 테스트 결과 및 평가**

프로젝트의 개념 증명을 위해 '고구마'라는 키워드와 3개의 이미지 생성을 조건으로 테스트를 진행했다.

* **실행 결과**:
  + '고구마'를 주제로 한 음성 파일과 이미지 3장이 성공적으로 생성되었다. (생성된 이미지 중 하나는 주제와 동떨어진 결과물을 보였다.)
  + 음성, 이미지, 자막이 합쳐진 최종 쇼츠 영상이 완성되었다.
* **자체 평가**:
  + 기능적으로는 성공했으나, 결과물의 품질이 "부끄러울 정도로 퀄리티가 떨어졌다"고 평가했다.
  + 이는 버전 1.0이 기능 구현에 초점을 맞춘 프로토타입이며, 상용 수준의 결과물을 위해서는 상당한 개선이 필요함을 시사한다.

**4. 버전 2.0 개선 계획**

개발자는 버전 1.0의 한계를 명확히 인지하고, 품질 향상을 위한 네 가지 구체적인 개선 계획을 제시했다.

**4.1. 대본 품질 향상**

* **문제점**: GPT에 대본 생성을 요청하는 프롬프트가 너무 단순하고 허술하여 대본의 질이 낮았다.
* **해결 방안**: 지속적인 수정을 통해 프롬프트를 더욱 정교하고 구체적으로 만들어 대본의 완성도를 높일 계획이다.

**4.2. 비용 절감 및 시각 자료 다각화**

* **문제점**: DALL-E API의 이미지 생성 비용("몇 번 테스트하면 금방 1달러씩 사라지는 게 너무 아까워서")에 대한 부담 때문에 이미지를 3개만 생성하는 등 시각 자료를 충분히 활용하지 못했다.
* **해결 방안**: Pixabay에서 제공하는 무료 API를 도입할 예정이다. 이를 통해 고품질의 무료 이미지와 동영상을 활용하여 AI 이미지 사용을 줄이고, 결과적으로 비용을 절감하며 더 역동적인 영상 제작을 시도할 것이다.

**4.3. 디자인 및 템플릿 개선**

* **문제점**: 자막에 기본 폰트를 사용하고 별도의 영상 템플릿이 없어 전체적인 디자인이 부실했다.
* **해결 방안**:
  + **폰트**: 가독성 높고 심미적으로 우수한 폰트를 찾아 적용한다.
  + **템플릿**: 디자인 감각 부족을 보완하기 위해 '비디오 스튜플릿 리스트' 등에서 제공하는 정보성, 인용구 등 잘 만들어진 템플릿을 활용할 계획이다. 템플릿은 "일종의 포장지"로서, 이것만 적용해도 영상의 전반적인 품질이 크게 향상될 것으로 기대된다.

**4.4. 완전 자동화 구현**

* **문제점**: 현재 자동화는 영상 파일을 생성하는 단계에서 멈춘다.
* **해결 방안**: 진정한 의미의 자동화를 위해, YouTube API를 연동하여 생성된 영상을 채널에 자동으로 업로드하는 기능까지 구현할 계획이다.

--------------------------------------------------------------------------------

**5. 결론**

'잠코딩'의 '쇼츠 메이커' 프로젝트는 유료 플랫폼에 의존하지 않고 파이썬과 공개 API만으로 유튜브 쇼츠 제작을 자동화할 수 있는 가능성을 명확히 보여주었다. 버전 1.0은 기능적 구현에 성공한 프로토타입으로서, 낮은 품질이라는 한계점을 드러냈지만, 이는 동시에 구체적인 개선 방향을 설정하는 중요한 기반이 되었다.

향후 프롬프트 엔지니어링, 무료 스톡 리소스 활용, 전문 템플릿 도입, 업로드 자동화 등 버전 2.0 계획이 성공적으로 이행된다면, 본 프로젝트는 비용 효율성과 콘텐츠 품질을 모두 만족시키는 강력한 개인 맞춤형 자동화 도구로 발전할 잠재력을 충분히 가지고 있다.