

# AI 자막 생성 및 번역을 활용한 웹 기반 강의 스트리밍 플랫폼 설계

송우승<sup>1</sup>, 정세엽<sup>2</sup>, 고은샘<sup>3</sup>, 천지윤<sup>4</sup>, 이미진<sup>5</sup>  
<sup>12345</sup> 경북대학교 컴퓨터공학부 심화컴퓨터전공 학부생

dntmd2061@gmail.com, einstein3028@knu.ac.kr, dmsamy@knu.ac.kr,  
 a47526175@gmail.com, wls4482@gmail.com

## Development of a Web-Based Lecture Streaming Platform Utilizing AI Subtitle Generation and Translation

Woo-Seung Song<sup>1</sup>, Sae-Yeop Jung<sup>2</sup>, Eun-Saem Ko<sup>3</sup>, Ji-Yoon Chun<sup>4</sup>, Me-Jin Lee<sup>5</sup>  
<sup>12345</sup>Dept. of Computer Science Engineering, Kyung-Pook National University

### 요 약

본 연구는 AI 를 활용하여 온라인 강의 영상의 음성을 자동으로 텍스트로 변환하고, 이를 다국어로 번역하여 자막을 생성 및 관리하는 자동화 시스템을 설계하였다. Faster Whisper 기반의 음성 인식 모델을 통해 강의 음성을 텍스트로 변환한 후, MarianNMT 기반의 텍스트 생성 모델을 활용하여 주요 언어(한국어, 영어, 일본어, 중국어 등)로 번역 자막을 생성하며, 생성된 자막은 서버에 저장된다. 사용자는 강의 시청 시 원하는 언어의 자막을 실시간 지연 없이 응답 받을 수 있으며, 관리자는 번역 품질이 낮거나 전문 용어가 부정확한 경우 직접 검수 및 수정할 수 있도록 하였다. 본 시스템은 실시간 번역 방식에 비해 서버 부하를 줄이고, 사용자 경험을 개선하며, 자막 품질을 일정 수준 이상으로 유지할 수 있다는 점에서 효율적인 대안이 될 수 있다.

### 1. 서론

최근 온라인 강의 플랫폼의 성장으로 다양한 국가의 사용자가 콘텐츠에 접근하게 되면서, 다국어 자막 제공은 필수 요소로 자리잡고 있다. 자막의 유무와 품질은 학습 효율과 몰입도에 큰 영향을 미치며, 특히 글로벌 환경에서는 자막 시스템이 콘텐츠 확장성과 사용자 만족도에 중요한 역할을 한다 [1]. 하지만 기존 자막 제작 방식은 수작업 또는 외부 번역 서비스에 의존해 시간과 비용이 많이 들고, 강의 수가 늘어날수록 품질 유지와 관리에 어려움이 있다 [2]. 이에 본 연구에서는 인공지능을 활용한 자동 자막 생성 및 번역 시스템을 제안한다.

본 연구에서는 이러한 문제를 해결하기 위해 인공지능(AI)을 활용한 자동 자막 생성 및 번역 시스템을 제안한다.

### 2. 설계 및 구현

.NET 기반으로 구축되며, 사용자 인증 및 강의 정보, 자막 데이터를 관리한다. AI 엔진을 사용하여 강의 음성을 텍스트로 변환한다. 그리고 사후 처리 후 다국어로 번역한다. SRT 파일은 로컬에 저장된다. 사용자는 회원가입 및 로그인 후 강의를 수강 신청할 수 있으며, 강의 시청 시 서버는 이미 저장된 SRT 파

일과 함께 강의를 제공한다. 이를 통해 사용자는 실시간 번역 요청 없이 다국어 자막이 포함된 강의를 원활하게 시청할 수 있다.

또한, 번역 품질이 낮은 경우 관리자에게 수정 요청할 수 있으며, 관리자는 검토 후 SRT 파일을 업데이트할 수 있다. 본 시스템은 관리자가 동영상에 대해 자막 요청 시 AI 엔진이 다국어 자막 생성을 자동 수행하도록 설계되어, 실시간 API 호출 방식보다 성능 최적화가 가능하며, 사용자 경험을 극대화할 수 있다.

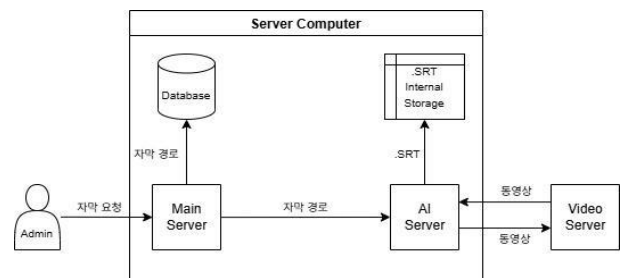


그림 1. 시스템 아키텍처.

그림 1은 관리자와 서버, AI 서버, 비디오 서버, DB 간의 상호작용을 시스템 아키텍처로 표현한 것으로, 한 번의 요청을 통해 다국어 자막 생성 제공까지의 전체 흐름을 직관적으로 나타낸다.

## 1) 기능 활용

화면	기능	설명 및 내용
시작 화면	사이트 정보 제공	사이트 소개와 사용자 기능을 보여준다.
로그인 화면	회원가입, 로그인	사용자 정보를 DB에 저장하고 관리한다.
내 강의실	강의 확인 및 시청	구해진 강의를 시청하고 학습 상태 확인한다.
강사	강사 확인	분야 별 강사 프로필을 확인한다. 강사를 선택하면 강사의 강의를 볼 수 있다.
커뮤니티	공지사항	관리자만 작성할 수 있어야 한다.
	질문과 답변	사용 권한이 있는 사용자가 질문을 작성하고 권한을 가진 관리자가 답변을 작성한다.
	자료실	관리자가 작성한다. 관리자는 파일을 첨부하고 사용자는 첨부한 파일을 다운로드할 수 있다.

표 1. 웹 미디어 강의 플랫폼 기능

사용자는 시작 화면을 통해 사이트 정보를 파악하고, 로그인 후에는 강의 수강, 자료 열람 등 다양한 학습 활동을 수행할 수 있다. 특히 ‘내 강의실’ 메뉴를 통해 자막이 포함된 강의 시청, 진도 확인 등의 핵심 학습 기능에 접근할 수 있도록 구성 되어있다.

## 1) AI 자막 추출

동영상을 업로드한 뒤 해당 영상의 자막 요청 시스템은 해당 동영상의 자막 파일을 저장할 절대 경로를 인식하고, 이 정보와 함께 AI 서버에 처리 요청을 전송한다.

서버 측에서는 먼저 Faster Whisper 모델을 활용하여 동영상의 오디오로부터 텍스트를 추출한다. 이 과정에서 시작 및 종료 시간이 포함된 원본 언어 SRT 자막 파일이 생성되며, 사후 처리 과정을 거친 뒤에 입력 받은 절대 경로에 저장된다.

이후 시스템은 다국어 번역을 위해 MarianNMT 모델을 활용한다. 이 모델은 원본 언어에서 영어, 중국어, 일본어, 베트남어 총 4 개의 목표 언어로 추출된 텍스트를 번역한다. 각 언어별 번역 결과는 시간 정보가 유지된 개별 SRT 파일로 변환된다.

생성된 SRT 파일은 정확성과 전문성을 높이기 위해 도메인 사전을 활용한 사후 처리 과정을 거친다. 이 과정에서 각 언어별 특화된 전문 용어 사전을 참조하여 번역된 텍스트 내의 전문 용어, 고유명사, 약어 등이 적절하게 교정된다. 사후 처리가 완료된 SRT 파일은 언어별로 지정된 절대 경로에 최종 저장된다.

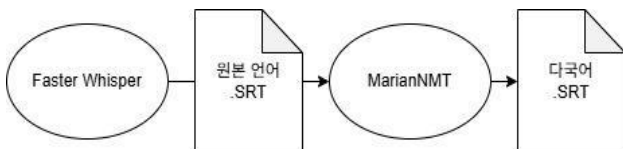


그림 2. AI 파이프라인.

그림 2 파이프라인을 통해 단일 동영상 업로드만으로 다양한 언어의 자막을 자동으로 생성할 수 있게 되어, 다국어 콘텐츠 제작 및 배포 과정의 효율성을

크게 향상시킬 수 있었다. 특히 도메인 사전을 활용한 사후 처리 단계의 도입으로 자동 생성된 자막의 전문성과 정확성이 대폭 개선되었다.

## 3. 결론 및 기대효과

본 논문에서는 온라인 강의 플랫폼을 대상으로 AI를 활용한 자동 자막 생성 및 다국어 번역 시스템을 제안하였다. 제안된 시스템은 Faster Whisper 기반의 음성 인식(STT)과 사후처리 기술, MarianNMT 기반의 다국어 번역 모델을 활용하여 강의 등록 시 자동으로 다국어 자막(SRT)을 생성하고 저장한다.

제안된 시스템은 기존의 YouTube 자동 자막 방식에 대해 다음과 같은 차별성을 지닌다. Youtube의 자동 자막 방식은 실시간 요청에 기반하여 동작하지만 제안된 시스템은 자막을 미리 생성하여 서버에 저장하는 방식을 채택하여 실시간 요청으로 인한 서버 부하 문제를 효과적으로 줄일 수 있을 것이다. 또한, YouTube 자동 자막은 일반적인 음성 데이터에 최적화되어 있어 전문적인 강의 내용에 대해서는 정확성이 떨어질 수 있다. 그러나 본 시스템은 도메인 특화 사전을 통한 사후처리 기술을 적용하여 전문 용어 및 표현의 정확도를 높였다.

YouTube 자동 자막의 경우 생성된 자막에 대한 사용자의 직접적인 피드백이나 수정 요청에 제한적으로 운영된다. 이에 비해 본 시스템은 사용자의 피드백을 통해 자막 품질을 관리하고 개선할 수 있는 구조를 갖추고 있다.

이러한 특성은 온라인 강의 플랫폼만 아니라, 다양한 분야로의 확장을 통해 글로벌 사용자 접근성과 콘텐츠 몰입도 향상을 기대한다.

## Acknowledgement

“이 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업 지원을 통해 수행되었음” (2021-0-01082)

## 참고문헌

- [1] Lei, S., "The Role of Videos' Subtitles in Second Language Acquisition," *Journal of Education, Humanities and Social Sciences*, vol. 13, pp. 12–20, 2023.
- [2] Striuk, A. M., and Hordiienko, V. V., "Research and Development of a Subtitle Management System Using Artificial Intelligence," *CS&SE@SW 2024: 7th Workshop for Young Scientists in Computer Science & Software Engineering*, Kryvyi Rih, Ukraine, December 27, 2024, pp. 415–427.