

Z 세대를 위한 AI 기반 대학 강의 숏폼 변환 및 상호작용적 학습 플랫폼

김건우¹, 홍성환¹, 이석우¹, 이한준¹, 양다원¹, 김인권^{2*}

¹한성대학교 컴퓨터공학부 학부생

²아이스큐브랩

riven5124@hansung.ac.kr, 2171115@hansung.ac.kr, 1971377@hansung.ac.kr, 2091083@hansung.ac.kr,
2271505@hansung.ac.kr, ingunbi@gmail.com

Short-form Transformation and Interactive Learning Platform for AI-Based University Lectures for Generation Z

Geonu Kim¹, Sung-Hwan Hong¹, Seok-Woo Lee¹, Han-Jun Lee¹, Da-Won Yang¹, In-Kwon Kim^{2*}

¹Dept. of Computer Engineering, Hansung University

²*Icecubelab. Inc

*Corresponding Author

요 약

Z 세대는 숏폼 콘텐츠 소비에 익숙하나, 기존 대학 LMS 는 긴 영상 중심으로 몰입도를 저하시킨다. 본 논문은 AI 를 활용해 긴 강의를 1 분 내외의 ‘강의 쇼츠’ 로 자동 변환하는 플랫폼 ‘E-Clips’ 를 제안한다. 멀티모달 AI 기반의 4 단계 자동 변환 파이프라인은 교수의 제작 부담을 줄이고, 친숙한 UI/UX 로 학습 접근성을 높인다. 또한 챗봇, 퀴즈 등 상호작용 기능을 통합해 능동적 학습을 유도한다. 본 연구는 Z 세대의 미디어 소비 특성을 교육에 긍정적으로 통합하여 학습 참여와 효과성을 증대시키는 AI 에듀테크 솔루션의 가능성을 제시한다.

1. 서론

[1] Z 세대는 디지털 네이티브로서, 틱톡(TikTok)이나 인스타그램 릴스(Instagram Reels)와 같은 숏폼 비디오 플랫폼을 통해 정보를 빠르고 시각적으로 소비하는 데 익숙하다. 이러한 미디어 소비 습관은 기존의 교육 방식에 중대한 도전을 제기한다. [2] 대부분의 대학에서 사용하는 전통적인 학습관리시스템(LMS)은 긴 길이의 강의 영상 위주로 구성되어 있어, Z 세대 학습자의 학습 동기와 몰입도를 이끌어내기 어렵다.

본 연구에서는 이러한 문제를 해결하기 위한 방안으로 ‘E-Clips’를 제안한다. E-Clips 는 AI 기술을 통해 고품질의 장편 교육 콘텐츠를 Z 세대에게 친숙한 숏폼 형식으로 재구성하고, 여기에 상호작용적 학습 요소를 결합하여 학습 효과를 극대화하는 혁신적인 교육 플랫폼이다. 본 플랫폼은 킬링타임 콘텐츠 형식을 강력한 교육 도구로 전환하는 것을 목표로 한다.

2. 관련 연구

[3] 고등교육 분야의 인공지능(AIED)은 주로 개인화된 학습 경로 추천, 지능형 튜터링 시스템 등 학습 과정을 지원해왔다. 이와 더불어, Z 세대의 학습 특성에 부합하는 마이크로러닝(Microlearning)이 주목받고 있다. [4] 마이크로러닝은 정보를 짧은 단위로 제공하여 학습자의 집중력을 유지시키고, 지식이 단기 기억에서 장기 기억으로 효과적으로 전환되도록 돕는다. 숏폼 비디오는 이러한 마이크로러닝 전략을 구현하는 효과적인 도구로, 짧은 시간 내 시각적, 청각적 요소를 활용해 학습자의 주의를 집중시키고 핵심 개념의 이해를 돕는다. E-Clips 는 Z 세대에게 친숙한 숏폼 형식을 통해 교육 콘텐츠의 ‘전달 방식’ 자체를 혁신한다는 점에서 기존 AIED 연구와 차별성을 가진다.

[5] 최근 거대언어 모델(LLM)과 멀티모달 LLM(M-LLM)의 발전은 자동 비디오 요약 기술에 큰 진전을 가져왔다. [6] LLMVS 와 같은 최신 연구들은 비디오 프레임과 텍스트를 함께 분석하여 핵심 내용을 요약하는 ‘local-to-global’ 접근법을 사용한다. E-Clips 의 AI

파이프라인 역시 강의 영상의 음성 텍스트를 LLM 으로 분석하는 유사한 구조를 채택하되, 단순 요약이 아닌 교육적 가치(핵심 개념, 정의, 예시)를 기준으로 학습 모듈을 재구성하도록 설계되었다. 이는 범용 요약 기술을 교육 도메인에 맞게 고도화한 새로운 시도이다.

3. E-Clips: AI 기반 교육 콘텐츠 변환 시스템

E-Clips: AI 기반 교육 콘텐츠 변환 시스템 E-Clips는 학생용 모바일 어플리케이션, 교수용 웹 대시보드, 백엔드 서버, 그리고 핵심 AI 처리 파이프라인으로 구성된 통합 시스템이다. 학생들은 모바일 앱을 통해 강의 쇼츠로 핵심 개념을 학습하고, 교수는 웹을 통해 강의 자료를 업로드 및 학습 현황을 관리한다. 백엔드 서버는 사용자 데이터와 학습 콘텐츠를 관리하며 AI 파이프라인과의 연동을 담당한다.

본 시스템의 핵심 기술은 긴 강의 영상을 교육용 숏폼으로 자동 변환하는 4 단계 AI 파이프라인이다. 첫 단계로, Naver Clova Speech 를 사용하여 원본 영상의 음성을 텍스트 대본으로 변환한다(STT). 이렇게 생성된 텍스트는 후속 자연어 처리 단계의 기초 입력 데이터로 활용된다. 파이프라인의 핵심 엔진 역할을 하는 OpenAI 의 GPT-4o 는 이 대본 전체를 분석하여, 교육적으로 중요한 핵심 주제를 추출하고 Z 세대의 흥미를 유발할 수 있는 숏폼 대본을 재창조한다. 예컨대, 딱딱한 전공 개념을 두 캐릭터가 문답을 주고받는 대화 형식으로 재구성하거나, 복잡한 이론을 시각적 비유를 통해 설명하는 방식이다.

콘텐츠 재창조 과정의 차별점은 정교한 프롬프트 엔지니어링 전략에 있다. GPT-4o 에 적용된 시스템 프롬프트는 ‘숏폼 비디오 스크립트 전문가’라는 명확한 역할을 부여하고, API 응답을 JSON 형식으로 강제하여 후속 모듈과의 안정적인 연동을 보장한다. 또한, 구체적인 예시를 제공하여 모델이 시스템이 추구하는 톤 앤 매너(Tone and Manner)를 구현하도록 유도하며, 한국어 대본 생성 후 이를 바탕으로 영어 영상 프롬프트를 생성하는 연쇄적 과업을 수행한다. 이는 영어 프롬프트에 더 정교하게 반응하는 Text-to-Video 모델의 특성을 활용한 전략이다.

이렇게 생성된 프롬프트를 바탕으로, 다양한 Text-to-Video 모델을 선택적으로 활용하여 실제 영상 클립을 생성한다. 대본은 인공지능 음성(TTS)으로 합성되며, 마지막으로 생성된 영상, 음성, 자막을 병합하고 숏폼 환경에 최적화된 9:16 비율로 편집하여 최종 ‘강의 쇼츠’를 완성한다. 또한, 수동적 시청을 방지하기 위해 [7] RAG 기술을 활용한 AI 챗봇과 각 쇼츠 내용 기반의 자동 생성 퀴즈 등 상호작용 기능을 제공하여 학습자의 능동적 참여와 몰입도를 극대화한다.

4. 결론 및 향후 연구

본 논문은 Z 세대 학습자의 미디어 소비 특성과 기존 대학 교육 방식 간의 괴리에서 발생하는 문제를 정의하고, 이를 해결하기 위한 AI 기반 교육 플랫폼 ‘E-Clips’를 제안했다. 제안된 시스템의 핵심인 4 단계 AI 파이프라인은 긴 강의 영상을 교육적 가치를 유지하면서 매력적인 숏폼 콘텐츠로 자동 변환할 수 있다. 또한 AI 챗봇 및 퀴즈와 같은 상호작용 기능으로 학습자의 능동적 참여를 유도한다. 이를 통해 E-Clips 는 Z 세대의 특성에 최적화되어 학습 장벽을 낮추는 가능성을 제시한다.

향후 연구로는 첫째, 실제 대학생들을 대상으로 E-Clips 플랫폼의 사용성 평가 및 학습 효과성 검증을 위한 사용자 연구를 수행할 계획이다. 이를 통해 전통적인 LMS 와 비교하여 학습 몰입도, 학업 성취도, 콘텐츠 이해도에 미치는 영향을 정량적으로 분석하고자 한다. 둘째, 장기적으로 대학 교육을 넘어 평생 교육 및 직업 훈련 분야로 플랫폼을 확장하여, 보다 폭넓은 학습자에게 맞춤형 교육 콘텐츠를 제공하는 것을 목표로 한다.

5. 사사

※ 본 논문은 과학기술정보통신부 대학디지털교육 역량강화사업의 지원을 통해 수행한 한이음 드림업 프로젝트 결과물입니다.

참고문헌

- [1] C. Seemiller and M. Grace, *Generation Z Goes to College*, San Francisco, Jossey-Bass, 2016
- [2] S. Ozkan and R. Koseler, "Multi-dimensional students' evaluation of e-learning systems in the higher education context: An empirical investigation," *Computers & Education*, Vol. 53, No. 4, pp. 1285-1296, 2009
- [3] L. Chen, P. Chen, and Z. Lin, "Artificial Intelligence in Education: A Review," *IEEE Access*, Vol. 8, pp. 75264-75278, 2020.
- [4] G. S. M. Okkol, "Microlearning an E-learning Approach (Advantages and Disadvantages)," *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, Vol. 10, No. 1, pp. 29-35, 2019.
- [5] H. Liu, et al., "Visual Instruction Tuning," *Advances in Neural Information Processing Systems (NeurIPS)*, Vol. 36, pp. 1-23, 2023.
- [6] J. Lee, et al., "Video Summarization with Large Language Models," in *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1-11, 2025.
- [7] P. Lewis, et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," *Advances in Neural Information Processing Systems (NeurIPS)*, Vol. 33, pp. 9459-9474, 2020.