

NLP 기반 소설 IP 확장 솔루션 개발

김혜수, 류영주, 유수연, 황성민
 덕성여자대학교 컴퓨터공학전공 학부생

20200733@duksung.ac.kr, yellowjam144@duksung.ac.kr, dbtndus0527@gmail.com,
 sungmin000720@gmail.com

NLP-based Solution for Expanding Novel IPs
into Multimedia Franchises

Hye-Soo Kim, Young-Joo Ryu, Su-Yeon Yoo, Sung-Min Hwang
 Dept. of Computer Engineering, Duksung Women's University

요약

소설, 웹툰이 영상물로 재탄생하는 사례는 신선한 소재와 검증된 지식재산권(IP)에 기인한다. 본 연구에서는 Confusion Resolution과 CSS 정보를 활용하여 소설을 대본으로 변환하고, Stable Diffusion 모델을 통해 스토리보드를 제작하는 알고리즘을 제안한다. 이 알고리즘은 대본과 스토리보드 생성을 자동화하여 시간과 인건비를 절감함으로써 영상물 제작 과정의 초기 단계의 효율성을 높이며, 게임 개발, 교육 자료 개발, 마케팅 콘텐츠 제작 등 다양한 분야에서 활용될 것으로 기대된다.

1. 서론

최근 소설, 웹툰을 기반으로 영화, 애니메이션으로 재탄생하는 사례가 증가하고 있다. 이는 신선한 소재와 스토리라인, 검증된 지식재산권(IP)에 기인한다. 소설의 영상화를 위한 대본 및 스토리보드 제작 업무 자동화의 필요성을 논하였다. 본 논문에서 자연어 처리를 통한 대본화 알고리즘, 이미지 생성 모델을 이용한 스토리보드 제작 알고리즘을 제안한다.

2. 소설 텍스트를 대본 형식으로 변환

2.1. 소설 속 등장인물 추출

가장 먼저 등장인물을 목록화하기 위해 상호참조해결(Coreference Resolution)을 수행하고 개체명 인식(NER, Named Entity Recognition)과정을 거친다. 상호참조해결이란 동일한 개체를 나타내는 다양한 명사구를 식별하고 그룹화하는 과정으로, ETRI API를 사용하여 장문 텍스트에 대해 작업하였다.

그룹화된 명사구에 대해 개체명 분류를 수행하여 등장인물을 지칭하지 않는 명사구를 등장인물 후보군에서 제외한다. 사람(PS_NAME), 동물(AM), 직위 및 인간관계 명칭(CV) 외의 명사를 제거하였다.

등장인물 추출 결과: 공주, 어머니, 계란, 집
 개체명 태깅 결과: 공주, 어머니

<예시1> 개체명 인식을 거쳐 등장인물 필터링

각 등장인물을 대표하는 명사구(중심어)를 선정한다. 선정 기준은 다음과 같다: ①소설에서 언급된 명사구의 개수가 가장 많은 경우 ②대명사가 아닌 명사구. 만약 한 등장인물에 대해 언급된 횟수가 동일하면, 처음 언급된 명사구를 선택한다.

왕자: [왕자, 그, 남자는, 남자, 왕자, 왕자님]
 중심어: 왕자

<예시2> 등장인물 중심어 선정

소설 속 등장인물을 지칭하는 대명사에 대해서도 중심어로 치환한다. konlpy를 사용하여 명사구를 형태소 단위로 분리하고, 대명사를 중심어로 치환한 후 다시 조사와 결합한다.

치환 전: 그는 공주를 바라보며 말했습니다.
 치환 후: 왕자는 공주를 바라보며 말했습니다.

<예시3> 대명사를 중심어로 치환

2.2. 등장인물과 대사 매칭

소설 내에서 인용문의 발화자는 주변 문장에서 언급된다는 특성에 기초하여, 특정 대사에 대해 발화자를 추정한다. 인용문 주변의 앞뒤 맥락 정보는 발화자와 인용문 사이의 정보 (CSS, Candidate Specific Segments)를 활용한다.[1] BERT 모델을 사용하여 각 후보자에 대해 거리 점수를 계산하고, 가장 확률이 높은 등장인물(중심어)을 매칭한다.

[발화자 추정 전]
 “하지만 **웬디**, 여긴 3층이야.”“**엄마**, 창문 주변에도 나뭇잎이 떨어져 있지 않나요? 그건 피터가 다녀간 표시예요.” 웬디의 말대로 창문 주변에도 나뭇잎들이 흩어져 있었습니다.

[발화자 추정 후]
엄마 하지만 웬디, 여긴 3층이야.
웬디 엄마, 창문 주변에도 나뭇잎이 떨어져 있지 않나요? 그건 피터가 다녀간 표시예요.

<예시 4> 동화 피터팬 발화자 추정

2.3. 비대사 지문으로부터 행위 지문 생성

소설 내에서 따옴표로 표현된 부분 외의 비대사에서, 한국 형태소 품사 태깅(POS Tagging)을 통해 구분된 동사 중에서도 인물의 행위만을 추출하기 위해, 주체를 나타내는 형태소를 모두 포함하는 동사구를 행위지문으로 분류하고 대사와 매칭한다.

2.4. 장소 및 소품 추출, 장소 기반 씬 단위 분할

인물과는 구별되는 주요한 명사를 인식하기 위해, NER(개체명 인식, Named Entity Recognition) 기법을 참고하여 소설 데이터 내에서 장소와 소품을 라벨링한 자체 데이터셋을 구축하고, 커스텀 NER 모델을 개발하였다.

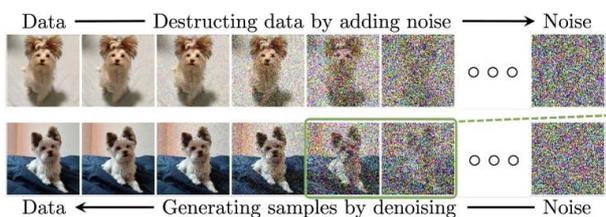
3. 대본을 스토리보드 형식으로 변환

3.1. 컷 단위로 분할 후 프롬프트 생성

인물이 나타난 그림을 묘사하기 위해 인물의 행위를 중심으로 컷을 분할하였고, 컷 단위로 가공한 자연어 처리 결과물을 프롬프트로 사용하였다.

3.2. 스토리보드 이미지 생성(Text to Image)

텍스트에서 이미지를 생성하는 과정에는 diffusion 모델을 채택하였다. Stable Diffusion은 주어진 랜덤 노이즈로부터 이미지 생성 과정을 역으로 진행하는 방식으로 작동한다. 초기 노이즈를 점진적으로 제거하면서 점차 고해상도의 이미지를 복원한다.



<그림1> diffusion 모델 원리

SDXL(Stable Diffusion XLarge) 모델은 대규모의 UNet 백본을 포함하고 있어 attention block과 cross-attention context에 의한 텍스트 인코더의 성

능이 향상된 모델이다.[2] 학습 규모와 용량을 고려하여 LoRA (Low-Rank Adaptation) 방식으로 학습된 UNet 사전을 통해 화풍을 적용하였다.

표1은 임의의 프롬프트에 대해 base model과 safetensors 종류를 달리하여 비교한 결과이다. (The red hat girl knocked on the door.) 최종적으로 본 논문에서는 대본을 스토리보드로 변환하기 위해 SDXL 1.0 모델에 Storyboard Sketch의 화풍을 적용하였다. 또한 이미지 파이프라인 파라미터를 검토하여 생성 시간을 단축할 수 있는 조합을 택하였다.

SDXL 1.0	SD 1.5	SD 1.5
Storyboard Sketch	doodle	dan_milligan
		

<표1> Stable Diffusion 이미지 생성 결과 비교

4. 결론

본문의 알고리즘을 통해 원작 소설 내용을 최대한 반영한 대본과 스토리보드를 제작하였다. 이 과정을 자동화함으로써 영화 및 TV 드라마 제작 과정의 초기 단계에서 투입할 자원을 절감하고, 비주얼 노벨 형태의 게임 개발 등에도 활용될 것으로 기대된다.



<그림2> 스토리보드 생성 결과

Acknowledgement

본 논문은 과학기술정보통신부 대학디지털교육역량강화 사업의 지원을 통해 수행한 ICT멘토링 프로젝트 결과물입니다.

참고문헌

[1] 전승호, 윤은, 이하원, 강민혁, 김정원 and 유길상. (2024). 텍스트 맥락 정보를 활용한 CSN기반 발화자 인식 시스템 개발. 한국정보기술학회논문지, 22(5), 151-163.

[2] Dustin P., Zion E., Kyle L., Andreas B., Tim D., Jonas M., Joe P., Robin R., (2023). "SDXL: Improving Latent Diffusion Models for High-Resolution Image Synthesis". Retrieved from <https://arxiv.org/abs/2307.01952>, 3.