

스포츠 중계 영상 내 문자인식 성능 향상을 위한 후처리 기법 연구

이채운¹, 장현겸¹, 김대한¹, 신윤호², 이웅희^{1*}

¹한성대학교, ²엘지유플러스

¹{rjh2436, hyungyumjang, kkkhaeun, whlee}@hansung.ac.kr, ² yoon731@gmail.com

A Study on Post-Processing Technique for Enhancing Text Recognition in Sports Broadcast Videos

Chae-Eun Lee¹, Hyungyum Jang¹, Daehan Kim¹, Yoonho Shin², Woonghee Lee^{1*}

¹Hansung University, ²LG Uplus

요약

본 연구는 KBO(한국야구위원회) 프로야구 하이라이트 영상에서의 문자인식 성능 개선을 위한 새로운 후처리 기법을 제안하고 검증한다. TRBA(TPS-ResNet-BiLSTM-Attn) 모델을 한국어 글자체 데이터셋으로 재학습하고, 중계 영상에 등장하는 주요 문자 범주를 정의하였다. 선수 및 팀 정보 크롤링을 통해 구축한 데이터베이스를 바탕으로 Jaccard 유사도, Cosine 유사도, N-gram 유사도 함수를 활용한 후처리 기법을 적용하였다. 실험 결과는 제안된 후처리 방법이 문자인식의 정확도를 크게 향상시킴을 보여준다.

I. 서론

최근 스포츠 중계 영상에서의 문자인식은 중요한 연구 영역으로 자리 잡고 있다. 중계 영상에는 점수판, 선수 이름, 통계 등의 다양한 정보가 포함되어 있어, 이 정보의 정확한 인식은 중계 영상의 가치를 크게 향상시킬 수 있다.

본 연구는 스포츠 중계 영상에서 문자인식의 성능을 개선하기 위한 후처리(Post-Processing) 기법에 초점을 둔다. 스포츠 중계와 같은 동적이고 복잡한 환경에서의 문자인식은 배경의 변화, 문자의 변형, 다양한 폰트 스타일 등으로 인해 정확도가 크게 저하될 우려가 있다.

이에 본 연구에서는 야구 중계 영상을 대상으로 크롤링 및 유사도 함수를 사용한 후처리 기법을 통해 문자인식 성능을 개선한다. 제안 방법의 성능 평가를 위해, 네이버 Clova AI에서 개발한 Scene Text Recognition TRBA (TPS-ResNet-BiLSTM-Attn)[1] 모델을 사용하였다. 분석 결과, 본 연구에서 제안된 후처리 기법을 적용함으로써 문자인식의 정확도가 향상됨을 확인할 수 있었다.

그림 1은 본 연구에서 제안하는 후처리 기법이 포함된 OCR 시스템의 전체 동작 과정을 보여준다. 일반적인 OCR 시스템에서 후처리 기법은 문자인식 과정 이후에 이루어지는데, 본 연구의 후처리 기법은 문자인식 단계와 병행하여 진행된다. 이 시스템은 문자인식이 이루어지는 동안 웹 크롤링(Web Crawling)을 통해 야구 경기에 등장하는 선수와 팀 등 관련 정보를 포함하는 데이터베이스를 구축한다. 이후 문자인식 모델의 결과가 데이터베이스 내의 항목과 일치하는지 검사한다. 결과값이 데이터베이스 내에 존재할 경우, 해당 인식이 정확하다고 판단하고 결과를 출력한다. 반면, 결과값이 데이터베이스에 없으면, 데이터베이스 내의 단어들과 유사도를 계산하여 가장 높은 유사도를 보이는 단어를 새로운 결과값으로 선택해 최종 출력값으로 교체하여 출력한다.

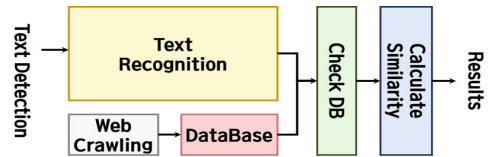


그림 1. OCR 시스템 전체 동작 과정

A. 환경 설정

본 연구는 KBO(한국야구위원회) 경기 영상

II. 제안방법

내의 문자인식을 위해 TRBA 모델을 한국어 글자체 데이터셋[2]으로 재학습시켰다. 추가로 2023년도 KBO 프로야구 경기 하이라이트를 분석하여 중계화면 스코어보드에 등장하는 주요 문자열들을 조사하였다[3]. 이 조사를 통해 투수 이름, 타순, 타자 이름, 투구 횟수, 팀 이름, 팀 점수, 이닝, 볼카운트 등 문자인식에 필요한 범주를 설정했다. 이 문자열들은 숫자를 값으로 가지는 경우와 한글을 값으로 가지는 경우로 구분하여, 각 범주별로 가능한 값들을 정의하였다.

B. 웹 크롤링 및 데이터베이스 구축

KBO 홈페이지에 게시된 선수 등록 정보를 기반으로 Python의 'Selenium' 라이브러리를 사용하여 선수 정보를 크롤링한다[4]. 이 정보를 활용하여 투수와 타자 이름의 범주를 정확히 설정할 수 있다. 수집된 날짜 및 팀별 1군 엔트리 정보는 그림 2와 같이 데이터베이스에 저장되어, 최종적으로 A과정에서 파악한 값들과 B과정에서 크롤링된 데이터를 종합하여 데이터베이스를 구축한다.

그림 2. 날짜 및 팀별 1군 엔트리 현황 크롤링 일부

C. 후처리

문자인식 모델이 반환하는 결과값이 데이터베이스에 없을 경우, 이를 잘못 인식된 값으로 간주하고 후처리 과정을 시작한다. 인식된 값과 데이터베이스 내 문자들 사이의 유사도를 계산하기 위해 Jaccard 유사도, Cosine 유사도, N_gram 유사도 함수를 사용한다. 이 중 가장 높은 유사도를 가지는 단어를 새로운 예측값으로 선택하고, 사용자에게 최종적으로 반환한다.

III. 실험 결과

그림 3에서는 KBO 프로야구 하이라이트 영상 내 선별된 장면 중, 스코어보드의 문자인식 과정에서 정확히 인식되지 못한 문자들의 원본 이미지들을 보여준다. 표 1의 기존 인식값에서 알 수 있듯이 박세혁을 박세로 예측하고, 류진욱을 류진욱으로 예측했다. 이후, 이 문자들에 대하여 후처리를 수행하였다. 후처리 과정에서 여러 유사도 함수들이 사용되었으며, 각각의 결과는 표 1에 기록되어 있다.

본 실험을 통해, Cosine 유사도와 Jaccard 유

사도 함수를 포함한 유사도 함수들을 적용한 결과, 문자들이 정확하게 인식되는 것을 확인할 수 있었다. N_gram 유사도의 경우 다른 유사도 함수에 비해 정확히 수정되지 않음을 확인하였다.



그림 3. 정확히 인식되지 않는 문자들의 원본 이미지들

기존 인식값	Cosine	Jaccard	N_gram
박세	박세혁	박세혁	박세혁
류진욱	류진욱	류진욱	류진욱
오윤석	오윤석	오윤석	박세혁
황재균	황재균	황재균	황재균

표 1. 유사도 함수에 따른 후처리 결과 비교

IV. 결론

본 연구에서는 스포츠 중계 영상, 특히 KBO 프로야구 하이라이트의 문자인식 문제를 해결하기 위한 후처리 기법을 제안하고, 그 효과를 실험을 통해 검증하였다. 크롤링을 통해 데이터베이스를 구축하고 Cosine, Jaccard, N-gram 유사도 함수를 활용한 후처리 접근법은 문자인식의 성능을 향상시켰다. 이러한 진전은 스포츠 중계 영상에서의 문자인식 기술 발전에 중요한 기여를 한다. 향후 연구에서는 본 후처리 기법을 다양한 스포츠 중계는 물론 다양한 동적 환경에 적용하고, 실시간 문자인식 기술 개발 등을 통해 기법의 범위와 효율성을 확장할 것이다.

ACKNOWLEDGMENT

본 연구는 한성대학교 학술연구비 지원과제임.

참고 문헌

[1] Baek, Jeonghun, et al. "What is wrong with scene text recognition model comparisons? dataset and model analysis." Proceedings of the IEEE/CVF international conference on computer vision. 2019.
 [2] Korean Character Font Images Dataset, AI Hub, accessed on 2019.
 [3] <https://www.koreabaseball.com/Schedule/Schedule.aspx>
 [4] <https://www.koreabaseball.com/Player/RegistrarAll.aspx>