

컴퓨터공학부 캡스톤디자인 중간평가 답변서

팀명: 4 조 유하

조원: 이태훈,이주형,이인평,김성수,김민재

심사의견 or 질문

향후 계획 내용을 보완하면 좋겠습니다. 예를 들어, 영화, 드라마 등 자막이 있는 동영상은 자막을 추출하여 욕설을 탐지하는 기능을 추가하면 좋겠습니다. 한글 문장 토큰화 방식으로 sentencePiece 방법을 적용하는 방법, 워드 임베딩 방법으로 GloVe 등의 방법을 이용하여 실험하는 것도 고려할 수 있겠습니다.

답변

1) 자막을 추출하여 욕설을 탐지하는 기능

저희가 자막을 추출하는 기능, 즉 이미지에서 글자를 추출하는 기능을 Tesseract-OCR 과 Google Cloud PlatForm 에서 제공하는 Vision API 의 OCR, 2 가지 툴로 구현해보았습니다. Tesseract-OCR 의 경우 무료 서비스이고 Google API 의 경우 유료 서비스이지만 정확도가 높고 제공하는 기능들이 강력하여 저희는 Google Vision API 의 OCR 을 선택했습니다.

2) 문장 토큰화, GloVe 적용

말씀해주신대로 Hanbert 및 Glove 등 다양한 툴로 실험을 해보았지만 정확도 측면에서의 성능과 소요되는 시간, 모델의 용량 등을 고려해보았을 때 현재 저희가 사용하는 형태소 분석과 FastText 가 가장 효율적이라고 판단하여 계속해서 사용하는 것으로 결론을 내렸습니다.

심사의견 or 질문

시연 효과 향상을 위해 음성 마스킹(빠~ 처리) 구현이 가능할지 고려해보기 바랍니다.

답변

1) 음성 마스킹 처리

말씀해주신대로 음성 마스킹 기능을 구현하였습니다. 먼저 저희는 Google Speech To Text API 를 채택하여 텍스트로 변환하고 욕설을 감지하는 시스템입니다. Google API 에서 텍스트를 추출할 때 해당 시간을 nanosecond (ns)단위로 Time Stamp 해주어서 해당 기능을 비교적 간편하게 구현할 수 있었습니다

심사의견 or 질문

자막을 자동으로 추출하는 기능등의 내용이 좀더 보완되면 좋을 것 같습니다.

답변

첫번째 피드백과 동일하게 저희가 자막을 추출하는 기능, 즉 이미지에서 글자를 추출하는 기능을 Tesseract-OCR 과 Google Cloud PlatForm 에서 제공하는 Vision API 의 OCR, 2 가지 툴로 구현해보았습니다. Tesseract-OCR 의 경우 무료 서비스이고 Google API 의 경우 유료 서비스이지만 정확도가 높고 제공하는 기능들이 강력하여 저희는 Google Vision API 의 OCR 을 선택했습니다.