

컴퓨터공학부 캡스톤디자인 중간평가 답변서

팀명: 4 조 YouHi

조원: 이태훈,이주형,이인평,김성수,김민재

심사의견 or 질문
과제범위가 넓은 것 같으니 설계-구현 과정에서 중요도를 정하여 범위를 축소 또는 중요한 주제로 제한할 필요가 있을 것 같음. 과제수행 단계를 1) 이미지에 대한 등급 및 유형 판단, 2) 동영상에 적용, 3) 유튜브 등 API 에 integration 등으로 구체화하면 좋겠음.
답변
1)이미지에 대한 등급 및 유형 판단 - 저희는 영상 판단 부분에서 프레임 추출시 너무 많은 프레임이 추출되어 서비스 속도가 매우 저하된다는 기술적 한계점을 갖고 있었습니다. 강승식 교수님의 피드백을 받아들여 먼저 이미지에 대한 판단을 진행하는데,그 과정은 다음과 같습니다. 전체 동영상에서 3 초에 1 프레임을 추출합니다. 추출된 프레임은 Object Detection 을 이용해 "칼, 담배, 총, 옷을 선정적으로 입은 여자"를 감지합니다. 2) 동영상에 적용 - Object Detection 을 통해 만약 위와 같은 유형의 Object 가 감지된다면, 해당 이미지가 추출된 구간(3 초)에 대해 영상 판단을 진행합니다. 3) Youtube 등 API 에 intergration 저희는 YouTube 에 적용시키겠다라는 현실적으로 불가능한 목표를 다음과 같이 수정했습니다. "YouTube 등 동영상 플랫폼에 적용시킬 수 있는 높은 정확도를 가진 영상 연령 제한 감지 시스템의 구현"

심사의견 or 질문
기간 내 개발 가능한 기술로 초점을 맞추어 프로젝트 범위를 한정하는 것이 필요하리라고 생각합니다.
답변
프로젝트 목표를 다음과 같이 수정했습니다. "YouTube 등 동영상 플랫폼에 적용시킬 수 있는 높은 정확도를 가진 영상 연령 제한 감지 시스템의 구현"

프로젝트 범위를 다음과 같이 한정했습니다.

(기존)

1. 폭력적인 장면

- 칼에 베이거나 찢려 피가 튀는 장면
- 총이 나오는 게임에서 사람 및 사람을 닮은 물체를 쏘을 때 피가 튀는 장면

2. 유해한 장면

- 담배를 피는 장면

3. 선정적인 장면

- 여성이 옷을 벗거나 속옷만 입은 장면

4. 저속한 언어가 포함된 장면

- 한국어로 표현된 욕설이 들어가있는 장면

(수정 내용)

1. 칼에 베이거나 찢려 피가 튀는 장면 -> 칼에 찢리는 장면이 부각되거나 명확한 장면

2. 총이 나오는 게임 -> 1 인칭 FPS 게임으로 제한

3. 담배를 피는 장면 -> 담배 연기가 충분히 표현된 장면

4. 여성이 옷을 벗거나 속옷만 입은 장면 -> 여성 혼자 상반신을 노출한 장면

심사의견 or 질문

기존 시스템(예: 유튜브)와의 차이점은 무엇인가요? 또는 차별화 전략이 있는지 궁금합니다.

답변

기존 시스템에 대해 먼저 설명드리겠습니다. YouTube 를 제외한 많은 동영상 업로드 공간에서 자동 영상 검열 시스템은 현재 존재하지 않습니다. 또한 YouTube 는 "노란 딱지"라는 자동 영상 검열 시스템 적용이 있는데, 이것은 19 세 연령 기준에 대한 검열이 아닌 광고가 적합한 영상인지 (예, 정치성 발언이 짙은 영상)으로 검열합니다. 또한 업로드 이후 진행되는 것이기 때문에, 업로드 이후 영상 검열이 진행되기 이전까지 많은 사람들에게 그대로 노출됩니다. 마지막으로 이 "노란 딱지"가 적용된 영상을 업로드한 업로더들은 자신의 영상에서 어떤 부분이 검열되었는지 알 수 없습니다.

따라서 저희는 다음과 같은 차별화 전략을 세웠습니다.

1. 영상을 업로드 하기 이전 영상에 대한 자동 검열(검열 시간을 최대한 짧게 하는 것이 목표)
2. YouTube 와 한국 방송 통신 위원회에서 규정한 내용에 대한 검열(정치성 발언 등 논란이 생길 수 있는 규정 제외)
3. 업로더가 자신의 영상 중 어떤 부분이 검열되었는지 해당 구간과 검열 기준을 알림

심사의견 or 질문

데이터 전처리 과정 부분에 대해서도 언급 부탁드립니다.

답변

데이터 전처리는 다음과 같이 이루어집니다.

동영상이 업로드되면, 3 초에 1 프레임씩 추출되어 폴더에 저장됩니다.

Object Detection 을 이용해 감지된 구간(3 초)가 있으면 해당 구간 영상(Subclip)을 검열하는데, 저희는 Two Stream Convolution Network 를 선택했기 때문에, RGB 프레임과 Optical Flow 로 표현된 Frame 모두 필요합니다. 따라서 RGB 프레임을 Optical Flow 로 변환해드립니다. 즉 RGB 를 Horizontal Flow, Vertical Flow 두 이미지로 분리합니다. 이후 (342, 256)크기로 Resized 를 진행한 후, Normalize 를 거쳐 검열합니다.

심사의견 or 질문

분류 모델의 precision/recall 등의 정량적인 수치를 정해서 이를 달성하는 것을 목표로 하는 것이 좋을 듯 합니다.

답변

Precision/Recall 에서 활용되는 True 와 False 를 다음과 같이 설정했습니다.

True: 유해한 장면

False: 유해하지 않은 장면

따라서 유해한 장면을 유해한 장면으로 판단(TT), 유해한 장면을 유해하지 않은 장면으로 판단(TF), 유해하지 않은 장면을 유해하다고 판단(FT), 유해하지 않은 장면을 유해하지 않다고 판단(FF)으로 구성해 Precision/Recall 을 활용할 것입니다.

또한 해당 수치는 Youtube 에서 총 100 개의 영상을 임의로 캡처 및 다운로드 받아 진행하여 표현할 예정입니다.

심사의견 or 질문

프로젝트 scope 를 좀 더 구체적이고 세부적으로 좁혔으면 더 좋은 결과가 나오지 않을까 생각합니다.

답변

이시윤 교수님 의견의 답변과 동일하게, 저희는 기존 프로젝트 범위에서 다음과 같이 구체적이고 세부적으로 좁혔습니다.

1. 칼에 베이거나 찢려 피가 튀는 장면 -> 칼에 찢리는 장면이 부각되거나 명확한 장면
2. 총이 나오는 게임 -> 1 인칭 FPS 게임으로 제한
3. 담배를 피는 장면 -> 담배 연기가 충분히 표현된 장면
4. 여성이 옷을 벗거나 속옷만 입은 장면 -> 여성 혼자 상반신을 노출한 장면입은 장면
-> 여성 혼자 상반신을 노출한 장면