

MRCNN을 이용한 영화속 등장인물 면적추출 방법

김영훈⁰, 유은순^{**}, 강수환^{***}, 박승보^{*}

⁰인하대학교 소프트웨어융합학과

^{**}인하대학교 인공지능 콘텐츠창작 연구센터

^{***}인하대학교 한국학연구소

e-mail: qorxjs@naver.com⁰, tesnier@naver.com^{**}, 22151147@inha.edu^{***}, molaal@naver.com^{*}

Extraction Method of Face Area in Movie Using MRCNN

Kim Yeonghuh⁰, You Eun Soon^{**}, Kang SooHwan^{***}, Seung-Bo Park^{*}

⁰Software Convergence Engineering, Inha University

^{**}AI Content Creation Research Center, Inha University

^{***}Center for korean studies, Inha University

● 요약 ●

본 연구는 영화에 대한 정량적 분석을 위해 MRCNN을 활용한 영화 속 등장인물의 얼굴 면적을 검출하였다. MRCNN을 선택한 이유는 기존 얼굴 인식 시스템이 갖는 한계(뒷모습, 누워있는 모습의 측정 오류)의 개선과 면밀한 계산을 하고자 함이었다. 영화 한편에서 주인공과 상대주인공이 함께 등장한 씬을 선별한 726개의 이미지 중 496개의 이미지가 마스크가 됨으로서 68%의 성능을 보였다. 반면에 230개의 이미지 파일에서는 다소 문제가 발견되어 32%의 오차가 발생했다. 오차를 개선하기 위해서 주요 인물을 학습시킨 뒤 마스크를 씌우는 작업을 함으로써 현 확률보다 높은 확률로 정상적으로 이미지가 추출될 수 있도록 시험해 볼 것이다.

키워드: MRCNN, Character, Story Analyze

I. Introduction

비디오 데이터가 폭발적으로 증가함에 따라 검색과 추천시스템의 성능 향상을 위한 비디오 데이터의 시맨틱 정보를 분석하는 것이 중요해지고 있다. 특히 영화는 스토리라는 고차원적(high-level)정보를 갖고 있기 때문에 스토리를 구성하는 다양한 요소들에 대한 분석이 필수적이다. 인물은 영화의 스토리를 이끌어가는 주체로서 주요 인물들 간에 발생하는 갈등은 스토리의 핵심이라고 할 수 있다. 이에 본 논문은 영화에 대한 정량적 분석을 위해 MRCNN을 이용한 영화속 등장인물의 얼굴면적 추출 방법을 제안한다. 얼굴은 영화에서 인물의 감정과 심리적인 상태와 같은 스토리 정보를 직접적으로 전달하는 역할을 한다. 따라서 얼굴 면적의 면밀한 측정이 필요하므로 MRCNN을 활용하였다. 페이스북 AI팀이 개발한 MRCNN은 각 이미지에 바운딩 박스의 접근이 아닌 픽셀 이미지 레벨의 접근이 가능하기 때문에 얼굴의 정면 측면 뒷면까지 측정하기 용이하다는 장점이 있다. 영화 속 얼굴면적 추출은 영화의 하이라이트 예고편 제작 등 다양한 활용도 가능하다.

II. Related works

이전 Faster RCNN까지는 속도 개선과 효율적인 처리를 기준으로 바운딩 박스를 이용하여 객체들을 식별 하였다면 MASK R CNN에서는 이 바운딩 박스가 아닌 픽셀 수준의 분류로 CNN을 확장시켰다. MRCNN은 Faster RCNN에 Binary Mask라 하는 각 픽셀이 오브젝트에 해당하는 것인지를 이진치를 구별하는 마스크 네트워크를 추가한 것으로 Classification, BBox Regression에 Mask를 더하여 3가지를 얻는 구조이다. Image Pixel level 분류는 바운딩 박스와 다르게 픽셀수준의 특이도를 요구하기에 RoIPooling에서 발생하는 오차를 감소시켜야했으므로 RoIAlign통해 RoIPool을 조정하여 정확한 정렬을 이룸으로써 이 문제를 풀 수 있었다. 결론적으로 RoIAlign을 사용함으로써 Object detection도 4%정도 RoIPool대비 올라가고 Image Pixel level에서의 분류가 가능하게 되었다. [2] 픽셀수준에서의 측정으로 얼굴면적의 면밀한 검출을 시도할 수 있었다.

III. Extraction Method and Evaluation of Face Area Base on MRCNN

본 논문은 영화 <월스트리트>의 주인공(버드)과 상대 주인공(고든)이 함께 등장하는 씬(scene)을 추출하여 이미지 파일을 만들었다. 얼굴 면적을 추출하는 프로그램을 사용하여 각 이미지 파일에 들어있는 두 인물의 얼굴 면적을 측정하였다. 그리고 MRCNN을 활용하여 이미지에 마스크를 입혔으며 얼굴 면적 안에서 마스크 된 픽셀값을 계산함으로써 씬 별 두 인물의 크기를 면밀하게 측정하였다. 726개의 이미지 중 496개의 이미지가 마스크가 됨으로서 68%의 성능을 보였다. 반면에 230개의 이미지 파일에서는 다소 문제가 발견되어 32%의 오차가 발생했다. 그 이유는 문제가 있는 이미지는 전반적으로 어두운 배경으로 인해 사람과 사물 간의 경계선이 모호해지거나, 두 주요 인물이 10명 이상의 많은 인파와 함께 한 장면에 잡힌 경우다. 또한 흐릿하게 보이는 사물이나 모형을 사람으로 인식하는 경우도 발견되었다.



Fig. 1. Normal Image

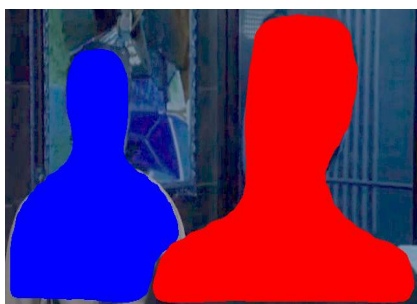


Fig. 2. MRCNN Image

IV. Conclusions

MRCNN을 활용하여 영화 속 인물들의 얼굴 면적을 측정함으로써 영화의 데이터를 정량적으로 측정하고 수집할 수 있게 되었다. 위에서 기술했듯이 마스크가 안된 이미지의 오차를 개선하기 위해서 주요 인물을 학습시킨 뒤 마스크를 씌우는 작업을 함으로써 현 확률보다 높은 확률로 정상적으로 이미지가 추출될 수 있도록 시험해 볼 것이다.

REFERENCES

- [1] Hyun Chul Song, Min-Sik Knag, Tae-Eun Kimg. "Object Detection based on Mask R-CNN from Infrared Camera." *Journal of Digital Contents Society*, Vol. 19 No. 6, pp. 1213-1218, Jun 2018
- [2] K. He, G. Gkioxari, P. Dollár, & R. Girshick, "Mask r-cnn" In *Computer Vision (ICCV), 2017 IEEE International Conference*, pp. 2980-2988, October 2017