

딥러닝 기반 방송 콘텐츠 클래스 분류 시스템 개발

*김신, **윤경로
건국대학교 컴퓨터공학과
*new.xin22@gmail.com

Development of Broadcast Content Class Classification System based on Deep Learning

*Shin, Kim, **Kyoungro, Yoon
Department of Computer Science Engineering, Konkuk University

요 약

최근 수 년간 비디오 콘텐츠 소비 공간이 인터넷으로 확장되며 지능적 비디오 콘텐츠 추천 기술 개발이 진행되어 왔다. 하지만 지능적 비디오 콘텐츠 추천 기술은 사용자의 기호나 업로드된 비디오 콘텐츠의 제목 등을 기반으로 하여 비디오 콘텐츠 클래스에 대한 분석 없이 유사한 비디오 콘텐츠를 탐색하고 추천해주는 기술이 대부분이다.

본 논문에서는 지능적 콘텐츠 추천을 위한 딥러닝 기반 방송 콘텐츠 클래스 분류 시스템을 제안한다. 방송 콘텐츠 내 영상 정보를 이용하여 방송 콘텐츠 클래스를 분류하며 높은 분류 정확도를 보여주는 것을 확인할 수 있다.

1. 서론

최근 수 년간 비디오 콘텐츠를 소비할 수 있는 공간이 인터넷 상으로 확장되면서 방송 콘텐츠 뿐만 아니라 인터넷 방송 콘텐츠도 상일적으로 소비할 수 있게 되었다. YouTube 와 같은 무료 동영상 공유 사이트에서 사용자가 특정 단어를 검색하여 비디오 콘텐츠를 검색할 때, 해당 비디오 콘텐츠와 연관된 추천 목록을 확인할 수 있다. 하지만 추천 비디오 콘텐츠는 사용자의 비디오 콘텐츠 선호도나 업로드된 비디오 콘텐츠 제목 등으로 구성된다. 즉, 재생되고 있는 비디오 콘텐츠에 대한 분석없이 추천 목록이 형성된다.

비디오 콘텐츠 클래스에 대한 분석 없이 비디오 콘텐츠 추천 목록이 형성되기 때문에 해당 콘텐츠 클래스와 연관 없는 추천 콘텐츠들이 포함되기도 한다. 따라서 지능적 비디오 콘텐츠 추천 시스템의 정확도를 올리기 위해서는 해당 비디오 콘텐츠 분석도 요구된다.

본 논문에서는 딥러닝 기반 방송 콘텐츠 분류 시스템을 제안한다. 현재까지는 방송 콘텐츠가 비디오 콘텐츠 중 가장 다양한 소비층을 가지고 있으며 파괴력이 강력하기 때문이다.

본 논문의 구성은 다음과 같다. 2 절에서는 방송 콘텐츠 클래스를 분류에 필요한 배경 지식에 대해 설명하며, 3 절에서는 방송 콘텐츠 클래스 분류 시스템을 제안한다. 4 절에서는 본 논문에서 제안하는 방송 콘텐츠 클래스 분류 시스템에 대한 실험 결과에 대해 서술하며 마지막으로 5 절에서는 본 논문의 결론을 맺는다.

2. 배경 지식

방송 콘텐츠 클래스 분류 시스템은 이미지 분류 기술을 필요로 한다. 이미지 분류 기술은 최근 딥러닝 기법 도입을 통해 엄청난 분류 정확도 향상을 가져왔다.

GoogLeNet[1]은 2014 년 ILSVRC(ImageNet Large Scale Visual Recognition Challenge)에서 1 등을 차지한 이미지 분류 모델로 통상적으로 Inception v1 으로 불린다. Inception 모듈은 CNN 망의 깊이를 깊게 만들어주되 연산량은 크게 증가시키지 않는 것이 가장 큰 특징이며 현재 Inception v3[2]가 일반적으로 가장 많이 사용되고 있는 이미지 분류 모델 중 하나이다.

ResNet[3]은 2015 년도 ILSVRC 에서 우승한 모델로 계층을 일부 뛰어넘는 기술인 Residual Connection 의 도입으로 계산량은 낮추면서 높은 정확도의 이미지 분류를 가능케 했다.

Inception-ResNet[4]은 2016 년 ILSVRC 에서 우승한 이미지 분류 모델로서 Inception 모듈과 Residual Connection 을 융합하여 보다 높은 이미지 분류 정확도를 보여준다. 세 모델의 이미지 분류 에러율 비교는 표 1 을 통해 확인할 수 있다.

표 1. 네트워크에 따른 이미지 분류 오류율 비교(퍼센트) [4]

Network	Resnet [3]	Inception v3 [2]	Inception-Resnet (v2) [4]
Top-5 Error (%)	7.8%	4.6%	4.1%

3. 방송 콘텐츠 클래스 분류

본 논문에서는 딥러닝 기반 방송 콘텐츠 클래스 분류 시스템을 제안한다. 방송 콘텐츠 클래스 분류를 위한 딥러닝 모델은 2016 년 ILSVRC 우승 모델인 Inception-Resnet [4]를 차용하였으며 transfer learning 을 통해 방송 콘텐츠 클래스 분류 모델을 학습하였다.

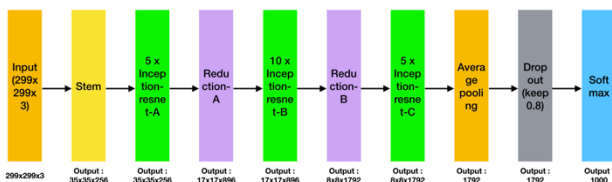


그림 1. Inception-Resnet 모델 기본 구조[4]

4. 실험 결과

본 논문에서는 실험을 위해 News, Entertainment, Documentary, Drama, Concert, Sport 로 구성된 6 개의 방송 콘텐츠 클래스를 정의하고 클래스에 해당하는 데이터 세트를 자체 구축하였으며 방송 콘텐츠 클래스 분류 정확도를 측정하였다. 방송 콘텐츠 클래스 분류 모델 훈련 및 실험 환경은 표 2 과 같다.

표 2. 방송 콘텐츠 클래스 분류 모델 훈련 및 실험 환경

Operating System	Ubuntu 16.04 LTS
GPU	GeForce GTX 980 ti
Framework	Keras, Tensorflow
Language	Python 3.6



그림 2. 방송 콘텐츠 클래스 분류를 위한 훈련 데이터 예시

방송 콘텐츠 클래스 분류 모델을 위한 데이터 세트는 각 방송 콘텐츠 클래스당 훈련용 이미지 940 장, validation 용 이미지 260 장, 분류 실험용 이미지 360 장으로 구성되어 있다. 스케일 조정, 이미지 좌우 상하 반전, 이미지 회전을 통해 훈련용 이미지 데이터를 증대시켰으며 그림 2 는 방송 콘텐츠 클래스 분류 모델을 훈련하기 위한 데이터의 일부 예제이다.

표 3 를 통해 방송 콘텐츠 클래스 분류 실험의 정확도를 확인할 수 있으며 실험 결과 Documentary 클래스를 제외한 모든 방송 콘텐츠에서 약 90%의 정확도로 클래스 분류가 가능한 것을 확인할 수 있다.

표 3. 방송 콘텐츠 클래스 분류 정확도(퍼센트)

Class	News	Entertainment	Documentary	Drama	Concert	Sport
Classification	80.3	93.1	55.8	89.7	87.8	97.8
Precision (%)						

5. 결론 및 향후 과제

본 논문에서는 딥러닝 기반의 방송 콘텐츠 클래스 분류 시스템을 제시하였다. 이미지 분류 알고리즘을 이용해 방송 콘텐츠 클래스를 분류하였으며 높은 분류 정확도를 보여주는 것을 확인할 수 있었다.

향후 연구 방향으로 방송 콘텐츠 클래스 분류를 통해 얻은 비디오 클래스 기반 지능적 비디오 콘텐츠 추천 엔진 및 시스템을 연구해야 할 것이다.

감사의 글

본 논문은 2018 년 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임(2017-0-00024, UHD 방송콘텐츠 기반 지능형 Dynamic Media 생성, 분배 및 소비 기술 개발)

참고 문헌

[1] Szegedy, Christian, et al. "Going deeper with convolutions." *Cvpr*, 2015.

[2] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.

[3] He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[4] Szegedy, Christian, et al. "Inception-v4, inception-resnet and the impact of residual connections on learning." *AAAI*. Vol. 4. 2017.