

인공지능 기반 동영상 장면 자동 분할 기술



김선중 (kimsj@etri.re.kr)



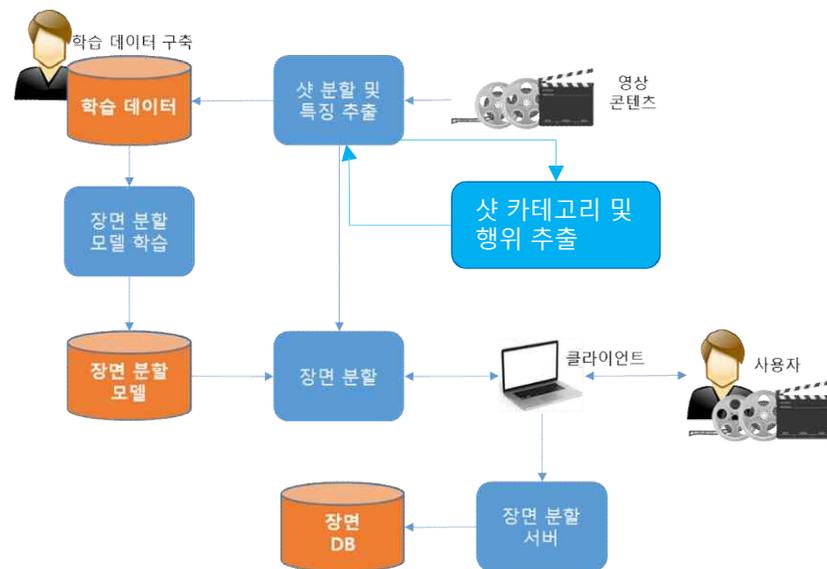
목 차

1. 기술의 개요
2. 기술이전 내용 및 범위
3. 경쟁기술과 비교
4. 기술의 사업성
 - 활용분야 및 기대효과
5. 국내외 시장 동향

1. 기술의 개요

인공 지능 기반 동영상 장면 자동 분할 기

- 술**
- ❖ 영상 콘텐츠를 대상으로 딥러닝 기술을 하여 장면을 생성하는 기술임. 색상 히스토그램의 변화를 기반으로 샷을 분할하고, 분할된 샷의 영상/음성 특징과 행위 정보, 샷 카테고리 정보를 추출 한 후, 이를 학습된 딥러닝 모델에 전달하여 장면을 분할하는 기술
 - ❖ 세 개의 *convolution* 레이어로 이루어진 영상 네트워크, 음성 네트워크와 *bidirectional RNN* 으로 구성된 행위 및 샷 카테고리 네트워크를 기반으로 이들을 묶는 2개의 *dense network* 로 이루어진 앙상블 구조를 가지는 장면 분할 모델을 단일 딥러닝 네트워크의 앙상블로 구성된 장면 분할 모델을 포함함
 - ❖ 본 기술에는 대용량의 장면의 경계지점이 태깅된 영상 콘텐츠 데이터를 입력으로 학습된 모델을 이용하여 장면을 분할함



<본 기술의 개념도>

2. 기술이전 내용 및 범위

□ 기술이전 내

용

- 영상 콘텐츠의 샷 분할 및 특징 추출 기능
- 샷 단위 카테고리 및 행위 추출 기술
- 학습된 장면 분할 딥러닝 모델을 이용한 장면 분할 기능
- 웹 기반의 영상 콘텐츠 장면 분할 관리 기능

□ 기술이전 범

위

- 인공지능 기반 장면 분할 기술 시스템 요구사항 정의서 1종
- 인공지능 기반 장면 분할 기술 상세설계서 1종
- 인공지능 기반 장면 분할 기술 시험 절차 및 결과서 1종
- 인공지능 기반 장면 분할 기술 프로그램 3종
 - 영상 콘텐츠의 자동 샷 분할 및 특징 추출기
 - 인공지능 기반 영상 콘텐츠 장면 자동 분할기
 - 장면 분할 웹 인터페이스

2. 기술이전 내용 및 범위

▣ 기술 개발 현황

❖ 기술성숙도(TRL : Technology Readiness Level) 단 (6)단계

구분	단계	정의	세부 설명
기초 연구 단계	1	기초 이론/실험	기초이론 정립 단계
	2	실용 목적의 아이디어, 특허 등 개념정립	기술개발 개념 정립 및 아이디어에 대한 특허 출원 단계
실험 단계	3	실험실 규모의 기본 성능 검증	실험실 환경에서 실험 또는 전산 시뮬레이션을 통해 기본 성능이 검증될 수 있는 단계 개발하려는 부품/시스템의 기본 설계도면을 확보하는 단계
	4	실험실 규모의 소재/부품/시스템 핵심 성능 평가	시험생물을 제작하여 핵심 성능에 대한 평가가 완료된 단계 3단계에서 도출된 다양한 결과 중에서 최적의 결과를 선택하려는 단계 컴퓨터 모사가 가능한 경우 최적화를 완료하는 단계
시작품 단계	5	확정된 소재/부품/시스템 시작품 제작 및 성능 평가	확정된 소재/부품/시스템의 실험실 시작품 제작 및 성능 평가가 완료된 단계 개발 대상의 생산을 고려하여 설계하나 실제 제작한 시작품 샘플은 1~수개 미만인 단계 경제성을 고려하지 않고 기술의 핵심 성능으로만 볼 때, 실제로 판매가 될 수 있는 정도로 목표 성능을 달성한 단계
	6	파일럿 규모 시작품 제작 및 성능 평가	파일럿 규모(복수 개~양산규모의 1/10정도)의 시작품 제작 및 평가가 완료된 단계 파일럿 규모 생산품에 대해 생산량, 생산용량, 불량률 등 제시 파일럿 생산을 위한 대규모 투자가 동반되는 단계 생산기업이 수요기업 적용 환경에 유사하게 자체 현장 테스트를 실시하여 목표 성능을 만족시킨 단계 성능 평가 결과에 대해 가능하면 공인인증 기관의 성적서 확보
실용화 단계	7	신뢰성 평가 및 수요기업 평가	실제 환경에서 성능 검증이 이루어지는 단계 부품 및 소재 개발의 경우 수요업체에서 직접 파일럿 시작품을 현장 평가(성능 및 신뢰성 평가) 가능하면 인증기관의 신뢰성 평가 결과 제출
	8	시제품 인증 및 표준화	표준화 및 인허가 취득 단계
사업화	9	사업화	본격적인 양산 및 사업화 단계 6-시그마 등 품질관리가 중요한 단계

3. 경쟁기술과 비교

▣ 기존 경쟁기술 대비 개량된 부분

- ❖ 영상 콘텐츠의 영상/음성/샷카테고리/행위 특징을 기반으로 딥러닝 학습 기술을 이용하여 장면 분할 모델을 구축
- ❖ 장르 단위/콘텐츠 단위 모델 학습을 위한 학습기 및 다수 콘텐츠의 동시 처리를 위한 웹 기반 클라이언트/서버 포함

기술분야	기존 기술	본 기술
영상 콘텐츠 장면 분할	<ul style="list-style-type: none"> ● 단일 특징 정보만을 이용한 장면 분할 기술 ● 비지도 학습 기반 장면 분할로 인해 장르/콘텐츠 제작사 등에 따른 특성 반영 불가 ● 의미 단위 장면이 아닌 샷 단위 분할 기술이 대부분 	<ul style="list-style-type: none"> ● 영상 콘텐츠의 영상/음성 특징 기반 관계 정보와 샷 카테고리/행위 특징 기반 흐름 정보를 모두 활용하는 딥러닝 모델 ● 학습 데이터 구축에 따라 장르/제작사에 따른 각기 다른 모델 생성 가능 ● 다중 특징을 활용한 의미기반 장면 분할 가능

4. 기술의 사업성

❖ 예상 응용 제품 및 서비스

예상 제품/서비스	예상 수요자(층)
자동 클립 배포 서비스	방송 사업자
VOD 클립 서비스	VOD 콘텐츠 사업자, 포탈 사업자 등
e-Learning 서비스	인터넷 교육 서비스 사업자

❖ 사업성

예상 제품 /서비스	예상단가 (천원)	이전기술의 비중(%)	잠재적/현재적 경쟁자와 가격,시장 등에서 경쟁상 유리한 점	판매 가능 시기
장면 자동 생성기	50,000	50 %	a. 가격경쟁력면: 기존제품 대비 가격 50% 절감 b. 시장환경면: 2020년 이후 연간 10억 예상	2020년 이후

4. 기술의 사업성

❖ 기술이전 업체 조건

- 1단계 : 기술 이해 및 안정화 (약 6개월 소요 추정)
 - 2단계 : 콘텐츠 확보 및 분할을 위한 서버 구축
(초기 6개월 소요 및 지속적 유지/보수)
 - 3단계 : 상용화 (1년 후 상용화 가능 추정)
- ※ 영상 기반 서비스 경험이 있거나 계획이 있는 업체

❖ 사업화시 제약 조건

애로점	극복(개선)방안
영화 콘텐츠 재활용을 위한 저작권 문제 발생 가능	영화 콘텐츠 및 클립 재활용 서비스를 위한 저작권 문제 발생이 가능한데, 이를 위해 영화 제작사와 사전에 긴밀히 협조하여 애로점을 극복할 필요가 있음.
영상 콘텐츠의 장면 태깅 데이터 구축 비용 발생 가능	장면 분할 모델의 적용 범위를 상위 장르를 기준으로 한정하는 것과 같은 방법을 통해 필요 데이터의 수를 줄여 비용 절감 가능.

5. 국내외 시장 동향

• 국내외 시장 전망

- 국내 시장의 경우 인터넷 포털 회사를 중심으로 사용자가 생산한 콘텐츠를 공유하고 재생산 할 수 있는 플랫폼 개발에 목표를 두고 플랫폼 기반 서비스의 연구가 활발히 진행되어지고 있음
- 현재 높은 기술 수준을 요구하는 동영상 처리 보다는, 이전 단계인 텍스트, 오디오, 이미지 중심의 서비스 중심으로 기술화가 진행되고 있음. 카카오의 경우, 카카오톡과 포털의 결합 형태인 '모바일 라이프 플랫폼'구축과 폐쇄형 동영상 공유 서비스 '잼(zap)'등의 다양한 형태로 콘텐츠 재생산에 집 중하고 있음

• 국내외 시장 규모

- 관련 제품/서비스의 국내외 시장 예상 규모(향후 5년) (단위:백만불, 억원)

관련 제품/서비스	시장	1차년도 (2020)	2차년도 (2021)	3차년도 (2022)	4차년도 (2023)	5차년도 (2024)	합계
장면 기반 광고 /커머스 서비스	해외	200	300	800	1,000	1,500	3,800
	국내	100	300	500	700	1,000	2,600

- 예상 제품/서비스의 예상 매출액(생산/판매부터 향후 5년간 추) (단위:백만불, 억원)

관련 제품/서비스	시장	1차년도 (2020)	2차년도 (2021)	3차년도 (2022)	4차년도 (2023)	5차년도 (2024)	합계
영화 장면 클립 검색 서비스	해외	-	1.5	8	20	45	247
	국내	0.5	3	15	28	60	377

감사합니다.



www.etri.re.kr

♣ 연락처 : 방송·미디어연구소, 김선중 (042-860-5402, kimsj@etri.re.kr)