

테크놀로지 리더를 위한

Media & IT(Information Technology)

#10. 인공지능과 미디어

강자원

컴퓨터시스템응용기술사

/ KBS MNC(Media Network Center)팀



연재 목차

- | | | |
|-----------------|------------------|----------------------|
| 1. 소프트웨어공학과 미디어 | 5. 소프트웨어 개발과 미디어 | 9. 클라우드와 미디어 |
| 2. 네트워크와 미디어 | 6. 애플리케이션과 미디어 | 10. 인공지능과 미디어 |
| 3. 보안과 미디어 | 7. 시스템 아키텍처와 미디어 | 11. 블록체인과 미디어 |
| 4. 데이터와 미디어 | 8. IT 운영과 미디어 | 12. 가상현실과 미디어 |

매년 새해가 되면 그해의 트렌드를 분석하는 책들이 서점에 배치된다. 그중에 미디어 트렌드 관련 책, ‘디지털미디어 인사이트 2024’의 부제는 ‘AI 시대의 생존전략, 미디어 패러다임을 바꿔라!’이다. 시작부터 책 이야기를 하는 이유는 우리 일상으로 스며든 생성형 인공지능(AI)에 대한 이야기를 하고 싶어서다. 부제에서 이미 드러났듯이 AI는 일상에 자연스럽게 스며들었고, 이제는 미디어 패러다임을 바꿀 만큼 더 가까이에 있다. 우리에게 생성형 인공지능(AI)에 대한 키워드가 낯설지 않은 것은 작년부터 방송, 신문, SNS를 통해 많이 들어봤기 때문일 것이다. AI는 등장부터 많은 위기와 기회의 아이콘으로 대두되면서 사람들의 입에 오르내리고 있다. 필자가 속한 직원들 사이에서도 “AI, 모르면 뒤쳐질 것 같아 불안해” 또 다른 누군가는 말한다. “AI, 써보니까 편하던데?” 지금 주변인들의 AI에 대한 반응은 대개 이 둘로 나뉜다. 이제 AI를 누구나 쉽게, 그리고 특별한 기술 없이 사용할 수 있다. 더 이상 사람들은 AI를 쓸지 말지 고민하지 않는다. 쓰지 않을 수 없다는 결론이 이미 나왔기 때문이다. 단지 어떻게 잘 배워서 어떻게 잘 활용할지 고민 중이다. 그리고 그것 때문에 불안해한다. 시대의 흐름에 뒤쳐질까봐. 이 걱정은 사용자 관점에서 또, 시스템을 기획하고 운영하는 엔지니어 관점에서도 마찬가지다. 생성형 인공지능(AI)과 함께 바뀌게 될 미디어 환경과 이를 맞이하는 우리의 자세에 대해 이야기 해 보고자 한다.

미디어에서 말하는 생성형 인공지능(AI)이란 무엇인가?

텍스트, 오디오, 이미지 등 기존 콘텐츠를 활용해 유사한 콘텐츠를 새롭게 만들어 내는 인공지능(AI) 기술을 가리키며 생성형 인공지능(AI)이라고 한다. 기존 AI가 데이터와 패턴을 학습해서 대상을 이해했다면 생성형 인공지능(AI)은 기존 데이터와 비교 학습을 통해 새로운 창작물을 탄생시킨다. 즉 콘텐츠의 생성자와 만들어진 콘텐츠를 평가하는 판별자가 끊임없이 서로 대립하고 경쟁하면서 새로운 콘텐츠를 만들어 낸다. 이를 통해 현실에 있을 법한 새로운 콘텐츠가 탄생하게 된다. 이미지 분야에서는 특정 작가의 화풍을 모사한 그림으로 사진을 재생성하거나 가짜 인간 얼굴을 무제한으로 생성할 수 있다. 음성 분야에서는 특정 장르의 음악을 작곡하거나 특정 노래를 원하는 가수의 음색으로 재생성할 수 있다. 텍스트 분야에서는 특정 소재로 시를 짓거나 소설을 창작할 수 있다. 이 같은 결과물은 쇼핑, 영화, 광고 등 산업에서 활용된다. 최근에는 글로 명령하면 이를 이미지나 비디오로 변환시키는 생성형 인공지능(AI)이 대두되고 있다. 메타는 문장을 입력하면 비디오로 만들어 주는 ‘Make a video’ 서비스를 선보였다. 구글도 텍스트를 동영상화하고 영상 콘텐츠를 생성할 수 있는 AI 비디오 생성기 ‘Imagen video’를 공개했다.

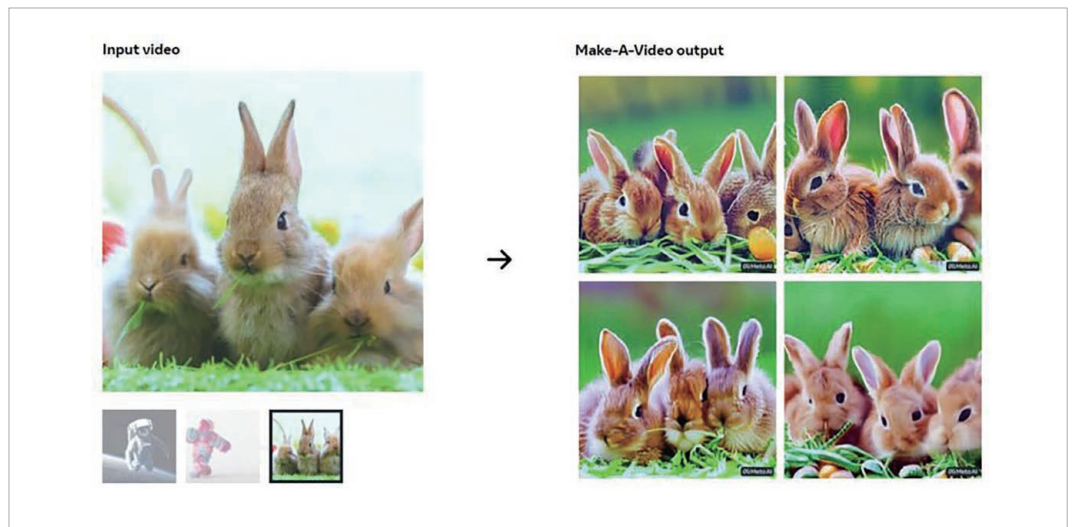


그림 1. 메타의 Make a Video를 통해 생성된 애니메이션 비디오

용어설명

메타(Meta) : 전 세계 최대의 소셜네트워크서비스(SNS)로, 2004년 2월 마크 저커버그가 ‘페이스북’이라는 명칭으로 출범시켰다. 그러다 창업 17년 만인 2021년 10월 회사 이름을 메타로 변경했다.

비즈니스 DNA의 일부가 된 인공지능(AI) 기술의 현주소

2차 산업처럼 어떤 물건 형태로 제품이 생산되는 게 아니라 정보나 콘텐츠 같은 추상적인 가치로 미디어 산업을 꾸려왔으니 이 부분은 기계가 대체할 수 없다고 장담해 왔는데 그 믿음이 산산이 부서지고 있는 형국이다. 생성형 AI라는 괴물은 인간의 전유물로 여겨져 왔던 ‘창작’의 부분까지 잠식해 들어오고 있다. 소셜도, 웹툰도, 하다못해 영화도 인공지능이 대신 만들어 버리는 세상이다. 지난 6월 미국에서 발생한 ‘넷플릭스·AI 파업’(미국작가조합 파업)은 인류가 역사적으로 AI로부터 일자리 위협을 느끼는 첫 사

레라 볼 수 있다. 미국 작가조합(WGA)은 제작사들이 AI를 활용해 이전에 작가들이 쓴 시나리오에서 새로운 스크립트를 생성하거나, AI가 만든 대본 초안을 작가들에게 손보라고 요구해서는 안 된다고 주장하고 있다. 스튜디오는 AI 사용을 할 수밖에 없다는 입장이었고 결국, 작가들은 자신들의 작품이 AI 학습 자료로 사용되는 것을 허용하는 대신 대본 작성에 AI가 개입하더라도 급여 삭감은 하지 않는다는 내용으로 ‘AI와의 공존’을 받아들이기로 했다.



그림 2. 넷플릭스 스튜디오 밖에서 파업 중인 미국작가조합(WGA)과 배우조합(SAG-AFTRA) 회원들의 피켓시위



그림 3. ZDNet Korea 2023.05.21.일자 기사 일부

작가들은 영화나 드라마, 방송 대본 등 콘텐츠 제작에 AI가 도입되면 작가들의 일자리를 상당 부분 뺏기게 될 것으로 우려해 왔다. WSJ에 따르면 할리우드 제작사들은 대본 요약부터 특수 효과, 홍보·마케팅에 이르기까지 모든 분야에서 AI 사용을 검토하고 있다. 조만간 우리의 제작 환경에도 다가올 현상으로 보인다.

넷플릭스의 이야기가 조금 멀게 느껴졌거나, 먼 나라 미국에서의 일이라 크게 와닿지 않는다면, 조금은 더 가까운 우리의 이야기를 해보자. BTS가 소속되어 있는 엔터테인먼트 기획사, 하이브가 신인가수 ‘미드나잇’을 공개했는데, 미드나잇의 음원에 AI 기술을 활용한 다국어 발음 교정 기술과 보이스 디자인 기술을 활용했다. 6개의 언어(한국어, 영어, 일본어, 중국어, 스페인어, 베트남어)로 가창한 노래 음원에 원어만 발음 데이터를 접목해 발음을 자연스럽게 교정하고, 아티스트 목소리에 AI 음성을 삽입했다. 보이스 디자인 기술은 남성의 목소리를 여성의 목소리로 바꿔줄 수도 있다. 이제 기술은 사용자에게 다양한 경험을 제공해 줄 수 있기에 충분히 발전했고, 창작의 영역은 더 이상 인간만이 할 수 있는 범위를 넘어섰다.

미디어 콘텐츠 제작 현장에서의 인공지능(AI)

☒ AI 기자가 작성한 뉴스 원고

포털에서 뉴스를 검색하면, ‘이 기사는 AI가 작성하고 OOO이 편집했습니다.’라고 적혀있는 기사들을 종종 볼 수 있을 것이다. 미국의 몇몇 언론사는 AI 기사를 도입했다. 초기엔 기사의 사실 확인 문제가 논란의 대상이 되었지만, 이제는 그 활용 방법의 현실적인 적정선을 찾은 것 같다. 베테랑 기자들은 AI가 작성한 문장을 부분적으로 사용하고, 개요 작성과 오타 수정에 AI를 활용하며, 기사에 유리한 제목이나 인터뷰 질문 목록을 만든다. AI와의 공존 방법을 찾은 것이다.

☒ AI 휴먼 생성 플랫폼을 통한 다양한 분야의 콘텐츠 제작

생성 인공지능(AI) 전문 씨앤에이아이는 AI 휴먼 생성 플랫폼 ‘폴라’를 개발했다. 텍스트(프롬프트)가 묘사한 이미지를 정확히 생성하는 Text2Image 기능과 복수의 기존 이미지를 결합해 새로운 이미지를 만드는 Image2Image 기능을 제공한다. 또한 기존 이미지를 바탕으로 사용자가 원하는 방식으로 맞춤형 이

미지 편집도 가능하다. Few Short Learning 기술이 탑재돼 있어, 적은 레퍼런스 데이터만으로도 다양한 콘텐츠를 생산해내며, AI 합성데이터 기술을 접목해 쉽고 빠르게 사용자가 원하는 프로필(성별, 나이, 키, 체중 등)을 설정해 나만의 가상인간을 만들 수도 있다. 가상인간 생성을 위한 파라미터 조정으로 언어나 목소리는 물론 표정, 동작 등 초현실, 초개인화 가상인간 생성이 가능하다.



그림 4. MBN의 AI 딥휴먼 기술로 탄생한 AI 가상전문기자 리나, 엘라, 진호, 태빈

언론사 MBN도 AI 가상 기자를 이미 2022년도에 시청자들에게 선보인 바 있다. 이들 네 명은 MBN 취재 기자들을 모델로 제작됐지만 고향, 나이, 취미, 전공, MBTI까지 각자 자기만의 정체성을 갖고 있다. 가상 기자가 단순히 뉴스 전달만 하는 것이 아니라 방송국의 로고송이나 캐릭터처럼 MBN 뉴스의 브랜드 정체성을 찾기 위한 시도였다 한다. 매우 참신하면서도 도전적인 시도다. 이렇게 누군가는 발 빠르게 인공지능과 공존하는 방법들을 스스로 찾고 시도하고 있다.

☒ 콘텐츠 메타데이터 태깅 강화를 통한 제작 생산성 향상

우리는 내부적으로 아카이브를 구축해 필요한 콘텐츠 자료를 검색한다. 이런 미디어 콘텐츠 저장, 검색관리에서 인공지능(AI)에 요구되는 주요 기능은 콘텐츠 평가 또는 인터뷰 음성의 텍스트 변환, 얼굴 인식, 감정 분석 및 장소 인식과 같이, 기존 콘텐츠에서 더 많은 콘텍스트를 제공하기 위한 자동 메타데이터의 강화이다. 인공지능을 통해 이런 반복적이고 복잡한 작업을 자동화하면 제작팀이 태그 지정과 같은 지루한 작업이 아닌 콘텐츠 가치 창출에 더욱더 집중할 수 있게 된다. 따라서 인공지능은 가치 사슬의 자동화 측면에서 계속 성장하고 있다. 또한 메타데이터 강화뿐만 아니라 오류 예측, 사용자 통계 및 관리 등의 측면에서도 도입하고 있다. 이는 인공지능 기술을 아주 효율적으로 활용하고 있는 예이다. 사람이 하는 단순 반복적이지만 사고가 필요한 일들을 기계가 대신하게 하면서 생산성을 향상할 수 있는 머신러닝 기술의 목적에 부합하는 사례가 된다.

용어설명

콘텍스트(Context) : 사물의 서로 잇닿아 있는 관계나 연관

메타데이터(Metadata) : 데이터를 효율적으로 이용하기 위해 구조화하는 데이터의 정보

☒ 고품질 스트리밍을 위한 인코딩 최적화

현재 폭발적으로 이용자가 늘고 있는 넷플릭스, 유튜브 및 디즈니 플러스 서비스들은 인공지능을 이용하여 대용량 동영상의 크기를 압축하고 있다. 동영상을 한 번에 전송하지 않고 여러 개로 작게 나눈 다음 대

역폭을 조정하여 사용자들의 컴퓨터 환경에 맞게 지능적으로 영상을 보여줄 수 있게 되었다. 하지만 무선 네트워크가 없는 지역에서 기존의 LTE보다 좀 더 빠른 5G가 보편적으로 사용될지라도 영상 스트리밍 대역폭을 조정하는 것은 기존 기술로는 쉬운 일이 아니다. 인프라의 처리 성능을 향상하거나 트랜스코딩 알고리즘을 최적화하는 방안으로 해결 방법을 모색해 볼 수 있을 것 같은데 이 또한 비용의 문제와 마주하게 되면 기술적으로도 가능하더라도 실현 불가능할 수 있다. 즉, 사용자가 부담하는 데이터 비용은 가능한 줄이면서, 콘텐츠의 영상 품질은 높게 유지하는 것. 최종적으로, 사용자의 좋은 경험을 유지하면서 비트레이트 개선을 위한 인코딩을 최적화할 수 있도록 인공지능(AI) 기술을 적용하는 것이 목표다.



그림 5. 구간별 동영상의 특징을 반영한 비트레이트 설정

영상을 구간별로 잘라서 모든 CRF 레벨과 모든 Maxrate 수치의 조합으로 인코딩을 실시한다. 그리고 품질을 측정한다. 정적인 화질을 유지하면서 사이즈를 가장 작게 할 수 있는 파라미터를 고른다. 영상의 특징과 효율적인 파라미터를 매칭하고, 딥러닝을 이용하여 학습시킨다.

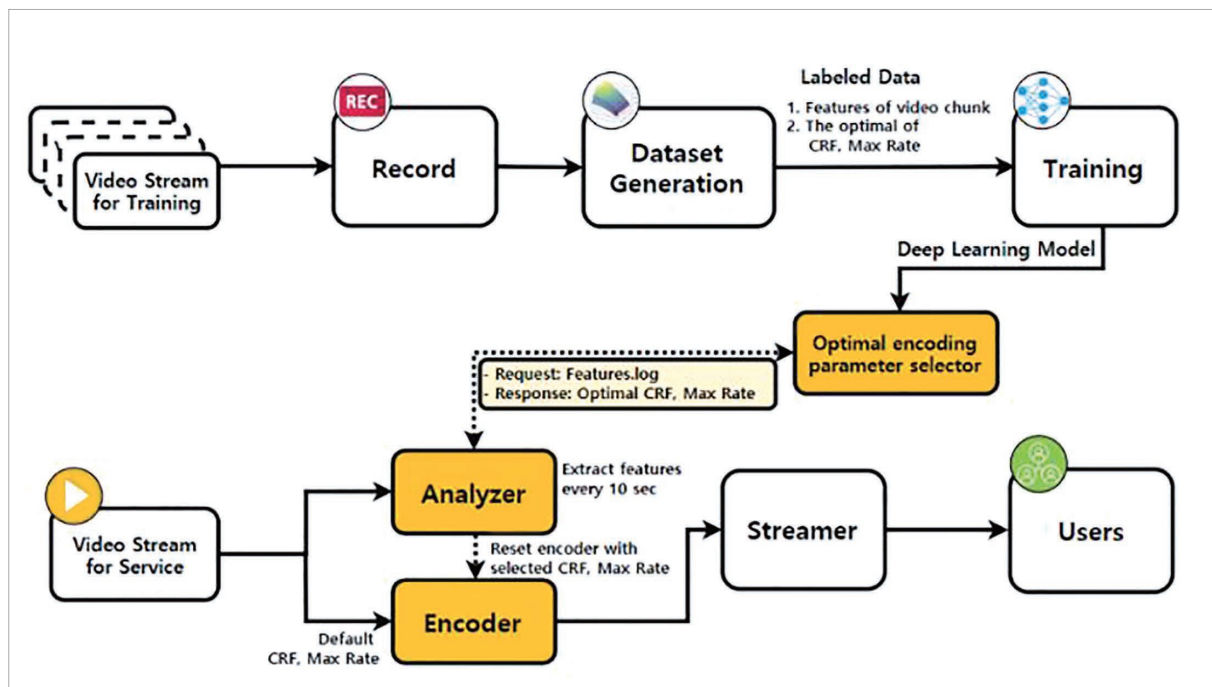


그림 6. 딥러닝을 이용한 인코딩 시스템 아키텍처

☑ 영상 이미지 처리 및 색보정 등의 후반작업 영상처리 기술 고도화

어도비에서 개발한 AI 프레임워크인, ‘센세이 GenAI’와 같은 도구는 크리에이티브가 비디오 제작 과정의 속도를 높이는 데 도움을 줄 수 있다. 신속하게 장면을 일치시키는 기능은 컬러 그레이딩 작업의 고통을 덜어주어 컬러리스트가 대조, 색조, 채도, 블랙 레이블 및 화이트 포인트와 같은 이미지의 다양한 속성을 동영상, 비디오 또는 정지 이미지에 반영하는 처리 속도를 향상할 수 있다. 어도비 사는 생성형 AI 기능을 어도비 플랫폼에 탑재해 기업에 전례 없는 속도와 생산성을 제공한다는 구상을 하고 있다.

인공지능(AI) 시대에 방송사의 생존전략

이렇게 다양한 미디어 산업에 인공지능(AI)이 적용되고 있고, 각각의 영역에서 어떻게 잘 활용할 수 있을 것인지에 대한 방법들을 모색하고 있다. 그러나 각각의 낱겉과 몸짓은 거시적 관점에서 볼 때, 좋지 않은 결과를 만들어 내는 방향으로 진행되기도 한다. 인공지능(AI)이 소셜미디어와 포털 등 플랫폼 종류를 가리지 않고 대격변을 일으키고 있는 가운데 플랫폼 성격이 서로 비슷해지는 ‘동질화(homogenation)’와 콘텐츠 품질이 떨어지는 ‘엔시트화(enshittification)’가 이뤄진다고 전문가들은 말한다. 엔시트화란, 결국 생산자들은 자극적인 제목으로 승부 걸지 않으면 이용자를 만날 수 없는 알고리즘 구조로 되는 것을 말한다. 이제는 너도나도 AI를 사용하고 서로가 서로를 학습하다 보면 결국 비슷해지게 되고 다 같이 품질이 평균 하향화가 되는 현상이 일어날 수도 있다. 이럴 때 전통적인 미디어는 어떤 기준을 세워야 발전적인 방향으로 일어설 수 있을까?

바로 신뢰도를 향상하는 일일 것이다. AI 기자, 이미지 생성 등 다양한 방식으로 콘텐츠를 생산해 낼 수 있지만 자칫 잘못하면 정크정보, 스팸을 생산해 내는 방향으로 미디어는 변모할 수도 있다. 이럴 때일수록 방송이 언론사로서의 경쟁력을 강화하기 위해서는 시청자에게 신뢰를 지키면서 미디어 콘텐츠 소비자들이 필요한 새로운 기능과 콘텐츠 전달에 대한 더 효율적인 운영, 콘텐츠 제작 관점에서 향상된 창의성을 제공해야 하며, 기존의 기술로는 달성하지 못한 것을 달성할 수 있도록 해야 한다.

그렇다면, 엔지니어는 인공지능(AI) 시대를 어떻게 준비해야 할까? 물론 생산성을 향상하는 기술로 사용하는 것은 가장 바람직할 것이며, 기술적인 관점뿐 아니라 시야와 관점을 좀 더 넓혀 콘텐츠를 창조적으로 생산하는 생산자의 관점에서 인공지능(AI)이 학습데이터에 대한 올바른 데이터 셋을 가질 수 있도록 프로세스를 만들고 시스템을 구축하고 준비해야 할 것이다. 결국, 인공지능(AI)은 데이터를 학습해서 결과를 만들어 낸다. 그런데, 학습하는 데이터들이 올바르지 않다면 의미 있는 생산활동을 하지 못할 것이다. 인공지능(AI)을 활용하게 될 전통 미디어인 방송에서 미디어의 신뢰성은 결국 AI를 학습시키는 데 사용될 신뢰할 수 있는 데이터 셋에 기인한다.

따라서 방송사의 역할은 그리고 우리 엔지니어의 역할은 품질 좋은 데이터 셋을 유지하고 축적하는 것이며, 공정하고 신뢰할 수 있는 알고리즘을 찾아 신뢰할 수 있는 방송·미디어로 진화해야 할 것이다. 다들 위기라고 말할 때, 그 안에서 기회를 찾을 방법이 무엇인지 절실히 고민해 봐야 할 시기다. 