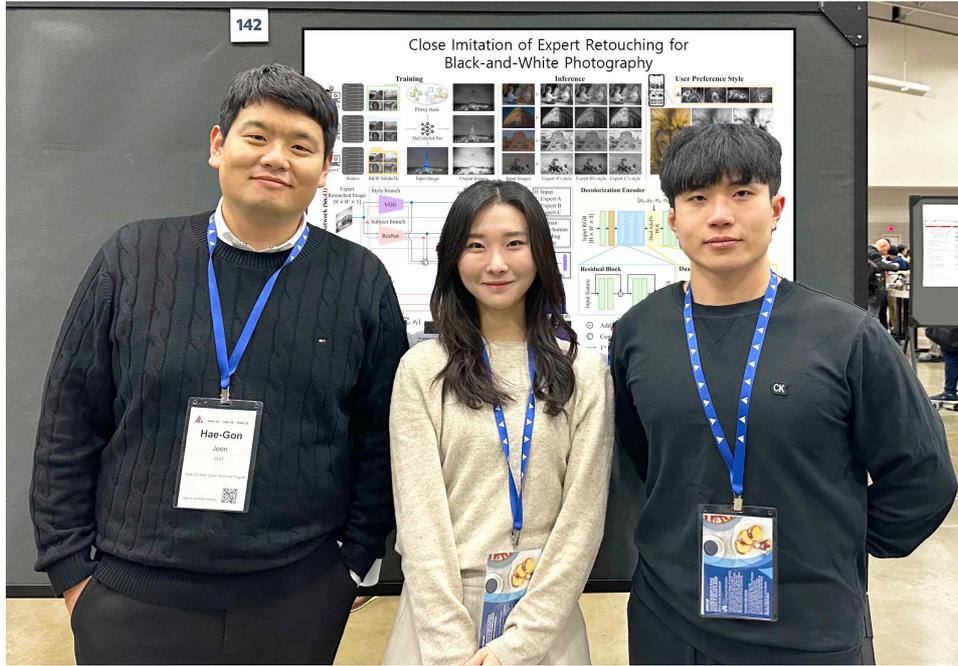


“AI가 변환한 흑백 사진·영화 너무 멋진데” GIST, 사진·영상을 전문가 수준으로 컬러→흑백 보정하는 AI 기술 개발

- AI대학원 전해곤 교수팀, 국내 저명 사진작가의 포트폴리오로부터 보정 기법·스타일 모사하는 알고리즘 개발.. 전문가 영역의 흑백 사진 누구나 생성, 영상에도 적용 가능
- 영화 '기생충' 오리지널 흑백 개봉판보다 연구성과 적용한 컬러→흑백 재생성 결과물 선호도가 더 높아.. 세계 최고 비전 학회「CVPR」에서 오는 6월 발표 예정



▲ (왼쪽부터) AI대학원 전해곤 교수, 신지수 석박통합과정생, 신승현 석박통합과정생

2011년 풀리처상을 수상한 바바라 데이비드슨(Barbara Davidson)은 “컬러 사진은 간혹 색감에 의해 작가의 의도가 훼손되는 경우가 있지만, 흑백 사진은 감정과 연출 의도를 더 집중하여 피사체에 담아 표현할 수 있다”며 흑백 사진의 장점을 언급한 바 있다.

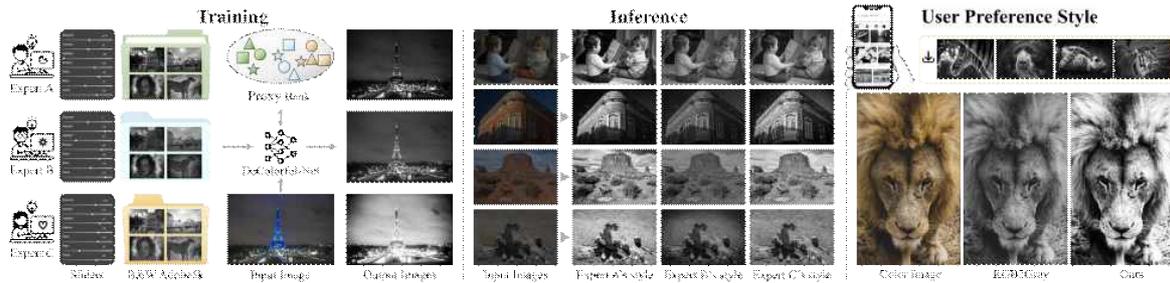
광주과학기술원(GIST, 총장 임기철)은 AI대학원 전해곤 교수 연구팀이 인공지능(AI) 기술을 활용해 흑백 사진작가들의 포트폴리오로부터 흑백 사진의 보정 기법을 모사하는 알고리즘을 개발했다고 밝혔다.

흑백 사진은 컬러 사진과 달리, 색으로부터 벗어나 질감, 선, 패턴, 대비 등에 초점을 맞춰 사진을 표현할 수 있다. 이 같은 특징을 살려 최근 인스타그램 등 소셜 네트워크 서비스(SNS)에서 심미적 효과를 연출하는 데 널리 활용되면서 흑백 사진이 새롭게 주목받고 있다.

그동안 컴퓨터 비전 분야의 연구에서 컬러→흑백 영상 변환은 심미적 완성도를 높이기 위한 것보다는 단순히 다른 과업을 수행하기 위한 전처리(preprocessing) 단계로서, 또한 변환 과정에서 정보를 잃지 않는 것을 목표로 다루어졌다.

심미적 목적의 흑백 사진을 얻기 위해서는 일반 카메라보다 훨씬 값비싼 흑백 사진 전용 카메라를 구입하거나 **전문가에게 보정을 의뢰**해야 한다. 따라서 일반적인 컬러 사진보다 더 많은 시간과 비용이 드는 작가 수준의 흑백 사진은 실용성 측면에서 아쉬운 점이 있었다.

이점에 착안해 연구팀은 **컬러 사진으로부터 전문 사진사의 미학적 요소를 매우 정밀하게 재현하는 흑백 사진 생성 알고리즘을 개발**하였다.



▲ **본 연구의 전체적인 파이프라인.** 전문가들로부터 사진 데이터셋을 구축하여 이들을 근접하게 모사하도록 학습한 후, 인터넷에서 다운로드한 어떤 스타일의 사진이든 흑백 사진으로 변환할 수 있다.

연구팀은 학습 데이터를 구축하기 위해 **3명의 국내 전문 사진작가를 섭외**해 **그들 고유의 스타일*로 컬러 사진을 흑백으로 변환하는 작업을 의뢰**하여 **작가 특유의 개성이 담긴 흑백 사진 데이터셋(작가별 5,000장, 총 15,000장)**을 구축했다.

연구팀은 이를 바탕으로 사진 데이터셋을 각각 어느 사진작가가 보정했는지와 작가별 사진 중에서도 대상 피사체가 무엇인지에 따른 **계층화 분류 작업을 추가로 수행**하였다.

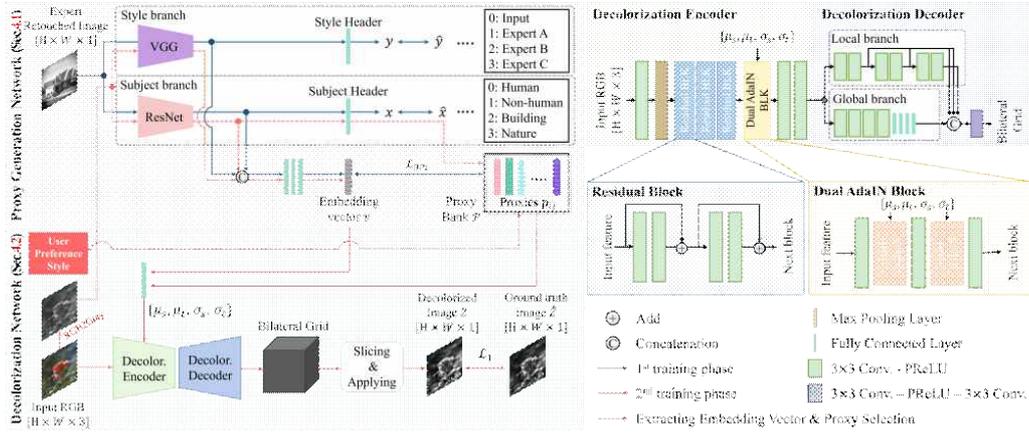
* **스타일(style):** 본 자료에서 사용되는 스타일이란 작가가 사진을 보정할 때 주관적으로 선호하는 결과물의 밝기, 대비, 추가 효과 등을 지칭한다.

이후 연구팀은 **딥메트릭러닝(Deep Metric Learning) 기반의 인공지능망을 적용하여 심미적으로 뛰어난 흑백 사진을 생성**하였다. 알고리즘은 사진작가의 포트폴리오로부터 **작가 고유의 스타일 특징을 추출하는 '제1신경망'**과 실제 전문가의 보정 방법에 매우 근접한 수준의 리터칭(retouching)을 **미학적 요소로 재현하는 '제2신경망'**으로 구성된다.

'제1신경망'은 각 계층을 대표하는 벡터인 **프록시***를 생성한다. 연구팀은 피사체별 세부 군집을 형성하는 프록시를 학습하도록 했다. 학습된 '제1신경망'은 사진작가 고유의 스타일 차이를 계층적 구조로 분류하고, 세부 군집은 피사체별 보정 스타일을 반영하는 벡터로서 활용하도록 알고리즘을 설계했다.

* **프록시(proxy):** 임베딩 공간 상에서 같은 군집으로 속하는 값들을 대표하는 벡터

* **임베딩 공간(embedding space):** 딥러닝 신경망에서 입력 영상이 주어졌을 때 신경망을 통과하던 중간 값을 나타내는 저차원 공간을 지칭한다.



▲ 본 연구에서 제안한 흑백 사진 생성 네트워크 구조. 사진 더미들로부터 스타일을 벡터로 추출하는 제1신경망과 추출한 벡터의 미학적 요소를 재현하여 흑백 사진을 만들어 내는 제2신경망으로 나뉜다.

‘제1신경망’은 각 계층을 대표하는 벡터인 **프록시***를 생성한다. 연구팀은 피사체별 세부 군집을 형성하는 프록시를 학습하도록 했다. 학습된 ‘제1신경망’은 사진작가 고유의 스타일 차이를 계층적 구조로 분류하고, 세부 군집은 피사체별 보정 스타일을 반영하는 벡터로서 활용하도록 알고리즘을 설계했다.

* **프록시(proxy)**: 임베딩 공간 상에서 같은 군집으로 속하는 값들을 대표하는 벡터

* **임베딩 공간(embedding space)**: 딥러닝 신경망에서 입력 영상이 주어졌을 때 신경망을 통과하던 중간 값을 나타내는 저차원 공간을 지칭한다.

다음으로 이전 과정에서 정의된 프록시가 입력 영상과 함께 ‘제2 신경망’에 제공되었을 때 **각 프록시가 내포하고 있는 스타일에 맞는 흑백 영상을 추출했다**. 이때 사진 보정이 사진 전체에 동일하게 적용되는 것뿐만 아니라 **각 영역별로도 다르게 진행되는 문제를 해결하기 위해 사진을 의미론적 부분으로 나누어 영역별 보정이 가능하도록 제2신경망을 설계하였다**.

이렇게 학습된 알고리즘은 사용자가 개인적으로 선호하는 사진들에 맞춰 **‘제1신경망’을 미세조정*** 할 수 있다. 즉, 제안한 데이터셋으로 학습한 네트워크의 파라미터를 기반으로 **새로운 사진작가의 스타일을 모사하는 것이 가능하다**.

* **미세조정(fine-tuning)**: 한 번 학습된 신경망을 추가적으로 학습하여 신경망의 변수값을 미세하게 조정하는 방법이다.

연구팀이 개발한 알고리즘을 통해 얻은 흑백 사진과 갤럭시/아이폰/인스타그램 필터들과 일반인 선호도를 조사한 결과, 대부분의 사람들이 연구팀에서 **제안한 알고리즘의 결과를 선호하는 것으로 확인됐다**.



▲ 입력 영상(a)를 기본 흑백변환 알고리즘을 활용한 결과물(b), 사진 전체에 동일한 보정 효과를 진행한 결과물(c), 영역별로 다른 보정 효과를 줄 수 있는 우리 알고리즘의 결과물(d)의 비교 그림. 본 알고리즘은 (e)에서 강조한 바와 같이 영역별로 다른 보정을 진행하여 심미적으로 더 아름다운 흑백 영상을 생성할 수 있다.

설문조사는 20~50대 남녀를 나이와 성별에 따라 10명씩 총 80명을 대상으로, 설문조사 플랫폼 'Amazon M-Turk'를 활용하여 진행되었으며, 20개의 사진을 각 필터로 변환하여 보여 주고 사진별로 맘에 드는 필터에 순위를 매기는 방식으로 진행한 결과, 연구팀에서 개발한 알고리즘이 가장 높았다. (최고점 4.58점 획득, 차점 iPhone14Pro 4.12점)

나아가 영화 '기생충'과 '매드맥스'를 대상으로 ▲ 두 영화의 원본 흑백판과 ▲ 연구팀의 알고리즘으로 두 영화의 컬러판을 재생성한 흑백 영상에 대한 선호도를 비교한 결과, 마찬가지로 연구팀이 개발한 알고리즘에 따른 결과물의 선호도가 압도적으로 높게 나왔다.

영화 영상비교 실험은 총 10개의 영상으로 진행되었으며, 흑백판 재개봉 영화와 연구팀이 개발한 알고리즘으로 생성된 영상이 더 나은지를 ▲ 매우 그렇다(5점)에서 ▲ 매우 아니다(-5점)로 점수를 부여하게 한 결과, 연구팀의 알고리즘으로 생산된 영상이 평균적으로 2.36점을 받아 원본 흑백영화 대비 훨씬 더 선호되는 결과를 확인할 수 있었다.



▲ 본 연구 알고리즘과 기존 사진 보정 어플리케이션의 필터들과의 비교 사진. 컬러 이미지(a)가 주어졌을 때 각각 아이폰(b), 갤럭시(c), 인스타그램(d), 틱톡(e), 포토샵(f), 사진가의 카메라(g)에서 지원하고 있는 각 흑백 필터를 적용한 결과와 본 연구의 알고리즘을 통해 보정된 결과(h)를 나타낸다.

전해곤 교수는 "이번 연구 성과를 통해 기존 사진에 대한 높은 배경지식과 고가의 카메라를 사용해야 얻을 수 있던 미학적으로 가치가 높은 흑백 사진을 우리 알고리즘을 통해 일반인도 쉽게 얻는 것이 가능해졌다"며 "향후 사진 보정 어플리케이션 및 미디어 산업계의 영상 후처리 과정 등에서 활용될 것으로 기대된다"고 말했다.

GIST AI대학원 전해곤 교수 연구실의 신승현 석박통합과정생 주도하고 인하대학교 심인욱 교수가 참여한 이번 연구는 과학기술정보통신부의 'AI혁신 허브'과제 및 '지역의 미래를 여는 과학기술 프로젝트'사업과 과학기술정보통신부(연구개발특구)와 광주광역시의 지원으로 GIST가 주관하는 '인공지능 기반 메타버스 구현을 위한 융·복합 문화 가상 스튜디오' 과제의 지원을 받았으며, 컴퓨터 비전 분야 세계 최고 학회인 'CVPR(Computer Vision and Pattern Recognition)'에서 2024년 6월 19일 발표될 예정이다.

논문의 주요 정보

1. 논문명, 저자정보

- 저널명 : The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2024 (CVPR) · 한국정보과학회 및 BK21+ 기준 최우수 학술대회
- 논문명 : Close Imitation of Expert Retouching for Black-and-White Photography
- 저자 정보 : 신승현(제1저자, GIST AI대학원), 신지수(제2저자, GIST AI대학원), 배지환(제3저자, 차의과학대학교), 심인욱(인하대학교), 전해곤(교신저자, GIST AI대학원)