

## 딥러닝을 활용한 감성 증명사진 제작 웹 애플리케이션

\*김도영, \*\*강인영, \*\*\*김연수, \*\*\*\*박구만  
서울과학기술대학교  
\*전기정보공학과 \*\*ITM 전공 \*\*\*컴퓨터공학과  
\*\*\*\*전자 IT 미디어 공학과, 교신저자  
\*kacel33@naver.com, \*\*kiyoog02@gmail.com

## Web Application for Creating Emotional ID Photos using Deep Learning

\*Do Young Kim, \*\*In Yeong Kang, \*\*\*Yeon Su Kim  
\*\*\*\*Goo man Park  
Seoul National University of Science and Technology  
\*Dept. of Electrical and Information Engineering  
\*\*Dept. of Information Technology and Management  
\*\*\*Dept. of Computer Science and Engineering  
\*\*\*\*Dept. of Electronics and IT Media Engineering

### 요 약

최근 본인에게 어울리는 색상을 배경으로 촬영하는 감성 증명사진이 유행하고 있다. 개인마다 퍼스널 컬러를 찾아 배경색에 적용하는 것은 시간, 비용, 인력적으로 어려움이 있으므로 자동으로 개인에 따른 배경색을 찾아서 사진을 합성하여 감성 증명사진을 제작해 주는 딥러닝 기반 시스템을 구축하였다. 본 논문에서는 Convolution Neural Network 를 기반으로 한 딥러닝 기술을 이용해 Image Matting 과 Multi-Label Classification 을 수행하여 기존 감성 증명사진들을 학습하여 모델을 구축하였으며, 해당 시스템으로 사용자에게 새로운 배경색이 적용된 감성 증명사진을 제공하는 웹 애플리케이션을 제안한다.

### 1. 서론

인물 사진을 촬영을 할 때에 인물 뒤에 배경과 함께 촬영되는데 촬영자의 의도에 따라 배경이 선택되기도 하고 추후 편집을 통해서 사진의 배경을 수정하기도 한다. 최근 '퍼스널 컬러'라는 표현과 함께 자신에게 가장 어울리는 색상, 질감 등을 찾고 화장품이나 옷 등에 적용하는 트렌드가 있다. '감성기록관', '시현하다'와

같은 사진관에서는 실제 본인의 선호 색상 또는 퍼스널 컬러를 배경으로 한 증명사진들을 촬영하여 기록으로 남기는 것이 유행하고 있다. 하지만 이러한 사진관에서의 촬영 또는 편집을 위해서는 시간과 비용 등이 많이 소요된다. 이에 주목하여 본 논문에서는 사용자의 기존 사진을 감성 증명사진으로 제작해 주는 딥러닝 시스템과 해당 시스템을 적용한 웹 애플리케이션을 제안한다. SNS 및

구글링을 통해 감성 증명사진들을 수집하였고, 이를 학습하여 사진에 따른 배경 색상 추론 모델을 구축하였다. 즉, 사용자가 입력하면 모델을 통해 사진의 배경색이 추론되고 이를 활용하여 사진을 새롭게 합성할 수 있다. 본 논문의 감성 증명사진 생성 시스템을 통해 사진사나 편집가의 손을 거치지 않고도 본인에게 어울리는 배경색이 적용된 감성 증명사진을 얻을 수 있다. 더하여 딥러닝을 활용한 취업 사진 생성 앱[1]과 같이 사진 촬영 및 편집과 같은 다수의 서비스는 앱으로 구현된다. 앱 다운로드에 대한 번거로움을 줄여 접근성을 높이고, 사람들이 쉽게 사진을 입력하여 빠르게 합성된 사진을 얻을 수 있도록 하기 위해 최종적으로 웹 애플리케이션의 형태로 감성 증명사진 생성기를 구현하였다.



(그림 1) 감성 증명사진의 예<sup>1</sup> (좌: 김해준, 우: 로지)

## 2. 관련 연구

### 2-1 Image Matting

Image Matting 을 통해 사진 속에서 관심 있는 물체를 추출할 수 있다. 전경 또는 배경으로 나뉘는 이진 이미지를 생성하는 Image Segmentation 과 달리 Image Matting 은 경계가 모호한 unknown 영역을 가지는 Trimap 에서 정확한 알파 값을 추정하여 전경과 배경 분리의 정밀도가 달라진다. 최근 Image Matting 을 수행하는 네트워크들 중에서 P3M-10k 데이터 셋을 활용한 P3M-Net 은 portrait matting 을 위해 학습되어 뛰어난 성능을 보여준다[3]. 사전 훈련된 Image Net 모델 ResNet-34 를 활용하였으며, 더 나은 trimap-free portrait matting 을 위해서 의미적인 인식과 상세한 매칭을 위한 통합된 프레임워크를 적용했다. 본 논문의 인물 증명사진을 생성이라는 목적과 함께 정밀한 사진을 위해 SAD, MSE, MAD 수치가 가장 높은 순위를 달성하고 있는 P3M-Net 을 사용하였다.

### 2-2. Multi-Label-Classification

Classification 은 크게 3 가지 타입 Binary Classification, Multi-Class-Classification, Multi-Label-Classification 으로 구분된다. Binary Classification 은 두 개의 고유한 클래스만 존재하고, 분류하려는 데이터가 해당 클래스 중 하나에만 독점적으로 속할 때 사용된다. 예를 들어, 제품에 대한 게시물의 긍정/부정 반응을 분류하기 위해 사용될 수 있다. Multi-Class-Classification 은

3 개 이상의 클래스가 존재하고, 분류하려는 데이터가 해당 클래스 중 하나에만 독점적으로 속할 때 사용된다. 예를 들어, 0~9 사이에 숫자를 분류하기 위해 사용될 수 있다. Multi-Label-Classification[2]은 두 개 이상의 클래스가 존재하고, 분류하려는 데이터가 클래스 중 어느 것에도 속하지 않거나 모든 클래스에 동시에 속할 때 사용된다. 예를 들어, 이미지에 하나의 객체만 있는 것이 아니라 여러 개의 객체가 있는 경우와 같이 각 클래스가 배타적이지 않을 때 사용될 수 있다. 본 논문에서는 이러한 Multi-Label-Classification 을 이용하여 원본 사진 배경의 RGB 값을 예측하는 데 사용하였다.

## 3. 감성 증명사진 제작 시스템 설계 및 구현

### 3-1. 사진 제작 시스템의 구성

본 논문에서 제안하는 웹 애플리케이션에 적용된 새로운 감성 증명사진 제작을 위한 시스템의 프로세스는 그림 2 와 같다.



(그림 2) 증명사진 제작 시스템 구축 과정

### 3-2. 데이터 수집 및 전처리

- 사진 크롤링 : SNS 와 구글 웹사이트에서 데이터 크롤링을 통해 배경색이 있는 인물 사진 42,898 장을 수집하였다. 여러 인물이 등장하거나 테두리가 있는 사진 등은 제거하였다.
- Image Matting: P3M-Net 은 사진 속에서 인물을 식별하고 높은 품질의 alpha matte 을 얻을 수 있다[3]. 이를 활용하여 사진 속 물체와 배경을 분리해 낼 수 있다. 추후에 입력으로 사용될 인물 사진을 배경과 분리하기 위해서도 필요한 기능이다.
- 배경색 추출 : 사진의 배경색(R, G, B 값)을 추출하여 txt 파일로 저장하였다. 배경색 데이터는 학습 시에 라벨 데이터로 사용된다. RGB 값을 255 로 나누고 정규화 하여 활용했으며, 학습 과정에서 라벨의 정답은 0 부터 1 사이의 값을 가지도록 하였다.



(그림 3) 감성 증명사진 데이터 셋

<sup>1</sup> 시현하다, <https://sihyunhada.com/>.

### 3-3. Inference

Multi-Label-Classification 으로 각 사진에 대한 배경색을 추정한다. EfficientNet 에서도 가장 가벼운 모델인 EfficientNet-B0 의 pretrained 모델을 사용하여 빠른 속도로 추론을 진행하고자 하였다[5]. 최종 output layer 를 추가하여 본 데이터에 맞는 추론을 할 수 있게 하였다. RandomHorizontalFlip, RandomCrop, RandomRotation 으로 데이터 증강을 하였으며, Loss Function 은 L1, MSE Loss 로 각각 실험하였다.

### 3-4 사진 합성

추정된 배경색을 사용하여 새로운 감성 증명사진의 배경을 합성한다. 앞선 Image Matting 과정을 통해 추출된 인물 사진에 배경색을 적용하여 jpg 타입으로 저장한다.



(그림 4) 배경 합성 전 사진과 배경 합성 후 사진

## 4. 웹 애플리케이션 개발

앞선 과정을 통해 개발된 시스템의 코드를 활용하여 사용자가 본인의 사진을 업로드하면 새로운 감성 증명사진을 생성해 주는 웹 애플리케이션을 구현하였다.

### 4-1. 시스템 아키텍처

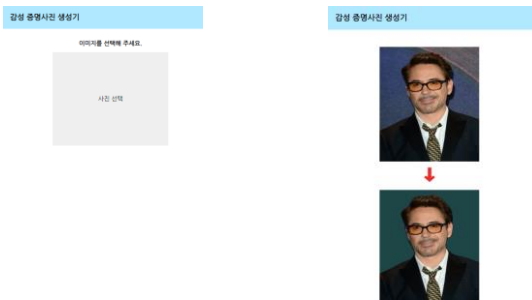
프론트엔드는 React JS 를 사용하였다. React JS 는 흔히 React 라고 불리며, 풍부한 UI 디자인을 위해 개발된 JavaScript 라이브러리이다[4]. 백엔드는 Flask[6], 데이터 저장을 위해서 Amazon S3(Amazon Simple Storage Service)[7]가 활용되었다.



(그림 5) 증명사진 제작 웹 애플리케이션 프로세스

### 4-2. 인물 사진 변환 웹 애플리케이션 UI

본 논문에서 제안하는 웹 애플리케이션의 UI는 다음과 같다.



(그림 5) 감성 증명사진 생성기 웹 화면 (그림 6) 감성 증명사진 생성이 완료된 웹 화면

사용자가 웹에 접속하면 그림 5 와 같은 화면이 제공된다. 사용자는 사진 선택을 클릭하여 변환하고 싶은 인물사진을 선택할 수 있다. 사진 선택이 완료되면 변환하기 버튼이 나타난다. 변환하기 버튼을 클릭하면 백엔드에서 학습된 모델을 통해 배경색이 추론되고, 해당 배경색으로 사진이 합성된다. 마지막으로 작업이 완료되면 그림 5 와 같이 새롭게 생성된 사진을 확인할 수 있다.

## 5. 실험 결과 및 분석

### 5-1. 실험 환경 사항

실험 환경으로 윈도우 OS 에서의 Python3.8 버전과 pytorch1.8.0+cu111 을 사용하였다. CPU 는 I9-9900KF, GPU 는 RTX3080 을 사용하였고, 데이터 셋으로는 SNS 와 구글 웹사이트를 크롤링 하여 수집한 사진들을 사용하였다.

### 5-2. 실험 내용

본 논문에서는 4 만 여장의 이미지 중에서 10%인 4,000 여 장을 Test Set 으로 사용하였고, 나머지 90%에서 20%인 7,200 여 장을 Validation Set, 80%인 28,800 여 장을 Train Set 으로 사용하였다. Loss 함수별로 얼마나 다른 차이를 나타내는지 알아보기 위해서 L1Loss, MSELoss 를 사용해서 각각 측정해 보았다. 정량적인 판단도 중요하지만 감성 증명사진은 사용자의 만족도를 우선으로 하기 때문에 사진에 대한 선호도를 조사하는 설문을 실시하였다.

설문은 Type A, Type B, Type C 의 경우로 나누어 진행하였다. Type A 의 경우에는 원본 사진과 유사하게 추론된 이미지들(n=10) 이고, Type B 의 경우에는 원본 사진과 많이 다르게 추론된 이미지들(n=5)이다. 마지막으로 Type C 의 경우에는 새롭게 크롤링 한 연예인 사진들을 모델에 적용하여 얻은 이미지(n=5)이다. 사람들(n=20)에게 위치를 세 가지 Type 별로 나누어진 서로 다른 설문 문항으로 설문을 진행하였다.

Type A 의 경우 '1. 사진이 얼마나 자연스럽다고 생각하시나요?(1~5 점)', '2. 두 사진 중에서 어떤 사진이 더 나은 사진이라고 생각하시나요?(0: 원본, 1: 결과)', '3. 왼쪽 사진과 오른쪽 사진을 얼마나 유사하다고 생각하시나요?(1~5 점)'로 설문을 진행하였다.

Type B 의 경우 '1. 왼쪽 사진을 감성 증명사진 생성기에 넣어 오른쪽 사진을 결과물로 얻었을 때, 결과물에 대해 얼마나 만족하시나요?(1~5 점)', '2. 사진의 배경색이 인물과 얼마나 적절한 조화를 이루고 있다고 생각하시나요?(1~5 점)', '3. 사진 속 배경과 인물의 경계가 자연스럽게 처리되었다고 생각하시나요?(1~5 점)', '4. 두 사진 중에서 어떤 사진을 더 선호하시나요?(0: 원본, 1: 결과)'로 설문을 진행하였다.

Type C 의 경우 '1. 왼쪽 사진을 감성 증명사진 생성기에 넣어 오른쪽 사진을 결과물로 얻었을 때, 결과물에 대해 얼마나 만족하

시나요?(1~5 점), '2. 사진의 배경색이 인물과 얼마나 적절한 조화를 이루고 있다고 생각하시나요?(1~5 점)', '3. 사진 속 배경과 인물의 경계가 자연스럽게 처리되었다고 생각하시나요?(1~5 점)'로 설문을 진행하였다.

### 5-3. 실험 결과

	train	validation	test
L1 Loss	0.023	0.034	0.274
MSE Loss	0.00093	0.00266	0.126

표 1. 감성 증명사진 배경 추론 모델 실험 결과

L1Loss 의 경우 Train 0.02, Validation 0.03 까지 줄어들었지만 test 에서 상대적으로 좋지 않은 성능이 나왔고, MSELoss 의 경우 Train 0.00093, Validation 0.00266 까지 줄어들지만 test 에서 상대적으로 좋지 않은 성능이 나왔다. 이는 크롤링 된 데이터의 크기가 조금씩 달랐던 것이 이유라고 생각된다.

설문 조사 결과 표 2와 같은 결과를 얻을 수 있었다.

	Type A			Type B				Type C		
	Q1	Q2	Q3	Q1	Q2	Q3	Q4	Q1	Q2	Q3
평균	3.33	0.33	3.58	3.71	3.77	4.18	0.66	3.17	3.4	2.96

표 2. 감성 증명사진에 대한 설문 조사 결과

### 5-4. 실험 결과 분석

레이블 값을 RGB 에서 각각 255 로 나눠주었기 때문에 L1Loss 에서 0.03(validation 기준)가 나왔다는 것은 전반적으로 모델의 추론 결과가 원본 사진과 크게 차이 나지 않는다고 볼 수 있다. 하지만 증명사진의 특성상 색이 조금만 달라져도 사람들이 보기에 선호되지 않을 수 있다. 그러한 이유로 본 논문에서는 사람들(n=20)에게 추가적인 설문을 진행하였다.

위에 표를 참고하면 Type A 의 Q2 는 0.33 으로 원래 사진보다 상대적으로 선호가 덜한 경향을 보였다. 반면 Type B 의 Q4 의 경우에는 0.66 으로 원래 사진보다 본 논문의 결과 사진을 더 선호하는 경향을 보였다.

이것으로 결과적으로 새롭게 생성된 감성 증명사진만 봤을 때 선호도가 높다는 점과 딥러닝 모델을 이용하여 사진 제작을 자동화함으로써 초당 몇 십장 이상을 생성할 수 있다는 것을 상당한 이점으로 볼 수 있기에 굉장히 뛰어난 성과이다.

## 6. 결론 및 향후 연구

본 논문은 딥러닝을 활용하여 개인의 사진에 따라 추천 배경색을 추론하고, 이를 활용해 새롭게 제작한 사진을 제공해 주는 웹 서비스를 제안하였다. 웹을 통해 사용자들은 사진관에 가는 번거로움을 줄이고, 사진 촬영 및 편집 비용 없이 사진을 생성하며 저장할 수 있다. 또 사진관에서 딥러닝을 이용하여 배경색을 추천하는 프로그램을 가지고 있다면 사진사가 별도의 훈련과정 없이도 감성 증명사진을 연출해 낼 수 있을 것이다.

그리고 더욱 발전된 Image Matting 기술을 활용하면 더 자연스럽게 이미지를 얻을 수 있기 때문에 스마트폰 앱에 이 기능을 적용한다면 일반 사람들이 더 쉽고 다양한 환경에서 사용할 수 있을 것이다.

또한, 사진 변환 이후 사용자의 만족도를 수집하고 이를 라벨 데이터로 활용하여 배경색 추론 모델 학습에 활용될 수 있을 것이다. 이는 생성된 사진이 사용자 개인에게 더 높은 만족도를 주도록 기여할 것이다.

## 감사의 글

이 논문은 2022 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No.2017-0-00217, 투명도와 레이어 가변형 실감 사이너지 기술 연구)

## 참고문헌

- [1] 강지연, 김서현, 김서영, 이경미.(2021).딥러닝을 이용한 통합형 취업사진 생성 앱 개발.한국디지털콘텐츠학회 논문지,22(12),1951-1958.
- [2] Bi, W., & Kwok, J. (2013, May). Efficient multi-label classification with many labels. In International conference on machine learning (pp. 405-413). PMLR.
- [3] Li, Jizhizi, et al. "Privacy-preserving portrait matting." Proceedings of the 29th ACM International Conference on Multimedia. 2021.
- [4] Maratkar, Pratik Sharad, and Pratibha Adkar. "React JS-An Emerging Frontend JavaScript Library." Iconic Research And Engineering Journal s 4.12 (2021): 99-102.
- [5] Tan, Mingxing, and Quoc Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." International conference on machine learning. PMLR, 2019.
- [6] Flask [Internet], <https://flask.palletsprojects.com/en/2.1.x/>
- [7] Amazon Simple Storage Service [Internet], <https://docs.aws.amazon.com/AmazonS3/latest/userguide/Welcome.html>