

# CNN 기반의 모델 학습을 통한 관계 분류 모델

## : AI 기반의 셀프사진관 포즈 추천 프레임워크

백강민<sup>1</sup>, 한연지<sup>†</sup>

<sup>1</sup>대전대학교 정보통신공학과 학부생

<sup>†</sup>한양대학교 컴퓨터·소프트웨어학과

gsh1023@naver.com, yeon9891@naver.com

### Relationship classification model through CNN-based model learning: AI-based Self-photo Studio Pose Recommendation Frameworks

Kang-Min Baek<sup>1</sup>, Yeon-Jee Han<sup>†</sup>

<sup>1</sup>Dept. of Information and Communication Engineering, Dae-Jeon University

<sup>†</sup>Dept. of Computer Science, Hanyang University

#### 요 약

소위 ‘인생네컷’이라 불리는 셀프사진관은 MZ 세대의 새로운 놀이 문화로 떠오르며 사용자 수가 나날이 증가하고 있다. 그러나 짧은 시간 내에 다양한 포즈를 취해야 하는 셀프사진관 특성상 촬영이 낯선 사람에게는 여전히 진입장벽이 존재한다. 더불어 매년 비슷한 포즈와 사진 결과물에 기존 사용자는 점차 흥미를 잃어가는 문제점도 발생하고 있다. 이에 본 연구에서는 셀프사진관 사용자의 관계를 분류하는 모델을 개발하여 관계에 따른 적합하고 다양한 포즈를 추천하는 프레임워크를 제안한다. 사용자의 관계를 ‘couple’, ‘family’, ‘female\_friend’, ‘female\_solo’, ‘male\_friend’, ‘male\_solo’ 총 6 개로 구분하였고 실제 현장과 유사하도록 단색 배경의 이미지를 우선으로 학습 데이터를 수집하여 모델의 성능을 높였다. 모델 학습 단계에서는 모델의 성능을 높이기 위해 여러 CNN 기반의 모델을 전이학습하여 각각의 정확도를 비교하였다. 결과적으로 195 장의 test\_set 에서 accuracy 0.91 의 성능 평가를 얻었다. 본 연구는 객체 인식보다 객체 간의 관계를 학습시켜 관계성을 추론하고자 하는 것을 목적으로, 연구 결과가 희박한 관계 분류에 대한 주제를 직접 연구하여 추후의 방향성이나 방법론과 같은 초석을 제안할 수 있다. 또한 관계 분류 모델을 CCTV 에 활용하여 미아 방지 혹은 추적과 구조 등에 활용하여 국가 치안을 한층 높이는 데 기대할 수 있다.

#### 1. 서론

끊임없이 변화하고 발전하는 현대 문화에서 ‘인생네컷’이라 불리는 셀프사진관은 새로운 놀이 문화로 급부상하였다. 장 당 2 천원이라는 합리적인 가격에 DSLR 의 퀄리티 높은 촬영 장비를 이용할 수 있다는

점은 MZ 세대에게 공감과 흥미를 이끌었다. 그 영향으로 셀프사진관의 전국 매장 수는 1,800 개 이상 생겨났고, 셀프사진관은 변화가를 넘어 위치에 영향을 받지 않고 대부분의 사람들이 접할 수 있게 되었다. 하지만 다양한 포즈를 통해 자기를 표현할 수 있는

<sup>†</sup> 교신저자(Corresponding author): KT 융합기술원

한연지(yeonjee.han@kt.com)

창의성과 개인화가 요구되는 서비스이기에 촬영이 낯선 사용자에게 어려움이 따른다. 10 초마다 각각의 포즈로 8 장의 사진을 찍어야한다는 점은 신규 사용자에게 장벽이 될 수 있고, 기존 사용자 입장에서는 매번 비슷한 포즈를 취함으로써 지루함을 느낄 수 있다.

본 연구에서는 모델 학습을 통해 사용자의 관계를 6 가지로 분류하고 이에 따른 맞춤형 포즈를 사용자에게 추천하고 제시한다. 또한 객체를 탐지하고 인식하는 기존 연구들과 달리 객체 간의 관계를 추론하고 분류할 수 있다는 점에서 차별화된 의의를 가진다.

## 2. 데이터셋

관계 분류 모델의 성능을 높이기 위해 셀프사진관 현장과 유사한 이미지를 우선으로 수집하였다. 또한 가족사진 혹은 학창시절 사진 등 앨범에 남아있는 일상적인 사진을 수집하여 부족한 데이터셋을 보강하였다. 데이터 수집은 공유하지 않는 목적으로 허락 하에 지인에게 이미지를 제공받았다. 결과적으로 셀프 사진관 데이터 723 장과 일상사진 1,281 장을 수집하였고 이미지 가공을 통해 셀프사진관 데이터 2,684 장 (671x4)과 1,090 장의 일상사진을 학습에 사용하였다. 전처리 과정에서 이미지 크기를 244x224 로 설정하였고 train\_set 2,723 장, val\_set 722 장, test\_set 195 장으로 데이터셋을 무작위로 분류하였다. 마지막으로 모델의 성능을 높이기 위해 여러 데이터 증강 기법을 사용하여 기존 데이터에 대한 다양한 변환을 적용하였다.

## 3. 관계 분류 모델

관계 분류 단계에서는 모델 성능을 높이기 위해 CNN 기반의 모델들을 사용하여 전이학습을 진행하였다. 분류할 클래스는 총 6 개로 couple=[0], family=[1], female\_friend=[2], female\_solo=[3], male\_friend=[4], male\_solo=[5]로 나누었다. 데이터 수는 female\_solo > female\_friend > couple > male\_solo > family > male\_friend 순으로 많았다. 학습에 사용한 일반 CNN, ResNet, VGG 및 GoogLeNet 의 모델 모두 동일한 데이터셋과 하이퍼파라미터를 적용하여 각각의 성능을 비교하였다. 학습 후 모델 성능 평가에서 test\_accuracy 값이 가장 높은 GoogLeNet 의 InceptionResNetv2 모델을 사용하였다(그림 2). 학습 과정에서 Overfitting 과 Vanishing

문제를 방지하기 위해 val\_loss 값이 7patient 동안 복귀되지 않을 경우 Early stopping 되도록 설정하였다. 모델은 10 개의 레이어로 구성된 일반 CNN 모델 1 개와 ResNet-50, VGG-16, InceptionResNetv2 3 개의 pre-trained 모델로 학습을 진행하였다.

일반 CNN 모델을 사용한 학습은 33 epoch 에서 Early stopping 이 발생하였고 validation 에서 accuracy 0.66, loss 0.92 로 모호한 성능을 보였다. ResNet-50 모델의 경우 accuracy 0.34, loss 1.32 로 가장 낮은 성능을 보였다. VGG-16 모델의 경우 accuracy 0.76, loss 0.65 로 전이학습에 대한 효과를 보였다. InceptionResNetv2 모델에서 가장 높은 성능을 보였는데 validation 과 test 에서 각각 accuracy 0.92-0.90, loss 0.22-0.38 의 성능을 보여 관계 분류 모델에 적합한 모델로 선정하였다<표 1>. 이 결과는 계획 단계에서 조사 및 참고했던 논문 [1]의 결과와 유사하였다.

test\_set 에 대한 클래스별 성능 평가를 보았을 때 데이터 수와 성능 평가가 비례하지 않는다는 것을 확인할 수 있다<표 2>. 하나의 이미지 파일에 속한 객체 수가 적은 클래스일수록 높은 성능 평가를 보였다.



(그림 2) 관계 분류 결과 사진.

<표 1> 모델에 대한 성능 평가 점수

	Accuracy	Recall	Precision	F1-Score
CNN	0.66	0.68	0.67	0.67
ResNet-50	0.34	0.37	0.34	0.35
VGG-16	0.76	0.74	0.75	0.77
*Inceptionv4	0.92	0.90	0.90	0.90

\*Inceptionv4=InceptionResNetv2

<표 2> 클래스별 성능 평가 점수

	Accuracy	Recall	Precision	F1-
--	----------	--------	-----------	-----

				Score
couple	0.94	0.85	0.85	0.90
family	0.83	0.83	0.95	0.89
female_friend	0.88	0.88	0.92	0.90
female_solo	0.95	0.95	0.93	0.94
male_friend	0.83	0.83	0.83	0.83
male_solo	0.92	0.92	0.92	0.92

#### 4. 결론

본 연구에서는 CNN 기반의 모델을 학습하여 사람 객체 간의 관계 분류에 대한 유의미한 결과를 얻어냈다. 또한 pre-trained 모델의 전이학습을 통해 모델에 최적화된 학습 방법을 연구하였다. 결과적으로 객체 간의 관계를 추론하는 모델로 기존에 없던 2 차원적인 개념에 대한 연구를 제시하고 검증하였다. 향후 관계 분류 모델을 CCTV 에 활용하여 미아 방지 혹은 추적과 구조에 활용하여 국가 치안을 높이는 데 도움이 될 수 있다.

#### 사사

본 논문은 과학기술정보통신부 정보통신창의인재 양성사업의 지원을 통해 수행한 ICT 멘토링 프로젝트 결과물입니다.

#### 참고문헌

- [1] SZEGEDY, Christian, et al. Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proceedings of the AAAI conference on artificial intelligence. 2017.