

# XAI 를 활용한 설명 가능한 요가 자세 이미지 분류 모델

박유림, 김현희  
동덕여자대학교 정보통계학과  
20190140@dongduk.ac.kr, heekim@dongduk.ac.kr

## Yoga Poses Image Classification and Interpretation Using Explainable AI (XAI)

Yu Rim Park, Hyon Hee Kim  
Dept. of Information and Statistics, Dongduk Women's University

### 요 약

최근 사람들의 건강에 대한 관심이 많아지고 다양한 운동 콘텐츠가 확산되면서 실내에서 운동을 할 수 있는 기회가 많아졌다. 하지만, 전문가의 도움없이 정확하지 않은 동작을 수행하다 큰 부상을 입을 위험성이 높다. 본 연구는 CNN 기반 요가 자세 분류 모델을 생성하고 설명가능 인공지능 기술을 적용하여 예측 결과에 대한 해석을 제시한다. 사용자에게 설명성과 신뢰성 있는 모델을 제공하여 자신에게 맞게 올바른 자세를 결정할 수 있고, 무리한 동작으로 부상을 입을 확률 또한 낮출 수 있을 것으로 보인다.

### 1. 서론

코로나-19 팬데믹 이후, 사람들의 건강에 대한 관심이 증가하면서 홈 트레이닝 열풍이 수년째 불고 있다. 사회적 거리두기로 인한 실외활동의 제한으로 홈 트레이닝을 선택하는 비중이 높아지면서 SNS 로 인한 수많은 운동 콘텐츠가 빠르게 확산되었다. 이에 사람들은 운동에 대한 장벽을 없애고 헬스, 요가, 필라테스, 스트레칭 등 다양한 분야에 쉽게 접근할 수 있었다.

그 중에서도, 100 여개가 넘는 요가 자세들을 전문가의 피드백 없이 정확하게 수행하는 것은 매우 어려운 일이다. 자신의 신체 상태를 파악하지 못한 채 무리한 동작을 이어가면 관절에 통증이 발생하거나 척추에 큰 부담을 가해 디스크가 생기는 등 큰 부상으로 이어질 가능성이 높다 [1]. 또한, 비슷한 요가 자세라 하더라도 사용하는 신체 부위에 따라 근육이나 관절에 부하되는 힘의 크기가 다르기 때문에 더욱 주의할 기울여야 할 부분이다.

본 연구는 요가 자세 이미지를 CNN 모델로 분류하고, 설명가능 인공지능 기술(XAI) [2]를 적용해 예측값에 설명성을 부여하도록 한다. 요가 자세 데이터를 4

개의 컨볼루션 층을 통과하여 5 개의 카테고리 분류하고 LIME 모델[3]을 적용하였다. 93.9%의 높은 성능을 기록한 CNN 모델이 요가 자세의 어떤 부분을 보고 분류 예측하게 되었는지 찾아낼 수 있었고, 신뢰성을 갖춘 모델을 구축할 수 있었다. 예측에 대한 설명과 해석이 가능해지면서 사용자들은 모델에 대한 신뢰도를 높일 수 있을 뿐만 아니라 검증된 분류 모델로 자신의 몸 상태에 맞는 요가 자세를 선택하여 올바른 의사결정을 할 수 있도록 할 것이다.

### 2. 데이터 수집 및 전처리

머신러닝 관련 서비스를 제공하는 공공 데이터 플랫폼 Kaggle 사이트를 이용하였다. 5 개의 요가 자세 파일(Downdog, Goddess, Plank, Tree, Warrior2)로 구성된 총 988 개의 이미지 데이터를 수집하여 파일 이름을 정답값(labels)으로 생성하였다. 다음으로 이미지 데이터를 각각 훈련 데이터 80%, 테스트 데이터 20%로 분리하고 각 데이터 집합에 원-핫 인코딩을 적용하였다. 또한, 이미지를 255 로 나누는 정규화를 수행하여 기계가 학습할 수 있는 데이터 형태로 전처리하였다.

### 3. CNN 모델 구축 및 성능 평가

#### 3.1. CNN 모델 구축

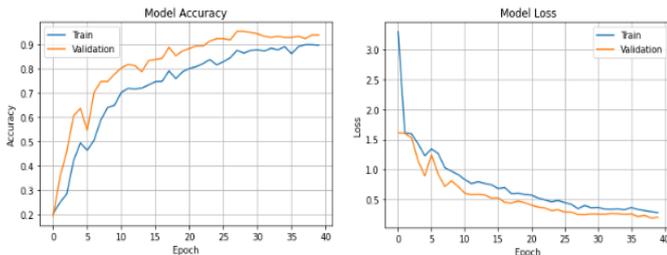
\* 본 연구는 2023 학년도 중소기업벤처부의 SW · 콘텐츠 인재 양성사업(벤처스타트업아카데미) 지원에 의한 연구임

170x170 크기의 rgb 채널 이미지를 입력 받아 총 4 개의 컨볼루션층과 일반 신경망 층을 통과하도록 구성하였다. 3x3 크기의 커널 32 개와 64 개를 지정하여 다중 특징맵을 추출하였다.

과잉적합 문제를 방지하기 위하여 2x2 크기의 최대 풀링층으로 특징맵 요약통계량을 추출하였고, 드롭아웃을 0.25 로 지정하여 은닉층의 노드를 임의로 제거하였다. 또한, 신경망이 충분히 큰 훈련 집합을 학습하도록 이미지 증대(Image Augmentation) 기법을 모델에 삽입하여 데이터에 대한 다양성을 높이고자 하였다. 요가 자세의 5 가지 카테고리로 알맞게 분류하기 위하여 출력층에는 소프트맥스(softmax) 활성화 함수와 categorical-crossentropy 손실함수를 사용하였고, 최적화 방법은 Adam 기법을 사용하였다.

### 3.2. CNN 모델 성능평가

그림 1. 는 훈련 데이터가 CNN 모델을 통과하면서 40 회의 학습(epochs)을 진행한 이후의 성능을 나타낸다. 정확도는 최대 93.9%를, 손실함수는 최저 0.201 을 기록하였다.



(그림 1.) CNN 모델의 정확도, 손실함수 그래프

## 4. LIME 기반 XAI 기법 적용

### 4.1. LIME 모델 구축

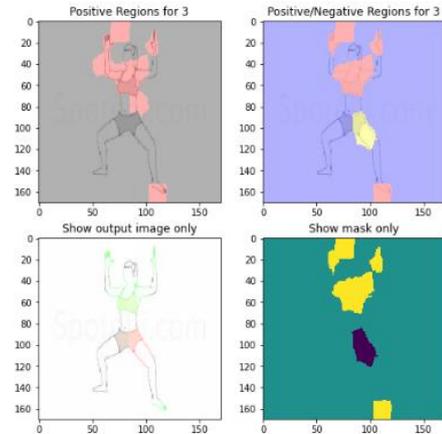
가장 먼저 LimeImageExplainer() 메소드로 LIME 설명체를 생성하였고, 분할 알고리즘 (Segmentation Algorithm)으로는 Slic 을 적용하여 이미지를 의미 있는 픽셀 단위로 그룹화하도록 하였다. 이후, explain\_instance() 메소드로 이미지에 대한 설명 가능성을 분석한 LIME 모델을 구축하였다. CNN 분류 모델의 예측 결과와 앞서 선택한 분할 알고리즘 Slic 을 호출하였고, 해석을 위한 샘플의 개수 (num\_samples) 로 1000 개를 지정하였다.

결과적으로, 모델이 결정한 예측값에 공헌도가 높았던 픽셀을 마스크로 시각화 할 수 있었고, 직관적인 이해가 가능하게 되어 이미지의 어떤 부분이 예측에 영향을 주었는지를 알 수 있었다.

### 4.2. LIME 모델 시각화 및 결과

테스트 데이터를 임의로 정하여 보았을 때, 상체와 발을 보고 여신 자세(Goddess)로 올바르게 분류하였을

을 알 수 있다. 네 개의 그림 모두 CNN 분류기가 데이터의 어떤 부분을 보고 분류 예측하였는지 마스크를 표시함으로써 개별 예측값에 대한 설명을 제공할 수 있도록 하였다.



(그림 2.) LIME 모델을 적용한 해석 결과 시각화

## 5. 결론

기존의 딥러닝 모델은 성능 평가를 하기 위한 방법으로 정확도(Accuracy), 정밀도(Precision), 재현율(Recall) 등을 계산한 혼동 행렬(Confusion Matrix)을 사용했지만, 그 모델이 예측하는 과정을 볼 수 없다는 단점이 있었다.

본 논문에서는 이러한 점을 보완하고자 LIME 모델을 적용하여 요가 자세 데이터를 5 가지로 분류하는 CNN 예측 모델을 구축하였다. 그 결과, 인공지능 모델이 분류하는 과정을 개별 이미지 데이터들의 마스크를 통해 해석이 가능하였고, 나아가 CNN 예측 모델에 대한 신뢰성도 증명할 수 있었다.

LIME 설명가능 인공지능(XAI)은 모델 독립적(model-agnostic)이기 때문에 데이터나 모델 형태 상관없이 유연한 적용이 가능하다. 따라서 요가 자세 뿐만 아니라 다른 도메인에서 사용하는 데이터, 혹은 다양한 모델들도 설명가능 인공지능 기술을 통한 신뢰성과 설명력을 사용자들에게 제공할 수 있을 것으로 기대된다.

### 참고문헌

- [1] 이수민, 이재유, 정희진, 치우웨이, 배지훈, 요가 자세 교정을 위한 개인 트레이닝 서비스에 대한 연구, 한국정보기술학회 종합학술발표논문집, 2022, 730-732
- [2] 안재현 저, 「XAI 설명 가능한 인공지능, 인공지능을 해부하다」, 위키북스, 2020
- [3] Ribeiro, M.T., Singh, S. and Guestrin, C. (2016) "Why Should I Trust You?: Explaining the Predictions of Any Classifier", Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations, 1135-1144