

CNN 기반 분변 이미지 분석을 통한 장 건강 평가 기술

장보석, 박미소, 이기원, 조태진, 신용구*

고려대학교

{bs_jang0629, miso419, assuuss1, microcho, *ygshin92} @korea.ac.kr

Gut Health Assessment Using CNN-Based Fecal Image Analysis

Bo-Seok Jang, Mi-So Park, Ki-Won Lee, Tae-Jin Cho, and Yong-Goo Shin

Korea University

요약

본 논문은 CNN (Convolutional Neural Network) 기반의 이미지 분류 인공지능 모델을 활용하여, 영유아의 분변 이미지를 브리스톨 분변 차트 (Bristol Stool Chart)에 따라 자동으로 분류하고 이를 통해 장 건강 상태를 평가하는 방법을 제안한다. 모델 학습 과정에서 발생하는 데이터 불균형으로 인한 성능 저하 문제를 완화하기 위해 다양한 데이터 증강 기법과 Focal Loss 함수를 적용했다. 실험을 통해 제안한 방법이 데이터 불균형 문제를 효과적으로 완화하고, 영유아 분변 이미지 분류 정확도 향상에 기여함을 확인했다.

I. 서론

이미지 분류 (Image Classification)는 인공지능 모델이 입력된 이미지를 분석하여 사전에 정의된 클래스에 따라 분류하는 문제로, 컴퓨터 비전 분야에서 오랜 기간 활발히 연구되어 온 핵심 과제 중 하나이다. 이러한 분류 작업에는 주로 컨볼루션 신경망 (Convolutional Neural Network, CNN) 기반 모델이 활용된다.

본 연구에서는 CNN 기반의 이미지 분류 모델을 활용하여 영유아의 분변 이미지를 분석 및 분류하고, 이를 통해 장 건강 상태를 평가하는 방법을 제안한다. 분변의 형태적 특징을 체계적으로 분류하기 위해, 본 연구는 브리스톨 분변 차트 (Bristol Stool Chart)[1]를 기준으로 사용한다. 이 차트는 대변의 형태를 7가지 유형으로 나누어 설명하며, 각 유형은 변비부터 정상, 설사까지의 상태를 나타낸다. 간결하고 직관적인 분류 방식으로 구성되어 있어, 의료 전문가뿐만 아니라 일반인도 쉽게 활용할 수 있으며, 장 건강 모니터링 및 질환 조기 진단에 유용하게 사용된다.

한편, 본 연구에 사용된 영유아 분변 이미지 데이터의 전체 개수가 현저히 적고 클래스 간 데이터 분포 또한 불균형하게 나타나, 학습 과정에서 성능이 저하되는 문제가 발생했다. 이에 본 연구에서는 데이터의 다양성과 학습 안정성을 확보하기 위해 이미지 증강 (Data Augmentation) 기법을 적용하고, 클래스 불균형에 효과적으로 대응할 수 있는 Focal Loss[2]를 도입하여 분류 성능을 향상시킨다.

II. 본론

2.1 Focal Loss 함수

기존의 분변 이미지 분류 연구들에서는 상대적으로 많은 양의 학습 데이터를 기반으로 3개 또는 4개의 간략화된 클래스 (예: 정상, 변비, 설사 등)로 분류하는 접근이 주를 이루었다[3, 4]. 그러나 실제 임상 환경이나 제한된 데이터 세트 활용하는 상황에서는 클래스 간 데이터 불균형이 심각하게 발생할 수 있다. 특히 본 연구에서는 각 분변 차트 유형별로 분포된 이미지 수의 편차가 크고, 수집된 전체 데이터 수도 현저히 적기 때문에 기존의 이미지 분류 모델 학습에서 주로 사용되는 Cross Entropy

	1번	2번	3번	4번	5번	6번
Train	28	62	62	123	25	20
Test	6	6	13	33	6	20

표 1. 영유아 분변 데이터 세트 구조

Loss 함수만으로는 성능 저하가 불가피하다. 이를 해결하기 위해, 본 연구에서는 소수 클래스에 더 집중하여 학습할 수 있도록 설계된 Focal Loss 함수를 도입한다.

$$L_{focal} = \alpha \cdot (1 - p_t)^\gamma \cdot L_{CE} \quad (1)$$

Focal Loss는 데이터 수가 많아 예측이 쉬운 클래스보다 학습이 어려운 클래스에 더 큰 학습 비중을 부여하여, 데이터 불균형 환경에서도 우수한 분류 성능을 기대할 수 있도록 설계된 Loss 함수이다. 이 함수는 각 Train 데이터의 정답 클래스에 대한 예측 확률 p_t 에 따라 Loss 값을 조절한다. p_t 가 큰 경우 Loss 기여도를 줄이고, p_t 가 작은 경우 Loss 기여도를 높인다. 여기서 γ 는 Focusing 파라미터로, 값이 커질수록 모델은 분류하기 어려운 데이터에 더욱 집중하게 된다. 또한, 클래스 불균형을 완화하기 위해 클래스별 데이터 개수에 비례하여 가중치 계수 α 를 적용하여, 상대적으로 적은 수의 클래스를 집중적으로 학습할 수 있도록 유도한다.

2.2 실험 환경

본 연구는 분변 차트에 기반하여 라벨링된 영유아 분변 이미지 데이터 세트로 분변 이미지 분류 모델을 학습하고 평가한다. 데이터 세트는 표 1과 같이 총 404장의 이미지로 구성되어 있으며, 1번부터 6번까지 총 6개의 클래스로 구성되어 있다. 4번 클래스 이미지 수가 많은 반면, 1번, 5번, 6번 클래스는 상대적으로 이미지 수가 현저히 적어, 데이터 불균형 문제가 나타난다. 이는 모델 학습 시 데이터 수가 적은 클래스에 대해 학습이 잘 이루어지지 않아, 전체적인 분류 성능 저하의 원인이 된다.

모델 Loss	ResNet18	ResNet50	ResNext50	Efficient-B0
CE Loss	53.5 %	51.19 %	54.76 %	57.14 %
Focal Loss	59.5 %	52.38 %	58.33 %	60.71 %

표 2. 모델 학습 결과 - Top Accuracy

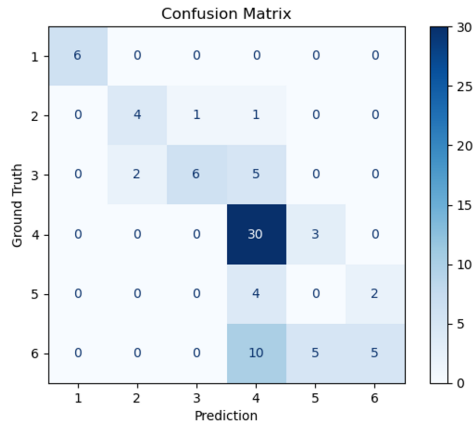


그림 1. 혼동행렬 (Confusion Matrix) 시각화

이를 완화하고 모델의 일반화 성능을 향상시키기 위해, 본 연구에서는 다양한 데이터 증강기법을 적용했다. 구체적으로는 이미지 좌우 반전 (flip), 90도 회전 (Rotation), 가우시안 블러 (Gaussian blur), 중심부 자르기 (Center crop), 수평 이동 (Horizontal shift)의 총 다섯 가지의 증강 기법을 활용하여 데이터 수를 증가시키고 다양성을 부여하여 안정적인 학습이 가능하다.

모델 학습은 ImageNet[5]으로 사전학습 된 모델을 기반으로 전이 학습 (Transfer learning)을 수행했다. 실험에 사용된 모델은 ResNet18[6], ResNet50[6], ResNeXt50[7], EfficientNet-B0[8]이며, 각 모델의 성능을 비교 분석했다. 전체 학습 반복 횟수 (Iteration)는 데이터 규모를 고려하여 5000으로 설정하였고, Optimizer는 AdamW[9]를 사용했다. 학습률 (Learning rate)은 초기값을 5×10^{-3} 으로 설정한 후, 1000회마다 점진적으로 감소시키는 방식으로 조정했다.

2.3 실험 결과

2의 실험 결과를 보면, 모든 CNN 모델에서 Cross-Entropy(CE) Loss 대비 Focal Loss를 적용하였을 때 더 우수한 성능을 나타냈다. 이는 전반적으로 데이터 수가 적어 학습이 어려운 클래스에 대한 분류 정확도가 향상되고, 모델이 안정적으로 학습할 수 있도록 영향을 준 것으로 볼 수 있다. 특히 EfficientNet-b0 모델에 Focal Loss를 적용했을 때 60.71%의 정확도를 기록하며 가장 높은 성능을 달성했다.

그림 1은 EfficientNet-b0 모델에 Focal Loss를 적용하여 학습한 결과에 대한 혼동행렬(Confusion Matrix)을 시각화한 것이다. 혼동행렬은 모델의 예측 결과와 실제 정답 간의 관계를 통해 분류 성능을 평가할 수 있는 도구로, 가로축은 예측 클래스, 세로축은 실제 정답 클래스를 나타낸다. 데이터 수가 많은 4번 클래스에 대해 높은 예측 정확도를 보이지만, 잘못 예측한 경우에는 정답 클래스와 시각적 특성이 유사한 인접 클래스로 예측된 경우가 많았다. 이는 정답 클래스 분변 유형과 인접한 클래스의 분변 유형 간 시각적 유사성이 모델의 예측 혼동을 유발한 것으로 판단된다.

III. 결 론

본 연구에서는 CNN 기반 이미지 분류 모델을 활용하여 영유아 분변 이미지를 분석하고, 데이터 증강 및 Focal Loss를 통해 데이터 불균형 문제를 완화함으로써 모델의 분류 성능을 효과적으로 향상시켰다. 실험 결과, 제안한 방법이 유의미한 성능 개선을 이루었으며, 분변 이미지를 분변 차트에 따라 분류함으로써 장 건강 평가에 활용될 수 있음을 확인했다. 일부 인접 클래스 간 혼동이 발생하는 문제에 대해서는 향후에 데이터 다양성 확보와 세분화된 모델 개선을 통해 정확도를 더욱 향상시킬 예정이다.

ACKNOWLEDGMENT

This work was supported by IITP grant funded by MSIT (Grant#: RS-2025-02263277, 50%). This work was supported by ITRC support program supervised by the IITP and funded by MSIT (Grant#: IITP-2025-RS-2023-00258971, 50%).

참 고 문 헌

- [1] S.J. Lewis and K.W. Heaton. "Stool form scale as a useful guide to intestinal transit time." *Scandinavian journal of gastroenterology* 32.9 (1997): 920-924.
- [2] TY. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar. "Focal loss for dense object detection." *Proceedings of the IEEE international conference on computer vision*. 2017.
- [3] J. Zhou, J. McNabb, N. DeCapite, JR. Ruiz, DA. Fisher, S. Grego, and K. Chakrabarty. "Stool image analysis for digital health monitoring by smart toilets." *IEEE Internet of Things Journal* 10.5 (2022): 3720-3734.
- [4] YP. Choy, G. Hu, and J. Chen. "Detection and classification of human stool using deep convolutional neural networks." *IEEE Access* 9 (2021): 160485-160496.
- [5] J. Deng, W. Dong, R. Socher, LJ. Li, K. Li, and L. Fei-Fei. "Imagenet: A large-scale hierarchical image database." *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009.
- [6] K. He, X. Zhang, S. Ren, and J. Sun. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [7] S. Xie, R. Girshick, P. Dollar, Z. Tu, and K. He. "Aggregated residual transformations for deep neural networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- [8] M. Tan and Q. Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." *International conference on machine learning*. PMLR, 2019.
- [9] I. Loshchilov and F. Hutter. "Decoupled weight decay regularization." *arXiv preprint arXiv:1711.05101* (2017).