

# ResNet을 이용한 평발 분류 모델 개발

Lee-SeungWoo

이 승 우

## Abstract

발자국 영상을 주었을 때 평발과 정상발의 형태를 구분하기 위해 데이터를 학습시켰다. 이는 평발에 대한 빠른 예방 및 대처를 목적으로 분류하게 되었다. 데이터 분류를 위해 전이학습 중에서도 모든 ResNet 모델을 사용하였다. 그 중 ResNet50을 사용했을 때 컴퓨터가 평발과 정상발을 높은 정확도로 분류할 수 있었다. 이 분류된 데이터를 사용해 조기진단으로 빠르게 평발과 정상발을 구분할 수 있어 효율적이게 예방을 할 수 있다. 또 이를 바탕으로 병원에서 더 정확한 진단을 받을 수 있을 것이다.

## 1. 서론

발자국 영상을 주었을 때 평발인지 정상발인지 알기 위해 학습을 수행했다. 우리가 서있을 때 발은 몸을 지탱해주고 걸음 이동을 도와주는 역할을 한다. 발바닥은 아치 형태에 따라 걷거나 달릴 때 충격완화를 시켜주는 정도가 다른데, 발의 아치 구조는 평발과 정상발 등으로 구분할 수 있다. 평발은 아치가 없는 발바닥의 형태로 골반과 척추 정렬 상태에 변화를 줄 수 있고 정상발은 충격 정도가 가장 이상적이며 앞서 말한 평발에 비해 피로가 덜하다[1]. 이렇게 발의 형태에 따라 몸의 구조가 크게 바뀔 수 있다. 따라서 집에서 조기진단을 통해 예방하려는 목적으로 발자국 영상이 주어질 때 평발과 정상발을 컴퓨터가 구분할 수 있게 모델을 구축하려 한다.

## 2. 본론

### 2.1 데이터 세트



그림 1 발자국 데이터 셋

발의 압력에 따라 찍힌 발자국 데이터를 정상발과 평발로 분류하여 데이터를 구축하였다. 트레이닝 데이터 개수는 평발과 정상발 각각 100개로 총 200개 데이터를 구축하였고 테스트 데이터 개수는 각각 42개로 총 84개의 데이터를 구축해 7:3 정도의 비율로 데이터를 구축하였다.

### 2.2 수행 환경

파이토치를 사용했고, 코랩으로 실험을 진행했다.

### 2.3 알고리즘

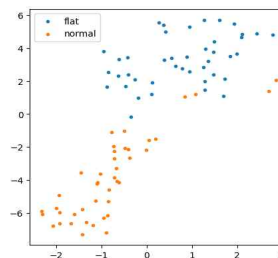
전이학습모델인 ResNet18, ResNet34, ResNet50, ResNet101, ResNet152 알고리즘을 사용하였고 학습환경을 Epoch : 10, Batch\_Size : 32로 동일하게 맞춰주어 ResNet 5개의 버전으로 layer층에 변화를 주어 학습시켰다. 각 모델의 layer는 아래 그림과 같다.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		$1.8 \times 10^9$	$3.6 \times 10^9$	$3.8 \times 10^9$	$7.6 \times 10^9$	$11.3 \times 10^9$

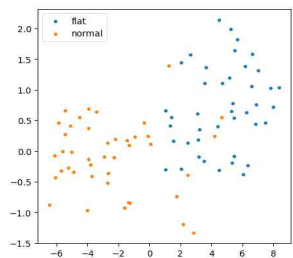
그림 2 ResNet 각 모델별 구조[2]

## 3. 결과

<ResNet 18>



<ResNet 34>



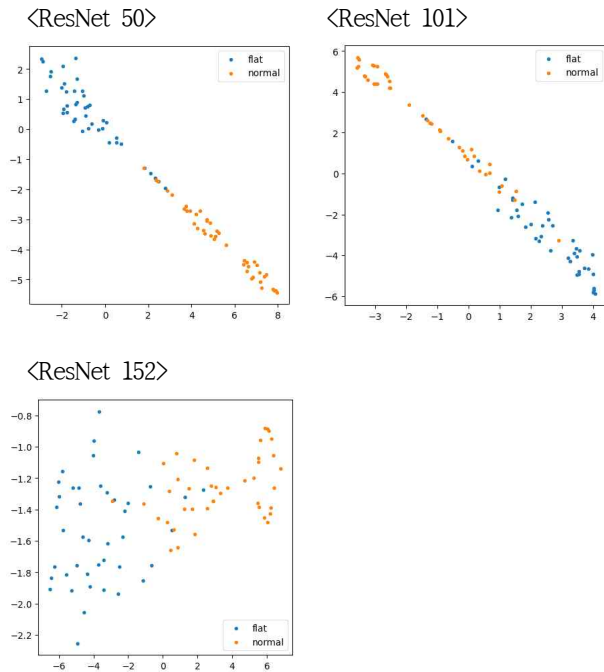


그림 3 ResNet 각 모델별 군집도 비교

표 1 ResNet 각 모델별 정확도 비교

Epoch : 10, Batch_size : 32		
	Training Best ACC	Validation Best ACC
ResNet18	0.9598	0.9404
ResNet34	0.9777	0.7976
ResNet50	0.9776	0.9166
ResNet101	0.9776	0.8690
ResNet152	0.9553	0.8214

#### 4. 고 찰

표 1을 보았을 때 정확도는 큰 차이가 없어 그림 1과 같이 군집도 형성 정도에 따라 모델을 해석하려고 한다. 그림 1에서 ResNet 5가지 모델 중 ResNet18, 34, 152에서의 군집도 보다 ResNet50, 101 모델에서의 군집도가 더 잘 형성이 되어있다. 실험 결과 평발과 정상발에서의 특성 추출은 낮은 layer로도 잘 분류할 수 있을 거라 생각했지만 그보다 더 깊은 layer에서 잘 분류가 되었다. 하지만 ResNet152 모델을 보았을 때 분류가 잘 안 되어있는 것을 볼 수 있다. 즉, layer가 깊으면 깊을수록 더 잘 분류되는 것이 아닌 것을 확인했다. 향후 발자국 데이터뿐만 아니라 실제 발바닥 사진을 촬영해 발바닥 아치 형태에 따른 명암차이 또한 분류할 수 있도록 깊은 layer를 사용하는 모델을 적용하는 방향이 좋을 것으로 보인다.

#### 5. 결 론

자신의 발이 평발인지 정상발인지 구분하는 것은 의사를 통해 진단받아야 하는 번거로움이 있어 집에서 쉽게 자신이 평발인지 정상발인지 조기진단을 할 수 있도록 이 실험을 진행하게 되었다. 즉 이를 바탕으로 병원에서 더 정확한 진단을 받을 수 있도록 도와주는 효과를 기대할 수 있다. 그리고 layer가 깊어질수록 더 잘 분류하는지 알아보기 위해 모든 ResNet모델을 비교하였다. 그 결과 ResNet50 모델이 가장 잘 분류하였고 layer가 가장 깊은 ResNet152는 예상과 다르게 분류가 눈에 띄게 잘되지 않아 layer가 깊어질수록 더 잘 분류하는 것은 아니라고 생각했다. 결론적으로 전이학습모델인 ResNet에서 평발과 정상발의 분류 정도가 높은 모델은 군집도가 제일 좋게 형성된 ResNet50 모델로 판단을 내렸다.

#### 참고문헌

- [1] 발 ! 인체정보 - 서울아산병원, “발 정의”, <https://www.amc.seoul.kr/asan/healthinfo/body/bodyDetail.do?bodyId=28>, (2023.05.28)
- [2] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).