

COSE474-2024F: Final Project Proposal

“ViT를 이용한 머리카락 이미지의 모발 유형 분류”

2022100017 컴퓨터학과 김세희

1. Introduction

모발 특성에 따라, 연출 가능한 헤어스타일의 종류가 다르다. 그에 따라 머리카락 이미지를 분류한다면, 사용자는 보다 쉽게 자신의 모발 유형과 같은 헤어스타일 사진을 참고하여 스타일링을 할 수 있다.

2. Problem definition & challenges

인간에게는 공통적인 모발 형질이 존재하기 때문에, 정해진 범주로 분류가 가능하다. 즉 이 문제 상황에서는 image classification 모델을 사용하는 것이 적합하다. 또한 도메인이 머리카락 이미지에 한정되므로, pre-trained 모델에서 머리카락 이미지를 통해 fine-tuning을 진행할 필요가 있다. 주어진 머리카락 이미지로 모발 유형을 올바르게 구분할 수 있는 수준으로 학습해야 한다.

3. Related Works

ViT는 기존의 RNN, CNN 기반 모델의 한계를 극복하고자 자연어 처리에 사용되던 transformer 모델을 image classification 문제에 적용했다. Transformer의 핵심 메커니즘인 Self-Attention은 이미지 내의 전역적 특징을 학습할 수 있도록 한다.(Dosovitskiy, 2020) 그러나 ViT 모델은 대규모 데이터셋 학습을 요구하기 때문에, 적은 데이터셋으로 학습한다면 과적합 문제가 발생할 수 있다. 따라서 소규모 데이터셋에서도 데이터 효율성을 극대화하는 ViT 기반의 DeiT도 중요한 참고 모델이 된다.(Touvron et al., 2021)

4. Datasets

Kaggle에 업로드되어 있는 Hair Type Dataset(Sree, 2024)을 사용해 fine-tuning을 진행할 예정이다. 모발 유형에 따른 5종의 라벨이 있으며, 각 라벨 별로 수백 개의 이미지 파일이 존재한다. 라벨에 따라 5개의 클래스로 분류하여 학습을 진행할 예정이므로, 해당 데이터셋과 클래스가 적다는 점에서 유사한 CIFAR-10 데이터셋을 활용한 SOTA 모델들을 참고했다.

5. State-of-the-art methods and baselines

Figure 1에서 가장 높은 정확도를 보인 모델은 eff-

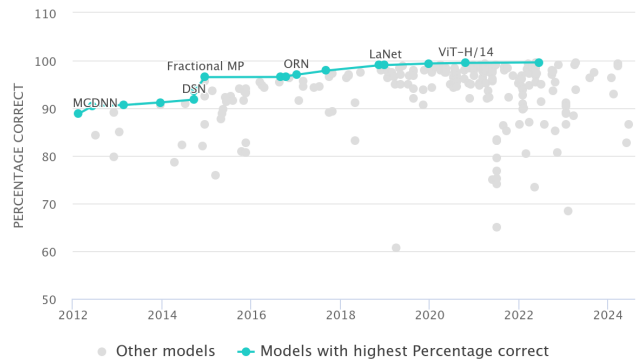


Figure 1. Image Classification on CIFAR-10 (Paperwithcode)

cient adaptive ensembling이지만, 해당 모델은 꽤 복잡하기 때문에, 뒤이어 가장 높은 정확도를 보인 ViT-H/14(Dosovitskiy, 2020)를 사용할 계획이다.

6. Schedule

11.3 모델 학습 및 fine-tuning 11.17 필요시 데이터셋 추가 수집 12.1 모델 성능 평가 및 개선 12.8 보고서 작성

References

- Dosovitskiy, A. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.
- Paperwithcode. Image classification on cifar-10. <https://paperswithcode.com/sota/image-classification-on-cifar-10>. Accessed: 2024-10-16.
- Sree. Hair type dataset. <https://www.kaggle.com/datasets/kavyasreeb/hair-type-dataset>, 2024. Accessed: 2024-10-16.
- Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., and Jégou, H. Training data-efficient image transformers & distillation through attention. In *International conference on machine learning*, pp. 10347–10357. PMLR, 2021.