CNN 기반 암세포 현미경 이미지 분류

김신, 윤경로
건국대학교 컴퓨터공학과
new.xin22@gmail.com, yoonk@konkuk.ac.kr

Cancer Histopathological Image Classification based on Convolutional Neural Network

Shin Kim, Kyoungro Yoon
Department of Computer Science Engineering, Konkuk University

요 약

최근 수 년간 뉴럴 네트워크 기반 이미지 분류 기법의 성능이 눈에 띄게 향상되었다. 특히 CNN 은 딥 러닝 기법을 도입하면서 이미지 분류 정확도가 향상되었으며, 이는 의학 분야 등 다른 분야에도 영향을 주게 되었다. 의학용 이미지의 분류 시스템의 경우, 오분류가 치명적인 결과를 초래할 수 있기 때문에 높은 정확도의 이미지 분류 시스템을 필요로 하게 된다.

본 논문에서는 CNN 기반 암세포 현미경 이미지 분류 기법에 대해 제안한다. 사전에 훈련된 뉴럴 네트워크의 가중치의 일부를 다시 계산하고, 재계산을 통해 얻은 가중치를 기반으로 암세포 현미경 이미지를 분류하며, 분류 결과 높은 정확도로 이미지를 분류하는 것을 확인할 수 있다.

1. 서론

최근 몇 년간 딥 러닝 기술 도입으로 인하여 이미지 분류 기술이 급격하게 발전하였다. 이미지 분류 정확도가 향상되면서 이미지 분류 기법을 응용하여 새로운 이미지 데이터를 분류하는 연구가 활발해졌으며, 이는 의학 분야에 많은 영향을 끼쳤다.

의학 분야는 MRI, 현미경 이미지, 일반 RGB 이미지 등 다양한 이미지 데이터를 가지고 있으며 이러한 데이터를 기반으로 질병을 판단하고 이에 맞는 치료법을 제공해야 한다. 또한 의학 분야에서는 False Negative 에 대한 비용이 높기 때문에 높은 정확도로 분류가 가능해야 한다.

본 논문에서는 CNN 기반 암세포 현미경 이미지 분류 기법을 제안한다. CNN 은 딥 러닝 기반 이미지 분류 기법으로 매우 높은 정확도로 이미지 분류가 가능하기 때문에 높은 분류 정확도를 필요로 하는 의학 분야에 사용하기 적합하다.

본 논문의 구성은 다음과 같다. 2 절에서는 본 논문에서 필요로 하는 배경지식에 대해 서술하며, 3 절에서는 Fine-Tuning 에 대해 기술한다. 4 절에서는 본 논문에서 제안하는 시스템에 대해 설명하며, 5 절에서는 제안한 기법의 실험 결과에 대해 서술하며 마지막으로 6 절에서는 본 논문에 대한 결론을 맺는다.

2. 배경 지식

CNN(Convolutional Neural Network)는 이미지 분류 기법 중에서 현재 제일 일반적으로 사용되고 있는 기법이다. 딥 러닝 기술의 도입으로 인해 CNN 을 이용한 이미지 분류 기법의 성능이 엄청나게 향상되었다.

GoogLeNet[1]은 2014 년 ILSVRC(ImageNet Large Scale Visual Recognition Challenge)에서 우승을 차지한 이미지 분류 네트워크다. GoogLeNet 의 핵심 모듈인 Inception Module 은 필터를 통해 얻은 이미지 특징 차원을 축소시켜 줌으로써 CNN 의 망을 깊게 하되 계산복잡도를 크게 증가시키지 않는 특징을 갖고 있다. 이후 Inception Module 의 내부 구조를 수정하면서 현재 이미지 분류 기법에서 많이 사용되는 Inception V3[2] 가 개발되었다.

Xception[3]은 Inception V3 와 동일한 구조를 가지는 네트워크로서 일반적인 Convolution 을 사용하는 Inception V3 와 다르게 Depthwise Separable Convolution 을 사용한다. Depthwise Separable Convolution 은 이미지 내 공간 정보와 채널 간 정보가 독립적이라는 가정 하에 Convolution 을 수행하는 방법으로 같은 크기의 이미지를 Convolution 한다고 할 때 일반적인 Convolution 보다 계산 복잡도가 낮아 효율적으로 파라미터를 계산할 수 있다. 또한 Residual Connection[4]을 추가하여 기존의 Inception V3 보다 분류 정확도를 향상시켰으며 Inception V3 와 Xception 의 분류 정확도는 표 1 와 같다.

표 1. 네트워크에 따른 이미지 분류 정확도

네트워크	Inception V3[2]	Xception[3]
Top-1 Accuracy	0.782	0.790
Top-5 Accuracy	0.941	0.945

3. Fine-Tuning

현재 높은 분류 정확도를 가지는 이미지 분류 네트워크는 많이 존재하나 이는 ImageNet 기반으로 훈련이 되어있기 때문에 ImageNet 과 다른 특징을 가지는 데이터에 적용하여 분류할 때는 분류 정확도가 떨어지는 문제점이 있다. 따라서 이러한 새로운 데이터에 대해 분류 정확도를 높이기 위해서는 새로운 데이터 세트를 가지고 기존 네트워크 모델을 재훈련해야 한다.

Fine-Tuning 이란 사전에 ImageNet 데이터로 훈련된 네트워크 모델의 가중치 일부분을 다시 계산하는 것을 의미한다. 자체적으로 구축한 트레이닝 데이터 세트 및 검증 데이터 세트를 기반으로 모델의 일부분을 재훈련하며, 재훈련시 트레이닝 데이터 세트 및 검증 데이터 세트를 높은 정확도로 분류하게 해주는 가중치를 찾아 계산한다.

4. 암세포 현미경 이미지 분류

본 논문에서는 암세포 현미경 이미지 분류를 위해 BreakHis[5] 데이터 세트를 사용하였다. BreakHis 데이터 세트는 현미경으로 촬영한 유방암 세포 이미지를 포함하고 있으며 있으며 확대 비율에 따라 x40, x100, x200, x400 으로 나뉘어져 있다. BreakHis 데이터를 구성하고 있는 이미지 데이터 개수는 아래 표 2 와 같으며 BreakHis 데이터의 이미지 예제는 그림 1 과 같다.

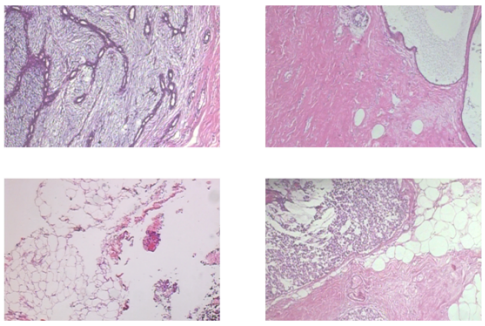

그림 1. BreakHis 데이터 예제

표 2. BreakHis 데이터 구성(단위 : 장)

확대 비율	양성	악성	총합
x40	625	1,370	1,995
x100	644	1,437	2,081
x200	623	1,390	2,013
x400	588	1,232	1,820

암세포 현미경 이미지 분류를 위하여 사전 훈련된 Xception 네트워크 모델을 차용하였으며 상위 4 개의 레이어를 Fine-Tuning, 이후 계산을 통해 얻은 가중치를 기반으로 암세포 현미경 이미지 분류 실험을 수행하였다. 암세포 현미경 이미지 분류 시스템의 출력 값은 각 레이블에 해당될 확률 값이며 softmax 함수를 통해 얻게 된다.

5. 실험 결과

본 논문에서 제안한 분류 기법을 이용하여 암세포 현미경 이미지를 악성 또는 양성으로 구분하는 실험을 수행하였으며, 암세포 이미지 분류 모델 훈련 및 이미지 분류 실험 환경은 아래 표 3 와 같다.

표 3. 암세포 현미경 이미지 분류 모델 훈련 및 실험 환경

운영체제	Ubuntu 16.04 LTS
GPU	GeForce GTX 980 ti
프레임워크	Keras, Tensorflow

Fine-Tuning 및 이미지 분류를 위해 BreakHis 데이터 세트를 Training Set, Validation Set, Test Set 으로 나눴으며, 각 데이터 세트는 이미지 데이터가 겹쳐지지 않도록 구성하였다. Test Set 에 대한 모델 평가도는 표 4 와 같으며 평균적으로 약 96%의 분류 정확도가 나오는 것을 확인할 수 있다.

표 4. 암세포 현미경 이미지 분류 정확도(단위: %)

확대 비율	x40	x100	x200	x400
분류 정확도	96.47	98.30	96.31	91.06

x400 의 확대 비율을 가진 암세포 이미지 같은 경우, 여타 확대 비율을 가진 이미지 데이터들에 비해 분류 정확도가 떨어지는 것을 확인할 수 있는데, 이는 이미지 데이터가 확대되어 이미지 자체에서 갖고 있는 정보가 다른 확대 비율 이미지에 비해 적기 때문으로 추정된다.

6. 결론 및 향후 과제

본 논문에서는 Fine-Tuning 을 통한 Xception 기반 암세포 현미경 이미지 분류 기법을 제안하였다. 제안한 암세포 이미지 분류 기법을 이용하여 암세포 현미경 이미지를 양성 또는 악성으로 분류하였으며 높은 정확도로 분류가 가능한

것을 확인할 수 있었다.

향후 연구 방향으로 암세포 현미경 이미지 뿐만 아니라 MRI 이미지나 X-Ray 이미지 등 다른 의학 이미지를 위한 분류 시스템 및 세부 질병까지 분류가 가능한 의학용 어플리케이션을 개발해야 할 것이다.

참 고 문 헌

[1] Szegedy, Christian, et al. "Going deeper with convolutions." Cvpr, 2015.

[2] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.

[3] Chollet, François. "Xception: Deep learning with depthwise separable convolutions." *arXiv preprint*, 2017.

[4] He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[5] Spanhol, Fabio Alexandre, et al. "Breast cancer histopathological image classification using convolutional neural networks." *Neural Networks (IJCNN), 2016 International Joint Conference on*. IEEE, 2016.