

병변 처리 전이학습을 활용한 당뇨병성 망막증 분류 성능 향상

목다현¹, 양희규², 추현승¹
¹성균관대학교 전자전기컴퓨터공학과
²성균관대학교 컨버전스연구소

{dahyun1025, huigyu, choo}@skku.edu

Improving Diabetic Retinopathy Classification Performance with Lesion Processing Transfer Learning

Dahyun Mok¹, Huigyu Yang², Hyunseung Choo¹
¹Dept. of Electrical and Computer Engineering, Sungkyunkwan University
²Convergence Research Institute, Sungkyunkwan University

요 약

이 연구는 딥러닝을 사용하여 당뇨병성 망막증을 탐지한다. ImageNet 에서 사전 학습된 Swin Transformer 모델을 IDRiD 데이터셋으로 미세 조정하여, 병변 분할 및 당뇨병성 망막증 분류에서 강건한 모델을 개발하였다. 이 모델은 93.2%의 높은 정확도를 기록하며 질병 진단에서의 효율성을 입증한다. 데이터 전처리와 최적 임계값 설정을 통해 모델의 정확성과 신뢰성을 더욱 향상시킨다.

1. 서론

당뇨병성 망막증(Diabetic Retinopathy, DR)은 전 세계적인 실명의 주요 원인 중 하나로, 유병률이 4.8%에 달한다[1]. 초기에는 증상이 거의 나타나지 않지만, 질병이 진행되면서 시력 손실을 유발할 수 있다. 조기 발견과 치료는 심각한 시력 손실을 50-60%까지 예방할 수 있어, DR의 조기 진단과 지속적인 모니터링이 매우 중요하다. 최근 의료 영역에서의 딥러닝 기술 도입은 DR와 같은 병변의 자동 탐지 정확도를 크게 향상시켰으며, 이를 통해 의사의 시간을 절약하고 환자의 경제적, 심리적 부담을 줄일 수 있다.

본 연구에서는 전이 학습을 활용해 미세동맥류, 출혈, 삼출물과 같은 병변에 대한 학습을 추가시켜 DR을 정확히 분류하고자 하며, 공공 데이터셋을 활용하여 모델의 일반화 성능을 극대화한다. ImageNet에서 사전 학습된 Swin Transformer 모델을 미세 조정하여 병변 분할 및 DR 분류에서 높은 성능을 목표로 한다.

2. 방법론

2.1 전이학습

전이학습은 어떤 태스크를 통해 얻은 지식과 정보를 다른 문제를 푸는데 사용하는 방식으로, 본 연구

에서는 ImageNet 데이터셋을 기반으로 사전 학습된 Swin Transformer Tiny (Swin-T) 모델을 사용해 IDRiD 데이터셋의 pixel-level 병변 주석을 학습한다. 이를 통해 Swin-T 모델은 일반적인 이미지 특징을 인식하는 것뿐만 아니라 미세동맥류, 출혈, 삼출물과 같은 작은 병리적 특징을 학습할 수 있게된다. 병변 분할 학습이 끝나면 학습된 인코더를 사용하여 EyePACS 데이터셋에서 병변의 유무를 구별하는 image-level의 학습을 하게 된다. 이렇게 학습된 모델은 이미지 분석에서 중요한 초기 단계의 특징 추출에 필요한 깊이 있는 이미지 특징을 내재하고 있으며, 이를 통해 당뇨병성 망막증의 정확한 미세한 병변 인식 기반을 제공한다.

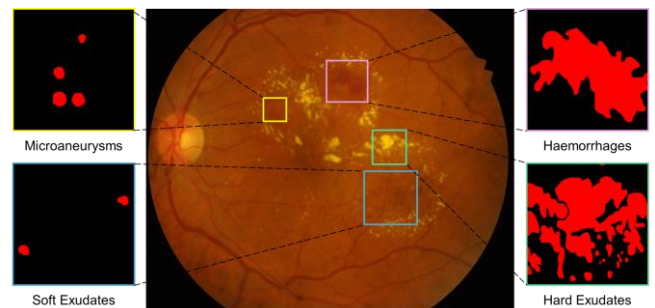


그림 1 IDRiD 데이터셋에서의 병변 맵 예시

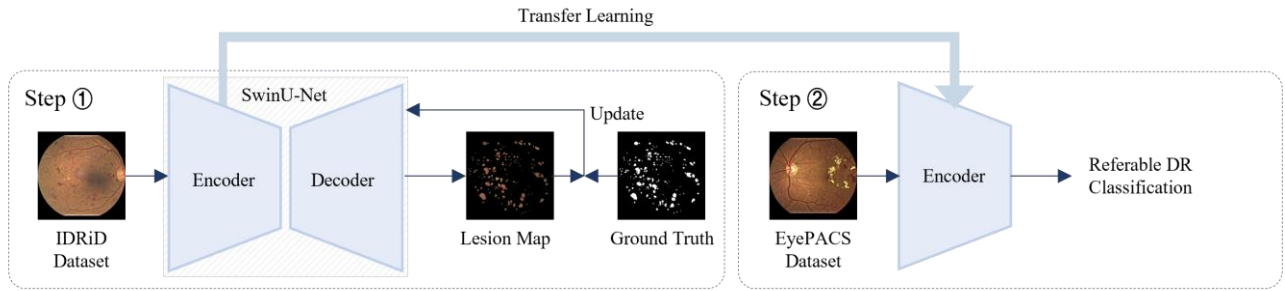


그림 2 제안 모델 구조

2.2 데이터 전처리와 증강기법

이 연구에서는 데이터셋의 일반화 성능을 높이기 위해 데이터 전처리와 증강 기법을 적용했다. 이러한 과정은 의료 영상 데이터의 다양한 특성으로 인해 필요하며, 이러한 전략을 통해 모델이 이미지를 잘 처리할 수 있도록 강건하고 신뢰성 있는 성능을 유지하게 된다. 이를 통해 자동 진단 도구로서 다양한 의료 환경에서 일관되게 높은 성능을 발휘할 수 있게 한다.

이미지 전처리 과정에서는 대조도 향상을 통해 당뇨병성 망막증의 미세한 병변 특징을 더 잘 추출할 수 있도록 하였다. CLAHE(Contrast Limited Adaptive Histogram Equalization) 기법을 사용해, 기존의 histogram equalization 보다 노이즈 과증폭을 방지하며 국소 영역별로 대조도를 선택적으로 향상시켰다. Input 이미지는 640X640 사이즈로 입력되며, 픽셀 정규화를 통해 모델 훈련 시 빠르고 안정적인 수렴을 돕는다.

데이터 증강 과정에서는 회전, 반전과 같은 변형을 도입하여 실제 임상에서 다양한 각도에서 촬영된 이미지를 학습하도록 했다. 밝기, 대조, 채도와 같은 색상 조정도 추가하여, 다양한 임상 환경에서의 이미지 품질 변화를 모델이 처리할 수 있도록 했다.

2.3 모델구조

모델은 Swin-T와 U-Net 프레임워크를 결합하여 병변 분할 학습에 사용된다. 인코더로 사용된 Swin-T는 window shift 메커니즘을 통해 유연한 수용 영역과 효율적인 주의 메커니즘을 제공하며, 이는 미세한 병변을 식별해야 하는 망막 이미지 분석에서 좋은 성능을 낼 수 있다. 이러한 복잡한 패턴과 의존성을 포착하는 능력은 당뇨병성 망막증과 같은 정밀한 병변을 처리하는 데 중요한 역할을 한다.

디코더는 U-Net 구조로, 업샘플링과 컨볼루션 레이어를 통해 출력을 점진적으로 세분화한다. 인코더에서 디코더로 이어지는 skip connection은 다운샘플링 과정에서 손실된 spatial 정보를 복구하는 데 도움을 주며, 이를 통해 미세동맥류와 출혈 같은 당뇨병성

망막증 병변을 정확하게 구분할 수 있다.

3. 실험 및 결과

3.1 데이터셋

본 연구에서는 두 개의 공공 데이터셋을 사용하였다: Kaggle의 EyePACS 데이터셋[2]과 Indian Diabetic Retinopathy Image Dataset(IDRiD)[3]이다. EyePACS 데이터는 분류 작업에, IDRiD 데이터는 병변 분할 작업에 사용되었다. IDRiD 데이터셋은 81개의 양성 이미지를 포함하며, 학습, 테스트에 각각 54, 27개 데이터를 사용하였다. EyePACS 데이터셋은 총 88,702장으로, 이 중 35,126장은 학습, 10,906장은 검증, 42,670장은 테스트에 활용되었다. EyePACS 데이터셋은 당뇨병성 망막증(DR)을 5단계로 분류하지만, 본 연구에서는 이진 분류를 위해 referable DR 방식을 사용하여, 0(음성)으로 no DR 및 mild DR, 1(양성)으로 moderate DR, severe DR, PDR를 relabeling 하였다.

이와 같은 데이터셋 구성을 통해 다양한 환자군에 대한 모델의 일반화 능력을 검증하였으며, 공공 데이터셋을 사용하여 연구의 일반성을 높이고, 향후 실험 재현성 또한 보장하였다. 이러한 접근 방식은 의료 영상 분석에서 다양한 실제 환경에 적용 가능한 높은 신뢰성을 제공한다.

3.2 학습

학습은 총 50번의 epoch을 통해 진행되며, Adam 옵티마이저를 사용하였다. Adam 옵티마이저는 최소한 그라디언트 문제를 잘 처리하고 학습률 조정에 적합하다. 초기 학습률은 0.001로 설정되었으며, 검증 손실 지표에 따라 이후 학습률을 조정하여 과적합을 방지하였다. 이를 통해 빠른 수렴과 최소 손실 영역을 놓치지 않는 균형을 유지할 수 있었다.

모델 성능은 Intersection over Union (IoU), Dice coefficient, 정확도, 민감도, 특이도, AUC 등의 주요 지표를 사용하여 모델의 성능을 정량적으로 평가하였다. 특히 AUC-ROC는 모델의 분별력을 종합적으로 평가하는 데 사용되었으며, 민감도와 특이도를 극대화할

수 있는 최적의 임계값을 선택하여 모델의 진단 정확도를 더욱 향상시켰다.

3.3 결과

실험 결과는 93.2%의 정확도, 94.8%의 AUC-ROC, 85.9%의 민감도, 93.0%의 특이도를 나타내며, 제안 scheme 의 강건함과 효율성을 입증하였다. ROC 곡선을 분석하여 다양한 임계값에 따른 True Positive Rate(TPR)와 False Positive Rate(FPR)을 평가하였으며, 최적의 임계값을 선택하여 모델이 임상 환경에서 DR 단계를 정확하게 구별할 수 있도록 하였다.

또한, 학습 시 높은 드롭아웃 비율을 적용한 경우 모델의 분류 성능이 향상된 것을 확인하였다. 이는 훈련 데이터의 다양성과 복잡성이 증가하면서 모델이 보이지 않는 샘플에 대해 더 잘 일반화할 수 있게 된 것으로, 특히 의료 영상 데이터셋에서의 변동성을 처리하는 데 있어 높은 성능을 보였다.

4. 결론

본 연구에서는 고급 딥러닝 기술을 활용하여 당뇨병성 망막증의 참조 가능한 분류를 실험하였다. Swin-T 와 U-Net 아키텍처의 결합과 최적화된 임계값 설정을 통해 높은 정확도의 강건한 모델을 개발하였으며, 전이 학습 및 특화된 데이터 전처리가 자동 의료 진단의 정확성과 신뢰성을 크게 향상시킬 수 있음을 입증하였다. 향후 연구에서는 더 다양한 데이터셋을 사용하여 모델의 일반화 능력을 개선하고, Grad-CAM 을 도입하여 모델의 의사결정 과정을 시각화할 계획이다.

Acknowledgement

본 연구는 정부(과학기술정보통신부)의 재원으로 4 단계 BK21 사업(50%), 정보통신기획평가원의 ICT 명품인재양성사업(RS-2020-II201821, 25%)과 안저영상분석기반 뇌혈관 질환(뇌졸중) 예측 모델 개발 지원사업(RS-2024-00459512, 25%)의 지원을 받아 수행된 연구임.

참고문헌

- [1] H. Buch et al., "Prevalence and causes of visual impairment according to world health organization and United States criteria in an aged, urban Scandinavian population: the Copenhagen city eye study," Ophthalmology, Vol. 108, No. 12, pp. 2347–2357, 2001.
- [2] EyePACS, "Diabetic retinopathy detection," 2015.
- [3] IDRID, "Indian diabetic retinopathy image dataset," 2018.