

웨이블릿-GLCM 을 이용한 EfficientNet 기반 흉부 X-ray 폐렴 분류 성능 평가

이재훈, 유재천*

성균관대학교 전자전기컴퓨터공학과

jaehoonlee@skku.edu, *yoojc@skku.edu

Performance Evaluation of EfficientNet-Based Chest X-ray Pneumonia Classification Using Wavelet-GLCM Features

Lee Jae Hoon, Yoo Jae Chern*

Department of Electrical and Computer Engineering, Sungkyunkwan University

요 약

본 논문은 흉부 X-ray 영상 기반의 폐렴 자동 분류 문제에서 미세조직 질감 패턴의 민감도 이슈에 대하여 웨이블릿 변환-기반 GLCM(gray-level co-occurrence matrix) 피처를 기존 CNN 구조와 융합하는 방식을 제안하고, 이의 성능을 평가하였다. Kaggle Chest X-ray Pneumonia 데이터(정상 1,583, 폐렴 4,273)를 환자 ID 단위로 8:1:1(학습·검증·시험) 분할한 뒤, 히스토그램 평활화·300 × 300 리사이즈·Albumentations 증강을 적용하고 Sym4 3-level 2-D DWT 로 LH·HL·HH 서브밴드를 추출하여 GLCM 통계량을 계산하였다. 베이스라인으로 EfficientNet-B3 모델을, 융합 구조로는 CNN 과 Wavelet-GLCM 피처를 병렬 입력으로 사용하는 경량 FusionNet 을 제안하였다. 그 결과, 베이스라인 모델은 AUC 0.9983, 정확도 98.53 %, 정밀도 0.9890, 재현율 0.9912, F1-score 0.9901 의 우수한 성능을 보였으며, Fusion 구조는 AUC 0.9992, 정밀도는 높았으나 재현율이 다소 낮아 F1-score 는 0.9820 으로 측정되었다.

I. 서 론

폐렴은 전 세계 소아 사망률 1 위의 감염 질환으로서, 흉부 X-ray 영상 판독의 신속성과 정확성은 임상적·사회적 파급 효과가 매우 크다. 최근 EfficientNet, ResNet 등의 합성곱 신경망(CNN)이 우수한 분류 성능을 보여 왔으나, 미세 조직(texture) 패턴에 대한 민감도가 낮아 정상·비정상 경계가 흐려지는 한계가 보고되고 있다. 웨이블릿 변환은 영상의 다중 주파수 성분을 분리하고, GLCM 은 공간적 공존 빈도를 통계량으로 요약함으로써 이러한 질감 정보를 정량화할 수 있다 [3]. 본 연구에서는 EfficientNet-B3 백본과 웨이블릿-GLCM 48 차원 벡터를 지연 융합(late-fusion)한 FusionNet 을 설계·검증하였다. 동일한 학습 조건에서 B3 단독 대비 분류 성능, 파라미터 증분, 오류 유형(혼동행렬)을 정량 비교하여 질감 보강의 실효성을 평가한다.

II. 본론

2.1 데이터셋 구성 및 전처리

Kaggle Chest X-ray Pneumonia 데이터(정상 1,583, 폐렴 4,273 장)를 환자 ID 단위로 8:1:1(학습·검증·시험)로 재분할하였다 [1]. 영상은 히스토그램 평활화 후 300 × 300 으로 리사이즈하고, Albumentations 기반 증강(수평 반전 50 %, 회전 $\pm 10^\circ$, 밝기·대비 $\pm 15\%$)을 적용하였다 [5]. 이어 Sym4 3-level 2-D 웨이블릿 분해로 LH·HL·HH 서브밴드를 추출하였으며 [4], 그림 1 에는 전처리된 원본 영상과 세 서브밴드가 2 × 2 패널로 제시된다. 각 서브밴드에서 64 × 64 패치를 선택해 16-gray 양자화를 수행한 뒤 GLCM($d = 1$, $\theta = 0^\circ$)을 계산하였으며, 정규화된 행렬은 그림 2 에 시각화하였다.

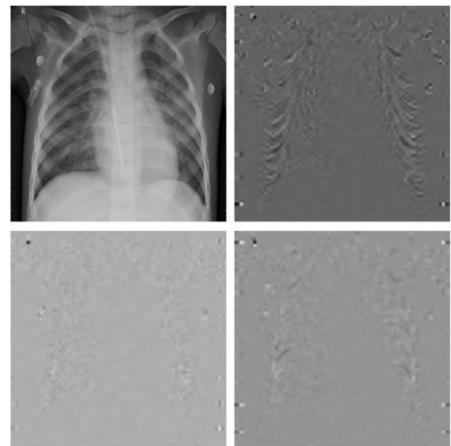


그림 1. 다중 스케일 웨이블릿 분해 예시

GLCM Matrix (from Wavelet Patch)

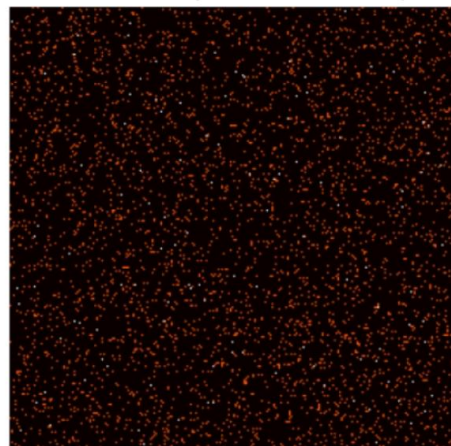


그림 2. 16-Gray GLCM 정규화 행렬

2.2 모델 구조 및 학습 설정

모델은 EfficientNet-B3 로 설정되었고 [2], 융합 모델은 CNN 출력과 Wavelet-GLCM 피처를 Fully Connected Layer 에서 병합하는 구조로 설계되었다. 손실 함수는 BCEWithLogitsLoss, 옵티마이저는 AdamW 를 사용하였으며, 러닝레이트는 $3e-4$, 에폭 수는 20 으로 설정하였다. 모든 학습은 PyTorch 기반으로 수행되었으며, FP16 학습을 위해 autocast 를 적용하였다.

2.3 실험 결과 및 시각화 비교

시험 세트 성능은 B3 Baseline 이 Accuracy 0.985, Precision 0.989, Recall 0.991, F1 0.990, ROC-AUC 0.998 을, FusionNet 이 Accuracy 0.974, Precision 0.998, Recall 0.967, F1 0.982, ROC-AUC 0.999 를 각각 달성하였다. FusionNet 은 False Positive 를 5 \rightarrow 1 로 감소시켜 정밀도를 향상시켰으나, False Negative 가 4 \rightarrow 15 로 증가해 재현율이 낮아지는 trade-off 가 관찰되었다.

그림 3 은 좌측 열(B3 Baseline)과 우측 열(FusionNet)로, 첫 행은 ROC Curve, 두 번째 행은 Precision-Recall Curve, 세 번째 행은 Confusion Matrix 를 나타낸다. 이를 통해 FusionNet 이 전반적인 판별 능력(ROC-AUC, PR-AUC)과 특이도(정밀도)는 높지만, 민감도(재현율)에서 손실이 발생함을 한눈에 확인할 수 있다.

Model	Accuracy	Precision	Recall	F1	AUC
B3	0.985	0.989	0.991	0.990	0.998
FusionNet	0.974	0.998	0.967	0.982	0.999

표 1. 모델별 정량 성능 지표

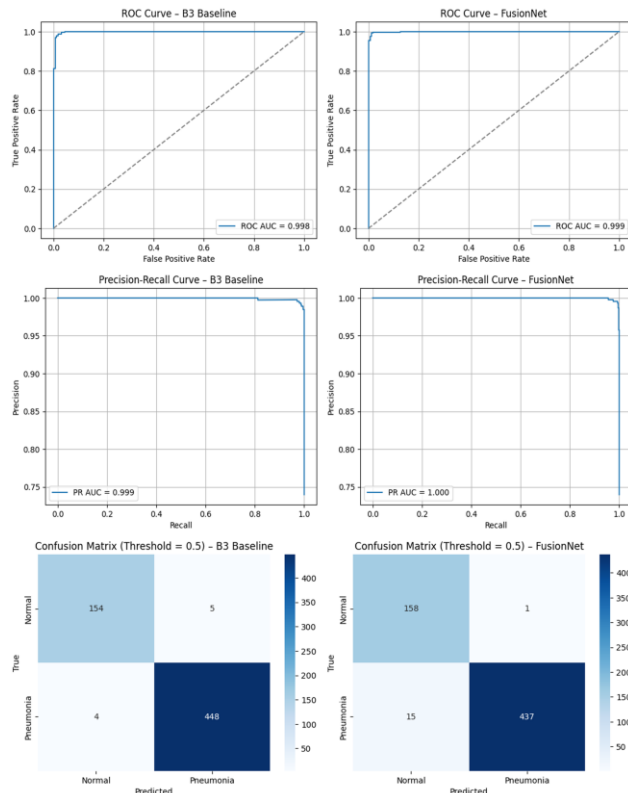


그림 3. B3 Baseline vs. FusionNet 시험 성능 비교: ROC-PR 곡선 및 혼동행렬

III. 결론

본 연구에서는 웨이블릿 기반 질감 피처를 CNN 구조에 융합하여 폐렴 분류 성능을 향상시키는 방법을 제안하였다. 데이터셋 구성의 신뢰성을 확보하기 위해 환자 ID 기반 재분할을 적용하였으며 [1], 이는 과적합 방지 및 일반화 성능 확보에 기여하였다. 베이스라인인 EfficientNet-B3 는 AUC 0.9983, 정확도 98.53 %로 매우 우수한 성능을 보였으며 [2], FusionNet 은 AUC 0.9992 로 더 높은 구분력을 보였으나, 재현율은 다소 낮아 정확도 측면에서는 오히려 소폭 하락하였다. 이는 wavelet 기반 피처가 정상/비정상 경계를 강화하는 데 효과적일 수 있으나, 클래스 민감도 균형 측면에서는 조정이 필요함을 시사한다. 또한 융합 구조가 질감 기반 피처의 기여를 효과적으로 반영하지 못했을 가능성도 존재하며, 단순한 FC 기반 융합보다는 attention, gating 또는 multi-branch 구조 도입이 필요한 것으로 판단된다 [4]. 향후 후속 연구에서는 이러한 구조적 보완과 함께 threshold 조정에 따른 성능 변화도 함께 분석할 예정이다.

참 고 문 헌

- [1] D. S. Kermany, K. Zhang, P. M. Lim, M. Goldbaum *et al.*, "Chest X-Ray Images (Pneumonia)," Kaggle Data Set, ver. 3, 2018. [Online]. Available: <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia> (Accessed: 12 Jun. 2025).
- [2] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proc. 36th Int. Conf. Mach. Learn. (ICML)*, Long Beach, CA, USA, Jun. 2019, pp. 6105–6114, doi: 10.48550/arXiv.1905.11946.
- [3] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973, doi: 10.1109/TSMC.1973.4309314.
- [4] S. Lyu, J. Wang, Z. Chen, and X. Yu, "Pulmonary Disease Classification on Chest X-Ray Images Using Wavelet-Based Texture Features," *Biomedical Engineering Online*, vol. 21, no. 1, pp. 1–17, Jan. 2022, doi: 10.1186/s12938-022-00937-8.
- [5] A. Buslaev, A. Parinov, E. Khvedchenya *et al.*, "Albumentations: Fast and Flexible Image Augmentations," *Information*, vol. 11, no. 2, Art. ID 125, Feb. 2020, doi: 10.3390/info11020125.