

Artificial Intelligence Term-Project

Cell Detection and Verification Model for Automated Cell Counting

2024.12.19.Thu

부산대학교 의생명융합공학부

데이터사이언스전공

김진

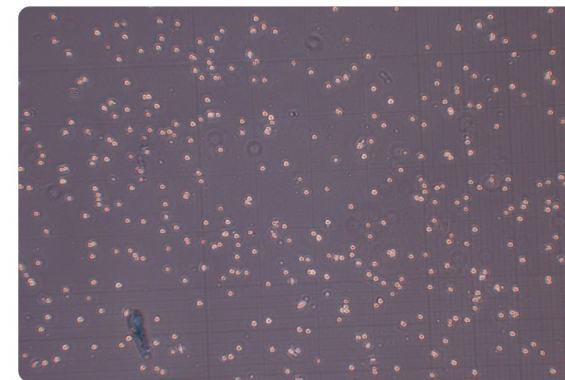
Contents

- ① Project Background and Necessity
- ② Initial Approach and Limitations
- ③ Model Architecture
- ④ Data Preprocessing
- ⑤ Training Process
- ⑥ Performance Evaluation and Comparison
- ⑦ Real Testing and Applicability
- ⑧ Conclusion and Suggestions

Cell Detection and Verification Model for Automated Cell Counting

Project Background and Necessity

- 의생명공학실험에서의 세포 수 세는 과정의 비효율성
 - ❖ 수작업으로 세포를 세는 데 많은 시간 소요
 - ❖ 반복 작업으로 인한 오류 발생 및 정확도 저하
- 세포 수를 측정해주는 모델 개발의 필요성을 느낌



세포가 몇개 있는지 알려줘



파일을 분석해서 세포의 개수를 확인하려면 이미지 처리 도구를 사용할 수 있습니다. 업로드하신 이미지 파일을 분석해보겠습니다. 잠시만 기다려 주세요.

(파일을 분석 후 결과를 알려드리겠습니다.)

이미지가 성공적으로 불러와졌으며, 해상도는 **1536x2048** 픽셀입니다. 이제 이미지 내 세포를 분석하여 개수를 추정하겠습니다. [-]

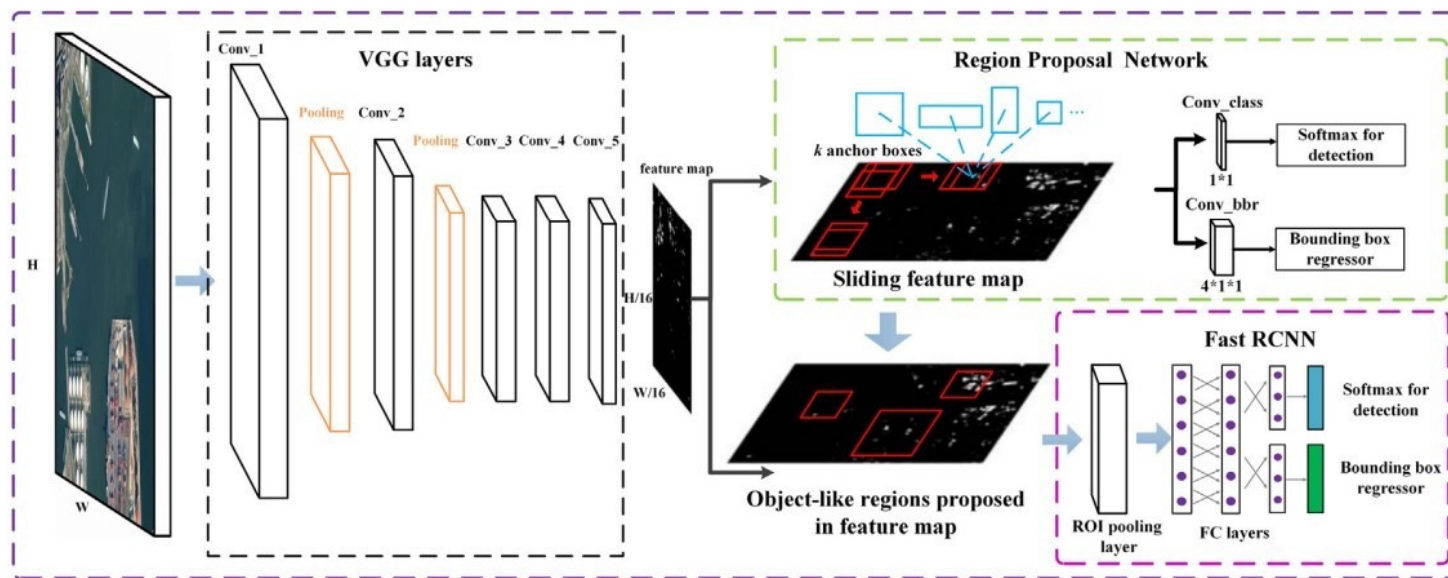
이미지에서 **1개**의 세포가 감지되었습니다. 추가 분석이나 세포의 특징에 대한 세부 정보가 필요하시면 말씀해 주세요! [-]

???

Cell Detection and Verification Model for Automated Cell Counting

Initial Approach and Limitations

- Faster-RCNN 기반으로 이미지에 존재하는 모든 세포 탐색 시도
 - 모델이 세포를 탐지하지 못하고 빈 결과를 반환
 - ❖ 훈련 이미지 수가 적고 객체 수가 과도하게 많아 학습이 어려운 것으로 예상 됨
 - ❖ 기존 Faster-RCNN 모델의 한계로 한번에 100개가 넘어가는 객체를 처리하지 못하는 것으로 예상 됨

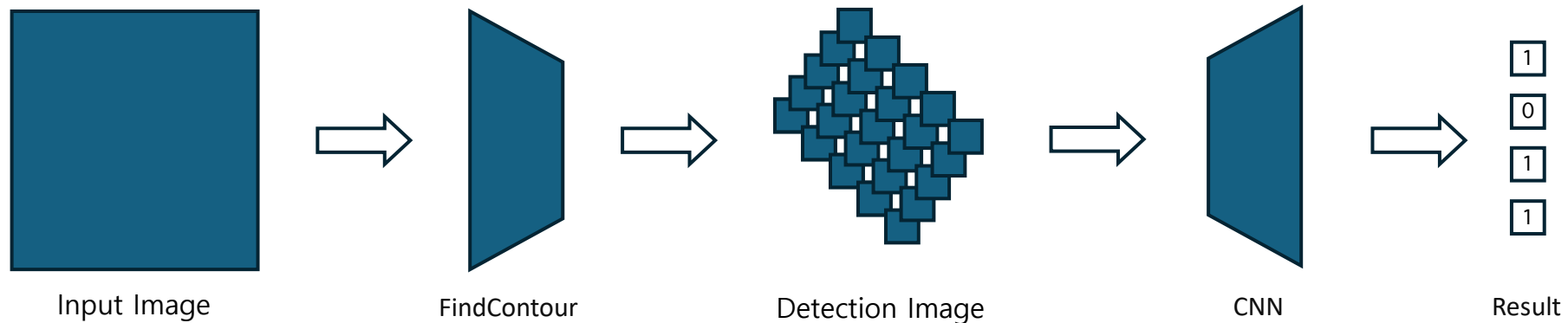


Faster-RCNN Architecture

Cell Detection and Verification Model for Automated Cell Counting

Model Architecture

- 모델을 두 파트로 나눔
 - 객체 인식
 - ❖ OpenCV의 **FindContour**를 활용해 이미지 내 모든 객체의 경계를 탐지
 - 세포 여부 판별
 - ❖ **CNN 이진 분류 모델**을 사용해 탐지된 객체가 세포인지 아닌지 판별

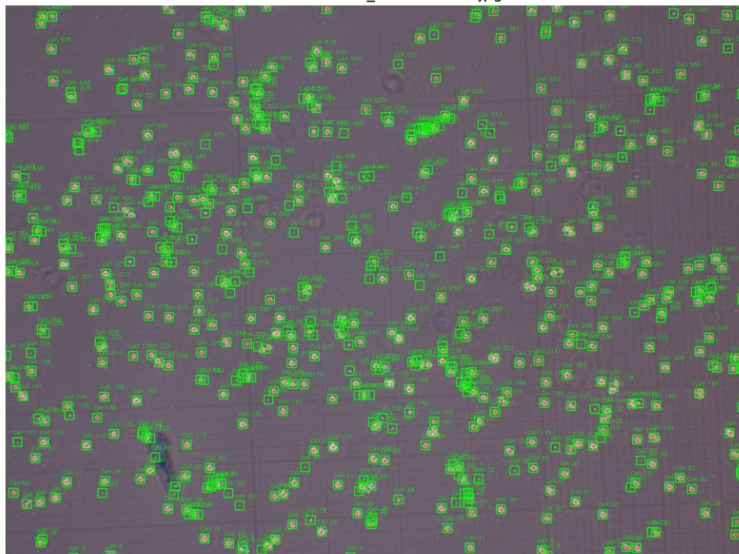


Cell Detection and Verification Model for Automated Cell Counting

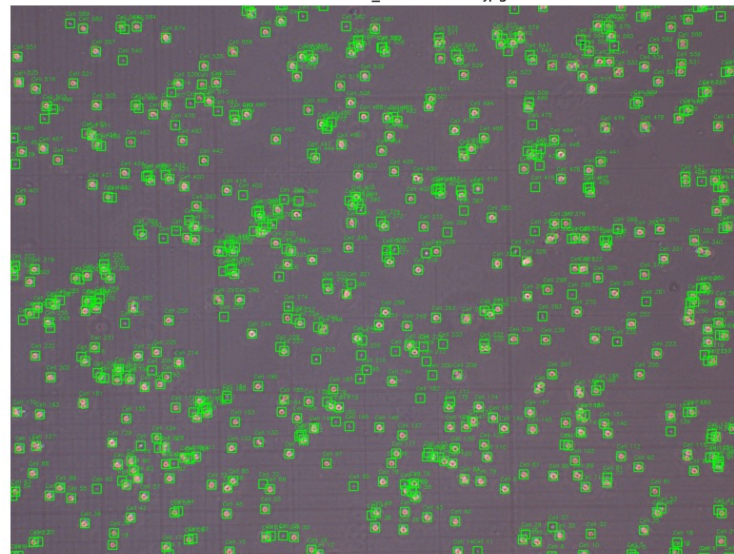
Data Preprocessing

- 객체 탐지 및 이미지 저장
 - ❖ OpenCV의 findContour를 활용해 원본 이미지에서 객체가 있을 법한 위치를 모두 탐지
 - ❖ 탐지된 객체를 별도의 이미지 파일로 저장

Visualization: 1_visualization.jpg



Visualization: 2_visualization.jpg



Data Preprocessing

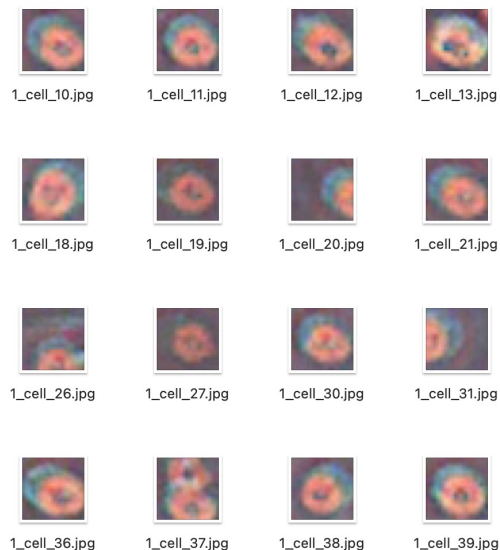
- 객체 라벨링

- ❖ 잘 인식된 이미지와 잘못 인식된 이미지를 **직접 확인하여** 별도의 파일에 저장

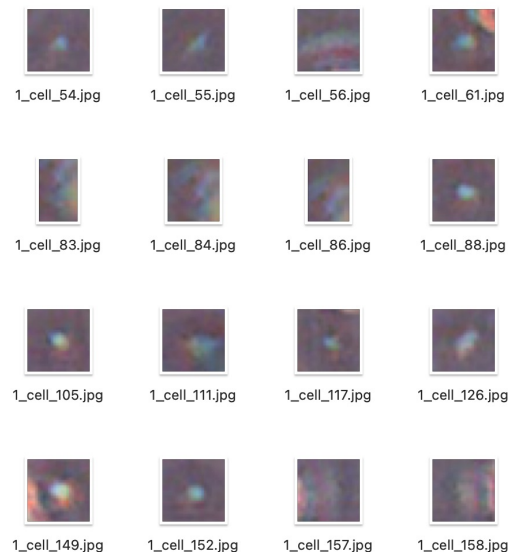
- 데이터 증강 수행

- ❖ 원본 데이터를 10배 증강하여 데이터셋 크기 확장

- 좌우 반전, 랜덤 회전, 밝기 및 대비 조정, 위치 변환



올바르게 인식한 객체

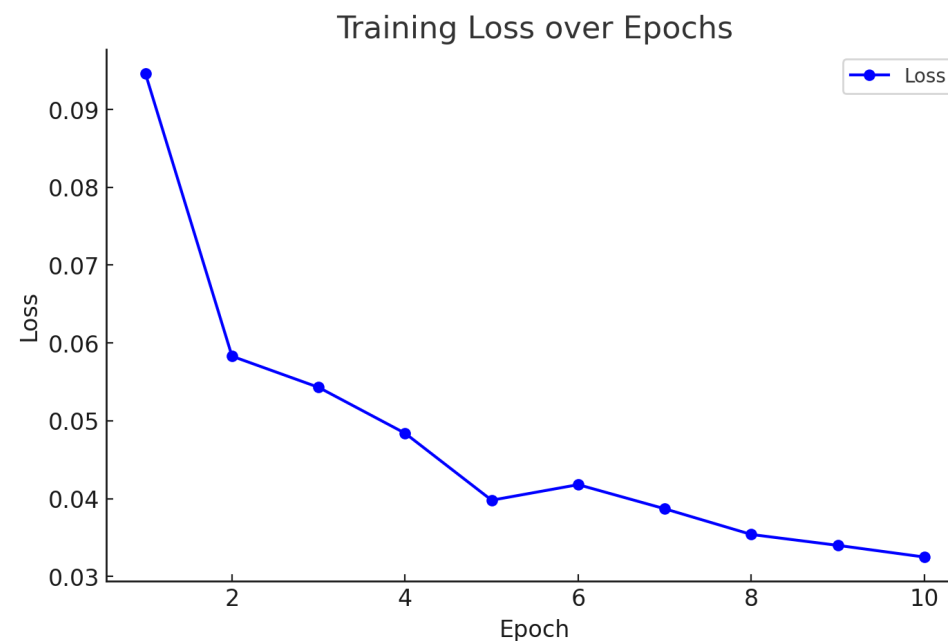
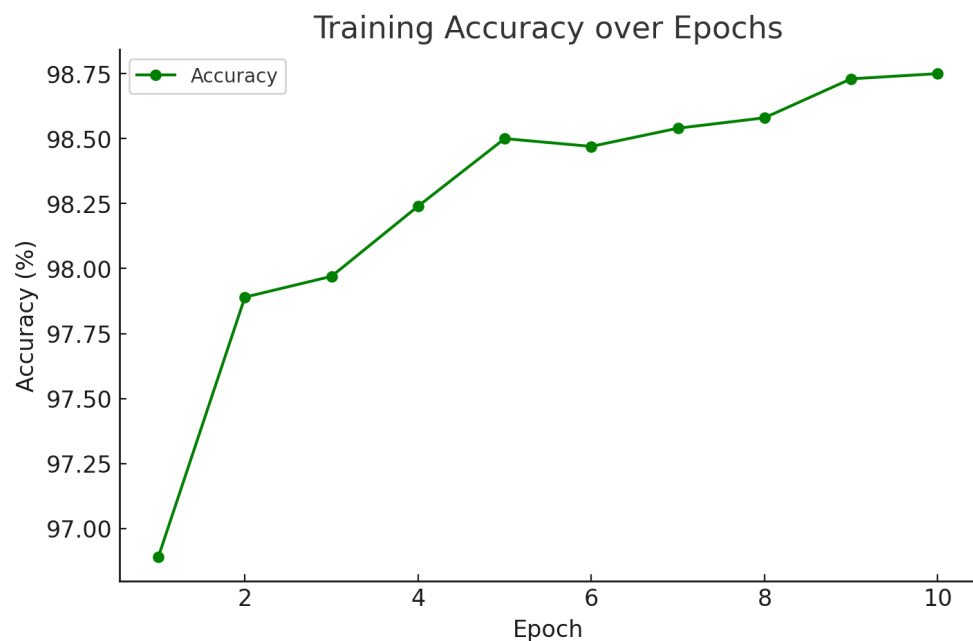


잘못 인식한 객체

Cell Detection and Verification Model for Automated Cell Counting

Training Process

- CNN 모델 학습 설정
 - ❖ ResNet18 기반 모델 사용
 - ❖ CrossEntropyLoss를 손실 함수로 사용
 - ❖ Optimizer는 Adam 학습률은 0.001 설정



Cell Detection and Verification Model for Automated Cell Counting

Performance Evaluation

- 모델 성능 평가 결과

- ❖ Accuracy: 99%

- ❖ Precision / Recall / F1-Score: Alive 클래스 1.00, Misclassified 클래스 0.97

- 한계점

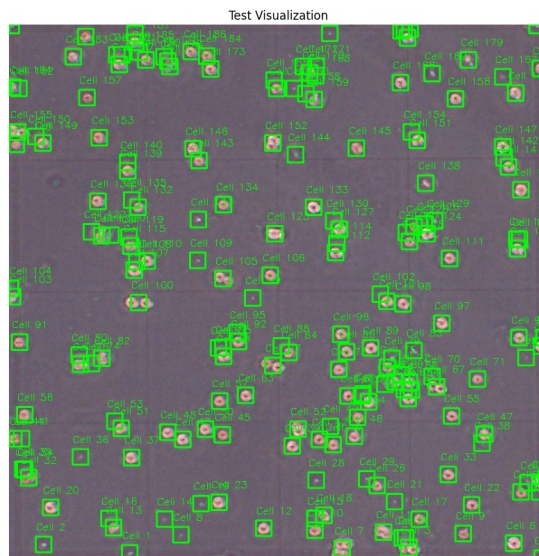
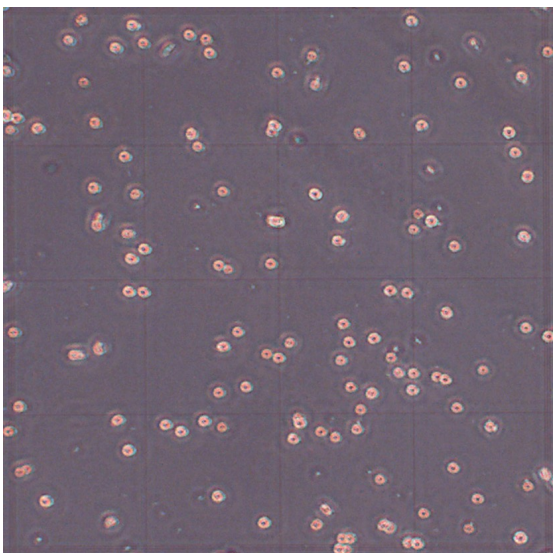
- ❖ 검증 데이터가 훈련 데이터에서 증강된 데이터이기 때문에 실제 일반화 성능을 평가하기에는 한계가 있음

	precision	recall	f1-score	support
alive	1.00	1.00	1.00	4391
misclassified	0.97	0.97	0.97	689
accuracy			0.99	5080
macro avg	0.98	0.98	0.98	5080
weighted avg	0.99	0.99	0.99	5080

Cell Detection and Verification Model for Automated Cell Counting

Real Testing and Applicability

- OpenCV의 findCountour 결과
 - ❖ 4x4 세포 이미지에서 197개의 객체를 탐지
- CNN 모델 판별 결과
 - ❖ 탐지된 197개 중 124개를 세포로 판단
- 실제 세포 수 검증 결과 119개로 측정 되었으며, 약간의 오차는 있지만 실제 사용에도 무리가 없음



테스트 데이터 분류 결과:
alive: 124개
misclassified: 73개

Conclusion and Suggestions

- 데이터셋 수집의 부족
 - 실험 초기에 다양한 데이터셋을 충분히 확보하지 못해 모델 학습에 한계가 있었음
- 다양한 환경 및 조도에서의 검증 부족
 - 현재 모델은 단일 환경에서만 테스트되었으며, 다른 조도나 환경에서의 성능을 검증하지 못함
- 모델 학습 및 일반화 여부 불확실
 - ❖ 검증 데이터가 훈련 데이터에서 증강된 것이기 때문에 과적합 가능성이 존재, 일반화 성능 평가가 부족함.