

폐암 진단을 위한 영상 분석 기반 암세포 검출 시스템 구현

이주형¹, 이민아¹, 권용현², 류병석³, 김영균⁴

¹강원대학교 의생명공학전공

²한양대학교 인텔리전스컴퓨팅학과

³연세대학교 화공생명공학과

⁴융합소프트웨어랩

wngud735@naver.com, otpy0528@naver.com, mrkyh380@hanyang.ac.kr,

bsryu@yonsei.ac.kr, ygkim-2004@hanmail.net

Implementation of Image Analysis based Cancer Cell Detection System for Lung Cancer Diagnosis

Juhyeong Lee¹, MinA Lee¹, YongHyun Kwon², Byeongseok Ryu³, YoungGyun Kim⁴

¹Dept. of Medical Biotechnology, Kangwon National University

²Dept. of Intelligence Computing, Hanyang University

³Dept. of Chemical and Biomolecular Engineering, Yonsei University

⁴Convergence Software Lab.

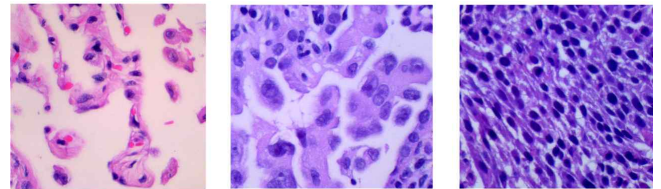
요 약

본 논문에서는 국내 사망 원인 1위 질환인 암 중 가장 큰 비중을 차지하는 폐암의 암 오진율 감소 및 정밀 진단을 위해 폐암세포를 검출 및 계수 할 수 있는 시스템을 구현하였다. 사용자가 관심 영역을 지정하면 H&E 염색 방식을 사용한 폐암세포 전처리 과정을 거쳐 검출 및 계수 할 수 있다. 본 시스템을 통해 병리학자가 단 시간에 폐암세포 검출 및 계수하여 객관적 진단 도구로 활용할 수 있으며, 디지털 기술과 융합하여 정밀 의료에 크게 기여할 수 있을 것으로 기대된다.

1. 서론

국내 사망 원인 1위 질환인 암 중 기관지에 발생하는 폐암으로 인한 사망자는 인구 10만 명 당 36.8명으로 가장 큰 비중을 차지하였다[1]. 2020년 기준 폐암의 발병률은 2만 8,949명(11.7%)으로 가장 높았으며, 생존율은 36.8%로 낮은 수치를 나타내었다[2].

폐암은 크게 소세포 암종과 비소세포 암종으로 분류되며, 그중 폐암 발병의 약 80%를 차지하는 비소세포 암종은 선암종, 편평세포암종, 대세포암종 등으로 구분된다[3]. 대개 암세포는 핵의 크기가 커지고 염색질 양이 증가하여 염색이 짙게 되는 과다 염색질(Hyperchromasia)을 보인다. 신체의 다른 부위로 침범하지 않고 천천히 자라는 양성종양과 달리, 악성 종양인 암은 빠르게 성장하여 림프관을 통해 전이되기 때문에 양성종양과 악성을 구분하는 것은 치료 및 예후에 있어 매우 중요하다[4]. 폐의 양성 조직과 악성 조직인 비소세포암종의 선암, 편평세포암을 H&E 염색한 슬라이드에서 확인할 수 있는 형태학적 특성은 그림1과 같다[5].



a 양성 조직 b 선암종 c 편평세포암종

(그림 1) 양성 조직과 악성 조직 비교 사진

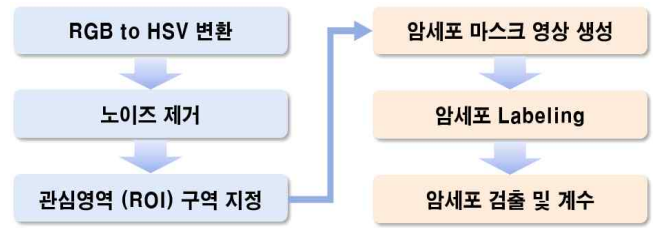
양성 조직(그림 1 a.)과 악성 조직은 핵 크기와 염색질의 양 차이를 통해 육안으로 비교할 수 있다. 폐 선암종(그림 1 b.)은 체액을 분비하는 기능을 하는 선세포에서 폐포 벽의 표면을 따라 증식하고, 편평세포암종(그림 1 c.)은 염색질의 구조를 파악할 수 없을 정도로 농축되고 불규칙한 핵과 진하게 각화된 세포질을 가진 울챙이의 형태를 보인다[6]. 종양마다 고유의 형태학적 특성을 가지고 있으나, 그림 1의 양성 조직처럼 암과 형태학적 특성이 비슷할 경우 구별이 어려워 영상판독에 오류가 발생할 수 있다. 이처럼 전문가의 병리학적 판독 과정에서, 주관적 판단에 의한 의학적 오류, 전문 인력 부족 등의 문

제와 환자의 상태에 따라 항상 정확한 상태를 유지할 수 없는 반정량적(semi-quantitative) 특성은 암 오진 발생의 주요 원인이 된다[7]. 전문가의 반정량적 소견에 대한 오류를 보완하기 위해 병리학적 데이터를 11단계 공정으로 획득하던 전통적 병리 검사법에서, 4차 산업혁명의 핵심 기술인 ‘ICT·빅데이터·AI’를 도입하여 6단계로 간소화된 디지털 병리학으로 발전하였다. 이를 통해 환자 맞춤형 의료 서비스 제공, 의료 인력의 업무 강도 완화를 가능하게 했을 뿐만 아니라 디지털 병리 시스템을 기반으로 한 소프트웨어 개발로 암 환자 정밀 진단 기여 등의 효과를 거둘 수 있을 것으로 기대된다. 따라서 컴퓨터를 이용한 병리 조직 영상 분석 기술은 보다 정확하고 객관적인 정보 추출, 해석을 가능하게 하며 기존의 현미경을 통한 조직 병리 진단 방법의 낮은 재현성 문제를 해결할 수 있고, 보조 진단 시스템으로 사용 시 암 진단의 정확성을 높일 수 있다[8]. 앞서 언급한 점을 바탕으로 암 오진율이 가장 높았던 폐암 조직의 H&E 염색 슬라이드에서 양성종양과 악성종양의 형태학적 특성과 색상 차이를 구분하여 폐암세포를 검출 및 계수하는 정밀 분석 소프트웨어를 구현하였다.

2. 구현

본 연구는 Python 3.9 환경에서 OpenCV 4.5.3 라이브러리를 사용하였으며, GUI는 PyQt5를 기반으로 구현하였다. 분석에는 표본 영상 관찰에 가장 보편적인 Hematoxylin(헤마톡실린) & Eosin(에오신) 염색 처리된 영상을 사용하였다[9]. 영상 전처리 과정의 첫 번째로, 염색된 특정 색상을 추출하기 위해 영상의 최소 단위인 픽셀의 색 속성을 모두 참고해야 하는 RGB 색 공간에서, 색상(Hue), 채도(Saturation), 명도(Value) 조합으로 순수한 색 정보를 가진 색상의 특성을 활용하여 보다 편리하게 색 추출이 가능한 HSV 색 공간으로 변환하였다. HSV 색 공간은 조명에 민감하지 않는 장점이 있다. 두 번째로 영상의 노이즈 제거 단계로, 이미지 내의 유사한 윈도우 영역들의 평균 값을 사용하는 Non-Local Means Denoising 방법으로 작업 시간은 조금 더 요구되나 우수한 결과를 갖는 FastNIMeansDenoising 노이즈 제거 방식을 사용하였다[10]. 다음으로 분석에 필요한 영역만을 추출하기 위해 관심 영역(ROI)을 지정할 수 있도록 하였으며, 이후 계수에 필요한 세포 핵을 세포질과 분리하기 위해 마스크 영상을 생성하여 전처리를 진행하였

다. 그 후 해당 영역에 대해 암세포 검출 및 계수를 진행하였다(그림 2).



(그림 2) 전체 흐름도

2.1. GUI

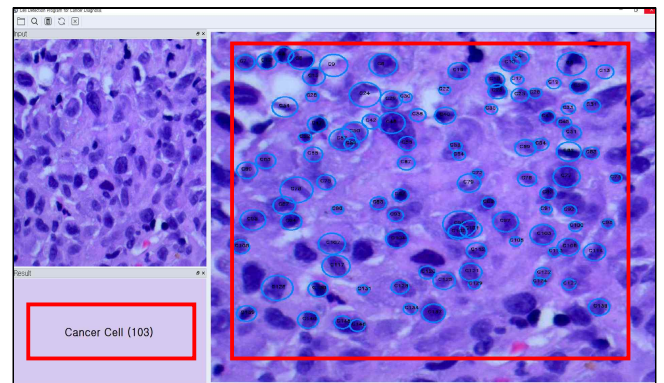
GUI를 통해 사용자는 병리학적 슬라이드에서 분석할 관심 영역(ROI)을 지정하여 암세포의 검출 및 계수를 진행할 수 있다. 검출된 암세포의 데이터가 사용자에게 제공되며, 특정 영역을 확대하여 자세하게 관찰할 수 있다(그림 3, 4).



(그림 3) GUI 초기화면

2.2 암세포 검출 및 계수

ROI 지정 관찰 범위 내에서 암세포만 Labeling 하여 검출 및 계수를 진행하였다(그림 4).



(그림 4) 암세포 Labeling, 검출 및 계수

3. 결과

Confusion matrix를 바탕으로 각 암종 별 10개 영상의 암세포를 직접 계수하여 전체 세포의 수를 구한 뒤, 알고리즘의 성능을 예상하기 위해 정밀도

(Precision), 재현율(Recall) 지표를 계산하였다. 20개의 암종 영상이 불균형한 데이터를 가지고 있기 때문에 정확도를 구하는 대신 F1 점수 지표로 나타내었다. 주어진 지표들을 계산하기 위해 TP(True Positive), FP(False Positive), FN(False Negative) 3가지의 값을 사용하여(표 1) 정밀도와 재현율 그리고 F1 점수를 계산하였다(표 2).

<표 1> Confusion Matrix 사용 값

	직접 계수	프로그램 계수
TP	암세포로 인식	암세포로 인식
FP	암세포로 인식 하지 않음	암세포로 인식
FN	암세포로 인식	암세포로 인식하지 않음

<표 2> 정밀도, 재현율, F1 점수 계산 공식

정밀도	$\frac{TP}{TP + FP}$
재현율	$\frac{TP}{TP + FN}$
F1 점수 *	$\frac{2 \times \text{정밀도} \times \text{재현율}}{\text{정밀도} + \text{재현율}}$

F1 점수* : 1에 수렴할 수록 정확함

선암종 6번 영상에서 정밀도 100%, 재현율 93%, F1 점수 0.96으로 가장 높았고, 편평세포암종은 2번 영상에서 정밀도 99%, 재현율 86%, F1 점수 0.92로 가장 높았다(표 3). 선암종 영상에선 평균 정밀도 99%, 재현율 77%, 그리고 0.86의 F1 점수를 나타내었다. 편평세포암종의 경우 평균 94%의 정밀도, 82%의 재현율 및 0.87의 F1 점수를 나타내었다.

<표 3> 암종별 정밀도, 재현율 및 F1 점수 비교

암종	암종별 정밀도, 재현율 및 F1 Score 비교							
	번호	all cells	TP	FN	FP	정밀도	재현율	F1 점수
선암종	1	129	69	60	0	1.00	0.53	0.69
	2	63	51	12	0	1.00	0.81	0.90
	3	91	68	23	0	1.00	0.75	0.86
	4	134	81	52	1	0.99	0.61	0.75
	5	75	49	26	0	1.00	0.65	0.79
	6	46	43	3	0	1.00	0.93	0.96
	7	52	47	5	0	1.00	0.91	0.95
	8	65	59	6	0	1.00	0.91	0.95
	9	60	48	7	5	0.91	0.87	0.89
	10	117	79	38	0	1.00	0.68	0.81
편평세포암종	1	35	26	6	3	0.90	0.81	0.85
	2	110	94	15	1	0.99	0.86	0.92
	3	38	31	6	1	0.97	0.84	0.90
	4	139	92	46	1	0.99	0.67	0.80
	5	37	27	7	3	0.90	0.80	0.85
	6	65	53	12	0	1.00	0.82	0.90
	7	103	81	10	12	0.87	0.89	0.88
	8	69	48	21	0	1.00	0.70	0.82
	9	52	42	4	6	0.88	0.91	0.89
	10	135	107	18	10	0.91	0.86	0.88

4. 결론 및 향후 연구

본 연구에서는 OpenCV 라이브러리를 이용하여 폐암 환자의 병리 조직 H&E 염색 슬라이드에서 암세포를 구별하고 계수하는 정밀 진단 시스템을 구현하였다. Confusion matrix를 바탕으로 선암종과 편평세포암종의 F1 점수 평균은 0.85, 0.87로 나타났다. 위 시스템을 진단 시스템에 도입한다면 기존의 병리학자들의 주관적 소견과 더불어 정량적이고 객관적인 수치를 함께 제공함으로써 진단 시 높은 신뢰도 확보에 기여할 수 있을 것이다. 향후 본 연구에서 사용된 검출 및 계수 정확도 향상과 다양한 암종에 활용 가능한 알고리즘 개발, AI 기술 적용을 통해 폐암 조기 진단을 위한 연구를 지속할 것이다. 이러한 연구를 통해 향후 디지털 융합 기술을 통해 정밀 의료에 크게 기여할 수 있을 것으로 기대된다.

참고문헌

- [1] 노형준, 서주희, 2021년 사망원인통계 결과 (통계청, 2022), pp. 59
- [2] 김한숙, 김난숙, 2020년 암등록 통계 및 2014-2018 지역별 암발생통계 보도자료:2020년 코로나19 유행 첫 해, 암 발생자수 감소 및 5년 단위 시군구별 암발생 통계 결과 발표 (보건복지부, 2022), pp. 12
- [3] Blandin Knight S, Crosbie PA, Balata H, Chudziak J, Hussell T, Dive C. Progress and prospects of early detection in lung cancer. Open Biol. 2017
- [4] Patel A. Benign vs Malignant Tumors. JAMA Oncol. 2020
- [5] Borkowski AA, Bui MM, Thomas LB, Wilson CP, DeLand LA, Mastorides SM. Lung and Colon Cancer Histopathological Image Dataset (LC25000). arXiv:1912.12142v1 [eess.IV], 2019 & Lung and Colon Cancer Histopathological Images <https://www.kaggle.com/datasets/andrewmvd/lung-and-colon-cancer-histopathological-images>
- [6] Zheng M. Classification and Pathology of Lung Cancer. Surg Oncol Clin N Am. 2016
- [7] 이승민, 전정배, 이동희, 영상의학의 오류와 불일치는 불가피한 것인가 혹은 피할 수 있는 것인가? 중추신경계 증상으로 촬영한 뇌 자기공명영상에서 간과된 청신경 종양 4예, J Korean Skull Base Society, 15권, 1호, pp.39-45, 2020.
- [8] 권혜미, 디지털 병리학, NICE 평가정보, 한국IR협의회, 2021
- [9] Fischer AH, Jacobson KA, Rose J, Zeller R. Hematoxylin and eosin staining of tissue and cell sections. CSH Protoc. 2008
- [10] 류병석, 권용현, 김소희, 김영균, “이미지 분석 기반 암세포의 림프관 전이 확인 시스템 구현”, 한국소프트웨어종합학술대회 논문집, 2021, pp. 347-349