

인공지능을 활용한 혀이미지 기반 암 진단의 해외 연구 동향 분석 : 스코핑리뷰

유광욱¹, 허서연¹, 맹현규¹, 이민건¹, 오용택¹, 김안나^{2*}

1. 우석대학교 한의과대학

2. 한국한의학연구원 한의약데이터부

Abstract

Analysis of Overseas Research Trends in Tongue Image-Based Cancer Diagnosis Using Artificial Intelligence

Kwang Wook Yu¹, Seo Yeon Heo¹, Hyun Gyu Maeng¹, Min Gun Lee¹, Yong Taek Oh¹, Anna Kim^{2*}

¹College of Korean Medicine, Woosuk University

²KM Data Division, Korea Institute of Oriental Medicine

Abstract

Tongue diagnosis has been a primary tool for disease diagnosis in traditional medicine. With recent advances in artificial intelligence technology and efforts to objectify image analysis, this study aims to systematically organize research on AI-based tongue diagnosis image analysis. By comparing disease groups, applied technologies, and predictive performance, it seeks to present the international status of cancer diagnosis research and the clinical applicability for early cancer detection. This study collected literature from PubMed using the terms 'Artificial Intelligence', 'Neoplasms', 'Tongue', 'image', related MeSH terms, and free text. Disease groups included gastric cancer, oral cancer, lung cancer, and tongue cancer, with machine learning and deep learning techniques such as CNN, ResNet, U-Net, DBN, and CGTO employed. This review examines the potential application of artificial intelligence in tongue-based cancer diagnosis, aiming to suggest the direction for next-generation radiology and early cancer diagnosis.

Key words

Artificial Intelligence, Tongue Diagnosis, Cancer Detection, Deep Learning, Medical Image Analysis

* 교신저자 : Anna Kim. KM Data Division, Korea Institute of Oriental Medicine. 1672 Yuseong-daero, Yuseong-gu, Daejeon, 34054, Republic of Korea.

Tel : +82-42-868-9452, E-mail : ankim2012@kiom.re.kr

• 원고접수일 : 2025.09.12 / 심사완료일 : 2025.09.26 / 게재결정일 : 2025.10.06

I. 서론

혀는 인체의 상태를 반영하는 지표로서 진단 학적 가치가 있다. 한의학에서는 혀의 상태를 관찰하여 혀의 색, 형태, 설태의 유무를 확인하고 이를 통해 인체의 기능적인 상태를 분별하며, 질병을 진단하는 방법이 체계화되어 있다. 서양 의학 연구에서도 혀의 상태가 특정 질환과의 연관성을 보인다는 결과가 축적되고 있다. 특히 혀의 상태를 보는 것은 비침습적이며 관찰이 용이하다는 특징 덕분에 환자 선별과 조기 진단을 위한 보조적 도구로서 높은 잠재력을 지닌다.

암은 전 세계적으로 주요 사망 원인 중 하나로, 조기 발견이 환자의 예후에 결정적인 영향을 미친다. 구강암을 비롯한 두경부암은 1~2기의 조기 진단 시 생존율이 크게 올라가는데, 혀 상태(혀 이미지) 분석을 통해 암 환자에게서 나타나는 색상 변화, 표면 질감 등의 미세한 특징이 확인되었으며, 이는 전통적인 시진(視診)보다 객관적이고 정량적인 진단 지표로 확인할 수 있다. 그러나, 기존 연구의 상당수는 육안적 관찰이나 제한된 형태의 영상 처리 기법에 머물러 있어, 복잡한 형태를 정밀하게 분석하기에는 한계가 있었다.

이러한 문제의식을 해결하기 위해 최근 인공지능 기술이 적극적으로 도입되고 있다. 특히 딥러닝 기반의 합성곱 신경망 (Convolutional Neural Network, CNN)과 같은 모델은 이미지 분석에서 탁월한 성능을 보여주며, 병리학적 슬라이드, 방사성 영상, 내시경 영상 등 다양한 의학 영상 분야에서 임상적 유용성이 입증되고 있다. 혀 이미지 분석 역시 이러한 흐름 속에서 주목받고 있으며, 인공지능을 통해 육안으로는 관찰하기 어려운 미세한 특징이나 패턴까지 스스로 학습하고 판별할 수 있다는 장점이 있다. 최근 해외에서는 인공지능 기반 혀이미지 분석을 통해 구강암, 위암, 폐암 등 다양한 암의 조기 진단 가능성에 대해 연구되고 있다.

암은 반복적인 이미지 패턴을 통해 학습하는 인공지능 특성상 암에 대한 조기 발견이 용이하고, 인공지능 분석을 적용하여 정량화 가능한 특징을 가지고 있어 조기 진단에 대한 효율

을 높일 수 있다.

이에 본 연구는 인공지능을 활용한 혀이미지 기반 암 진단 연구의 해외 연구를 체계적으로 검토해보자 하였다. 발표된 연구들을 분석하여 연구 설계, 데이터셋 특성, 성능 지표, 기술 완성도, 예측 성능 등을 종합적으로 정리한다. 이를 통해 혀 이미지와 인공지능을 접목한 암 진단 연구의 국제적 현황을 분석하고, 차세대 정밀의학과 암 조기 진단에 기여할 수 있는 방향성을 제시하고자 한다.

II. 방법 및 본론

본 연구는 Arksey와 O'Malley(2005)의 스코핑 리뷰 방법론적 틀을 기반으로 수행하였다. 이에 따라 (1) 연구 질문 설정, (2) 관련 문헌 검색, (3) 문헌 선정, (4) 데이터 추출 및 정리, (5) 결과 요약의 5단계를 거쳤다.

질적 평가(Quality Appraisal)는 본 연구의 목적이 기술적 동향 파악에 있으므로 수행하지 않았다. 다만, 이로 인해 포함 논문의 질적 이질성을 통제하지 못한 점은 본 연구의 제한점으로 고려하였다.

1. 데이터 베이스

본 연구에서는 혀이미지를 기반으로 한 암 진단에 인공지능을 활용한 연구를 검색하기 위해 생의학 및 생명과학 분야에서 가장 광범위하게 활용되는 데이터베이스인 PubMed(NCBI, National Library of Medicine 제공)를 데이터베이스로 선정하였다.

2. 검색 키워드

본 연구에서는 혀 이미지 기반 암 진단에 사용되는 인공지능 기술의 활용 현황을 검토해보자 한다. 주요한 검색 키워드는 ‘인공지능’, ‘혀’, ‘암’, ‘이미지’이며, 인공지능 검색식 구성 시 PubMed의 주요 MeSH 용어(“Artificial Intelligence”, “Machine Learning”, “Deep Learning”, “Neural Networks, Computer”, “Natural Language Processing”, “Expert Systems”, “Pattern Recognition, Automated”)와 자유어(“artificial intelligence”, “machi

ne learning", "deep learning", "neural network*", "AI" 등)를 함께 활용하였다. 키워드 선정은 Liang et al.(2024), Ma et al.(2023), Baybars et al.(2025), Zhu et al.(2022) 등 선행 연구에서 실제 사용된 기술 사례를 참고하여, 개별 모델명 대신 상위 개념어를 중심으로 포괄적으로 구성하였다.

암 검색식은 PubMed MeSH 용어 'Neoplasm s'를 중심으로 대표적 암종(MeSH: Carcinoma, Adenocarcinoma, Sarcoma, Leukemia, Lymphoma, Melanoma 등)을 추가하였다. 또한 인덱싱 전 논문을 보완하기 위해 제목·초록에서 자주 사용되는 'cancer, tumor, malignant, neoplasm*' 등의 자유어를 포함하였다. 구강암을 반영하기 위해 'Mouth Neoplasms', 'Oral Cancer'를 보강하여 최종 검색 블록을 구성하였다. 2025년 8월 23일, 발행일(Publication date)을 10년으로 설정하여 발행되어 있는 논문을 기준으로 하였다.

최종 검색식은 다음과 같다.

1) 인공지능 관련어

("Artificial Intelligence"[MeSH] OR "Machine Learning"[MeSH] OR "Deep Learning"[MeSH] OR "Neural Networks, Computer"[MeSH] OR "Natural Language Processing"[MeSH] OR "Expert Systems"[MeSH] OR "Pattern Recognition, Automated"[MeSH] OR "Artificial Intelligence"[tiab] OR "Machine Learning"[tiab] OR "Deep Learning"[tiab] OR "Neural Network*"[tiab] OR "Natural Language Processing"[tiab] OR "Expert System*"[tiab] OR "Pattern Recognition"[tiab] OR AI[tiab])

2) 혀 관련어

("Tongue"[MeSH] OR "Tongue Diseases"[MeSH] OR Tongue[tiab] OR Lingual[tiab] OR "Tongue Diagnosis"[tiab])

3) 이미지 관련어

(Image*[tiab] OR Imaging[tiab] OR Photog raph*[tiab] OR Picture*[tiab] OR "Image Analysis"[MeSH] OR Photo*[tiab] OR "Digital Image"[tiab])

4) 암 관련어

("Neoplasms"[MeSH] OR "Carcinoma"[MeSH] OR "Adenocarcinoma"[MeSH] OR "Squamous Cell Carcinoma"[MeSH] OR "Sarcoma"[MeSH] OR "Leukemia"[MeSH] OR "Lymphoma"[MeSH] OR "Melanoma"[MeSH] OR Neoplasm*[tiab] OR Cancer[tiab] OR Carc inoma[tiab] OR Adenocarcinoma[tiab] OR "Squamous cell carcinoma"[tiab] OR Sarco ma[tiab] OR Leukemia[tiab] OR Lymphoma [tiab] OR Melanoma[tiab] OR Malignant*[ti ab] OR Tumor*[tiab] OR "Mouth Neoplasm s"[MeSH] OR "Oral Cancer"[tiab])

3. 포함 및 배제 기준

본 연구의 포함 기준은 다음과 같다.

1) PubMed(NCBI, National Library of Medicine 제공)에서 검색 키워드를 통해 검색 된 논문

2) 인공지능 기술을 사용한 연구

3) 혀 이미지가 포함된 연구

4) 암에 대한 연구

위의 네 가지 조건을 모두 포함한 논문만을 선정했으며, 리뷰 논문은 제외하였다. 2025년 8 월 22일 4인의 연구자가 검색어로 검색된 논문 69편의 제목과 초록을 기준으로 1차 스크리닝 하였으며, 1차 선정된 논문의 원문을 검토하여 2차 스크리닝하였다. 스크리닝 과정 및 결과는 다음과 같다(Figure 1.).

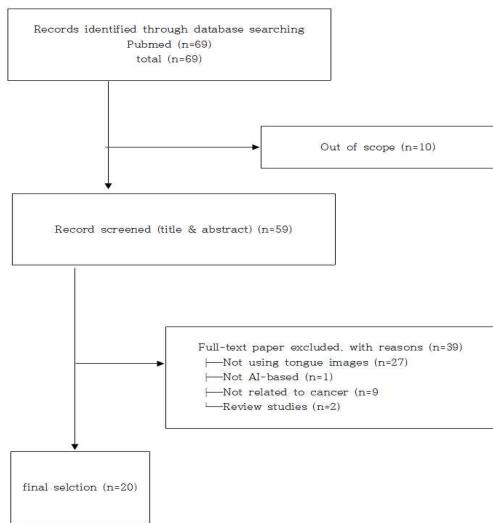


Figure 1. Paper search and selection process

4. 자료 추출

검색된 논문 중에서 설진 이미지에 기반하여 인공지능 기술을 통해 암을 예측하거나 분류한 논문을 선정하였다. 논문은 암, 사용된 인공지능 기술, 완성도(모델 구조 및 임상적용 가능성), 예측 성능(정확도, AUC 등), 데이터의 출처 및 규모, 기술의 연도 및 트렌드를 기반으로 정리했다.

III. 연구 결과

주요하게 다뤄진 암종은 구강암, 위암, 폐암, 설암 등이었으며, 각 연구들은 데이터셋 확보 방식, 적용된 인공지능 모델의 종류, 성능 지표, 설명가능성 적용 여부, 임상적 활용 가능성 등에서 다양한 특징을 보였다. 이를 암종에 따라 정리하면 다음과 같다.

1. 구강암

구강암은 혀이미지 기반 인공지능 연구에서 가장 활발하게 다뤄진 분야이다. 초기 연구들은 주로 CNN 계열의 분류 모델을 중심으로 진행되었다. DenseNet, ResNet, VGG, MobileNet, EfficientNet 등 다양한 아키텍처가 시험되었으며, 특히 DenseNet169 모델은 AUROC 0.895, AUPRC 0.918을 기록하여 일반 임상의보다 높은 성능을 보였다. 모델의 민감도는 81.1%, 특

이도는 86.8%로, 종양 전문의와 유사한 수준의 진단 능력을 보여주었다. 전문가와의 판독 일치도를 평가한 Kappa 지수도 0.685로, 인공지능 모델이 전문 의료진의 판단을 대체하거나 보완할 수 있는 잠재력을 입증하였다.

또한, 경량화 모델을 통한 실용화 가능성이 주목받고 있다. EfficientNet-B0 기반 연구에서는 소형 CNN 구조를 사용하여 정확도 85%, AUC 0.91을 달성하였다. 이는 제한된 연산 자원에서도 빠르게 구동될 수 있어, 스마트폰이나 임상용 소형 장비에 적용 가능성을 보여준다. 실제로 스마트폰으로 촬영한 임상 사진을 기반으로 개발된 모델은 AUC 0.96, 민감도 0.91, 특이도 0.81을 기록하였다. 일부 구강 해부학적 위치에서는 AUC 1.00에 달하는 성능을 보여, 원격진료나 1차 의료 현장에서 저비용·비침습적 선별 검사 도구로 활용될 가능성이 크다.

설명가능성(XAI) 적용은 아직 제한적이다. 일부 연구에서는 Grad-CAM을 통해 병변 부위에 집중하는 시각적 패턴을 확인하였지만, 대부분의 연구는 여전히 블랙박스 모델로 남아 있다. 다만 일부 논문에서는 “향후 Grad-CAM 적용”과 같은 계획을 언급하여, 해석 가능성 확보에 대한 필요성을 인식하고 있었다. 전반적으로 구강암 연구는 데이터셋 규모와 모델 구조의 다양성이 높고, 실제 임상 전문가와 성능을 비교하여 타당성을 검증하려는 시도가 활발하다는 점에서 다른 암종에 비해 연구 성숙도가 높은 것으로 평가된다.

2. 위암

위암 연구는 특히 아시아 지역에서 활발히 이루어지고 있으며, 전암성 병변(PLGC, Precancerous Lesions of Gastric Cancer) 환자를 대상으로 한 비침습적 선별 모델이 다수 보고되었다. ResNet50 기반 연구에서는 스마트폰 카메라로 촬영한 2D RGB 혀이미지를 분석하여 위암 고위험군을 예측하였다. 해당 모델은 AUROC 0.91, 민감도 82%, 특이도 88%를 기록하여 임상적 활용 가능성을 높였다.

특히 위암 연구의 강점은 대규모 다기관 데이터셋을 활용했다는 점이다. 중국 내 10개 센터에서 수천 명의 데이터를 수집하고, 7개 센터에서 외부 검증을 수행하여 모델의 일반화 성능

을 입증한 사례가 보고되었다. 이는 단일 기관 연구에 비해 과적합 위험을 줄이고 임상적 신뢰도를 크게 높이는 중요한 전략이다.

또한 멀티모달 융합 연구가 증가하고 있다. 혜이미지 단독 분석 외에도, 혈액 지표, 미생물(16S rDNA) 데이터, 생활습관 설문 등을 결합한 모델들이 보고되었다. 이미지 기반 모델이 AUC 0.83-0.88 수준인 반면, 미생물 기반 모델은 genus 단위에서 0.94, species 단위에서 0.95 까지 도달하였다. 이는 혜이미지 분석이 단순히 외형적 특징뿐 아니라, 인체 내부 대사·면역 환경과 밀접하게 연관되어 있음을 보여준다.

최근 연구에서는 CNN 외에도 Transformer 기반 모델(APINet, TransFG)이 도입되며 기술적 진보가 빠르게 이루어지고 있다. 예를 들어 APINet은 혜 특정 부위의 feature activation 을 강조하여 진단에 기여하는 특징을 확인할 수 있었고, TransFG는 attention 시각화를 통해 모델이 집중하는 영역을 설명할 수 있었다. 이는 임상적 신뢰성을 확보하는 데 중요한 진전이다.

3. 폐암

폐암을 대상으로 한 혜이미지 기반 연구는 상대적으로 적지만, 중요한 가능성을 제시하고 있다. 대표 연구에서는 TFDA-1 디지털 설진 장비를 사용하여 표준화된 혜이미지를 수집하고, 총 862명의 데이터를 분석하였다. 이 중 폐암 환자 263명, 양성 폐 결절 환자 292명, 건강 대조군 307명이 포함되었다.

모델은 랜덤 포레스트와 로지스틱 회귀를 중심으로 설계되었으며, 정확도 90.5%, 민감도 88.2%, 특이도 92.1%, AUC 0.94를 기록하였다. 이 결과는 혜이미지가 단순히 소화기 질환뿐 아니라 호흡기계 암종의 조기 경고에도 활용될 수 있음을 보여준다. 특히 저선량 CT 스크리닝은 비용과 방사선 노출 문제로 인해 보편적 적용이 어렵다는 점에서, 혜이미지를 활용한 저비용·비침습적 선별검사가 중요한 대안이 될 수 있다.

폐암 연구에서는 CNN보다는 전통적 머신러닝 기법이 중심을 이루었으며, 이는 아직 이 분야가 초기 단계임을 반영한다. 그러나 데이터 표

준화와 기계학습 기법의 단순함에도 불구하고 높은 AUC를 기록했다는 점에서, 향후 더 정교한 딥러닝 모델 도입 시 성능 향상 여지가 크다고 할 수 있다.

4. 설암

설암 연구는 혜이미지를 직접적으로 활용할 수 있는 대표적 암종으로, 진단뿐 아니라 수술 중 보조와 예후 예측에도 꼭넓게 적용되고 있다.

예후 예측 연구에서는 Radiomics 기반 모델이 다수 확인되었다. intraoral 초음파 영상, contrast-enhanced MRI, CT 영상 등 다양한 영상 데이터를 활용하여 림프절 전이, 국소 재발, 생존 기간 예측이 시도되었다. 예를 들어 intraoral 초음파 기반 Radiomics 연구에서는 NTB(Neural Tanh Boost) 모델이 AUC 0.967, 정확도 95%를 기록하여 설암 환자의 높은 시점 림프절 전이 예측에 탁월한 성능을 보였다.

또한, 병리 슬라이드 Whole Slide Image(WSI)를 기반으로 한 연구에서는 WSI 단독 모델이 AUC 0.826, 임상정보 단독 모델이 AUC 0.835였던 반면, 두 데이터를 융합한 모델은 AUC 0.991까지 향상되었다. 이는 데이터 융합 전략이 단일 영상 분석보다 훨씬 높은 성능을 달성할 수 있음을 잘 보여준다.

전반적으로 설암 연구는 “진단 보조 → 수술 중 보조 → 예후 예측”으로 이어지는 연속적 흐름을 형성하고 있으며, 이는 다른 암종보다 임상 적용 단계에 한 발 더 다가서 있는 것으로 평가된다.

5. 암종별 연구 동향의 종합적 비교

분석된 연구들에서 사용된 데이터셋은 암종에 따라 큰 차이를 보였다. 구강암은 5,000장 이상의 대규모 임상 내시경 이미지부터 100여 장 수준의 공개 데이터셋까지 다양했으며, 위암 연구는 1,000명 이상의 환자를 포함한 다기관 데이터가 특징이었다. 설암은 대부분 단일 기관 환자를 기반으로 하여 영상 모달리티의 다양성이 두드러졌으며, HSI·MRI·CT 영상 등 고해상도 영상 자료가 활용되었다. 폐암은 TFDA-1 디지털 장비를 통한 표준화 촬영이 장점이었으나, 데이터 규모는 상대적으로 제한적이었다.

AI 모델의 적용 양상 역시 암종에 따라 구분되었다. 구강암은 전통적인 CNN 계열이 주류를 이루었고, 위암은 최신 딥러닝 트렌드를 반영하여 Transformer 기반 모델이 결합되는 양상을 보였다. 설암 연구는 segmentation에 특화된 U-Net 계열과 예후 예측에 Radiomics·머신러닝 모델을 결합한 접근이 많았다. 폐암은 아직 전통적 머신러닝 기법에 의존하는 단계로, 향후 심층학습 적용 여지가 크다.

전체적으로 구강암과 위암 연구는 AUROC 0.90 이상을 다수 기록하며 조기 진단 도구로서의 가능성을 입증하였다. 설암 연구는 예후 예측 및 수술 중 보조 영역에서 Dice 계수 0.95, AUC 0.95에 달하는 성능을 기록하며 활용 범위를 확장하였다. 폐암 연구는 비록 소수지만 AUC 0.94라는 높은 성능을 보여 향후 발전 가능성을 제시하였다.

6. 설진 기반 AI 연구의 확장 동향

혀 이미지를 활용한 인공지능 연구는 암 진단 영역에 국한되지 않고, 설진 자체의 현대적 재해석과 임상 응용 범위를 넓히는 다양한 방향으로 확장되고 있다. 우선 설진은 전통 한의학에서 환자의 전신 상태를 파악하는 핵심 진단 법이었으나, 관찰자의 주관적 경험에 크게 의존하여 진단 결과가 불일치하는 문제가 있었다. 이러한 한계를 극복하기 위해 혀이미지를 디지털화하고, 영상처리 및 인공지능 기법을 접목하여 설진을 객관적이고 정량화된 체계로 발전시키려는 노력이 활발히 이루어지고 있다.

최근에는 혀 이미지를 활용해 암 여부를 판별하는 연구를 넘어, 나이·성별·체중 등 개인의 기본 생리 정보를 예측하는 시도가 보고되고 있다. 예컨대 Wiley(2024)의 연구에서는 ResNeXt 기반 모델을 적용하여 나이 예측에서 상관계수 0.71, 평균 절대오차(MAE) 8.5년을 달성하였으며, 성별 분류 정확도는 80%, AUC는 88%를 기록하였다. 체중 예측에서도 상관계수 0.39, MAE 9.06kg이라는 성과를 보였다. 이러한 결과는 혀가 단순한 진단 보조 지표를 넘어, 생리적 상태를 반영하는 정량적 생체 신호로 기능할 수 있음을 보여준다.

연구 범위는 광학 사진에 국한되지 않고 MRI

와 초음파 같은 다양한 영상 모달리티로 확대되고 있다. MRI 기반 연구에서는 U-Net을 이용해 ALS와 PBP 환자의 혀 부피를 자동 분할한 결과, PBP 환자의 혀 부피가 유의하게 감소한 사실이 확인되었다. 이는 혀의 형태가 신경 근육 질환의 진행 상황을 반영할 수 있음을 시사한다. 또한 초음파 영상을 활용한 TongueTransUNet 연구에서는 U-Net과 Vision Transformer를 결합하여 혀 윤곽을 고해상도로 분할하는 데 성공함으로써, 설진을 정적 관찰에서 동적·3차원적 분석으로 확장할 가능성을 보여주었다.

더 나아가, 혀 조직 수준의 병리학적 변화를 디지털 데이터로 구축하려는 시도도 진행되고 있다. 구강 상피 이형성(OED)을 대상으로 한 OralEpitheliumDB는 최초로 공개된 혀 조직 병리 영상 데이터셋으로, 총 456개의 이미지와 전문가 라벨을 포함한다. 이 데이터셋을 기반으로 U-Net과 ResNet-50을 결합한 세포핵 분할 모델은 F1 점수 0.83을 기록하였으며, Random Forest 분류는 정확도 94.22%를 달성하였다. 이러한 성과는 혀 조직의 미세 병리 변화를 디지털화하고 정량화함으로써, 조기 구강암 진단 및 설진의 병리학적 기초 마련에 기여할 수 있음을 보여준다.

종합하면, 설진 기반 AI 연구는 암 진단을 넘어서 생리적 상태 예측, 신경근육 질환 평가, 조직 병리 분석 등으로 빠르게 확장되고 있다. 이는 설진이 전통적 진단 도구에서 벗어나, 인공지능을 통해 현대 의학과 융합될 수 있는 새로운 정밀의학적 플랫폼으로 진화할 수 있음을 시사한다.

IV. 고찰

본 연구는 인공지능을 활용한 혀이미지 기반 암 진단 연구를 구강암, 위암, 폐암, 설암의 네 가지 암종을 중심으로 고찰하였다. 이를 통해 혀이미지가 암 조기 진단과 예후 예측에서 임상적으로 높은 잠재력을 가지고 있음을 확인할 수 있었다.

먼저, 구강암과 위암 연구는 임상 적용에 가장 가까운 성과를 보여주었다. 구강암은 CNN 계

열 모델을 중심으로 다수의 연구가 진행되었으며, AUROC 0.90 내외의 높은 성능을 기록하고 일부 연구에서는 전문가 수준의 진단 일치도를 달성하였다. 특히 경량화 모델과 모바일 기반 접근은 향후 1차 의료기관과 원격진료 환경에서 활용 가능성을 열어주었다. 위암 연구는 대규모 다기관 데이터를 바탕으로 수행된 사례가 많았고, PLGC (Precancerous Lesions of Gastric Cancer) 환자 조기 예측에서 우수한 결과를 보였다. 또한 혜이미지와 임상정보·미생물 데이터를 통합한 멀티모달 모델은 단일 영상 기반 분석을 능가하는 성능을 보여, 임상적 신뢰도를 높이는 방향으로 발전하고 있다.

설암 연구는 단순한 진단을 넘어 수술 중 보조와 예후 예측으로까지 확장된 점이 특징적이다. Hyperspectral Imaging과 U-Net을 활용한 연구는 수술 절제연 평가에 활용될 가능성을 보였고, Radiomics 기반 머신러닝 모델은 림프절 전이와 재발 예측에서 의미 있는 성과를 거두었다. 이는 혜이미지와 AI가 단순 진단 단계를 넘어 치료 및 환자 관리 과정 전반에 기여할 수 있음을 시사한다. 폐암 연구는 상대적으로 적었지만, 표준화 장비를 통한 혜이미지 수집과 AUC 0.94라는 높은 성능을 보여주어 새로운 조기 경고 도구로서 가능성을 보여주었다.

하지만 이러한 연구들에는 여전히 몇 가지 한계가 존재한다. 첫째, 데이터셋의 규모와 표준화가 부족하다. 구강암과 위암 일부 연구를 제외하면 대부분 단일 기관 소규모 데이터에 의존하고 있어 외부 검증력이 제한적이다. 둘째, 설명가능성 적용이 충분하지 않다. Grad-CAM을 도입한 사례가 있으나, 여전히 다수 연구는 블랙박스 형태의 모델에 머물러 있다. 임상 현장에서 실제 활용되기 위해서는 예측 근거를 명확히 제시할 수 있는 설명가능 AI 기법이 필수적이다. 셋째, 실제 임상 적용까지 이어지기 위해서는 실시간 분석 속도와 사용자 친화적인 인터페이스 개발이 병행되어야 한다. 현재 대부분의 연구는 후향적 데이터 분석에 머물러 있으며, 임상 현장에서 즉각적인 활용은 아직 제한적이다.

한편, 일부 연구는 혜이미지를 활용해 나이·성

별·체중 예측, MRI·초음파 기반 분석, 병리 영상 데이터 구축 등 암 이외의 영역으로 연구를 확장하고 있다. 이러한 시도들은 설진이 전통적 진단법을 넘어 정량적 생체 신호로 발전할 수 있음을 보여주지만, 본 연구의 주제인 암 진단과 비교했을 때 보조적 의미를 가진다.

종합하면, 혜이미지 기반 인공지능 연구는 암 조기 진단에서 이미 의미 있는 성과를 보이고 있으며, 구강암과 위암은 임상 적용 가능성이 높고, 설암은 수술 및 예후 관리 영역까지 확장 가능성을 제시하였다. 다만 데이터셋 표준화, 설명가능성 확보, 실시간 임상 적용이라는 과제가 남아 있으며, 이를 극복한다면 설진 기반 AI 분석은 향후 정밀의학 시대의 암 조기 진단에 중요한 도구로 자리매김할 수 있을 것이다.

V. 결론

본 논문은 인공지능을 활용한 혜이미지 기반 암 진단 연구의 해외 동향을 구강암, 위암, 폐암, 설암을 중심으로 고찰하였다. 그 결과, 혜이미지는 비침습적이고 비용 효율적인 특성을 바탕으로 암 조기 진단의 보조적 도구로 활용될 수 있는 높은 잠재력을 지니고 있음을 확인하였다.

구강암, 위암, 설암의 민감도, 특이도, auc의 중위값은 다음과 같다(table 1.).

구분	정확도	민감도	특이도	AUC
구강암	85.0%(85.0%-85.0%)	86.25%(84.5%-88.0%)	89.53%(86.7%-92.37%)	0.919(0.91-0.928)
위암	79.49%(73.33%-85.65%)	82%(82%-82%)	88%(88%-88%)	0.8(0.69-0.91)
설암	86.95%(74.1%-99.8%)	79.9%(60%-99.8%)	87.9%(76%-99.8%)	0.917(0.868-0.967)

table 1. The median values of sensitivity, specificity, and AUC for oral cancer, gastric cancer, and tongue cancer

VI. 향후 과제

구강암 연구는 CNN 계열 모델을 중심으로 높은 진단 정확도를 보였으며, 일부 연구는 전문가 수준의 판독 성능에 도달하였다. 위암 연구는 대규모 다기관 데이터와 멀티모달 융합을 통해 임상적 신뢰성을 확보하였고, 설암 연구는 수술 중 보조와 예후 예측으로까지 적용 범위를 확장하였다. 폐암 연구는 비록 소수였으나 표준화 장비를 활용하여 높은 AUC를 달성함으로써 향후 발전 가능성을 보여주었다.

다만, 데이터셋 표준화와 외부 검증 부족, 설명가능성(XAI) 확보 미흡, 임상 적용까지의 간극은 여전히 해결해야 할 과제로 남아 있다. 따라서 향후 연구는 ① 대규모 다기관 협력을 통한 표준 데이터 구축, ② 혜이미지와 임상·오믹스 데이터를 통합한 멀티모달 모델 개발, ③ 설명가능성을 강화한 AI 기법 도입, ④ 임상 현장에서 활용 가능한 실시간 분석 시스템 구현에 집중되어야 한다.

본 연구는 질적 평가를 수행하지 않았으므로, 논문의 결과를 비판적으로 해석해야 하며, 향후 연구에서는 설계를 통한 검증이 필요하다. 또한 현재 연구 분야는 연구 성과를 객관적이고 종합적으로 평가할 수 있는 표준화된 지표가 미비하여 연구 동향을 제대로 포착하지 못한다는 근본적인 한계를 가진다.

종합하면, 혜이미지 기반 인공지능 연구는 암 조기 진단과 환자 관리 전반에 걸쳐 임상적 가치를 높일 수 있는 중요한 분야이다. 이러한 연구들이 체계적으로 축적되고 임상 검증을 거친다면, 설진은 전통적 진단법을 넘어 현대 정밀 의학의 일환으로 자리잡아 암 진단과 예후 관리에 실질적 기여를 할 수 있을 것으로 기대된다.

VII. 감사의 글

본 연구는 한국한의학연구원 기본사업(KSN 2511021)의 지원을 받아 수행되었습니다.

VIII. 참고문헌

1. Heo J, Lim JH, Lee HR, Jang JY, Shin Y S, Kim D, Lim JY, Park YM, Koh YW, Ahn SH, Chung EJ, Lee DY, Seok J, Kim CH. Deep learning model for tongue cancer diagnosis using endoscopic images. *Sci Rep.* 2022;12(1):6281.
2. Jubair F, Al-Karadsheh O, Malamos D, Al Mahdi S, Saad Y, et al. A novel light weight deep convolutional neural network for early detection of oral cancer. *Oral Diseases.* 2022;28(4):1123-1130.
3. Rabinovici-Cohen S, Fridman N, Weinbaum M, Melul E, Hexter E, Rosen-Zvi M, et al. From pixels to diagnosis: algorithmic analysis of clinical oral photos for early detection of oral squamous cell carcinoma. *Cancers (Basel).* 2024;16(5):1019.
4. Rabinovici-Cohen S, Fridman N, Weinbaum M, Melul E, Hexter E, Rosen-Zvi M, et al. From pixels to diagnosis: algorithmic analysis of clinical oral photos for early detection of oral squamous cell carcinoma. *Cancers (Basel).* 2024;16(5):1019.
5. Heo J, Lim JH, Lee HR, Jang JY, Shin Y S, et al. Deep learning model for tongue cancer diagnosis using endoscopic images. *Scientific Reports.* 2022;12(1):6281.
6. Ma C, Zhang P, Du S, Li Y, Li S, et al. Construction of tongue image-based machine learning model for screening patients with gastric precancerous lesions. *Journal of Personalized Medicine.* 2023;13(2):271.
7. Ma C, Zhang P, Du S, Li Y, Li S, et al. Construction of tongue image-based machine learning model for screening patients with gastric precancerous lesions. *Journal of Personalized Medicine.* 2023;13(2):271.
8. Yuan L, Yang L, Zhang S, Xu Z, Qin J, et al. Development of a tongue image-based machine learning tool for the diagnosis of gastric cancer: a prospective multicentre clinical cohort study. *EClinicalMedicine.* 2023;57:101834.
9. Yuan L, Yang L, Zhang S, Xu Z, Qin J, et al. Development of a tongue image-based machine learning tool for the diagnosis of gastric cancer: a prospective

- e multicentre clinical cohort study. *EClinicalMedicine*. 2023;57:101834.
10. Shi Y, Guo D, Chun Y, Liu J, Liu L, et al. A lung cancer risk warning model based on tongue images. *Frontiers in Physiology*. 2023;14:1154294.
 11. Shi Y, Guo D, Chun Y, Liu J, Liu L, et al. A lung cancer risk warning model based on tongue images. *Frontiers in Physiology*. 2023;14:1154294.
 12. Konishi M, Kakimoto N. Radiomics analysis of intraoral ultrasound images for prediction of late cervical lymph node metastasis in patients with tongue cancer. *Head & Neck*. 2023;45(10):2619-2626.
 13. Adachi M, Sakashita S, Saito T, et al. Predicting lymph node recurrence in cT1-2N0 tongue squamous cell carcinoma: collaboration between artificial intelligence and pathologists. *J Pathol Clin Res*. 2024;10(1):1-10. doi: 10.1002/2056-4538.12392.
 14. Heo J, Lim JH, Lee HR, Jang JY, Shin Y S, et al. Deep learning model for tongue cancer diagnosis using endoscopic images. *Scientific Reports*. 2022;12(1):6281.
 15. Wang W, Liu Y, Wu J, et al. Early diagnosis of oral cancer using a hybrid arrangement of deep belief network and combined group teaching algorithm. *Scientific Reports*. 2023;13(1):22073.
 16. Shi Y, Guo D, Chun Y, Liu J, Liu L, et al. A lung cancer risk warning model based on tongue images. *Frontiers in Physiology*. 2023;14:1154294.
 17. Yu S, Jiang H, Xia J, Gu J, Chen M, et al. Construction of machine learning-based models for screening the high-risk patients with gastric precancerous lesions. *Chinese Medicine*. 2025;20:7.
 18. Silva AB, Martins AS, Tosta TA, Loyola AM, Cardoso SV, Neves LA, de Faria PR, do Nascimento MZ. OralEpithelium DB: A Dataset for Oral Epithelial Dysplasia Image Segmentation and Classification. *J Imaging Inform Med*. 2024;37(4):1691-1710.
 19. Konishi M, Kakimoto N. Radiomics analysis of intraoral ultrasound images for prediction of late cervical lymph node metastasis in patients with tongue cancer. *Head & Neck*. 2023;45(10):2619-2626.
 20. Shi Y, Guo D, Chun Y, Liu J, Liu L, et al. A lung cancer risk warning model based on tongue images. *Frontiers in Physiology*. 2023;14:1154294.
 21. Li X, Li L, Wei J, Zhang P, Turchenko V, Vempala N, Kabakov E, Habib F, Gupta A, Huang H, Lee K. Using advanced convolutional neural network approaches to reveal patient age, gender, and weight based on tongue images. *BioMed Res Int*. 2024;2024:5551209.
 22. Vernikouskaya I, Houssard A, Felblinger J, Vuissoz P-A, Laprie Y, Turpaul N. AI-assisted automatic MRI-based tongue volume evaluation in motor neuron disease (MND). *Int J Comput Assist Radiol Surg*. 2024 Aug;19(8):1579-1587.
 23. Al-Hammuri K, Gebali F, Kanan A, Chelvan IT. TongueTransUNet: toward effective tongue contour segmentation using well-managed dataset. *Med Biol Eng Comput*. 2025 Aug;63(8):2295-2309.
 24. Silva AB, Martins AS, Tosta TA, Loyola AM, Cardoso SV, Neves LA, de Faria PR, do Nascimento MZ. OralEpithelium DB: A Dataset for Oral Epithelial Dysplasia Image Segmentation and Classification. *J Imaging Inform Med*. 2024;37(4):1691-1710.
 25. Yuan L, Yang L, Zhang S, Xu Z, Qin J, et al. Development of a tongue image-based machine learning tool for the diagnosis of gastric cancer: a prospective multicentre clinical cohort study. *EClinicalMedicine*. 2023;57:101834.
 26. Trajanovski S, Shan C, Weijtmans PJ C, Brouwer de Koning SG, Ruers TJM, et al. Tongue tumor detection in hyperspectral images using deep learning semantic segmentation. *IEEE Transactions on Biomedical Engineering*. 2021;68(4):1330-1340.
 27. Li W, Li Y, Wang L, Yang M, Iikubo M, et al. Evaluating fusion models for predicting occult lymph node metastasis in tongue squamous cell carcinoma. *European Radiology*. 2025;35(9):5228-5238.
 28. Shi Y, Guo D, Chun Y, Liu J, Liu L, et al. A lung cancer risk warning model based on tongue images. *Frontiers in Physiology*. 2023;14:1154294.
 29. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*. 2005;8(1):19-32.