

# 조직세포의 자동화된 암 진단을 위한 병리지식 기반의 온톨로지 진단프레임워크에 관한 연구

송재원, 이주홍

인하대학교 컴퓨터 정보공학과

[sjw@datamining.inha.ac.kr](mailto:sjw@datamining.inha.ac.kr), [juhong@inha.ac.kr](mailto:juhong@inha.ac.kr)

## A study for ontology-diagnosis framework research based on pathology-knowledge for automated cancer diagnosis of biopsy samples

Jae-Won Song, Ju-hong Lee

Dept of Computer Science & Information Engineering, Inha University

### 요 약

본 논문은 병리진단지식을 활용한 조직세포 영상의 암진단을 위한 온톨로지 기반의 진단 프레임워크를 제안한다. 병리진단 분야는 환자로부터 획득한 조직샘플을 전자현미경을 이용하여 조직의 구조적 특징과 형태학적특징을 기반으로 진단을 한다. 이러한 형태의 진단은 의사의 주관적인 경험에 많이 의존하기 때문에 같은 병증에 대해서도 의사들마다 다른 진단을 하게 된다. 최근 이러한 주관적인 경험에 의한 오진을 줄이고자 주어진 조직세포 영상의 형태학적 특징들의 정량적인 수치들을 이용하는 컴퓨터 보조진단(CAD)시스템들이 많이 이용되고 있다. 그러나 이러한 진단 시스템의 요소기법들은 하나의 병증만을 진단하는데 활용되기 때문에 구성기술의 재사용성은 매우 떨어진다. 따라서 본 논문은 요소기술들의 재활용성을 높이고, 객관화된 병리진단을 위한 온톨로지 기반의 진단 프레임워크를 제시한다.

### 1. 서론

병리학자들은 암의 다양성과 복잡한 형태의 병변으로 인해 암 진단에 많은 어려움을 겪고 있다. 병리적인 암 진단은 진단의 많은 부분을 병리학자들 자신의 경험에 의존하여 진단하기 때문에 뚜렷이 객관화된 표준 진단 기준이 미비한 실정이다. 때문에 진단하는 병리의사에 따라 같은 병변에 대해서도 서로 다른 소견을 보이는 경우도 많다.[1,2] 이러한 개인적인 경험에 의존한 주관적인 진단을 보다 포괄적이고 객관적으로 진단하기 위해 병리학자들은 정기적인 학술적 모임을 통해 서로의 진단 지식을 공유하고 진단방법을 일반화 하고 있다. 그러나 공유된 지식을 체계적으로 축적하고 관리하기 위한 프레임워크가 존재하지 않기 때문에, 개별적인 차원의 지식축적과 관리가 이루어지고 있다. 따라서 병리학자들 간에 서로의 병리진단 지식을 체계적으로 관리하고, 공유하며, 검색, 관리할 수 있는 병리학회 차원의 통합 병리진단 지식관리 시스템 필요성이 인식되고 있다. 최근 온톨로지 기술의 발달은 지식을 체계적으로 표현하고 관리하는데 많은 기여를 함으로써, 각 분야의 지식관리 시스템 구축을 촉진 시키고 있다. 특히 온톨로지는 특성상 컴퓨터가 의미를 이해할 수 있기 때문에, 지식과 컴퓨터 기술이 융합된 다양한 응용이 가능하다. 따라서 병리진단 지식관리 시스템 구축과 향후 컴퓨터 기술이 융합된 다양한 응용을 위해 온톨로지 기술을

활용하는 것은 자연스러운 기술적 융합으로 볼 수 있다.

본 연구에서 제시하는 진단 프레임워크는 전문적인 진단지식을 컴퓨터가 이해할 수 있는 온톨로지로 축적하고, 축적된 진단지식을 영상처리 모듈과 맵핑함으로써 병리의사의 진단과정을 컴퓨터 진단에서 활용할 수 있도록 한다. 본 논문에서 제안하는 진단 프레임워크는 병리진단 지식을 획득하고 컴퓨터가 이해할 수 있는 온톨로지로 변환하는 지식획득 지원 시스템, 온톨로지로 표현된 진단지식들을 관리하기 위한 병리진단 지식관리 시스템, 주어진 조직세포영상의 영상처리의 형태학적인 특징추출을 위한 영상처리 서비스 시스템, 온톨로지로 표현된 진단지식 기반으로 진단기를 구성 하는 진단기 구성 시스템, 구성된 진단기에 기반을 두어 최종진단을 수행하고 진단결과를 피드백하는 병리진단 시스템으로 구성된다.

### 2. 제안된 병리지식 기반의 진단프레임워크

#### 2.1 지식획득 지원시스템

지식획득 지원시스템은 병리의사들의 진단 지식을 명세화 하는 것을 목적으로 한다. 이를 위해 먼저 선행되어야 할 과정은 진단 지식을 획득하는 것이며, 여러 가지 경로를 통해서 가능하다. 대표적인 진단지식 획득 방법과 같다.

- 전문의와 인터뷰

- 진단 증례 기록 (EMR 시스템)
- 의학 문헌 (서적 및 논문)

그러나 이 과정에서의 문제점은 같은 의학 진단이라 하더라도 임상이나 기술하는 필자의 용어 선호도에 따라 같은 의미를 가지는 다른 용어를 사용할 수 있다. 이는 체계적이고 일관된 진단행위의 명세를 어렵게 만든다. 이러한 문제점을 해결하기 위해서는 사용되는 용어를 통제와 용어들의 개념기반 해석이 필요하다[3]. 본 연구에서 제시되는 프레임워크는 진단명세 기술의 정확성과 일관성을 보장하기 위해 통제 임상용어사전[4,5,6] 도입과 개념수준 맵핑을 위한 용어저장소[7,8]를 도입한다. 기존 지식들은 사용된 용어들의 불일치성 문제를 줄이기 위해 용어저장소의 개념들과 맵핑한다. 그리고 새로 기술되는 지식들은 표현의 정확성일과 일관성을 유지하기 위해 임상용어 사전을 통해 통제된 어휘들만을 사용한다.

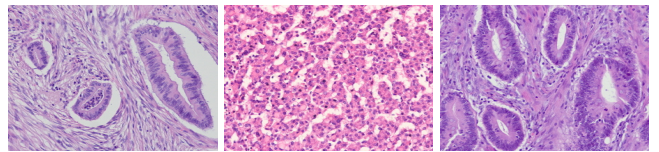
## 2.2. 병리진단 지식관리 시스템

병리진단지식을 관리하기 위한 기본 프레임워크로 온톨로지 지식 베이스 기반의 병리진단 지식관리 시스템을 구축한다. 온톨로지의 중요한 기능중의 하나는 사람이 보유한 지식을 사람과 컴퓨터, 컴퓨터와 컴퓨터 사이에 정보를 고유할 수 있도록 표현되었다는 점이다.

온톨로지로 표현된 지식을 저장하고, 추론하고, 검색하기 위해 본 시스템은 다음과 같은 구성요소들로 구성된다: OWL 지식베이스, 온톨로지 추론엔진, 온톨로지 매칭엔진. OWL지식베이스는 OWL로 표현된 진단지식을 저장하기 위한 저장소이다. 온톨로지 추론 엔진은 OWL지식베이스에 축적된 지식들로부터 새로운 진단 지식을 추론한다. 이는 진단 시 주어진 정보가 부족할 때 기존 축적된 지식들을 통해 진단이 가능하도록 돕는다. 온톨로지 매칭엔진은 OWL 지식베이스에 있는 진단지식을 검색하는데 이용된다.

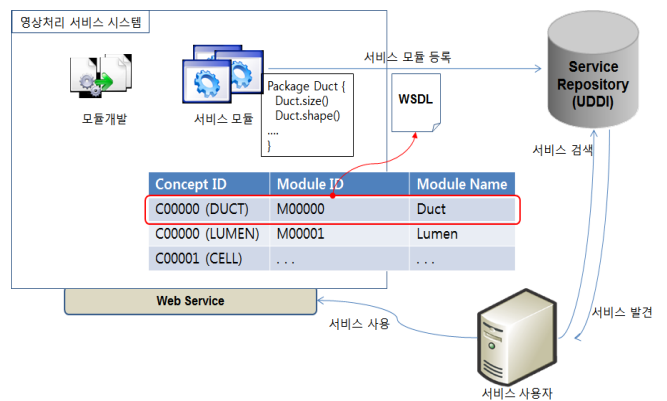
## 2.3. 영상처리 서비스 시스템

조직세포 이미지의 영상처리 기법들의 재사용성을 위해 각 오브젝트별 영상처리 기법들을 모듈별로 구성한다. 모듈화는 각 영상처리 기법들이 한 문제를 해결하기 위해 종속되지 않고 비슷한 유형을 문제들을 해결하는데 중복 없이 사용되는 장점이 있다. 아래 그림은 췌장, 간, 대장 부위의 조직세포 이미지 영상이다. 전체적인 영상의 형태는 비록 다르지만 영상의 구성요소인 핵들은 서로 비슷하게 보임을 확인할 수 있다. 또한 췌장과 대장에 있는 관영역도 상당히 비슷한 형태학적 특징을 보이고 있다.



(그림 1) 췌장암(왼쪽), 간암(가운데), 대장암(오른쪽)에 대한 조직, 세포이미지

위와 같이 부위에 관계없이 공통적으로 사용 가능한 모듈들을 조합하고 각 부위별로 필요한 세밀한 파라미터 조정을 통해 특화된 특징들의 상위 모듈로 재구성하는 것이 가능하다. 각 영상 처리 모듈은 진단기 구성 시 서비스 형태로 제공되기 위해, 통제 어휘 시스템의 개념에 맵핑된 후 서비스 모듈로 등록된다. 그림 2는 모듈의 생성으로부터 서비스 모듈 등록과 사용에 이르는 전체적인 과정을 보여주고 있다.



(그림 2) 영상처리 서비스 시스템

## 2.4. 진단기 구성 시스템

진단기 구성 시스템은 병리진단 시스템으로부터 입력된 환자 정보를 이용하여 진단기를 구성한다. 구성된 진단기는 하나의 진단 플랫폼 안에서 다양한 진단을 할 수 있도록 플러그인 형태로 제공된다. 새로운 형태의 이러한 접근 방법은 병리의사의 진단 지식과 경험을 진단 시스템에 직접적으로 반영함으로써 기존의 진단 시스템에서 수행될 수 없었던 다양한 형태의 진단과 진단에 대한 근거 제시가 가능해진다. 또한 각 처리 서비스 모듈들의 조합을 통해 진단기의 기능 변경 및 새로운 기능 추가에 효과적인 대처가 가능하다.

병리진단 지식관리 시스템에 등록된 진단 지식 명세는 온톨로지 표현 언어인 OWL로 표현되어 있기 때문에, 컴퓨터 시스템이 진단 지식 명세의 의미를 이해하고 위의 진단기 구성과정을 자동화 하는 것이 가능하다. 본 단계에서 생성된 진단기는 병리진단 시스템에 플러그인 된다. 즉, 병리진단 시스템은 상황에 따라 여러 가지 진단기들을 플러그인 함으로써 다양한 병리진단을 하고 병리 의사가 종합적으로 최종 판단을 하도록 도와줄 수 있다.

### 3. 진단시스템의 진단 결과를 이용한 최종 진단

병리의사는 "진단기 구성 시스템"으로부터 구성된 진단기를 최종 진단을 수행하기 위한 인터페이스인 Front-end 시스템에 플러그인하고, 축적된 병리진단 지식에 기반한 최종 진단을 수행한다. 최종진단을 위한 과정은 다음의 5단계 거친다.

1. 입력단계
  - 입력부로부터 진단하고자 하는 환자의 조직세포 이미지와 환자 정보를 입력 받는다.
2. 전처리단계
  - 입력부로부터 획득된 환자정보를 이용하여 진단 범위 설정
  - 진단기 구성기 시스템에 진단기를 요청.
3. 진단기 플러그인 단계
  - 진단기 구성기로부터 획득된 진단기들을 플러그인 한다.
4. 시스템 진단 단계
  - 플러그인된 진단기들은 각각 환자의 조직세포 영상을 분석
  - 각각의 진단기들은 영상 진단 결과와 진단 근거를 제시
5. 최종 진단 및 증례 기록
  - 병리의사는 시스템의 진단결과와 근거를 확인 후 최종 진단을 한다.
  - 최종 진단은 검증을 거쳐 새로운 진단 지식으로 등록되거나 기존 지식을 업데이트 한다.

### 4. 결론 및 향후 연구과제.

본 연구에서는 병리의사나 의학 문헌으로부터 획득된 진단 지식을 통해 진단과정을 온톨로지로 모델링하고 진단에 활용하는 방법을 제시한다. 따라서 제안된 시스템은 실제 병리의사의 진단 경험이 모델링되어 시스템 진단에 활용되기 때문에 진단근거 제시와 정확한 결론을 제시할 수 있으며, 축적된 진단 지식은 의사의 미경험이나 부주의로 인해 놓치기 쉬운 희귀질환에 대해서도 광범위한 진단을 할 수 있다. 또한 진단 지식 자체를 이용하여 진단기를 구성하기 때문에 새로운 진단방법이나 변경되는 진단방법 효과적으로 대응할 수 있다.

현재 연구는 채장을 중심으로 "영상처리 서비스 시스템"의 영상처리 모듈과 형태학적 특징추출 모듈을 개발 중에 있으며, 다음 진행 단계로 통제어휘의 각 개념에 "영상처리 서비스 시스템" 영상처리 모듈과 연동을 위한 연구를 진행할 예정이다.

### 참고문헌

[1] A. Andrion, C. Magnani, P.G. Betta, A. Donna, F. Mollo, M. Scelsi, P. Bernardi, M. Botta, B. Terracini, Malignant mesothelioma of the pleura: interobserver variability, J. Clin. Pathol. 48(1995) 856-860  
 [2] S.M. Ismail, A.B. Colclough, J.S. Dinnen, D. Eakins, D.M. Evans, E. Gradwell, J.P. O'Sullivan, J.M.

Summerell, R.G. Newcombe, Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia, Br. Med. J. 298(1989) 707-710  
 [3] 윤지현 외 5명, "EMR 시스템의 임상용어 통합과 관리를 위한 임상용어 사전의 개발," 대한의료정보학회지 제15권 제4호, 2009, p.411~421  
 [4] Rector A, "Coordinating Taxonomies: key to re-usable concept representations," Artificial Intelligence Medical 1995:17-28.  
 [5] Cimino JJ, Patel VL, Kushniruk AW, "Studying the human computer-terminology interface," Journal of American Medical Informatics Association, 2001;8(2):163-173  
 [6] Cimino JJ. "Desiderata for controlled medical vocabularies in the 21st century," Methods of Information in Medicine, 1998; 37(4-5):394-403  
 [7] Rocha RA, Huff SM, Haug PJ, Warner HR. "Designing a controlled medical vocabulary server: the voser project," Computers and Biomedical Research, 1994;27:472-507  
 [8] Cimino JJ, "From Data to knowledge through conceptoriented terminologies: experience with the medical entities dictionary," Journal of American Medical Informatics Association 2000;7(3): 288-297