

## Computer Aided Diagnosis System based on Performance Evaluation Agent Model

Hyun-Sook Rhee\*

### Abstract

In this paper, we present a performance evaluation agent based on fuzzy cluster analysis and validity measures. The proposed agent consists of three modules, fuzzy cluster analyzer, performance evaluation measures, and feature ranking algorithm for feature selection step in CAD system. Feature selection is an important step commonly used to create more accurate system to help human experts. Through this agent, we get the feature ranking on the dataset of mass and calcification lesions extracted from the public real world mammogram database DDSM. Also we design a CAD system incorporating the agent and apply five different feature combinations to the system. Experimental results proposed approach has higher classification accuracy and shows the feasibility as a diagnosis supporting tool.

▶ Keyword : CAD system, feature selection, performance evaluation agent, feature ranking

### I. Introduction

컴퓨터 보조 진단 시스템은 일반적으로 주어진 응용분야의 데이터를 획득하는 전처리과정, 특징선택방법, 의사결정메카니즘 등의 3가지 주된 과정으로 이루어져 있다[1-2]. 전처리과정은 원 데이터의 종류에 따라 처리과정이 다양하지만 보통 데이터를 분할하여 문제해결의 열쇠를 쥐고 있는 속성이나 해당 분야의 의사결정을 위하여 축적된 데이터를 분석하여 필요한 속성을 추출해 내는 과정으로 주어진 응용분야에 따라 여러 방법을 사용하고 있다. 특징선택은 이렇게 추출한 문제영역에서 관찰된 다차원 속성데이터 중에서 의사결정에 결정적인 역할을 수행하는 특징을 선택하여 실제 시스템에 적용하기 위한 효과적인 응용데이터 셋을 구성하는 것이다[3-4]. 이 과정은 시스템의 성능향상에 중요한 구성요소로서 상관관계기법, 차원축소 및 상호 정보처리 등의 정보이론이나 통계학적 접근방법으로 연구되어왔다. 특히 이 분야의 연구는 다루는 데이터의 양이 방대해지고 복잡해지면서 더욱 중요시 되어 인공지능분야의 휴리스틱 방법이 적용되기도 하였다. 의사결정 메카니즘에서는 전

문가에게 유용한 정보를 제공하는 정확한 시스템을 구성하기 위한 핵심 과정으로서 준비된 특징데이터를 가지고 패턴찾기, 분류 등의 기법을 활용하고 있다. 이 과정에서 가장 일반적으로 사용하는 방법은 주어진 데이터를 활용한 입력데이터로부터 출력데이터로의 매핑 방법을 연구하여 적용하는 것이다. 기존의 일반적인 전문가시스템을 포함하는 지능형시스템 구성요소의 중요한 연구대상으로 주어진 데이터의 학습을 통하여 구현되고 있다. 최근 각광받고 있는 데이터마이닝의 연구도 이 매핑방법을 기준으로 연구되었으며 이미 알려진 데이터를 학습하는 과정에서 최적의 모델을 만들어 비슷한 상황에서 지능적인 결정을 할 수 있는 방법을 정립하는 것이 핵심기술이다. 여기에 포함된 학습방법 중에서 대표적으로 사용된 기법은 신경망을 이용한 클러스터링 방법론이다[5]. 클러스터링은 비교사 학습(Unsupervised learning) 방법으로 속성이 비슷한 것들끼리 묶어 나누는 것으로, 분석하고자 하는 데이터가 너무 많아 전체를 파악하기 어려울 때, 부분을 살펴 전체의 윤곽을 잡도록 해준다. 이와 같은 클러스터링은 전문가시스템, 패턴인식, 영상처

---

• First Author: Hyun-Sook Rhee, Corresponding Author: Hyun-Sook Rhee  
\*Hyun-Sook Rhee(hsrhee@dongyang.ac.kr), Dept. of Software Engineering, Donyang Mirae University  
• Received: 2015. 12. 10, Revised: 2015. 12. 24, Accepted: 2016. 01. 18.  
• This work was supported by Dongyang Mirae University Research Grant.

리, 음성인식 등 학습이 필요한 모든 영역에서 포함하는 기본 단계이다. 특히 요즘 인터넷을 통한 검색엔진의 탐색과정에도 활용되어 비슷한 정보를 그룹핑하여 보여주고 있다. 그러나 기존의 클러스터링 방법론은 대부분 주어진 데이터 사이의 경계가 명확하다는 가정에서 각 패턴을 하나의 클래스에 소속시키는 방법이다. 그러나 이 모델은 우리가 다루는 데이터의 경계가 대부분 불명확하므로 실제 데이터 상호간의 군집성을 묘사하기에 부적절하며 주어진 데이터 분포의 성질을 잃어버리는 결과를 가져온다. 본 논문에서는 퍼지 클러스터 분석을 바탕으로 하는 OFUN-NET[6]을 적용하여 시스템을 구성한다. 이렇게 제안된 시스템을 유방 x-선 영상 진단에 활용하기 위하여 다음의 세 가지 방법으로 데이터를 획득하고자한다. 첫째는 영상 자체에 대한 데이터이다. 보통 픽셀의 수, 히스토그램, 평균 grey값, 평균 경계 grey값, 대비값, 비대칭값 등 영상자체가 갖는 데이터를 말한다. 이는 컴퓨터의 영상 처리 과정에서 자동적으로 추출되는 정보이다. x-선 판독전문의가의 시력의 한계에 도움을 주어 유용한 정보를 제공하며 기계적이기 때문에 항상 일관된 정보를 제공해 주는 장점이 있다. 두 번째는 x-선 판독전문가에 의한 영상 해석 정보이다. 전문가의 경험지식을 바탕으로 결정에 영향을 주는 요소를 수치적으로 인코딩한 데이터이다. 널리 사용되어온 방법이나 전문가의 지식 및 해석 수준에 대한 기준 등의 여러 변수가 있다. 세 번째는 환자의 개인정보이다. 환자의 개인정보에 대한 해석과 수집에 대한 어려움 때문에 시스템에서는 대부분 환자의 나이 이외의 데이터는 사용하지 않고 의사 결정에 내리는데 도움이 되는 경험적 데이터로 사용되고 있다. 본 연구에서는 실세계 의료 기관으로부터 수집되고 공개적으로 제공되는 맘모그램 데이터베이스 DDSM[7]으로부터 수집한 환부의 종류를 종괴와 석회로 나누어 각각 두 유형을 묘사하는 특징 5가지와 환자의 나이 정보를 사용하였다. 이 때 사용된 특징은 x-선 영상 판독 전문가의 의견을 BI-RADS 기준[8]에 의한 수치 값으로 변경하여 사용하였다. 또한 데이터와 분류에 영향을 끼치는 영상 처리 특징 데이터로 알려진 엔트로피(entropy), 표준편차(standard deviation), 픽셀의수(number of pixels) 등의 14개의 그레이레벨 특징데이터를 사용하였다[9-10].

본 논문에서는 이렇게 획득한 데이터를 가지고 퍼지클러스터분석과 타당성척도에 기반을 둔 성능측정에이전트를 제시하고 DDSM으로부터 추출된 데이터의 특징을 서열화한다. 또한 제안된 에이전트를 결합한 컴퓨터보조진단 시스템을 설계하고 5가지 유형의 특징조합을 가지고 시스템에 적용하여 테스트한다. 실험 결과는 제안된 시스템의 분류정확도를 가지고 진단 보조 도구로서의 가능성을 보여주고 있다.

본 논문의 구성은 다음과 같다. II장에서는 그 동안 연구되어 온 관련연구로서 클러스터 타당성 측정자와 특징선택방법에 대하여 요약하고 본 논문의 연구대상인 유방암진단을 위한 컴퓨터 보조진단 시스템에 대하여 고찰함으로 연구배경을 서술하고자한다. III장에서는 본 논문에서 제안하는 성능측정 에이전트

기반 컴퓨터보조진단 시스템의 구성도를 보여주며 에이전트 모델이 가지는 성능측정자를 이용한 특징서열화 방법을 서술한다. IV장에서는 DDSM 공개 데이터베이스로부터 획득한 입력 데이터를 준비하여 5가지 유형의 특징조합을 가지고 시스템에 적용하여 테스트한 결과 제안된 시스템의 진단 보조 도구로서의 가능성을 보여준다. 마지막으로 V장에서는 본 논문의 연구 내용을 요약하고 앞으로의 연구방향을 제시한다.

## II. Related Works

본 논문에서 제안한 성능측정에이전트 모델기반 시스템과 관련된 특징선택방법과 클러스터 타당성 측정자에 대한 기존의 연구를 소개하고 제안된 시스템의 가능성을 측정하기 위한 문제영역으로 선택한 유방 x-선 영상 진단을 위한 컴퓨터 보조진단 시스템에 대하여 고찰함으로 연구배경을 서술하고자한다.

### 1. Feature Selection Method and Cluster Validity Measures

특징선택 방법은 문제 영역에서 관찰된 다차원데이터로부터 데이터가 묘사하는 구조를 잘 반영하는 속성을 선택하여 효과적인 정제된 데이터 셋을 구성하는 시스템의 핵심부분이다[3-4]. 이 과정은 문서분류, 영상인식, 유전자 선택 분야에서의 같은 분류시스템의 성능향상에 중요한 구성요소로서 정보이론이나 통계학의 접근방법을 중심으로 연구되어왔다. 이와 같은 특징 선택 분야의 연구는 다루는 데이터의 양이 방대해지고 복잡해지면서 더욱 중요시 되어 인공지능 분야의 휴리스틱 방법이 적용되기도 하였다. 그러나 지금까지의 연구는 시스템의 목적이나, 응용 영역에 따라 상이한 데이터의 특성을 반영하지 못하는 이론적인 연구이며 무엇보다 계속 변화하는 새로운 데이터에 대한 적응력이 부족하다. 본 논문에서는 데이터가 가지는 특성을 반영하면서 새로운 데이터를 학습하고 일반화할 수 있는 특징선택방법으로 퍼지 클러스터 분석 방법을 이용한다. 이와 같은 클러스터 분석을 특징선택에 적용하기 위해서는 그 결과가 주어진 입력 데이터의 구조를 얼마나 잘 반영하고 있는가를 측정하는 척도가 필요하다. 이에 관련된 연구 영역을 "cluster validity problem"이라하며 이를 위하여 다음과 같이 클러스터 타당성 측정 함수를 정의하였다[11]. 이와 같은 타당성 측정 함수들은 클러스터링의 결과 형성된 소속 값 안에 포함된 정보를 하나의 값으로 간략하게 요약해 준다. 그러나 이들 타당성 측정 함수들은 그들이 산출하는 값이 주어진 데이터가 가지는 기하학적인 성질을 직접 반영하고 있지 않으며 클러스터의 수가 증가함에 따라 클러스터링 결과와 관계없이 단조 감소하는 값을 산출해 내는 단점을 가지고 있다. 그러므로 이와 같은 측정 함수를 클러스터링 알고리즘에 결합시킨 방법은 경험적 성질을 가지며 임의의 임

계치를 설정해 주어야 한다. 타당성 측정을 위한 함수가 단조 감소하는 경향을 극복하기 위하여 그 함수에 정규화(normalization)나 통계적 표준화(statistical standardization)를 적용하는 방법이 제안되었다[12]. 또한 주어진 데이터의 기하학적인 형태를 고려한 타당성 측정 함수가 제안하기도 하였다. 이러한 방법은 기하학적으로 같은 클래스의 데이터는 군집해 있고 다른 클래스 사이의 데이터는 멀리 위치하고 있는지를 알아내기 위하여 고안되었으나 퍼지 클러스터링 알고리즘에 직접 적용할 수 없다. 우선 클러스터링의 결과를 어느 특정 소속이 분명한 경우로 변경한 후에 적용해야 하는데 그 변경과정은 여러 방법이 있을 수 있으므로 하나의 결과를 얻을 수 없는 단점을 가지고 있다.

- Partition Coefficient(PC) :

$$PC(U;c) = \frac{\sum_{j=1}^n \sum_{i=1}^c (u_{ij})^2}{n}$$

- Classification Entropy(CH) :

$$CE(U;c) = \frac{-\sum_{j=1}^n \sum_{i=1}^c u_{ij} \log_a u_{ij}}{n}$$

- Proportion Exponent(PE) :

$$PE(U;c) = \log_c \left[ \prod_{j=1}^n \left\{ \sum_{l=1}^c (-1)^{j+1} \binom{c}{j} (1 - ju_l)^{c-1} \right\} \right]$$

, where  $u_i = \sum_j u_{ij}$

이와 같은 클러스터 타당성 척도 함수들은 클러스터 분석 방법에 의하여 형성된 결과가 얼마나 의사결정에 유용하게 사용될 수 있는지를 체크하는 과정에도 사용될 수 있다. 그러므로 특징선택과정에서 특징을 서열화하기 위한 기준으로도 활용될 수 있다.

## 2. Computer Aided Diagnosis System

컴퓨터 보조 진단 시스템의 응용분야는 인간 전문가가 관여하는 전체를 포괄할 정도로 광범위하다. 특히 의료장비의 발전과 패턴인식분야의 연구결과를 접목하여 컴퓨터를 활용한 디지털 유방 X선 영상 진단 시스템과 같은 의료 정보화 분야에서 괄목할만한 의미 있는 결과를 도출하고 있다[1]. 이러한 연구는 진단의 중요한 단서가 되는 환부 영상은 인간 시력의 한계와 감정적인 개입으로 객관적인 정보를 얻기 어려울 수 없는 경우가 있기 때문에 더욱 발전할 것이다. 또한 그 동안 축적된 한 개인의 데이터 뿐 아니라 여러 전문가의 경험사례와 지식을 데이터베이스에 저장하여 활용할 수 있는 시스템의 구축은 인

간 전문가의 의사결정을 돕기 위한 보조 수단으로 더욱 발전할 것이다. 특히 본 논문의 실험에서 사용한 맘모그램(유방디지털 x선영상)에 나타난 종괴와 석회 환부는 진단을 위한 중요한 데이터로서 여러 가지 형태의 컴퓨터 보조 진단 시스템이 제안되어왔다. 시스템의 진단과정은 크게 x-선 영상으로부터의 특징선택과정과 분류과정으로 나눌 수 있다. 이러한 과정에서 기존에 연구 되어 온 패턴인식과 컴퓨터 비전의 연구 결과가 의료영상의 해석과 분석에 활용되어왔다. 또한 분류를 위하여 지지벡터기계 등의 통계학적인 방법에서부터 이를 변형한 방법, 신경망과 퍼지이론에 이르기까지 광범위한 지능형 데이터 처리기법들이 활용되어왔다[13].

환부조직의 구조와 모양이 다양해짐에 따라 분류가 어려워지면서 Wu 등은 앙상블 네트워크 모델을 제안하여 분류의 정확도를 높이려는 연구를 하였다. 제안된 모델은 종괴 환부를 분류하기 위하여 가중치 평균과 퍼셉트론 평균의 두 가지 알고리즘을 사용하였다. [14]에서는 정해진 특징을 사용하여 특징 추출 과정에서 분류에 영향을 주는 특징을 찾아가는 과정을 통하여 역전파 신경망기반의 컴퓨터 보조진단시스템을 구성하였다. 8192번의 실험을 통하여 신경망의 구조를 결정하고 미세석회 패턴을 탐지하는 퍼지 논리 기반 특징결합방법을 제안하였다. 제안된 시스템은 특징을 선택하고 평가하는 과정에서부터 분류하는 과정까지 퍼지 신경망의 원리를 체계적으로 적용하였으나 20개의 양성환부와 20개의 악성환부 모두 40개의 미세석회환부가 학습에 사용되어 그 데이터가 작기 때문에 연구의 신뢰도가 높지 않은 것으로 판단된다. 다양한 특징 결합의 시도로부터 형태학적인 정보를 가지는 영상처리 데이터와 x-선 영상 판독 전문가에 의한 데이터가 적절히 결합되었을 때 성능이 우수한 것으로 알려졌다. 이러한 연구를 바탕으로 [6]에서는 두 가지 종류의 데이터에 각각 클러스터 분석 기법을 적용하여 진단시스템을 제안하기도 하였다. 제안된 듀얼 OFUN-NET을 활용한 진단시스템은 퍼지 클러스터링 학습방법을 사용하여 하나를 사용한 경우보다 높은 분류정확도를 나타냄을 보이면서 그 타당성을 보이고 성능에 영향을 주는 특징을 선택할 수 있는 자동화된 방법을 향후 연구과제로 제시하기도 하였다. 또한 최근 DWT(Discrete Wavelet Transformation)를 사용하여 영상으로부터 잡음을 제거한 후 ROI(region of interest)를 선택하여 20개의 GLCM(Gray Level Co-occurrence Matrix) 특징데이터를 가지고 여러 분류알고리즘에 적용한 분류율을 비교한 연구도 진행되었다[15]. 진단오류를 줄이기 위하여 두가지 관점에서의 맘모영상의 분석을 통하여 얻은 정보를 융합하는 방법에 대하여 연구하면서 [16]에서는 DDSM 데이터셋으로부터 선택된 152 케이스의 두가지 관점의 영상으로부터 304개의 ROI를 가지고 실험한 결과 하나의 관점에서의 영상보다 높은 분류정확도를 나타낸 것을 확인하였다. 또한 [17]는 여러 분류기에 의한 다중 에이전트 방법을 이용한 다각도 정보 융합 알고리즘을 제안하여 DDSM으로부터의 64 케이스에 대한 128개의 ROI를 가지고 실험하였다. 그 결과 분류정확도는 높이고

FPR(False Positive Rate)는 낮추는 실험 결과로서 그 방법의 타당성을 입증하였다.

### III. Performance Evaluation Agent for Computer-Aided System

컴퓨터 보조 진단시스템은 그동안 인공지능 연구결과를 특정전문영역에 접목한 지능형 의사결정시스템을 구현한 것이다 [1][13]. 이러한 의사결정시스템은 감정의 개입 없는 객관적 데이터를 바탕으로 인간 전문가의 특정영역의 작업을 대신 수행해 주는 시스템을 말한다. 이러한 시스템은 의사결정에 사용될 데이터를 구성하는 과정과 의사결정의 근거가 되는 데이터가 가지고 있는 패턴을 찾아내는 방법이나 분류방법이 그 핵심 구성요소가 될 것이다. 이 과정에서 가장 일반적으로 사용하는 방법은 주어진 데이터를 활용한 입력데이터로부터 출력데이터로의 매핑 방법을 연구하여 적용하는 것이다. 기존의 일반적인 전문가시스템을 포함하는 지능형시스템 구성요소의 중요한 연구대상으로 주어진 데이터의 학습을 통하여 구현되고 있다. 최근 각광받고 있는 데이터마이닝의 연구도 이 매핑방법을 기준으로 연구되었으며 이미 알려진 데이터를 학습하는 과정에서 최적의 모델을 만들어 비슷한 상황에서 지능적인 결정을 할 수 있는 방법을 정립하는 것이 핵심기술이다. 여기에 포함된 학습 방법 중에서 대표적으로 사용된 기법은 신경망을 이용한 클러스터링 방법론이다. 물론 시스템의 전체성과 효율성은 학습할 데이터를 구성하는 데이터준비과정에 의하여 좌우된다는 것은 잘 알려진 사실이다. 입력데이터의 준비과정 중 특징선택은 문제영역에서 관찰된 다차원데이터로부터 데이터를 묘사하는 구조를 잘 반영하는 특징을 추출하는 것이 핵심기능이다. 일반적으로 특징 추출은 같은 범주에 속하는 데이터들의 공통적인 속성(intraset features)과 서로 다른 범주에 속하는 데이터를 구별하는 속성(interset features)으로 나누어 고려되어 왔다. 이러한 특징선택은 어떤 시스템이든 근본적인 성능과 유용성을 결정짓는 중요한 영역으로 상관관계기법, 차원축소 및 상호정보보처리 등의 정보이론이나 통계학적인 접근방법을 중심으로 연구되어왔다. 하지만 점차 다루는 데이터의 양이 방대해지고 복잡해지면서 인공지능분야의 휴리스틱 방법론이 적용되어 연구영역을 확대하고 있다[3-4].

본 논문에서는 준비된 각 특징값의 클러스터링 결과를 가지고 각 특징의 클러스터에 대한 군집성과 분리성의 정도를 측정하여 그 값에 따라 특징을 선택하는 성능측정에이전트 모델을 제안한다. m개의 각 속성 데이터에 대하여 n개의 준비된 데이터는 퍼지클러스터링 분석(FCA)[5,6]에 의하여 만들어진 클러스터의 정보를 이용하여 그 특징데이터의 타당성을 측정한다.

이를 위하여 클러스터 중심점과 소속함수값을 이용하여 그 특징데이터에 의해 만들어진 클러스터의 성능을 측정하는 모델

을 구성하는 기준으로 주어진 특징데이터의 군집성 측정자(compactness index)  $G$  와 분리성 측정자(separation index)  $S$  를 정의하였다. 형성된 클러스터링의 성능은 같은 클러스터에 속하는 데이터사이에는 군집성이 있고 다른 클러스터에 속하는 데이터 사이에는 분리성이 있으나에 좌우된다는 것이다. 즉  $G$  의 값이 작을수록,  $S$  의 값이 클수록 특징데이터가 분류에 적합한 속성임을 알 수 있다. 이를 이용하여 다음과 같은 정의를 바탕으로 각 특징데이터의 클러스터 성능측정자  $I(d)$  를 군집성 측정자  $G$  에 대한 분리성 측정자  $S$  의 비율로 정의하였다.

[정의 1 : 군집성 측정자] n개의 데이터가 c개의 클러스터를 형성한 경우, 임의의 두 데이터,  $f_j$ 와  $f_k$ , 사이의 거리의 평균으로서 다음의 식 (1)과 같이 정의된다. 이때  $\omega_1$ 은 임의의 두 데이터가 같은 클러스터에 속할 가능성정도를 나타낸다

$$G = \frac{2}{n(n-1)} \sum_{j=1}^{n-1} \sum_{k=j+1}^n (f_j - f_k)^2 \omega_1 \quad (1)$$

, where  $\omega_1 = \min_i \{u_{ij}, u_{ik}\}$

[정의 2 : 분리성 측정자] 서로 다른 클러스터에 속하는 임의의 두 데이터 사이의 거리의 평균으로서 다음의 식 (2)과 같이 정의된다.  $u_{i,j}$  는  $f_j$ 가 가장 큰 소속정도를 가지는 클러스터를  $i_1$  이라하고  $f_k$ 가  $i_1$ 이 아닌 가장 큰 소속정도를 가지는 클러스터를  $i_2$ 라고 하자. 이때  $\omega_2$ 은 임의의 두 데이터가 서로 다른 클러스터에 속할 가능성정도를 나타낸다.

$$S = \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n (f_j - f_k)^2 \omega_2 \quad (2)$$

, where  $\omega_2 = \min \{u_{i_1,j}, u_{i_2,k}\}$

이러한 정의를 기반으로 클러스터의 성능을 측정하는 방법을 결합한 [특징 서열화 알고리즘]을 기술하면 다음과 같다.

[단계1] m개의 속성을 가지는 n개의 특징데이터를 준비한다.

[단계2]  $d = 1, \dots, m$  에 대하여

(2-1) n개의 d번째 특징 데이터에 대하여 퍼지클러스터분석 알고리즘에 의해 p개의 클러스터를 만든다. 이를 통하여 p개의 중심값,  $c_1, \dots, c_p$  와  $p \times n$ 의 퍼지소속 정보  $u_{i,j}(d)$  를 얻게된다.

(2-2) 특징 d의 군집성과 분리성을 측정하여 G와 S를 계산한다

(2-3) G와 S를 이용하여

①  $G \neq 0$  이면  $I(d) = S/G$  를 계산하여 저장한다

- ②  $G=0$ 이고  $S \neq 0$  인 경우  
속성  $d$ 는 특징으로 선택된다
- ③  $G=0$ 이고  $S=0$  인 경우  
속성  $d$ 는 특징이 될 수 없다

[단계3] [단계2]의 (2-3)-②에서 선택된 경우와  $I(d)$ 의 값이 큰 순서로 특징을 정해진 개수  $l$ 개의 특징을 선택하여  $n \times l$ 의 정제된 데이터 집합(cleaned data set)을 얻는다.

$I(d)$ 는 같은 클러스터 안에서의 군집성  $G$ 가 작을수록, 다른 클러스터 사이의 분리성  $S$ 가 클수록 속성  $d$ 가 만들어 낸 클러스터가 타당함을 나타내는 측정자가 된다. 즉  $I(d)$ 는 밀집성과 분리성의 비율을 나타내므로 데이터분포와 클러스터 사이의 관계를 반영한 값을 산출하며 데이터집합 사이의 상대적인 비교가 가능하게 하였다. 보통 성능측정자가 가지는 클러스터의 수가 많아짐에 따라 단조 감소하는 경향이 적으며 클러스터 대푯값 보다는 실제 데이터를 적용하여 위의 알고리즘의 [단계2]의 (2-3)과 같이 분석할 수 있다. [단계2]의 (2-3)-②는 같은 그룹에 속해 있는 데이터는 모두 같은 값을 가지고 있는 경우로 특징데이터로 선택될 수 있다. (2-3)-③의 경우는 모든 데이터가 같은 값을 가지고 있는 경우이므로 분류에 도움을 주지 못하는 속성이므로 특징데이터로 선택될 수 없다. 이와 같은 ②와 ③의 케이스는 극히 예외적이며 대부분의 데이터는 모든 속성에 대하여 얻은 측정값  $I(d)$ 를 기준으로 그 값이 큰 순서로 속성을 서열화 하여  $l$ 개의 특징을 선택할 수 있는 메커니즘을 제공한다.

[그림 1]은 본 논문에서 제안하는 성능측정 에이전트 모델의 기본구성도로 실제계로부터 수집된 데이터 셋으로부터  $m$ 개의 퍼지 클러스터 분석기(FCA<sub>1</sub>, ..., FCA<sub>m</sub>), 성능측정자  $I(d)$ 를 계산하는 PE Measures, 그리고 위의 서술한 [특징 서열화 알고리즘]을 통해 처리에 사용할 정제된 데이터를 추출하는 Feature Selection 모듈로 구성되어 있음을 보여주고 있다.

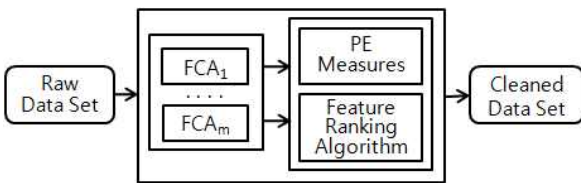


Fig. 1. Performance Evaluation Agent

이렇게 마련된 정제된 특징 데이터 셋은 분류를 목적으로 하는 데이터마이닝 방법을 적용하여 데이터를 분류한다. 데이터 마이닝은 방대한 양의 데이터를 탐색하여 숨겨진 정보와 규칙 또는 요약정보를 꺼내는 작업을 말한다. 이를 위하여 여러 방법론이 적용되어왔지만 클러스터분석방법은 가장 일반적으로 적용되는 방법이다. 이는 비교사학습방법으로 비슷한 그룹으로 묶어가며 그 군집에 속하는 대푯값을 찾아가는 방법으로 새로운 데이터에 대하여 적용하기 위한 신경망모델이 보편적으로

사용되고 있다.

본 논문에서는 데이터 분류를 위하여 클러스터링 결과가 만들어내는 거리의 합을 표현하는 퍼지함수의 값  $J_m$ 을 최소로 하도록 학습을 유도하는 퍼지신경망 OFUN-NET을 사용한다 [6]. OFUN-NET은 다음의 식 (3)과 (4)를 반복적으로 적용하여 클러스터의 대푯값 ( $c_1, c_2, \dots, c_p$ )과 각 데이터의 클러스터 대푯값과의 일치도를 측정하는 소속함수값  $u_{ij}$ 를 학습해 간다.

$$J = \sum_{j=1}^n \sum_{i=1}^p (u_{ij})^2 \| \mathbf{f}_j - \mathbf{c}_i \|^2 \quad (3)$$

$$u_{ij} = \frac{1}{\sum_{k=1}^p \left\{ \frac{\| \mathbf{f}_j - \mathbf{c}_i \|^2}{\| \mathbf{f}_j - \mathbf{c}_k \|^2} \right\}} \quad (4)$$

이러한 학습 알고리즘은 클러스터링의 결과가 만들어 내는 오류 값을 요약하는 퍼지 함수를 설정한 후 그 값이 최소가 되도록 학습의 방향을 유도하는 메커니즘에 의해 진행된다. 또한 제안된 방법은 입력과 출력 사이의 관계를 기술하기 어려운 경우도 쉽게 처리하는 비교사 학습신경망의 장점도 함께 가지고 있다. 이러한 OFUN-NET의 분류결과는 학습된 지식으로 저장되어 진단을 필요로 하는 테스트 데이터의 분류에 사용된다.

## IV. Experiment Results

본 장에서는 성능측정자에 의한 특징서열화 알고리즘을 주된 모듈로 하는 III장에서 제안한 성능측정 에이전트 모델을 포함하는 컴퓨터 보조진단시스템을 구성하여 여러 환자의 데이터에 적용하여 그 타당성을 검증해 보고자한다.

### 1. Preparation of Experimental Data

본 논문에서는 [표 1]에 요약된 것과 같은 공개적으로 널리 알려진 DDSM의 영상[7][18] 으로부터 데이터를 사용하였다. 실험에 사용한 데이터는 MGH(Massachusetts General Hospital), WFUSM(Wake Forest University School of Medicine), WU(Washington University of St. Louis Medical Center) 으로부터 수집한 43개의 볼륨안에 2620개의 케이스로부터 수집한 데이터이다. 준비된 각 환자의 디지털 x-선 영상의 종괴(mass)와 석회(calcification) 환부로부터 진단에 필요한 데이터를 추출하여 시스템의 입력데이터를 만들기 위하여 각 영상으로부터 2가지 방법으로 특징을 추출한다. 우선 American College of Radiology에서 제안한 BI-RADS 표준에 따라 x-선 영상 판독 전문가에 의한 J. Y. Lo가 제시한

rank ordering 방법을 활용하여 특징 데이터로 표현한다[8]. DDSM 안의 각 영상은 .ics 와 .OVERLAY 파일과 연결되어 있으며 이를 이용하여 분류에 필요한 데이터를 획득할 수 있다. .ics 파일로부터 환자의 나이(patient\_age)와 조직의 밀도(density)를 얻을 수 있으며 OVERLAY 파일로부터 각 환부 유형(lesion type)별 각 각 4 가지의 특징데이터를 얻을 수 있다. 또한 [9-10]에서 연구한 결과인 영상처리에 의한 통계적인 14 개의 특징데이터를 추출하여 사용하였다. [표 1]에 요약된 600 개의 케이스 데이터는 학습을 위하여 사용되고, 환자유형별 테스트를 위한 데이터 셋으로서 종괴타입을 위하여 TDBM, 석회타입을 위하여 TDBC를 구성하였다. TDBM, TDBC는 클러스터 데이터베이스 구축을 위하여 학습에 사용된 데이터로부터 양성과 악성 데이터 각각 50개씩 선택하고, 학습에 사용되지 않은 데이터로부터 각 각 50개의 데이터를 선택하여 200개의 케이스를 테스트 데이터로 구성하였다.

Table 1. Data information

| institution | Mass Type |           | Calcification Type |           |
|-------------|-----------|-----------|--------------------|-----------|
|             | Benign    | Malignant | Benign             | Malignant |
| MGH         | 50        | 50        | 38                 | 52        |
| WFUSM       | 50        | 40        | 52                 | 58        |
| WU          | 40        | 60        | 40                 | 60        |
| Total       | 150       | 150       | 130                | 170       |

2. Experiment and Analysis

준비된 데이터를 활용하여 22개 각 속성데이터에 대하여 클러스터분석 알고리즘에 의하여 클러스터를 만들어 각 속성의 성능 측정치 PE 값을 계산한 후 III장에서 제안한 [특징 서열화 알고리즘]을 적용한 결과는 [표 2]와 같다. 이러한 연구의 결과를 컴퓨터 보조진단에 활용하여 그 타당성을 보이기 위하여 [표 3]과 같이 다양한 형태의 실제 학습과 테스트에 사용되는 정제된 데이터 셋을 5가지 준비한다. 환부타입별로 테스트를 위하여 준비한 테스트 데이터 셋 TDBM과 TDBC 각각에 대한 5가지의 실험에서 여러 조합의 특징 데이터를 사용하여 정제된 데이터 셋을 구성하여 실험의 타당성을 확보하고자 하였다. 진단 시스템으로서의 타당성을 보이기 위한 대표적인 방법으로 테스트 케이스의 평균적인 진단결과를 비교하여 [표 3]과 같이 분류정확도로 정리하였다. 처음 세 가지 실험, E1), E2), E3)의 경우는 기존에 연구된 특징을 사용한 경우이며 나이를 포함하는 BI-RADS 기반의 특징 데이터 셋의 경우는 영상처리에 의한 그레이 레벨 기준의 통계적 데이터로 형성된 데이터 셋의 경우보다 분류정확도가 높은 성능을 나타낸 것을 볼 수 있다. E4)의 경우는 본 논문에서 제안한 성능측정 척도에 의하여 서열이 높은 선택된 10개의 특징을 가지고 정제된 데이터 셋을 만들어 실험한 경우로서 기존의 방법보다 높은 분류정확도를 보여주고 있다. 특히 E5)의 경우 특징선택여부를 새로운 특징을 첨가하였을 때 성능척도 값이 감소하는 경우 그 특징을 배

제하고 다음 특징을 선택하는 방법으로 10개의 특징을 선택하였다. 이러한 5가지의 실험 결과는 [표3]과 같은 분류정확도를 나타내며 본 논문에서 제안한 성능측정방법이 개입된 E4)와 E5)의 경우 더 높은 분류정확도를 보여주고 있음을 알 수 있다. 특히 특징으로 선택하면서 제안된 성능척도값이 증가하는 경우 선택하는 E5)의 경우 가장 높은 분류정확도를 보여주었음을 통하여 제안된 방법의 타당성을 확인하였다.

Table 2. Features and Ranking Based on Performance Evaluation

| rank | features                 |
|------|--------------------------|
| 1    | mass shape               |
| 2    | assess                   |
| 3    | age                      |
| 4    | avg. hist.               |
| 5    | num. of pixels           |
| 6    | calc. type               |
| 7    | avg. gray level          |
| 8    | mass margin              |
| 9    | avg. boundary gray level |
| 10   | density                  |
| 11   | mod. entropy             |
| 12   | cal. dist.               |
| 13   | subtley                  |
| 14   | mod. energy              |
| 15   | difference               |
| 16   | entropy                  |
| 17   | mod. skew                |
| 18   | contrast                 |
| 19   | energy                   |
| 20   | skew                     |
| 21   | standard dev.            |
| 22   |                          |

Table 3. Comparison of Classification Accuracy

|   | TDBM | TDBC |
|---|------|------|
| E1) 14 gray level                         | 86.8 | 81.3 |
| E2) 6 BI-RADS                             | 88.5 | 83.2 |
| E3) ALL Features                          | 87.2 | 83.1 |
| E4) Top 10 selected                       | 89.1 | 85.2 |
| E5) Stepwise selection of top 10 features | 91.7 | 88.4 |

V. Conclusions

컴퓨터 진단 시스템의 발전은 인간의 판독오류와 해석과정의 불일치를 줄이기 위하여 영상처리기술과 정보이론이 적용되어 전문가의 의사결정을 도울 수 있는 체계적인 도구로서 의료 정보화에 기여하고 있다.

본 논문에서는 퍼지클러스터분석과 타당성척도에 기반을 둔 성능측정에이전트 모델을 제시하고 실제세계의료기관으로부터 수집되고 공개적으로 제공되는 맘모그램 데이터베이스 DDSM

으로부터 추출된 데이터의 특징을 서열화하였다. 또한 제안된 에이전트를 결합한 컴퓨터 보조진단 시스템을 설계하고 5가지 유형의 특징조합을 가지고 시스템에 적용하여 테스트 하였다. 실험 결과는 제안된 시스템이 보다 정확한 분류정확도를 나타내어 진단 보조 도구로서의 가능성을 보여주었다. 이러한 시스템의 성능은 디지털라이저의 처리과정이나 데이터 형성 과정 등에 크게 좌우되므로 서로 다른 상황에서의 연구 결과와 정량적인 비교는 큰 의미가 없을 수도 있다. 그러나 본 연구는 클러스터분석 타당성 척도와 특징서열화를 연결시켜 특징선택을 위한 일반적인 메카니즘을 만들기 위한 중요한 연구의 시작이 될 수 있을 것이다.

앞으로 타당성척도와 특징서열화 알고리즘에 대한 면밀한 분석을 바탕으로 다양한 데이터베이스에서 수집된 학습 데이터 셋과 테스트 데이터 셋을 가지고 여러 형태의 특징조합을 만들어 다양한 실험을 시도할 필요가 있다. 이러한 과정에서 일반화시킬 부분은 독립모듈로 정립하고 보다 체계적인 시스템으로 발전시켜 나아가야 할 것이다. 또한 대부분의 의사결정이 필요한 시스템에 사용 가능하지만 특히 의료정보화 분야와 같이 제안된 시스템을 보조적으로 사용하여 영향력을 끼칠 수 있는 문제 영역에 대하여 연구가 집중되어야 할 것이다.

## REFERENCE

- [1] Howard Lee and Yi-Ping P. C., "Image based computer aided diagnosis system for cancer detection", *Expert Systems with Applications*, Vol.42, Issue 12, pp. 5356-5365, July 2015.
- [2] N. Petrick and B. Sahinerb, "Evaluation of computer-aided detection and diagnosis systems", *Med Phys*. Vol. 40, No. 8, Aug. 2013.
- [3] Isabelle Guyon and Andre Elisseeff, "An Introduction to Variable and Feature Selection", *Journal of Machine Learning Research* 3, 2003.
- [4] Huan Liu, "Evolving Feature Selection", *IEEE Intelligent Systems and Their Applications* Vol. 20, Issue 4 Nov.-Dec. 2005.
- [5] Z. Vlad, M. D. Ofelia, and T-A. Maria, "Fuzzy Clustering in an Intelligent Agent for Diagnosis Establishment", *Scientific Bulletin of the Petru Maior University of Tirgu Mures* Vol. 6, 2009.
- [6] H. S. Rhee and S. M. Yoon, "Fuzzy Cluster Based Diagnosis System for Digital Mammogram", *Korea Information Processing Society*, Vol. 16-B, No. 2, April, 2009.
- [7] Heath, M., Bowyer, K., Kopans, D., Moore, R., & Jr. Kegelmeyer, P., *The Digital Database for Screening Mammography*, 2001, Medical Physics Publishing.
- [8] J.Y. Lo, et al., "Computer-aided classification of breast microcalcification clusters: Merging of features from image processing and radiologists", 2003.
- [9] R. Panchal and B. Verma, "Characterization of breast abnormality patterns in digital mammograms using autoassociator neural network," in *International Conference on Image Processing 2006, Part III, LNCS*, vol. 4234, pp. 127-136, Springer-Verlag, 2006.
- [10] S. Kim and S. Yoon, "Mass lesions classification in digital mammography using optimal subset of bi-rads and gray level features," in *IEEE proceedings of ITAB2007*, pp99-102, 2007.
- [11] Qinpei Zhao, *Cluster Validity in Clustering Methods*, Dissertations in Forestry and Natural Sciences, Publications of University of Eastern Finland, June, 2012.
- [12] Csaba L., Sandor, J. and Attila B., "Cluster Validity Measurement Techniques", *proc. of the 5th WSEAS International Conference on AI, KE and DB*, 2006.
- [13] Nabih A., Yamina T-G and Nawel Z., "A Computer-Aided Diagnosis System for Breast Cancer Combining Features Complementarily New Scheme of SVM Classifiers Fusion", *International Journal of Multimedia and Ubiquitous Engineering*, Vol.8,No.4, July, 2013.
- [14] R. Panchal and B. Verma, "Characterization of breast abnormality patterns in digital mammograms using autoassociator neural network," in *International Conference on Image Processing 2006, Part III, LNCS*, vol. 4234, pp. 127-136, Springer-Verlag, 2006.
- [15] M. Radovic, M. Djokovic, A. Peulic, and N. Filipovic, "Application of Data Mining Algorithms for Mammogram Classification", 13th conf. on Bioinformatics and Bioengineering(BIBE), 2013.
- [16] L.Sun, L. Li, W. Xu, W. Liu, J. Zhang, and G. Shao, "A Novel Classification Scheme for Breast Masses Based on Multi-view Information Fusion", 4th Int. Conf. on Bioinformatics & Biomedical Engineering (iCBBE), 2010.
- [17] H. Zhao, W. Xu, L. Li, and J. Zhang, "Classification of Breast Masses Based on Multi-view Information Fusion Using Multi-Agent Method", 5th Int. Conf. on Bioinformatics & Biomedical Engineering(iCBBE), 2011.
- [18] Reference URL : <http://marathon.csee.usf.edu/Mammography/Database.html>

### Authors



Hyun Sook Rhee received the B.S., M.S. and Ph.D. degrees in Computer Science and Engineering from Sogang University, POSTECH, Sogang University in 1989, 1991 and 1997, respectively.

Dr. Rhee joined the faculty of the Department of Software Engineering at Dongyang Mirae University, Seoul, Korea, in 1997. She is currently a Professor in the Department of Software Engineering at Dongyang Mirae University. She is interested in intelligent information system, data mining, and software development methodology.