

# RBF와 LVQ 인공신경망을 이용한 尿 덱스틱 선별검사에서의 요로감염 분류

민경기 강명서 신기영 이상식 문정환

## Classification of UTI Using RBF and LVQ Artificial Neural Network in Urine Dipstick Screening Test

K. K. Min M. S. Kang K. Y. Shin S. S. Lee J. H. Mun

### Abstract

Dipstick urinalysis is used as a routine test for a screening test of UTI (urinary tract infection) in primary practice because urine dipstick test is simple. The result of dipstick urinalysis brings medical professionals to make a microscopic examination and urine culture for exact UTI diagnosis, therefore it is emphasized on a role of screening test. The objective of this study was to the classification between UTI patients and normal subjects using hybrid neural network classifier with enhanced clustering performance in urine dipstick screening test. In order to propose a classifier, we made a hybrid neural network which combines with RBF layer, summation & normalization layer and LVQ artificial neural network layer. For the demonstration of proposed hybrid neural network, we compared proposed classifier with various artificial neural networks such as back-propagation, RBFNN and PNN method. As a result, classification performance of proposed classifier was able to classify 95.81% of the normal subjects and 83.87% of the UTI patients, total average 90.72% according to validation dataset. The proposed classifier confirms better performance than other classifiers. Therefore the application of such a proposed classifier expect to utilize telemedicine to classify between UTI patients and normal subjects in the future.

**Keywords :** RBF-LVQ hybrid artificial neural network, UTI classification, Dipstick urinalysis, Neural network classifier

### 1. 서론

농업인은 힘겨운 노동과 농약중독 등으로 인하여 각종 질병에 노출되어 있으나 지리적 여건상 제때에 치료를 받지 못하고 있다. 또한 농촌의 고령화 비율은 2000년에는 15.6%로 도시의 6.7%에 비해 2배 이상으로 이미 고령화 사회에 이르고 있다. 따라서 농업인의 발병률이 높은 질병에 대하여 신속한 진단을 할 수 있는 원격의료서비스의 운영이 필요하다. 농촌 환경에서 주로 발생할 수 있는 유병율이 높은 질병들은 관

절염, 요통, 좌골통, 신경통 등 근골격계관련 만성질환들이 주를 이루고 있다. 그러나 농업인은 다른 직종에 종사하는 사람들 보다 만성과 급성질환에서 모두 유병율이 높게 나타나고 있으며 직업 관련성도 상당히 높은 것으로 보고되고 있어 (Lee, 2006), 신속한 진단과 치료가 필요한 급성질환에 대한 연구도 필요한 실정이다. 따라서 본 연구에서는 급성질환들 중의 하나인 요로감염(urinary tract infection)을 선택하여, 선별검사 방법을 제안함으로써 질병을 사전에 예방하여 만성질환으로 이어지는 것을 방지하기위하여 연구하였다.

This study was supported by a grant of the Korea Healthcare technology R&D Project, Ministry for Health, Welfare and Family Affairs, Republic of Korea (A060484). The article was submitted for publication on 2008-07-31, reviewed on 2008-08-26, and approved for publication by editorial board of KSAM on 2008-09-17. The authors are Kyoung Kee Min, KSAM member, Ph. D., Dept. of Biomechatronic Engineering, Sungkyunkwan University, Myung-Seo Kang, Professor, Dept. of Laboratory Medicine, College of Medicine, Pochon CHA University, Ki Young Shin, KSAM member, Ph. D. Candidate, Sang Sik Lee, KSAM member, Research Associate Professor, and Joung Hwan Mun, Associate Professor, Dept. of Biomechatronic Engineering, Sungkyunkwan University. Corresponding author: J. H. Mun, Associate Professor, Dept. of Biomechatronic Engineering, Sungkyunkwan University, Suwon, 440-746, Korea; Tel: +82-31-299-4822; Fax: +82-31-299-4825; E-mail: <jmun@skku.ac.kr>.

요로감염은 요를 배설하는 기관인 신장, 요관, 방광, 요도의 요로계에 여러 가지 원인으로 균이 침입하여 발생하는 질병이다. 여성과 노인에게 흔한 감염증으로 여성의 경우는 약 40~50%가 일생 중 요로감염을 경험하는 것으로 알려져 있으며, 발병률이 높은 비뇨기 질환의 하나로 조기 선별검사가 중요한 질병이다. 요로감염의 검사방법에는 요 배양검사와 현미경을 이용한 요 침사검사 그리고 생화학적 검사법인 요 덤스틱 검사 등이 주로 이용되고 있다. 요 배양검사는 요로감염 증의 원인균까지 규명할 수 있기 때문에 표준검사의 역할을 하여왔다. 그러나 요 배양검사에서 요로감염의 확진으로 볼 수 있는 의미있는 세균농도( Colony-forming units) 시  $10^5/\text{mL}$  이상으로 규정되어 있으나 실제 검체의 검사치는  $10^4/\text{mL}$  미만의 결과를 보이고 있어서 민감도가 떨어진다는(Bae 등, 2000). 그리고 현미경 검사는 백혈구의 개수를 세어야 하고 적혈구와 세균을 검사해야 한다. 따라서 이러한 방법들은 시간과 비용이 소모되는 단점이 있으며, 또한 선별검사 방법으로도 적합하지 않기 때문에 1차 진료기관의 선별검사에서는 시행이 간단한 덤스틱 요검사가 많이 활용되고 있다(Wilson 등, 1991; Choi 등, 2000; Jeon 등, 2003). 덤스틱 요검사는 여러 검사항목들이 이용된다. 검사항목들은 일반적으로 잠혈, 빌리루빈, 유로빌리노젠, 케톤체, 단백질, 아질산염, 포도당, pH, 비중 그리고 백혈구가 이용되고 있다(Jeon 등, 2003; Kim 등, 2006). 그리고 이러한 요 덤스틱 항목들의 화학적 반응인 정색반응의 정도를 통하여 의료전문가들은 각종 질병들을 선별검사할 수 있게 된다.

요로감염을 예측하기 위한 기존의 연구들(Wigton 등, 1985; Leibovici 등, 1989; Heckerling 등, 2007)은 환자의 증상, 징후, 의사의 임상적 소견 그리고 요검사 등을 통하여 요로감염 유발에 관계가 있다고 판단되는 많은 인자들을 선택하였다. 그리고 선택된 인자들에 대하여 단계적 판별분석과 로지스틱 분석 등의 변수 선택방법을 적용하여, 요로감염에 영향을 미치는 최적의 변수들을 찾은 후 결정규칙과 예측모델을 개발하였다. 또한 이런 과정을 통하여 요배양검사와 현미경검사 등에서 소모되는 시간과 비용을 절약하는 이점도 함께 추구하였다. 기존의 선행 연구들에서 요로감염을 분류하기 위하여 최종으로 선택된 변수들은 현미경적 농도, 혈뇨, 세균뇨, 요 덤스틱의 백혈구 에스테라제, 세균과 상피세포 등으로 다양하게 나타났다. 그러나 선행연구들에서 선택된 변수들은 여전히 요 배양과 현미경 그리고 덤스틱 검사가 함께 사용되어야 검사할 수 있는 인자들이다. 따라서 이러한 결과들은 1차 진료기관에서 각종 질병을 조기에 발견하고 검사하기 위한 선별검사의 목적에 적용하기에는 유용하지 못하며, 시간과 비용 면에서도 효율적이지 않다. 그리고 선행연구들(Wigton 등,

1985; Leibovici 등, 1989)의 주된 관점은 요로감염에 영향을 미치는 주요 인자들을 통계적 분석방법들을 이용하여 추출하는 것으로서, 인공신경망 분류기를 적용하여 요로감염을 분류하는 연구는 미흡한 실정이다.

본 연구에서는 생체정보와 같은 비선형 데이터를 분류하는데 있어서 통계적 분석방법 보다 인공신경망 방법이 성능 면에서 우수한 것으로 보고되고 있기 때문에(Li 등, 2000; Lisboa 등, 2006), 분류기로 인공신경망을 적용하였다. 기존의 문헌에서 인공신경망 방법을 적용하여 요로감염을 분류한 사례가 있으나(Heckerling 등, 2007), 이 문헌은 역전파 신경망을 이용하여 요 배양검사 결과를 목표값으로 학습시킨 경우이다. 그러나 본 연구는 요 덤스틱 검사만을 이용하여 의사가 질병을 검사한 진단결과를 사용하여 요로감염을 분류한 점에서 기존의 선행 연구들(Wigton 등, 1985; Leibovici 등, 1989; Heckerling 등, 2007)과 차별화 된다.

따라서 본 연구의 목적은 군집능력이 향상된 혼합형 인공신경망을 이용하여 요 덤스틱 선별검사서 요로 감염증 환자와 정상 피험자를 분류하려 하였다. 이를 위하여 RBF(Radial Basis Function)층, 합산 및 정규화층 그리고 LVQ(Learning Vector Quantization) 신경망층을 혼합한 4단 구조의 분류기를 제안하였고 성능을 검증하기 위하여 역전파 신경망법, RBFNN(Radial Basis Function Neural Network)법 그리고 PNN(Probabilistic Neural Network)법과 같은 일반적인 방법들의 결과와 비교하여 검증하였다.

## 2. 재료 및 방법

### 가. 데이터 수집 및 전처리

데이터 획득기기는 항목이 10개인 덤스틱(유리스캔 10 에스지엘 시험지, 영동제약, 한국) 방식의 요 화학 분석기(Urine Chemistry Analyzer, UriScan PROPlus, 영동제약, 한국)를 이용하였다. 임상데이터는 10년간 수집된 국내 C병원의 데이터들이다. 획득된 데이터들은 요 화학 분석기로부터 피험자들의 데이터를 획득한 후 의료전문가에 의해 정상과 요로감염증으로 진단된 요 덤스틱 임상결과 데이터들이다. 전체 임상데이터들은 1,164개로 요로감염증으로 진단된 496명의 환자 데이터와 건강검진을 위해 내원한 668명의 정상 대조군 데이터이다. 데이터 전처리 과정은 요 덤스틱의 10개 항목에 대하여 각 항목별 최대값으로 정규화를 적용한 후 적용된 인공신경망 분류기들의 입력 값으로 사용하였다.

정규화 과정은 표 1과 같이 요 덤스틱의 원래 시험부분의 항목별 민감도에 따른 레벨값들을 비색단계 별로 정규화(불

드체)를 한 변환표를 만들어 사용하였다. 덱스틱의 10개 항목들은 잠혈(occult blood), 빌리루빈(bilirubin), 유로빌리노겐(urobilinogen), 케톤체(ketones), 단백질(protein), 아질산염(nitrite), 포도당(glucose), pH, 비중(specific gravity) 그리고 백혈구(leukocytes)를 의미한다. 그리고 인공신경망 분류기들의 데이터 집합의 구성은 훈련데이터 집합은 582개(정상 대조군: 334명, 요로감염 환자군: 248명)이며, 검증 데이터집합도 훈련 데이터집합과 같은 수로 구성하였다.

본 연구에서 요 덱스틱의 10개 항목을 사용한 이유는 요로감염증이 증상과 농도와 세균농도 사이에 비일관적인 관계가 있어서 임상에서 진단이 어려운 질병으로 알려져 있기 때문이다(Sultana, 2001). 몇 가지 증상들과 징후들이 요로감염의 가능성을 예측할 수 있도록 해주지만 진단을 위한 규칙에 대해서는 아직도 많은 연구가 진행 중에 있다(Bent 등, 2002; Heckerling, 2007). 문헌상 나타나는 요 덱스틱 검사만의 인자들은 혈뇨, 백혈구, 단백뇨, 아질산염 등으로 나타나지만, 요로감염을 유발하는 모든 원인균(uropathogen)들이 아질산염을 발생시키는 것은 아니며, 다른 인자들도 요로감염 외의

질병에서도 검출되기 때문에 요로감염에 대한 확진인자는 아직까지는 백혈구가 유일한 것으로 알려져 있다(Sultana, 2001). 따라서 덱스틱 항목들 중에서 요로감염의 유발인자들만을 분류기에 적용하는 것이 분명하지 않은 점들이 있으며, 또한 선행연구들에 비해 분류율을 향상시키기 위하여 덱스틱의 10개 항목들을 사용하였다.

나. 기존의 인공신경망 분류기

제안된 4단 혼합 인공신경망의 성능을 검증하기 위하여 기존의 역전파 신경망, RBFNN, PNN 인공신경망 방법들을 비교 분류기로 선택한 이유는 다음과 같다. 역전파 신경망은 입력패턴에 대한 분류 에러를 줄이는 특성이 있기 때문에(Oh, 1996), 패턴들을 분류하는데 효과적일 것으로 판단되어 적용하였다. 그리고 역전파 신경망은 대표적인 MLP(Multi Layer Perceptron) 방식으로 제안된 4단 혼합 인공신경망 분류기가 RBF 함수를 기반으로 하기 때문에 서로 비교가 되기도 한다. RBFNN은 역전파 신경망에 비하여 연산속도가 빠르고 은닉층이 한 개이므로 구조적으로 단순하며, 가우시안 커널함수

Table 1 Normalization conversion table according to color chart levels

Factors	Sensitivity	Color chart levels					
		-	±	+	++	+++	++++
Occult blood	5 RBC/μl, 15 μg Hemoglobin/dL	0		10	50	250	
	<b>Normalize</b>	<b>0</b>		<b>0.04</b>	<b>0.20</b>	<b>1.00</b>	
Bilirubin	0.5 mg/dL	0		0.50	1.00	3.00	
	<b>Normalize</b>	<b>0</b>		<b>0.17</b>	<b>0.33</b>	<b>1.00</b>	
Urobilinogen	0.1 unit/dL		0.1	1	4	8	12
	<b>Normalize</b>		<b>0.01</b>	<b>0.08</b>	<b>0.33</b>	<b>0.67</b>	<b>1.00</b>
Ketones	5 mg/dL acetoacetic acid	0	5	10	50	100	
	<b>Normalize</b>	<b>0</b>	<b>0.05</b>	<b>0.10</b>	<b>0.50</b>	<b>1.00</b>	
Protein	10 mg/dL albumin	0	10	30	100	300	1000
	<b>Normalize</b>	<b>0</b>	<b>0.01</b>	<b>0.03</b>	<b>0.10</b>	<b>0.30</b>	<b>1.00</b>
Nitrite	0.05 mg/dL	0		0.05			
	<b>Normalize</b>	<b>0</b>		<b>1.00</b>			
Glucose	50 mg/dL	0	100	250	500	1000	2000
	<b>Normalize</b>	<b>0</b>	<b>0.05</b>	<b>0.13</b>	<b>0.25</b>	<b>0.50</b>	<b>1.00</b>
pH	±1 pH unit	5	6	7	8	9	
	<b>Normalize</b>	<b>0.56</b>	<b>0.67</b>	<b>0.78</b>	<b>0.89</b>	<b>1.00</b>	
Specific gravity	±0.005	1.00	1.01	1.01	1.02	1.03	1.03
	<b>Normalize</b>	<b>0.97</b>	<b>0.98</b>	<b>0.98</b>	<b>0.99</b>	<b>1.00</b>	<b>1.00</b>
Leukocytes	10 WBC/μl	0		25	75	500	
	<b>Normalize</b>	<b>0</b>		<b>0.05</b>	<b>0.15</b>	<b>1.00</b>	

에 의하여 비선형성이 강한 군집화의 기능을 하기 때문에 (Hines, 1997; Shin 등, 2001), 데이터 분류가 우수하다고 판단되어 사용하였다. PNN은 반복적으로 훈련되는 인공신경망인 역전파 신경망과 일반화된 회귀신경망인 GRNN(Generalized Regression Neural Network) 보다 분류가 잘 수행될 수 있고, 학습속도가 빠른 것으로 보고되어(Boniatis 등, 2007) 사용하였다.

**다. 개발 알고리즘**

제안된 4단 혼합 인공신경망 방법은 RBF 함수층, 합산과 정규화층 그리고 LVQ 네트워크를 조합함으로써 분류할 입력데이터들에 대해 이중 군집화를 수행하여 신경망의 군집능력이 향상되는 구조를 갖도록 고안하였다. 따라서 제안된 방법의 독창성은 이중 군집화를 위하여 설계된 인공신경망의 다단 구조와 성능에 있다. 기존의 방법들 중에서 알고리즘의 구성면에서 본 연구에서 제안한 방법과 똑같은 구성을 갖는 알고리즘은 문헌상 찾지 못하였다. 그리고 기존의 RBFNN, PNN, RBPNN(Radial Basis Probabilistic Neural Network), GRNN, LVQ 방법들은 신경망의 네트워크 구조 중 일부가 본 연구에서 제안한 방법과 부분적으로 비슷한 점은 있지만 기존의 방법들은 2단 또는 3단 구조(Hines, 1997; Huang, 1999)를 갖는 반면 제안된 방법은 4단 구조를 갖는 점이 분명 다르다. 그리고 성능면에서 제안된 분류기는 4단 구조에서 RBF 함수층과 합산 및 정규화층은 전처리 층으로서 군집화와 특징추출의 기능을 하며, 이어지는 LVQ 네트워크는 군집화를 한번 더 수행하여 이중 군집화를 하도록 한점이 기존의 방법들과 차별화된다. 그림 1은 제안된 4단 혼합 인공신경망의 개괄적인 구성을 블록 다이어그램으로 나타낸 것이다.

네트워크의 구조와 기능면에서 제안된 분류기의 설계원리는 다음과 같다. 첫 번째 단은 RBFNN 신경망의 후단인 선형층을 제거하여 RBF 함수층의 고유 기능인 군집기능을 갖게

하였다. 두 번째 단은 입력패턴들에 대한 특징을 추출하기 위하여 합산과 정규화를 수행하는 층을 연결하였다. 이는 RBFNN의 선형층이 감독 학습의 기능을 하기 때문에 선형층을 제거함으로써 RBF 함수의 다음에 연결되는 네트워크를 비감독 학습을 하는 합산 및 정규화 층을 붙이기 위해서다. 감독학습은 입력과 목표값이 주어진 상태에서 목표값에 대해 학습자료를 학습하는 것이고, 비감독 학습은 목표값이 없는 상태에서 스스로 주어진 입력자료를 학습한다. 그리고 LVQ 네트워크의 연결을 통하여 군집을 한번 더 수행함으로써 군집능력이 한층 더 향상되도록 설계하였다. 그림 2는 본 연구에서 제안한 4단 혼합 인공신경망의 전체구조이다.

그림 2를 이용하여 제안된 4단 혼합 인공신경망의 연산과정에 대한 전개식을 설명하면 다음과 같다.

**1) 입력층**

입력층은 비감독 학습 네트워크로서 요 덩스틱의 10개 항목들이 입력되며, 입력층과 첫 번째 은닉층( $H_1$ )은 연결강도  $\omega_1$  을 갖는다.

**2) 첫 번째 은닉층**

첫 번째 은닉층( $H_1$ )은 가우시안 커널함수를 갖는 비선형 처리층으로 RBFNN의 첫단과 같다. 식 (1)은 클러스터링 역할을 하는 RBF 함수의 출력식이다.

$$H_{1j} = \exp\left(-\frac{\|x - c_j\|^2}{\sigma_j}\right) \tag{1}$$

여기서,  $H_{1j}$ 는 첫 번째 은닉층의  $j$ 번째 출력,  $x$ 는 입력패턴,  $c_j$ 는 채택필드 영역의 중심,  $\sigma_j$ 는 채택군집의 범위(표준편차)를 의미한다(Hines, 1997).

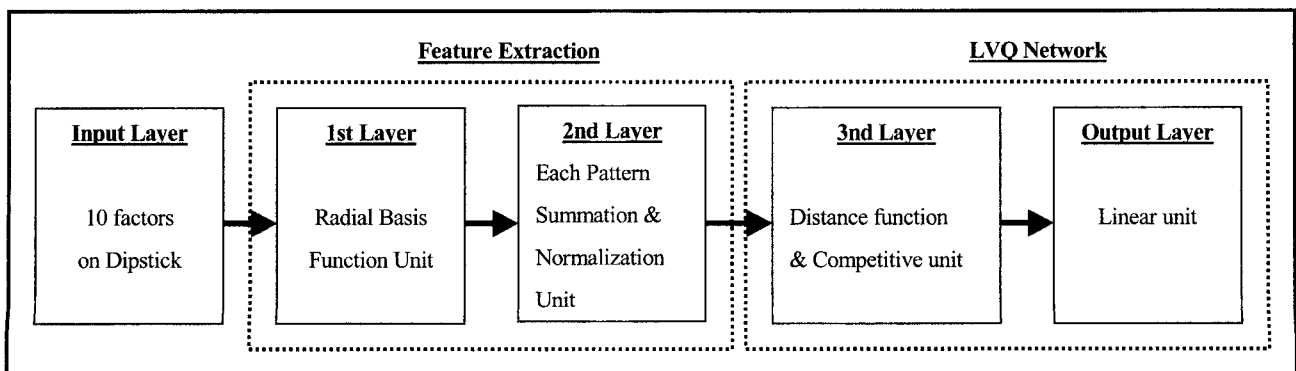


Fig. 1 The block diagram of proposed 4 stages hybrid artificial neural network.

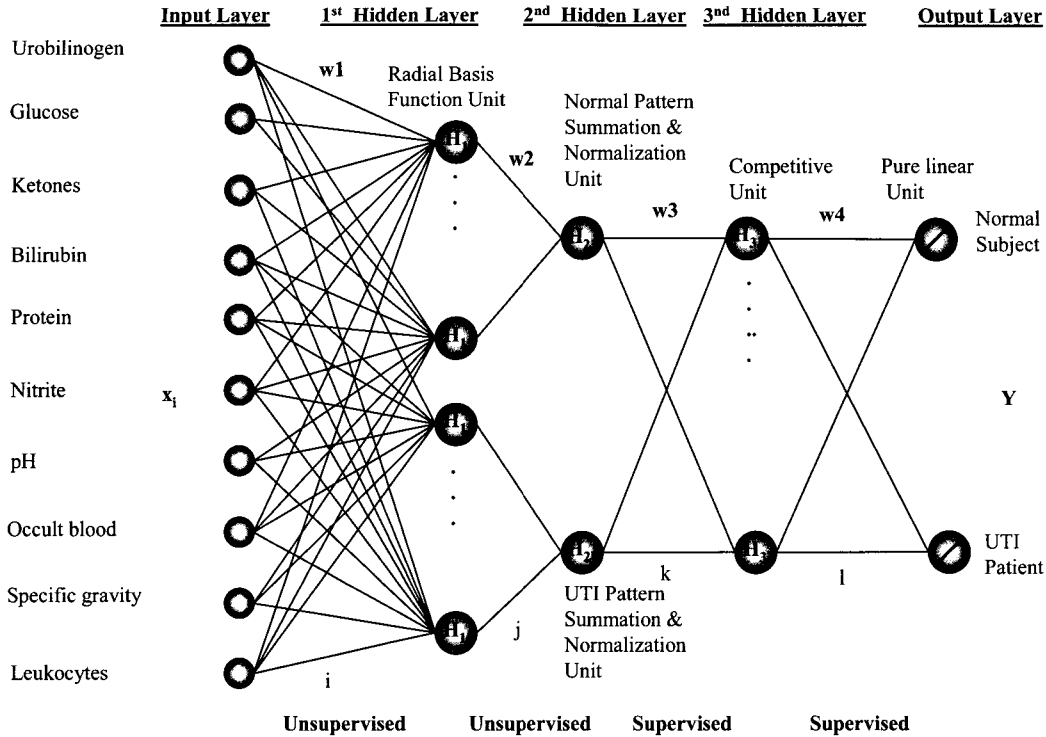


Fig. 2 The structure of proposed 4 stages hybrid neural network.

3) 두 번째 은닉층

두 번째 은닉층( $H_2$ )은 비감독 학습 네트워크로 합산 및 정규화 기능을 하는 층이다. 연결강도( $\omega_2$ )는 입력패턴들의 군집을 클래스별로 구분하기 위하여 가우시안 확률함수에 의해 0과 1로 군집이 이루어진다(Hines, 1997). 두 번째 은닉층의 연결강도( $\omega_2$ )는 연산층면에서 분류하고자 하는 패턴들을 군집별로 나누기 위한 것으로 비감독 학습의 역할을 한다. 두 번째 은닉층( $H_2$ )의 결과는 첫 번째 은닉층( $H_1$ )의 출력결과와 연결강도( $\omega_2$ )의 곱으로 나타나는 식 (2)가 된다.

$$H_2(x) = H_1 \cdot \omega_2 \tag{2}$$

그리고 두 번째 은닉층은 합산과 정규화 연산을 하므로, 식 (2)의 결과에서 먼저 합산기능은 식 (3)으로 표현된다.

$$\text{sum}(H_{2pj}(x)) = \sum_{j=1}^M H_{2pj}(x) \tag{3}$$

여기서,  $\text{sum}(H_{2pj}(x))$ 는 정상과 요로감염에 대한 각각의 입력패턴별 합산결과이며,  $M$ 은 첫 번째 은닉층의 출력 패턴 수로 제안된 4단 혼합 인공신경망의 출력층의 노드수와 같다. 식 (4)는 합산 후 0과 1사이의 값으로 정규화까지 시킨 두 번째 은닉층( $H_2$ )의 출력결과이다.

$$H_{2k}(x) = \text{sum}(H_{2pj}(x)) / \max(H_{2pj}(x)) \tag{4}$$

여기서,  $H_{2k}(x)$ 는 두 번째 은닉층( $H_2$ )의 최종 출력결과이다.

4) 세 번째 은닉층

세 번째 은닉층( $H_3$ )은 LVQ 네트워크의 첫 번째 입력층으로 먼저 훈련벡터의 거리를 계산한 후 가장 가까운 유닛을 승자로 취하게 되는 경쟁적 신경망이다. 식 (5)와 같이  $N$ 차원의 입력벡터  $X$ 가 주어졌을 때 연결강도 벡터  $\omega_3$ 와의 유클리드 거리( $D$ )의 계산은 식 (6)으로 표현되며 입력벡터들의 거리를 계산한다.

$$X = (x_1, x_2, \dots, x_N) \tag{5}$$

$$D = - \| H_{2k}(x) - \omega_3 \| \tag{6}$$

여기서,  $H_{2k}(x)$ 는 두 번째 은닉층( $H_2$ )의 최종 출력결과,  $\omega_3$ 는 세 번째 은닉층의 연결강도 벡터,  $\| \|$ 은 뉴클리드 거리를 의미한다.

LVQ 네트워크의 학습규칙은 입력벡터( $X$ )와 연결강도 벡터( $\omega$ )의 차이( $D$ )를 유클리드 거리로 계산한 후에  $D$ 의 값이 가장 최소인 뉴런을 승자뉴런으로 결정한다. 이후에 진행되는 학습과정은 승자뉴런이 참조패턴에 대해 기준뉴런으로 결

정되어 연결강도 벡터에 더해지는 것으로 학습이 진행된다. 이를 수식으로 표현하면 식 (7)이 되어  $D$ 값이 해당 클래스 ( $C$ )에 들면 연결강도를 조정하고, 그렇지 않으면 연결강도는 조정되지 않는다.

$$\begin{aligned} X \in C_j \text{이면, } Z(t+1) &= Z(t) + \alpha(t) \cdot [X - Z(t)], \\ X \notin C_j \text{ 이면 } Z(t+1) &= Z(t) \end{aligned} \quad (7)$$

여기서,  $\alpha(t)$ 는 네트워크의 학습율이고 시간에 따라 감소하게 된다. 학습의 초기단계에서는 경쟁층 내의 모든 뉴런들이 포함될 수 있으며, 이웃 반경은 훈련이 진행됨에 따라 서서히 줄어들어 점점 적은 개수의 뉴런들이 학습을 하게 된다. 따라서 최종적으로 승자 뉴런만이 연결강도를 조정하게 되어 학습이 끝난 후의 각 뉴런의 연결강도 벡터( $\omega_3$ )는 기하학적으로 입력패턴들과 유사성을 가지게 된다(Oh, 1996).

### 5) 출력층

출력층은 최종 패턴 분류를 위한 선형층으로 식 (8)로 표현된다.

$$Y = X \cdot \omega_4 \quad (8)$$

여기서,  $Y$ 은 제안된 4단 혼합 인공신경망의 최종 출력식이며,  $X$ 는 LVQ 네트워크에 입력되는 두 번째 층의 출력벡터이다. 그리고  $\omega_4$ 는 네 번째 층의 연결강도이다.

### 라. 분류기들의 네트워크 구조와 파라미터

적용된 인공신경망 분류기들의 네트워크 파라미터들은 표 2에 나타내었다. 표 2의 파라미터들은 적용된 분류기들이 모두 훈련단계에서 학습율이 100%가 되도록 조정된 최적의 값들로서 제안된 4단 혼합 인공신경망 분류기와 성능을 서로 비교하기 위하여 수동으로 최적화를 행한 값들이다(Fig. 3).

적용된 분류기들의 구현에 대한 개괄적인 설명은 다음과 같다. 역전파 신경망의 구성은 10개의 노드를 갖는 입력층,

12-8 노드를 갖는 2개의 은닉층과 1개의 노드를 갖는 출력층으로 구성하였다. 전달함수는 모든 계층에 대해 tangent sigmoid 함수를 사용하였다. systematic error rate은 목표값을 0.0001로 설정 하였다. 그리고 제안된 4단 혼합 인공신경망과 RBFNN 과 PNN의 spread 값은 사전에 최적값을 수동으로 찾아 가장 우수한 값인 0.0015를 선택하였다. 적용된 분류기들은 학습 후 분류를 위하여 임계값을 0.5값으로 적용하여 임계값 이하 이면 정상으로 이상이면 요로감염으로 분류하였다. 그리고 제안된 4단 혼합 인공신경망의 후단인 LVQ 네트워크에서 은닉층 노드의 수는 10개, 패턴별 비율치(class percentages)는 정상과 요로감염 패턴에 대해 0.57과 0.43을 각각 주었다.

## 3. 결과 및 고찰

### 가. 적용된 신경망 분류기들의 분류결과

제안된 4단 혼합 인공신경망의 성능을 입증하기 위하여 역전파 신경망, RBFNN, PNN 신경망들을 이용하여 분류성능을 서로 비교하였다. 표 3은 훈련과 검증 데이터집합에 대한 분류기들의 분류결과를 나타내었다. 훈련 데이터집합에 대해서 분류기들 모두 100%로 학습되었다. 그리고 검증 데이터집합에 대해서 역전파 신경망은 정상을 285개 분류하여 85.33%, 요로감염은 182개 분류하여 73.39%의 분류율을 보여 평균 분류율은 80.24%를 나타냈다. PNN 분류기는 정상을 287개 분류하여 85.93%, 요로감염은 185개 분류하여 74.60%의 분류율을 보여 평균 분류율은 81.10%를 나타냈다. RBFNN 분류기는 정상을 273개 분류하여 81.74%, 요로감염은 51개 분류하여 20.56%의 분류율을 보여 평균 분류율은 55.67%를 나타냈다. 제안된 4단 혼합 인공신경망 분류기는 정상을 316개 분류하여 95.81%, 요로감염은 211개 분류하여 83.87%의 분류율을 보여 평균 분류율은 90.72%를 나타냈다. 따라서 분류기들의 분류 결과를 통하여 볼 때 제안된 4단 혼합 인공신경망의 분류율이 가장 높은 것으로 나타났다.

그림 3은 검증 데이터집합에 제안된 4단 혼합 인공신경망을 적용한 표 3의 분류결과이다.

본 연구의 실험결과에 대한 주요 고찰 사항들은 다음과 같다. 본 연구는 유병율이 높은 것으로 알려져 있는 요로감염증을 선별하는 질병분류 알고리즘을 개발하여 향후 요 화학 분석기를 원격의료시스템에 적용하기 위하여 연구하였다. 현재 시판되고 있는 요 화학 분석기들은 여러 회사의 제품들이 있다. 그러나 시판되는 요 화학 분석기들은 요 디스틱의 화학적 반응을 통하여 검사항목들의 개별 성분들을 레벨 값으로만 제시할 뿐이며 질병의 분류는 해주지 않는다. 이러한 사실은

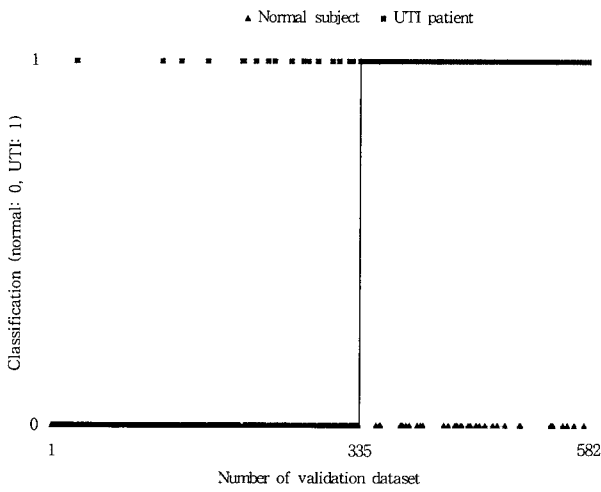
**Table 2** The network parameters of applied neural network classifiers

Network Parameters	Type of classifiers			
	Hybrid	RBFNN	PNN	Back-propagation
spread	0.0015	0.0015	0.0015	-
epoch	5	-	-	3,000
learning rate	0.1	-	-	0.1
momentum	-	-	-	0.95
threshold	0.5	0.5	0.5	0.5

**Table 3** The comparison of classification rate according to each classifiers

(Norm: Normal, Avr: Average)

Type of classifier	Classification Rate (%)					
	Training Dataset			Validation Dataset		
	Norm	UTI	Avr	Norm	UTI	Avr
Hybrid	100	100	100	95.81	83.87	90.72
RBFNN	100	100	100	81.74	20.56	55.67
PNN	100	100	100	85.93	74.60	81.10
Back-propagation	100	100	100	85.33	73.39	80.24



**Fig. 3** Validation result of proposed 4-staged hybrid neural network classifier.

검사자의 시각에 의존하여 요 디스틱의 반응된 색을 보아야 하기 때문에 고령화 사회로 가는 농촌의 현실에 맞지 않는다. 따라서 디스틱 요검사를 이용하여 요로감염을 선별검사하는 실용적인 취지에서 검사자의 시각검사에 의존하지 않고, 디스틱 요검사만을 이용하여 각종 질환들을 동시에 선별검사하는 자동화된 방법이 향후 요구된다. 이러한 자동화된 방법을 통하여 질병을 조기에 발견하여 농촌의 의료비 감소와 건강 증진 등의 긍정적인 효과를 기대할 수 있을 것으로 판단된다.

본 연구에서 디스틱 요검사 방법만을 이용하여 요로감염을 분류한 이유는 요배양 및 현미경 검사 방법과 비교시 디스틱 요검사의 효용성을 보고한 문헌들이 있기 때문이다. 디스틱 요검사의 효용성을 주장한 선행연구들(Choi 등, 2000; Jeon 등, 2003)은 요 배양 또는 현미경 검사 결과에 대하여 요 디스틱 검사 결과를 서로 비교하였다. 이점에 대하여 요 디스틱의 분류결과를 중심으로 설명하면 다음과 같다. 요로감염 선별검사서 백혈구와 아질산염 항목을 이용하여 요 디스틱 검사가 요배양과 현미경적 요검사를 대신할 수 있는지를 연구한 문헌(Choi 등, 2000)이 있다. 이 문헌에서는 16세 이상의 성인에 대하여 요 디스틱 검사가 현미경적 검사의 분류율

(73.6%)과 함께 나타나서 현미경적 요검사를 대신 할 수 있다고 보고하였다. 그리고 요로감염의 진단과 치료를 위하여 세균뇨를 신속히 발견하기 위한 방법을 제안한 연구(Sultana 등, 2001)에서 그래프 염색법과 요 디스틱의 백혈구와 아질산염 항목을 함께 적용하였다. 이 문헌에서는 요배양 검사와 요 디스틱검사 결과를 서로 비교하였다. 연구결과는 요 디스틱의 아질산염과 백혈구 검사 각각의 결과를 요배양 검사와 비교하였을 때 아질산염 항목은 민감도가 27.0%, 특이도는 97.4%로 나왔다. 그리고 백혈구 항목은 민감도가 45.9%, 특이도는 91.5%로 나와 두 항목 모두에서 민감도가 매우 낮게 나왔다. 따라서 아질산염과 백혈구 항목검사 단독으로는 요로감염 선별검사에 사용하기는 어렵다고 보고하였다. 따라서 본 연구에서는 요 디스틱의 10개 항목을 사용하여 분류율을 향상시키는 방법을 선택하였다. 그밖에 요디스틱 방법을 사용한 문헌들(Wigton 등, 1985; Leibovici 등, 1989; Little 등, 2006; Heckerling 등, 2007)은 있으나, 요 디스틱 항목들만을 적용한 연구라기 보다는 요배양과 현미경 그리고 요 디스틱 방법들을 함께 이용한 경우이다. 이러한 문헌들은 요로감염을 예측하기 위하여 결정규칙 및 임상모델 개발 그리고 예측인자들을 제안한 연구들이다. 본 연구는 요 디스틱 항목들만을 이용하여 요로감염을 예측하기위한 방법을 제시하고 이를 임상에 적용하기위한 연구라는 점에서 기존의 문헌들과 같은 목적을 갖는다.

#### 4. 요약 및 결론

본 연구는 요 디스틱 임상데이터들을 이용하여 요로감염증 환자와 정상 피험자를 선별검사하기 위한 연구를 수행하였다. 그리고 이를 위하여 RBF 함수층과 합산 및 정규화층 그리고 LVQ 네트워크를 혼합한 4단 구조의 인공신경망 분류기를 제안하고 그 성능을 검증하였다. 주요 연구결과를 요약하면 다음과 같다.

- (1) 제안된 4단 혼합 인공지능망의 검증데이터에 대한 분류율은 정상이 95.81%, 요로감염은 83.87%, 전체 평균은 90.72%로 나와 적용된 분류기들 중에서 가장 높은 분류성능을 나타내었다.
- (2) 요로감염에 대한 분류율이 83.87%로서 매우 양호한 결과는 아니지만 정밀검사가 아닌 선별검사용으로서는 적절한 것으로 판단된다.
- (3) 본 연구는 요배양 검사나 현미경검사 방법이 아닌 요 디스틱 검사만으로도 요로감염의 예측이 가능함을 제안된 4단 혼합 인공지능망의 적용을 통하여 입증하였다.
- (4) 검사에 드는 시간과 비용을 감소시킬 수 있다는 점에서 제안된 4단 혼합 인공지능망 분류기가 향후 요로감염과 정상을 선별검사하기 위한 요 화학 분석기와 농촌의 원격의료 서비스에도 유용하게 적용될 수 있을 것으로 기대된다.

## 참 고 문 헌

1. Bae, S. I., H. C. Lee, S. Y. Lim, K. D. Kim and B. C. Jung. 2000. Cytocentrifuge gram stain method and urine dipstick test as a screening test of bacteriuria. *Korean J. Clin. Pathol.* 20(4):410-414.
2. Bent, S., B. K. Nallamothu, D. L. Simel, S. D. Fihn and S. Saint. 2002. Does this women have an acute uncomplicated urinary tract infection?. *JAMA* 287(20):2701-2710.
3. Boniatis, I., L. Costaridou, D. Cavouras, I. Kalatzis, E. Panagiotopoulos and G. Panayiotakis. 2007. Assessing hip osteoarthritis severity utilizing a probabilistic neural network based classification scheme. *Medical Engineering & Physics* 29(2): 227-237.
4. Choi, S. D., H. J. Cho, D. Y. Cho, B. Y. Yu and K. H. Kim. 2000. Diagnostic value of dipstick urinalysis as a screening test for urinary tract infection. *J. Korean Acad. Fam. Med.* 21(6):772-781.
5. Heckerling, P. S., G. J. Canaris, S. D. Flach, T. G. Tape, R. S. Wigton and B. S. Gerber. 2007. Predictors of urinary tract infection based on artificial neural networks and genetic algorithms. *Int. J. of Medical Informatics* 76(4):289-296.
6. Hines, J. W. 1997. *MATLAB Supplement to Fuzzy and Neural Approaches in Engineering*. John Wiley and Sons, New York, USA.
7. Huang, D. S. 1999. Radial basis probabilistic neural networks: model and application. *Int. J. Pattern Recogn. Artif. Intell.* 13(7):1083-1101.
8. Jeon, G. R., G. R. Kim, S. Y. Ye, C. H. Kim, D. U. Jeong and J. H. Cho. 2003. A study on the design of classifier for urine analysis system. *J. of KOSOMBE* 24(3):193-201.
9. Kim, K. D., S. H. Koo, E. C. Kim, J. M. Kim, J. H. Kim, J. Q. Kim, H. J. Kim, D. S. Moon, W. K. Min, K. Y. Soo, Y. L., J. J. Lee, C. H. Jeon. M. E. Cho and S. S. Cho. 2006. Annual report on external quality assessment in clinical chemistry in Korea. *J. Lab. Med. Qual. Assur.* 28(1):63-89.
10. Lee, G. S. 2006. *Study on the Management System for Occupational Disease and Injury of Farmers*. Seoul National University Doctor Thesis.
11. Leibovici, L., A. Gershon, A. Laor, O. Kalter-Leibovici and Y. Danon. 1989. A clinical model for diagnosis of urinary tract infection in young women, *Arch. Intern. Med.* 149(9): 2048-2050.
12. Li, Y. C., L. Li, W. T. Chiu and W. S. Jian. 2000. Neural network modeling for surgical decisions on traumatic brain injury patients. *Int. J. Med. Inform.* 57(1):1-9.
13. Little, P., S. Turner, K. Rumsby, G. Warner, M. Moore, J. Lowes, H. Smith, C. Hawke and M. Mullee. 2006. Developing clinical rules to predict urinary tract infection in primary care settings: sensitivity and specificity of near patient tests (dipsticks) and clinical scores. *British Journal of General Practice* 56(529): 606-612.
14. Lisboa, P. J. and F. G. Taktak. 2006. The use of artificial neural networks in decision support in cancer: A systematic review. *Neural Networks* 19(4):408-415.
15. Oh, C. S. 1996. *Neuro Computer*. Jeeseung Press, Seoul, Korea.
16. Sultana, R. V., S. Zalstein. P. Cameron and D. Campbell. 2001. Dipstick urinalysis and the accuracy of the clinical diagnosis of urinary tract infection. *J. of Emergency Medicine* 20(1): 13-19.
17. Shin, S. S., S. B. Lee and Y. H. Cho. 2001. Recognition of disease in medical image. *J. of Contents Association* 1(1): 8-14.
18. Wigton, R. S., V. L. Hoellerich. J. P. Ornato. V. Leu, L. A. Mazzotta and I. H. C. Chen. 1985. Use of clinical findings in the diagnosis of urinary tract infection in women. *Arch. Intern. Med.* 145(12):2222-2227.
19. Wilson, J. D. 1991. *Harrison's Principles of Internal Medicine*. 12th edition, Vol. 2, McGRAW-HILL, International Edition, New York, USA.