# 자기조직화지도를 이용한 동적 분류기 선택[1]

이관희° 이일병*
연세대학교 컴퓨터 과학과
*Biometrics Engineering Research Center
{kwanhee°, yblee}@csai.yonsei.ac.kr

# Dynamic Classifier Selection Using Self-Organizing Maps

Kwanhee Lee° Yillbyung Lee*
Dept. of Computer Science, Yonsei University
*Biometrics Engineering Research Center

**요 약**

패턴 인식 분야에서 다중 분류기 시스템은 여러 분류기의 결과들을 조합하여 전체 성능을 향상 시키는 시스템이다. 다중 분류기를 사용함으로써 단일 분류기 보다 더 나은 결과를 얻을 수 있음은 이미 널리 알려진 사실이다. 서로 다른 구조를 갖는 분류기들은 상호 보완적인 정보를 제공하기 때문에 각 분류기마다 입력 공간에 대해서 지역적으로 좋은 성능을 갖는다. 본 논문에서는 지역적으로 가장 좋은 성능을 보이는 분류기 선택 방법을 제안한다. 제안하는 방법은 주어진 입력 공간에 대해 각 분류기들을 학습하는 과정에서 자기조직화지도를 생성하고 각 노드별로 평가함으로써 입력이 주어지면, 해당 노드에서 가장 성능이 좋은 분류기를 선택하여 전체 성능을 향상시키는 시스템이다.

## 1. 서 론

패턴 인식 분야에서, 다중 분류기 시스템은 서로 다른 분류기들의 결과를 조합하여 시스템의 성능을 향상시키는 방법이다[1]. 다중 분류기는 그룹적인 의견 종합이라는 측면에서 단일 분류기의 사용보다 일반적으로 좋은 성능을 보인다[2, 3]. 그 이유는 다수의 서로 다른 내부 구조와 학습 알고리즘을 갖는 분류기들은 상호 보완적인 정보를 제공하기 때문이다. 따라서 이들의 의견을 조합하는 것은 각 분류기들의 단점을 극복하여 전체 시스템의 성능을 향상시킨다[3].

본 논문에서 제안하는 시스템은 자기조직화지도를 이용하여 기반 분류기를 동적으로 선택하는 방법이다. 이때, 자기 조직화지도의 각 뉴런마다 분류기 가중치 망이라는 새로운 개념을 도입하여 분류기를 동적으로 선택하는 단계에서 이를 이용한다. 분류기 가중치 망의 학습은 평가 과정에서 이루어지며, 해당 뉴런에서 기반 분류기들을 평가하여 그에 따라 가중치가 조정된다. 따라서 분류기 가중치 망을 통해 입력 공간에 대한 지역적인 전문가를 쉽게 찾는 장점을 갖는다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 관련 연구에 대해서 살펴보고, 3장에서는 제안하는 방법에 대해 설명한다. 4장에서는 제안하는 시스템의 효용성 검증을 위해 실제 데이터에 대한 실험 및 비교 결과를 보인다. 마지막으로 5장에서는 결론 및 향후 연구 계획에 대해서 이야기한다.

## 2. 관련 연구

지금까지 연구된 다중 분류기 시스템은 구조적인 측면에서 순차 다중 분류기와 병렬 결합 다중 분류기로 나눌 수 있다[3].

순차 다중 분류기 방법은 각 분류기를 순차적으로 배열 한 후, 주어진 입력에 대해 전진 배치된 분류기가 분류에 성공하면 결과를 배출하고 시스템의 진행을 멈춘다. 이때, 분류에 실패하면 다음 분류기에게 입력을 넘기는 방식을 선택한다.

이와 비교하여 병렬 결합 다중 분류기 방법은 주어진 입력 공간에 대해 모든 분류기가 동시에 동작하여 결과를 배출하는 방식으로, 이것은 다시 분류기 혼합과 분류기 선택 방법으로 나눌 수 있다[3, 4].

분류기 혼합 방법은 크게 분류기 결합과 메타 학습으로 나뉜다. 분류기 결합 방법은 다수결 방법(Majority Voting), 베이시안 추론(Bayesian), 뎀스터-쉐퍼방법(Dempster-Shafer)등을 이용하여 각 분류기들의 결과를 결합하는 방법이다[5]. 메타 학습 방법은 각 분류기가 분류한 결과를 다른 분류기 통해 다시 학습하고 분류하는 과정을 거쳐 최종 결과를 배출하는 방법이다. 메타 학

[1] 이 연구는 부분적으로 과기부 지정 생체인식 연구센터의 지원을 받아 이루어 졌습니다.

습 방법으로는 Wolpert(1992)가 제안한 Stacked Generalization 이 있다[6].

분류기 선택 방법은 주어진 입력 공간에 대해 지역적인 전문가를 선택하는 방식으로, 이는 다시 동적 분류기 선택 방법과 정적 분류기 선택방법으로 나뉘게 된다[7]. 동적 분류기 선택 방법은 새로운 입력이 들어왔을 때 사전 확률과 사후 확률에 따라 높은 성능을 보이는 분류기를 선택하는 방식으로 Giacinto G.와 Roli F.이 제안하였다[8]. 군집화와 선택이라고도 불리는 정적 분류기 선택 방법은 학습 공간을 군집화하고 그 분할된 지역에 대해 성능이 가장 좋은 분류기를 선택하여, 새로운 입력이 들어왔을 때 해당 군집에서 선택된 분류기로 결과를 배출하는 방식이다[7, 9].

그 외에 전문가망[12], 배깅[13], 아킹[14] 등 다양한 연구가 진행되고 있다.

## 3. 자기조직화지도를 이용한 동적 분류기 선택 시스템

본 연구를 위해서는 모든 데이터 집합은 학습, 평가, 검증 데이터로 분할한다.

제안하는 시스템의 구조는 아래와 같다.

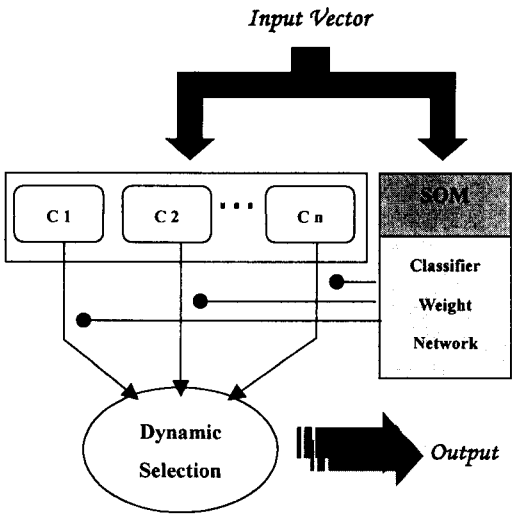


그림 1. 자기조직화지도를 이용한 동적 분류기 선택 시스템 구조

위의 그림에서 볼 수 있듯이 제안하는 시스템은 크게 두 부분으로 나뉘어 진다.

첫번째 부분은 여러 개의 독립적인 분류기들로 구성된 기반 분류기 모듈이다. 각 분류기들은 서로 다른 특성의 내부 구조와 학습 알고리즘을 갖는다.

두번째 부분은 경쟁 학습을 통해 생성된 자기 조직화 지도이다. 경쟁 학습은 주어진 입력 공간에 대해 승리한 출력 층의 뉴런만이 자신과 자신의 이웃 입력 층과의 연결 강도를 조정할 수 있도록 하는 승자 독점(Winer-Take-All)방법을 사용한다[10]. 자기 조직화 지도의 각 뉴런들은 각 분류기들에게 가중치 망을 제공하여 각 분류기들의 평가 시 출력 신호를 보내는 해당 뉴런에 대해 분류기 가중치를 학습한다. 실제 입력 패턴이 주어지면 자기 조직화 지도에서 출력 신호를 보내는 뉴런의 분류기 가중치가 가장 큰 분류기를 선택하여 최종 결과를 배출하게 된다.

### 3.1 학습

각 기반 분류기들은 서로 다른 구조나 학습 알고리즘을 갖도록 구성 한다.

학습은 두 단계에 걸쳐 일어난다.

첫번째는 학습 데이터를 이용해 기반 분류기들을 학습하고 자기 조직화 지도를 생성하는 과정이다.

두번째 학습은 평가 데이터에 대해 각 분류기들을 평가하고, 평가된 결과에 따라 자기 조직화 지도에서 출력 신호를 보내는 뉴런의 분류기 가중치 망을 학습한다. 따라서 각 뉴런의 분류기 가중치를 통해 해당 뉴런에 대한 각 분류기들의 성능을 알 수 있다.

### 3.2 실행

입력 패턴이 주어지면 자기 조직화 지도에서 출력 신호를 보내는 뉴런을 찾는다. 해당 뉴런에서 분류기 가중치가 가장 높은 분류기를 찾아 분류하여 최종 결과값으로 배출한다. 이때 분류기 가중치가 높은 분류기가 2개 이상이거나 출력 신호를 보내는 뉴런에서 분류기 가중치가 학습되지 않았다면, 다수결 방법을 통해 최종 결과값을 설정한다.

## 4. 실험 결과

본 논문에서 제안하는 시스템의 성능을 검증하기 위하여 UCI에서 제공하는 4개의 데이터 집합에 대해 실험하였다[11]. 실험에 사용된 데이터 집합의 속성은 아래와 같다.

표 1. 데이터 속성

| Data Set | #Variables | #Classes | #Train Set | #Test Set |
|---|---|---|---|---|
| Image Segment | 19 | 7(1~7) | 1848 | 462 |
| Waveform | 21 | 3(0~2) | 4000 | 1000 |
| Satellite Image | 36 | 6(1~5,7) | 4435 | 2000 |
| Shuttle | 9 | 7(1~7) | 43500 | 14500 |

위 [표 1]에서의 각 데이터 집합은 학습 데이터와 검증 데이터를 8:2의 비율로 나누었다. 제안하는 시스템에서는 [표 1]의 학

습 데이터를 다시 63:37%의 비율로 나누어 학습/평가 데이터로 사용하였다. 아래 [표 2]는 제안하는 시스템에서 사용된 데이터의 수를 나타낸다.

표 2. 제안하는 시스템에서의 데이터 비율

| Data Set | #Train Set | #Validation Set | #Test Set |
|---|---|---|---|
| Image Segment | 1164 | 684 | 462 |
| Waveform | 2520 | 1480 | 1000 |
| Satellite Image | 2794 | 1641 | 2000 |
| Shuttle | 27405 | 16095 | 14500 |

본 실험에서 독립적인 분류기 구성을 위해 결정트리(CHAID, CART, C4.5)와 오류 역전파 알고리즘(Backpropagation)을 사용하는 각기 다른 수의 은닉 노드(5개, 10개, 15개, 20개, 25개)를 가지는 신경망을 기반 분류기로 사용하였다.

위 [표 1]의 데이터 집합에 대한 독립적인 분류기들의 오분류율은 아래와 같다.

표 3. 결정 트리의 오분류율(%)

| Data Set | CHAID | CART | C4.5 |
|---|---|---|---|
| Image Segment | 7.36 | 7.57 | 4.76 |
| Waveform | x | 24.2 | 24.2 |
| Satellite Image | 18.8 | 16.75 | 15.5 |
| Shuttle | 7.400 | 0.400 | 0.360* |

표 4. 신경망의 오분류율(%)

| Data Set | NN1 | NN2 | NN3 | NN4 | NN5 |
|---|---|---|---|---|---|
| Image Segment | 4.33 | 3.46 | 2.38* | 2.81 | 2.81 |
| Waveform | 18.1 | 16.9* | 17.3 | 17.3 | 17.7 |
| Satellite Image | 14.25 | 11.85 | 11.7 | 10.9 | 10.25* |
| Shuttle | 0.538 | 0.407 | 0.400 | 0.407 | 0.414 |

위 [표 3]와 [표 4]에서 *는 해당 데이터 집합에서 가장 높은 분류 성능을 나타낸 분류기를 의미한다. 본 시스템의 기반 분류기들은 위의 실험에서 사용된 같은 알고리즘과 구조를 가진 8개(waveform의 경우 7개)의 분류기를 사용하여 실험하였다. 아래의 [표 5]는 가장 높은 성능의 독립 분류기와 다수결 결합 알고리즘, 그리고 제안하는 시스템의 실험 결과를 비교한 것이다.

표 5. 타 알고리즘과 제안하는 시스템의 오분류율(%) 비교

| Data Set | Best_Classifier | Majority Vote | 제안하는 시스템 |
|---|---|---|---|
| Image Segment | 2.38 | 2.38 | 1.95 |
| Waveform | 16.9 | 16.5 | 15.9 |
| Satellite Image | 10.25 | 11.4 | 9.9 |
| Shuttle | 0.360 | 0.317 | 0.124 |

위의 [표 5]에서 알 수 있듯이 제안하는 방법이 독립적인 분류기와 결합 알고리즘보다 향상된 결과를 나타내고 있다.

## 5. 결론

본 논문에서는 자기 조직화 지도를 이용하여 입력 공간을 분할하고 해당 지역별로 가장 높은 성능을 보이는 분류기를 선택함으로써 전체 성능을 향상시키는 시스템을 제안하였다. 자기 조직화 지도의 특성을 사용하여 입력 공간을 분할하고, 분할된 지역에서 동적 분류기 선택 방법의 장점을 이용하는 방법을 이용하였다. 제안하는 방법은 분류기 가중치 망이라는 모듈을 통해 각 분류기의 지역적 능력을 평가하고 저장함으로써 이를 실행 모드에서 활용하여 향상된 결과를 나타내었다.

본 연구의 향후 연구 계획은 다음과 같다. 첫째로 제안하는 방법론에 대해서 더 많은 분류기를 추가하여 실험한다. 둘째로, 다양한 데이터 집합에 대한 실험 및 성능 검증을 폭 넓게 수행한다.

## 6. 참고 문헌

[1]J. Kittler et al., On Combining Classifiers, IEEE Trans. Vol. 20, No. 3, pp. 226-239, 1998

[2]G. Valentini. and F. Masulli., Ensembles of learning machines, Lecture Notes in Computer Science, Vol.2486, Issue , pp. 3-19 , 2002

[3]Eunju Kim. Wooju Kim and Yillbyung Lee. Classifier Fusion using Local Confidence, Lecture Notes in Artificial Intelligence, Vol. 2366, pp. 583-591, 2002

[4]G. Giacinto and F. Roli, Adaptive Selection Of Image Classifiers, ICIAP ' 97, Lecture Notes in Computer Science, Vol. 1310, pp. 38-45, 1997

[5]Soyoung Sohn. Sungho Lee, Data Fusion, Ensemble and Clustering for the Severity Classification of Road Traffic Accident in Korea, Journal of the Korean Institute of Industrial Engineers, Vol. 26, No. 4, pp. 354-362, 2000

[6]Wolpert, D.H., Stacked generalization, Neural Networks 5, pp. 241-259, 1992

[7]Puuronen S. Terziyan V. Tsymbal A., A Dynamic Integration Algorithm for an Ensemble of Classifiers, Lecture Notes in Artificial Intelligence, Vol. 1609, pp. 592-600, 1999

[8] Giacinto G. Roli F., Dynamic Classifier Selection, Lecture Notes in Computer Science, Vol 1857, pp. 177-189, 2000

[9]Kuncheva L.I, Cluster-and-selection method for classifier combination, Proc. 4th International Conference on Knowledge-Based Intelligent Engineering Systems & Allied Technologies (KES'2000), 185-188, 2000

[10]Kohonen T, The self-organizing map, Proceedings of IEEE, Vol. 78, Issue. 9, pp. 1464-1480, 1990

[11]E. Bauer. E. Keogh. and C.J. Merz., UCI repository of Machine Learning database, http://www.ics.uci.edu/~mlearn/MLRepository.html, 1999

[12]R. A. Jacobs et al., Methods for Combining Experts' Probability Assessments, Neural Computation, Vol 7, pp. 867-888, 1995

[13]L. Breiman, Bagging Predictors, Machine Learning, Vol 24, pp. 123-140, 1996

[14] L. Breiman, Arcing classifiers, Annals of Statistics, Vol 26, pp. 801-824, 1998