

論文2002-39CI-3-4

결정경계 수직벡터의 해석적 계산을 통한 신경망 결정경계 특징추출 알고리즘의 성능 개선

(Improving the Performance of Decision Boundary Feature Extraction for Neural Networks by Calculating Normal Vector of Decision Boundary Analytically)

高續穢*, 李哲熙*

(Jinwook Go and Chulhee Lee)

요약

본 논문에서는 결정경계(decision boundary)를 이용한 신경망의 특징추출을 해석적으로 구현할 수 있는 방법을 제안한다. 최근 발표된 신경망의 결정경계 기반의 특징추출 방법은 기존의 특징추출 방법보다 우수한 성능을 보여 주었다. 이러한 결정경계 특징추출 방법은 패턴 분류기(pattern classifier)의 결정경계에 수직한 벡터가 패턴 클래스(class)간을 분류하는데 유용한 정보를 포함한다는 사실을 기반으로 원래의 데이터로부터 분류에 필요한 정보들만을 추출하게 된다. 그러나 기존의 결정경계 특징추출 알고리즘은 신경망 결정경계의 수직벡터를 구하기 위해 결정경계의 변화율(gradient) 근사 방법을 사용하였다. 그 결과 결정경계 수직 벡터가 부정확하게 계산될 가능성이 있고 계산 시간이 길어지는 문제점이 존재한다. 본 논문에서는 이러한 문제점을 해결하기 위해 수직벡터를 하나의 방정식으로부터 해석적으로 계산하는 방법을 제안한다. 제안된 방법을 원격탐사 데이터의 패턴분류에 적용하여 그 성능을 확인한 결과 특징추출에 필요한 연산 시간을 대폭 줄일 수 있고 또한 더 향상된 특징추출 성능을 얻음을 확인하였다.

Abstract

In this paper, we present an analytical method for decision boundary feature extraction for neural networks. It has been shown that all the features necessary to achieve the same classification accuracy xxxas in the original space can be obtained from the vectors normal to decision boundaries. However, the vector normal to the decision boundary of a neural network has been calculated numerically using a gradient approximation. This process is time-consuming and the normal vector may be inaccurately estimated. In this paper, we propose a method to improve the performance of the previous decision boundary feature extraction for neural networks by accurately calculating the normal vector. When the normal vectors are computed analytically, it is possible to reduce the processing time significantly and improve the performance of the previous implementation that employs numerical approximation.

I. 서론

신경망은 패턴 인식^[1] 동적 시스템 제어, 통신 시스템 등과 같은 많은 비선형 신호처리 문제에 성공적으로 적용되었고, 그 중 입력층, 다수의 은닉층(hidden layer), 출력층으로 구성되는 다층 퍼셉트론(perceptron)

* 正會員, 延世大學校 電氣·電子工學科

(Dept. Electrical and Electronic Eng., Yonsei Univ.)

※ 본 연구는 한국과학재단 연구비 지원으로 이루어졌습니다.

接受日字:2001年3月6日, 수정완료일:2002年4月15日

신경망이 널리 사용되고 있다. 그러나 다층 신경망을 적용하여 해결하고자 하는 문제가 복잡해질수록 다층 신경망을 구성하는 전체 연결가중치(connection weight)의 수는 급격히 증가하게 된다. 즉 다층 신경망 연결가중치의 개수는 입력패턴 벡터의 차원이 증가함에 따라 은닉층 뉴런(neuron)의 개수만큼 증가되므로 패턴분류(pattern classification) 시 시스템의 연산시간과 복잡도를 증가시키게 된다. 따라서 원래의 데이터로부터 분류에 필요한 정보들만을 추출하는 특징추출(feature extraction) 과정이 요구된다. 효율적인 특징추출 알고리즘은 패턴인식 시스템에서 패턴 분류에 필요한 정보의 손실을 최소화하는 동시에 가능한 적은 수의 특징들을 추출할 수 있어야 한다.

특징추출은 패턴분류나 패턴인식 문제에 있어서 매우 중요하게 다루어진 주제로서 오랫동안 많은 연구자들에 의해 폭넓게 연구되어 왔다.^[4~6] 그러나 현재 신경망에 적용할 수 있는 특징추출 알고리즘은 다소 부족한 형편이다. 기존의 대표적인 특징추출 알고리즘인 PCA(principal component analysis) 방법과 CA(canonical analysis) 방법을 신경망의 특징추출에 적용할 수는 있지만,^[2,3] 이러한 방법들은 복잡한 결정경계(decision boundary)를 형성하는 신경망의 특징을 이용하지 않으므로 신경망의 특징추출에는 만족할 만한 성능을 제공하기 어렵다.

최근 발표된 신경망 결정경계 특징추출(decision boundary feature extraction) 알고리즘에서는 원(original) 데이터로부터 얻을 수 있는 분류 정확도와 동일한 분류 정확도를 얻기 위한 특징벡터(feature vector)는 신경망이 정의하는 결정경계에 대한 수직벡터(normal vector)로부터 구할 수 있다는 이론을 제시하였다 [6]. 특히, 신경망 결정경계 특징추출 알고리즘은 결정경계로부터 직접 특징벡터를 추출함으로써 확률밀도함수(probability density function)의 가정 없이 임의의 결정경계를 형성하는 신경망의 중요한 특징을 충분히 이용할 수 있다. 이러한 결정경계 특징추출 알고리즘을 224 채널(channel)의 고차원 데이터인 AVIRIS(Airborne Visible-Infrared Imaging Spectrometer) 데이터에 적용한 결과 기존의 특징추출 알고리즘보다 우수한 성능을 보여주었다.^[8]

그러나 기존의 결정경계 특징추출 알고리즘은 신경망의 결정경계로부터 수직벡터를 구하기 위해 결정경계의 변화율 근사(gradient approximation) 방법을 사용

하였다. 그 결과 결정경계의 수직벡터를 부정확하게 구하는 경우가 발생할 수 있고, 특징추출에 소요되는 연산 시간이 길어지는 단점이 존재하였다. 본 논문에서는 이러한 결정경계의 수직벡터를 해석적으로 구하는 방법을 제안하고 계산에 필요한 식을 유도한다. 이러한 해석적 결정경계 특징추출(analytical decision boundary feature extraction) 방법을 통해 기존의 신경망 결정경계 특징추출 알고리즘의 성능을 향상시킬 수 있음을 확인하였다.

본 논문의 구성은 제II단락에서 결정경계 특징추출 알고리즘을 소개하고, 이 알고리즘을 신경망에 확장 적용시킨 신경망 특징추출 알고리즘을 설명한다. 제III단락에서는 본 논문에서 제안하는 신경망 결정경계의 수직벡터를 해석적으로 구하는 방법을 구체적으로 기술한다. 제IV단락에서는 실험결과 및 기존 알고리즘들과의 패턴 분류시의 특징추출 성능을 비교 평가하고 마지막으로 제V단락에서 결론을 맺는다.

II. 신경망 결정경계 특징추출 알고리즘

결정경계 특징추출 알고리즘은 패턴 클래스간의 결정경계로부터 패턴 분류에 유용한 정보를 제공하는 특징(feature)을 추출하는 방법이다.^[7] 이러한 결정경계 특징추출 알고리즘은 순방향 다층 신경망의 특징추출에 성공적으로 적용되었다.^[6] 결정경계 특징추출 알고리즘을 신경망에 적용하기 위해서는 먼저 신경망의 결정경계가 정의되어야 한다. 패턴 분류를 위한 클래스 결정 규칙(decision rule)은 일반적으로 많이 사용되는 가장 큰 출력 값을 갖는 출력층 뉴런에 대응되는 클래스를 선택하는 것으로 가정한다.

1. 결정경계 특징추출 알고리즘

결정경계 특징추출 알고리즘은 패턴 클래스의 분류에 유용한 정보를 제공하는 특징(discriminantly informative feature)과 불필요한 특징(discriminantly redundant feature)은 분류기가 정의하는 결정경계와 관계됨을 보여주었다.^[7] 특히, 분류에 유용한 특징벡터들은 결정경계 상의 적어도 한 점에서 결정경계에 대한 수직성분을 가지고 있고, 분류에 불필요한 특징벡터들은 결정경계 상의 모든 수직 성분과 직교함을 증명하였다. 분류에 유용한 특징벡터를 구하기 위하여 다음과 같은 결정경계 특징행렬(decision boundary feature matrix)

이 정의된다.^[7]

정의 1. 결정경계 특징행렬: 결정경계 상의 한 점 X 에서 결정경계에 수직한 단위벡터(unit vector)를 $N(X)$ 라 할 때, 결정경계 특징행렬은 다음과 같이 정의된다.

$$\Sigma_{DBFM} = \frac{1}{K} \int_S N(X)N^t(X)p(X)dX$$

이 때, $p(X)$ 는 X 의 확률밀도함수, $K = \int_S p(X)dX$, S 는 결정경계, 그리고 적분은 전체 결정경계에 대해 수행된다.

결정경계 특징행렬은 결정경계에 수직인 성분들로 이루어져 있으므로 결정경계 특징행렬의 고유벡터(eigenvector)를 구함으로 패턴 클래스를 분류할 수 있는 특징벡터를 얻을 수 있다. 그리고 이와 같이 구한 고유벡터에 대한 고유치(eigenvalue)의 크기가 클수록 그 고유벡터는 클래스를 더 잘 분류할 수 있고, 고유치의 크기가 0인 고유벡터는 클래스를 분류하는데 불필요한 벡터이다. 즉, 임의의 패턴 분류 문제에서 결정경계 특징행렬 Σ_{DBFM} 의 rank는 원래 데이터로부터 얻을 수 있는 분류정확도를 획득할 수 있는 최소의 특징 수를 나타낸다. 이와 같은 관찰에서 다음 두 가지 이론이 유도되었다.

이론 1. 임의의 한 패턴분류 문제에서 결정경계 특징행렬 Σ_{DBFM} 의 rank는 원래의 데이터 공간에서 얻을 수 있는 분류 정확도를 얻기 위한 최소 차원이다.

이론 2. 임의의 한 패턴분류 문제에서 결정경계 특징행렬 Σ_{DBFM} 의 0이 아닌 고유치에 대한 고유벡터들은 원래의 데이터 공간에서 얻을 수 있는 분류 정확도를 얻기 위한 특징벡터들이다.

2. 순방향 다층 신경망의 결정경계

2클래스간에 정의되는 신경망의 결정경계는 입력공간(input space)상에서 두 클래스에 상응하는 각 출력층 뉴런의 크기가 같은 점들의 궤적이다. 즉, 두 패턴 클래스의 결정경계는 다음과 같이 정의된다.

$$OUT_1(X) = OUT_2(X)$$

이 때 $OUT_i(X)$ 는 i 번째 출력층 뉴런의 출력 값을 나타내고, X 는 신경망의 입력 벡터이다. 신경망 출력층 뉴런의 출력 값들의 차를 $h(X)$ 라 하면 결정경계는 다음과 같이 나타낼 수 있다.

$$\{X | h(X) = OUT_1(X) - OUT_2(X) = 0\} \quad (1)$$

그림 1은 2클래스 패턴분류를 위한 두 개의 출력을 갖는 2층 신경망의 예를 보여준다. 신경망의 출력인 $OUT_1(X)$ 과 $OUT_2(X)$ 는 다음과 같이 구할 수 있다.

$$OUT_1(X) = F(W_h^1 \cdot F(W_i \cdot X))$$

$$OUT_2(X) = F(W_h^2 \cdot F(W_i \cdot X))$$

이 때, W_i 는 입력층과 은닉층간의 연결강도 행렬이고, W_h^1 과 W_h^2 는 은닉층과 출력층간의 연결강도 벡터이며 F 는 신경망 활성화함수를 나타낸다. 그러므로 식 (1)의 신경망 결정경계는 다음과 같이 표현될 수 있다.

$$\{X | F(W_h^1 \cdot F(W_i \cdot X)) - F(W_h^2 \cdot F(W_i \cdot X)) = 0\}$$

신경망의 결정경계가 결정되면 입력벡터 X 에서 결정경계 수직벡터는 다음 식으로 계산된다.

$$\nabla h(X) = -\frac{\partial h}{\partial x_1} \vec{x}_1 + \frac{\partial h}{\partial x_2} \vec{x}_2 + \frac{\partial h}{\partial x_3} \vec{x}_3 + \dots + \frac{\partial h}{\partial x_N} \vec{x}_N \quad (2)$$

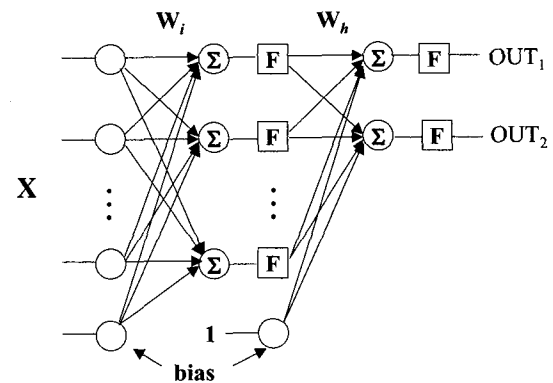


그림 1. 2클래스 패턴분류를 위한 순방향 2층 신경망의 예

Fig. 1. An example of a two-layer feedforward neural networks (2 pattern classes).

이 때, 신경망의 입력층의 뉴런 개수는 N 개이다. 최근 발표된 신경망 결정경계 특징추출 알고리즘에서는

다음 식과 같은 변화율(gradient) 근사 방법을 사용하여 식 (2)의 $\nabla h(X)$ 를 구하였다.

$$\nabla h(X) \approx \frac{\Delta h}{\Delta x_1} \vec{x}_1 + \frac{\Delta h}{\Delta x_2} \vec{x}_2 + \frac{\Delta h}{\Delta x_3} \vec{x}_3 + \dots + \frac{\Delta h}{\Delta x_N} \vec{x}_N$$

3. 신경망 결정경계 특징추출 알고리즘

2클래스 분류 문제의 경우 신경망 결정경계 특징추출 알고리즘은 다음과 같은 과정을 통해 구한다. 또한 이 과정은 다중 클래스 문제들에도 쉽게 적용할 수 있다.^[6]

- 단계 1: 모든 특징을 사용하여 신경망을 학습시킨다.
- 단계 2: 클래스 ω_1 의 맞게 분류된 각각의 학습 샘플들에 대해, 클래스 ω_2 로 맞게 분류된 학습 샘플들에서 거리가 가장 가까운 샘플을 찾는다. 같은 과정을 클래스 ω_2 로 맞게 분류된 학습 샘플들에 대해 반복한다.
- 단계 3: 단계 2에서 찾은 한 쌍의 샘플들을 연결하는 선(line)은 신경망의 결정경계를 통과하게 된다. 그러므로 이 선을 따라 이동함으로써 결정경계상의 한 점이나 결정경계에 인접한 점을 찾을 수 있다.
- 단계 4: 단계 3에서 찾은 각각의 점들에 대해 결정경계 수직벡터 N_i 를 다음과 같이 구한다.

$$N_i = \frac{\nabla h(X)}{|\nabla h(X)|}$$

단계 5: 단계 4에서 계산한 결정경계 수직벡터로부터 결정경계 특징행렬(decision boundary feature matrix) Σ_{DBFM} 을 추정한다.

$$\Sigma_{DBFM} = \frac{1}{L} \sum_{i=1}^L N_i N_i^t$$

이 때, L 은 올바르게 분류된 샘플 개수이다.

단계 6: 결정경계 특징행렬(decision boundary feature matrix)로부터 구한 크기가 0이 아닌 고유치(eigenvalue)에 대한 고유벡터(eigenvector)를 새로운 특징벡터(feature vector)로 선택한다.

2개 이상의 클래스가 존재하는 다중 클래스 분류 문제에 대해서는 전체 클래스에 대해 학습을 수행한 후

위의 과정을 각 쌍의 클래스들에 대해 반복함으로써 구할 수 있다. 즉, N 클래스 분류 문제에 대해 $N C_2 = \frac{N(N-1)}{2}$ 개의 결정경계 특징행렬들을 얻을 수 있다. 전체 결정경계 특징행렬은 각 쌍의 클래스들의 결정경계 특징행렬들을 평균함으로써 다음과 같이 계산된다.

$$\Sigma_{DBFM} = \sum_{i=1}^N \sum_{j=i+1}^N P(\omega_i)P(\omega_j)\Sigma_{DBFM}^{ij}$$

이 때, Σ_{DBFM}^{ij} 는 클래스 ω_i 와 클래스 ω_j 간의 결정경계 특징행렬이고, $P(\omega_i)$ 는 클래스 ω_i 의 사전 확률(prior probability)이다. 각 클래스의 사전 확률을 알 수 없는 경우에는 $P(\omega_i)$ 을 $\frac{1}{N}$ 로 설정한다.

III. 해석적 결정경계 특징추출 알고리즘

본 논문에서는 오류 역전파(error backpropagation) 알고리즘으로 학습된 순방향 다층 신경망에 대한 특징추출을 고려한다. 오류 역전파 학습 알고리즘은 출력층의 각 뉴런(neuron)의 자승 오차의 합인 오차함수(error function)를 정의하여 이 오차함수를 감소시키도록 연결강도를 변경시키는 학습 방법으로 활성화함수는 크기 범위가 제한되고 미분 가능하여야 한다.^[3] 이러한 조건을 만족하는 함수들 중에서 시그모이드(sigmoid) 함수가 널리 사용되고 있고, 시그모이드 함수는 그림 2와 같이 S자 형태의 단조증가(monotonically increasing) 함수로서 0과 1사이의 값을 갖는 다음 함수로 표현된다.

$$F(x) = \frac{1}{1 + e^{-x}}$$

이 때, 단조증가 함수의 특징에 따라 $F(x_1) \geq F(x_2)$ 와 $x_1 \geq x_2$ 는 등가이다.

본 논문에서는 식 (2)의 신경망 결정경계의 수직벡터를 해석적으로 계산하기 위해 시그모이드 함수의 단조증가 특성을 이용한다. 즉, 신경망의 학습이 종료된 경우에는 출력층 뉴런의 시그모이드 함수의 유무가 패턴 분류 정확도에 영향을 미치지 않으므로, 신경망 결정경계의 정의로부터 출력층 뉴런의 시그모이드 함수를 제거할 수 있다. 이러한 특성을 이용하면 수직벡터의 유도 과정을 용이하게 할 수 있다.

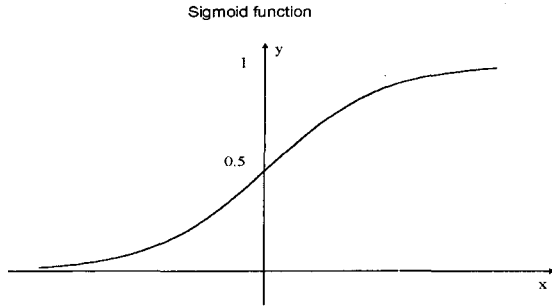


그림 2. 시그모이드 함수
Fig. 2. Sigmoid function.

시그모이드 함수의 단조증가 특성을 통해 결정경계는 다음 식과 같은 등가 결정경계로 표현될 수 있다.

$$\{X \mid W_h^1 \cdot F(W_i \cdot X) - W_h^2 \cdot F(W_i \cdot X) = 0\} \quad (3)$$

은닉층 뉴런의 개수가 M 개라고 가정할 때, 식 (3)은 다음과 같이 나타낼 수 있다.

$$\frac{w_{1,1}^h - w_{2,1}^h}{1 + e^{-\phi_1^h X}} + \frac{w_{1,2}^h - w_{2,2}^h}{1 + e^{-\phi_2^h X}} + \dots + \frac{w_{1,M}^h - w_{2,M}^h}{1 + e^{-\phi_M^h X}} = 0$$

이 때, ϕ_j 는 입력층과 j 번째 은닉층 뉴런간의 $N \times 1$ 연결강도 열벡터(column vector)이고, $w_{k,j}^h$ 는 j 번째 은닉층 뉴런과 k 번째 출력층 뉴런간의 연결강도의 크기이다. 그리고 j 번째 은닉층 뉴런의 전체(net) 입력은 다음 식과 같이 주어진다.

$$z_j' = \phi_j^T X = \phi_{j,1}x_1 + \phi_{j,2}x_2 + \dots + \phi_{j,N}x_N \quad (j = 1, \dots, M)$$

이 때, $\phi_{j,i}$ 는 i 번째 입력층 뉴런과 j 번째 은닉층 뉴런간의 연결강도의 크기이다. 또한 j 번째 은닉층 뉴런의 출력은 다음 식과 같다.

$$z_j = \frac{1}{1 + e^{-z_j'}} = \frac{1}{1 + e^{-\phi_j^T X}} \quad (4)$$

식 (4)를 i 번째 입력 x_i 에 대해 편미분하면 다음과 같다.

$$\begin{aligned} \frac{\partial z_j}{\partial x_i} &= \frac{\partial z_j}{\partial z_j'} \cdot \frac{\partial z_j'}{\partial x_i} = \frac{\partial z_j}{\partial z_j'} \cdot \phi_{j,i} x_i = \frac{e^{-z_j'}}{(1 + e^{-z_j'})^2} \phi_{j,i} x_i \\ &= \frac{1}{1 + e^{-z_j'}} \left(1 - \frac{1}{1 + e^{-z_j'}} \right) \phi_{j,i} x_i \\ &= z_j(1 - z_j) \phi_{j,i} \end{aligned}$$

최종적으로 i 번째 입력 x_i 에 대한 $h(X)$ 의 편미분은

다음 식으로 표현된다.

$$\frac{\partial h}{\partial x_i} = \sum_{j=1}^M z_j(1 - z_j) \phi_{j,i} (w_{1,j}^h - w_{2,j}^h)$$

IV. 실험결과 및 고찰

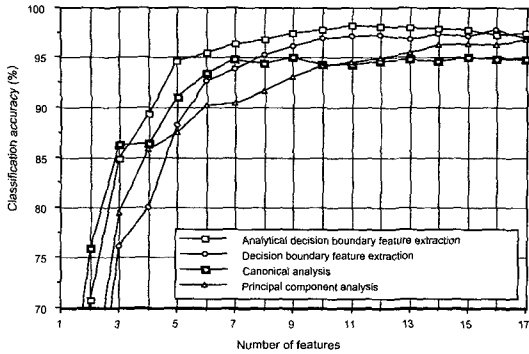
본 논문에서 제안한 특징추출 알고리즘의 성능을 평가하기 위하여 실제 원격탐사(remote sensing) 데이터를 사용하여 실험을 수행하였다. 사용된 데이터는 농업 지역에서 field spectrometer system (FSS)으로부터 채집된 다중스펙트럴(multispectral) 데이터로서 60개의 스펙트럴 채널(channel)로 구성되었다.^[9] 제안된 알고리즘과 성능비교를 위해, canonical analysis 방법, principal component analysis 방법, 변화율(gradient) 근사 방법을 사용한 기존의 결정경계 특징추출 알고리즘을 사용하였다. 모든 실험에서 신경망의 은닉층 뉴런의 개수가 입력벡터 차원의 3배를 갖는 2층 신경망을 사용하였고, 신경망은 0.1의 고정된 학습률(learning rate)을 갖는 오류 역전파 알고리즘을 사용하여 학습되었다.^[10]

표 1. 8개의 패턴 클래스 정보

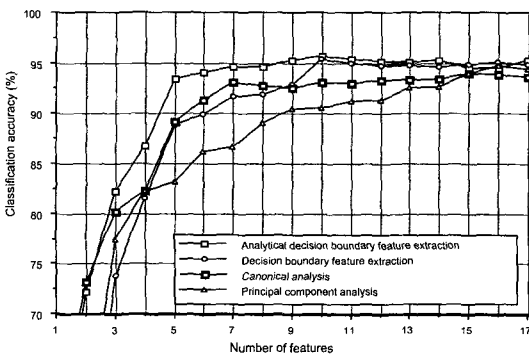
Table 1. Information of the 8 pattern classes.

No. class	Date	Location	Species	No. data
1	770308	Finney Co. KS.	Winter Wheat	691
2	770626	Finney Co. KS.	Winter Wheat	677
3	771018	Hand Co. SD.	Winter Wheat	660
4	780515	Hand Co. SD.	Spring Wheat	474
5	780816	Hand Co. SD.	Spring Wheat	464
6	760928	Finney Co. KS.	Summer Fallow	411
7	760928	Finney Co. KS.	Grain Sorghum	277
8	780602	Hand Co. SD.	Oats	259

첫 번째 실험으로 FSS 데이터에서 8개의 클래스를 선택하고 클래스 데이터의 인접 채널들을 합하여 특징 수를 17개로 축소하였다. 표 1은 8개 패턴 클래스의 정보를 보여준다. 각각의 클래스로부터 무작위로 추출된 150개의 샘플을 학습 데이터로 사용하였고 나머지 데이터를 시험 데이터로 사용하였다. 특징추출 과정 없이 17개의 특징(feature)을 모두 사용하여 10000번 동안 신경망을 학습시켰을 때, 학습 데이터와 시험 데이터의



(a)



(b)

그림 3. 특징추출 성능 비교 (8개 패턴 분류 문제). (a) 학습 데이터 (b) 시험 데이터
Fig. 3. Performance comparison (8 pattern classes). (a) training data, (b) test data.

분류 정확도는 각각 96.9%와 94.8%이다. 그림 3은 8개 패턴 클래스 분류 문제에 대한 4가지 특징추출 알고리즘의 성능 비교를 보여준다. Canonical analysis 알고리즘은 적은 개수의 특징을 사용하였을 때 학습 데이터와 시험 데이터 모두에 대해 가장 우수한 성능을 보여주지만, 2~3개의 특징을 사용하여 얻는 분류 정확도는 17개의 특징을 사용하였을 때의 분류 정확도와 비교하여 매우 큰 차이를 보여 주고 있다. 제안된 해석적 결정경계 특징추출 알고리즘은 4개 이상의 특징이 사용된 경우 학습 데이터와 시험 데이터 모두에 대해서 다른 3가지 특징추출 알고리즘보다 우수한 성능을 보여 주고 있다. 특히, 시험 데이터의 경우 해석적 결정경계 특징추출 알고리즘은 5개의 특징을 사용하여 원래의 데이터로부터 얻을 수 있는 분류 정확도인 94.8%와 유사한 93.4%의 분류 정확도를 제공하고 있다. 이에 반해 canonical analysis 방법, principal component analysis 방법, 변화율(gradient) 근사 방법을 사용한 기존의 결

정경계 특징추출 알고리즘은 각각 89.2%, 83.1%, 88.8%의 분류 정확도를 나타낸다. 이는 제안된 특징추출 알고리즘을 통해 원래 17차원 공간에서 획득할 수 있는 분류 정확도를 5~6 차원의 데이터로부터 얻을 수 있음을 의미한다. 또한, 그림 3은 정확한 해석적 계산에 의해 구한 결정경계 수직벡터를 통해 더욱 우수한 특징을 찾을 수 있음을 보여준다.

표 2. 10개의 패턴 클래스 정보

Table 2. Information of the 10 pattern classes.

No. class	Date	Location	Species	No. data
1	770308	Finney Co. KS.	Winter Wheat	691
2	770626	Finney Co. KS.	Winter Wheat	677
3	771018	Hand Co. SD.	Winter Wheat	660
4	770503	Hand Co. SD.	Winter Wheat	657
5	770626	Finney Co. KS.	Summer Fallow	643
6	780726	Hand Co. SD.	Spring Wheat	515
7	780921	Hand Co. SD.	Spring Wheat	469
8	781026	Hand Co. SD.	Spring Wheat	441
9	760928	Finney Co. KS.	Summer Fallow	411
10	780515	Hand Co. SD.	Winter Wheat	223

다음 실험에서는 FSS 데이터에서 10개, 20개의 클래스를 무작위로 선택하고 클래스 데이터의 인접 채널들을 합하여 특징 수를 17개로 조정하였다. 표 2와 3은 각각 10개와 20개의 패턴 클래스의 정보를 보여준다. 각각의 클래스로부터 무작위로 추출된 150개의 샘플을 학습 데이터로 사용하였고 나머지 데이터를 시험 데이터로 사용하였다. 그림 4와 5는 각각 10개와 20개 패턴 클래스 분류 문제에 대한 4가지 특징추출 알고리즘의 성능 비교를 보여준다. 그림 4에서 볼 수 있듯이 10개의 패턴 분류 문제에 대해 canonical analysis 알고리즘은 특징의 개수가 증가하더라도 더 이상의 성능 향상을 나타내지 않고 있다. 또한 그림 5의 20개의 클래스 분류 문제에 대해서는 특징의 개수 증가와 함께 분류 정확도는 향상되지만 그 향상이 미비하여 다른 특징추출 알고리즘들에 비해 낮은 분류 정확도를 보여주고 있다. 반면 제안된 해석적 결정경계 특징추출 알고리즘은 canonical analysis 방법과 principal component analysis 방법과 비교하여 매우 우수한 성능을 보여 주

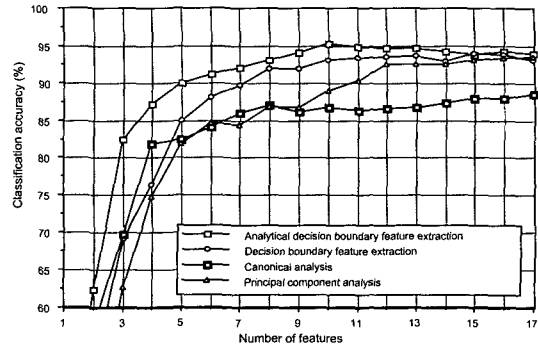
고 있으며, 근사 방법을 사용한 기존의 결정경계 특징추출 알고리즘보다도 현저한 성능 향상을 보여주고 있다. 예로 그림 5의 20개의 패턴분류문제에서, 해석적 결정경계 특징추출 알고리즘은 특징 6개에서 76%, 9개에서 79%의 정확도(시험 데이터)를 보여주고 있으나 기존의 결정경계 특징추출 알고리즘은 같은 수의 특징에서 71%, 76%의 정확도를 보여 주고 있다.

표 3. 20개의 패턴 클래스 정보
Table 3. Information of the 20 pattern classes.

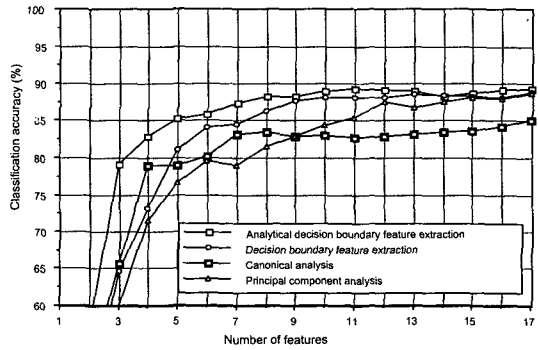
No. class	Date	Location	Species	No. data
1	770308	Finney Co. KS.	Winter Wheat	691
2	770626	Finney Co. KS.	Winter Wheat	677
3	771018	Hand Co. SD.	Winter Wheat	660
4	770503	Hand Co. SD.	Winter Wheat	657
5	770626	Finney Co. KS.	Summer Fallow	643
6	780726	Hand Co. SD.	Spring Wheat	515
7	780602	Hand Co. SD.	Spring Wheat	515
8	780515	Hand Co. SD.	Spring Wheat	474
9	780921	Hand Co. SD.	Spring Wheat	469
10	780816	Hand Co. SD.	Spring Wheat	464
11	780709	Hand Co. SD.	Spring Wheat	454
12	781026	Hand Co. SD.	Spring Wheat	441
13	760928	Finney Co. KS.	Summer Fallow	411
14	770920	Hand Co. SD.	Winter Wheat	292
15	780921	Hand Co. SD.	Winter Wheat	292
16	770308	Finney Co. KS.	Grain Sorghum	279
17	780602	Hand Co. SD.	Oats	259
18	780515	Hand Co. SD.	Winter Wheat	223
19	781026	Hand Co. SD.	Pasture	217
20	780816	Hand Co. SD.	Summer Fallow	216

표 4. 특징추출 연산시간 비교
Table 4. Comparison of processing times.

특징추출 방법 데이터	결정경계 수직벡터의 해석적 연산	결정경계 수직벡터의 변화율 근사 방법	시간 단축 비율 (변화율 근사/해석적)
17차원 8 클래스 (150 샘플/클래스)	7.96 초	31.48초	3.95
17차원 10 클래스 (150 샘플/클래스)	12.87초	51.37초	3.99



(a)

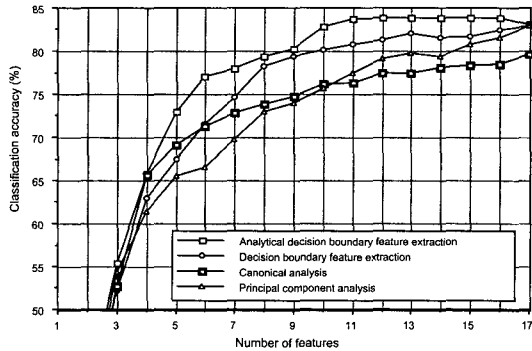


(b)

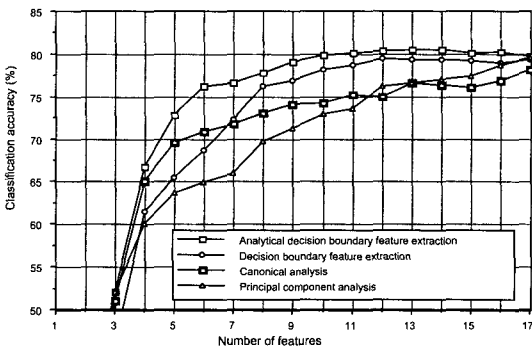
그림 4. 특징추출 성능 비교 (10개 패턴 분류 문제).
(a) 학습 데이터 (b) 시험 데이터
Fig. 4. Performance comparison (10 pattern classes).
(a) training data, (b) test data.

다음에서는 해석적 결정경계 특징추출 알고리즘의 특징추출을 위한 연산시간을 변화율(gradient) 근사 방법을 사용한 결정경계 특징추출 알고리즘과 비교한다. 표 4는 표 1-2의 8개 클래스와 10클래스에 대해 각 클래스당 150개의 학습 데이터를 사용한 경우의 특징추출 연산시간을 보여준다. 연산 환경은 리눅스(Linux) 커널(kernel) 버전 2.0.32를 운영체제로 탑재한

Pentium-III 1GHz processor 컴퓨터에서 실험하였고, 프로그래밍 언어는 GNU C++를 사용하였다. 표 4로부터 결정경계 수직벡터의 해석적 계산을 통해 특징추출에 필요한 시간을 약 4배정도 단축할 수 있음을 볼 수 있다.



(a)



(b)

그림 5. 특징추출 성능 비교 (20개 패턴 분류 문제).
(a) 학습 데이터 (b) 시험 데이터
Fig. 5. Performance comparison (20 pattern classes).
(a) training data, (b) test data.

IV. 결론

본 논문에서는 최근에 제안된 신경망 결정경계 특징추출 알고리즘에서 결정경계 수직벡터의 해석적 계산을 통해 특징추출 성능과 연산시간을 단축하는 방법을 제안하였다. 결정경계 특징추출 알고리즘은 신경망의 결정경계를 정의하고 이 결정경계의 수직벡터를 통해 새로운 특징벡터들을 계산한다. 하지만 기존 알고리즘에서는 신경망의 결정경계 수직벡터를 변화율 근사 방법을 통해 구하였고, 이러한 방법은 결정경계에 대한 정확한 수직벡터를 구할 수 없고 불필요한 많은 연산

을 필요로 한다. 이러한 단점을 보완하기 위해 본 논문에서는 신경망에 대한 결정경계 수직벡터를 유도하였다. 또한 시그모이드 함수의 단조증가 특성으로부터 수직벡터를 구하는데 필요한 연산량을 단축하였다. 해석적 결정경계 특징추출 알고리즘을 원격탐사 데이터의 패턴분류 문제에 적용한 실험 결과 기존의 결정경계 특징추출 알고리즘의 부정확한 수직벡터의 추정 문제를 해결할 수 있고 또한 연산시간을 4배 단축할 수 있음을 확인하였다.

참고 문헌

- [1] F. Fukushima and M. Wake, "Handwritten alphanumeric character recognition by the neocognitron," *IEEE Trans. Neural Networks*, Vol. 2, pp. 355~365, Mar, 1991.
- [2] J. A. Richards, *Remote Sensing Digital Image Analysis*. Springer-Verlag, 1993.
- [3] S. Haykin, *Neural Networks*, New York: Macmillan, 1994.
- [4] D. H. Foley and J. W. Sammon, "An Optimal Set of Discriminant Vectors," *IEEE Trans. Computer*, Vol. C-24, No. 3, pp. 281~289, 1975.
- [5] E. Choi and C. Lee, "Optimizing Feature Extraction for Multiclass Problems," *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 39, No. 3, Mar, 2001.
- [6] C. Lee and D. A. Landgrebe, "Decision boundary feature extraction for neural networks," *IEEE Trans. Neural Networks*, Vol. 8, No. 1, January 1997.
- [7] C. Lee and D. A. Landgrebe, "Feature extraction based on decision boundaries," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 4, pp. 388~400, 1993.
- [8] J. A. Benediktsson, J. R. Sveinsson, and K. Arnason, "Classification and feature extraction of AVIRIS data," *IEEE Trans. Geoscience and Remote Sensing*, Vol. 33, No. 5, Sept. 1995.
- [9] L. L. Biel and et. al., "A Crops and Soils Data Base For Scene Radiation Research," *Proc. Machine Process. of Remotely Sensed Data*

Symp., West Lafayette, Indiana, 1982.

[10] R. P. Lippmann, "An Introduction to Computing with Neural Nets," *IEEE Acoust. Signal Processing Mag.* Vol. 4, pp. 4~22, Feb. 1987.

저 자 소 개



高纘穢(正會員)
1996년 2월: 건국대학교 전자공학과 졸업 (공학사). 1997년 3월~1999년 2월: 연세대학교 전자공학과 졸업 (공학석사). 1999년 9월~현재: 연세대학교 전기·전자공학과 박사과정.
<주관심분야: 신경회로망, 패턴인식,

음성신호처리>



李 晳 熙(正會員)
1980년 3월~1984년 2월: 서울대학교 전자공학과 졸업 (공학사). 1984년 3월~1986년 2월: 서울대학교 대학원 전자공학과 졸업 (공학석사). 1986년 9월~1987년 3월: Technical University of Denmark (Researcher).

1987년 8월~1992년 12월: Purdue University Electrical Engineering (Ph. D.). 1993년 7월~1996년 8월: National Institutes of Health, Maryland, USA. (Visiting Fellow). 1996년 9월~1999년 8월: 연세대학교 기계전자공학부 조교수. 1999년 9월~현재: 연세대학교 기계전자공학부 부교수. <주관심분야: 영상신호처리, 음성신호처리, 패턴인식, 신경회로망>