

# 준감독 학습 알고리즘을 위한 능동적 레이블 데이터 선택

## Active Selection of Label Data for Semi-Supervised Learning Algorithm

한 지 호\*, 박은애\*, 박동철\*\*, 이윤식\*\*, 민수영\*\*

Ji-Ho Han\*, Eun-Ae Park\*, Dong-Chul Park\*\*, Yunsik Lee\*\*, Soo-Young Min\*\*

### Abstract

The choice of labeled data in semi-supervised learning algorithm can result in effects on the performance of the resultant classifier. In order to select labeled data required for the training of a semi-supervised learning algorithm, VCNN(Vector Centroid Neural Network) is proposed in this paper. The proposed selection method of label data is evaluated on UCI dataset and caltech dataset. Experiments and results show that the proposed selection method outperforms conventional methods in terms of classification accuracy and minimum error rate.

### 요약

본 논문에서는 준감독 학습 알고리즘(Semi-Supervised Learning Algorithm)의 학습데이터에 필요한 소수의 레이블 데이터를 능동적으로 선택하기 위한 무감독경쟁학습 알고리즘인 VCNN(Vector Centroid Neural Network)을 제안한다. 준감독 학습 알고리즘에서 레이블 데이터의 선택은 학습 결과 큰 영향을 미치고, 레이블 데이터를 선택하는데 있어 많은 비용과 전문적인 지식이 필요하다. 본 논문에서 능동적이고 효율적인 레이블 데이터 선택을 검증하기 위하여 UCI database 와 caltech dataset 을 이용하여 실험한 결과, 기존의 레이블 데이터 선택 방법과 비교하여 안정된 분류 결과와 최소의 오차율을 나타냈다.

*Key words : semi-supervised learning, label data, centroid neural network, data selection, classification*

## 1. 서론

지능형 학습 알고리즘의 종류는 학습 데이터에서 레이블 데이터의 유무에 따라 크게 두 가지 학습방식

으로 구분된다. 감독 학습 알고리즘은 레이블이 존재하는 학습데이터를 사용하고, 무감독 학습 알고리즘은 레이블이 존재하지 않은 학습데이터를 사용한다. 감독 학습 알고리즘은 연산량이 많고 수식이 복잡하여 낮은 학습 속도와 구현의 어려움이 문제로 제기되지만 무감독 학습 알고리즘 보다 우수한 성능의 이유로 학습데이터의 레이블을 획득하기 위한 높은 비용과 많은 시간을 소요하면서도 여러 분야에서 사용되고 있다. 반면, 무감독 학습 알고리즘은 연산량이 적고 수식이 간단하여 학습 속도가 빠르며 구현이 간단하지만 감독 학습 알고리즘보다 성능에서 차이를 보이기 때문에 이를 극복하기 위한 연구가 진행되고 있다.

최근에는 감독 학습 알고리즘과 무감독 학습 알고

\*Dept. of Electronics Eng., Myong Ji University

\*\*System IC R&D Division, KETI

★ Corresponding author, [parkd@mju.ac.kr](mailto:parkd@mju.ac.kr), 031-330-6756

※This work was supported by the Korean Government (2010-0009655) and by the 2012 IT R&D program of The MKE/KEIT (10040191, The development of Automotive Synchronous Ethernet combined IVN/OVN and Safety control system for 1Gbps class).

Manuscript received Jul. 18, 2013; revised Aug. 14, 2013; accepted Aug. 21, 2013

리즘의 단점을 극복하기 위해 중간 형태를 갖고 있는 준 감독 학습 알고리즘(Semi-Supervised Learning Algorithm)이 주목받고 있다. 준 감독 학습 알고리즘은 소수의 레이블이 있는 학습데이터와 다수의 레이블이 없는 학습데이터를 이용하여 학습하며 기존의 감독 학습 알고리즘의 높은 성능과 무감독 학습 알고리즘의 빠른 학습속도의 특징을 보여준다. 지금까지 준 감독 학습 알고리즘은 감독 학습 알고리즘에서 소수의 레이블 데이터만 사용하고 레이블이 없는 데이터 처리를 기존의 함수에 추가하거나 무감독 학습 알고리즘에 레이블 데이터 처리를 추가하는 방향으로 연구가 진행되어왔다. 레이블 데이터가 감소한 만큼 레이블 처리에 소요되는 시간과 비용은 감소하지만 실제 데이터양은 과거에 비해서 매우 증가되었기 때문에 소수의 레이블 데이터 선택의 신뢰성이 매우 중요하다 ([1]-[5]).

이러한 문제를 해결하기 위하여, 본 논문에서는 준 감독 학습 알고리즘을 위한 레이블 데이터 선택에 있어서 무감독경쟁 학습 알고리즘인 VCNN(Vector Centroid Neural Network)을 제안한다. VCNN은 기존의 CNN(Centroid Neural Network) 알고리즘 연산에서 추가적인 대푯값이 입력될 때, 벡터 이론이 추가 되어 패자뉴런의 발생을 최소화하고 반복적인 학습에도 기존의 CNN 보다 안정된 학습결과를 얻기 위하여 제안된다. 이를 입증하기 위해 준 감독 학습 알고리즘으로 무감독 학습 알고리즘을 기초로 하여 제안된 SSSVaD (Semi-Supervised Spatially Variant Dissimilarity Measure) 와 감독 학습 알고리즘 SVM을 준감독 학습 방식으로 실험을 정리했다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 사용하는 CNN 알고리즘과 제안된 VCNN 알고리즘 소개와 레이블 데이터 선택방법을 소개하고, 3장에서는 실험에 사용된 영상에서의 특징 추출방법을 소개하고, 4장에서는 기존의 레이블 선택 방법과 본 논문에서 제안된 방법과 데이터 분류 실험 결과를 비교하고, 5장에서는 결론을 기술했다.

## II. 본론

### 1. CNN(Centroid Neural Network)

CNN 알고리즘은 k-means와 SOM 같은 기존 알고리즘이 매 학습마다 연결 가중치를 수정하는 것과 달리, 승자와 패자의 연결 가중치를 각각 설정하여 현재와 이전 학습에서의 뉴런의 상태에 대한 정보를 통해 가중치를 변화시킨다. CNN의 목적함수는 다음의 식(1)와 같이 표현된다.

$$\min(W)\{J(W; X) = \sum_{i=1}^M \sum_{j=1}^{N_i} \|X_i(j) - W_i\|^2\} \quad (1)$$

분류되지 않은 데이터 집합을  $X = (x_1, x_2, \dots, x_N)$ , 뉴런  $i$ 의 연결 가중치는  $w_i = (w_{i1}, w_{i2}, \dots, w_{iN})$ 로 표현했을 때 승자뉴런  $j$ 와 패자뉴런  $i$ 의 연결 가중치 갱신은 다음과 같다.

$$w_j(n+1) = w_j(n) + \frac{1}{n_j+1} [x(n) - w_j(n)] \quad (2)$$

$$w_j(n+1) = w_j(n) - \frac{1}{n_j-1} [x(n) - w_j(n)] \quad (3)$$

여기서  $w_j(n)$ 과  $w_i(n)$ 은 승자 뉴런과 패자 뉴런의 연결 가중치를 표현하며,  $N_j$ 와  $N_i$ 는 현재 학습에서 군집  $j, i$ 에 속한 데이터의 수를 의미한다.

CNN알고리즘은 기존의 무감독 경쟁학습과 비교하여 여러 장점을 가지고 있다. 초기 설정 값에 있어서 학습반복수를 설정할 필요가 없고, 초기 중심 값에 대한 문제점도 해결할 수 있다. 학습반복수는 패자뉴런이 없을 때까지 반복되며, 학습속도 또한 기존의 알고리즘보다 월등히 빠름을 확인할 수 있다([6]).

### 2. 벡터를 이용한 CNN (VCNN)

CNN 알고리즘의 학습 단계에서 패자뉴런이 발생되지 않는 경우 새로운 군집 분류를 위한 대푯값이 추가된다. 기존에는 식(4)과 같이 그룹별로 데이터와 대푯값의 모든 거리의 합이 가장 큰 그룹의 대푯값 중심에 임의의 값을 설정하여 새로운 대푯값을 추가했다. 식(4)에서  $c$ 는 대푯값의 총 개수를 의미하고,  $N_j$ 는  $i$ 번째 대푯값을 승자뉴런으로 선택한 데이터의 총 개수를 의미한다. 이전 대푯값의 이동된 경로 내에 추가 되는 경우 다량의 패자뉴런이 발생하여 패자뉴런 갱신 연산의 증가와 학습 결과의 오차가 발생했다. 하지만 대푯값이 이동되는 방향을 고려하여 이동된 방향의 영역에 추가하는 경우 패자뉴런의 발생이 최소화 되어 연산량이 줄고 최소의 오차로 학습을 마무리 할 수 있다.

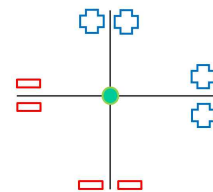


Fig. 1. Determination of the sign of each dimension of additional cluster in 2 dimension data.

그림 1. 2차원 데이터의 추가 대푯값의 차원별 부호 결정

$$E = \sum_{i=1}^c \sum_{j=1}^{N_j} \|x_{i(j)} - w_i\|^2 \quad (4)$$

$$\vec{w}_{add(m)} = \Delta \vec{w}_{(m)} + \vec{w}_{final(m)} \quad (5)$$

식 (5)에서  $m$ 은 차원을 의미하며,  $\vec{w}_{final}$ 는 오차가 가장 큰 그룹의 대푯값을 의미한다.  $\Delta w$ 의 차원별 부호에 따라 추가되는 대푯값의 위치를 결정할 수 있다. CNN과 같은 방법으로 초기 대푯값을 설정하고 그 위치를 기억한다. 추가 대푯값이 요구될 때 현재의 대푯값들의 이동된 위치와 초기 대푯값의 위치를 식(5)를 이용하여 해당 대푯값의 변화량을 구할 수 있다. 그림1에서 2차원 데이터에 대한 변화량에 따른 위치 지정을 예로 나타냈다. 고차원 데이터에서는 차원별 변화량의 부호에 따라 추가 되는 대푯값의 차원별 부호를 결정 한다. 최종적으로 이동된 방향을 찾았을 때 벡터를 구한 후 이동된 거리의 일정 비율로 대푯값을 추가하여 대푯값이 군집내에 추가 되도록 설정한다.

### 3. 레이블 데이터의 선택 방법

준감독 학습 알고리즘의 데이터 전체를 VCNN을 이용하여 학습을 수행하면, 설정한 개수만큼 대푯값을 얻을 수 있다. 대푯값은 입력된 데이터의 분할된 군집 중앙에 위치하기 때문에 대푯값 주변에는 해당 군집의 데이터들이 분포된다. 학습에 사용된 데이터 중에서 준감독 학습 알고리즘의 레이블 데이터 선택은 유클리디언 거리방식(Euclidean distance)으로 대푯값과 데이터간의 거리를 구할 수 있다.

식(6)에서  $w$ 는 대푯값을 의미하고,  $x$ 는 데이터,  $c$ 는 대푯값의 개수,  $j$ 는  $i$ 번째 해당하는 대푯값을 winner로 선택한 데이터의 개수이다.

$$distance = \sum_{j=1}^{N_j} \|w_i - x_j\|^2, \quad i = 1, \dots, c \quad (6)$$

식(6)를 이용하여 구해진 거리 중에서 가까운 거리를 기준으로 설정된 레이블 데이터 개수만큼 군집별로 선택하여 해당 대푯값의 레이블을 전달하여 선택한다. 위와 같은 방법으로 선택데이터는 데이터 군집의 중앙이나 군집간의 사이 경계에 존재하는 데이터들을 레이블 데이터로 선택하게 되어 준감독 학습 알고리즘의 오분류를 최소화 할 수 있다.

## III. 특징추출 방법

### 1. ULBP(Uniform Local Binary Pattern)

기본적인 국부이진패턴(Local Binary Pattern) 히스토그램은 1996년 Ojala에 의해 흑백영상에 대한 텍스처 기술자로 제안됐다([7][8]). LBP는 국부영역에서의 질감 형태에 따라 세밀한 특징을 표현할 수 있고 조명변화에 강건하기 때문에 얼굴 인식등의 분야에서 많이 사용된다.

$$LBP(i_c) = \sum_{n=0}^7 s(i_n - i_c) \cdot 2^n \quad (7)$$

$$s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

식 (7)에서  $i_c$ 는 LBP를 구하는 중심 픽셀이고,  $i_n$ 은  $i_c$ 와 이웃하는 픽셀들을 나타낸다.

기본적인 LBP 는 28 = 256 종류의 패턴을 가지고 이를 카운트 하여 256차원의 텍스처 히스토그램으로 표현할 수 있다.

### 2. HSV

HSV 색공간은 색을 표현하는 하나의 방법이자, 그 방법에 따라 배치하는 방식으로 영상의 RGB 색공간을 색상(Hue), 채도(Saturation), 명도(Value or Brightness)의 좌표로 써서 색을 표현하는 방법이다 ([9]).

본 논문에서는 HSV 색공간의 각 성분을 히스토그램으로 만든 후 차례로 이어 하나의 히스토그램으로 작성했다. 색상은 전체 360°를 모두 사용하지 않고 2° 범위를 하나의 색으로 가정하여 총 180차원의 히스토그램으로 표현하고, 채도와 명도는 각각 백분율로 표현하나 10% 마다 하나의 단계로 묶어 각 10차원의 히스토그램을 생성한다. 즉, 색상(H), 채도(S), 명도(V)를 차례로 이어 총 200차원의 HSV 색상 히스토그램을 생성했다.

### 3. DCT

DCT(Discrete Cosine Transform)는 JPEG, H.261, MPEG 등 영상의 압축 및 복원에 널리 사용되어 좋은 성능을 보여주며, DCT 계수(Coefficient)는 영상 주파수 특성을 나타내는 특징벡터로 사용되어 영상의 인식, 분류에도 좋은 결과를 나타낸다 ([10]).

본 논문에서는 입력영상의 가로와 세로가 8×8 픽셀이 중복되지 않는 블록으로 나누어 블록 DCT 연산을 한 후 좌측 상단에 위치하는 DC component로부터 지그재그 스캐닝을 통해 1차원 행렬로 정렬했다.

전체 64차원의 변환계수 행렬로부터 저주파성분 32차원을 저장하고 이렇게 저장된 각 블록의 32차원 계수들의 평균과 분산을 이어 하나의 영상을 표현하는 64차원의 특징벡터를 생성했다.

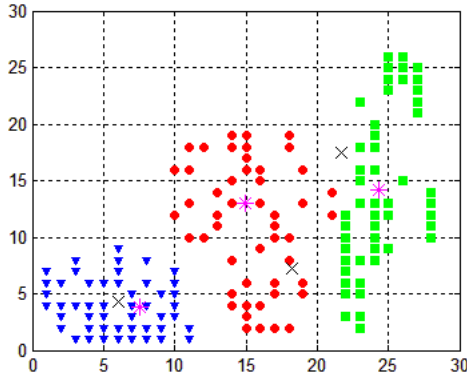


Fig. 2. VCNN result in 2D data point  
그림 2. 2차원 데이터에서의 VCNN 결과

#### IV. 실험 및 결과

##### 1. 2차원 데이터에 대한 실험

VCNN을 이용한 레이블 데이터 선택 방법에 대한 성능을 입증하기 위해 두 가지 실험을 비교했다. 첫째로 기존의 방법으로 학습 데이터 중에서 그룹별로 임의의 데이터를 찾아 해당 그룹의 레이블 데이터를 지정하여 SSSVaD 알고리즘과 준감독 학습 방식의 SVM의 입력 데이터로 사용했고, 두 번째 방법으로는 VCNN을 이용하여 군집의 대푯값을 찾아 대푯값을 승자누턴으로 선택한 데이터 중근집한 데이터를 기준으로 해당 그룹의 레이블 데이터를 선택하여 위와 같이 두 가지 알고리즘의 입력 데이터로 사용했다.

레이블 데이터 선택의 안정성을 확인하기 위해 그림 2에 표현된 2차원 데이터를 이용해서 VCNN의 레이블 데이터 선택 성능을 확인했다. 정확한 레이블 데이터 선택을 위해서 VCNN의 학습 결과는 각 군집의 중앙에 대푯값이 위치하여 대푯값 주변 데이터를 선택할 때 다른 군집의 데이터 선택을 방지해야 한다. \*형태의 대푯값은 VCNN을 이용하여 군집에서의 최적의 중심 값이며 × 형태의 대푯값은 기존의 CNN 학습 결과에서 오차가 발생한 중심 값이다. × 형태가 대푯값으로 선택되는 경우 이웃하는 데이터를 선택하는데 있어서 다른 그룹의 데이터가 선택되는 경우가 발생되기 때문에 준감독 학습 알고리즘에서 분류 정확도를 낮추는 결과를 나타낸다. 뿐만 아니라 학습 결과의 신뢰성이 낮으면 반복되는 실험에서의 오차는 증가한다. 그림 2의 데이터에서 임의의 레이블 데

이터를 선택하여 준감독 학습 방식의 Support Vector Machine ([11]) 과 SSSVaD 의 입력 데이터로 사용했을 때 분류 결과는 73.00% ~ 76.66% 이며, VCNN을 이용하여 레이블 데이터를 선택한 경우 준감독 학습 알고리즘의 분류 결과는 90.66% ~ 92.33%로 나

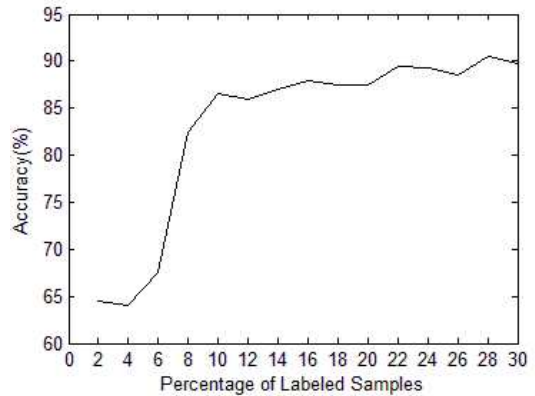


Fig. 3. Classification accuracy vs. number of label data samples

그림 3. 레이블 데이터 개수에 따른 분류정확도

타났다. 가장 높은 분류 결과에도 큰 차이가 나타나지만 VCNN을 이용하여 레이블 데이터를 선택한 경우 기존의 방식보다 오차율이 1.99% 감소했다.

##### 2. UCI database & Caltech database 실험

4.1에서 소개한 2차원 데이터를 이용한 실험에서 입력 데이터의 카테고리 별로 임의의 레이블 데이터를 선택하는 방법과 무 감독 학습 알고리즘을 이용한 레이블 데이터 선택 방법에 있어 특히, VCNN 알고리즘을 이용하는 경우 준감독 학습 알고리즘의 안정적인 분류 결과를 확인 할 수 있었다. 2차원 데이터의 경우 레이블 선택의 가능성을 확인하기 위해 인위적으로 제작된 데이터이기 때문에 객관성 있는 실험을 위해 UCI database([12])와 Caltech database([13])를 이용하여 추가 실험을 수행했다. UCI database 중에서 iris, wine, vehicle, sonar, ionosphere data와 Caltech database에서는 airplane, cars, faces, leopards, motobikes를 이용하여 논문에서 제안한 방법을 확인했다. UCI database 경우 특별한 전처리 없이 실험에 사용이 가능하지만 Caltech 이미지 데이터 경우에는 특징 추출을 위해 HSV, ULBP, DCT의 특징 추출 방법을 이용했다. 추출된 특징은 HSV 200차원, ULBP 59차원, DCT 64차원으로 구성되며, 각 특징별 차원을 하나의 벡터로 연결시켜 총 323차원의 입력데이터를 구성했다.

실험 방법으로 데이터 마다 10%의 데이터를 검증

Table 1. Classification results of SSSVaD

표 1. SSSVaD의 분류 실험 결과

data	SSSVaD	VCNN+ SSSVaD
iris	76.67%±10%	88.66%±6%
wine	79.15%±5%	91.33%±2%
vehicle	57.83%±3.7%	62.82%±1.2%
sonar	58.77%±3.5%	68.66%±5%
ionoshere	77.06%±2.7%	79.77%±0.5%
<b>average</b>	<b>69.90%±5.0%</b>	<b>78.33%±3.6%</b>
Caltech	66.00%±5.0%	68.00%±3.0%

데이터로 남겨놓고, 나머지 90%의 데이터를 이용하여 학습을 수행했다. 이와 같은 방법으로 검증 데이터와 학습 데이터를 다르게 하여 10번의 실험을 수행했고, 실험 결과의 최소, 최대 분류 결과와 모든 데이터에 대한 결과의 평균으로 실험 결과를 정리했다. 레이블 데이터는 전체 데이터의 20%를 선택했고, 임의의 레이블 데이터 선택을 위해 특징을 추출하기 전에 이미지에서 임의의 레이블을 지정하여 기존의 방법으로 레이블 데이터를 선택했다. VCNN을 이용한 레이블 데이터 선택방법은 323차원의 모든 데이터에 대해서 해당 database 마다 그룹의 개수를 지정하여 학습을 했고, 2.3절에서 소개했듯이 각 군집의 중심에 위치한 대푯값을 승자뉴턴으로 선택한 데이터 중에서 가까운 순위로 20%를 선택했다. 그림 3에 레이블 데이터 개수에 따른 분류정확도를 iris 데이터의 실험 결과로 표현했다. SSSVaD의 경우 위의 소개한 실험 방법으로 실험이 진행되었으며, SVM의 경우 SVM\_light를 이용하여 실험을 하였으며 레이블 데이터를 제외한 데이터는 목적값을 0으로 처리하여 레이블이 없이 업데이터를 진행했고, 이진분류기인 SVM을 멀티클래스로 진행하기 위해 one-Vs-rest 방식을 사용했다.

표 1에서는 SSSVaD 알고리즘에 대해 UCI database, Caltech database의 실험 결과를 나타내며 표 2에서는 SVM 알고리즘을 이용한 실험 결과를 보여준다. 2차원 데이터의 실험과 같이 VCNN을 이용한 레이블 데이터 선택은 준감독 학습 알고리즘의 분류정확도에서 향상된 결과를 확인할 수 있다. 기존의 레이블 선택 방법에서는 최대 10%의 오차율이 발생하지만 VCNN을 이용한 레이블 데이터 선택 방법에서는 최대 6%의 오차율이 발생되었고, 최소 0.5%의 오차율을 통해 안정된 결과를 확인할 수 있다.

## V. 결론

본 논문의 실험을 통해서 준감독 학습 알고리즘의 입력 데이터 중 레이블 데이터가 전체적인 분류도 향

Table 2. Classification accuracy of Semi SVM

표 2. Semi SVM의 분류 정확도

data	SVM	VCNN+ SVM
iris	74.44%±8%	90.66%±6%
wine	84.44%±4%	91.33%±2.5%
vehicle	61.00%±4.2%	62.82%±2.0%
sonar	63.77%±5%	70.66%±5%
ionoshere	80.00%±2%	82.00%±2%
<b>average</b>	<b>72.73%±4.6%</b>	<b>79.49%±3.5%</b>
Caltech	71.00%±5.0%	71.00%±3.0%

상에 있어 중요성을 확인했다. 준감독 학습 알고리즘은 감독 학습 알고리즘과 다르게 소수의 레이블 데이터를 사용하기 때문에 레이블 데이터 선택에서의 오차는 분류 결과에 큰 오차를 발생시키고, 정확한 레이블 데이터 선택은 안정된 분류 결과를 나타냈다. 이에 대해 본 논문에서 제안한 VCNN을 이용한 레이블 데이터 선택은 안정화된 무감독 학습 알고리즘의 분류 결과를 확인 했고, 기존의 방법보다 레이블 데이터 선택의 소요 시간이 빠르고 높은 분류도를 보였다. 실험 결과 무 감독 학습 알고리즘을 이용하여 준감독 학습 알고리즘의 레이블 데이터 선택의 가능성을 확인할 수 있었고, 향후 다양한 방법을 응용하여 무감독 학습 방식을 준 감독 학습으로 변형시킬 수 있는 것을 기대할 수 있다.

## References

- [1] H. Zeng, Y. Cheung, "Semi-Supervised Maximum Margin Clustering with Pairwise Constraints", *IEEE Tr. on Knowledge and Data Engineering*, vol.24, no.5, pp.926-939, 2012
- [2] M. Mathia et. al, "Semisupervised Least Squares Support Vector Machine", *IEEE Tr. on Neural Network*, vol.20, no.12, pp.1858-1870, 2009
- [3] P. Mallapragada et. al, "SemiBoost: Boosting for Semi-Supervised Learning", *IEEE Tr. on PAMI*, vol.31, no.11, pp.2000-2014, 2009
- [4] N. Kumar, K. Kumamuru, "Semi supervised Clustering with Metric Learning Using Relative Comparisons", *IEEE Tr. on Knowledge and Data Engineering*, vol.20, no.4, pp.496-503, 2008
- [5] K. Chen, S. Wang, "Semi-Supervised Learning via Regularized Boosting Working on Multiple Semi-Supervised Assumptions", *IEEE Tr. on PAMI*, vol.33, no.1, pp.129-143, 2011

- [6] Dong-Chul Park, "Centroid Neural Network for Unsupervised Competitive Learning", *IEEE Trans on Neural Network*, vol.11, no.2, pp.520-528, 2000
- [7] T. Ojala et. al, "Performance Evaluation of Texture Measures with Classification Based on Kullback Discrimination of Distributions", *Proc. of ICPR*, vol.1, pp.582-585, 1994
- [8] T. Ahonen et. al, "Face Description with Local Binary Patterns: Application to Face Recognition", *IEEE Tr. on PAMI*, vol.28, no.12, pp.2037-2041, 2006
- [9] C. Novak, S. Shafer, "Anatomy of a Color Histogram", *IEEE Conf on Computer Society*, pp.599-605, 1992
- [10] N. Anmed et. al, "Discrete Cosine Transform", *IEEE Tr. on Computer*, vol.23, no.1, pp.90-93, 1974
- [11] SVM-light, <http://svmlight.joachims.org>
- [12] Caltech-101 dataset, <http://www.vision.caltech.edu/>
- [13] UCI dataset, <http://archive.ics.uci.edu/ml/datasets.html>

---

## BIOGRAPHY

---

### Ji-ho Han (Student Member)



2011, B.S. Dept. of Inf. Eng.  
Myong Ji. University  
2013, M.S. Dept. of Electronics  
Eng. Myong Ji. University  
2013 - Present, Ph.D. program,  
Dept. of Electronics Eng. Myong  
Ji. University

<Areas of Interest> Intelligent computing, Image processing

### Eun-Ae Park (Student Member)



2012, B.S. Dept. of Inf. Eng.  
Myong Ji. University  
2012 - Present, M.S. program,  
Dept. of Electronics Eng. Myong  
Ji. University

<Areas of Interest> Intelligent computing, Image processing

### Dong-Chul Park (Member)



1990, Ph.D. Dept. of Electrical  
Eng. Univ. of Washington, USA  
1990 - 1994, Asst. Prof. Dept. of  
Elect. and Comp. Eng, FIU,  
Miami, FL, USA

1994 - Present, Prof. Dept. of  
Electronics Eng. MyongJi. Univ.

<Areas of Interest> Intelligent computing, Pattern Recognition

### Yunsik Lee (Member)



1993, Ph.D. Dept. of Computer  
Eng., Univ. of South Florida.

2005 - Present : Senior Research  
Staff, System IC R&D Division,  
KETI

<Areas of Interest>

Semiconductor design automation,  
FPGA, Signal processing

### Soo-Young Min (Member)



1987, B.S. Dept. of Electronics  
Eng., Inha University

1993- Present : Senior Research  
Staff, System IC R&D Division,  
KETI

<Areas of Interest> Mobile

communication network, Wireless  
communication protocol, System software