

PCA 기반 특징 되먹임을 이용한 중요 영역 추출

이승현*, 김도연*, 최상일**, 정구민***

Extraction of Important Areas Using Feature Feedback
Based on PCA

Seung-Hyeon Lee*, Do-Yun Kim*, Sang-Il Choi**, Gu-Min Jeong***

요약 본 논문에서는 손글씨 숫자 데이터셋, 얼굴 데이터셋의 중요영역 추출을 위한 PCA 기반의 특징되먹임방법을 제안한다. 이전의 LDA 기반의 특징되먹임 방법을 확장하여 PCA 기반 특징되먹임 방법이 제안된다. 제안된 방법에서 데이터에 차원 축소 머신러닝 알고리즘 중 하나인 PCA 기법을 적용하여 데이터를 중요한 특징 차원들로 축소한다. 차원 축소과정에서 도출되는 weight를 통해 축소된 각 차원 축에서의 데이터 중요 지점을 확인한다. 각 차원 축은 축의 고유값의 크기에 따라 전체 데이터에서의 가중치가 다르다. 이에 각 차원 축의 고유값의 크기에 비례하는 가중치를 부여하여 각 차원 축에서의 데이터 중요 지점을 합하는 연산 과정을 거친다. 연산 과정을 통해 얻어진 데이터에 Threshold를 적용하여 데이터의 중요 영역을 구한다. 그 후 도출된 데이터의 중요 영역에 원본데이터로 역매핑을 유도하여 원본 데이터 공간에서 중요영역을 선택한다. MNIST 데이터셋에 대한 실험 결과를 확인하고 기존의 LDA 기반의 특징되먹임 방법을 통한 결과와 비교를 하여 PCA기반 특징되먹임을 기반한 패턴 인식 방법의 유효성과 가능성을 확인한다.

Abstract In this paper, we propose a PCA-based feature feedback method for extracting important areas of handwritten numeric data sets and face data sets. A PCA-based feature feedback method is proposed by extending the previous LDA-based feature feedback method. In the proposed method, the data is reduced to important feature dimensions by applying the PCA technique, one of the dimension reduction machine learning algorithms. Through the weights derived during the dimensional reduction process, the important points of data in each reduced dimensional axis are identified. Each dimension axis has a different weight in the total data according to the size of the eigenvalue of the axis. Accordingly, a weight proportional to the size of the eigenvalues of each dimension axis is given, and an operation process is performed to add important points of data in each dimension axis. The critical area of the data is calculated by applying a threshold to the data obtained through the calculation process. After that, induces reverse mapping to the original data in the important area of the derived data, and selects the important area in the original data space. The results of the experiment on the MNIST dataset are checked, and the effectiveness and possibility of the pattern recognition method based on PCA-based feature feedback are verified by comparing the results with the existing LDA-based feature feedback method

Key Words : Eigenface, Eigenmask, Feature Feedback, MNIST, PCA

This paper is a research conducted by the government (Ministry of Trade, Industry and Energy) in 2020 with support of the core technology development project of the robotics industry (No. 10080615). and This research was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MSIP)(NRF-2018R1D1A1A09083894).

*Electronic Engineering, Kookmin University

**Computer engineering, Dankook University

***Corresponding Author : Electronic Engineering, Kookmin University (gm1004@kookmin.ac.kr)

Received October 10, 2020

Revised October 26, 2020

Accepted November 03, 2020

1. 서론

패턴 인식 기술과 인공지능 기술의 발전으로 영상, 신호 등 다양한 데이터의 인식률이 높아지고 있다. 최근 관련 연구에서는 인식률을 높이는 측면과 함께 중요 영역을 추출하는 방법[4][5][3], 계산량을 절감하는 방법 등도 중요해지고 있다.

특징 되먹임 기반 패턴 인식 방법 [4]은 주어진 데이터 중에서 중요한 영역을 추출하기 위해서 제안되었다. 패턴인식으로 추출된 특징을 데이터 영역으로 역사상하여 중요한 영역을 추출할 수 있도록 하였다. [5]에서는 PCA+LDA 기반으로 얼굴 영상에 대해서 중요영역을 추출하였으며, [6]에서는 전자코 데이터에 대해서 클래스 내 데이터를 정규화한 LDA방식인 RLDA(Regularized Linear Discriminant Analysis)[3]를 기반으로 중요영역을 추출하였다. 특징 되먹임 기반 패턴인식 방법에서는 중요 영역 추출을 통한 데이터양의 절감과 실제 임베디드 시스템 적용을 위한 계산량 절감 효과를 얻을 수 있다. 최근 딥러닝에서도 뉴럴네트워크를 출력단에서 입력단으로 역전파 및 기여도를 재분배하여 데이터 학습 기여도를 판단하는 LRP(Layer-wise Relevance Propagation)[8], ReLU(Rectified Linear Unit)를 적용하여 기존 CAM(Class Activation Mapping)을 일반화하여 중요 부분을 추출하는 Grad-CAM(Gradient-weighted Class Activation Mapping)[9] 등 다양한 설명가능한 AI를 통해서 인식에 영향을 미치는 중요한 영역을 추출하기 위한 연구가 계속되고 있다. 특징되먹임 기반 패턴인식 방법은 LRP, Grad-CAM 등 딥러닝 기반 방법에 비해서, 방법이 간단하고, 계산량이 적은 장점을 가진다.

이 논문에서는 기존의 특징 되먹임 기반 패턴 인식 방법을 확장하여 PCA를 이용한 중요 영역 추출 방법을 제안한다. 일반적으로 LDA 기반 방법이 PCA에 비해서 분류율이 높지만, 일부 데이터에 대해서는 PCA가 더 효율이 좋은 경우가 있다. 제안 방법에서는 PCA로 얻어진 특징을 되먹

임하여 중요영역을 추출하고 데이터량과 계산량을 절감할 수 있도록 한다. 실험에서는 손글씨 데이터 베이스인 MNIST 데이터셋[10]을 통하여 제안 방법의 효용성을 검증한다.

이 논문의 구성은 다음과 같다. 2장에서는 기존의 LDA 기반 특징 되먹임 알고리즘을 정리한다. 3장에서는 PCA 기반 특징 되먹임 알고리즘을 제안한다. 4장에서는 실험 결과를 정리하고, 5장에서는 결론을 맺는다.

2. 관련 연구

데이터에는 중요한 영역과 중요하지 않은 노이즈 영역이 존재한다. 이에 데이터에서 중요한 영역만을 추출하여 보다 적은 데이터 특징만으로 원본 데이터와 비슷한 성능을 얻는 특징되먹임을 이용하여 원본데이터에서 노이즈가 제거된 데이터를 추출할 수 있다. [4][5]에서는 데이터의 중요영역을 추출하기 위한 PCA+LDA기반 특징 되먹임 알고리즘을 제안하였다.

2.1 PCA(Principal Component Analysis)

PCA는 패턴인식[4], 신호처리 등에서 대표적으로 사용되는 차원축소 알고리즘으로 데이터의 분산을 고려하여 차원을 축소한다. 이로써 데이터가 가지고 있는 주요정보는 보존한 채 데이터의 불필요한 노이즈를 제거하여 데이터 처리 시간을 단축시킨다.

이러한 PCA는 오리지널 데이터의 전처리 과정을 수행하고 이를 공분산시켜 분산 데이터를 추출한다.

N개의 이미지 데이터에 PCA를 적용 시 전체 데이터는 $\{x_1, x_2, x_3, \dots, x_N\}$ 로 구성이 되며, 식(1)에서 각 오리지널 데이터 x_i , 전체 데이터의 평균 u 를 이용하여 데이터의 공분산 데이터인 S_T 를 도출한다. 예를 들어, (60000, 784)의 전체 이미지 크기를 가지는 MNIST 학습데이터를 PCA 적용 시 각 x_i 는 (784)의 크기를 가지는 데이터로

형성되며, 전체 데이터 평균 μ 도 (784)의 크기로 구성된다. 이후 도출한 공분산 데이터를 기반으로 데이터의 고유값, 고유벡터를 추출하면 식 (2)가 나오게 되는데 데이터의 고유값 λ_{PCA} , 데이터의 고유벡터 W_{PCA} 의 상관관계를 확인할 수가 있다. W_{PCA} 는 PCA를 통해 얻어진 각 고유값 λ_{PCA} 에 매칭되는 축 데이터로 이를 PCA의 주성분이라고 부른다. 매칭된 고유값의 크기에 비례하여 W_{PCA} 가 구성하는 축 데이터가 큰 분산치를 가지게 된다.

이후 축소하고자 하는 차원의 크기에 따라 (feature, feature) 크기를 가지는 고유벡터 W_{PCA} 의 차원을 축소한다.

$$S_T = \sum_{i=1}^N (x_i - \mu)(x_i - \mu)^T \quad (1)$$

$$\lambda_{PCA} = W_{PCA}^T S_T W_{PCA} \quad (2)$$

PCA는 최종적으로 원본데이터의 차원 크기를 원하는 차원의 크기만큼 축소를 진행시켜야 하고 이는 $Z = XW_{PCA}$ 와 같이 나타낼 수 있다. 이는 원본데이터 X 와 W_{PCA} 의 곱을 통해 (데이터의 개수, 축소된 차원의 크기) 데이터 형태를 지니는 차원이 축소된 데이터인 Z 를 도출하는 것을 의미한다.

2.2 LDA(Linear Discriminant Analysis)

LDA는 PCA와 같이 데이터 차원을 축소하는 알고리즘으로 데이터의 각 클래스 정보를 고려하여 차원 축소를 진행한다. LDA 방식을 살펴보면 클래스 간 분산(S_W)은 최소화하고, 서로 다른 클래스 간 분산(S_B)은 최대화하여 서로 다른 클래스 간의 거리를 최대한 멀리 구성하고 같은 클래스 내 거리를 최대한 가깝게 축소 차원을 구성한다. 이러한 S_W , S_B 공식 및 이를 통한 LDA 목적 함수는 아래와 같이 구성 가능하다.

$$S_B = \sum_{i=1}^c (\mu_i - \mu)(\mu_i - \mu)^T \quad (3)$$

$$S_W = \sum_{i=1}^N \sum_{x_k \in c_i} (x_k - \mu)(x_k - \mu)^T \quad (4)$$

$$\lambda_{LDA} = \frac{W^T S_B W}{W^T S_W W} \quad (5)$$

식(3)의 서로다른 클래스간 분산(S_B)식은 각 클래스 구성 데이터의 평균값 μ_i 에 전체데이터의 평균값 μ 를 뺀 값의 공분산 처리한 값으로 구성된다. 식(4)의 동일한 클래스 간 분산(S_W)은 각 클래스의 데이터 x_k 에 전체데이터의 평균값 μ 를 뺀 값의 공분산 처리한 값으로 구성된다. 식(3), (4)로 서로 다른 클래스 간 분산과 동일 클래스간 분산을 구하고 식(5)를 통해 S_B 과 S_W 의 역행렬을 곱하여 S_B 값은 최대화하고 S_W 값은 최소화한 차원 축 데이터를 도출한다.

2.3 PCA+LDA를 이용한 특징되먹임

[4][5]에서는 중요 영역을 추출하기 위한 PCA+LDA 특징 되먹임 알고리즘을 제안하였으며 Yale데이터를 기반한 LDA 패턴 인식[5]을 적용하였다. 이 과정을 간단하게 그림1, 2로 표현이 가능하다.

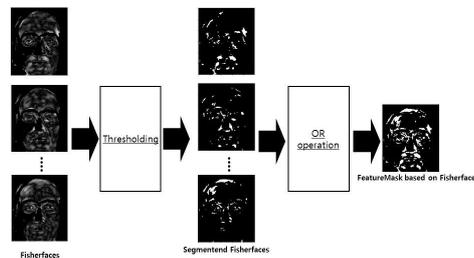


그림 1. PCA+LDA 특징마스킹 추출
Fig. 1. Extract Feature mask

위의 과정을 간단하게 살펴보면 우선 PCA를 통해 고유벡터인 W_{PCA} 을 평균값을 기준으로 이진화하여 대략적인 데이터의 중요 영역을 추출한다. 이때 고유값이 가장 큰 첫 번째 차원의 Eigenface를 적용하여 데이터의 대략적인 노이즈를 제거한다.

이후 노이즈를 제거한 데이터에 PCA를 적용하여 해당 데이터의 차원의 크기를 데이터 개수보다 작게 줄여 SSS(Small Sample Size)문제[7]를 회피한다. 이후 원본데이터 개수 이하의 차원을 가진 데이터에 LDA를 적용시켜 Fisherface를 추출하고 Fisherface의 평균값을 기준으로 Fisherface의 중요지점을 분류하여 이진화시킨다. 이렇게 얻은 Featuremask를 원본 데이터에 적용하여 정확도 분류를 진행하고 이를 통해 특징되먹임의 성능을 확인한다.

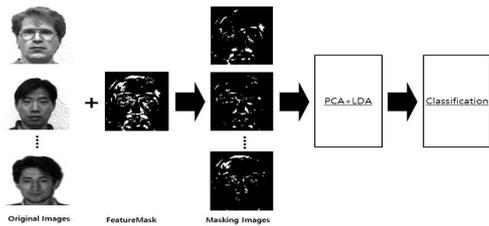


그림 2. 특징 되먹임 기반 얼굴 인식
Fig. 2. Face recognition based on Feature feedback

3. PCA를 이용한 특징 되먹임

이 논문에서는 PCA를 이용한 Eigenface를 기반으로 데이터의 중요 특징지점을 추출하는 방법을 제안한다.

PCA를 이용해 차원을 축소하고 축소된 각 차원을 이루는 벡터 공간인 Eigenface를 추출한다. Eigenface는 원본 데이터의 공분산 데이터의 고유값의 크기에 Eigenface의 얼굴 특성이 비례한다. 이에 각 Eigenface에 할당된 고유값의 크기가 작아질수록 Eigenface의 얼굴 특성이 작아지게 된다.

이 논문에서 제안하는 알고리즘은 다음과 같다.

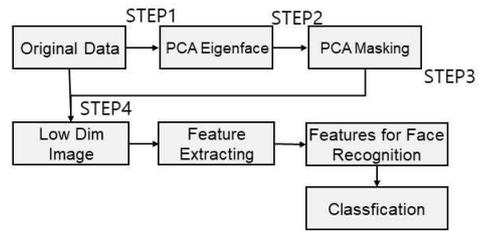


그림 3. PCA 특징되먹임 순서도
Fig. 3. PCA Flow chart of Feature Feedback

STEP1 : 원 데이터에 PCA하여 고유벡터 W_{PCA} 를 추출 및 평균값을 기준으로 이진화시켜 Eigenface를 생성한다.

STEP2 : 각 Eigenface에 고유값을 통한 가중치를 부여해서 합하여 Eigenmask를 생성한다.

STEP3 : Eigenmask를 임의의 기준값으로 이진화시켜 최종적인 PCA 특징마스크를 생성한다.

STEP4 : PCA특징마스크를 원본데이터에 적용하여 특징 데이터를 생성한다.

이상의 과정을 거쳐 데이터의 특징만 가지게 되는 특징데이터를 추출하게 된다.

STEP1과정에서 원본데이터에 PCA를 적용하여 고유값 λ 과 고유벡터 W_{PCA} 를 추출한다. 그리고 W_{PCA} 를 식(6)에 따라 평균값을 기준으로 이진화시켜 데이터의 중요한 부분(EI : Eigenface Important)과 중요하지 않은 부분(EU : Eigenface Unimportant) 획득할 수 있다. T_{PCA_i} 는 W_{PCA_i} 를 이진화시키는 기준점 값으로 W_{PCA_i} 의 평균값으로 구성한다. W_{PCA_i} 는 각 고유값에 해당하는 고유벡터이고 이를 구성하는 값 중 W_{PCA_i} 의 평균값(T_{PCA_i})보다 큰 값들은 W_{PCA_i} 에서 중요한 값으로 판단하여 EI에 해당되고, 평균값보다 작은 값들은 중요하지 않은 값으로 판단하여 EU에 해당하게 된다.

$$\begin{cases} W_{PCA} \in EI & \text{if } \|W_{PCA}\| \geq T_{PCA} \\ W_{PCA} \in EU & \text{otherwise} \end{cases} \quad (6)$$

MNIST데이터에 PCA를 사용한 식(6)를 적용하

여 MNIST데이터의 Eigenface를 5개 추출하면 그림4과 같은 이미지가 추출될 수 있다.



그림 4. MNIST Eigenface 이미지
Fig. 4. MNIST Eigenface Image

STEP2과정에서 각 Eigenface에 고유값 λ 에 비례하여 가중치를 부여하여 식(7)과 같이 Eigenmask를 구성할 수가 있다. 이 과정을 통해 각각의 Eigenface이미지의 중요도까지 고려를 하여 최종적인 PCA특징마스크를 구성하는데 사용이 될 수 있다. 식(7)과정으로 도출된 MNIST의 Eigenmask(E_m)는 그림5 같은 이미지로 표현이 된다. E_m 은 식(6)에서 도출한 l 개의 EI 합으로 구성한다. 하지만 각 EI와 매칭된 고유값의 크기가 다르고 이에 전체 데이터에서 각 EI가 가지는 중요도도 다르므로 이를 고려하여 각 EI에 $\sqrt{\lambda}$ 를 곱하여 각 EI의 중요도를 고려한 식을 구성한다.

$$E_m = \sqrt{\lambda_1}EI_1 + \sqrt{\lambda_2}EI_2 + \dots + \sqrt{\lambda_l}EI_l \quad (7)$$

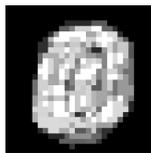


그림 5. MNIST Eigenmask 이미지
Fig. 5. MNIST Eigenmask Image

STEP3과정에서는 STEP2이 과정에서 획득한 E_m 을 임의의 기준점값 T 를 기준으로 이진화시켜 Eigenmask의 중요한 부분 EI(Eigenmask Important), 중요하지않은 부분 EU(Eigenmask Unimportant)를 획득할 수 있다.

$$\begin{cases} E_m \in EI & \text{if } \|E_m\| \geq T \\ E_m \in EU & \text{otherwise} \end{cases} \quad (8)$$

본 논문에서는 기준점값 T 를 데이터의 픽셀개수를 일정량으로 기준점을 잡아서 기준을 지정한 다. 예로 MNIST데이터에서 기준점 T 를 데이터 크기에서 값이 큰 상위 30%를 잡도록 설정 시 EI(Eigenmask Important)를 도출하면 그림6과 같은 이미지로 표현할 수 있다.



그림 6. Eigenmask 상위 30%픽셀 중요부분
Fig. 6. Eigenmask TOP 30% important pixel

이러한 STEP1 ~ STEP3의 과정을 전체적인 그림으로 표현하면 그림7과 같은 과정을 보인다.

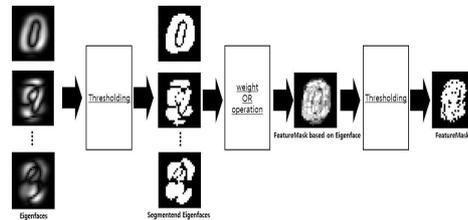


그림 7. MNIST PCA 특징마스크 추출
Fig. 7. Extract MNIST PCA Featuremask

STEP4에서는 STEP3에서 추출한 Eigenmask Important를 원본 데이터에 적용을 시켜 데이터에서 중요한 부분만 살리고 중요하지않은 부분은 제거한 데이터로 바꾼다. 이 과정을 STEP3의 상위 30% 픽셀기준으로 추출한 EI로 적용시 그림8과 같은 MNIST 원본이미지의 변화를 표현할 수 있다.

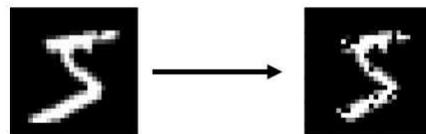


그림 8. MNIST 원본이미지 중요영역 필터링
Fig. 8. MNIST Original data Filtering

해당 PCA 특징되먹임 과정들을 다 진행하고

획득한 특징데이터를 PCA를 통해 분류를 진행하여 특징되먹임 정확도 비교를 진행한다.

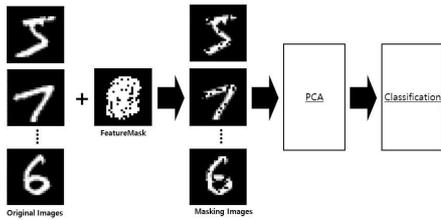


그림 9. PCA 특징되먹임 기반 정확도 분류
Fig. 9. PCA Featurefeedback Classification

4. PCA 특징되먹임 실험 결과

MNIST의 특징되먹임을 하기 위해 우선 PCA, LDA를 이용한 MNIST데이터 정확도 분류를 진행한다. 해당 과정을 통해 MNIST데이터 정확도 분류 성능이 좋은 차원 축소 알고리즘을 확인한다. 해당 정확도를 분류 시 분산량을 기준으로 분류를 진행한다. 이때 분산량은 축소된 차원이 보유한 고유값의 비율로 정확도 분류 시 분산량은 PCA, LDA 방식 모두 93%로 설정을 하였다. 이때 PCA를 통한 정확도 분류는 97.26%, LDA를 이용한 분류는 89.74%가 나오는 것을 확인할 수가 있다. 따라서, MNIST와 같은 데이터셋은 PCA 기반의 특징되먹임이 LDA 기반 특징 되먹임[4][5]에 비해서 성능이 좋을 수 있다.

표 1. PCA와 LDA기반 MNIST 정확도 분류
Table 1. PCA & LDA base MNIST Classification

데이터 분산량	MNIST-PCA	MNIST-LDA
93%	97.26	89.74

제안된 PCA를 이용한 데이터의 중요지점 판단 알고리즘을 MNIST데이터셋에 적용하여 성능을 평가한다. MNIST데이터셋은 사람 손글씨 데이터로 손글씨로 쓰여진 0 ~ 9숫자에 대한 70000개의 데이터로 구성된다. 해당 데이터는 28x28의 이미지로 구성이 되어있어서 데이터의 개수보다 작은 형태로 구성되어있다. 이에 MNIST는 LDA

를 통한 정확도 분류 진행할 때 차원을 미리 축소를 할 필요가 없게 되어 PCALDA를 통해 정확도 분류나 fisherface를 추출하지않고 LDA만 이용하여 과정을 수행한다. 본 실험에서는 PCA를 이용한 특징마스킹 특징점 확인, LDA를 이용한 특징마스킹 특징점 확인 결과를 비교하여 PCA 특징마스킹의 성능을 확인한다.

이때 실험조건은 PCA, LDA에서 각 마스크를 추출하고 마스크값에서 상위 30%(EI, FI), 하위 30%(EU, FU)를 추출한 뒤 이를 원본 데이터에 적용시켰을 때의 정확도 분류를 구하여 표2, 그림6으로 구성을 하였다.

이때 PCA를 통한 정확도 분류시 93%의 분산을 기준으로 분류를 진행하였다.

실험을 통해 PCA를 이용한 특징되먹임 EI가 전체 데이터의 30%만 사용했음에도 원본 데이터와 유사한 정확도를 보임을 확인하며 기존의 머신러닝을 통한 특징되먹임 방식인 LDA 적용 방식보다 2.23%의 성능 우위가 있음을 확인한다.

표 2. 특징마스킹과 분산량에 따른 인식률 변화
Table 2. Recognition rate of Feature mask and the Dispersion rate

데이터 분산량	Original Data	EI Featuremask	EU Featuremask	FI Featuremask [4]	FU Featuremask [4]
50%	93.74	91.39	15.95	84.81	57.90
93%	97.26	97.30	17.16	95.07	81.40

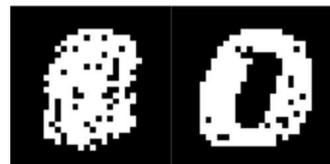


그림 10. EI, FI PCA 특징마스킹 이미지
Fig. 10. EI, FI PCA Featuremask image

표3은 특징마스킹의 데이터 이진화 비율을 상위10%(78개) ~ 90%(705개)까지 구분하여 정확도 분류를 진행한 실험 결과이다. 해당 실험을 통해 각 이진화 비율에서의 EI, FI 및 추가적으

로 단순히 데이터의 평균값에서 데이터의 중요치를 뽑아내는 픽셀 카운팅을 비교하여 정확도 분류를 진행한다. 픽셀 카운팅은 단순히 각 픽셀의 값을 더하여 가장 많은 순서로 데이터를 뽑아낸 방법이다. 이때 PCA를 통한 정확도 분류 시 93%의 분산을 기준으로 분류를 진행하였다. 이 결과로 MNIST데이터에서 EI, FI, 픽셀 카운팅간의 정확도 분류 시 제안된 PCA를 적용한 특징되먹임 방식의 평균 정확도는 96.99%, 기존의 LDA를 적용한 특징되먹임 방식의 평균정확도는 95.09%, 픽셀 카운팅의 평균 정확도는 68.62%가 나오는 것을 확인할 수 있다. 이를 통해 MNIST데이터의 특징되먹임은 PCA > LDA > 픽셀 카운팅 순으로 높은 성능을 보임을 확인한다.

표 3. 특징마스크와 이진화 비율에 따른 인식을 변화
Table 3. Recognition rate of Feature mask and Binarization rate

Binarization rate	EI Feature mask	FI Feature mask[4]	pixel counting
10% (78개)	95.04	91.72	53.63
20% (156개)	96.92	94.24	54.32
30% (235개)	97.33	95.07	54.59
40% (313개)	97.34	95.32	55.08
50% (392개)	97.27	95.59	57.94
60% (470개)	97.24	95.48	71.28
70% (548개)	97.25	95.71	83.12
80% (627개)	97.24	95.70	91.61
90% (705개)	97.24	96.95	96.00
평균	96.99	95.09	68.62

5. 결론

이 논문에서는 PCA를 이용한 특징되먹임 방법을 제시하고 MNIST데이터셋으로 PCA 특징되먹임의 성능과 기존의 LDA를 이용한 특징되먹임의 성능 비교를 진행했다. 이 논문의 연구는 MNIST데이터에서 데이터 차원 축소시의 동일한 데이터 보존량을 기준으로 PCA가 LDA보다 성능이 높다는 것을 기반으로 MNIST데이터에서 PCA기반 특징되먹임의 높은 인식률을 예상하고 진행했다. 해당 실험은 PCA, LDA모두 93% 가량 이상의 분산량을 가질 때 기반으로 Eigenface, Fisherface를 추출하여 각 차원 축소 알고리즘의 특징 마스크를 추출했다. 두가지 실험결과 MNIST데이터의 경우 PCA를 이용한 특징마스크 적용 시 LDA를 이용한 특징마스크보다 확실한 성능 우위를 보인다. 이러한 결과를 보았을 때 MNIST데이터의 특징되먹임을 진행 시 PCA가 LDA보다 좋은 성능을 보임을 확인하였다.

이를 통해 데이터셋의 종류에 따라 기존의 LDA 특징되먹임 뿐이 아닌 제안된 PCA 특징되먹임도 적용을하여 보다 확실한 데이터 중요 지점을 판단할 수 있음을 확인할 수 있다.

향후에는 MNIST뿐만 아닌 추가적인 데이터셋들을 통한 연구를 진행하여 PCA, LDA 특징되먹임 성능 차이에 대한 추가적인 연구와 더불어 특징되먹임 뿐만 아닌 데이터 복원 등 다른 요소에서의 PCA, LDA의 성능 비교 연구가 필요할 것으로 생각된다.

REFERENCES

- [1] M. Turk and A. Pentland, "Eigenfaces for recognition," J. Cognitive Neurosci. vol. 3, no. 1, pp. 71-86, 1991
- [2] W. Lu, KN. Plataniotis, and N. Venetsanopoulos, "Face recognition using LDA -based algorithm," IEEE Trans. Neural Network, vol. 14, no. 1, pp. 195-200, 2003.
- [3] T. V. Bandos, L. Bruzzone and G.

Camps-Valls, "Classification of hyperspectral images with regularized linear discriminant analysis", IEEE Trans. Geosci. Remote Sensing, vol. 47, no. 3, pp. 862-873, 2009.

[4] G.-M. Jeong, H.-S. Ahn, S.-I. Choi, N. Kwak, C. Moon Pattern recognition using feature feedback: application to face recognition International Journal of Control, Automation and Systems, 8, pp. 141-148, 2010

[5] Su-Hyun Kim, Sang-Il Choi, Sung-Han Bae, Young-Dae Lee, Gu-Min Jeong, "Pattern Recognition using Feature Feedback : Performance Evaluation for Feature Mask" The Institute of Internet, Broadcasting and Communication Vol 10.5, pp 179-185, October 2010

[6] Sang-Il Choi, Su-Hyun Kim, Yoonseok Yang, Gu-Min Jeong, "Data Refinement and Channel Selection for a Portable E-Nose System by the Use of Feature Feedback" Sensors 10387-10400 October 2010

[7] Alok Sharma, Kuldip K. Paliwal "Linear discriminant analysis for the small sample size problem: an overview", International Journal of Machine Learning and Cybernetics, 6, pp 443-454, 2015

[8] Sebastian Bach, Alexander Binder, Grégoire Montavon, Frederick Klauschen, Klaus-Robert Müller, Wojciech Samek "On pixel-wise explanations for non-linear classifier decisions by layerwise relevance propagation" PloS One e0130140 July 2015

[9] Selvaraju R.R, Cogswell M, Das A, Vedantam R, Parikh D, Batra D "Grad-CAM : Visual Explanations from Deep Networks via Gradient-based Localization" Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp 618-626 2017

[10] Y. LeCun, "The MNIST database of handwrittendigits," <http://yann.lecun.com/exdb/mnist>

저자약력

이 승 현(Seung-Hyeon Lee)

[일반회원]



- 현재 국민대학교 전자공학과 석사과정

〈관심분야〉 차량용 소프트웨어, 인공지능, 딥러닝, 영상처리, 자율주행

김 도 연(Do-Yun Kim)

[일반회원]



- 현재 국민대학교 전자공학과 석사과정

〈관심분야〉 차량용 소프트웨어, 인공지능, 딥러닝, 영상처리, 패턴인식

최 상 일(Sang-Il Choi)

[일반회원]



- 2005년 서강대학교 전자공학과(학사)
- 2010년 서울대학교 전기·컴퓨터공학부 (석박사통합)
- 2010년~2011년 Post Doctoral Fellow, BK21 Information Technology, Seoul National University, Seoul, Korea
- 2011년 Post Doctoral Fellow, Computer Science Department, University of Southern California, CA, USA
- 2011년~현재 단국대학교 컴퓨터학과 조교수

〈관심분야〉 Pattern Recognition & Computer Vision

정 구 민(Gu-Min Jeong)

[일반회원]



- 1995년 서울대학교, 제어계측공학과 학사
- 1997년 서울대학교, 제어계측공학과 석사
- 2001년 서울대학교, 전기컴퓨터공학부 박사
- 2001년~2004년 (주) 네오엠텔 기반 기술팀, 팀장(co-founder)
- 2005년~현재 국민대학교 전자공학부 교수
- 2013년~현재 (주) 유비벨룩스 사외이사
- 2015년~현재 국가기술표준원 자동차전기전자 및 통신 전문위원회 위원장
- 2019년~현재 (주) 휴맥스 사외이사

〈관심분야〉 차량용 마이컴, 차량용 소프트웨어, 커넥티드카, 자율주행