

## 얼굴 표정 기반의 극심한 통증 검출

기민송<sup>1</sup> · 최영우<sup>2\*</sup><sup>1</sup>연세대학교 컴퓨터과학과 박사과정<sup>2</sup>숙명여자대학교 컴퓨터과학과 교수

## Extreme Pain Detection based on Human Facial Expressions

Minsong Ki<sup>1</sup> · Yeongwoo Choi<sup>2\*</sup><sup>1</sup>P.HD student, Department of Computer Science, Yonsei University, Seoul, Korea<sup>2</sup>Professor, Department of Computer Science, Sookmyung Women's University, Seoul, Korea

### [요약]

다양한 환경에서 사람의 안전을 위해 비전 기술을 이용하여 극심한 통증을 감지하는 연구가 필요하다. 예를 들어 운전 중에 갑작스런 고통이나 심 정지 등의 위급한 상황으로 인해서 대형사고가 발생할 수 있으므로, 본 논문에서는 인간의 위급한 상황을 검출하기 위해 얼굴 표정 인식 기반의 통증 검출 시스템을 제안한다. 제안하는 시스템은 극심한 통증을 느꼈을 때의 얼굴 표정 클래스를 별도로 구성하고 LeNet 모델을 수정하여 사용하였다. 또한, 학습 데이터 내 노이즈를 해결하기 위한 리샘플링 과정을 추가하고, 샘플이 적고 분류하기 어려운 통증(Painful) 클래스를 위해 특징 공간에서 표정 클래스 별로 클러스터링이 잘 되도록 링 손실 함수를 사용하였다. 7가지 얼굴 표정 관련 공인데이터인 FER2013[1]과 Pain Expressions [2]에서 추출한 통증 클래스를 추가하여 사용한 검증과 테스트에서 각각 63.3%, 60.4%에 해당하는 정확도를 보여 인식 성능이 개선된 것을 확인하였다. 이러한 연구 결과는 추후 운전자의 갑작스런 통증으로 인한 대형사고를 미연에 방지하는 시스템의 개발에 활용될 수 있다.

### [Abstract]

We need research to detect human extreme pains using vision technologies in various environments for safety. For example, a car driver may be in a state of emergencies such as sudden pain or cardiac arrest. Therefore, we propose an emergency detection system based on human facial expressions to detect human emergent states. We organize the data for painful facial expression classes separately and propose a modified LeNet to use as a baseline. We add a resampling process to solve the noise in the training data. In addition, for Painful class with few samples and difficult to classify, we apply ring loss with softmax for clustering by facial expression class in feature space. We show accuracies of 63.3% and 60.4% for validation and testset with 8 expression classes both from 7 expression classes of FER2013[1] and an added pain class extracted from Pain Expression [4]. These results can hopefully be used to develop a system that can prevent terrible car accidents due to a sudden pain of the car drivers.

색인어 : 얼굴 표정 인식, 위급 상황 검출, 통증 검출, 통증 클래스, 링 손실 함수

**Key word:** Facial expression recognition, Emergency detection, Pain detection, Painful class, Ring loss, etc

<http://dx.doi.org/10.9728/dcs.2020.21.2.415>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 24 December 2019; Revised 13 January 2020

Accepted 25 February 2020

\*Corresponding Author; Yeongwoo Choi

Tel: 

E-mail: [ywchoi@sookmyung.ac.kr](mailto:ywchoi@sookmyung.ac.kr)

## I. 서론

최근 사람에게 유용한 인공지능 기술에 대한 연구가 다양하게 진행되고 있다. 이는 모바일, 웹, 차량 등에 IT 기술을 융합하여 자동으로 사람에게 안전성과 편의성을 제공해 준다. 가장 중요한 안전성을 위해 부주의한 행동 및 심 정지 등의 갑작스런 상황으로 사람이 정상적으로 행동 할 수 없는 경우에 사고 감지를 위해 사람의 위급 상황을 실시간으로 검출하는 기술 개발이 중요하다. 표정은 사람의 상태 및 감정에 대한 가장 직접적이고 시각적인 표현이다.

따라서, 본 논문은 사람의 감정 및 상태가 표현되는 얼굴 표정을 자동으로 인식하여 위급 상황을 감지 할 수 있는 시스템을 제안한다. 이는 지능형 자동차, 지능형 감시 카메라 등에 요소 기술로서 다양하게 활용될 수 있다.

일반적으로 얼굴표정 인식은 얼굴 영역 검출, 얼굴 특징 추출 및 표정 분류의 3단계로 진행된다[5]. 얼굴 영역 검출 과정은 복잡한 배경으로부터 얼굴 영역의 위치를 자동으로 찾는다. 검출한 얼굴 영역에서의 특징 추출 방법은 외관(appearance)기반의 방법과[6,7] 각각의 표정동작 단위를 의미하는 AU(action units)[8]를 통해 움직임의 정도를 파악하는 방법으로 크게 분류된다. 마지막으로 추출한 특징을 이용한 표정 분류에는 SVM(support vector machine), KNN(k-nearest neighbor), LDA(linear discriminant analysis) 등이 일반적으로 많이 사용된다[9]. 또한, 최근에는 대량의 학습 데이터를 이용한 딥러닝 기반의 방법도 활발히 연구되면서 얼굴 인식 및 표정 인식 분야에 서 좋은 성능을 만들어 내고 있다[10, 11, 12, 18, 19].

지금까지의 연구에서[10, 11, 12, 20, 21, 22] 사용된 데이터 셋은[1,3,4] 통증을 느끼는 표정의 클래스를 포함하지 않았다. 본 논문에서는 FER2013[1] 데이터 셋의 7개의 얼굴 표정 클래스에 통증을 느꼈을 때의 표정 클래스를 추가하여 총 8개의 클래스로 얼굴 표정을 분류한다. 통증에 해당되는 학습 데이터는 Pain Expressions[2] 데이터 셋에서 통증을 심하게 느끼는 부분의 얼굴 표정 데이터를 발췌해서 사용한다. FER2013 데이터 셋은 7개의 클래스 별로 데이터 수의 차이가 심한 불균형 문제점과 화남 및 역겨움 등 일부 클래스 간에는 얼굴 표정의 차이가 크지 않다. 본 논문에서는 이러한 문제점들을 해결하여 인식 성능을 높이고자 했으며, 분류된 8개의 클래스를 바탕으로 사람의 통증으로 인한 위급한 상태를 검출하고자 한다. 본 논문의 II 장에서는 제안하는 시스템의 개요와 방법을 설명하고, III장에서 실험 및 분석결과를 정량적, 정성적 평가를 통해 제시한다. 결론 및 향후 연구는 IV장에서 기술한다.

## II. 제안하는 방법

제안하는 시스템은 사람의 갑작스런 통증으로 인한 위급한 상태 여부를 감지하는 것으로서 전체적인 흐름은 그림 1과 같

다. 입력된 영상에서 얼굴 영역을 검출하고, 검출된 얼굴 영역을 개선된 LeNet[15] 신경망에 입력하여 얼굴을 8개의 표정으로 분류한다. 8개 표정에 대한 결과값들을 종합하여 위급(Emergency) 상태와 비위급 상태(Normal)을 판단한다.

### 2-1 얼굴 검출

입력된 얼굴 영상에서 정확한 얼굴 위치를 검출하기 위해서 Viola와 Jones가 제안한 안정적이고 검출 성능이 우수한 Adaboost 기반의 얼굴검출 알고리즘을[13] 사용하였다. 이 방법은 Viola와 Jones가 제안한 Haar 특징을 기반으로 얼굴을 검출하며, Haar 특징 값은 밝은 부분에 해당하는 영상 픽셀들의 밝기 합에서 어두운 부분의 밝기 합을 뺀 차로 계산한다. 이 밝기 값의 차가 일정 임계값 이상일 때 검출한다[14]. 결과로서 검출된 얼굴의 위치를 박스 형태로 출력한다.

### 2-2 개선한 LeNet 모델

앞에서 검출한 얼굴 영역의 표정을 인식하기 위해 본 논문에서는 LeCun[15]이 설계한 LeNet을 변형하여 사용한다. LeNet은 컨볼루션(convolution) 연산과 서브 샘플링(subsampling)을 반복적으로 수행하면서 마지막 층은 완전 연결(fully-connected)층으로 주어진 입력을 분류하는데, 3개의 컨볼루션 층과 2개의 서브 샘플링 층, 그리고 1개의 완전 연결 층으로 구성되어 있다.

컨볼루션 층은 학습 가능한 여러 종류의 필터들을 이용하여 이미지에서 특징을 추출하고, 서브 샘플링 층은 컨볼루션 층에서 검출된 특징들의 부분적인 이동변환에 무감하도록 샘플링하는 과정으로서 해상도를 줄여주는 역할도 한다. 비교적 파라미터 수가 적고 소수의 층으로 구성되어 있어서 차량에 탑재하기에 적당한 크기의 모델이다. 본 논문에서는 48x48 크기의 얼굴 해상도 이미지를 사용하므로 그림 2와 같이 이미지 크기를 줄여주는 과정을 생략하고, 6개의 컨볼루션 층으로 5x5 크기의 필터를 사용한다. 각 노드에 입력된 값들은 PReLU 활성화 함수를 통과 시킨 후에 다음 층으로 전달한다. 기존의 ReLU 함수는 0보다 작은 값에 대해서는 기울기가 0이어서 뉴런이 비활성화되는 단점이 있지만, PReLU 함수는 새로운 파라미터를 추가하여 0보다 작은 값에 대해서도 기울기를 학습할 수 있는 장점을 갖고 있다.

### 2-3 불균형한 데이터 및 데이터 내 노이즈 처리

구성된 모델에 데이터를 학습할 때 반복 1회 당 사용하는 데이터의 셋을 배치(Batch)라고 하며, 본 논문에서는 배치 크기를 32로 사용한다. 학습에 사용한 FER2013 데이터 셋은 7개의 클래스로서 그림 3과 같이 데이터가 고르게 분포되어 있지 않으며, 클래스가 틀리게 레이블링된 데이터도 다수 존재한다. 이러한 문제점들을 노이즈로 간주하며 학습 데이터의 노이즈로

인해서 분류 또는 인식 성능이 하락할 수 있다[16]. 이를 해결하기 위해서 전체 데이터 셋에 대해서 틀리게 레이블링된 데이터들을 정제하고, 또한 데이터를 구성할 때 리샘플링 과정을 추가하여 데이터 수의 분포를 고르게 하였다.

리샘플링 과정은 가장 많은 수의 샘플을 갖는 클래스에서는 샘플을 제거하고(Under-sampling), 샘플 수가 적은 수의 클래스에서는 샘플을 추가하는 것 (Over-sampling)으로 구성된다. 본 논문에서는 소수 클래스에 샘플을 추가하여 각 클래스별 균형을 이루도록 하며 이를 위해 학습시에 먼저 클래스 별로 이미지 수를 계산하여 클래스 데이터의 발생 확률을 계산한다. 매 에폭 (Epoch)마다 전체 데이터 셋을 샘플링하며 확률에 반비례하도록 가중치를 주어 리샘플링한다. 이 과정을 통해 초기에 샘플 수가 적은 클래스는 데이터가 추가되면서 데이터의 클래스별 균형을 맞출 수 있게 된다.

**2-4 모델 학습**

균형 있는 데이터 셋을 구성한 후에 수정한 LeNet 모델의 최종 출력층에 결과가 나오면 그 결과에 대한 손실(Loss)을 계산하며, 학습의 최종 목표는 손실이 최소화된 모델을 만드는 것이다. 본 논문에서는 식1과 같이 설계한 손실 함수를 사용한다. 그림 1에서 완전 연결 층을 거치고 소프트맥스 함수를 통과하면 0과 1사이의 정규화된 값이 출력된다. 따라서 학습은 소프트맥스 함수가 예측한 값과 실제 값 사이의 차이를 계산하여 이 차이가 줄어드는 방향으로 진행된다.

FER2013 데이터 셋은 클래스마다 얼굴 표정의 차이가 뚜렷하지 않은 경우도 많아서 기존의 소프트맥스 함수를 통한 계산한 손실만으로는 정확한 분류가 어렵다. 소프트맥스 함수는 멀티 클래스 분류를 쉽게 구현해 주는 장점이 있지만, 데이터 샘플의 분포에 편향되는 단점을 갖는다. 즉 주어진 미니 배치에서 모든 샘플의 조건부 확률을 최대화시켜서 샘플이 적고 분류하기 어려운 경우를 고려하지 않는 단점이 있다.

이러한 문제점을 해결하며 특징 공간에서 표정 클래스 별로 클러스터링을 잘하기 위해서 특징 정규화(Feature normalization) 과정이 중요하다. 본 논문에서는 전체 샘플 특징들을 정규화하기 위해 링 손실값[17]을 적용하였다. 링 손실값은 단독으로 사용하기엔 값이 작아서 소프트맥스 손실과 더하여 최종 손실값으로 사용한다. 이를 사용하면 클래스 별로 특징 분리가 잘되며, norm constraint를 주기 때문에 저해상도 이미지에서도 인식 성능을 확보할 수 있는 장점이 있다. 링 손실값  $L_R$ 은 식 1과 같고, 여기서  $F(x)$ 는 어떠한 샘플  $x_i$ 에 대해 모델을 통과하여 나온 특징들이다.  $m$ 은 배치 사이즈이고,  $R$ 은 타겟 놈(Target norm)으로서 학습이 가능하다.  $\lambda$ 값은 가중치를 주어 기존 손실값과 추가된 링 손실값의 비율을 결정할 수 있게 해준다. 본 논문에서는 실험을 통해  $\lambda$ 값을 1로 설정하였다.

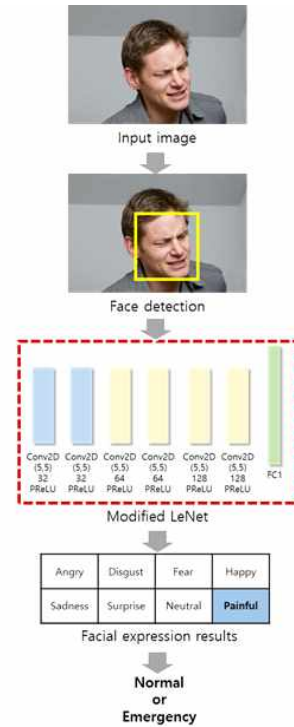


그림 1. 제안하는 시스템 흐름도  
Fig. 1. System Architecture

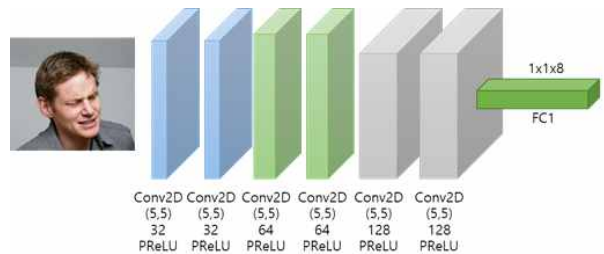


그림 2. 개선된 LeNet[15] 네트워크 구조  
Fig. 2. Modified LeNet[15] Network Architecture

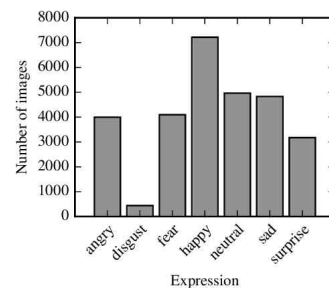


그림 3. FER2013[1] 데이터의 얼굴 표정 별 클래스 분포  
Fig. 3. Class distribution of each facial expression of FER2013[1] dataset

$$L_R = \frac{\lambda}{2m} \sum_{i=1}^m (\|F(x_i)\|_2 - R)^2 \quad (1)$$

### 2-5 얼굴 표정 판단

본 논문의 8개 얼굴 표정 클래스는 행복, 무표정, 화남, 역겨움, 무서움, 슬픔, 놀람, 통증이며, 이 중에서 행복과 무표정을 비 응급인 Normal 카테고리, 나머지를 응급인 Emergency 카테고리로 구분하였다. 이에 대한 입력된 얼굴의 최종 판단 결과 예는 그림 4와 같다. 이 중에서 통증 클래스에 해당되는 경우는 그림 5처럼 입이 벌어지거나 미간을 찌뿌리는 경우 등이 많이 나타나는데, 몇 종류의 표정 클래스들이 복합적으로 나타난다. 즉 행복, 무표정을 제외한 나머지 클래스에서 보여지는 특징들이 복합적으로 포함된다. 이 점에 착안하여 본 논문에서는 통증으로 인한 위급상황을 위와 같이 설정하였다. 응급 정도에 따라 카테고리 구분을 재 조정할 수도 있다. 그림 4에서 보듯이 입력 영상에 대해서 먼저 8개 클래스의 개별 확률이 출력되며, 클래스 확률 값의 총합은 1이 된다. 획득한 8개 클래스의 확률 값을 Emergency와 Normal 카테고리 별로 합산한 해당 값을 합계 점수라고 출력하여 큰 쪽으로 최종 판단한다.

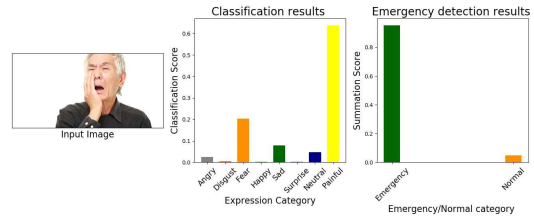


그림 4. 얼굴 표정 판단 예시  
 Fig. 4. Example of facial expression decision



그림 5. 통증 클래스 내의 얼굴 이미지 예시  
 Fig. 5. Example of facial images in painful class

### III. 실험 및 고찰

실험 환경은 Pycharm에서 PyTorch와 OpenCV 라이브러리를 사용하여 Python으로 구현하였다. 실험 데이터는 실제로 통증을 느끼는 데이터를 확보하기에 한계가 있어 48x48 크기의 해상도를 갖는 FER2013 데이터와 720x576 해상도의 Pain Expressions 데이터 및 웹에서 직접 수집한 다양한 해상도의 얼굴 이미지 100장을 함께 사용하였다.

#### 3-1 학습 데이터

학습에 사용한 데이터 셋의 예는 그림 6과 같다. 왼쪽 상단부터 FER2013 데이터 셋에 포함되어 있는 화남, 역겨움, 무서움, 행복함, 슬픔, 놀람, 무표정 7가지 클래스와 오른쪽 하단의 Pain Expressions 데이터 셋에서 얼굴 영역만 검출한 통증 클래스이다. Pain Expressions 데이터 셋에서 통증을 확실하게 느끼고 있는 얼굴 이미지들만 선별하여 그림 7처럼 총 300장의 이미지를 학습에 사용하였다. Pain Expressions 데이터 셋은 배경을 포함하는 RGB 컬러의 얼굴 영상이며, 따라서 학습 시 별도로 명도 이미지를 변환한 후 얼굴 영역을 검출하였다.

FER2013의 얼굴 데이터는 48x48 크기 해상도의 명도 이미지로서 학습 데이터 28,709장, 검증 데이터 3,589장 및 3,589장의 테스트 데이터로 구성된다. 표정 클래스가 틀리게 레이블링되어 있는 경우가 많은데 이를 올바른 클래스로 변경해서 사용했다.



그림 6. 각 클래스 별 학습 얼굴 이미지 예시  
 Fig. 6. Example of training images in each class



그림 7. Pain Expressions[2] 데이터 내 학습 이미지 예시  
 Fig. 7. Example of training images in Pain Expressions dataset[2]

#### 3-2 정량 및 정성적 평가

앞 절에서 설명한 것과 같이 클래스 별 데이터 분포를 고르게 하기 위해서 리샘플링을 먼저 수행하고, 전체 샘플 특징들을 정규화하기 위해 링 손실값[17]을 적용하여 기존의 소프트맥스 손실과 더한 후 최종 손실값으로 사용하였다.

표 1은 제안하는 시스템의 분류 정확성을 베이스라인 방법과 비교하여 보여주며, 배치(Batch) 사이즈 32, 학습률(Learning rate) 0.01의 동일한 조건으로 성능을 측정하였다. 베이스라인은 기존 소프트맥스 손실 함수만 사용한 방법이다. FER2013의



7개 표정에 통증(Painful) 클래스를 포함한 8개 클래스에 대한 정량적 정확도 평가 결과로서, 검증 및 테스트 데이터 모두에서 제안하는 방법의 성능이 향상된 것을 확인할 수 있다.

표 2는 각 과정의 성능 향상 정도를 비교하기 위해 단계 별로 과정을 추가하면서 측정한 정확도이다. 기준은 2.3절에서 제안한 수정된 LeNet으로서 소프트맥스 손실함수를 사용하였다. 두 번째 행은 기준 네트워크에 리샘플링 과정을 추가하여 클래스간의 데이터 균형을 맞춘 결과이고, 세 번째 행은 특징 정규화를 위해서 링 손실값을 추가하여 손실 함수로 사용한 결과이다. 마지막 행은 리샘플링 과정과 두 가지 손실값을 합하여 사용한 방법이다. 제안한 시스템의 성능으로서 과정 별 성능 향상을 통해 최종적으로 기준 네트워크에 비해서 약 5%의 정확도가 향상된 것을 보여준다. 링 손실값을 추가적으로 사용한 경우 클래스 별로 특징 분리가 잘 되어 유사한 클래스 간의 인식에 도움이 된 것을 확인할 수 있다.

표 3은 본 논문에서 제안한 8가지의 표정을 먼저 분류한 후 위급 상황 여부를 판단한 계층적 분류와 처음부터 응급 및 비응급 2개의 클래스로 표정을 분류한 성능을 비교하였다. 두 실험 모두 본 논문에서 제안한 수정된 LeNet 모델을 사용하고 동일한 조건에서 정확성을 측정하였다. 제안하는 방법은 두 단계를 거쳐 최종 위급 상황을 판단하므로 계층적 분류로 표기하였고, 단일층 분류는 Emergency 카테고리로서 6가지(화남, 역겨움, 무서움, 슬픔, 놀람, 통증) 클래스를 포함하고, Normal 카테고리는 나머지 2가지(행복함, 무표정) 클래스를 포함한 두 클래스 분류 결과이다. 정확성 측정은 최종 Emergency와 Normal 카테고리에 대한 측정 결과로서 FER2013[1] 데이터 셋에 통증 클래스를 추가한 테스트 데이터에 대한 결과이다. 제안하는 계층적 분류 방법이 좀 더 세분화된 표정 클래스를 먼저 인식한 후에 최종 위급 상황을 결정하여 단일층 분류 방법에 비해 보다 정확히 검출하는 것을 확인할 수 있었다.

연구에서 많이 사용하는 공인된 데이터 이외에도 웹에서 수집한 얼굴 영상으로 실험한 결과를 정성적인 관점에서 보면 아래와 같다. 그림 8은 옳게 분류한 영상 3개와 틀리게 분류한 영상에 대한 예이다. 처음 3개의 얼굴 영상은 Emergency와 Normal 카테고리를 잘 분류한 예를 보여준다. 4번째 얼굴 영상은 행복(Happy) 클래스에서 일반적으로 나타나는 입 꼬리가 올라가는 형태와 슬픔(Sad), 화남(Angry) 클래스에서 보이는 미간 찌푸림이 공통으로 나타나 오분류된 것으로 파악된다. 5번째 영상은 공포(Fear) 클래스를 무표정(Neutral)으로 처음 단계에서 오분류하여 최종 Normal 카테고리 틀리게 분류된 것을 보여준다.

표 1. 베이스라인과 제안하는 방법의 성능 비교(FER2013 + Painful class = 8 classes)

Table 1. Performance comparison between baseline and the proposed system (FER2013 + Painful class = 8 classes)

Accuracy(%)	Validation	Test
Baseline	58.20%	56.34%
<b>Proposed</b>	<b>63.30%</b>	<b>60.41%</b>

표 2. FER2013과 Pain Expressions 데이터에서의 과정 별 성능 비교(FER2013 + Painful class = 8 classes)

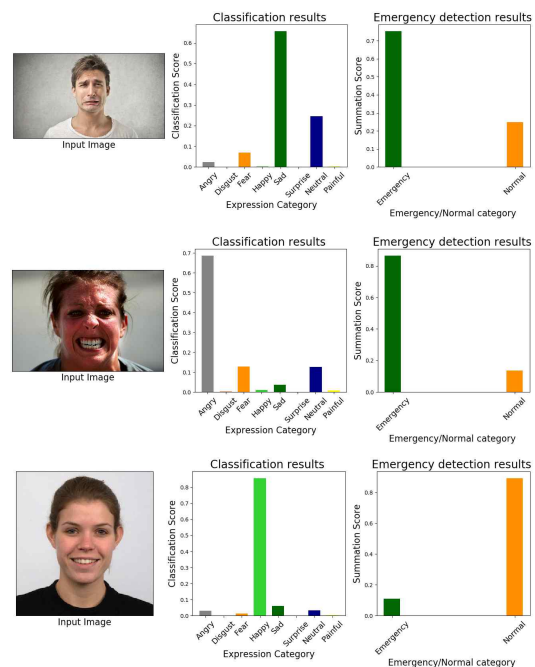
Table 2. Ablation test results on FER2013 and Pain Expressions dataset (FER2013 + Painful class = 8 classes)

Accuracy(%)	Validation	Test
Baseline (Softmax loss)	58.20%	56.34%
Softmax loss + resampling	58.42%	56.53%
Softmax loss + ring loss	63.11%	59.86%
<b>Proposed</b>	<b>63.30%</b>	<b>60.41%</b>

표 3. 위급 상황과 일반 상태를 결정하는 방법의 성능 비교 (Emergent/Normal = 2 classes)

Table 3. Performance comparison of deciding emergent or normal states. (Emergent/Normal = 2 classes)

Emergency Detection (Emergent/Normal)	Accuracy(%)
Single Classification	73.14%
<b>Proposed (Hierarchy Classification)</b>	<b>78.25%</b>



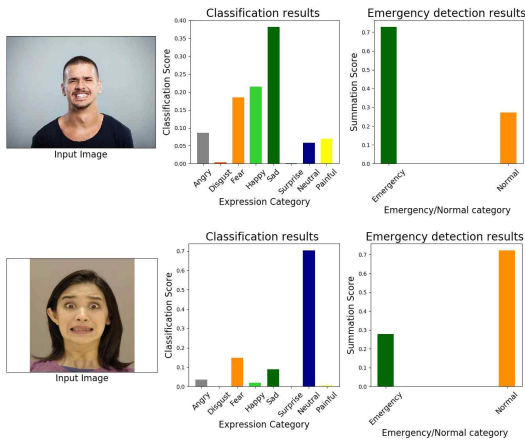


그림 8. 웹 이미지에서의 제안하는 방법 실험 결과  
**Fig. 8.** Results of the proposed system on web image dataset

#### IV. 결론

본 논문은 사람이 느끼는 갑작스런 통증으로 인한 위급상태를 파악하기 위해서 표정을 인식하여 분류하는 시스템을 제안하였다. 기존의 연구에서는 고통스런 얼굴 표정 클래스를 다루는 연구가 거의 없으며, 따라서 공인된 2개의 데이터 셋을 통합하여 연구를 수행하였다. 초기 단계의 연구로서 통증 클래스에 대한 충분한 데이터 확보 등 개선할 사항들이 많지만 초기 연구 결과로서 만족스런 결과를 얻었다고 판단한다.

향후 연구로는 얼굴 표정 및 위급 상황 발생에서 사람에게 나타나는 신체 특징인 머리의 방향 및 각도, 제스처 인식을 추가하여 응급상황에 대한 분류 정확성을 높이고자 한다. 또한, 이를 차량 운전자의 응급상황 판단에 적용하기 위해서 데이터 수집이 어렵겠지만 조명 변화가 다양한 응급상황의 운전자 얼굴 및 신체 실 데이터를 수집하여 운전자의 위급 상황을 정확히 검출할 수 있는 지능형 차량 안전시스템에 적용하고자 한다.

#### 감사의 글

본 연구는 2020년도 한국연구재단의 지원에 (No.NRF-2017R1D1A1B04035633) 의하여 이루어진 연구로서, 관계부처에 감사드립니다.

#### 참고문헌

[1] IJ Goodfellow, D Erhan et. al, "Challenges in representation learning: A report on three machine learning contests," International Conference on Neural Information Processing, pp. 117-124, 2013

[2] Psychology, School of Natural Sciences University of Stirling, 2D face sets, Available: <http://pics.psych.stir.ac.uk>

[3] A Dhall, R Goecke, S Ghosh et. al, "From individual to group-level emotion recognition: Emotiw 5.0," Proceedings of the 19th ACM International Conference on Multimodal Interaction, pp. 524-528, 2017

[4] A Dhall, OV Ramana Murthy et. al, "Video and image based emotion recognition challenges in the wild: Emotiw 2015," Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, pp. 423-426, 2015

[5] B Fasel and J Luetttin, "Automatic facial expression analysis: a survey," *Pattern recognition*, Vol. 36, No. 1, pp. 259-275, 2003

[6] C Shan, S Gong and P. W McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study, " *Image and Vision Computing*, pp.803-816, 2009

[7] J Whitehill and C. W Omlin, "Haar features for faces au recognition," 7th International Conference on Automatic Face and Gesture Recognition, 2006

[8] M Liu, S Li, S Shan, X Chen, "AU-aware Deep Networks for Facial Expression Recognition," 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), 2013

[9] Z Zeng, M Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions, " *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp.39-58, 2009

[10] B.-K. Kim, J Roh, S.-Y. Dong, and S.-Y. Lee, "Hierarchical committee of deep convolutional neural networks for robust facial expression recognition," *Journal on Multimodal User Interfaces*, pp.1-17, 2016

[11] G Levi and T Hassner, "Emotion recognition in the wild via convolutional neural networks and mapped binary patterns," In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, pp. 503-510

[12] A Mollahosseini, D Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," IEEE Winter Conference on Applications of Computer Vision (WACV), pp.1-10, 2016

[13] P Viola and M Jones, "Rapid object detection using a boosted cascade of simple features," In Proceedings of the IEEE Computer Society Conference, volume 1, pp. 1-1, 2001

[14] P Viola and M Jones, "Robust Real-Time Face Detection, " *International Journal of Computer Vision*, Vol. 57, No. 2, pp.137-154, 2004

- [15] Y LeCun, L Bottou, Y Bengio and P Haffner, "Gradient Based Learning Applied to Document Recognition," *In Proceeding of the IEEE*, pp.2278-2324, 1998
- [16] F Wang, L Chen et. al, "The devil of face recognition is in the noise," The European Conference on Computer Vision (ECCV), pp.765-780, 2018
- [17] Y Zheng, DK Pal, M Savvides, "Ring Loss: Convex Feature Normalization for Face Recognition," The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.5089-5097, 2018
- [18] Wonjae Choi, Taehwan Kim, Hyunseok Bae, Jungsun Kim, "Comparison of Emotion Prediction Performance of CNN Architectures," Proc. of the KIISE Conference, pp. 1,029-1,031, 2019
- [19] Hai-Duong Nguyen et. al, "Facial Expression Recognition Based on Deep Learning," Proc. of the KIISE Conference, pp. 873-875, 2018
- [20] Daeok Kim, Jongkwang Hong, Hyeran Byun, "Face Recognition Based on Facial Landmark Feature Descriptor in Unconstrained Environments," *Journal of KIISE*, Vol. 41, No. 9, pp. 666-673, 2014
- [21] Sae-han-seul Yi, Youngsang Woo et. al, "Efficient Local Binary Pattern Based Face Recognition Using OpenCL on the Embedded GPU," *Journal of KIISE*, Vol. 40, No. 6, pp. 257-265, 2013
- [22] Jae Hun Bang, Sungyoung Lee, "Adaptive Speech Emotion Recognition Framework Using Prompted Labeling Technique," *Journal of KIISE*, Vol. 21, No. 2, pp. 160-165, 2015



**기민송(Minsong Ki)**

2014년 : 덕성여자대학교 컴퓨터과학과 (학사)

2014년~현 재: 연세대학교 컴퓨터과학과 석박통합과정

※ 관심분야 : 인공지능(AI), 객체검출(Object detection), 얼굴인식(Face recognition), 지능형 자동차(IV) 등



**최영우(Yeongwoo Choi)**

1985년 : 연세대학교 전자공학과(학사)

1986년 : University of Southern California 컴퓨터공학과(석사)

1994년 : University of Southern California 컴퓨터공학과(박사)

1994년~1997년: LG전자기술원 선임연구원

1997년~현 재: 숙명여자대학교 컴퓨터과학과 교수

※ 관심분야 : 인공지능(AI), 머신러닝(ML), 영상처리(IP), 지능형자동차(IV) 등