

딥러닝을 이용한 연속 시간 감정 상태 추론 시스템 설계

Design of Emotional State Estimation System in Continuous Time Using Deep Learning

저자 (Authors)	심희린, 심귀보 Heereen Shim, Kwee-Bo Sim
출처 (Source)	한국지능시스템학회 논문지 29(1) , 2019.2, 76-81(6 pages) Journal of Korean Institute of Intelligent Systems 29(1) , 2019.2, 76-81(6 pages)
발행처 (Publisher)	한국지능시스템학회 Korean Institute of Intelligent Systems
URL	http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07617430
APA Style	심희린, 심귀보 (2019). 딥러닝을 이용한 연속 시간 감정 상태 추론 시스템 설계. 한국지능시스템학회 논문지, 29(1), 76-81
이용정보 (Accessed)	연세대학교 165.***.14.104 2019/11/23 22:34 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.



딥러닝을 이용한 연속 시간 감정 상태 추론 시스템 설계

Design of Emotional State Estimation System in Continuous Time Using Deep Learning

심희린·심귀보[†]

Heereen Shim and Kwee-Bo Sim[†]

중앙대학교 전자전기공학부

Department of Electrical and Electronics Engineering, Chung-Ang University

요약

표정은 사람이 감정을 표현하는 직접적인 방법 중 하나이다. 사람은 연속시간 상에서 표정의 변화를 통해 감정 상태를 나타내지만 기존의 기술은 대부분 정지영상에서만 감정 상태를 인식하였다. 본 연구는 연속 시간상에서 감정 상태를 추론하기 위한 딥러닝 시스템을 제안한다. 제안된 연속 시간 감정 상태 추론 시스템은 세 단계로 이루어져 있다. Convolutional Neural Network (CNN) 모델을 사용한 특징 추출 단계, 활성화 지도를 사용한 얼굴영역 검출 및 관심 영역 풀링 단계, 그리고 Recurrent Neural Network (RNN) 모델을 사용한 시계열 추론 단계. 본 논문에서는 짧은 표정 동영상상을 통해 감정 상태를 추론하는 실험을 수행하였다. 실험을 통해 제안된 시스템이 기존의 CNN 기반 물체 검출 모델보다 빠르게 얼굴영역 검출을 수행하며, 연속 시간상에서 감정 상태 추론이 가능함을 보였다.

키워드 : 감정 상태 판단, 표정 인식, Convolutional neural network, neural networks.

Abstract

Facial expressions are one of the efficient methods of expressing emotions. Most of the existing facial expression analysis techniques recognize emotional state only in still images. This study proposes a deep learning system for estimating emotional states in continuous time. The proposed emotional state estimation system in continuous time consists of three processes. Feature extraction using Convolutional Neural Network (CNN) model, Activation Map (AM)-based facial region detection and region of interest (RoI) Pooling, and sequential estimation using Recurrent Neural Network (RNN) model. In this paper, we performed an experiment to estimate the emotional state from in short video clips of facial expressions. Experimental results show that the proposed system can perform facial region detection faster than conventional CNN-based object detection models and emotional state estimation in continuous time.

Key Words : Emotional state estimation, Facial Expression Recognition, Convolutional neural network, Recurrent neural network.

Received: May, 24, 2018

Revised: Jun, 29, 2018

Accepted: Jun, 30, 2018

[†]Corresponding authors

kbsim@cau.ac.kr

1. 서론

감정 표현은 빠르고 직관적인 의사소통 방식이다. 다윈은 사람을 포함한 몇몇의 포유류는 생존과 본능에 관련된 감정을 안면 근육의 움직임을 통해 전달하도록 진화했다고 주장한다[1]. 에크만은 실험을 통해 문화와 언어에 관계없이 모든 인류는 특정한 표정으로부터 보편적인 감정을 인식 할 수 있음을 밝혔다[2]. 또한 연구를 통해 감정 표현은 사람 간 사회적 관계 형성에 중요한 역할을 함이 밝혀졌다[3].

사람로봇 상호작용 (Human-Robot Interaction, HRI)은 다양한 의사소통 채널을 통한 상호작용을 연구하는 분야이다. HRI의 목표는 사람과 로봇사이 인지적 또는 정서적 상호작용을 가능하게 하는 것이다. 이를 위해서는 로봇이 사람을 인식하고 적절하게 반응할 수 있어야 한다. 감정 인식 기술은 로봇이 사람의 감정 상태를 인식할 수 있도록 하는 기술이다. 감정 인식 기술은 사람과 로봇사이의 상호작용의 질을 향상시킬 수 있다 [4]. 이러한 사실을 기반으로, 사람의 표정에서 감정 상태를 추론하는 시스템을 설계하여 HRI 기술을 향상시키고자 본 연구가 진행되었다.

본 연구는 동영상을 통한 감정 상태 추론 시스템 개발을 목표로 한다. 본 논문의 구성은 다음과 같다. 제 2장에서는 제안 시스템을 구성하는 기술과 관련된 배경 이론을 소개한다. 제 3장에서 본 논문의

본 논문은 본 학회 2018 춘계학술대회에서 선정된 우수논문입니다.

본 논문은 한국연구재단 중견연구지원사업(No. 2012-000872)에서 지원하여 연구하였으며 연구비 지원에 감사드립니다.

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.

제안 시스템의 세부 구성 요소를 기술하고, 제 4장에서 제안 기술의 우수성을 실험 결과를 통해 입증한다. 마지막 5장에서는 결론 및 향후 진행 예정인 연구 방향에 대하여 논의한다.

2 배경 이론

2.1 Convolutional Neural Network

인공 신경망(Artificial Neural Network, ANN)은 인간의 뇌가 감각기관을 통해 입력 받은 정보를 토대로 정보를 인식하는 과정을 모사한 계산 모델이다. ANN은 분류 및 회귀 문제에 적합한 특성을 가지고 있다. 최근 ANN을 깊은 구조로 쌓은 심층 신경망(Deep Neural Network, DNN)이 영상인식, 음성인식, 자연어처리와 같은 다양한 분야에서 뛰어난 성능을 보이고 있다. 이중 컨볼루션 신경망(Convolutional Neural Network, CNN)은 격자 모양의 위상을 가지는 데이터를 처리하기 위한 특수한 종류의 ANN이다. 초기의 CNN은 손글씨 인식을 위해 제안되었으나 [5], 현재는 다양한 분야의 영상 인식에 널리 사용되고 있다 [6].

CNN 모델은 컨볼루션 레이어와 풀링 레이어로 이루어져 있으며, 이를 통해 데이터로부터 공간적인 특징을 추출할 수 있다. 모든 입력 뉴런들이 모든 출력 뉴런들과 연결되어 있는 ANN 모델과 달리, CNN 모델의 컨볼루션 레이어에서 출력 뉴런은 인접한 입력 뉴런하고만 연결되어 있다. 이러한 특징을 **희소 연결**이라 하며, 이를 통해 CNN은 입력 데이터의 공간적 특징을 유지하고, 비교적 적은 개수의 매개변수를 필요로 하게 된다. 기본적인 CNN 모델에서 컨볼루션 레이어 이후에는 비선형 활성화 함수가 적용되며 풀링 레이어가 뒤 따른다. 풀링 레이어는 레이어의 총 출력을 요약하는 기능을 하며 다운 샘플링으로 볼 수 있다. 이러한 풀링 레이어 또한 CNN 모델에 **노이즈** 및 **변화에 강인한** 특성을 부여하게 된다. 이러한 특성으로 인해 CNN 모델은 **이미지 데이터로부터 유의미한 특징을 추출하는 능력**을 가진다. 본 논문에서는 이러한 특성을 활용하여 표정기반 감정 상태 추론을 위해 CNN 모델을 사용하여 입력 영상으로부터 특징을 추출하였다.

2.2 Weakly-supervised object localization

CNN은 특징 추출에 뛰어나기 때문에 이미지 분류에서 높은 성능을 보이지만 분류하고자 하는 해당 물체의 위치를 검출하는 능력은 없다. 이는 CNN 모델에서 분류기 역할을 하는 최종 완전 연결 (fully-connected) 레이어 때문이다. 앞서 살펴본 것과 같이 CNN 모델은 다수의 컨볼루션 레이어와 풀링 레이어를 통해 입력 데이터로부터 특징 텐서를 출력한다. 이렇게 추출된 특징 텐서는 **3D 형태**를 가지며 공간적인 특성을 유지하고 있는 상태이지만, 최종 완전 연결 레이어의 입력으로 들어가는 과정에서 **2D 형태의 벡터**로 변형된다. 이 과정을

통해 특징 텐서는 공간적 정보를 잃게 된다.

최근 제안된 **weakly-supervised object localization** 방식은 컨볼루션 레이어에서 출력된 특징 텐서를 활용하여 **물체의 위치를 추정**한다 [7]. 분류 과정에서 입력 데이터의 어느 영역이 주로 활성화 되었는지를 기준으로 영상 내에 해당 물체가 위치함을 추정할 수 있을 것이라는 가정이다. 이때 컨볼루션 레이어에서 출력된 특징 텐서는 여러 개의 채널을 가진 특징 지도로 생각할 수 있다. 따라서 주어진 특징 지도를 이용하면 CNN의 최종 완전 연결 레이어에서 입력 영상의 어떠한 영역이 주로 영향을 끼쳤는지 해석이 가능하다. 본 논문에서는 특징 지도를 이용하여 영상 내 물체 위치 추정 방식을 응용하여 **입력 영상 내 얼굴 영역 검출을 수행**하였다.

3. 연속 시간 감정 상태 추론 시스템

본 연구에서는 딥러닝을 사용한 연속 시간 감정 상태 추론 시스템을 제안한다. 제안한 연속 시간 감정 상태 추론 시스템은 세 단계로 이루어져 있다. 첫 번째 단계에서는 CNN 모델을 사용하여 입력 영상에서 표정 분석을 위한 특징을 추출한다. 두 번째 단계에서는 추출된 특징 텐서의 활성화영역을 RoI (Region of Interest)로 검출한다. 마지막 단계에서는 특징 텐서 중 두 번째 단계를 통해 얻은 특징 텐서의 활성화 부분을 RNN 모델의 입력으로 받아 시계열 추론을 진행한다. 제안 시스템의 개요는 그림 1과 같다.

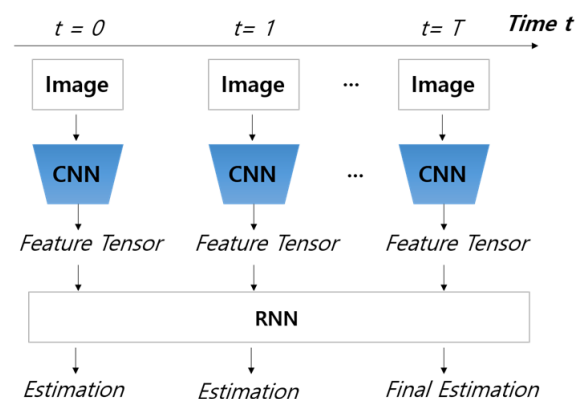


그림 1. 연속 시간 감정 상태 추론 개요

Fig. 1. Overview of the sequential emotional state estimation

3.1 특징 추출

본 논문에서는 CNN 모델을 사용하여 입력 영상으로부터 특징 텐서를 추출한다. 제안 시스템의 CNN 모델은 표정인식 분류를 위해 학습된 모델을 사용하지만 **실제 분류는 진행하지 않으며 최종 완전 연결 레이어 이전에서 특징 텐서를 출력하도록 설계**되었다.

3.2 얼굴 영역 검출 및 관심영역 풀링

연속 시간 상 감정 상태 추론을 위해 입력 영상으로부터 얼굴영역을 찾고 특징맵에서 해당하는 영역의 부분만 시계열 모델의 입력으로 사용한다. 본 논문에서는 활성화 지도를 사용하여 얼굴 영역의 위치를 검출하였다. 추가의 얼굴 탐지 모델을 사용하지 않고, CNN 모델을 사용한 특징 추출 과정에서 얻은 특징 텐서를 재사용 한다는 점과 추출된 특징 텐서에서 얼굴 영역에 해당하는 부분만을 사용한다는 것이 핵심이다. 제안하는 얼굴영역 검출 방법의 구성도는 그림 2와 같다.

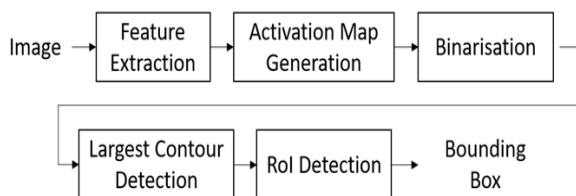


그림 2. 활성화 지도 기반 얼굴영역 검출 프로세스

Fig. 2. AM-based facial region detection process

제안된 얼굴영역 검출 프로세스는 활성화 지도의 어느 영역이 주로 활성화 되었는가를 참조하여 얼굴 영역을 검출한다. 활성화 부분 중 가장 큰 영역을 얼굴이 위치하고 있는 영역이라 가정하며 해당 영역의 경계 상자 (bounding box)를 관심 영역 (Region of Interest, RoI)으로 지정한다. 표정 분류에 중요한 영향을 끼치는 부분은 표정 마다 상이할 수 있으나, 영향을 끼칠 수 있는 모든 영역은 얼굴 안에 위치하기 때문이다.

얼굴영역 검출 프로세스의 관심영역 검출은 영상 처리 알고리즘을 사용하여 수행한다[8]. 활성화 부분의 가장 큰 영역을 찾기 위해 이진화, 최대 윤곽선 찾기, 관심 영역 검출을 순차적으로 진행한다. 각 단계 별 결과 예시 및 입력 이미지와 결과 이미지는 그림 3에 나타나 있다.

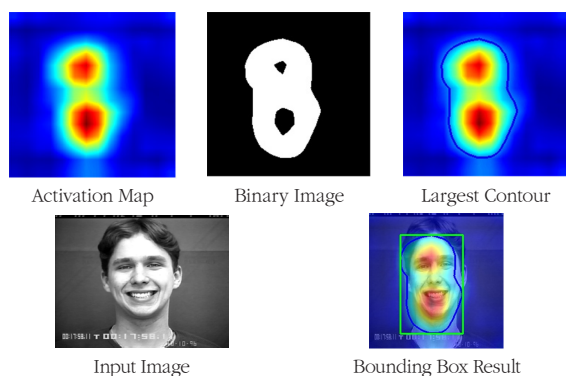


그림 3. 활성화 지도 기반 얼굴 영역 검출 프로세스 결과

Fig. 3. Results of AM-based region detection process

검출된 얼굴영역 정보를 사용하여 추출된 특징 텐서에서 얼굴영역에

해당하는 관심영역 부분만을 선택한다. 이렇게 선택된 관심영역 풀링 결과물은 RNN 모델의 입력으로 사용된다.

3.3 연속 시간 상 감정 상태 추론

제안 시스템의 RNN 모델은 각 시각 스텝에서 추출된 특징 텐서를 입력으로 받아 감정 상태를 추론한다. RNN 모델의 hidden Layer의 unit 들은 이전 시각 스텝의 정보를 모두 함축하고 있으며, 현재 시각 스텝의 출력에도 영향을 끼친다. 제안 시스템의 RNN 모델은 각 시각 스텝에서 감정 상태를 추론하고, 마지막 타임 스텝의 결과를 최종 결과로 삼는다.

4. 실험 및 결과

4.1 데이터베이스

본 실험에서는 Extended Cohn-Kanade (CK+) Facial Expression 벤치마크 데이터베이스를 사용하였다 [9]. CK+ 데이터베이스는 210 명의 피실험자로 이루어진 592개의 이미지 시퀀스로 이루어져있다. 본 실험에서는 총 309개의 이미지 시퀀스가 사용되었으며, 학습에는 전체 데이터 셋의 80%가 사용되었고, 나머지 20% 는 테스트용으로 사용되었다. 각 이미지 시퀀스는 그림 4. (a)와 같이 중립 상태에서 시작하여 특정 표정이 가장 강하게 표현된 시점에서 끝난다. 각 시퀀스에서 표현된 기본 6정서 중 하나에 해당한다. 예시 이미지는 그림 4. (b)와 같다.



그림 4. CK+ 데이터베이스

(a) 이미지 시퀀스 예시 (b) 기본 6 정서 표정 예시

Fig. 4. CK+ data base (a) Examples of image sequences (b) Examples of facial expressions of basic 6 emotions

4.2 학습 데이터 구축

본 실험에서는 CNN 모델을 학습을 위해 데이터에 변화를 주는 **데이터 증대 작업을 진행**하였다. 데이터 증대는 학습용 데이터에만 적용하여 총 학습데이터 양을 증가하였다. 변형된 이미지는 그림 5와 같다.

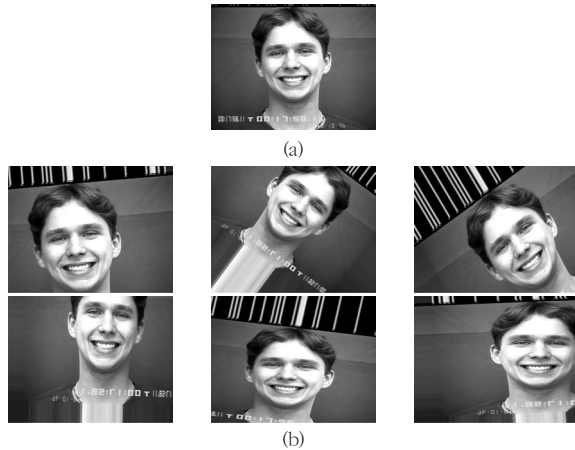


그림 5. 데이터 증대 작업 예시 (a) 원본 이미지 (b) 변형된 이미지

Fig. 5. Examples of data augmentation
(a) original image (b) transformed images

실험에 사용된 CK+ 데이터베이스는 얼굴 영역에 대한 정보를 제공하지 않는다. 따라서 본 실험에서는 다음과 같이 얼굴영역을 정의하였다.

규칙 1: 턱 끝부터 이마선 위까지로 지정한다.

규칙 2: 귀는 얼굴 영역에 포함되지 않는다.

규칙 3: 위 두 규칙을 따르면서 최소 여백을 가지는 사각형을 얼굴 영역으로 지정한다.

위 규칙을 따라 지정된 얼굴영역은 그림 6과 같다.



그림 6. 얼굴영역 예시

Fig. 6. (a) Example of facial region

4.3 실험 환경

실험 순서는 다음과 같다. CNN 모델의 표정 분류 학습을 진행한다. 그 다음으로 활성화 지도를 사용한 영역 검출 테스트의 정밀도와 속도를 확인한다. 마지막으로, RNN 모델 학습을 후 시계열 추론 정확도를 확인한다.

CNN 모델은 동영상의 마지막 프레임만 사용하여 표정 분류를 학습한다. 본 실험에서는 VGG-16 네트워크와 Inception v3, 네트워크를 기본 구조로 사용하였고 최종 fully-connected 레이어 이전 max pooling

레이어는 global average pooling 레이어로 변경되었다. Object function으로는 식 (1)을 사용하였다. CNN 모델 학습을 위해 사용한 하이퍼 파라미터 값을 표 1에 나타내었다.

$$L_i = \sum_j t_{i,j} \log(p_{i,j}) \quad (1)$$

, p are the predictions, t are the targets, I denotes the data point and j denotes the class

표 1. CNN 학습을 위한 하이퍼 파라미터들

Table 1. Hyper parameters for training CNN

Architecture	Parameter	Value
VGG-16	Learning Rate	0.0001
	Learning	Gradient Descent
Inception v.3	Learning Rate	0.0001
	Learning	Gradient Descent

활성화 지도를 사용한 영역 검출 테스트는 앞서 학습된 CNN 모델을 사용하여 정밀도를 계산한다. 정확도는 식 (2)과 같으며, Intersection of Union (IoU)이 0.5 이상인 검출 결과를 true positive로 사용하고, 그 이하의 검출 결과는 false positive로 사용한다. IoU 값 계산 방법은 식 (3)에 나타나 있다.

$$Precision = \frac{true\ positive}{true\ positive + false\ positive} \quad (2)$$

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \quad (3)$$

얼굴 영역 검출 속도는 알고리즘 수행에 걸리는 연산 시간을 기준으로 한다. 이때의 알고리즘 수행시간은 입력 이미지로부터 CNN 모델을 사용하여 특징 텐서를 추출하고 RoI 영역 추정에 소요되는 시간의 총 합이다.

RNN 모델을 사용한 시계열 추론은 이미지 시퀀스 전체를 입력으로 받고 마지막 프레임에서의 출력을 기준으로 학습을 진행한다. 학습을 위한 object function은 위의 CNN 모델 학습과 동일하며 그 외 하이퍼 파라미터는 표 2와 같다.

표 2. RNN 학습을 위한 하이퍼 파라미터들

Table 2. Hyper parameters for training RNN

Architecture	Parameter	Value
LSTM	Learning Rate	0.0001
	Hidden units	32
	Learning	RMSProp
GRU	Learning Rate	0.0001
	Hidden units	32
	Learning	RMSProp

학습된 RNN 모델은 매 순간 입력 프레임에 대해 추론 결과를

출력하고, 시퀀스 마지막 출력 결과를 기준으로 식 (4)을 사용하여 시계열 추론 정확도를 계산한다.

$$Accuracy = \frac{\# \text{ of correct predictions}}{\# \text{ of test samples}} \quad (4)$$

본 실험에 사용된 장비의 상세 정보는 다음과 같다.

CPU: AMD Ryzen 7 1800X

GPU: NVIDIA 1080 Ti with 11.2 GB memory

OS: Ubuntu 16.04 LTS

Parallel computing: CUDA v8.0

Framework: Keras + Tensorflow (backend)

4.4 실험 결과

그림 7은 CNN 모델의 학습 모니터링 결과이다. 두 모델 모두 학습이 진행됨에 따라 오류가 감소하는 추세를 확인할 수 있다. 단, Inception v.3 모델의 경우 VGG-16 모델보다 깊고 학습할 파라미터의 개수가 월등히 많기 때문에 수렴하기 까지 더 많은 시간이 소요됨을 확인할 수 있다.

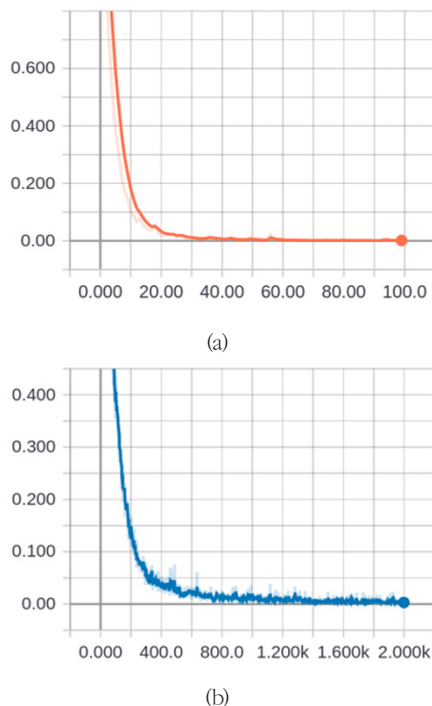


그림 7. Cross-entropy loss (a) VGG-16 (b) Inception v.3

Fig. 7. Cross-entropy loss (a) VGG-16 (b) Inception v.3

표 3는 활성화 지도를 사용한 얼굴 영역 검출 속도 비교 결과이다. 결과에서와 같이 제안 방식이 다른 CNN 기반의 검출 모델보다 빠른 검출 속도를 보임을 알 수 있다. 또한 제안된 방식은 학습과정에서 얼굴영역에 대한 정보가 필요하지 않기 때문에 학습 데이터 구축에

표 3. 다른 검출 모델과 비교

Table 3. Comparison with other detection models

Model	Accuracy (%)	Speed (fps)
Faster R-CNN [10]	72.9	5
YOLO v.2 [11]	98.4	21
Proposed	98.2	28

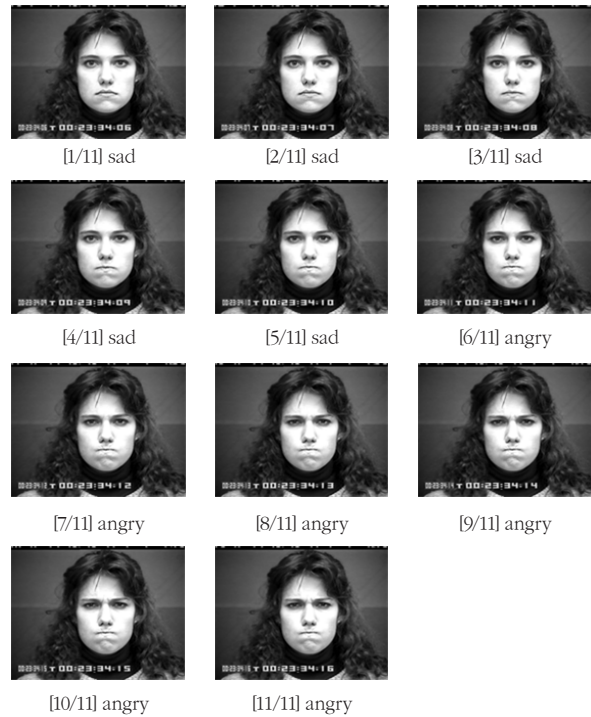


그림 8. 연속시간 감정 상태 추론 결과 예시

Fig. 8. Examples of sequential emotional state estimation.

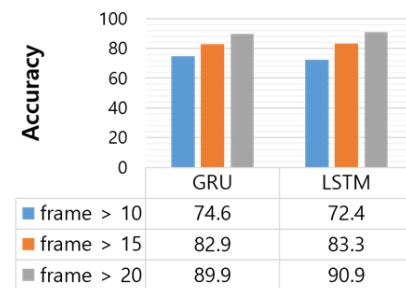


그림 9. 시계열 추론 결과

Fig. 9. Result of sequential estimation

소요되는 시간과 노동력을 절감할 수 있다는 이점이 있다.

이미지 시퀀스에서 표정을 통해 감정 상태를 추론한 결과는 그림 8과 같다. 제시된 이미지 시퀀스의 정답 라벨은 화남이며 마지막 프레임에서의 추론 결과 또한 화남이다. 결과를 통해 제안된 시스템이 시간의 흐름에 따라 감정 상태 추론 결과를 업데이트 해나감을 확인할 수 있다.

시계열 추론 정확도는 그림 9에 표시되어 있다. 실험에 사용된 두

종류의 모델 모두 영상의 프레임 길이가 길수록 정확도가 향상되는 추세를 보인다. 이는 프레임이 길수록 감정 상태를 추론하기 위한 표정 영상의 정보가 늘어나기 때문으로 보여 진다.

5. 결론 및 향후 연구

본 연구에서는 연속 시간에서 감정 상태 추론을 위한 딥러닝 시스템을 제안하였다. 실험을 통해 제안된 시스템이 기존의 CNN 기반 물체 검출 모델보다 빠르게 얼굴영역 검출을 수행하며, 연속 시간상에서 감정 상태 추론이 가능함을 보였다. 제안된 시스템에서 CNN과 RNN 모델의 학습은 따로 이루어 졌으며, 추후 **end-to-end training에 대한 연구가 진행될 예정이다**. 이를 통하여 성능을 향상시킬 수 있을 것으로 예상된다.

References

- [1] C. Darwin, and P. Prodger, "The expression of the emotions in man and animals," Oxford University Press, USA, 1998.
- [2] P. Ekman, and H. Oster, "Facial expressions of emotion," Annual review of psychology, vol.30, no.1, pp. 527-554, Feb. 1979.
- [3] F. Foroni, and G. R. Semin, "Language that puts you in touch with your bodily feelings the multimodal responsiveness of affective expressions," Psychological Science, vol. 20, no. 8, pp. 974-980, 2009.
- [4] P. Rani and N. Sarkar, "Emotion-sensitive robots - a new paradigm for human robot interaction," in Proc. 4th Intl. Conf. IEEE/RAS Humanoid Robots, Santa Monica, CA, USA, vol. 1, pp. 149-167. 2004.
- [5] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp.2278-2324, 1998.
- [6] I. H. Jang, K. E. Ko and K. H. Cho, "Automated System for Shoe-upper Pattern Recognition using Deep Neural Network," *Journal of Institute of Control, Robotics and Systems*, vol. 23, no. 9, pp. 705-710, 2017.
- [7] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," *Computer Vision and Pattern Recognition (CVPR)*, pp. 2921-2929, 2016.
- [8] H. R. Shim, K. E. Ko, K. H. Cho, I. H. Jang and K. B. Sim, "Real-time Shoe-pattern Analysis System using Deep Neural Network," *Journal of Korean Institute of Intelligent Systems*, vol. 28, no. 1, pp.20-27, 2018.
- [9] P. Lucey et al., "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in Proc. CVPRW, IEEE Conf. Computer Society, San Francisco, CA, USA, pp. 94-101, 2010.
- [10] S Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, pp.91-99, 2015.
- [11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779-788, 2016.

저자 소개



심희린(Heereen Shim)

2016년 : 중앙대학교 전자전기공학부 공학사

2018년 : 중앙대학교 대학원 전자전기공학과
공학석사

관심분야 : 인공지능, 기계학습, 감정인식, 소셜 로봇 등.

Phone : +82-2-820-5319

E-mail : abdeep@cau.ac.kr



심귀보(Kwee-Bo Sim)

1990년 : The University of Tokyo 전자공학과
공학박사

1991년~현재 : 중앙대학교 전자전기공학부 교수

2006년~2007년 : 한국지능시스템학회 회장

관심분야 : 인공생명, 뇌-컴퓨터 인터페이스, 감정인식, 의도인식, 유비쿼터스 지능형로봇, 지능시스템, 컴퓨테이셔널 인텔리전스, 유비쿼터스 컴퓨팅 및 센서 네트워크, 소프트웨어 컴퓨팅(신경망, 퍼지, 진화연산), 인공지능시스템, 지능형 감시시스템, 사물인터넷(IoT), 빅데이터, 딥러닝, 패턴 인식, 지능형 센서 등.

Phone : +82-2-820-5319

E-mail : kbsim@cau.ac.kr

Homepage URL : http://alife.cau.ac.kr