

누설 ReLU를 적용한 심층학습 기반 얼굴 표정 인식 성능 향상

신영미*, 장수지*, 박장식*

* 경성대학교 전자공학과

e-mail : sym0020@ks.ac.kr, neco57@ks.ac.kr, jsipark@ks.ac.kr

Performance Improvement of Facial Expression Recognition Based on Deep Learning with Leaky ReLU

Young-Mi Shin*, Su-Ji Jang*, Jangsik Park*

*Dept of Electronic Engineering, Kyungsung University

요 약

본 논문에서는 심층학습의 활성화함수에 따라서 얼굴 표정 인식 성능에 대하여 분석한다. 심층학습의 활성화함수로 ReLU를 널리 사용되고 있지만, 응용에 따라서 활성화함수의 영향을 받는다. 본 논문에서는 얼굴 표정 인식에서 ReLU와 누설 ReLU를 적용하여 비교 분석한다. 다양한 얼굴 표정 인식 데이터셋을 활용한 시뮬레이션 결과 누설 ReLU를 적용한 결과가 약 3% 정도 개선되는 것을 확인한다.

1. 서론

얼굴 표정 인식 (FER, Facial Expression Recognition)[1,2]는 얼굴인식과 달리 각 개인 간의 차이보다는 모든 사람들의 공통된 표정변화의 차이를 인식하는 것을 목적으로 하는데, 광범위한 응용 분야에 적용되고 있다. 사람의 얼굴 표정은 주로 입, 눈, 눈썹 등의 조합적인 형태 변화로 나타내며, 표정의 종류도 각 지역이나 문화에 따라 다양하나 6가지 표정(기쁨, 슬픔, 분노, 혐오, 공포, 놀람)이 전 세계 문화에 동일하게 인지되는 것으로 알려지고 있다[3]. FER은 다양한 머리 자세와 조명 상태, 해상도 등 다양한 원인으로 인한 인식률 저하에 관한 문제점을 다루고 있다[4-7]. FER을 위하여 최근 심층학습(deep learning) 모델이 적용되고 있으며, 활성화함수(activation function)로 ReLU 함수가 주로 활용되고 있다.

본 논문에서는 활성화함수에 따른 얼굴 표정 인식 성능을 비교 분석한다. ReLU가 종래의 시그모이드 함수(sigmoid function)에 비해서 우수한 성능을 보이지만, 응용에 따라서 활성화함수에 따른 성능이 변화가 있다. 활성화함수를 기존의 ReLU와 누설 ReLU(leaky ReLU) 함수를 각각 적용하여 인식률 성능을 비교 분석한다.

2. 심층학습 활성화함수

심층학습 대표적인 문제점은 기울기 소실(Gradient Vanishing)이다. 은닉층(Hidden Layer)이 적을 때에는 문제가 생기지 않지만 네트워크가 복잡할수록 이러한 기울기 소실 문제가 많이 발생한다. 기울기 소실이란 기울기가

사라지는 문제를 말하는 것이며, 이는 시그모이드(Sigmoid) 함수에서 기울기는 0에서 1사이의 값이 출력된다. 심층학습의 역전파(Back-propagation)를 하면서 여러 계층을 걸치면서 계속해서 기울기를 곱하게 되는데, 이는 기울기가 점점 작아져서 0에 수렴하게 되는 것이다. 관련 연구에 의하면, 은닉층이 9개 또는 10개 정도가 되면 네트워크가 잘 작동하지 않는다는 것을 알 수 있다. 식 (1)에서는 시그모이드 함수이다.

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

ReLU는 기울기 소실 문제를 어느 정도 개선하였다. ReLU 함수는 식 (2)로 ReLU는 기울기가 0 또는 1이면서 함수 자체도 아주 간단해서 시그모이드 함수를 계산하면서 걸리던 시간을 단축시킬 수 있다.

$$f(x) = \text{Max}(0, x) \quad (2)$$

그러나, ReLU가 음수들을 모두 0으로 처리하기 때문에 한번 음수가 나오면 더 이상 그 노드는 학습되지 않는다는 단점이 있다. 따라서 간단한 네트워크에 대해서는 오히려 좋지 않은 성능을 낼 수도 있다. 이러한 문제를 개선하기 위하여 식 (3)으로 표현되는 Leaky ReLU 함수가 활용되고 있다.

$$f(x) = \begin{cases} \alpha x, & x < 0, \\ x, & x \geq 0 \end{cases} \quad (3)$$

α : very small value constant

3. 실험 결과 및 검토

얼굴 표정 인식 분야에서 일반적으로 자주 사용되고 있는 3 가지 데이터셋, CK+, MMI, KDEF으로 활성화 함수에 따른 얼굴 표정 인식 성능을 비교 분석한다.

ReLU와 누설 ReLU를 활성화 함수(Activation Function)로 적용하여 인식을 비교한다. 그림 4와 같이 AlexNet은 256×256 차원의 입력을 사용하기 때문에 모든 학습 이미지의 크기는 정규화 한다. 학습의 세대(epoch), 배치 크기(batch size), 학습률(learning rate), 학습 감쇠율 등 다양한 매개 변수를 사용하여 AlexNet 모델로 시뮬레이션을 수행한다.

각 데이터셋은 6 가지 클래스로 분류한다. 6가지 클래스는 화난(Angry), 혐오(Disgust), 두려움(Fear), 행복(Happy), 슬픔(Sad), 놀람(Surprise)으로 구성되어 있다. 학습 과정에서 전체 Dataset의 60%를 사용하였으며, 유효성 검사에 20%, 테스트에서는 20%를 사용하였다.

CK+ 데이터셋은 Angry 300장, Disgust 395장, Fear 260장, Happy 476장, Sad 402장, Surprise 402장으로 구성되어 있고, 데이터셋의 영상은 계조 영상(gray scale)이다. MMI 데이터셋은 Angry 33장, Frown 37장, Sad 74장, Smile 70장, Surprise 53장으로 구성되어 있고, 색조 영상(color image)이다. KDEF 데이터셋은 Angry 300장, Disgust 140장, Fear 140장, Happy 140장, Sad 140장, Surprise 140장으로 구성하였고, 색조 영상이다.

표 1은 활성화 함수에 따른 각 데이터셋에 대한 인식을 비교한 것이다. ReLU 함수보다 Leaky ReLU 함수가 약 2.64%의 인식이 우수하다.

표 2 활성화 함수에 따른 데이터베이스 인식을 비교

	ReLU	Leaky ReLU
CK+	95.98%	97.98%
MMI	67.92%	72.00%
KDEF	81.44%	83.28%
Average	81.78%	84.42%

3. 심층학습 기반 얼굴표정 인식

얼굴 표정 인식을 위한 본 논문에서는 기본적인 심층학습 모델인 AlexNet 모델을 적용한다. AlexNet 모델은 그림 1과 같이 5개의 컨벌루션 계층, 3개의 풀링 계층, 3개의 완전연결 계층으로 구성되어 있다. 6가지 감정을 분류하는데 그 목적이 있으므로 완전연결 계층을 6가지 클래스로 설정한다.

4. 결론

본 논문에서는 AlexNet 심층학습 모델의 활성화 함수로

ReLU 와 누설 ReLU를 적용하여 얼굴 표정 인식 성능을 비교 분석하였다. ReLU 함수보다 Leaky ReLU 함수가 약 2.64% 정도의 성능 향상이 나타난 것을 확인할 수 있음을 알 수 있었다.

감사의 글

본 논문은 이 논문은 2019년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 지역신산업 선도인력양성사업(No. 2016H1D5A1910536)과 2018년 BB21+ 사업에 지원되었음

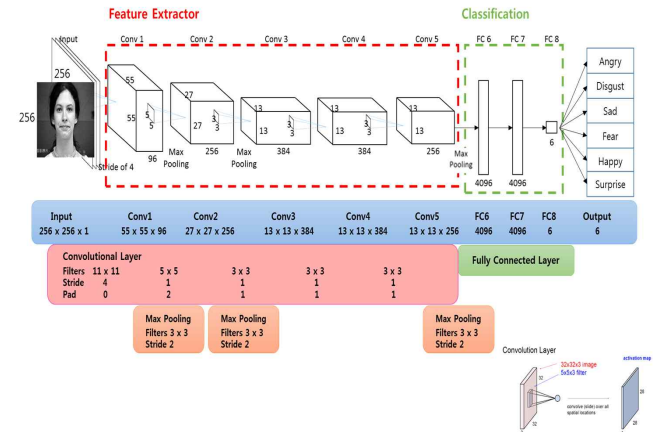


그림 1. AlexNet 모델 기반 얼굴 표정 인식 구성

참고문헌

- [1] Shan Li and Weihong Deng, "Deep Facial Expression Recognition: A Survey", arXiv: 1804.08348v2 [cs.CV] 22 Oct 2018.
- [2] I.Michael Revina, W.R. Sam Emmanuel, "A Survey on Human Face Expression Recognition Techniques", Journal of King Saud University - Computer and Information Sciences, 2018.
- [3] C. Darwin, et al., 1872, The Expression of Emotion in Man and Animals, John Murry, London.
- [4] Y. Kim, B. Yoo, Y. Kwak, C. Choi, and J. Kim. Deep generative-contrastive networks for facial expression recognition. arXiv preprint arXiv:1703.07140, 2017.
- [5] Z. Meng, P. Liu, J. Cai, S. Han, and Y. Tong. Identity-aware convolutional neural network for facial expression recognition. In Automatic Face & Gesture Recognition (FG 2017), 2017 12th IEEE International Conference on, pages 558 - 565. IEEE, 2017.
- [6] B. Martinez, M. F. Valstar, B. Jiang, and M. Pantic, "Automatic Analysis of Facial Actions: A Survey", IEEE. Transactions On Affective Computing, Vol. 13, No. 9, pp1-22, Issue. 99, 2017.
- [7] Huiyuan Yang, Umur Ciftci and Lijun Yin, "Facial Expression Recognition by De-expression Residue Learning", CVPR paper IEEE Xplore, pp.2168-2177, 2018.