

Prepoznavanje emocija korišćenjem dubokih neuronskih mreža

Seminarski rad u okviru kursa
Računarska inteligencija
Matematički fakultet

Denis Aličić
a.denis96@gmail.com

Sažetak

U ovom radu je predstavljena duboka neuronska mreža koja rešava problem klasifikacije emocija na osnovu fotografije. Ulaz u neuronsku mrežu predstavlja vektor realnih brojeva dužine 128 (eng. *face embedding*) [6] koji preslikava sliku lica u euklidski vektorski prostor. Zatim je nad tim vektorima trenirana mreža. Za implementaciju modela je korišćeno okruženje Keras [4], koji olakšava upotrebu moćne biblioteke Tensorflow [2].

1 Uvod

Prepoznavanje emocija predstavlja jedan od poznatih i još uvek u potpunosti nerešenih problema. Razvoj mašinskog učenja, konkretno oblasti konvolutivnih neuronskih mreža, doveo je do poboljšanja postojećih rešenja. Ovaj problem ima bitnu ulogu u oblasti međuljudskih odnosa [3], kao i odnosa čoveka i računara. Popularnosti ovog problema doprinelo je Kaggle [1] takmičenje otvoreno 2013. godine, čiji je nagradni fond iznosio 500 dolara. U takmičenju je učestvovalo 56 timova. Većina postojećih rešenja koristi treniranje konvolutivnih mreža izgrađenih od nekoliko slojeva. S obzirom da sam pročitao rad [6] koji govori o preslikavanju lica u euklidski vektorski prostor dimenzije 128, kao i da svaki takav vektor (eng. *face embedding*) jedinstveno određuje lice, odlučio sam da to iskoristim kao ulaz u neuronsku mrežu. Za treniranje i testiranje mreže je korišćen

skup podataka „fer2013” [5].

2 Opis skupa podataka

Skup podataka „fer2013” [5] sadrži sive (eng. *grayscale*) slike lica veličine 48x48 piksela. Cilj je klasifikovati svako lice u jednu od 7 kategorija:

0. Ljutnja (eng. *angry*)
1. Gađenje (eng. *disgust*)
2. Strah (eng. *fear*)
3. Sreća (eng. *happy*)
4. Tuga (eng. *sad*)
5. Iznenadenje (eng. *surprise*)
6. Neutralno (eng. *neutral*)

Skup sadrži 28709 slika u delu za trening i 3589 slika u delu za testiranje. Skupovi za trening i test su dati u jednom csv fajlu koji sadrži piksele slike, kategoriju i informaciju o tome u kom skupu se nalazi dati red.

Literatura

- [1] Challenges in Representation Learning: Facial Expression Recognition Challenge . <https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression>
Accessed: 2019-05-23.

- [2] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [3] Vinay Bettadapura. Face expression recognition and analysis: The state of the art. *CoRR*, abs/1203.6722, 2012.
- [4] François Chollet et al. Keras. <https://keras.io>, 2015.
- [5] Ian Goodfellow, Dumitru Erhan, Pierre-Luc Carrier, Aaron Courville, Mehdi Mirza, Ben Hamner, Will Cukierski, Yichuan Tang, David Thaler, Dong-Hyun Lee, Yingbo Zhou, Chetan Ramaiah, Fangxiang Feng, Ruifan Li, Xiaojie Wang, Dimitris Athanasakis, John Shawe-Taylor, Maxim Milakov, John Park, Radu Ionescu, Marius Popescu, Cristian Grozea, James Bergstra, Jingjing Xie, Lukasz Romaszko, Bing Xu, Zhang Chuang, and Yoshua Bengio. Challenges in representation learning: A report on three machine learning contests, 2013.
- [6] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. *CoRR*, abs/1503.03832, 2015.