CS 766 Project Proposal

Jiangang Chen(jiangang.chen@wisc.edu)

Introudction

Emotion recognition has attracted major attention in numerous fields because of its relevant applications in the contemporary world. It is possible to recognize emotion in several ways. Facial Expression Recognition(FER) is an important visual recognition technology to detect emotions given the input to the intelligent system in a facial image.

This project focuses on the state-of-art FER methods and reimplements FER algorithms (including Conventional and Deep Learning-based approaches). Two groups will be tested and compared on several datasets(In-the-Lab/ In-the-Wild). The project will also discover and contribute to the migration of two approaches from visible light to infrared images platform.

Motivation

FER in the human-computer interactive system is one of the challenging research topics in the field of artificial intelligence which has drawn plenty of attention in recent years. However, it is difficult to achieve natural and harmonious emotional interaction with traditional interaction methods such as keyboard, mouse, screen, and pattern, which is far from meeting the requirements for artificial intelligence.

Human expression is the most important carrier of inspirational perception and the most direct and obvious way of expressing emotions. Thus, FER has important theoretical significance for improving the emotional interaction ability of computers. Furthermore, facial expression is arguably the most natural, powerful and immediate signal to communicate emotional states and intentions.

However, automatic FER is still difficult in an unconstrained real-life situation with the widespread use of deep learning techniques. It encounters various challenges caused by occlusion, face pose variations, illumination changes, head motion, expressions ambiguity, and so on. An ideal automatic FER system is supposed to be able to tackle these challenges.

State-of-Art Approaches

In this part, we first categorize the existing FER methods into two main groups, ie, conventional approaches and deep learning-based approaches.

Most conventional approaches are based on engineered features, such as Histogram of Oriented Gradients(HOG), Local Binary Pattern Histogram(LBPH), and Gabor. Since facial expressions result from facial muscle movements or deformations, and HOG is very sensitive to object deformations, work [1] applies the HOG to encode these facial components as features. A linear SVM is then trained to perform the facial expression classification. Paper [2] proposes es-LBP (expression-specific LBP) which could better capture the local information of faces on important fiducial points. However, it has the disadvantage of being sensitive to noise.

Due to an increase in the ready availability of computational power and increasingly large training databases to work with, the machine learning technique of neural networks has seen a resurgence in popularity. Work [3] presents a new deep neural network architecture for FER which increased classification accuracy on both subject independent and cross-database evaluation scenarios. The addition of Inception layers increases the depth and width of the network while keeping the computational budget constant. Paper [4] instead of using the whole face region, three kinds of active regions are applied to classify facial expression through a decision-level fusion strategy. Work [5] proposes a deep convolutional fusion network, which addresses the FER task through discriminative spatial features learning and temporal dependencies accumulating.

Implementation

This project will reimplement the conventional[1] and deep learning-based[6] approaches to FER.

There are some drawbacks to concern. For example, acquiring images from an uncontrolled environment, where there may be movement, harsh lighting conditions, different poses, great distances, among other factors. This also involves the difference between the in-the-lab dataset and the in-the-wild dataset. In the meantime, considering the sensitive attributes such as gender, age, and ethnicity in the dataset, bias[7] may be a problem for FER. This project will try to address some concerns.

Evaluation

The evaluation includes FER accuracy comparison among two approaches groups on both in-the-lab and in-the-wild datasets. Results from other state-of-arts FER methods will be added to the comparison.

In-the-Lab: CK+; JAFFE; FER2013

In-the-Wild: AffectNet; FER+

Oulu-CASIA Near-Infrared and NVIE datasets may be considered for the infrared conditions.

The project will also examine some limitations for FER methods, such as head poses, illumination conditions, resolutions, occlusions, age, gender, ethnic background.

TimeLine

Feb 24: Project Proposal	Apr 7: Test on In-the-Lab datasets
Mar 10: Implement Conventional Approach	Apr 21: Test on In-the-Wild datasets
Mar 24: Implement DL Based Approach	May 5: Work and Test on Infrared datasets
Apr 5: Mid-Term Report / Website	May 5: Final presentation / Website

Reference

[1] Chen, Junkai, et al. "Facial expression recognition based on facial components detection and hog features." *International workshops on electrical and computer engineering subfields.* 2014.

[2] Chao, Wei-Lun, Jian-Jiun Ding, and Jun-Zuo Liu. "Facial expression recognition based on improved local binary pattern and class-regularized locality preserving projection." *Signal Processing* 117 (2015): 1-10.

[3] A. Mollahosseini, D. Chan and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2016, pp. 1-10, doi: 10.1109/WACV.2016.7477450.

[4] Sun, Ai, et al. "Facial expression recognition using optimized active regions." *Human-centric Computing and Information Sciences* 8.1 (2018): 1-24.

[5] Liang, Dandan, et al. "Deep convolutional BiLSTM fusion network for facial expression recognition." *The Visual Computer* 36.3 (2020): 499-508.

[6] Siqueira, Henrique, Sven Magg, and Stefan Wermter. "Efficient facial feature learning with wide ensemble-based convolutional neural networks." Proceedings of the AAAI conference on artificial intelligence. Vol. 34. No. 04. 2020.

[7] Xu, T., White, J., Kalkan, S., & Gunes, H. (2020). Investigating Bias and Fairness in Facial Expression Recognition.. *ECCV Workshops* (6), 12540 506-523. https://doi.org/10.17863/CAM.56933