

# Facial Expression Classification Proposal

Katherine Fu  
jfu57@wisc.edu

Lynette Gao  
qgao38@wisc.edu

Flossie Wan  
kwan9@wisc.edu

September 2020

## Abstract

Image processing and recognition are essential and widely used in different areas of studies. Classifying facial and emotional expression is of great interest in the deep learning field as a high-performance recognition model could lead to more complex classification tasks. Facial and emotional expression classification could be commonly applied in the real world. For example, businesses could perform facial recognition on users' selfies and analyze their rate of satisfaction on the products. More advanced applications could be the development of humanoid robots with the integration of facial and emotional recognition functionality. Thus, we need to have a better understanding of its basic implementation. In this project, we will utilize the image samples from the Real-world Affective Faces (RAF) Database. Its database contains 15339 facial images with 12271 training samples and 3068 testing samples and 37 landmark location labels. The images are grouped into 7 emotions including surprise, fear, disgust, happiness, sadness, anger, and neutral. We will implement multiple Convolutional Neural Network (CNN) architectures and evaluate their performance using accuracy metric.

## 1 Significance

Detecting facial expression promotes communication between individuals. For humans, recognizing facial expressions and emotions is a basic skill that is learned at an early age. We can look at a person's face and can quickly recognize the common emotions of anger, happiness, surprise, disgust, sadness, and

fear. However, to transfer this skill to a machine is a complex task. Meanwhile, there are many practical uses of the emotion detection, if the facial expression algorithm can be successfully implemented. For example, the US Department of Transportation claims that driving-related errors cause around 95 percent fatal road accidents [2]. In this case, facial emotion detection can find us find subtle changes in facial micro-expressions that precedes drowsiness and send personalized alerts to the driver. Therefore, we are interested in developing a deep learning algorithm to automatically understand human's emotion given an image.

## 2 Current work

Researchers have devoted decades of engineering time to writing computer programs that recognize a feature with accuracy. In the field of computer vision and machine learning, various facial expression recognition systems have been explored to encode expression information from facial representations. The majority of the traditional methods have used hand-crafted features or shallow learning, such as local binary pattern. Recently, with large amount of data available, deep learning has recently become a hot research topic and has achieved state-of-the-art performance. Some of the example algorithm includes Convolutional neural network (CNN), Deep belief network (DBN), Deep autoencoder (DAE) and Generative Adversarial Network (GAN) [6].

### 3 Planning

Oct 1 - Oct 15: Reading reference materials and start to build a benchmark model

Oct 16 - Nov 3: Implementing network and algorithms

Nov 3 - Dec 2: Training and improving models. Prepare for presentation and final write-up

## 4 Experiments

### 4.1 Image processing

Our image data have already been aligned based on the two-eye locations and the position of the mouths and resized to 100\*100 pixels. Since our samples include faces from different genders and races, it becomes important for us to take this information into consideration. Faces in some photos are pointing into different angles, which could result in various light and color intensity. Since the upright faces are of great interest, we will perform basic rotations to those images.

In some photos, people wear glasses. Those images would not be a big issue for our project since these glasses are transparent and we can still be able to detect the facial features.

It is also essential to include the images that show neutral emotion in our samples. We will use those images to improve the facial detection ability.

### 4.2 Network Training

There are many existing architectures and techniques that can be applied to facial expression recognition. We plan to implement multiple models using Convolutional Neural Network (CNN) architectures, including VGGNet [8], AlexNet [5], GoogLeNet [9] and ResNet [3]. CNN is one of the most widely-used technique for facial expression recognition, as it has been proved to be robust to face location and good at normalization [1]. All of the models we choose are competitors in ImageNet Large Scale Visual Recognition Challenge (ILSVRC), with different training methods and achievements [4][6]. We want to experiment on

the parameters of those models. By manipulating parameters like batch size, layer type, etc., we hope to achieve best performance of each model, and evaluate the advantages and shortcomings of our models.

### 4.3 Evaluation

To evaluate the performance of our model, we primarily plan to use their accuracy of recognizing facial expressions. Confusion matrix can also be used to evaluate whether our model is biased toward one or more classes (emotions, in our case). We can also detect over-fitting issue by evaluating validation metric, like loss. The turning point of loss function from going down to going up indicates over-fitting in our model.

## 5 Resources

Our project will be written in python with basic python libraries. We are using Real-world Affective Faces Database (RAF-DB) [7] for labeled facial expressions. All our models can be implemented with TensorFlow Keras API. TensorBoard will be used for analyze the performance of our model by evaluating loss and accuracy. All the training, testing and analyzing processes will be done on our personal laptops.

## References

- [1] Beat Fasel. Robust face analysis using convolutional neural networks. volume 2, pages 40 – 43 vol.2, 02 2002.
- [2] Shashank Gupta. Practical Use Cases of Facial Emotion Detection Using Artificial Intelligence - DZone AI, 2018.
- [3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015.
- [4] Asifullah Khan, Anabia Sohail, Umme Zahoora, and Aqsa Saeed Qureshi. A survey of the

recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, 53(8):5455–5516, Apr 2020.

- [5] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.
- [6] Shan Li and Weihong Deng. Deep facial expression recognition: A survey. *IEEE Transactions on Affective Computing*, page 1–1, 2020.
- [7] Shan Li, Weihong Deng, and JunPing Du. Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. In *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*, pages 2584–2593. IEEE, 2017.
- [8] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2014.
- [9] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions, 2014.  
<https://dzone.com/articles/practical-use-cases-of-facial-emotion-detection-us>  
<https://dl.acm.org/doi/pdf/10.1145/3357384.3357956>  
<https://medium.com/@abhinaya08/google-landmark-recognition-274aab3c71ae>