# MATLAB BASED FACIAL EXPRESSION RECOGNITION

**Jianyang Guo[1], Chunhui Xu[1]**

[1]Southern University of Science and Technology, Shenzhen, China

## ABSTRACT

*This project realizes facial expression recognition in image recognition based on MATLAB. Through the input of facial expression pictures, to analyze the possibility of its being different expressions. This project is related to deep learning, neural networks and computer vision. Focusing on the advantages of MATLAB's efficient scientific computation, it realizes a simple facial expression recognition.*

Keywords: MATLAB, Facial expression recognition

## 1. INTRODUCTION

MATLAB is a business mathematics software that can be used in areas such as data analysis, deep learning, image processing and computer vision. Facial expression recognition, as a hot field in the field of image recognition, has been widely discussed and studied in recent years. However, most facial expression recognition is based on PyTorch or TensorFlow, two deep learning frameworks, and implemented in Python language. So this project wants to use MATLAB to realize facial expression recognition. Specifically, the user provides a picture of a person, and our system recognizes the facial expression of the person in the picture and gives a probability of predicting the expression. At the same time, we also provide a GUI interface to provide a better interactive experience for users.

## 2. MATERIALS AND METHODS

In this Project, we use CNN network and MATLAB toolboxes to implement a facial expression recognition system through deep learning technology.

### 2.1 Dataset

We use CK+ dataset which stands for Extended Cohn-Kanade dataset, is a dataset of 593 video sequences. The videos came from 123 different subjects, ranging in age from 18 to 50, with different genders and bloodlines. Each video shows a facial transition from a neutral expression to a directional extreme expression, recorded at 30 frames per second (FPS) at a resolution of 640 x 490 or 640 x 480 pixels. Of those videos, 327 were labeled in one of seven expression categories: anger, contempt, disgust, fear, happiness, sadness, and surprise. The CK+ dataset is widely considered to be the most widely used laboratory-controlled facial expression classification database and is used in major facial expression classification methods.

### 2.2 Network architecture

MobileNet is a lightweight CNN model proposed by Google, trained on the ImageNet dataset and designed to enable efficient image recognition and other computer vision tasks in resource-constrained environments such as mobile devices. By combining the power of deep learning with lightweight model design, MobileNet provides a high-performance and computationally efficient solution for mobile applications. MobileNetV2 is an evolutionary version of the MobileNet family of models, introduced in 2018. While maintaining the characteristics of lightweight and high efficiency, MobileNetV2 further improves the performance indicators of the model, such as accuracy and convergence speed, and is more suitable for practical application scenarios. The MobileNetV2 network architecture includes:

- Inverted Residuals
- Linear Bottlenecks

The following figure shows the MobileNetV2 network configuration diagram. Where t represents the multiplicity of 1×1 1 \times 11×1 convolution ascending dimension in the Inverted Residuals structure (compared to the input channel), and c is the depth channel of the output feature matrix. n indicates the number of times that a bottleneck (that is, an Inverted Residuals structure) repeats. s represents the step, but only the step of DW convolution in the first bottleneck. The stride that follows a bottleneck is equal to 1.[5]

| Input | Operator | $t$ | $c$ | $n$ | $s$ |
|---|---|---|---|---|---|
| $224^2 \times 3$ | conv2d | - | 32 | 1 | 2 |
| $112^2 \times 32$ | bottleneck | 1 | 16 | 1 | 1 |
| $112^2 \times 16$ | bottleneck | 6 | 24 | 2 | 2 |
| $56^2 \times 24$ | bottleneck | 6 | 32 | 3 | 2 |
| $28^2 \times 32$ | bottleneck | 6 | 64 | 4 | 2 |
| $14^2 \times 64$ | bottleneck | 6 | 96 | 3 | 1 |
| $14^2 \times 96$ | bottleneck | 6 | 160 | 3 | 2 |
| $7^2 \times 160$ | bottleneck | 6 | 320 | 1 | 1 |
| $7^2 \times 320$ | conv2d 1x1 | - | 1280 | 1 | 1 |
| $7^2 \times 1280$ | avgpool 7x7 | - | - | 1 | - |
| $1 \times 1 \times 1280$ | conv2d 1x1 | - | k | - | |

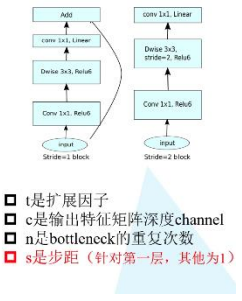


**FIGURE 1:** Architecture of the MobileNetV2 Network

### 2.3 Setup

The following are MATLAB toolboxes we used:

- Computer Vision Toolbox
- Deep Learning Toolbox
- Deep Learning Toolbox Model for MobileNet-v2 Network
- Image Processing Toolbox
- Parallel Computing Toolbox

### 2.4 Implementation

The following are the training options of our model:

- Execution environment: GPU
- Mini batch size: 64
- Max epochs:300
- Initial learn rate: 1e-4
- Shuffle: every-epoch
- Validation data: imdsValid
- Validation frequency: 60
- Verbose: true
- Plots: training-progress
- Output network: best-validation-loss

Based on the pre-trained MobileNet-v2 network structure provided in MATLAB, we use 70% samples in CK+ dataset as the training set, 15% samples as the valid set, 15% samples as the test set and use the above Settings to conduct network training and fine-tune the model parameters. We also accelerated it with Parallel Computing capabilities provided by the Parallel Computing Toolbox and GPU, eventually applying our model to facial expression recognition.

## 3. RESULTS AND DISCUSSION

### Results

This is the GUI of our project. Here is just the result display. After inputting the images that you want to recognize, the GUI will display the grayscale image. Then, through the network computation, the probability ratio of each expression type is obtained and a bar chart is drawn to display it.

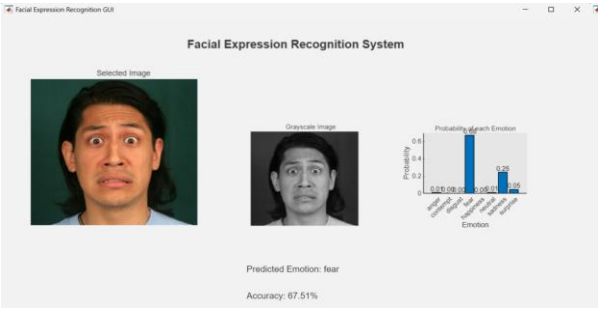


**FIGURE 2:** Graphic User Interface of Our Project

However, we also encountered some limitations in the analysis of the results. Since we only trained on the CK+ dataset, the results calculated by the model will be highly consistent with the classification benchmark of CK+.

This means that if we choose an angry face that humans can distinguish, the model may classify it as a disgusting face based on the dense wrinkles on the face.

In short, there are some limitations in the depth of the model. Compared with humans who perceive expressions from emotions, the model classifies facial expressions more mechanically based on features.

### Discussion

In our research, we closely integrate MATLAB with what we learn, which includes both traditional mathematical modeling and analysis, and modern machine learning frameworks such as TensorFlow and PyTorch. As a commercial mathematics software, MATLAB has shown strong advantages in parallel computing.

First, MATLAB is outstanding in parallel computing optimization. By using Parallel Computation Tools and GPU computing, we observed significant improvements in efficiency, even up to ten times of traditional CPU computing. This capability makes MATLAB a powerful choice when it comes to large-scale data processing and complex algorithm implementation.

Secondly, MATLAB's ease of use is also one of its attractions. Although its underlying implementation is relatively opaque, its good encapsulation and user-friendly interface make it easy to get started even by those who are not familiar with programming. For professionals in mathematical computing and data analysis, MATLAB provides a stable and efficient working environment, allowing them to focus on the essence of the problem rather than the complexity of programming.

In summary, through this exploration, we have deeply experienced the superiority of MATLAB in parallel computing and user-friendliness. This not only helps us deepen our understanding of the tool itself, but also provides strong support and guidance for future research and work.

## 4. CONCLUSION

In this project, we used MATLAB and relevant toolbox to train and fine-tune the MobileNet-v2 model on CK+ data set, and finally realized a facial expression recognition system. The system can recognize face images, analyze and predict facial expressions and give the prediction probabilities of various expressions.


## ACKNOWLEDGEMENTS

This project would like to thank Ms. Wei Yan for her advice and support. Thank her for leading us into the door of MATLAB. Instead of boringly learning the theoretical knowledge of the language, she advocates applying the knowledge learned in the practice of our own knowledge field. This makes us more interested in MATLAB and gives us more inspiration for the design of the project.


## REFERENCES


[1] A. Mollahosseini, D. Chan, and M. H. Mahoor. *Going deeper in facial expression recognition using deep neural networks*. In 2016 IEEE Winter conference on applications of computer vision (WACV), pages 1–10. IEEE, March 2016.

[2] H. Yang, U. Ciftci, and L. Yin. *Facial expression recognition by de-expression residue learning*. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 2168–2177, 2018.

[3] T. Zhang, W. Zheng, Z. Cui, Y. Zong, J. Yan, and K. Yan. *A deep neural network-driven feature learning method for multi-view facial expression recognition*. IEEE Transactions on Multimedia, 18(12):2528–2536, 2016.

[4] 腾讯云博客-全栈君. Pytorch *实现基于卷积神经网络的面部表情识别*. https://cloud.tencent.com/developer/article/2085474, 2022

[5] 木卯_THU. (2022, 11 16). 深度学习之图像分类（十一） --MobileNetV2 网络结构. https://blog.csdn.net/baidu_36913330/article/details/12007909