Proposal: Deep Learning Based Face Mask Recognition

ELEC-435: Deep Learning: Foundations, Applications, and Algorithms 2022 Win

JiaqiGuo: JGR9647; ManzhuWang: MWU2863; HuiwenHu: HHX5994; ChenxiLiu: CLS9757

Summary of the Proposal

Due to the highly contagious nature of COVID-19, personal protection has become extremely important. For the purpose of controlling the spread of the epidemic, every one of us has the obligation and responsibility to wear masks. The objective of our project is to proposed a system that can monitor people's mask wearing status (Correct, Incorrect and No mask). To be more specific, we will propose a dedicate deep learning framework to finish this classification problem. Besides, we will compare our results with several other existing deep learning milestone models (e.g. Resnet, LeNet) to prove the feasibility of our model

Background

Until February 2, there have been 386,548,962 confirmed cases of COVID-19 worldwide and droplet transmission is the main cause of transmission. Traditional labor-based supervision system will cause a great waste of labour resource. Thanks to advances in computer science and artificial intelligence in recent years, we can use image-based deep learning techniques to solve this problem. That is, collecting facial images through cameras and performing masking wearing recognition with specific deep learning frameworks. In this way, labor and time costs can be saved, and the spread of the virus will be more effectively controlled.

Goal and Objectives

The main objective of this project is to propose a deep learning framework to classify a set of face mask images. Through the deep learning framework proposed in this project, we will finally classify this data set into 3 categories, which is illustrated as below:

- **Correct:** The person in the picture is wearing a mask correctly.
- **Incorrect:** The person in the picture is not wearing the mask properly, for example by exposing his nose and mouth to the outside.
- **No Mask:** The person in the picture is not wearing a mask at all

Our network will eventually achieve recognition accuracy of more than 85% and show high robustness to various disturbances (e.g. light, skin color, face orientation). Besides, we will further compare our proposed model with a some deep learning milestones, critically analyze the advantages and disadvantages. And the result will be summarized in the final submitted demo video and Jupyter notebook.

Methodologies

This project mainly implement two deep learning models: multilayer perceptron and convolutional neural network (CNN) as a comparison.

Data Preporcessing: We use some image preprocessing methods to extract features from the original image. And compare the results with using original images to explore the efficiency of these methods.

- **Feature Extraction:** Dense local descriptors (HOG or local binary pattern); Encoding (local coordinate coding or Gaussian model super-vector coding).
- **Dimensional Reduction:** Pooling (weighted pooling or max-pooling with spatial pyramid matching); principal component analysis; linear discriminant analysis; canonical correlation analysis; non-negative matrix factorization.

Multilayer Perceptron: An MLP consists of at least three layers of nodes: an input layer, a hidden layer and an output layer. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP utilizes a supervised learning technique called backpropagation for training.

Convolutional Neural Network: Compared to MLP, the CNN is more powerful to deal with images. In order to make comparison with MLP, we will build a lightweight CNN with a few layers to keep the amount of parameters of MLP and CNN in the same level. Furthermore, we will conduct some state-of-the-art CNN models like LeNet and ResNet. A CNN model is very easy to be conducted using some deep learning libraries like Pytorch, Tensorflow and Sklearn.