



Identification of emotions through digital image processing

Kehan Li

Tutor: Carlos Hernandez

Cotutor: Jose Javier

Trabajo Fin de Grado presentado en la Escuela Técnica Superior de Ingenieros de Telecomunicación de la Universitat Politècnica de València, para la obtención del Título de Graduado en Ingeniería de Tecnologías y Servicios de Telecomunicación

Curso 2017-18

Valencia, 16 de abril de 2019



Abstract

Image Processing technology has a wide range of applications in all walks of life including the well-known emotion recognition technology. Emotion recognition is the process of identifying human emotion, most from facial expressions and verbal expressions. This is both something that humans do automatically but computational methodologies have also been developed.

In this project I focus on facial expressions and use TensorFlow, a free and open-source software library, to train models by inputting substantial images of people with different emotions. The information of facial expression is typically concentrate on facial expression information regions, which are the mouth, eyes and eyebrow regions. In order to identify it, the TensorFlow model should be trained and get ability to distinguish the position and shape of the above organs to find diverse parameters of the decision tree for recognizing the facial expressions.

However, mostly the input images are normal people and the program will give wrong results when it faces people with mental illness or mentally handicapped whose facial expression is different from others. To have better performance of using model to recognize the emotions in this situation, I went to Center “La Torre” and collected many images of people with intellectual disability and discussed with the professors there. According to the advice they gave to me, I labelled and added this part of the images to the training dataset and adjusted some of the parameters of the model. After several rounds of training and testing, the model finally achieved good results and I used TFlite, a program provided by TensorFlow, to put the trained model into an .apk file which means it can be used on both computer and android phone.



Índice

Capítol 1. Introduction	2
1.1 Background and purpose of this project	2
1.2 Implementation of this project	2
Capítol 2. Background	4
2.1 Not finished yet.....	4
2.2.....	4
2.3.....	4
2.4.....	4
2.5.....	4
Capítol 3. Design and Implementation.....	5



Capítol 1. Introduction

1.1 Background and purpose of this project

As a famous branch of image processing, the emotion recognition technology has already entered our life. When a person is driving, the camera inside the car can automatically recognize the emotion on the driver's face, determine whether he is distracted, anxious or in some unfavourable emotions and issues a warning. When people shop in the mall, the camera can automatically recognize the expression on the customer's face and determine whether they are happy, excited or other emotion for a certain product to help mall understand the requirement of customers. Besides, it also applied in collecting the video game players' experience, assisting health care providers in assessing patient health and so on, the areas around emotion recognition continue to grow. Amazon, Microsoft and IBM also promote "emotional analysis" as the main function of their facial recognition products.

However, emotion recognition is based on machine learning technology, training the program through a large number of images of known content, then let the program analysis the other images according to the learning experience. While these applications focus on normal people, they may give low accuracy results under some specific conditions. People with mental illness or mentally handicapped often have facial expression different from normal people although they are in the same emotion.

This project aims to implement a program which have high accuracy of recognition when faces to the people with mentally handicapped.

1.2 Implementation of this project

After learned the knowledge of facial emotion recognition and machine learning and discussed with my supervisor, I decided to use TensorFlow in this project while its high efficiency and strong compatibility match the project requirements and supports mainstream programming languages C and Python. TensorFlow is a free and open-source software library for distinguishing between data streams and programming of various tasks and can also be used in machine learning applications such as neural networks.

To use Tensorflow, I installed the Ubuntu 18.04 system while it provides lots of relevant softwares, and virtual environment which can make me have higher working efficiency. During preparing the training image dataset, I used a software named labelIMG, it can open an image and select an area of the image then input the words to describe what's in the area. The area and the words I input combined and became the label of this image. After labelled the image, I converted all the file in one data folder into .tfrecord file, which is the standard format required by TensorFlow training, by using Python script. The last step before start training is choosing and preparing the model and maybe it's the hardest part. There are many models in TensorFlow library and have different characteristics and advantages depended on the application and the device such as training time, accuracy and inference speed. After balance the inference speed and accuracy, I chose SSD (SingleShot-Box-Detector) with mobilenet-v1 model which has been optimised for the use on mobile device so the inference speed is really fast. Then modified the configuration file of the model and activated the python script, the model training was started. The training time depends on the scale of the used image dataset and the number of steps, the larger dataset and number of steps, the longer training time.

In this project, I totally established three image dataset each contains about 1000, 3000 and 5000 images. The first dataset is combined by the public images of normal people I found from internet. The second and third dataset not only contained images from normal people but also contained images from people with mentally handicapped which provided by "La Torre". "La Torre", an occupational center in Valencia, Spain which helps people with mentally handicapped to work and socialize. I went to the center with my supervisor to get images and discussed with the



professors working. These professors gave me many advices to improve the performance of the model and hope this program can help them understand the people in the center more and make them better communicate with each other.

After several rounds of training and testing, the model improved a lot comparing with the first time and is able to give high accuracy results while facing people with mentally handicapped. By using the TFlite function, a program provided by TensorFlow, I put the trained model into an .apk file which means it can be used on both computer and android phone directly and efficiently.



Capítol 2. Background

2.1 Not finished yet.

2.2

2.3

2.4

2.5



Capítol 3. Design and Implementation

Not finished yet.