BABEŞ BOLYAI UNIVERSITY, CLUJ NAPOCA, ROMÂNIA
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Emotional recognition using artificial intelligence

– MIRPR report –

Team members

Name, specialisation, group

Sbengheci Vlad-Andrei, Informatica Romana, 236

Săfticu Daniel-Marian, Informatica Romana, 236

Stegaru Sebastian, Informatica Engleza, 937

Tăuteanu Andrei, Informatica Engleza, 937

Abstract

Ensuring that a piece of software is easy and intuitive to use is vital to its success. While there are many guidelines for ensuring a pleasant user experience, the most important part of improving an application is user feedback. Unfortunately, feedback is as hard to gather as it is important, especially when the product in question is addressed to very young children, who can not read, write, or describe their experience in detail. This issue can be solved using artificial intelligence. By using emotion recognition on the user, an artificial intelligence algorithm can compile and classify data relating to their experience, revealing valuable information about the usability of the application. Face recognition should also be applied, in order to identify the user.

Contents

Chapter 1

Introduction

1.1 What? Why? How?

Because in today's word the need for computers keeps on growing, the software has to be more intuitive than ever. To ensure the ease of use, applications are tested on children and their reaction are stored. But without a use of an expert, the accuracy could not be as great. So appeared a need of a program that determine and records the facial expression of young children, without a need of a human observer. We try to resolve this problem with an application that let the child login into their account intuitively and with ease, and starts in the background to accurately detect and records their reaction using a neuronal network.

1.2 Paper structure and original contribution(s)

The research presented in this paper advances the theory, design, and implementation of several particular models.

The main contribution of this report is to present an intelligent algorithm for solving the problem of detecting the facial expression of a child in a short amount of time and with the best accuracy possible.

The second contribution of this report consists of building an intuitive, easy-to-use and user friendly software application. Our aim is to build an algorithm that will help removing the need of an expert in the process of detecting and recording the facial expressions for the test of intuitive programs.

The third contribution of this thesis consists of detecting the facial expression of a child in a short amount of time and with the best accuracy possible.

Chapter 2

Scientific Problem

2.1 Problem definition

Our bigger problem is how can a program can accurately detect facial expression, without using a lot of resources and in a short amount of time. To be more specifically, we need to give a image from the camera and return a list of probability for each facial emotion, everything happening in real time. This problem require the use of an intelligent algorithm because of how abstract the problem is. Another used method of resolving this problem is manually extract specific shapes for every emotion and then the program would detect. This method is not as good as an intelligent algorithm because the overall accuracy is lower and determining the shapes would take a lot of time.

This problem require to imitate into a computer program one of the basic function that a human brain can do intuitively.

Chapter 3

State of art/Related work

A similar application of emotion recognition has been proposed in Emotion Recognition and its Application in Software Engineering, where 8 different scenarios are presented where artificial intelligence can be used to assess the quality of human-computer interaction. The scenarios are:

- 1)First impression test: where the user's facial expressions upon first seeing the interface is used to rate its design and usability, differentiating between the user's excitement or boredom/disgust.
- 2)Task based usability test: here a larger array of data is used: visual, biometric and an analysis of keystrokes, while the user is walked through typical tasks. The main comparison being made here is between feelings of empowerment and frustration.
- 3)Free interaction test: this method is mainly veered towards entertainment platforms, letting the user navigate with no predefined task, and distinguishing between encouragement and discouragement.
- 4)Tele and office working comparison: this method relates to comparing software developers' emotions, when working from home or in the office, on a PAD scale.

While there are 4 other scenarios presented, they pertain to productivity more than user experience, therefore being less relevant to the subject of our project. However, the 4 presented scenarios can easily be translated to the issue of assessing the quality of the interaction between young children and software. Both interface design and the setting in which the application is used can have great effect on the youngsters' experience and emotion recognition can be used to quantify this effect.

Chapter 4

Proposed approach

4.1 Approach 1: Facial Expression Detection Techniques: Based on Viola and Jones algorithm and Principal Component Analysis

The Viola-Jones algorithm is a widely used mechanism for object detection. The main property of this algorithm is that training is slow, but detection is fast. This algorithm uses Haar basis feature filters, so it does not use multiplications. In this approach the local features (like nose, eyes) of the face are found. Then these features are segmented then they are used as the input data for structural classifier.

4.2 Approach 2: Landmark-Based Support Vector Classifier

The feature that the algorithm uses to classify the image is the position of different landmarks on the face ("the moving parts"). Using the relative position of the landmarks from a "center of mass".

Before the landmarks are extracted, some preprocessing is applied to the image. First, the image is resized and converted to grayscale. Then the contrast is improved using a Contrast Limited Adaptive Histogram Equalization algorithm.

The extracted landmarks are then fed into an Support Vector Classifier with a linear kernel function.

4.3 Approach 3: Automatic emotion recognition through facial expression analysis in merged images based on an Artificial Neural Network

The preprocessed image is fed to a fully connected ANN. The preprocessing consist of firstly extracting only the eyes and mouth of the face, then resizing to a specific size, converted to greyscale and than converting to a binary image, with the pixel over 230 having the value 0 and |(pixel - 255) / 255| otherwise.

4.4 Approach 4: Affective Information Processing and Recognizing Human Emotion

The Neuronal Network is fed the image and detects independent Action Units(AUs). Expressions can be described by combining multiple AUs. Each expression is given a score that is made up of the list of AUs.

4.5 Approach 5: Emotion recognition based on image processing with Convolutional Neural Network

The preprocessed image is fed to a CNN. The preprocessing consist of resizing the image to a specific size

4.6 Approach 6: Landmark-based Artificial Neural Network

This approach is using same process to as Approach 2 to extract the landmarks, the landmarks are normalized then feed to a fully connected ANN

Chapter 5

Application (numerical validation)

5.1 Methodology

The data set is randomly split into training data and test data, with a proportion of 80-20. This is done every iterations, for 10 iterations in the case of this particular test trial.

5.2 Data

The dataset use was the CAFE dataset, containing labeled pictures of children's facial expressions.

5.3 Approach 2: Landmark-Based Support Vector Classifier

5.3.1 Results

In the first test, with all 7 classes, the algorithm has achieved an average accuracy of 65% on the test data, over 10 iterations.

Following a hypothesis that the 'disgust' part of the dataset causes overfitting, disgust was removed from the emotion model, and the algorithm achieved an accuracy of 73% over test data.

5.4 Approach 3: Automatic emotion recognition through facial expression analysis in merged images based on an Artificial Neural Network

5.4.1 Results

In the first test, only a small portion of the test data was use(100 in total for the training data and 80 for the test data) and for an ANN with 5 fully connected layers, the accuracy reached 60%.

The next tests consist in increasing the data and the number of layers, each time the accuracy dropping.

5.5 Approach 5: Emotion recognition based on image processing with Convolutional Neural Network

5.5.1 Results

The first tests consist in modifying the CNN size and parameters, for training using only 972 photo samples and 237 for validation, the best accuracy obtained was around 45% but the loss kept increasing on validation.

The next tests consist in increasing the data using translations and gaussian blur to images, increasing the number of training data to around 23000, modifying the structure of the network more the accuracy increase to 56% which was the best result for this result

5.6 Approach 6: Landmark-based Artificial Neural Network

5.6.1 Results

In the first tests the ANN was composed of 10 fully connected layers, and the accuracy reached 60%

In the future tests, after decreasing ANN size to 6 fully connected layers, using dropout layers to reduce the chance of overfitting, and using only sigmoid and relu activation function the accuracy reached 71%

The confusion matrix show that this model is classifing most of the sad emotion as neutral

Chapter 6

Conclusion and future work

Every approach gave significant results, but the 2nd one, the Landmark-Based Support Vector Classifier approach has the most potential.

For now, the best result was the accuracy of 73%, but with more data and more iterations, we believe that this result can be better.

What we succeeded in doing was creating a method that creates acceptable results in a short amount of time. What we demonstrated with our results is that it is possible to have a good accuracy on detecting facial expressions with a neuronal network that was trained on a relative small amount of data.

Bibliography

[1] Facial Expression Detection Techniques: Based on Viola and Jones algorithm and Principal Component Analysis

[2] Emotion Recognition and its Application in Software Engineering

[3] Automatic emotion recognition through facial expression analysis in merged images based on an Artificial Neural Network

[4] Affective Information Processing and Recognizing Human Emotion