TRƯỜNG ĐẠI HỌC
SƯ PHẠM KỸ THUẬT TP. HỒ CHÍ MINH
HCMC University of Technology and Education
HCMUTE

KHOA ĐÀO TẠO CHẤT LƯỢNG CAO
FACULTY FOR HIGH QUALITY TRAINING
FHQ

EMOTION, GENDER, AGE RECOGNITION BASE ON HUMAN FACE USING CONVOLUTIONAL NEURAL NETWORK AND BOUNDING BOX REGRESSION.

Nguyễn Công Thành[1].

1. HCMC University of Technology and Education, Ho Chi Minh City, Vietnam.

ARTICLE INFO

Article history:

Received 23 June 2022

Keywords:

CNN

Bouding box

Regression

Deep learning

Python

Realtime

face detection

Age,gender,emotion

ABSTRACT

This research is about age, gender and facial emotion recogniton using Convolution Neural Network (CNN) model and human face localization using bouding box regression , applying on human face for customer service analysis, marketing, surveying, security and crimes investigation

I. Introduction

Currently, the development of computer vision technology and AI is helping the retail and customer service industry to find optimal solutions to analyze how to approach consumers. Not only that, AI technology is also applied in the fields of security, military, education, healthcare and countless other practical applications that we cannot list.

One of the real life applications of AI is human age,gender and emotion prediction for analyzing customer attitudes towards products and services or security investigation

This solution is based on the method of deep learning: Convolutional neural network(CNN) and the the detection techniqe of human face using bouding box regression to extract the region of interest(ROI)

In this article, I will disscus the method in detail

II. Methodology

1. Image labeling

The goal of image labeling, a type of data labeling, is to recognize and tag particular details in an image.

There are many image labeling tool, which Labelimg is one of them

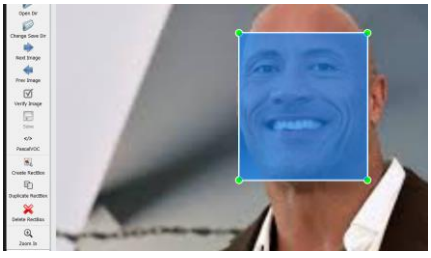


Fig.1 Image labeling using Labelimg

2. Face Detection

Face detection is a computer technology being used in a variety of applications that identifies human faces in digital images.[Face detection also refers to the psychological process by which humans locate and attend to faces in a visual scene

Face-detection algorithms focus on the detection of frontal human faces. It is analogous to image detection in which the image of a person is matched bit by bit. Image matches with the image stores in database. Any facial feature changes in the database will invalidate the matching process.

3. Bouding box regression algorithm

Bounding-box regression is a popular technique to refine or predict localization boxes in recent object detection approaches. Typically,bounding-box regressors are trained to regress from either region proposals or fixed anchor boxes to nearby bounding boxes of a pre-defined target object classes.

5. Binary and multiclass classification

The logistic classification model (or logit model) is a binary classification model in which the conditional probability of one of the two possible realizations of the output variable is assumed to be equal to a linear combination of the input variables, transformed by the logistic function. Losgitic classification is used to classify the human genders based on the facial apperance using the loss function binary crossentropy and sigmoid activation function

Multiclass classification provide a natural extension to the multi-class problem. Instead of just having one neuron in the output layer, with binary output, one could have N binary neurons leading to multi-class classification. In practice, the last layer of a neural network is usually a softmax function layer, which is the algebraic simplification of N logistic classifiers, normalized per class by the sum of the N-1 other logistic classifiers. Softmax classification is used to classify the facial recognitions based on the facial apperance using softmax activation function and categorial crossentropy

4. Age Regression

The algorithm which using neural network regression to estimate the age of human based on facial apperance using loss function mean squared error(mse) or mean absolute error(mae) to calcultion the loss

III. CNN Model and algorithm

1. Datasets

Our dataset is taken from UTKface . The image was labeled into PascaVOC format.

Fig.2 Images data and annotation files

2. Algorithm and model

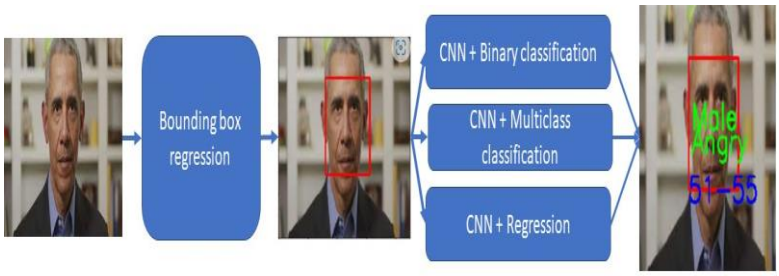


Fig.3 Algorithm

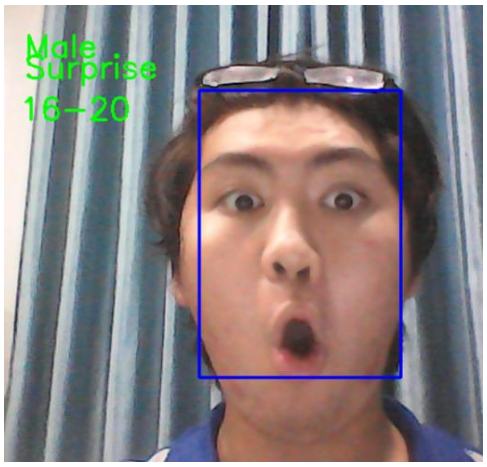
The human face was detected by the bouding box regressor. The extracted ROI was taken into the CNN model to predict the output.

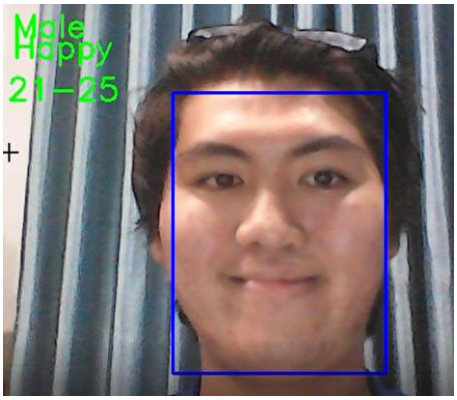
The CNN model with losgitic classification is used for genders classifying

The CNN model with softmax classification is used for emotion recognition

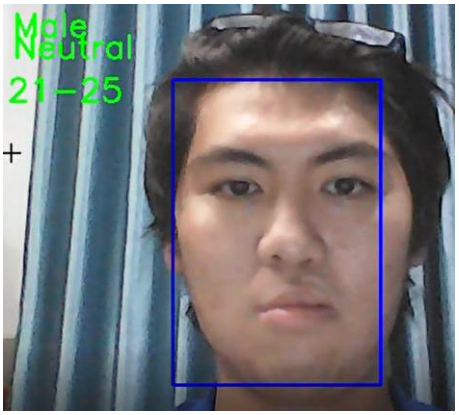
The CNN model with regression is used for age estimation

IV. Results




Male
Happy
21-25


Male
Angry
21-25


Male
Nautral
21-25