MEDINFO 2021: One World, One Health – Global Partnership for Digital Innovation
P. Otero et al. (Eds.)

doi:10.3233/SHTI220302

Artificial Intelligence-Based Mobile Application for Emotion Sensing for Children Through Art

Nashva Ali[a], Alaa Abdrazaq[a], Zubair Shah[a], Mowafa Househ[a]

[a] Division of Information and Computing Technology, College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar

Abstract

In this paper, we develop an artificial intelligence (A.I.) based Emotion Sensing Recognition App (ESRA) to help parents and teachers understand the emotions of children by analyzing their drawings. Four different experiments were conducted using a combination of two datasets. The deep learning model was trained using the Fastai library in Python. The model classifies the drawings into positive or negative emotions. The model accuracy ranged from 55% to 79% in the four experiments.

Introduction

Over the past five years, there has been a steady growth in the use of technology and Artificial Intelligence to analyze emotions. A dearth of research has been conducted in using Artificial Intelligence to understand the emotions of children through art. We only found one study where a researcher developed a deep learning model to detect the emotions of children through their drawings by extracting features representing five emotional states: happiness, anxiety and depression, anger, and violence [1]. The model achieved an accuracy of 87% showing that this technique can help psychology professionals detect child emotions through art [1]. The major limitation of the study was not involving Art Therapy professionals to help label the data. This major limitation, we believe, may have led to the improper classification of the drawings and yielded biased results. Furthermore, the researcher also trained and validated the model using a small dataset, which was taken from children with symptoms related to Post Traumatic Stress Disorders (PTSD) [1]. Given that little research work has been conducted in this area of research, in this paper, we study the use of A.I. to assess the emotional state of children through Art. We anticipate that our work will be helpful to mental health professionals, parents, and social workers in understanding the emotions of children through art.

Methodology

Data Collection

This study included Primary (local) and Secondary (online) data sources for child art drawings. Our primary source was data collected from a local school based in Doha. We collected 102 drawings to test the effectiveness of art therapy-based emotion-sensing. Our secondary data source was searching online where we collected a total of 521 drawings through Google and Instagram.

Data labelling

Due to the limited number of drawings, an Art Therapist suggested to have only two high-level emotional classifications. The positive class included emotions such as happiness, excitement, pleasant, energetic, and hope. Likewise, the negative class included emotions such as sad, angry, hate, fear, abuse, and disappointment. Thus, the data was labelled by the Art Therapist into two different groups: positive emotions (n=365) and negative emotions (n=258). The drawings in the negative class were labelled as 0 and drawings in the positive class were labelled as 1 for the model.

Deep learning model

In this study, ResNet50 convolutional neural network model was implemented for classification of drawings. The model was implemented with Python using the open source Fastai (https://www.fast.ai/) framework. The parameters required for ResNet50 were set by default by the Fastai library and did not require the researcher to optimize the parameters in order to obtain the results. This study employed transfer learning to train the convolutional neural network for classifying the drawings as positive or negative. Transfer learning uses a pre-trained model built upon a very large dataset which is then customized for the problem at hand. This technique is useful for a small dataset, which is the case in this study.

Data processing

The drawings were generated on a per-batch basis, with a batch size of 64. This means that the algorithm takes the first 64 drawings from the training set and trains the network. All the drawings were reshaped into 224x224 pixels. The Fastai library has a set of standard transformations designed for photos by default, which were used in this experiment. In addition, data augmentation was applied on the training data where more data can be produced from the existing dataset. For every image in the training set, the model created 9 augmented drawings, which were created from original image using various augmentation techniques such as cropping, zooming, rotating, flipping and tilting. Three pre-trained models were used for training the model: ResNet18, ResNet34, and ResNet50

Experimental design

Four different experiments were tried out using our primary and secondary datasets. Initially, only one experiment was tried in which the model was trained on our secondary data and tested on our primary dataset. But this model did not yield good results. As a result, we decided to conduct four different experiments with the two datasets.

Experiment 1

This model was trained and tested on the primary (local dataset) which contains the school drawings using ResNet34. This was our smallest dataset containing only 102 drawings collected from the school. The data was split into 80-20% for training and test set respectively. Approximately, 20% of the data from both positive and negative classes constituted the test set, which included a total of 21 drawings. The training set had 65 drawings while the validation set had 16 drawings.

Experiment 2

This model was trained and tested on the secondary (online dataset). This set constituted a total of 521 drawings collected from both Google drawings and Instagram. We split the data to 80-20% respectively for training and the test set. The test set had a total of 104 drawings, which was 20% of the data from both the positive and negative classes. The training set and validation set had 334 and 83 drawings respectively.

Experiment 3

This model was trained on the local dataset but tested on the online dataset. The local dataset had 102 drawings from the school. Twenty percent of the drawings were included in the validation set which contained 21 drawings and the remaining 81 drawings were included in the training set. Randomly 40 drawings from the online dataset were selected to represent the test set.

Experiment 4

This model was trained on the online dataset and tested on the local dataset. The online dataset had 521 drawings in total out of which 417 drawings constituted the training set and 104 drawings the validation set. The entire local dataset was used as the test set for this experiment.

Intermediate Layer Activations

The basic function of our deep learning model was to classify the given drawings into two primary classes: positive emotions or negative emotions. We also wanted to highlight the areas in the image to demonstrate how the algorithm selected a particular classification. Therefore, in addition to the basic function of our deep learning model, we implemented an algorithm that helps us illustrate the layer activations that classify an image as having positive or negative emotions. A heatmap was produced highlighting the layers that find patterns that lead to the activation of the image as being classified as positive or negative. In sum, our deep learning model not only classified an image into one of the two primary groups, but also can explain, through the activation feature, the reasons behind the classification decision.

Results

The first model was trained and tested on the primary dataset. The accuracy achieved for the model was 76% (Table 1). The second model was trained and tested on the secondary dataset. The model achieved 79% test accuracy, which was the highest among the four experiments. The third model was trained on the local dataset but tested on the online dataset. The accuracy achieved for the model was 62% which unsatisfactory. The final model was trained on the online dataset and tested on the local dataset. The accuracy achieved for this model was 55% which was almost the same as random sampling.

Experiments	Training dataset	Testing dataset	Model accuracy	Model	Epochs
Experiment 1	Primary	Primary	76%	ResNet 34	10
Experiment 2	Secondary	Secondary	79%	ResNet 50	20
Experiment 3	Primary	Secondary	62%	ResNet 34	15
Experiment 4	Secondary	Primary	55%	ResNet 18	25

Table 1: Summary of the classification results

Intermediate layer activations

The intermediate layer activations were developed to draw a heatmap on top of the image to illustrate or explain the reasons behind the model's prediction. The corresponding algorithm to obtain the intermediate layer activations and heatmap illustration was used only for experiment two. The algorithm was not implemented in experiment one due to the inconsistency among the primary dataset. The school drawings varied largely from each other in terms of shapes or colors. Experiment three and experiment four did not employ this algorithm as the accuracies obtained for these two experiments were not sufficient.

With the layer activations, the particular pixels of the image that was responsible for the prediction of that image as positive or negative was highlighted. These highlighted areas illustrate what shape or color had been studied by the model and classified as either positive or negative.

One obvious pattern that was observed from the heatmaps was that of the detection of color yellow as positive and red color as negative.

These results illustrate that this model can classify an image into positive or negative emotions along with highlighting the layers activated in order to be classified into one of the two groups.

Conclusion

Emotion Sensing Recognition App (ESRA) has the potential to help parents and teachers to understand the emotions of children through analyzing their drawings using artificial intelligence (AI). However, the underlying algorithm proposed in this work needs to be trained and evaluated using more drawings to improve its current accuracy. Further, more works are needed to improve ESRA so that it can identify more emotions (e.g., happy, sad, depressed) and to consider cultural nuances in the representation of the drawings.

References

[1] Martinez VR. How Artificial Intelligence can detect emotions in children's drawings. 2019 [November 11, 2020]; Available from: https://medium.com/datadriveninvestor/how-artificial-intelligence-can-detect-emotions-in-childrens-drawings-4359cf51ab3d.