Affective computing and

emotion detection

Anna PRECZYŃSKA, Mariusz PISARSKI,

Yura SHCHEOHOLIEV, Michał SULIBORSKI

*Information Technology,*

*Lodz University of Technology, International Faculty of Engineering*

**Abstract** - *Current methods concerning emotion detection using technological devices are either inaccurate or require advanced apparatus and instruments. They almost entirely do not give the possibility to be implemented in the mobile devices, since they base on many aspects of emotion recognition, including internal processes of an organism, visual or vocal feedback. In this project we combined many different methods of emotion detection during the creation process, in order to implement the solution on the mobile device and increase the accuracy of the resulting product. We used Android Studio as IDE and for detecting emotion we used library called TensorFlow Lite. During the process of creating the project we used deep learning method, which combined with “Keggle Face Dataset” allowed us to considerably improve its reliability. In this article we present the output of our work, which by using different technologies and libraries resulted in creation of a mobile-device-application to recognise emotions.*

1. Introduction
   1. Background information

Detecting emotions is an everyday task that average person faces many times a day while taking part in social interactions. Most people however do not realize that for some it can be a struggle. Among people with an identified spectrum of autism, problems with social skills are clearly noticeable, including the difficulty in recognizing emotions. This leads to difficulties in communications, making new friends and everyday interactions with other people.

* 1. Problem finding

In order to find out, if nowadays technology allows us to find out the emotions we have conducted a research in order to find as many methods of detecting emotions as possible. The main problem was the lack of a generally available, effective method, guaranteeing 100% of accuracy . In conclusion to this research was decided to define the problem statement: “ Inaccuracy in emotion detection process, due to lack of verification methods”.

1. Idea finding
   1. State of the art

After finding the problem statement, our research proceeded in the direction of finding the information about emotion detection. First step was to describe emotions themselves and what processes are involved in feeling a certain emotion, which led us to biofeedback.

Biofeedback is a therapy, which results in gaining control over particular physical processes in human organism. The most relevant aspect of biofeedback are instruments used during the therapy. Different experiments prove a high accuracy of emotion detection, when using biofeedback devices.

Next part of the research considered finding possible ways of emotion detection and emotional state by using methods, which resemble human senses. This idea focused on hearing and sight, for those ones might be easily substituted by camera or microphone. According to a study, led by Jeffrey F. Cohn, Tomas Simon Kruez, et al. such an approach proved its accuracy in 79%. They used two microphones, as well as two cameras recording face and shoulders positioned approximately 15 degrees to the left and right, one camera recording face and one for full body recording.

* 1. Innovative ideas

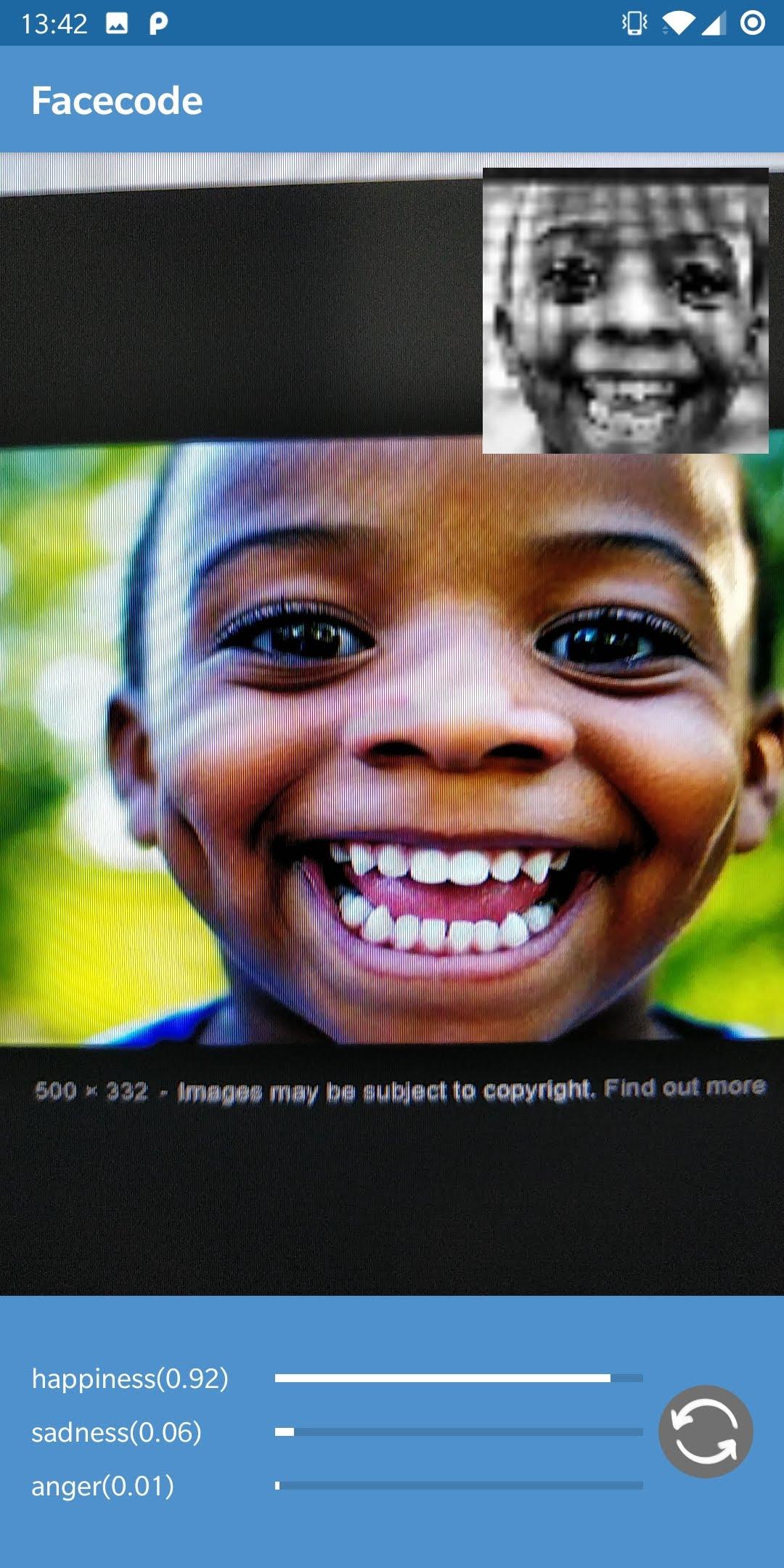
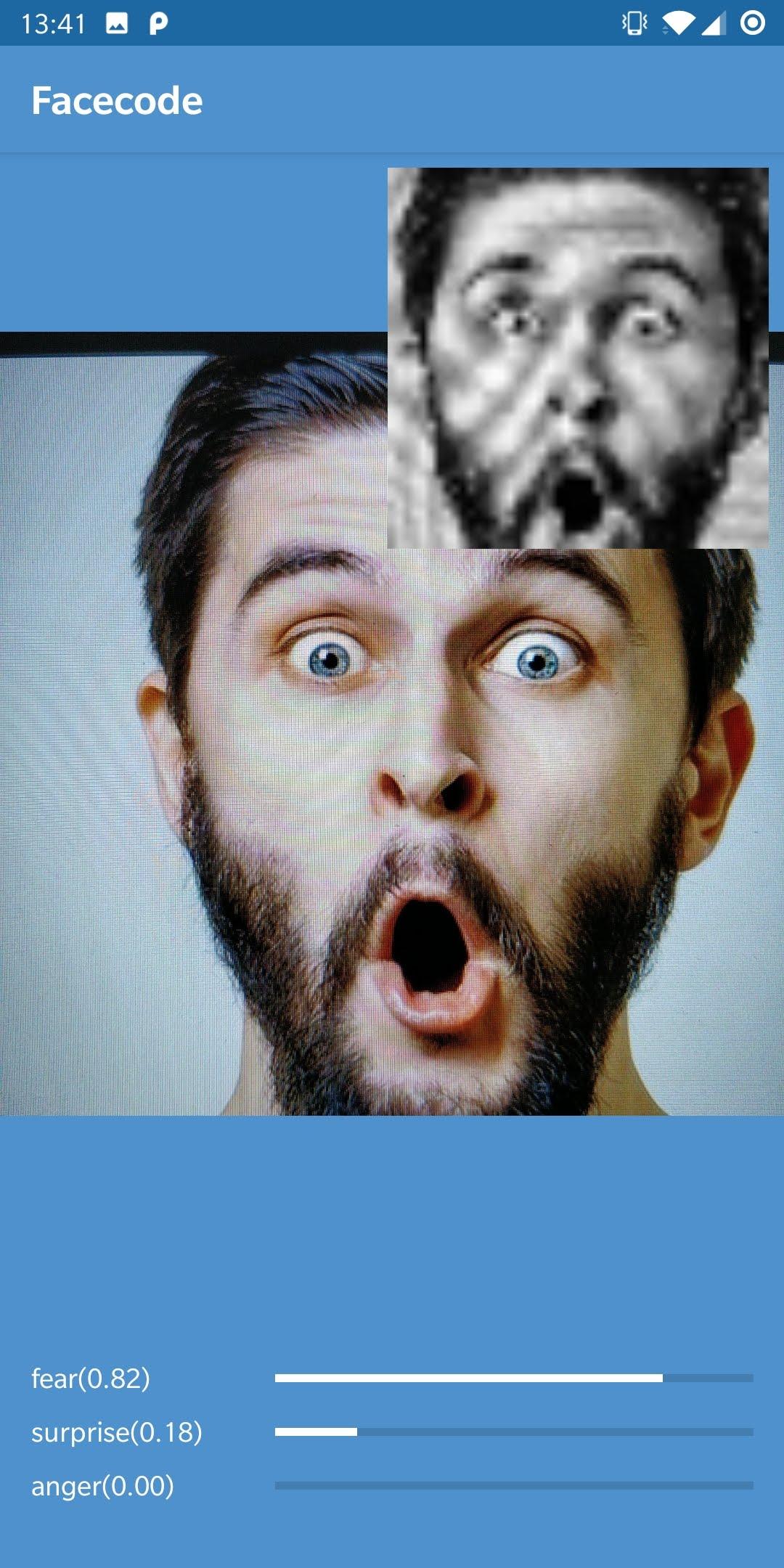
Our idea of solution to the problem, was to combine facial recognition of emotions with biofeedback instruments, in order to verify obtained results. Google vision provide an already created face detector. We also considered using deep learning method, which is technical representation of human-based learning method. It memorises patterns of previously encountered objects and uses them to recognise other objects with similar patterns. With a large database the program would be able to recognise facial expression with high accuracy. In order to increase reliability even more, emotions learned via database would be verified by using biofeedback instruments.

* 1. Main idea selection and justification

When choosing the most appropriate solution to the problem, we had to take into consideration many aspects. We decided to make an application, because all of the team members had skills to produce one. Our main group of stakeholders were autistic people. Therefore, we knew that the applications should be as simplistic to use as possible. We also had to fit in our limited budget, while producing the solution.

Finally we decided on making the mobile device application, which by using the camera is able to recognise emotions from the facial expression of a person. This application would use Google Vision, in order to detect face patterns, and would be based on deep learning method. We also wanted to increase the reliability of our implementation by verifying the obtained results using biofeedback instruments, however resigned from this idea because of the high prices of needed apparatus. The accuracy could be easily improved by implementing an external database using machine learning. Many different databases were available, and mostly free to use. In fact, we found little of a problem in finding a database with different facial expressions in amount of over 28 000 photographs. We also aborted the idea of using multiple cameras and microphones, for they would not fit in the limited budget of ours as well. Moreover, They did not provide enough mobility, which we considered of high importance. According to discussions, which we have made, the possibility to use an application anywhere and at any time, might return better results, when treating autism disorder. Face detector from Google Vision also required additional costs, however these were low enough for us to cover.

1. Solution implementation



* 1. Detailed solution description

After defining the main idea as a solution of the stated problem our group started to work on the technical aspects of the implementation.

First we need to state the number of functionalities supported in our project, in order to clarify each members responsibility and to be able to finish the project in a given term.

From functionalities stated the application should be capable of detecting the human face form a picture taken from front or back camera of the phone, then analyse and detect the emotion of the person. It should offer two modes: one working on a given photo, second working in realtime. Finally it should implement nice, clean and functional design.

Because our project is designed to be in a form of mobile application, one important part of creating the solution was an interface design. In process of the interface design there were discussed different versions and propositions concerning the color gama, number and position of interactive buttons. The prototype was made in “Adobe XD” and the interface itself was made in ”Android Studio”. To check the efficiency, simplicity and understanding of an interface from user point of view our team visited the ASPI Foundation to gather feedback from the main target group of our application. This meeting helped us to highlight the problem present in a prototype and create a new version with all the requests and suggestions taken into account.

As for the internal implementation, the whole code for the application was written in Java in order to be able to launch the app on mobile devices supported by Android operating system. This decision also simplified the peripheral communication between code and the device by providing a powerful APIs. In particular using both google.hardware.camera and google.hardware.camera2 APIs to get an image from the cameras of the phone and save it as a bitmap (array of pixels).

In order to detect emotions from this image, few extra steps had to be done. The processed image must match the dataset images used to train our neural network. We decided to use Keggle Face Dataset, since it is free for educational purposes and provides over 28,000 pictures of faces with dimensions 48 by 48 pixels, all categorised into seven groups (anger, fear, disgust, happiness, neutral, sadness and surprise).

This dataset allowed us to use technique called retraining. It is essentially the process of using “knowledge” of pretrained model and data form new database. Model is then limited to recognise only labels from new dataset. We used Google’s pretrained model, called MobileNet V2, that specializes in classifying various objects. This classifier already saw millions of photos and is able to successfully classify them, what makes it is perfect for the job. The model has been retrained in Google Colab Notebook in Python using TensorFlow library and converted to .tflite format. The accuracy of emotion recognition was directly proportional to the amount and quality of images and amount of epochs (training cycles) of the dataset.

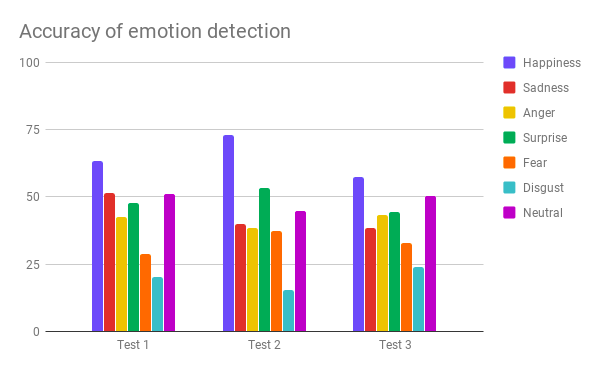
Having our classifier trained and ready to process data we had to prepare taken image to match dataset. We used the Google Vision API and tool called FaceDetector (com.google.android.gms.vision.facedetector) to extract face out of the bitmap. If the face is detected we extract it to a new bitmap (in case of multiple faces detector chooses the most foreground face). Because FaceDetector is not 100% accurate , in case the program does not detect face, it will still process image in order to find emotion. Then the bitmap is cropped to 1:1 ration, converted to grayscale and scaled to 48x48 resolution to match image properties in our dataset.

As it was mentioned above the efficient way to detect and classify emotions of an image is to use artificial intelligence and deep learning technologies. To use those we have chosen a library developed by Google called TensorFlow Lite. It allows to use trained classifier to detect emotions out of the photo.

Finally program takes proper bitmap, analysis it and outputs three most suitable labels out of seven available with its certainty. Then the labels are displayed alongside with bar visualizing the accuracy.

Program allows users to choose whether they want to analyze single photo on demand (photo mode) or they want to process camera feedback constantly (realtime mode)

The chart estimating the accuracy for different emotional facial expressions:



Unfortunately because of an unequal amount of images in datasets for different emotions the accuracy of the analysis differs a lot. As we can see from the chart, because the database has about 7000 images of happiness and barely 400 images of disgust, the algorithm chooses to output happiness much more frequently. The accuracy is also heavily impacted by the small size of the analyzed photos.

* 1. Ways of verification

We decided to test our applications utility at the stage of creating it. First, we started with conversations with an employee of the ASPI Foundation to find out whether such a need really exists. Then we created prototypes of applications and verified their appearance and usefulness thanks to the opinions of target groups, with which we had the opportunity to meet and talk about their expectations in relation to our product. As a result of this meeting we received positive feedback with some remarks and suggestions concerning the look of an app.

At this point we are going to call a new meeting with ASPI Foundation and target groups to present the finished product.

1. Conclusions and perspectives

The solution we present provides a powerful affective computing tools to detect and distinguish emotions. The efficiency of this solution depends on an application area.

Because of a nonideal accuracy our application cannot be used as a security system or an artificial psychologist, but still it can be integrated even to those if we increase a dataset and add additional functionalities.

However it is a perfect tool which can help to teach children with autism to distinguish emotions in an interactive and simple way.

1. Acknowledgements

Full gratitude to the supervisor of our team dr hab. inż. Piotr Napieralski

Many thanks to the contributors for making the dataset used public. The dataset is a subset taken from:

*Keggle Face Dataset*

<<https://www.kaggle.com/c/face-recognition2>>

Michael J. Lyons, Miyuki Kamachi, Jiro Gyoba.

Japanese Female Facial Expressions (JAFFE), Database of digital images (1997).

Information about the JAFFE database is available at:

<<http://www.kasrl.org/jaffe.html>>

References

1. **Hariharan, A., & Adam, M. T. P.** (2015). *Blended Emotion Detection for Decision Support.* IEEE Transactions on Human-Machine Systems*, 45(4), 510–517.*doi:10.1109/thms.2015.2418231

<<https://sci-hub.se/https://ieeexplore.ieee.org/abstract/document/7090962/>>,

1. **Kim, D., Kim, J., Lee, E., Whang, M., & Cho, Y.** (2011). *Interactive emotional content communications system using portable wireless biofeedback device.* IEEE Transactions on Consumer Electronics, 57(4), 1929–1936.

<<https://sci-hub.se/https://ieeexplore.ieee.org/abstract/document/6131173/>>

1. **Rajalakshmi Kandaswamy** (2017) *Autism Cure: Neurological Understanding of the Brain-Gut-Energy Fields Consciousness Quartet Reveals the Solution*. Nessa Journal of Neurology and Neurodisorders.
2. **Magda Kowalska and Monika Wróbel** (2017): *Basic Emotions*. Faculty of Psychology, University of Social Sciences and Humanities, Warszawa, Poland Institute of Psychology, University of Lodz, Łódź, Poland
3. **Jeffrey F. Cohn, Tomas Simon Kruez, Iain Matthews, Ying Yang, Minh Hoai Nguyen, Margara Tejera Padilla, Feng Zhou, and Fernando De la Torre**: *Detecting Depression from Facial Actions and Vocal Prosody*.
4. **D Erhan, PL Carrier, A Courville, M Mirza, B**

**Hamner, W Cukierski, Y Tang, DH Lee, Y Zhou, C Ramaiah, F Feng, R Li,**

**X Wang, D Athanasakis, J Shawe-Taylor, M Milakov, J Park, R Ionescu,**

**M Popescu, C Grozea, J Bergstra, J Xie, L Romaszko, B Xu, Z Chuang, and**

**Y. Bengio:** *Challenges in Representation Learning*: *A report on three machine learning contests.* I Goodfellow,. arXiv 2013.

1. Google Vision, *“AI & Machine learning products”*

<<https://cloud.google.com/vision/>>