Computer Vision for detection of body expressions of children with cerebral palsy

Cristhian Rosales, Luis Jácome, Jorge Carrión, Carlos Jaramillo, Mario Palma
Systems Engineering Career
Universidad Nacional de Loja
Loja, Ecuador
cerosalesr@unl.edu.ec, lrjacome@unl.edu.ec, jorgetulio.carrion@unl.edu.ec, litosjc2013@gmail.com, mapalma@unl.edu.ec

***Abstract*—This article is the result of an investigation to improve the communication with a case study that suffers Cerebral Palsy through the use of Computer Vision. At present, there is a 15% of the world's population that suffers some form of disability which prevents them from complying with the common activities of a normal person in a social environment. The technology can help to facilitate the implementation of processes that support people with special needs to improve their lifestyle; in this project the main objective is to improve the communication with the patient in order to facilitate the patient care. For conducting the investigation it was necessary the development of a prototype that detects body expressions using the OpenCV library with the Python programming language. The results are promising because the computer vision system is able to detect with high accuracy the following body patterns: headache 77%, happiness 75%, hunger 82%, fear 88% and recreation 77%. Finally, when a body pattern is detected it is communicated to the patient's caregiver through a mobile application.**

Keywords— Computer Vision, Machine Learning, Cerebral Palsy, Artificial Intelligence, OpenCV

I. Introduction

Cerebral Palsy (CP) is the main cause of disability in children which prevents the development of movement and posture of the child who suffers it. "The motor disorder of cerebral palsy is frequently accompanied in sensory, cognitive, communication and perceptual disorders of conduct, or by epilepsy" [1].

The CP is a common global problem so there are some families who have one member with special disabilities, having an incidence of 2 to 2.5 cases per 1,000 live births [2]. Several of the children born under this condition do not communicate easily due to several factors, what comes to hinder direct understanding of the needs of the child; resulting in stress for both the family and the child with CP.

Since there are several types of cerebral palsy (CP), it can affect in different ways on each person who possesses it; CP is not as a disease but as a set of syndromes [3], so it is advisable to investigate a specific case to get ideas and get the necessary conclusions for the present project.

A. Types of cerebral palsy

In order to classify the types of cerebral palsy, the most characteristic aspect of the syndrome is taken into consideration: the neuromotor manifestations, in which the following criteria apply:

- *Topographic criteria:* It refers to the anatomical area.
- *Nosological criteria:* In reference to the neurological symptoms with regard to the muscle tone, the features of the movements, the balance, the reflexes and the postural patterns.
- *Funtional-motor criterion:* Referred to the overall degree of the neuromotor affectation.

B. Related work

Many efforts have been done to improve the quality of people's life using computer science, among them, we have:

In [4], authors refer to works of vision and speech stimulation for children with cerebral palsy. The survey summarizes researches of computer science, neural networks, electrical engineering and psychology. It also contains guidelines for construction of vision and speech stimulation tools which support for young and visually inattentive children.

A Vision based Interface is presented in [5] in which it detects and tracks the movement of the hand, foot or head of the user. The system is implemented for the interaction between the children and computer. The evaluation of the system is based on the HAAT (Human Activity Assistive Technology) model.

In [6], they built up a prediction method based on video, the system is projected to be less intrusive and cheaper than other solutions. It works in the following way, first, the motions of different body parts are separated, then, the motion features are extracted and used to classify infants to healthy or affected. The results show that visually obtained motion data allows detection of cerebral palsy, being as accurate as the state-of-the-art detection using electromagnetic sensor data.

The "Camera Mouse" system presented in [7], has been developed to provide computer access for people with severe disabilities. The system tracks the computer user's movements with a video camera and translates them into the movements of the mouse pointer on the screen.

In [8], they developed a visual monitoring system that passively observes moving objects in a site and learns patterns of activity from those observations. The system is based on motion tracking, camera coordination, activity classification and event detection. The system is useful for classifying sequences, as well as individual instances of activities in a site.

In [9], converted a video into a sequence of frames which is used for prediction of human behavior by correlating frames. The system analyzes faces that are detected determining the ROI of the expressions (Region of Interest), it also uses Viola-Jones patterns, and the AdaBoost method for filtering colors to detect people's faces, noses, eyes, mouth, or upper body. These parts have higher entropy for emotion detection.

A system that recognizes three facial expressions is proposed in [10]; the Geometric feature approach is used for feature extraction, the system uses the Multilayer Perceptron (MLP) neural network algorithm with backpropagation for classification; the facial expressions to be detected are neutral, happiness, and surprised. The recognition rate achieved is 93.33%.

A facial emotions recognizing method is presented in [11]; it is based on ASM (Active Shape Model) for facial feature extraction, and Radial Basis Function Network (RBFN) to recognize the facial emotion expressions with accuracy of 90.73%.

In [12], they introduce a tree-augmented naive-Bayes (TAN) classifier that learns the dependencies between facial features; it also provides an algorithm for finding the best TAN structure. The results show that using the TAN structure provides significantly better results than using simpler NB-classifiers.

In [13], they propose the concept of manifold of facial expression based on the observation of images of a subject's facial expressions that define a smooth manifold in the high dimensional image space. They use active wavelet networks (AWN) on the image sequences for facial feature localization, locally linear embedding (LLE) and Lipschitz embedding are used for classification.

A Cauchy Naive Bayes classifier is introduced in [14] which uses the Cauchy distribution as the model distribution, and provides a framework for choosing the best model distribution. The results show that the Cauchy distribution typically provides better results than the Gaussian distribution.

In [15], it is presented a new approach for facial expression decomposition. The higher-order singular value decomposition (HOSVD), in which the system learns the expression subspace and the person subspace from a corpus of images showing seven basic facial expressions. The algorithm is able to classify the given image into one of the seven basic facial expression categories, and then other facial expressions of the new person can be synthesized using the learned expression subspace model. The system is also able to recognize simultaneously face and facial expressions as a result of facial expression decomposition.

In [16], they developed a computer vision system that automatically recognizes individual action units or action unit combinations in the upper face using hidden Markov models (HMMs). They used three approaches to extract facial expression information: 1) facial feature point tracking 2) dense flow tracking with principal component analysis (PCA) 3) high gradient component detection. The recognition results of the upper face expressions are: feature point tracking 85%, dense flow tracking 93% and high gradient component detection 85%.

A facial expression recognition is presented in [17], it is based on a previous work of the development of a multimodal animated avatar. In both cases the work is oriented to FACS (Facial Action Coding System) which enables them to be easy integrated. The resulting animated system is able to maintain nonverbal and verbal bidirectional communication, enhancing the interaction with users.

In [18], they present a framework for real time face recognition and face emotion detection, using the open source computer vision (OpenCV) library and machine learning with python. Machine learning algorithms are used for recognition and classification of different classes of face emotions.

An emotion recognition system is presented in [19], that utilizes Histograms of oriented gradients (HoG) features and Support Vector Machines with Radial Basis Function kernels for learning the emotion patterns. The overall classification performance of the emotions detected is 70%.

In [20], they describe an emotion recognition system based on facial expression that is based on three steps: face detection, facial characteristic extraction and facial expression classification. The classification step is based on SVM method (Support Vector Machines), the system has an emotion recognition rate higher than 90% in real time. This approach is used to control a music player based on the variation of the emotional states of the user.

Regarding the present project and seeing the various existing problems, we have considered the proposal of a prototype that will help to improve the understanding and communication with children that suffer cerebral palsy, which raises the following hypothesis of research:

Can we improve the communication with children who have cerebral palsy utilizing a computer vision prototype of body patterns detection?

II. Materials and Methods

First, we have to highlight that the present research project was approved by the academic commission of the Systems Engineering career of the "Universidad Nacional de Loja". On the other hand, the main motivation for conducting the present project is the fact that one of the authors is a close relative of the patient, having an enormous interest of improving the patient’s quality of life.

At the start of the project, it was carried out an interview with the patient's representative in which we asked if she would like to understand better what the patient wants to express by using a mobile application; the answer was positive and the representative mentioned that the problem itself is that they do not understand what the child wants or feels. The interview and the test results were validated and endorsed with the signature of the patient's representative.

The patient’s diagnose prescribed by the specialized basic education school "Don Bosco" of the Zaruma City describes that the patient has infantile cerebral palsy of type ataxica (nosological) associated with low vision. The patient currently is 8 years and 6 months, her height is 1.16 meters with a weight of 18 kg; having a percentage of disability of 92%.

The patient presents several specific problems mentioned below:

- The caregiver does not directly understand the needs of the child, so it is difficult to please her.
- Difficulty vocalizing the words what prevents the child from communicating with her family.
- Language therapies require much time and effort so the child is stressed and not cooperates.
- The body movements are awkward so it is difficult to grab objects with precision or perform a particular action.
- Does not have a tool to express herself what creates stress in the family and makes it difficult to take care of her.

A. Skills consolidated of the case study

Through the report of consolidated skills of the early initial stimulation and language test held by the school for children with disabilities "Don Bosco" of Zaruma city, in the Table I, it can be seen the different evaluation parameters for the case study

TABLE I. EVALUATION PARAMETERS

Evaluation parameters	
Parameter	***Meaning***
Self-assistance	Being able to do things by herself
Socialization	The patient relates with others
Language	Understability of the spoken language
Knowledge	Being able to remember things
Gross motor skills	Coordination of the arms, legs and other large body parts and movements
Fine motor skills	Smaller movements that occur in the wrists, hands, fingers, feet and toes

These parameters are qualified by the following levels of achievements:

- A: reached skill.
- B: skill performed with aid.
- C: skill based on attempts.
- D: skill based on strength or not reached.

The results of this report are represented in the following way:

1) Self-assistance: The results obtained concerning auto evaluated assistance of the girl with cerebral palsy, demonstrate that there is an achievement of 39% achieved by itself, which tells us that it is not sufficiently independent to make their personal needs. Fig. 1. shows the results achieved by the patient.

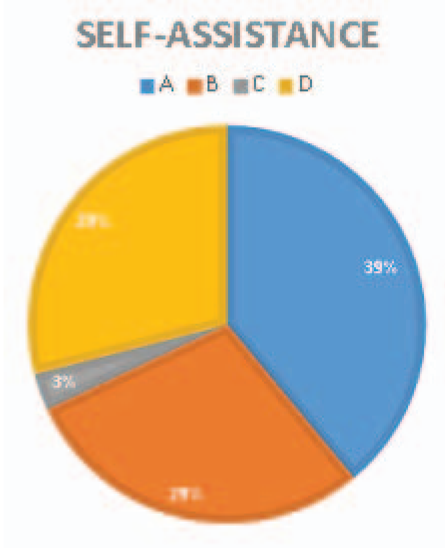

Fig. 1. Results of self-assistance achievement of the patient

2) Socialization: In this parameter the results indicate the girl with CP has a 52% of skill achieved in terms of their level of socialization. This shows that her relationship with third parties is essential for her personal development.

3) Language: In terms of the skills related to the language, the patient with CP shows 41% of achievement in this aspect, and the report uses a 20 word vocabulary with a dexterity based on attempts. The type of CP that the patient has makes it difficult to express herself through speech with her environment.

4) Knowledge: The patient gets 48% of achievement accomplished by her and 48% of achievement that is assigned to unreached skills. Comparing these results it is concluded that the patient in consequence of her type of CP does not have a good development of this parameter.

5) Gross motor function: The gross motor area refers with the body position changes and the ability to maintain the balance, in which the patient presents 25% achievement; but by the type of CP presents 53% of skills not achieved for this parameter.

6) Fine motor: Fine motor is related to the coordination of fine movements which is movements that require precision, for example movement of eyes and hands. In this parameter the patient presents 58% of skills obtained by herself; which helps us to get information of certain movements that seek to communicate a need or an emotion.

B. Expressions to detect

According to the psychomotor development report issued by the patient's doctor; in which it says that after an assessment of physical therapy and early stimulation, it is considered that the patient has presented improvements during her therapeutic process of rehabilitation in schematic behaviors of development like:

- Language: improve understanding.
- Gross motor: increased tonicity, body stability, and slight head control.
- Social: knows people by voice, and is integrated into games.
- Cognitive: knows body scheme and notions of up and down.
- Coordination: carries out voluntary movements such as: manipulate objects with greater precision and the passes from one hand to another, coordinates marching alternating feet, climbs and goes down stairs with help, introduces large objects to boxes, works with plasteline, etc.

Through the observations made, it has been possible to distinguish certain expressions that the patient tries to communicate, and which are more likely to be recognized through the use of Computer Vision. The patterns of the body expressions are explained in the Table II.

TABLE II. BODY EXPRESSIONS

Body expressions	
Movement	***Wishes to express***
Smile on the face	Happiness
Shrink the body	Fear
Stick out the tongue when she wants to drink or eat something	Hunger
It supports the hand on the head	Headache
Aim with the hand when she wants an object, or wants to leave the house	Recreation

C. The prototpye

This research seeks to obtain body patterns of the patient showing an emotion, need or pain; to do so, it was necessary to build classifiers based on machine learning and computer vision.

The prototype was developed using the Iconix methodology which helped to get the analysis of requirements, the preliminary analysis, design and implementation; moreover, it was necessary to build the intelligent agent responsible for capturing the body expressions through an IP camera, and the mobile application whose function is to inform about those body expressions. The Table III describes the software used for building the prototype.

TABLE III. SOFTWARE USED IN THE PROJECT

Software used in the project		
Software	***Description***	***Version***
Windows	Operating system	10 pro
LiClipse	Editor based on Eclipse for the development of the intelligent agent	2.5.4
OpenCV	Library used for computer vision	3.0.0
Python	Programming language used to program the intelligent agent	2.7.11
Android Studio	Integrated Development Environment used for building the mobile application	1.5.1

In order to build the classifiers it was necessary to capture a large number of images; for doing it, we used an IP camera with the following features: IHOMECAM brand, ICAM-608 model and 720P(HD) Megapixel resolution. The Table IV shows the number of images used for building the classifiers.

TABLE IV. NUMBER OF IMAGES USED FOR BUILDING THE CLASSIFIERS

Number of images used for building the classifiers			
Body pattern	***# of positive images (training)***	***# of negative images(training)***	***# of images (test)***
Headache	161	1200	50
Happiness	124	1200	50
Hunger	161	1200	50
Fear	211	1200	50
Recreation	209	1200	50

The characteristics of the positive images focus on the patient and her respective body expressions that we want to detect; on the other hand, the negative images have common objects that surround the environment of the patient.

In order to differentiate the body patterns, we used the Boosting technique called AdaBoost in which a cascade of classifiers, built by stages, forms an adapted classifier, that will make higher and better level detections.

In the Viola-Jones algorithm, AdaBoost algorithm is able to choose from a large set of filters, the Haar features, in order to select at any time which one fits better to do a satisfactory classification. The operation of the algorithm focuses on the idea of whether the sample being analyzed belongs to one type or another. For each sample is needed:

a) To train the classifier on the basis of a distribution example.

b) To compute the rates of error.

c) To compute the indices of weight of each sample.

d) To update the distribution of the initial samples.

This procedure is obtained after some iteration where the Haar classifiers are tested for each of the samples, the classifier separates positive and negative samples. First, a detection is generated with a classifier in which poorly classified elements increase their weight in order that the next involved classifier give more importance to the classification of the elements with higher weight. After working with several classifiers a single classifier is obtained as a combination of the previous classifiers that can classify correctly with an acceptable value of error.

To know the performance of classifiers the Confusion matrix is used, which enables to see the level of confusion of a classifier. The tests of a classifier are done by binary classification that means that the result of the classifier will tell if there is or not what is wanted to be detected.

D. *Limitations of the study*

The use of the prototype is limited to illuminated places in order to increase the performance of the classifiers in the detection of expressions of the case study. The minimum accepted resolution of the images is 300 pixels / inch; if the resolution is lower than the minimum resolution, the body expression will not be detected.

The evaluation of the body expressions and its data capture had a high level of difficulty due to the percentage of disability of the patient and the little control of movements or expressions that we want to detect using this prototype. Also, it is considerable the time involved waiting for the patient to make the body expressions we want to evaluate.

Fig. 2 describes the Architecture of the prototype.

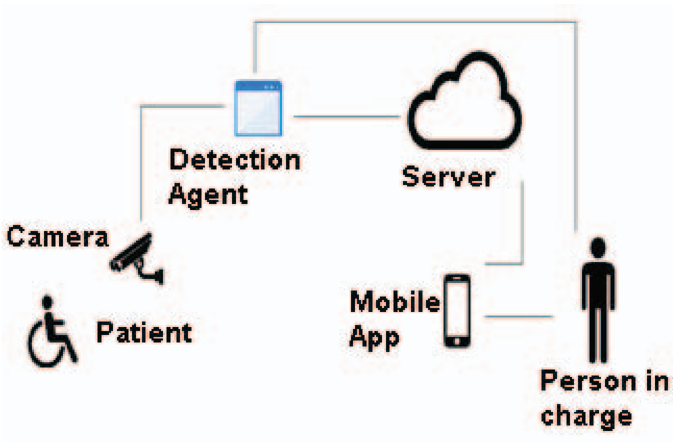

Fig. 2. Architecture of the prototype

III. RESULTS

For testing the performance of the classifiers we used the Confusion matrix to get the following metrics: accuracy, precision, sensitivity and specificity of each classifier for each body pattern.

Fig. 3 shows the different body patterns of the patient to be identified by the intelligent agent. The performance of the classifier is shown in the Table V.

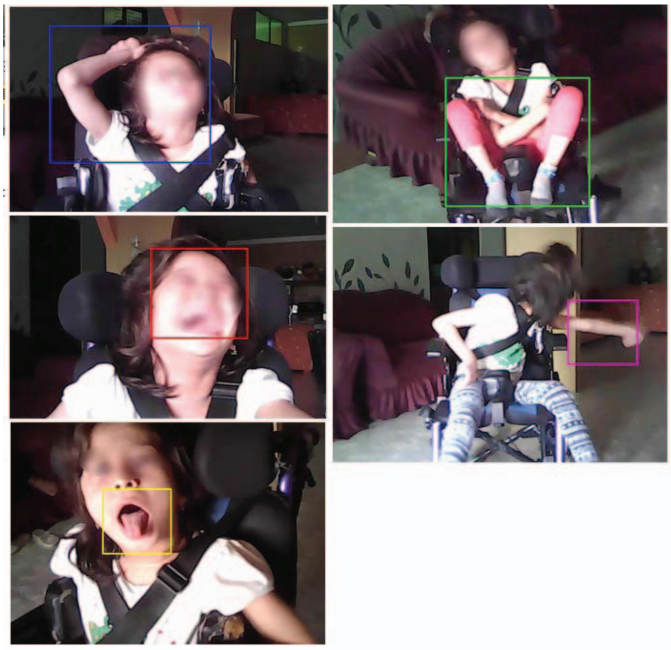

Fig. 3. Body patterns of the patient to be detected by the intelligent agent

TABLE V. HEADACHE EXPRESSION

	Headache	
Body pattern	***Metric***	***Percentage***
Headache	Accuracy	77%
Headache	Precision	100%
Headache	Sensitivity	30%
Headache	Specificity	100%
Happiness	Accuracy	75%
Happiness	Precision	85%
Happiness	Sensitivity	80%
Happiness	Specificity	58%
Hunger	Accuracy	82%
Hunger	Precision	85%
Hunger	Sensitivity	90%
Hunger	Specificity	62%
Fear	Accuracy	88%
Fear	Precision	97%
Fear	Sensitivity	87%
Fear	Specificity	90%
Recreation	Accuracy	77%
Recreation	Precision	100%
Recreation	Sensitivity	30%
Recreation	Specificity	100%

The use of several classifiers tends to increase the response time of the algorithm, so it is recommendable to utilize a machine with high processing characteristics or to reduce the size of the captured image. It is also important to obtain sufficient positive and negative images samples in order to improve the training of classifiers, reaffirming that the higher number of positive images, the better the performance of classifiers. Having a tool to select the area of interest on the positive images is helpful for the process of the training of classifiers, since it gives us the exact coordinates of the location of the object that we want to detect, decreasing significantly the need for processing. Finally, to prevent the people who take care of the patient, we develop a mobile application that sends alerts when a body expression is detected. This application is available for the Android operating system. Fig. 4 shows a screen of the mobile app.

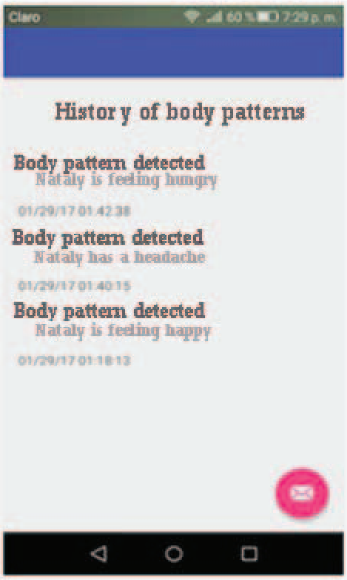

Fig. 4. Screen of an alert in the mobile application when a body expression has been detected

CONCLUSIONS

With the development of this prototype, it has been demonstrated that the communication with children with cerebral palsy can be improved by the use of Computer Vision. For this research we used a case study with CP to build a classifier for detection of body expressions using an intelligent agent that could report what the child with CP wishes to express. The results show high accuracy of the detection of body patterns, having headache 77%, happiness 75%, hunger 82%, fear 88% and recreation 77%. This prototype is a starting point that seeks to help people, especially children who suffer from a disability; by means of the use of technology to improve their lifestyle. As we can see, the results are promising and the applications of detection of body patterns using Computer Vision can be applied to an increasing number of fields and situations.

REFERENCES

[1] P. Póo Argüelles, "Parálisis cerebral infantil," 2008.

[2] I. C. Valdovinos Lopez, "Evaluación diagnóstica del niño con parálisis cerebral en el tercer nivel de atención," 2009.

[3] J. M. Valdez, Medicina : organo de la Sociedad Argentina de Investigación Clínica. La Sociedad.

[4] G. Sumathy and A. Renjith, "A survey of vision and speech stimulation for cerebral palsy rehabilitation," in 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014, pp. 1315–1319.

[5] M. González, D. Mulet, E. Perez, C. Soria, and V. Mut, "Vision based interface: an alternative tool for children with cerebral palsy," in 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, 2010, pp. 5895–5898.

[6] H. Rahmati, O. M. Aamo, O. Stavdahl, R. Dragon, and L. Adde, "Video-based early cerebral palsy prediction using motion segmentation," in 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2014, pp. 3779–3783.

[7] M. Betke, J. Gips, and P. Fleming, "The Camera Mouse: visual tracking of body features to provide computer access for people with severe disabilities," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 10, no. 1, pp. 1–10, Mar. 2002.

[8] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 747–757, 2000.

[9] S. Shakya, S. Sharma, and A. Basnet, "Human behavior prediction using facial expression analysis," in 2016 International Conference on Computing, Communication and Automation (ICCCA), 2016, pp. 399–404.

[10] J. V. Patil and P. Bailke, "Real time facial expression recognition using RealSense camera and ANN," in 2016 International Conference on Inventive Computation Technologies (ICICT), 2016, pp. 1–6.

[11] E. Setyati, Y. K. Suprapto, and M. H. Purnomo, "Facial emotional expressions recognition based on Active Shape Model and Radial Basis Function Network," in 2012 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA) Proceedings, 2012, pp. 41–46.

[12] I. Cohen, N. Sebe, A. Garg, M. S. Lew, and T. S. Huang, "Facial expression recognition from video sequences," in Proceedings. IEEE International Conference on Multimedia and Expo, vol. 2, pp. 121–124.

[13] Y. Chang, C. Hu, and M. Turk, "Manifold of facial expression," in 2003 IEEE International SOI Conference. Proceedings (Cat. No.03CH37443), pp. 28–35.

[14] N. Sebe, M. S. Lew, I. Cohen, A. Garg, and T. S. Huang, "Emotion recognition using a Cauchy Naive Bayes classifier," in Object recognition supported by user interaction for service robots, vol. 1, pp. 17–20.

[15] Hongcheng Wang and Ahuja, "Facial expression decomposition," in Proceedings Ninth IEEE International Conference on Computer Vision, 2003, pp. 958–965 vol.2.

[16] J. J. Lien, T. Kanade, J. F. Cohn, and Ching-Chung Li, "Automated facial expression recognition based on FACS action units," in Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 390–395.

[17] S. Marcos, J. Gómez-García-Bermejo, E. Zalama, and J. López, "Nonverbal communication with a multimodal agent via facial expression recognition," in Proceedings - IEEE International Conference on Robotics and Automation, 2011, pp. 1199–1204.

[18] K. M. Rajesh and M. Naveenkumar, "A robust method for face recognition and face emotion detection system using support vector machines," in 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), 2016, pp. 1–5.

[19] M. Dahmane and J. Meunier, "Emotion recognition using dynamic grid-based HoG features," in Face and Gesture 2011, 2011, pp. 884–888.

[20] F. Abdat, C. Maaoui, and A. Pruski, "Human-Computer Interaction Using Emotion Recognition from Facial Expression," in 2011 UKSim 5th European Symposium on Computer Modeling and Simulation, 2011, pp. 196–201.