INTELLIGENT TOOL FOR DEVELOPING STUDENT'S SOCIAL SKILLS

R. Kupriyanov[1], A. Uvarov[2], M. Makarova[3], P. Unruh[3], G. Vodopyan[3], I. Levin[4]

[1]Moscow City University (RUSSIAN FEDERATION)
[2]Higher School of Economics, FRC IC RAS, Moscow (RUSSIAN FEDERATION)
[3]ORT de Gunzburg School, St. Petersburg (RUSSIAN FEDERATION)
[4]Tel Aviv University (ISRAEL)

Abstract

Developing social skills is one of the basic aims of the contemporary school system. The process of mastering such skills by students is time-consuming for teachers. Video recording of the students' activities, followed by the analysis of the video and the corresponding future planning of the skills development can reduce the time of learning thereby increasing its effectiveness. However, in the personalized learning process, such a straightforward video recording has to comprise recording of each of the students, which is almost impossible. The paper proposes a solution to the above problem. The solution uses an intelligent machine learning tool. The proposed tool is based on real-time intelligent processing of the 360 video records of the students' cooperation activities. Such processing enables us to identify each student's video from the global video stream and to process it separately. It allows recognizing patterns of behavior that the student demonstrates during the exercise. The intelligent tool finds and highlights representative patterns that are appropriate and/or inappropriate according to the role the student has to perform during the activity. The intelligent video processing allows specifying sets of specific actions that each member of the group demonstrates according to his/her roles in the team. After finishing the activity, each student immediately gets personal feedback that reflects his/her success and failures during the activity. The proposed deep learning-based approach of pattern recognition was implemented as a software application. This application uses quite advanced hardware and software resources that enable the corresponding recording and processing. Nevertheless, students use conventional equipment (smartphone, laptop) to get feedback. Preliminary experimental results of our study are positive. They demonstrate encouraging prospects of integration of intelligent tools to support effective reflection of the students' teamwork in the elementary school.

Keywords: Social skills, elementary school, teamwork, machine learning.

1 INTRODUCTION

Cooperation in training, especially in small groups with positive interdependence of the participants, was first mentioned in the Talmud. After the works of E. Cohen, R. Johnson, D. Johnson and others [6,10], cooperative teaching techniques have become part of the everyday educational practice of schools in many countries of the world.

Today, providing schoolchildren with teamwork skills has become an indispensable part of general education. However, training schoolchildren for accepted teamwork standards, for successful communication, and for effective using of various templates for interacting with partners in solving various problems remains a complicated and time-consuming task. Such training highly time-consuming task, which requires extraordinary efforts of participants during several years.

According to the experience of educators of the ORT de Gunzburg School (https://portal.ort.spb.ru), it is especially difficult to provide a group processing i.e. "how effectively group members are working together and how the group's effectiveness can be continuously improved" [10]. In a class of around thirty students, where 7 to 8 groups of students work simultaneously, the teacher has not enough time to concentrate on each group. Therefore, the teacher is unable to evaluate effectively the teamwork and to produce correct pedagogical recommendations for each of student. The idea to transfer this work to the group members, especially in the primary school, is unproductive since students cannot always remember and recall individual situations that occur during the teamwork, they cannot assess correctly their actions. Using the video, which allows the students to see their behaviour during the work of the group, is also not practical. Viewing the footage video by the group members takes too much time; moreover, analysing individual situations by the group members is ineffective enough without external teacher's assessment. Our study was initiated by the emergence of the recent achievements in the field of AI. These two facts together have opened the way for effectively analyzing the video stream in real-time to highlight the desirable fragments. Our main idea is to develop an intelligent assistant (an educational robot-consultant named UROK), as an additional intelligent team member. The main function of this assistant is to support the students during group processing. Physically, such an intelligent assistant (UROK) is equipped with an acoustic system with a video camera with a circular overlook. The assistant transmits the recorded data for processing in a neural network connected to cloud libraries. UROK is located in the center of the table at which the team works. Members of the team can communicate with UROK using its voice interface (as with another member of the group), as well as using a mobile-specific application. The ability of UROK to learn provides the group members an opportunity to teach their robot putting themselves into a teaching position. We consider this teaching activity is an innovative and highly important component for mastering the students' teamwork skills.

Our paper describes the first steps of the study. The ultimate goal of the study is to develop a new approach for teaching social skills based on using the learning robot as an intellectual tool and to study pedagogical effectiveness of the proposed approach that, according to our hypothesis, would help in teaching elementary school students effective communication and teamwork skills.

2 BACKGROUND

The idea of creating an intellectual UROK assistant for developing teamwork skills in students was initiated by emerging achievements in the field of machine learning. These significant achievements are based on using neural networks recognition of non-verbal signals and response to such signals during the teamwork. In turn, the above allows assessing a students' involvement in the educational process using analysis of their gazes, positions of their head and body. Further, the idea completed by using the computer analysis of interactions of schoolchildren working in small groups. We also relied on the research on machine learning, which carried out at Moscow City University [12].

As a research framework, we consider using a data-driven social signal prediction framework [7] that may be used to describe sub-tasks considering various channels of input and output social signals in teamwork communication. We also relied on the well-known works [5,1] devoted to the construction of neural networks for analyzing video records of the lessons to assess the involvement of students in the educational work based on directions of their gaze. We relied as well on [14] for tracking the eye and the pupil during the time of reading hypertext documents to produce data on the degree of reading comprehension. An interesting approach based on using neuron networks for evaluation of attention in a class during a frontal lesson was described in [9]. Such an approach allows tracking the angle of the head rotation and inclination, the position of the tip of the mouth, and of the eyes.

An important parameter for analyzing a student's involvement in any academic work, in addition to the body position, is the student's body movement. In [13] suggested using a system of cameras that allows tracking not only the trajectory of the student's gaze but also how the student moves during a traditional training activity. An interesting and successful experiment describes in [3], where an inexpensive web camera and a trained neural network allow recognizing a nonverbal student's behavior, his/her facial expression, head, and eye movements and emotions were automatically recognized and evaluated.

There are a number of studies on machine analysis of students' interaction while working in small groups. One of the experiments [15] takes into account the trajectory of gaze, body position, and exchange of oral conversations, which affect the productivity of the group interaction. In [4,8] the learning of students in small groups is analyzed. The student's involvement in the teamwork was evaluated in various ways. The computer vision monitored the frequency of schoolchildren's hands' movements, the tilt of the body, the duration of gazing the monitor, the students' positive and neutral emotions were recognized and analyzed. The researcher was able to achieve pattern recognition with an accuracy of more than 90%.

In the experiments reported in [4,17], the speech of participants in teamwork was recognized and analyzed. The collected data provided an assessment of the degree of cooperation in the team and a summary of the content of the participants' conversations.

Studies of machine analysis of the teamwork (both of its process and results) are rapidly multiplying. In our study, it was important to ensure that modern methods of machine learning allow successfully constructing neural networks for analyzing both individual and teamwork of schoolchildren. Analyzing relevant literature indicates that capabilities of publicly affordable round-robin video cameras are sufficient in forming a video stream for recognizing behavioral patterns of students in the teamwork. At the same time, the question "whether the extent to which the available technological means are sufficient for real-time analysis of the video stream" requires a special study.

3 METODOLOGY AND EXPERIMENT

The first step of our study comprised two experiments. In the first experiment, when the students trained a robot, we have studied how students of the elementary school interact with the robot, how they overcome the effect of novelty, how serious they are when teaching the robot, how they embed it in the teamwork. In the second experiment, we studied various organizational and technical problems that occurred when recording videos of the group, handling the streaming video and recognition of special features of the students' behavior in the teamwork.

3.1 Robot training

The experiment involved 16 children of the first grade of a primary school (6 - 7 years old). Their parents agreed to their kids' participation and submitted photos of the children to the school, for pre-training of the neural network. Parents of other children of the class didn't agree to submit such photographs. The prototype of the UROK robot was a local computer running MS Windows 10. On the upper edge of the computer's monitor, two webcams with a resolution of 640x480 were attached. One camera was used for face recognition, and the other - for recording what was happening (the records were saved on the local disk of the computer.) This arrangement allowed eliminating a conflict of two tasks over the computing resources. For face recognition, two pre-trained neural networks were used, cooperating with Open Computer Vision libraries for the Python language. The first neural network distinguished a face from the background, while the second one compared the face with a known sample. The first task was solved quite reliably (accuracy over 90%). The second task was solved less reliably due to the low resolution of the photos.

The scenario of the experiment is as follows. After the robot recognizes the face of the child standing in front of the camera, he says “hello” with pronouncing the name of the child, and asks if he recognized the child correctly. For this, a library of speech synthesis is used. The child can respond using the keyboard or a microphone, which was also used for performing a task "remember the name of a new face".

The participants of the experiment were divided into four teams of four members each. The children had a task to train the robot to congratulate the team on the upcoming New Year holiday. The reinforced learning was used for this purpose. The experiment comprised a number of series of two exercises. In the first exercise, the teacher showed to the children that robot may incorrectly call somebody (for example, when somebody's head was slightly turned before the camera). Each team tried to teach the robot to properly name/call each participant, no matter whether the participant looks into the camera directly or not. In the second exercise, the children also taught the robot to recognize the members of their team under various emotional states (laughing, sad, angry, etc.). Each team was given 25 minutes to complete the series. The teamwork was recorded using video recording means, for subsequent analysis. The results have shown that elementary school students interact quite rationally with the robot and solve the task successfully. Signs of the novelty effect ceased to be observed after just a few exercises. The children actively discussed various features of the robot training, proposed their own original ideas to improve the training. Actually, they considered the trained robot as a full member of their team and prepared the robot to participate in a New Year Eve performance.

3.2 Recognizing the student’s behavior in a group work

The second experiment was aimed to answer two research questions:

1. Do modern face recognition algorithms successful in the task of primary school children's teamwork analysis?
2. Is it appropriate to use neural networks, which have been trained on general datasets, to analyze primary school children's emotions?

We used the following methodology in order to respond to these questions. A number of 360-degree videos of primary school children group work were recorded. The videos were processed with AI-tools in order to recognize students' faces, emotions, and behavior features. A number of experts in primary school children psychology and pedagogics analyzed the results to check the revealed patterns of team behavior.

We used two special software applications based on the machine learning algorithms [12]. The first application was focused on the recognition of primary school children's emotions. The second one - on the recognition of primary school children's behavior. The first application is based on Microsoft Azure Cognitive Services (MS ACS) and allows recognizing age, gender, head pose, facial landmarks, indicators of lipstick, glasses, mustache, sideburns, beard, and recognize emotions such as anger, contempt, disgust, fear, happiness, neutral, sadness and surprise [16; 12].

In order to use the MS ACS, automatic preprocessing of the initial video algorithm was implemented. This application cuts a video into separate frames at a specified frequency sends frames to MS ACS and receives the corresponding recognized features for future use. The result of the students' face and emotions recognition is presented in Figure 1.

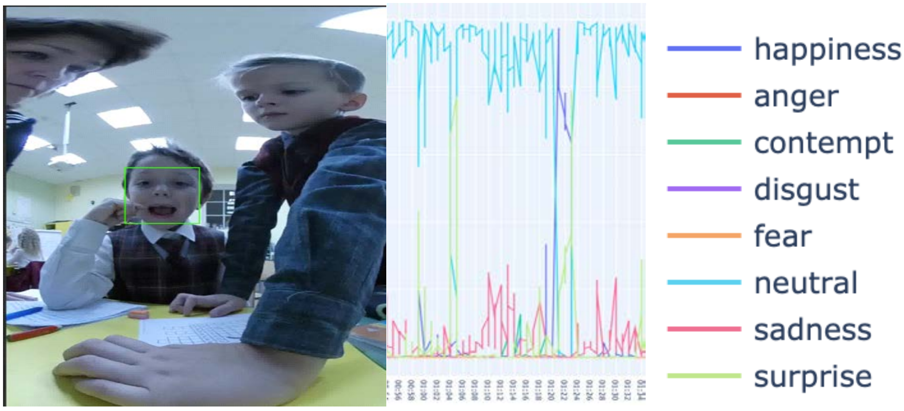

Figure 1. The result of the students' face and emotions recognition.

Figure 1 demonstrates that MS ACS successfully recognized a child's face. We analyzed all recorded videos and found the same effect. The second application was used for recognizing and tracking the students' eye gazes and head poses. Similarly, the emotion recognition module for the preprocessing of the initial video was used. However, as the task of 360 degrees video processing is a high-loaded process, we implemented an additional preprocessing algorithm of segmenting a high-resolution image to reduce calculation time in subsequent functional blocks. As a result, the developed application has become capable to determine the person's area with quite a high accuracy and speed up to 10 fps..

The application allowed us to find and to track the head pose; the facial landmarks (eyes, nose, and mouth areas) and the eye gazes. The recognized child's eye directions and the head position were used to determine the eye gazes of each child. Examples of the students' eye gaze and head pose recognition are presented in Figure 2.

Figure 2. The students' eye gazes and head pose recognition

4 CONCLUSIONS

Experimental results of our study demonstrate that the idea of teaching the robot to primary school students turned out to be quite workable. Since the primary goal of elementary school is to form a full-fledged activity for students, a robot taught by schoolchildren may become one of the main tools for children's independent activities (including writing, counting, etc.). Students of the elementary school efficiently interact with the robot. The novelty effect disappears quite fast. The students actively discuss the features of the robot training by their own initiative, as well as look for ways to intensify the training.

Preliminary results of our study show that known modern algorithms for recognition of the face, head poses, and eye gaze can be successfully used for the students' teamwork analysis. The model and the algorithms implemented in the MCU's software applications provide accuracy and speed of the 360-degree video stream processing sufficient for the analysis of students' teamwork leaning. In contrast, we have found that known neural networks, which were developed and trained on general datasets are incapable to analyze emotions of primary schoolchildren. It looks that the difference between a children’s and an adult's emotions are pretty significant. This problem could be solved, for example, by enriching a pool of the neural network training samples or by training a special neural network aimed at the task of the recognition of primary school children's emotions.

Summarizing, the newly proposed approach for developing teamwork social skills in elementary school is promising. The positive preliminary results of our study open the way for future research and development. We believe that the proposed approach will appear fruitful both in the theoretical and practical spheres.

REFERENCES

[1] Aung A., Ramakrishnan A., Witehill J. Who are they looking at? Automatic Eye Gaze Following for Classroom Observation Video Analysis, *Proceedings of the 11th International Conference on Educational Data Mining*, pp.252-259, (2018).

[2] Avramides K., Cukurova M., Luckin R., Mavrikis M., Spikol D., An Analysis Framework for Collaborative Problem Solving in Practice-based Learning Activities: A Mixed-method Approach, *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge*, (2016).

[3] Behera A., Canning S., Fang H., Keidel A., Matthew P., Vangorp P., Associating Facial Expressions and Upper-Body Gestures with Learning Tasks for Enhancing Intelligent Tutoring Systems, (2020). https://link.springer.com/article/10.1007/s40593-020-00195-2

[4] Bett M., Finke M., Stiefelhagen R., Waibel A., Meeting browser tracking and summarizing meeting, *Proceedings of Broadcast News Transcription and Understanding Workshop*, pp. 281 – 286, (1998).

[5] Bidwell J. Classroom Analytics: Measuring Student Engagement with Automated Gaze Tracking, (2019). https://chili.epfl.ch/page-92073-en-html/page-92240-en-html/page-126025-en-html/

[6] Cohen, E. Designing groupwork, *New York: Teachers College Press*, (1986).

[7] Cikara M., Joo H., Sheikh Y., Simon T. Towards Social Artificial Intelligence: Nonverbal Social Signal Prediction in A Triadic Interaction*, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10873-10883, (2019).

[8] Farzaneh A. H., Kim Y., Zhou M., Qi X. Developing a Deep Learning-Based Affect Recognition System for Young, *AIED 2019 Chicago*, (2019).

[9] Gerjets P., Draghetti L., Kasneci E., Kubler T., Santini T., Trautwein U., Wagner W., Automatic Mapping of Remote Crowd Gaze to Stimuli in the Classroom, (2019). http://www.ti.uni-tuebingen.de/uploads/tx_timitarbeiter/etel2017-classroom_camera-ready_01.pdf

[10] Holubec E., Johnson, D. W., Johnson, R., Cooperation in the classroom (9th ed.). Edina, MN: Interaction Book Company, (2013).

[11] Kalenichenko D. and Philbin J., Schroff F., FaceNet: A unified embedding for face recognition and clustering, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 815-823, (2015).

[12] Kupriyanov R. B., Application of computer vision technologies for automatic data collecrion about emotions of students during group work, *«Informatics and Education»*, in press 2020.

[13] Raca M. Classroom Attention Monitoring, (2019). https://chili.epfl.ch/page-92073-en-html/page-92240-en-html/page-126025-en-html/.

[14] Rajendran R., Kümar A., Carter K., Levin D., Biswas G. Predicting Learning by Analyzing Eye-Gaze Data of Reading Behavior, *Proceedings of the 11th International Conference on Educational Data Mining*, pp.455-462, (2018).

[15] Ravenell M., Reilly J., Schneider B., Exploring Collaboration Using Motion Sensors and Multi-Modal Learning Analytics, *Proceedings of the 11th International Conference on Educational Data Mining*, PP.333-341, (2018).

[16] Taigman Y., Yang M., Ranzato M. and Wolf L., DeepFace: Closing the Gap to Human-Level Performance in Face Verification, *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1701-1708. (2014). https://research.facebook.com/publications/480567225376225/deepface-closing-the-gap-to-human-level-performance-in-face-verification/

[17] VanLehn K., Viswanathan S. A., Collaboration Detection that Preserves Privacy of Students' Speech, *Proceedings AIED 2019*, (2019).