Assisted therapeutic system based on reinforcement learning for children with autism

Minjia Li*, Xue Li*, Lun Xie#, Jing Liu#, Feifei Wang# and Zhiliang Wang#

*aSchool of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing, China; #Peking University Institute of Mental Health, Beijing, China

ABSTRACT
Assisted therapy is increasingly used in autism spectrum disorders (ASD) for improving social interaction and communication skills in recent years. A lot of studies have proven that the form of interactive games for therapy has a good effect on children with autism. Thus, our study provided an assisted therapeutic system based on Reinforcement Learning (RL) for children with ASD, which has five interactive subgames. As is well known, it is necessary to establish and maintain compelling interactions in therapeutic process. Therefore, we aim to adjust the interactive content according to the emotions of children with autism. However, due to the atypical and unusually differences in children with autism, most systems are based on off-line training of small samples of individuals and online recognition, so the existing assisted systems are limited in their ability to automatically update system parameters of new mappings. The integration of RL and Convolutional Neural Network (CNN)-Support Vector Regression (SVR) was used to deal with the updating online of prediction model’s weights. The normalized emotion labels were evaluated by the therapists. Eleven children with autism as subjects were invited in this experiment and captured facial video images. The experiment lasted for five weeks of intermittent assisted therapy, and the results were evaluated for the system and the therapy effect. Finally, we achieved a general reduction in the root mean square error of the model prediction results and labels. Although there is no significant difference in Social Responsiveness Scale (SRS) scores before and after assisted therapy (p value = 0.60), in individual subjects (Sub. 1, Sub. 2 and Sub.3), the SRS total score is significantly reduced (Average drop of 19 points). These results demonstrate the effectiveness of prediction model based on RL and show the feasibility of assisted therapeutic system in children with autism.

KEYWORDS
Assisted therapeutic system; reinforcement learning; convolutional neural network; support vector regression

Introduction
In recent years, many assisted systems for health care have been clinical interventions [1]. As a new therapy mode in ASD rehabilitation, assisted system is applied to autism children’s coordination and interaction abilities training for communicating with neurotypical people naturally [2].

Nowadays, many therapists encourage children to learn through games [3], and treat the toys as an open interactive partner, such as social robots. The AuRoRA (Autonomous Robotic Platform as a Remedial Tool for Children with Autism) project is the first assistant robot research project for children with autism. Patients could learn and master social skills, such as eye contact, behavior mimicry, joint attention, and body avoidance in Human-Robot Interaction (HRI), to enhance the cognitive affective abilities [4,5]. The IROMEC (Interactive Robotic Social Mediators as Companions) project [6] was used in building social relationship with autism children gradually launched. Ben Robins and Ester Ferrari’s research contributed to the development of the IROMEC robot, and promoted the communication from function interaction to emotion and awareness. A portable emotion detector for autistic children has been presented by K. G. Smitha et al. to ensure usability in on-the-move applications [7]. Besides, a non-negligible work was provided by Ognjen et al. [8] to achieve personalized robotic emotional perception using contextual and personal characteristics of autistic children, which provided a foundation for adjuvant therapy.

A number of studies have been devoted to the realization of emotional perception of children
through multiple modalities in an automatic mode. For example, Kim et al. [9] used audio dataset to assess children’s emotion in interactive process. Stefano et al. [10] used a set of 2D and 3D full-body movement features to classify emotions. Hui-Chuan et al. [11] proposed a facial expression-based emotion recognition method with transition detection. But most therapeutic systems in these studies were trained offline previously. They treat some new mappings (never happened in offline training) as outliers and predict the wrong results. Thus, in the case of realistic assisted therapy, offline sample collection cannot cover new mappings that occur during long-term use, and the systems don’t have the ability to automatically update to identify new particular samples.

The sensory overloads and social rejection of children with autism lead to depression and anxiety, and therefore, it is necessary to adjust interactive content based on emotion of children. In our study, using the participatory approach of therapists’ manual intervention, we established an assisted therapeutic system based on reinforcement learning [12] for children with autism called RLAT. The system could predict the ASD children’s emotion by facial expression and then adjust game strategy for enhancing the interactive effects. Based on the child’s facial expressions, behaviors, including stereotyped one, the therapist performs assessments intermittently of the child’s qualitative emotional state. The assessments are used as ground truth (the gold standard labels). The overall system contains two parts, namely, manual intervention and automatic mode as shown in Figure 1. The whole process can be regarded as reinforcement learning. Since SVR can complement the generalization ability of CNN even without a large number of samples [13], the combination of CNN and SVR is used as the prediction model. Moreover, the therapy adopts semi-autonomous mode, and the therapist could intervene in the experiment by the tablet for controlling the situation and improving the schedule.

The structure of the paper is organized as follows: We will describe the materials of RLAT in the section ‘Materials’, the methodology including CNN-SVR model, reinforcement learning process, and SVR weights adjustment in the section ‘Method overview’. The experimental results presented in the section ‘Results’, will be discussed in the section ‘Discussion’.

Materials

System design

The client (robot and tablet)-server architecture is implemented in this system. The robot used in this project is developed by our research group, which is a child-sized cartoon robot with a height of 40 cm and

![Figure 1. The assisted therapeutic system for children with ASD: The top panel is process of manual intervention. And the bottom panel represents the automatic mode.](image-url)
weight of approximately 5.6 kg. The robot has 10 DOF and 100 kinds of cartoon facial expressions, so it could express abundant behaviours and emotions for social skill therapy, and according to the emotional state and scores during the interaction, it shows several fixed expression patterns that have been set in advance.

The system of the robot and the tablet computer select the Android operating system, and use the socket communication through WiFi to complete the data interaction with the server, and finally the robot is fed to select the corresponding game difficulty and process according to the interaction result. Among them, the robot uses the Corten-A8 processor, the processor model S5PV210 (clocked at 1 GHz), and is equipped with PowerVR SGX540 parallel processing, also with miniHDMI high-definition output, CMOS camera, WiFi, 7-inch LCD touch screen, and other equipments. The operating system of the server system development platform is 2.5 GHz, 4 core, Inter(R) Core (TM) i7-6500U, 8 G memory, and the algorithm uses Matlab_R2015b. The background development framework is Spring, the IDE development tool eclipse, and the database is MySQL.

The server accumulates the raw data in the therapy including the duration of response, attention and standstill, the assessment scale of parent and doctor, training data and so on. In addition, the server is responsible for producing measured emotional labels and sent them to a client for controlling.

Game design

The interactive software program is designed for enhancing autistic children social skills. There are already some games for this kind of autism assisted therapy, such as FaceSay™ [14], which has been proven effects to improve the ability of autistic children to recognize faces, facial expressions and emotions by teaching the children where to look for facial cues such as an eye gaze or a facial expression. In this paper, we provide five subgames in order, namely Emoticon Learning, Expression Test, Expression Imitation, Emoticon Puzzle, and Scenario Test. Figure 2 shows the examples of game interface. All subgames are accompanied by a therapist. In addition to emoticon imitation, the remaining subgames are

| Emoticon Learning | Expression Test | Scenario Test |
|-------------------|-----------------|---------------|
| Emoticon Puzzle   | Expression Imitation | Emoticon Learning |

Figure 2. The examples of game interface.
automatically scored by the system. In the absence of therapist intervention, the game is adjusted automatically by the predicted emotional state: When the child is in a negative mood, the game is adjusted to a level that is easier to complete.

- **Emoticon Learning**: According to the test question displayed on the touch screen, the child selects an emoticon picture that meets the requirement of the question from the three expression pictures to be selected.

- **Expression Test**: When the child performs the expression test, according to the test question displayed on the touch screen, an emoticon image is selected that meets the requirement of the topic from the three expression pictures.

- **Expression Imitation**: On the touch screen of the robot, a facial expression picture that is to be imitated by the child will be displayed in a random manner. The child carefully observes and learns the expression on the facial expression picture, and the therapist judges the accuracy of the expression imitation result.

- **Emoticon Puzzle**: The subgame is divided into four difficulty levels. The first two difficulty levels are for the child to sort out four out-of-order emoticon picture fragments and select four from six unordered fragments to arrange them in order. The fragments are placed in the correct position and they are put together into a complete and correct emoticon picture; The latter two difficulty levels are to select the correct fragments of the facial expression from the emoticon picture of the two eyes and the two mouths, and then to match the emoticon picture lacking eyes or mouth displayed on the touch screen.

- **Scenario Test**: Combine multiple scenes familiar to people’s daily life to perform more realistic and stereoscopic expression recognition training. The scenario test mode is divided into three levels of difficulty. In this mode, a scene picture to be observed and recognized by the child is randomly displayed on the touch screen, and the specific information of the scene is described by means of voice playing, and the child is required to select the emoticon that best matches the given scenario in the question from four, six and eight expression pictures.

### Study subjects

In order to evaluate our system and methods on children with autism, 11 children with autism were recruited from the outpatient clinic in the Institute of Mental Health, Peking University, China. All children met the inclusion criteria including (1) age 6–12 years old. (2) Diagnosed with autism by the experienced child psychiatrists according to the criteria of autistic disorder in Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV). (3) Childhood Autism Rating Scale (CARS) total score $\geq 30$. (4) No other serious mental disorders and major physical or neurological diseases. (5) No taking psychotropic medication. (6) Having the ability to cooperate with therapy.

### Ethical considerations

The research procedure was approved by the Ethics Committee of Peking University Institute of Mental Health. The participants and their guardians were fully informed about the research and agreed to participate. The experiment was completed with the informed consent in writing from guardians of subjects.

### Experimental process

The entire experimental process took five weeks. During this time, these children subjects received at least one therapy per week. Before experiments, each child had been captured facial images with the labels (ground truth) assessed by the therapist simultaneously. Due to the speed of perceiving the facial expression by observers, the therapist gave one label and continued to use this label representing features until the next label being assessed. We used the self-assessment manikins [15] for valence value assessment. This offline dataset has been used for the system preliminary training. The detailed process as follows:

1. At baseline (before therapy) and at the end of the last week of this experiment, the subjects are assessed with Social Responsiveness Scale (SRS).
2. Before formal therapy, selected emotional videos are shown to each subjects for offline dataset. This part experiment is stopped by obtaining images and labels of sufficient emotional expression from each subject without time limits. There is an average of 30 samples per subject.
3. In therapeutic process, each subject plays the game in the following order: Emoticon Learning, Expression Test, Expression Imitation, Emoticon Puzzle, and Scenario Test. Subjects can stop or finish all the questions at any time in every subgame. During the period, subjects play games through the robot touch screen, while the
therapist synchronously uses the tablet to make feedback on children’s emotions and game adjustments. The therapist only gives feedback when the emotional recognition results do not match the observed state.

4. At the last week of this experiment, select a period for the therapist to perform a simultaneous assessment as quickly as possible, as ground truth for the model comparison. In this period, the therapist doesn’t need to provide feedback.

Method overview

Now we describe our method based on the integration of RL and CNN-SVR, by which the generalization of personalized facial expression recognition during the interaction process can be improved according to external feedback. The whole is divided into two steps: 1. Using original CNN to obtain the initial CNN parameters for CNN feature extraction and then training SVR model; 2. Using RL to adjust weights of SVR and fine-tune weights of CNN. Step 1 is training process by offline samples, which has been presented by the existing works [13,16]. In this paper, we focus on the Step 2.

CNN-SVR model

CNN performs great to mine the target spatial relationship in natural image, similar to the filters (receptive field) in the visual cortex that are locally sensitive to external input. Thus, the image feature extraction is performed in CNN part. As shown in Figure 3, CNN-SVR architecture in this paper is designed by integration of SVR and the classical LetNet-5 [17] without the last output layer and the last convolution layer. SVR as decision layer can compensate the limits of the generalization in CNN for that a few samples can’t obtain good training results. The objective function of original CNN based on the concept of cross entropy is:

$$\text{Loss} = \frac{1}{N} \sum_{i=1}^{N} y_i L(x_i, y_i) + \lambda \sum_{j=1}^{L} \text{sum} \left( ||W_j||^2 \right)$$  \hspace{1cm} (1)

where $N$ the number of samples, $L$ the number of layers, $W_j$ the weights of $l$th layer, and $L(x_i, y_i)$ the negative conditional log-likelihood of training data.

In general, the parameters of CNN are measured from original CNN architecture training previously, and the error feedback of SVR cannot pass to the upper level in CNN. We treat the SVR output as one neuron, and this output layer sensitivity is:

$$\delta^l = \sum_{i=1}^{N} \left( \alpha_i - \alpha_i^* \right) x_i \cdot \gamma$$ \hspace{1cm} (2)

$$\gamma = \begin{cases} 0 & ||\hat{y}_i - y_i|| < \varepsilon \\ \frac{\hat{y}_i - y_i}{||\hat{y}_i - y_i||} & \text{otherwise} \end{cases}$$ \hspace{1cm} (3)

where $\alpha_i$ and $\alpha_i^*$ are the Lagrange multipliers and $\hat{y}_i$ is the ground truth. $\varepsilon$ is the threshold and is set to 0.1. Formula (2) is used in RL process for fine-tuning CNN weights with 0.01 learning rate.

Besides, the parameters of CNN-SVR are set as follows: the number of iterations in CNN model is 10000, the input layer size D1 is 32, C1 convolution layer with 6 filter kernels(size: $5 \times 5$), C2 convolution layer with 16 filter kernels(size: $5 \times 5$), all pooling layers with $2 \times 2$ size and full connection layer with 100 neurons. The network input are $32 \times 32 \times 1$ pre-processed facial images and corresponding labels from the same subject.

Reinforcement learning

We use the reinforcement learning method that strives to achieve the model parameters online updating considering about new samples from an interactive environment, proposed to compensate the limits of the...
less offline training sample data and achieve the effect of correcting the deviation of the regression result. The key elements of RL include environment, reward, action or state.

- The part of environment is ground truth from therapist’s assessment results of the interactive child with ASD, which give the current observed emotional state of the child (the therapist does not agree with the real-time results displayed on the tablet). The corresponding features and the predicted labels are determined by the previous second (considering the operation delay) that receives the assessment feedback.
- The reward is given from environment which compares ground truth (assessment feedback) and predicted label. It can be defined as:

\[
\text{reward} = \begin{cases} 
0 & \|\hat{y}_i - y_i\| < \epsilon \\
1 & \text{otherwise} 
\end{cases}
\]  

(4)

- If the reward from current state is 0, the CNN-SVR model will not make any changes. And the system is in automatic mode. But if reward is 1, the model will be trained online for performance improvement.

It is noticed that the reward is set as 0 by default without the assessment feedbacks. We assume that this is because the therapist agrees with the system’s automatic prediction results.

SVR weights adjustment

The epsilon-Support Vector Regression (\(\varepsilon\)-SVR) is adopted in this study to perform continuous mapping. Only sample points outside the \(\varepsilon\) pipe wall and on the \(\varepsilon\) pipe contribute to the solution, which is called support vector. The complexity of the SV function representation is independent of the dimension of the input space \(X\) and depends only on the number of SVs. For this regression model, the per-example convex loss function is:

\[
L(x_i, y_i, \theta) = \begin{cases} 
0 & \|\hat{y}_i - y_i\| < \varepsilon \\
\|\hat{y}_i - y_i\| - \varepsilon & \text{otherwise} 
\end{cases}
\]  

(5)

where \(\theta\) is the weight parameters of SVR.

The regularized convex optimization problem for linear regression:

\[
\theta^* = \frac{1}{N} \sum_{i=1}^{N} L(x_i, y_i, \theta) + \arg\min_{\theta \in \mathbb{R}^d} \frac{\tau}{2} \|\theta\|^2
\]  

(6)

Similar to CNN weights calculation, the SVR weights adjustment can be obtained from stochastic gradient descent (SGD). The gradient is the derivation of loss function with respect to weights:

\[
\text{gradient} = \nabla \frac{1}{N} \sum_{i=1}^{N} L(x_i, y_i, \theta)
\]  

(7)

It’s worth noting that Formula (2) is derivation with respect to bias while Formula (7) with respect to weights. The former is calculated to obtain the upper layer sensitivity in the BP algorithm for the weight update of the network, while the latter is directly dependent on the gradient descent method. Thus, at time \(t\), the weights are updated as:

\[
\theta_{t+1} = \theta_t - \eta (\nabla h(x_t, y_t, \theta_t) + \tau t)
\]  

(8)

where \(\eta\) is learning rate, and we decide to use 0.01 learning rate after multiple convergence test. It can be treated as the action of RL method.

Results

CNN-SVR based on RL results

As for our proposed method, we used CNN-SVR regression based on RL to predict emotion. The valence value is used to represent the state of emotion. For verifying the effectiveness of RL, the CNN-SVR model trained by offline dataset is compared with the model based on RL after five weeks. The offline dataset has a sample size of not less than 40 per subject. In addition, our method is capable of processing 4 times prediction per second on our server.

Figure 4 describes one example’s details of predicted emotion from fragment of Subject 2. The predicted valence from RL model is closer to ground truth than from offline model, but the predicted values of both are not as smooth as ground-truth. The overall predicted value of offline model is lower than RL, which proves that manual intervention leads to hyperplane deflection through reinforcement learning. Besides, we found that in a few frames, the subject habitually bit the finger and blocked part of the face. The corresponding errors of these frames increase. In RL model, the obvious errors (absolute value >0.2) are similar to the form of a pulse signal. Therefore, usually the corresponding game with the level caused by these short-term errors will not be presented in the interactive interface.

For analysis of the overall sample, the Figure 5 shows the distribution of subjects’ features in two-dimensional space by the method of t-distributed stochastic neighbor embedding (t-SNE) where high-dimensional features have been projected into low dimension. The perplexity is selected as 100. We can
find that the subject’s features have significant individual differences and each subject’s data has become aggregated. But for Subject 10, feature distribution is dispersed.

Root Mean Square Error (RMSE) is used to measure the deviation of the predicted value from the ground truth. Boxplots in Figure 6 show more detailed distribution of RMSE for each subject. For comparison, labels are also derived from the therapist’s assessment. The RMSE mean of RL model under condition of each subject is both less than the offline model. The upper limits of RL are generally lower than offline method’s, which means the worst results of RL is still better than the other; For each subject, the degree of dispersion of the RMSE is various, which depends on validity of data, such as the reliability of self-assessment and children’s facial expression habits. And for eleven children subjects, the dispersion degree of the RMSE of the RL model is less than offline model’s, which verifies that RL method has better generalization ability. But this result is not true for Subject 9 and Subject 10 et al. One of the reasons may be that some features dispersed in the eigenspace.

Therapeutic results

In our study, SRS was chosen to evaluate the change of symptoms of children with autism after the training. SRS includes 5 sub-scales (awareness, cognition, communication, social motivation, autistic mannerism) which is often used to evaluate social skill deficits in children with autism. The higher the scores, the more severe of social interaction deficit. The efficacy was analyzed by comparing SRS total score and its factor scores at baseline and at the end of therapy. The Table 1 showed the Demographic information of study subjects and scale assessment results at baseline.

We recorded the interactive time which subjects spent on looking at robot and screens during the assisted therapy. The mean and SD of interactive time are shown in Table 2. Each child has spent an average of 29 minutes per interaction, and there is not much deviation. The number of failures of games was recorded also, and we find that the subgame
Figure 6. The RMSE of predicted emotion from each child subject.

Table 1. Demographic information and scale assessment results at baseline.

| Sub. | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | M   | SD  |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Age (years) | 7   | 8   | 10  | 11  | 7   | 8   | 9   | 10  | 8   | 11  | 9.1 | 1.45|
| CARS total score | 32  | 32  | 31  | 40  | 37  | 33  | 32  | 33  | 39  | 31  | 34.18| 3.37|
| SRS total score | 115 | 109 | 66  | 85  | 41  | 93  | 104 | 81  | 99  | 71  | 89.82| 24.23|

Table 2. The interactive details of subjects in assisted therapy.

| Subject | Interactive time/minutes | Emoticon Learning | Emoticon Puzzle | Scenario Test | Expression Test | Number of failures (at Week 5/ at Week 0) |
|---------|--------------------------|-------------------|----------------|---------------|----------------|------------------------------------------|
|         | Mean SD                  | Mean SD           | Mean SD        | Mean SD       | Mean SD        | Mean SD                                  |
| 1       | 29.4/2.1                 | 0.51/0.68         | 0.17/0.36      | 1.04/1.01     | 0.66/0.81      | 0.66/0.81                                |
| 2       | 28.7/3.2                 | 0.23/0.37         | 1.65/1.84      | 0.66/0.81     | 0.11/0.08      | 0.82/0.56                                |
| 3       | 29.2/2.7                 | 0.12/0.07         | 0.14/0.11      | 0.21/0.52     | 0.51/0.76      | 0.86/0.82                                |
| 4       | 31/2.2                   | 0.35/0.51         | 0.17/0.35      | 0.24/0.51     | 0.86/1.03      | 1.15/0.36                                |
| 5       | 27.4/4.5                 | 0.69/0.87         | 0.56/0.73      | 0.69/0.83     | 0.65/0.97      | 1/1.03                                   |
| 6       | 27.8/1.6                 | 0.14/0.36         | 0.35/0.19      | 0.62/0.75     | 0.52/0.78      | 1.13/0.89                                |
| 7       | 28.8/3.2                 | 0.23/0.41         | 0.21/0.44      | 0.65/0.79     | 0.67/0.77      | 0.56/0.46                                |
| 8       | 30.4/1.6                 | 0.41/0.28         | 0.123/1.39     | 0.34/0.25     | 0.63/0.49      | 0.72/0.68                                |
| 9       | 29.7/2.4                 | 0.74/0.95         | 0.76/0.87      | 0.56/0.66     | 0.79/0.16      | 0.34/0.48                                |
| 10      | 26.2/0.7                 | 0.34/0.51         | 0.34/0.27      | 0.46/0.51     | 0.93/1.11      | 0.96/1.03                                |
| 11      | 30.1/1.9                 | 0.28/0.48         | 0.23/0.49      | 0.78/0.59     | 0.34/0.58      | 1.02/0.73                                |
| Mean    | 28.97/2.37               | 0.37/0.50         | 0.53/0.64      | 0.46/0.46     | 0.61/0.77      | 0.81/0.70                                |
| SD      | 1.34/0.97                | 0.20/0.24         | 0.48/0.51      | 0.23/0.20     | 0.22/0.28      | 0.28/0.22                                |
Emoticon Learning has the lowest failure number. From the overall mean and standard deviation, we can conclude that there are fewer errors (approximately equal to 1) in each group of games. Besides, the average number of failures of each subgame at week 5 is less than the one at week 0. These validate the engagement of child subjects.

The same parent filled in SRS for subjects at baseline and endpoint of the therapy. The assessment results are shown in Table 3 and Figure 7. Before the therapeutic stage (at week 0), mean of SRS assessment was 89.82 (SD = 24.23); after 5 weeks' therapy, the mean was 84 (SD = 27.69). The non-parametric Kruskal-Wallis (K-W) test between the SRS at week 0 and at week 5 is performed to test significance of therapeutic effect. The results show that there is a non-significant difference between SRS total score at baseline and endpoint, with Chi-sq 0.28, p value 0.60. For 5 subscales, social awareness with Chi-sq 0.18, p value 0.67, social cognition with the Chi-sq 0.91, p value 0.34, social communication with Chi-sq 0.13, p value 0.72, social motivation with Chi-sq 0.09, p value 0.77, and autistic mannerism with Chi-sq 0.03, p value 0.87. In short-term therapy, the change in SRS scores was not significant for the subjects as a whole. But when each autistic subject score before and after the therapy is treated as a categorical group, SRS total score is different significantly due to the Chi-sq 19.07 and p value 0.039 < 0.05. Therefore, the therapeutic effect varies from child to child. Most of the subjects' changes were small (increase for 2 subjects, drop slightly for 6 subjects) and the three subjects' SRS results presented drop obviously (subject-1 from 115 to 104, subject-2 from 109 to 89, subject-3 from 66 to 40) in general. In particular, unlike the remaining subjects, the most of sub-scale scores and the total score of the Subject 7 become higher, which affects the assessment of the overall sample. Although the mean value of the decline of social cognition is higher than social awareness, the deviation is significantly larger.
(social cognition 4.13, social awareness 2.84). We can see that nine subjects’ social awareness has improvement.

Discussion

Effectiveness of CNN-SVR based on RL

This work considers about the reinforcement learning for personalizing the prediction model in the therapy progresses, for the sessions of that samples could be drawn for training the models, and these samples are further apart in time and less likely to be visible. The prediction model uses CNN-SVR. CNN’s learning algorithms are based on empirical risk minimization, while SVMs are designed to use structural risk minimization principles. The former attempts to minimize errors in the training set, while the latter minimizes the generalization error of unseen data with a fixed distribution for training set. Backpropagation algorithms tend to fall into local minima and stop training. Therefore, CNN’s generalization ability is lower than SVR. The SVR calculates the global optimal solution by solving the quadratic programming problem. Therefore, we use the generalization ability of SVR to maximize the prediction error of the hybrid model. CNN-SVR model performance is also subject to the confidence of the subject sample and the reliability of the label. RL method is used to improve the problems of the former for that it increases the new mappings of sample to label. RMSE results from all subjects have proven the improvement by RL method.

Therapeutic effect

We conducted a comparison between the SRS details’ scores at week 0 and at week 5 to verify the efficacy of assisted therapy on children with autism. Overall, the difference between SRS scores at baseline and endpoint is not significant. But a slight decrease in the SRS total score in most subjects could be seen. In addition, the therapeutic effect varies greatly from person to person. For example, the extent of autistic social impairment of Subject 7 did not improve but worsened, while three subjects’ extent has improved obviously (Subject 1, Subject 2 and Subject 3). So the results suggest that there are some factors impact the results, maybe the subject’s state (impaired language, age, cognitive skills), culture and environment factors [18], small sample size or short-term therapy. So more research considering above factors is needed to clarify the efficacy of the assisted therapy on children with autism. Although the efficacy of assisted therapy on autism is needed to verify in the future, our results show that the improvement of social skill was concentrated in the social awareness in children with autism, so it is suggested the assisted therapy maybe a promising method for the therapy of deficit of social awareness in children with autism.

System feasibility

The ultimate objective of the research is to empirically test the feasibility of the assisted therapy system. And the factors that measure the feasibility are engagement, therapeutic effect of children, and whether the system can support the work of the therapist. In the interactive process, Children with ASD have lower game failures which verify the engagement of children. For therapeutic effect, the results for individual have a significant change, and most subjects’ SRS total score is slightly lower at endpoint than at baseline. The possibility of treating deficit of social awareness has been discussed. Besides, the effectiveness of CNN-SVR based on RL verifies that this system can help support therapists’ working with children with autism, because accurate prediction and recording of children’s emotions can help therapists quickly extract patterns of changes between children’s emotions and behaviors to get some medical statistics from these. A system could excel at detecting or tracking small behavioral changes to help therapists observe the point of time for small progress.

Limitations and future work

Our study has some limitations in several ways. For machine learning model, the predicted results are based entirely on facial expressions, which may cause the superficiality and one-sidedness of emotional judgment. The reliability of facial and other modalities characterizing emotions of autistic children are not described in this paper. But multimodal fusion and mutual exclusion enhance reliability of prediction [19].

For therapeutic effects, a small number of experimental subjects and short-term therapeutic time are a limit that cannot be ignored. In addition, cultural differences and the family environment can also cause uncontrollable effects on children. In future work, we will try our best to invite more subjects and reduce the external factors to make the impact of the therapy effect as single as possible. Cultural difference will be considered too in the design of system. Moreover, the system can be designed as a partner to autistic children for emotional interaction, which also needs the
ability to generate and express emotions in addition to emotional perception.

Conclusions
This paper described the development and application of an assisted therapy system for children with ASD, integrated RL and CNN-SVR to predict emotion from facial expression. The RLAT system contains two parts, namely, manual intervention and automatic mode. Manual intervention can be treated as external feedback from environment in RL process. And the results demonstrate the effectiveness of prediction model based on RL while demonstrate the feasibility of assisted therapeutic system in children with autism. In the future, we will take into account multimodality to compensate for the confidence of single mode and expand the number of subjects to further determine the efficacy of assisted therapy on children with autism.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
The work is supported by the National Key Research and Development Program of China [Grant No. 2017YFB1002804], and National Natural Science Foundation of China [No. 61672093, No. 61432004].

References
[1] Riek LD. Healthcare robotics. Commun ACM. 2017;60:68–78.
[2] Feil-Seifer D, Mataric MJ. B3IA: a control architecture for autonomous robot-assisted behavior intervention for children with autism spectrum disorders. Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication, 2008 RO-MAN 2008; 2008 Aug 1–3; Munich, Germany: IEEE; 2008. p. 328–333.
[3] Krebs Hl, Palazzolo JJ, Dipietro L, et al. Rehabilitation robotics: performance-based progressive robot-assisted therapy. Auton Robot. 2003;15:7–20.
[4] Xu QL, Zhou F, Jiao J. Design for user experience: an affective-cognitive modeling perspective. Proceedings of the 2010 IEEE International Conference on Management of Innovation and Technology (ICMIT); 2010 Jun 2–5; Singapore: IEEE; 2010. p. 1019–1024.
[5] The AuRoRA Project [Internet]; 2010 Mar 26 [cited 2018 Mar 28]. Available from: http://www.aurora-project.com/
[6] Klein T, Gelderblom GJ, de Witte L, et al. Evaluation of short term effects of the IROMEC robotic toy for children with developmental disabilities. Proceedings of the 2011 IEEE International Conference on Rehabilitation Robotics (ICORR); 2011 Jun 29–Jul 1; Zurich: IEEE; 2011. p. 1–5.
[7] Smitha KG, Vinod AP. Facial emotion recognition system for autistic children: a feasible study based on FPGA implementation. Med Biol Eng Comput. 2015;53:1221–1229.
[8] RUDOVIC, Ognjen, et al. Personalized machine learning for robot perception of affect and engagement in autism therapy. arXiv preprint arXiv:1802.01186, 2018.
[9] Kim JC, et al. Audio-based emotion estimation for interactive robotic therapy for children with autism spectrum disorder. Proceedings of the 2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI). IEEE; 2017. p. 39–44.
[10] Piana S, Staglianò A, Camurri A, et al. A set of full-body movement features for emotion recognition to help children affected by autism spectrum condition. Proceedings of the IDGEl International Workshop; 2013 May 4; Chania, Crete, Greece. 2013.
[11] Chu H-C, et al. Facial emotion recognition with transition detection for students with high-functioning autism in adaptive e-learning. Soft Comput. 2017;22:1–27.
[12] Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: a survey. JAIR. 1996;4:237–285.
[13] You W, Shen C, Guo X, et al. A hybrid technique based on convolutional neural network and support vector regression for intelligent diagnosis of rotating machinery. Adv Mech Eng. 2017;9:168781401770414.
[14] FACESAY™ SOCIAL SKILLS SOFTWARE GAMES [Internet]; 2013 Dec 13 [cited 2018 Mar 28]. Available from: www.facesay.com/
[15] Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential. J Behav Ther Exp Psychiatry. 1994;25:49–59.
[16] Zhou J, Hong X, Su F, et al. Recurrent convolutional neural network regression for continuous pain intensity estimation in video. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops; 2016 Jun 26–Jul 1; Las Vegas. 2016. p. 84–92.
[17] Lecun Y, et al. LeNet-5, convolutional neural networks; 2015 [Cited 2018 May 14]. Available from: http://yann.lecun.com/exdb/lenet.
[18] Hus V, Bishop S, Gotham K, et al. Factors influencing scores on the social responsiveness scale. J Child Psychol Psychiatry. 2013;54:216–224.
[19] Kryszczuk K, Richiardi J, Prodanov P, et al. Reliability-based decision fusion in multimodal biometric verification systems. EURASIP J Adv Signal Process. 2007; 2007:086572.