Supporting Children's Metacognition with a Facial Emotion Recognition based Intelligent Tutor System

Citation for published version:
Ruan, X, Palansuriya, C, Constantin, A & Tsiakas, K 2023, Supporting Children's Metacognition with a Facial Emotion Recognition based Intelligent Tutor System. in *IDC '23: Proceedings of the 22nd Annual ACM Interaction Design and Children Conference.* ACM, pp. 502-506, 2023 ACM Interaction Design and Children (IDC), Chicago, Illinois, United States, 19/06/23. https://doi.org/10.1145/3585088.3593882

Digital Object Identifier (DOI):
10.1145/3585088.3593882

Published In:
IDC '23: Proceedings of the 22nd Annual ACM Interaction Design and Children Conference

Supporting Children's Metacognition with a Facial Emotion Recognition based Intelligent Tutor System

XINGRAN RUAN*, The University of Edinburgh, UK
CHARAKA PALANSURIYA†, The University of Edinburgh, UK
AURORA CONSTANTIN†, The University of Edinburgh, UK
KONSTANTINOS TSIAKAS†, Delft University of Technology, Netherlands

The present study aims to investigate the relationship between emotions experienced during learning and metacognition in typically developing (TD) children and those with autism spectrum disorder (ASD). This will assist us in using machine learning (ML) to develop a facial emotion recognition (FER) based intelligent tutor system (ITS) to support children's metacognitive monitoring process in order to enhance their learning outcomes. In this paper, we first report the results of our preliminary research, which utilized an ML-based FER algorithm to detect four spontaneous epistemic emotions (i.e., neutral, confused, frustrated, and boredom) and six spontaneous basic emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise). Subsequently, we adapted an application ('BrainHood') to create the 'Meta-BrainHood', that embedded our proposed ML-based FER algorithm to examine the relationship between facial emotion expressions and metacognitive monitoring performance in TD children and those with ASD. Finally, we outline the future steps in our research, which adopts the outcomes of the first two steps to construct an ITS to improve children's metacognitive monitoring performance and learning outcomes.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: Intelligent tutor system, facial emotion recognition, metacognitive monitoring process, learning outcomes

ACM Reference Format:
Xingran Ruan, Charaka Palansuriya, Aurora Constantin, and Konstantinos Tsiakas. 2018. Supporting Children's Metacognition with a Facial Emotion Recognition based Intelligent Tutor System. In *Interaction Design and Children (IDC '23), June 19–23, 2023, Chicago, IL, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/https://doi.org/10.1145/3585088.3593882

1 INTRODUCTION

Self-regulated learning (SRL) is a multifaceted educational construct that pertains to how learners manage their cognitive, metacognitive, behavioral, motivational, and emotional processes to attain educational goals [35]. Within the SRL framework, metacognition is a vital aspect and is often described as 'thinking about one's thinking' [35]. One of the key components of metacognition is metacognitive monitoring, which involves evaluating one's learning process or current state of knowledge (e.g., assessing the accuracy of one's answers). Research evidence suggests that supporting metacognitive monitoring skills is the most cost-effective method of improving children's learning outcomes [19].

However, studies have demonstrated that some children with ASD exhibit deficits in metacognition, particularly in the area of metacognitive monitoring. Significant differences in metacognitive monitoring scores have been observed between groups of children with ASD and TD children [3, 5, 17, 29]. Conventional educational interventions have been effective in improving the learning outcomes of both TD children and children with ASD by supporting their metacognitive monitoring skills (e.g., by providing feedback and supportive comments) [19, 24]. For instance, Cogliano et al. [7] reported that undergraduate students who received metacognitive retrieval practice training performed better on the final exam compared to the control group. Similarly, Maras et al. [23] showed that the average score of children with ASD (mean age: 13.3) who received metacognitive support through a system called 'Math Challenge' was significantly higher than that of the control group. Nonetheless, these conventional interventions have their limitations, as they rely on children's answers to questions and do not provide real-time support based on children's performance/activity during the learning process.

Recent research that focused on the relationship between emotions and metacognition, highlighted a strong correlation between epistemic emotions (i.e. confusion, frustration, boredom, surprise, and delight) and metacognitive monitoring performance [6, 16, 30, 31]. For instance, surprise and delight have been found to have a positive correlation with subjects' metacognitive monitoring performance [4, 30], while boredom has a negative correlation [6]. However, this field of study is still emerging, particularly with regard to children with ASD, who exhibit heterogeneous facial emotion expressions. Furthermore, most FER techniques in conventional interventions focus on six basic emotions [21] instead of epistemic emotions which occur considerably more often when compared with basic emotions. Consequently, state-of-the-art conventional interventions are unable to provide effective feedback to children in real time, thus limiting their ability to support the metacognitive monitoring process of both TD children and children with ASD.

Our research aims to address three gaps in the current literature. Firstly, the current automatic FER techniques are limited to identifying only basic emotions [13]. Secondly, the relationship between emotions that occur during learning and metacognitive monitoring performance in children with ASD is not fully understood [4, 30]. Thirdly, conventional interventions do not provide real-time support to enhance the metacognitive monitoring [16, 33]. Our research seeks to explore the development of an automatic machine learning-based ITS that can interpret emotions that arise during learning and provide real-time feedback to support the metacognitive monitoring process of both TD children and children with ASD.

*Corresponding author.

IDC '23, June 19–23, 2023, Chicago, IL, USA

2 RELATED WORK

Currently, state-of-the-art ITS utilize both non-automatic and automatic approaches to recognize learners' facial emotion expressions and improve their learning outcomes.

Non-automatic FER approaches implemented in ITS involve the use of self-report methods (e.g. learners themselves reporting their emotional state) [15, 32] or annotator-report methods (e.g. involve experts to annotate learners' emotions based on their observed behaviour) [8, 11] to recognize learners' emotions. For example, in [15], subjects learn computer literacy with the help of 'AutoTutor'. During the learning process, researchers record subjects' facial expressions, and the experienced emotions of the subjects in the videos are recognized by the learners themselves in a post-learning session. Similarly, in [32], a questionnaire about the subjects' affective status is implemented at the end of each game round, and children's emotions are recognized based on their answers. On the other hand, the annotator-report method is used to recognise emotions. For example, Craig et al. [8] recognise learners' emotions during learning activities by a team of trained observers. Additionally, D'Mello et al. [11] used an "Emote-Aloud" method to recognize learners' emotions, which required participants to verbalize their feelings while interacting with 'AutoTutor'.

In contrast to non-automatic approaches in ITS, numerous automatic FER methods have been proposed to objectively recognize learners' emotions [20, 27, 28, 36]. For example, Ekman and Oster's Facial Action Coding System (FACS) [14] identifies emotions by 46 unique Action Units (AU) expressed on the face. The well-known application iMotion [20] applies FACS for automatic FER, and it has been integrated into 'MetaTutor' to evaluate learners' affective states. However, FACS requires manual human coding and is, therefore, less reliable and objective [2]. In contrast, ML-based methods can produce more objective predictions by training on emotion datasets. For instance, Peng et al. [27] use the ResNet10 neural network to extract textures and recognize basic emotions in images. Xue et al. [36] propose a vision-transformer (ViT)-based neural network to recognize emotions when some parts of the face are obscured, and it outperforms state-of-the-art FER techniques in terms of accuracy of emotion recognition.

By combining automatic FER techniques with ITS, it is possible to respond to learners' emotions during learning. For example, Savchenko et al. [28] design an ML-based FER framework to classify students' emotions and engagement levels in an online learning scenario. D'Mello et al. [10] develop a multimodal emotion classifier in 'Affective AutoTutor' to improve students' learning outcomes based on their emotions.

While related work has attempted to improve children's learning outcomes by interpreting their emotions using automatic or non-automatic FER methods, the important relationship between emotions and metacognitive monitoring performance has been overlooked in design of a state-of-the-art ITS. As a result, these ITS are unable to provide real-time support for the metacognitive monitoring of both TD children and those with ASD, thereby limiting the effectiveness of feedback provided for optimal learning outcomes.

3 RESEARCH AIM AND OBJECTIVES

In this research project, we aim to develop an automatic ML-based ITS for both TD children and children with ASD in primary school (ages 7 to 11) to enhance their learning outcomes by providing real-time support to their metacognitive monitoring (as depicted in Fig.1). To accomplish this goal, we show the following objectives and this paper illustrates our work to achieve the first objective:

1. Identify the emotions (including basic emotions and epistemic emotions) that are significantly correlated to the metacognitive monitoring performance in TD children and children with ASD.

2. Design the ITS that provides feedback to support the metacognitive monitoring process of TD children and children with ASD by responding to their emotions in real time.

3. Evaluate the performance of ITS by children's metacognitive monitoring performance and learning outcomes.

4 FACIAL EMOTION RECOGNITION (FER) ALGORITHMS FOR ITS

Epistemic emotions, which are produced during the metacognitive process, are significantly correlated with the performance of metacognitive monitoring. However, state-of-the-art FER techniques for identifying epistemic emotions are limited due to the scarcity of training data. In order to improve the recognition rate of an ML-based FER algorithm for classifying students' epistemic emotions during learning, we proposed a new loss function called the Affective Dynamic Loss (AD-Loss), which was designed based on the Control Value Theory [26].

We evaluated the proposed method on the PUZZLED dataset [22] [1]. As shown in Table 1, the results showed that the network trained by the AD-Loss function outperforms other three state-of-the-art loss functions for FER: Cross Entropy (CE) [1], Additive Margin Softmax (AM) [34], and ArcFace (AF) [9] and provides better recognition rate to recognise epistemic emotions. In addition, we have also conducted experiments to evaluate the performance of the FER algorithm in recognizing the six basic emotions of TD children. We used a convolutional-based neural network [18] and tested it on the ChildEFES dataset [2] [25]. Although the validation set showed a promising accuracy with an average of 91.76% (In Fig.2, we show the attention maps of some discriminative features learned by the convolutional neural network on the validation set), the average accuracy on the test set was only 48%. This result motivated us to explore the attention mechanism, such as the use of the transformer model [12], to learn more discriminative features from children's facial expressions. Psychological studies have shown that facial emotion expressions are characterized by dynamic motions of certain facial parts, such as the eyes, nose, and mouth, which contain discriminative features for representing different emotions [21]. To improve the accuracy of the ML-based FER algorithm, we trained a transformer-based neural network, ViT [12], to recognize facial

[1]The PUZZLED consists of 10 videos of students when they are watching educational videos. Their emotional annotations have 4 values: Neutral, Confused, Frustrated, and Boredom.

[2]ChildEFES is a photo and video database of 4-to-6-year-olds expressing the seven induced and posed universal emotions (happy, disgust, surprise, fear, sad, anger, contempt) and a neutral expression.

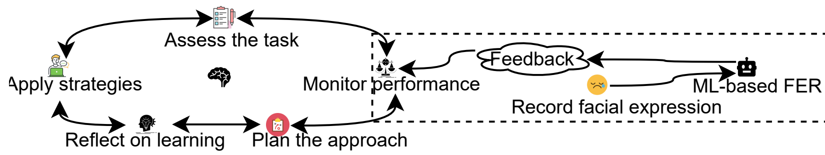


Fig. 1. Our expected interaction between ITS and children's metacognition workflow.

Loss Function	Weighted F1 score (Ave)	Accuracy (Ave)	Loss Function	Weighted F1 score (Ave)	Accuracy (Ave)
CE	0.4866	0.5299	**CE + AD**	**0.5296**	**0.5541**
AM	0.5194	0.6376	**AM + AD**	**0.5216**	**0.6419**
AF	0.5055	0.6027	**AF + AD**	**0.5320**	**0.6414**

Table 1. Recognition Rate (Accuracy) and F1 score of AD-Loss (our proposed) and three state-of-art loss functions on the test set (10% of the subjects) of PUZZLED. (**CE**: Cross Entropy; **AD**: AD-Loss (we proposed); **AM**: Additive Margin Softmax; **AF**: ArcFace)

(a) Anger (b) Disgust (c) Fear (d) Happy (e) Sad (f) Surprised

Fig. 2. Neural network's attention on six basic emotions from the validation set of the ChildEFES. The chin raiser in anger expression in Fig.2a; the lip corner depressor and the lower lip depressor in disgust expression in Fig.2b; the brow lower and upper lid raiser in fear expression in Fig.2c; the cheek raiser and lip corner puller in happy expression in Fig.2d; the brow lower and lip corner depressor in sad expression in Fig.2e; the jaw drop in surprise expression in Fig.2f.

Base Model	Numbers of patches on AUs	Weighted F1 score (Ave)	Accuracy (Ave)
ResNet50 [18]	0 (Fig.3a)	0.47	0.47
ViT [12]	0 (Fig.3a)	0.46	0.49
AU-ViT (our proposed)	31 (Fig.3b)	0.56	0.60
AU-ViT (our proposed)	76 (Fig.3c)	0.63	0.68

Table 2. Recognition Rate and F1 score of AU-based ViT (AU-ViT) on the test set of the ChildEFES.

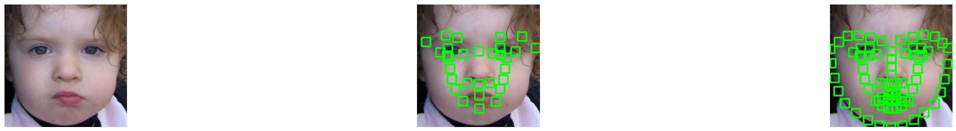
emotion expressions depending on action units (AUs) areas on the face, see Fig.3. Specifically, we proposed an AU-based ViT (AU-ViT) to recognize six basic emotions of TD children.

Compared with training on origin images (Fig.3a) which achieves 0.49 accuracy, the accuracy of AU-ViT on the test set (10% of subjects) of ChildEFES is presented in Table.2. It demonstrates that training on typical AUs' areas (Fig.3b,3c) instead of the entire facial expression (Fig.3a) improves the recognition rate. The experimental results demonstrate the superiority of our proposed FER algorithms in recognizing both epistemic and basic emotions, as compared to state-of-the-art algorithms. This success motivates us to apply our FER techniques to identify emotions that exhibit a significant correlation with metacognitive monitoring performance in both TD children and those with ASD.

5 THE 'META-BRAINHOOD' PROTOTYPE APPLICATION

We have developed the 'Meta-BrainHood' by adapting a prototype ('BrainHood') which includes a set of cognitive games for self-regulated game-based learning experiments [32]. Our adapted version is designed to be used by both TD children and children with ASD to investigate the relationship between their facial emotional expressions and their performance in the metacognitive monitoring process. To achieve this, we applied our proposed FER algorithms to 'Meta-BrainHood' to recognize facial emotion expressions collected from the children.

Compared with 'BrainHood', our implementations of 'Meta-BrainHood' is illustrated in Fig.4. Firstly, we connected the webcam to 'Meta-BrainHood' to collect facial emotion expressions from children (see the webcam in Fig.4), which is activated when children log in. We then simplified the game flow by creating a welcome and information page for children, both TD and with ASD. In specific, we provided three pre-defined game scenarios on the welcome page (easy, medium, and hard) and moved the game configuration to the 'customized' game page (1 and 4 in Fig.4 respectively). This modification aims to reduce the cognitive load on children by simplifying the information presented on the welcome page and facilitating familiarity with the game by allowing children to play a few rounds



(a) Origin facial expression. (b) 31 numbers of patches on AU area. (c) 76 numbers of patches on AU area.

Fig. 3. The example of different numbers of patches on facial emotion expressions: Origin image in (a) and different numbers of patches in (b) and (c).

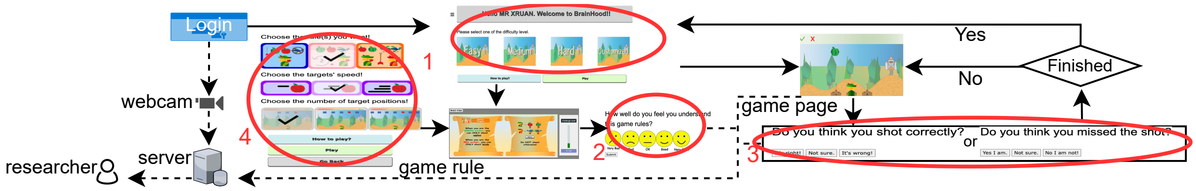


Fig. 4. A play flow of 'Meta-BrainHood'. 1: welcome page; 2: the JOL questionnaire of game rules; 3: the JOL questionnaire of the current action (shoot or do nothing); 4: customized game; Dash lines to server: transmit children's behaviour (emotion and actions in game); Dash line to researcher: transmit results of emotion recognition and JOL answers.

before beginning the actual experiment. We also enabled the application to collect Judgement-Of-Learning (JOL) answers and game performance data from children (the dash lines from 2, 3 and game page to the server in Fig.4). These data are stored on the server, as shown by the dash lines in Fig.4. Finally, our proposed FER algorithms will be used to recognize basic emotions and epistemic emotions of children that occurred during play. The results of FER and answers of JOL of TD children and children with ASD will be transmitted to researchers, see the dash line between server to a researcher in Fig.4.

The 'Meta-BrainHood' interface has been formatively evaluated with two experts in user-experience and autism. They explored the application and then answered a semi-structured interview regarding the ease of use of the system and its appropriateness for TD and ASD children. The experts feedback has been used to refine the interfaces, e.g., some of the labels were changed to be easier to understand for children, the language used for information provided to children has been simplified. Overall, the experts concluded that the interface is easy to use and appropriate for TD and ASD children.

6 DISCUSSION AND NEXT STEP

To date, our research has focused on the first stage, aimed at addressing the initial research question of identifying the emotions (including basic emotions and epistemic emotions) that exhibit a significant correlation with the metacognitive monitoring performance in TD children and children with ASD. We have developed two FER algorithms, namely AD-Loss-based convolutional network and AU-ViT, for recognizing epistemic emotions (i.e., neutral, confused, frustrated, and boredom) and six basic emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise). Our experimental results, obtained from various datasets, demonstrated that our proposed FER algorithms outperform state-of-the-art methods. Based on these promising findings, we incorporated our FER algorithms into the 'Meta-BrainHood' to examine the relationship between emotions and metacognitive monitoring performance. Given the varying facial emotion expressions among subjects, particularly children with ASD, we will fine-tune our FER algorithms on each subject before they are included in the 'Meta-BrainHood' study.

In the subsequent stage of our study, we will recruit participants (TD children and children with ASD) to play 'Meta-BrainHood', and we will analyse the collected data, including video recordings of facial emotion expressions and responses to JOL questions. This analysis will enable us to investigate the relationship between emotional changes and metacognitive monitoring performance in both TD children and children with ASD. Upon examining the relationship between facial emotion expressions and metacognition in children with ASD, we will develop a FER-based ITS to support their metacognitive monitoring process (as illustrated in Fig.1).

A comparative study will be conducted to evaluate the FER-based ITS against conventional interventions. The final ITS is expected to provide feedback in real-time to TD children and children with ASD to improve their metacognitive monitoring process and enhance their learning outcomes.

REFERENCES

[1] MAH Akhand, Shuvendu Roy, Nazmul Siddique, Md Abdus Samad Kamal, and Tetsuya Shimamura. 2021. Facial emotion recognition using transfer learning in the deep CNN. *Electronics* 10, 9 (2021), 1036.
[2] Kellen Briot, Adrien Pizano, Manuel Bouvard, and Anouck Amestoy. 2021. New Technologies as Promising Tools for Assessing Facial Emotion Expressions Impairments in ASD: A Systematic Review. *Frontiers in psychiatry* 12 (2021), 530.
[3] Mark Brosnan, Hilary Johnson, Beate Grawemeyer, Emma Chapman, Konstantina Antoniadou, and Melissa Hollinworth. 2016. Deficits in metacognitive monitoring in mathematics assessments in learners with autism spectrum disorder. *Autism* 20, 4 (2016), 463–472.
[4] Stephanie Buono, Ana Zdravkovic, Milan Lazic, and Earl Woodruff. 2020. The effect of emotions on self-regulated-learning (SRL) and story comprehension in emerging readers. In *Frontiers in Education*, Vol. 5. Frontiers Media SA, 588043.
[5] Katie L Carpenter, David M Williams, and Toby Nicholson. 2019. Putting your money where your mouth is: examining metacognition in ASD using post-decision wagering. *Journal of autism and developmental disorders* 49, 10 (2019), 4268–4279.
[6] Elizabeth B Cloude, Franz Wortha, Daryn A Dever, and Roger Azevedo. 2020. How do Emotions Change during Learning with an Intelligent Tutoring System? Metacognitive Monitoring and Performance with MetaTutor.. In *CogSci*.

[7] MeganClaire Cogliano, Matthew L Bernacki, and CarolAnne M Kardash. 2020. A metacognitive retrieval practice intervention to improve undergraduates' monitoring and control processes and use of performance feedback for classroom learning. *Journal of Educational Psychology* (2020).

[8] Scotty Craig, Arthur Graesser, Jeremiah Sullins, and Barry Gholson. 2004. Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. *Journal of educational media* 29, 3 (2004), 241–250.

[9] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. 2019. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 4690–4699.

[10] Sidney D'mello and Art Graesser. 2013. AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 2, 4 (2013), 1–39.

[11] Sidney K D'Mello, Scotty D Craig, Jeremiah Sullins, and Arthur C Graesser. 2006. Predicting affective states expressed through an emote-aloud procedure from AutoTutor's mixed-initiative dialogue. *International Journal of Artificial Intelligence in Education* 16, 1 (2006), 3–28.

[12] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. 2020. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929* (2020).

[13] Andrius Dzedzickis, Artūras Kaklauskas, and Vytautas Bucinskas. 2020. Human emotion recognition: Review of sensors and methods. *Sensors* 20, 3 (2020), 592.

[14] Paul Ekman and Wallace V Friesen. 1978. Facial action coding system. *Environmental Psychology & Nonverbal Behavior* (1978).

[15] Arthur Graesser, Patrick Chipman, Brandon King, Bethany McDaniel, and Sidney D'Mello. 2007. Emotions and learning with auto tutor. *Frontiers in Artificial Intelligence and Applications* 158 (2007), 569.

[16] Arthur C Graesser. 2020. Emotions are the experiential glue of learning environments in the 21st century. *Learning and Instruction* 70 (2020), 101212.

[17] Catherine Grainger, David M Williams, and Sophie E Lind. 2016. Metacognitive monitoring and control processes in children with autism spectrum disorder: Diminished judgement of confidence accuracy. *Consciousness and Cognition* 42 (2016), 65–74.

[18] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[19] Steve Higgins, Maria Katsipataki, AB Villanueva-Aguilera, Robbie Coleman, P Henderson, LE Major, R Coe, and Danielle Mason. 2016. The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit. (2016).

[20] iMontion. 2018. Attention tool. https://imotions.com/. [Online; accessed 28-December-2022].

[21] Shan Li and Weihong Deng. 2020. Deep facial expression recognition: A survey. *IEEE transactions on affective computing* (2020).

[22] Adam Linson, Yucheng Xu, Andrea R English, and Robert B Fisher. 2022. Identifying Student Struggle by Analyzing Facial Movement During Asynchronous Video Lecture Viewing: Towards an Automated Tool to Support Instructors. In *International Conference on Artificial Intelligence in Education*. Springer, 53–65.

[23] Katie Maras, Tim Gamble, and Mark Brosnan. 2019. Supporting metacognitive monitoring in mathematics learning for young people with autism spectrum disorder: A classroom-based study. *Autism* 23, 1 (2019), 60–70.

[24] Napoleon A Montero et al. 2021. The Impact of a Metacognitive Intervention using IMPROVE Model on Grade 7 Students' Metacognitive Awareness in Mathematics. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12, 3 (2021), 3881–3894.

[25] Juliana Gioia Negrão, Ana Alexandra Caldas Osorio, Rinaldo Focaccia Siciliano, Vivian Renne Gerber Lederman, Elisa Harumi Kozasa, Maria Eloisa Famá D'Antino, Anderson Tamborim, Vitor Santos, David Leonardo Barsand de Leucas, Paulo Sergio Camargo, et al. 2021. The child emotion facial expression set: a database for emotion recognition in children. *Frontiers in psychology* 12 (2021), 1352.

[26] Reinhard Pekrun. 2006. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review* 18, 4 (2006), 315–341.

[27] Min Peng, Zhan Wu, Zhihao Zhang, and Tong Chen. 2018. From macro to micro expression recognition: Deep learning on small datasets using transfer learning. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. IEEE, 657–661.

[28] Andrey V Savchenko, Lyudmila V Savchenko, and Ilya Makarov. 2022. Classifying emotions and engagement in online learning based on a single facial expression recognition neural network. *IEEE Transactions on Affective Computing* 13, 4 (2022), 2132–2143.

[29] Alyssa CP Sawyer, Paul Williamson, and Robyn Young. 2014. Metacognitive processes in emotion recognition: Are they different in adults with Asperger's disorder? *Journal of Autism and Developmental Disorders* 44, 6 (2014), 1373–1382.

[30] Michelle Taub, Roger Azevedo, Ramkumar Rajendran, Elizabeth B Cloude, Gautam Biswas, and Megan J Price. 2021. How are students' emotions related to the accuracy of cognitive and metacognitive processes during learning with an intelligent tutoring system? *Learning and Instruction* 72 (2021), 101200.

[31] Gabriella Tisza, Konstantinos Tsiakas, and Panos Markopoulos. 2022. Exploring the Relationship between Self-Regulation and Fun in Learning. *Available at SSRN 4047740* (2022).

[32] Konstantinos Tsiakas, Emilia Barakova, Javed Vassilis Khan, and Panos Markopoulos. 2020. BrainHood: towards an explainable recommendation system for self-regulated cognitive training in children. In *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. 1–6.

[33] Elisabeth Vogl, Reinhard Pekrun, and Kristina Loderer. 2021. Epistemic Emotions and Metacognitive Feelings. In *Trends and Prospects in Metacognition Research across the Life Span*. Springer, 41–58.

[34] Feng Wang, Weiyang Liu, Haijun Liu, and Jian Cheng. 2018. Additive Margin Softmax for Face Verification. *arXiv preprint arXiv:1801.05599* (2018).

[35] Philip H Winne. 2011. A cognitive and metacognitive analysis of self-regulated learning. *Handbook of self-regulation of learning and performance* (2011), 15–32.

[36] Fanglei Xue, Qiangchang Wang, and Guodong Guo. 2021. Transfer: Learning relation-aware facial expression representations with transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 3601–3610.