VALIDATING BABY FACEREADER TO ANALYZE INFANT AFFECTIVE AND COGNITIVE RESPONSES

A. Maroulis[1], J.J.M. Theuws[2], L.P.J.J. Noldus[2]

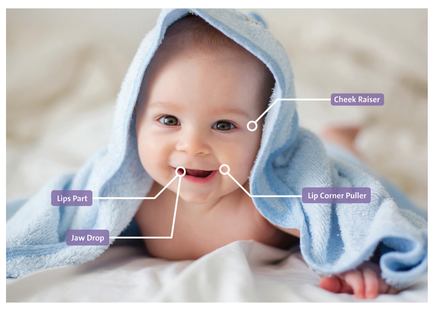
[1]VicarVision, Amsterdam, The Netherlands

[2]Noldus Information Technology BV, Wageningen, The Netherlands {l.noldus@noldus.nl]

INTRODUCTION

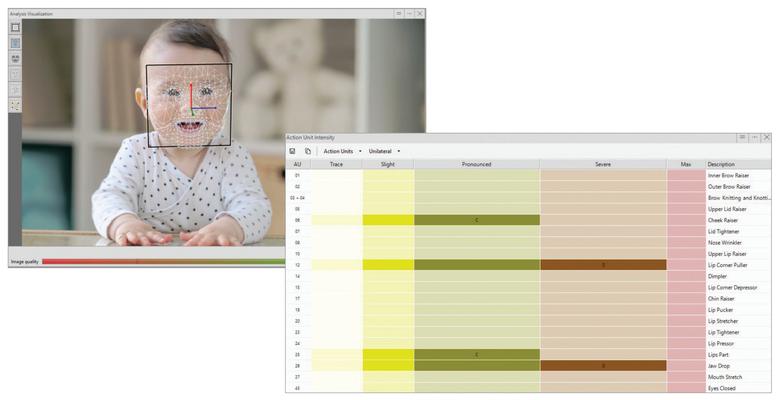
Facial expression (FE) analysis is one way to evaluate affective and cognitive responses of infants *(Desrosiers, Harrison, & Letham, 2015; Hannigen, 2014; Nakayama, 2013; Oster, 2005)*. Infants often do not provide coherent verbal feedback, making FE analysis particularly insightful when trying to evaluate responses to audio, visual, or sensory stimuli. The most comprehensive method to measure infant FEs is *Oster's BabyFACS (2016)*, detailing distinct facial Action Units (AUs) infants can produce.

Manually coding AUs is a cumbersome process. To overcome this, a number of computer vision solutions have been developed to automate this process for adult faces *(McDuff, Mahmoud, et al.,2016; Noldus, 2019; De laTorre, Chu, et al., 2015; Bartlett, Littlewort, et al., 2006)*. Similar solutions for infant FE have received little attention *(Zamzmi, Pai et al., 2016; Messinger, Mahoor, Chow, & Cohn, 2009)* largely due to challenges in the physiology of the infant face: less visible eyebrows and less wrinkles due to subcutaneous baby-fat, make the detection of FEs challenging. This paper presents the validation results of a solution to automatically measure infant FEs, Baby FaceReader (BFR).



METHODS

BFR automatically classifies AUs associated to typical expressions (e.g. smiling - AU6+12, crying - AU20). We evaluated the performance of BFR using a set of 67 images present in the BabyFACS manual. BFR uses neural networks trained on over 10.000 images from university studies and internet sources, excluding images from the BabyFACS manual.



Baby FaceReader with Action Unit Classification.

RESULTS

We evaluated BFR AU classifications calculating recall, precision, F1 and accuracy scores for all BabyFACS manual images. A brief description of these terms:

- **RECALL**
  A ratio of correct detection versus annotations of AUs
  For example a recall of 0.75 for AU15 indicates that 75% of the annotated images with AU15 are classified as such by BFR.
- **PRECISION**
  Ratio correct detection versus total detection of AUs
  For example, in the case of AU1 the FaceReader classification is correct 70% of the time.
- **F1**
  Measured as 2 * ((precision * recall) / (precision + recall))
- **ACCURACY**
  Represents the percentage of correct classifications, computed by dividing the number of correctly classified images (both positive and negative) by the total number of images.

This table shows the scores per AU. "Present" indicates the frequency of the AU in the dataset.

AU	Present	Recall	Precision	F1	Accuracy
1	10	0,438	0,700	0,538	0,821
2	11	0,333	0,182	0,235	0,806
3+4	9	0,375	1,000	0,545	0,776
5	10	0,583	0,700	0,636	0,881
6	30	0,636	0,933	0,757	0,731
7	20	0,581	0,900	0,706	0,776
9	8	0,333	0,375	0,353	0,836
12	11	0,306	1,000	0,468	0,627
15	8	0,750	0,375	0,500	0,910
17	13	0,556	0,769	0,645	0,836
20	11	0,583	0,636	0,609	0,866
25	46	0,929	0,848	0,886	0,851
26	21	0,600	0,857	0,706	0,776
27	12	0,846	0,917	0,880	0,955
43	11	0,579	1,000	0,733	0,881
Average	**15,4**	**0,562**	**0,746**	**0,613**	**0,822**

DISCUSSION

This study aims to establish BFR's utility in studying affect in infants and in research areas such as developmental disorders. The results indicate that F1 performance is reaching 0.7 for a number of AUs, a standard set by FACS *(Ekman, Friesen & Hager, 2002)*. Future work aims to further improve classifications and validate unilaterally expressed AUs. Unilateral AUs are expressed on one side of the face and can be detected by BFR.

BFR not only automatically detects AUs, but also head orientation and gaze direction. These measurements could indicate if an infant is looking at a stimulus at the time of presentation. This can greatly reduce the time needed to analyze data from eye tracking studies with infants.

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 777394. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and AUTISM SPEAKS, Autistica, SFARI.