Teaching Emotion Expressions to a Human Companion Robot using Deep Neural Architectures

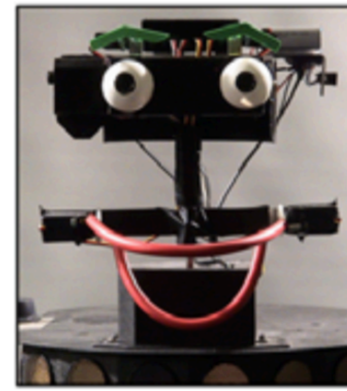
Nikhil Churamani, Matthias Kerzel, Erik Strahl, Pablo Barros and Stefan Wermter

Knowledge Technology Institute
Department of Informatics
Universität Hamburg

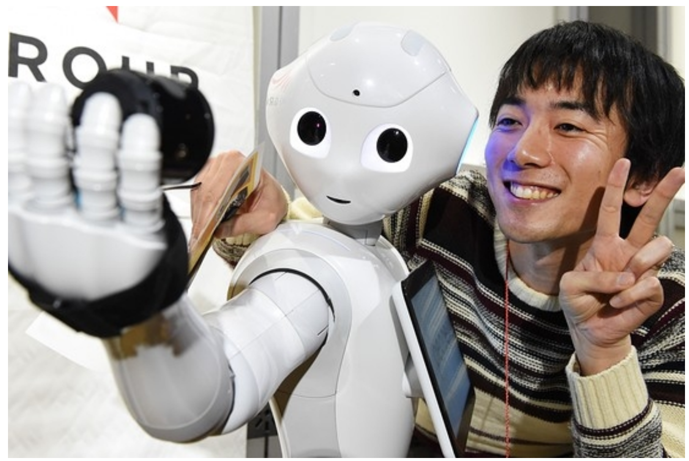
http://www.informatik.uni-hamburg.de/WTM/

Expressions in Agents

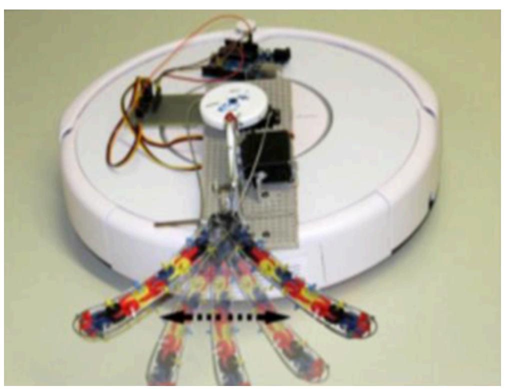
Rea et al., 2012



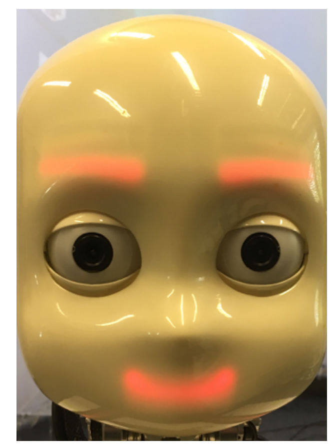
MIRA: The Robot Head



Pepper by Aldebaran Robotics and SoftBank

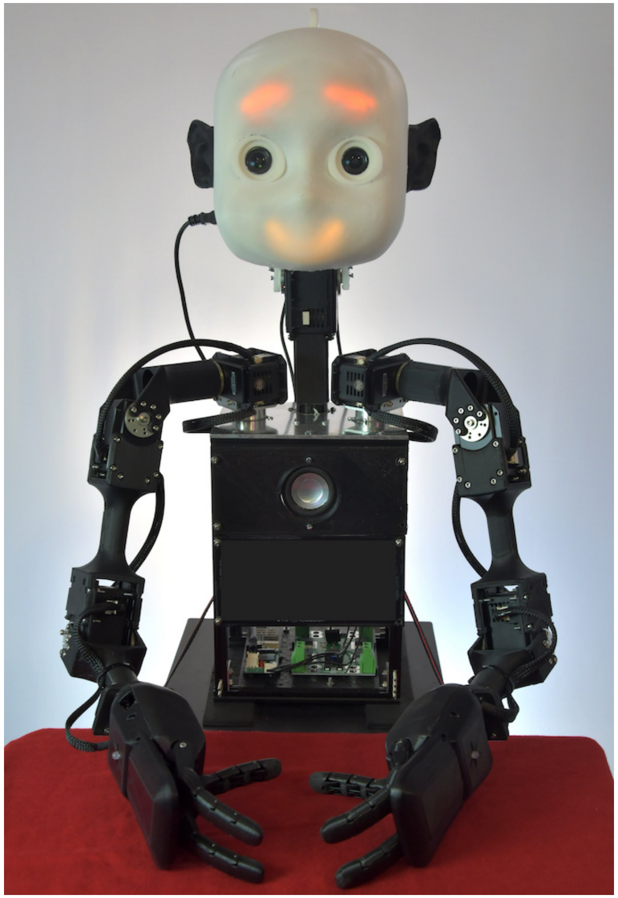


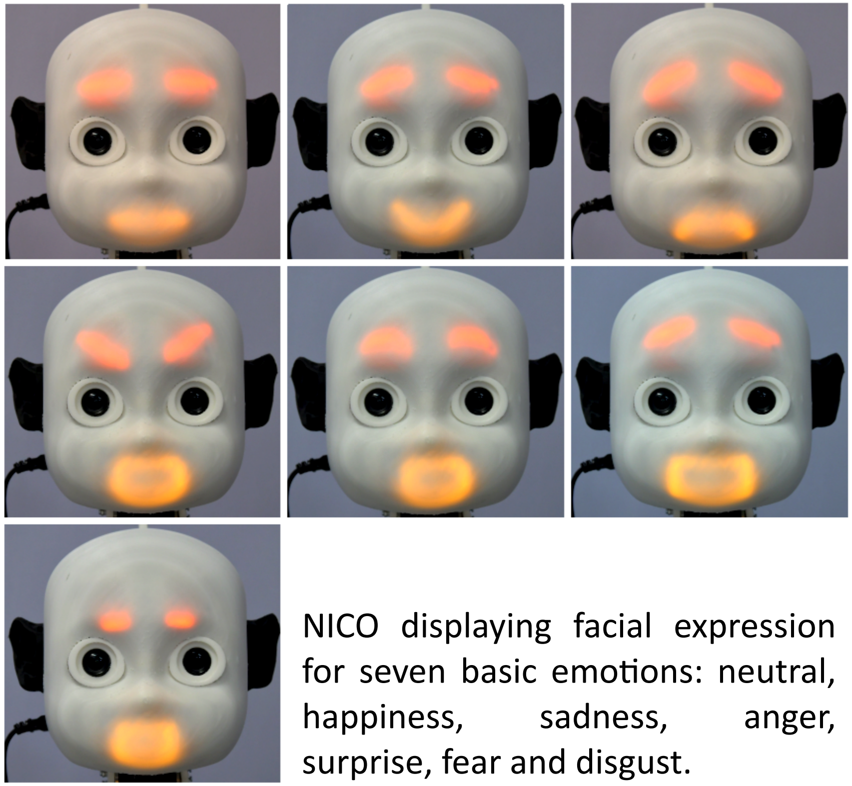
Singh and Young, 2012



iCub Robot Head

NICO: Neuro-Inspired Companion

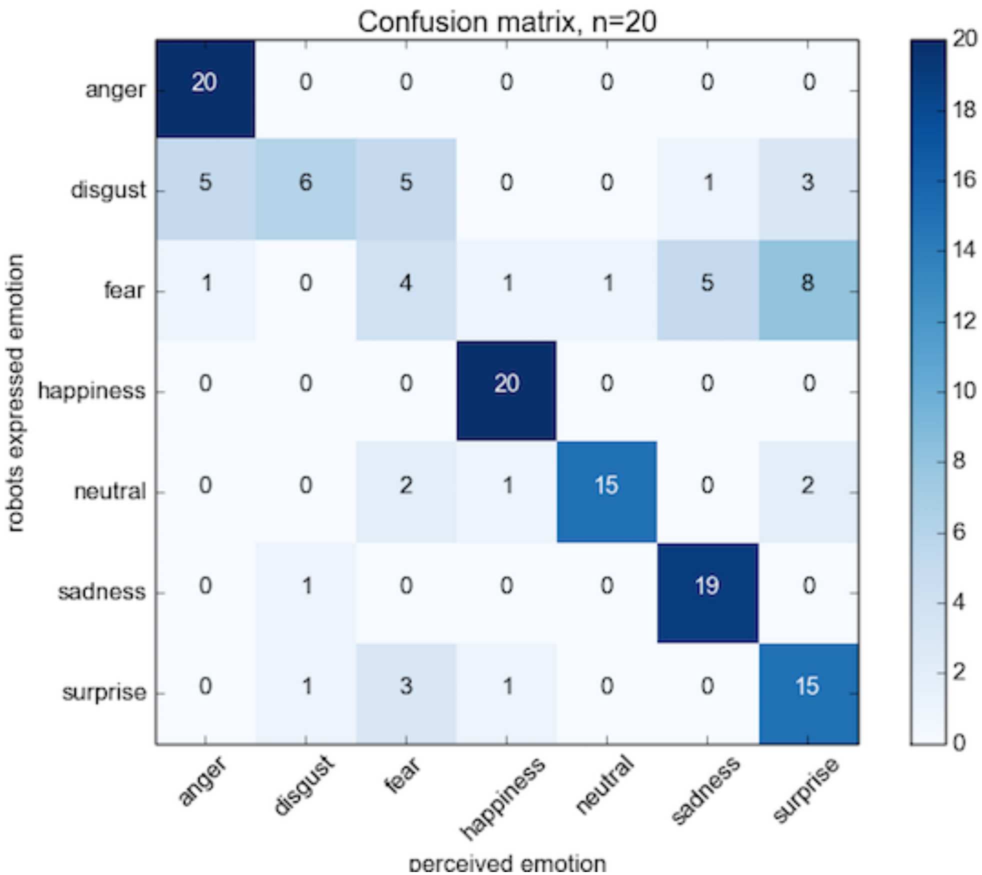




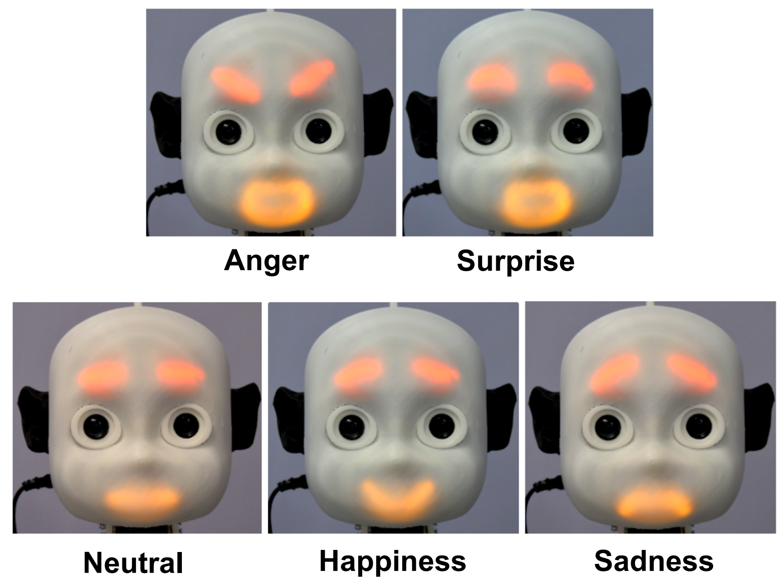
NICO displaying facial expression for seven basic emotions: neutral, happiness, sadness, anger, surprise, fear and disgust.

Ekman, 1992

NICO: Neuro-Inspired Companion

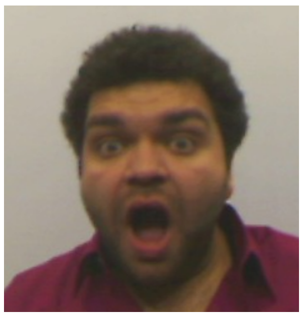


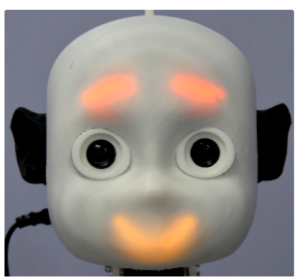
Confusion Matrix for 20 users' labelling of the emotion expressed by NICO



Five expressions chosen for the study based on users' perception.

User Interaction Scenario





Input Phase

- User Express an Emotion
- Robot selects and performs an expression

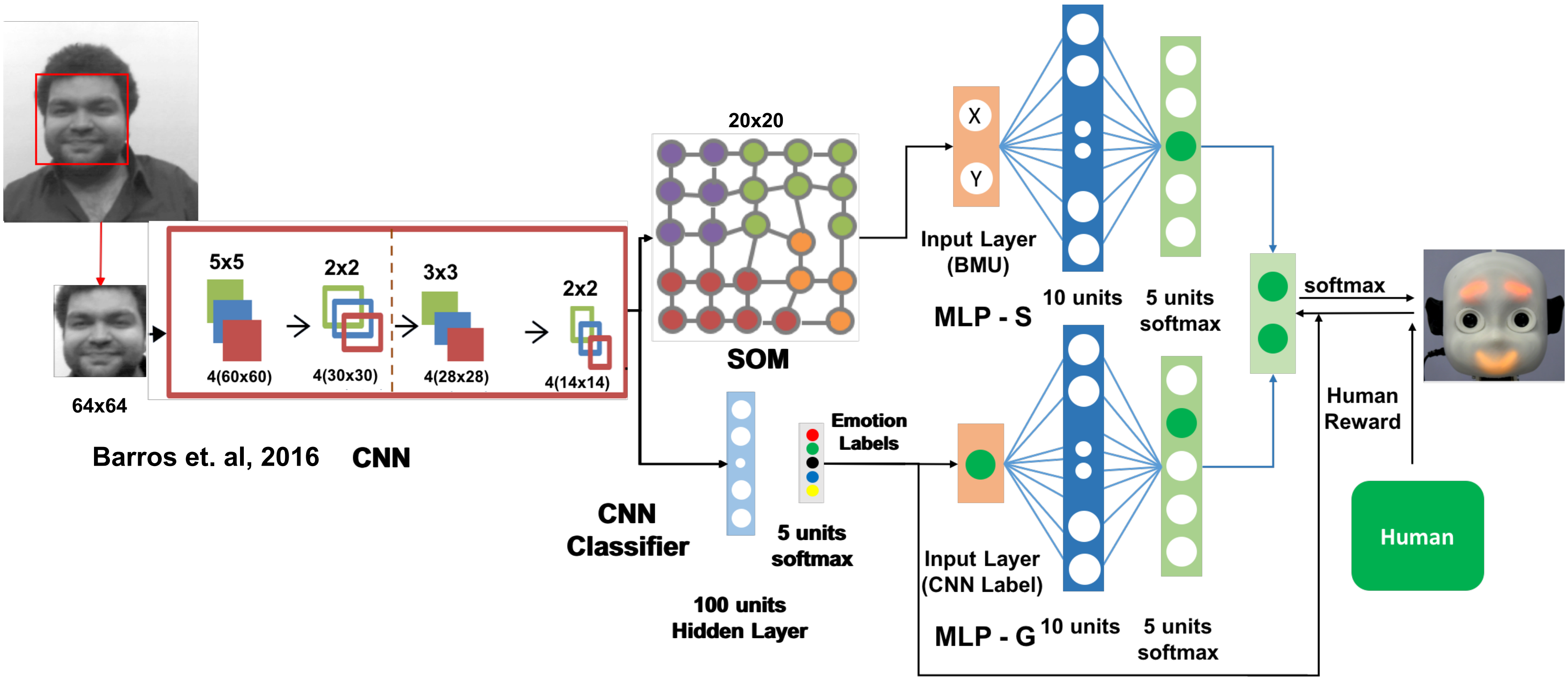
Reward Phase

- User then evaluates the expression performed
- Robot learns emotion-expression association

Sample Interactions

Sample Inputs for Five emotions

Proposed Model

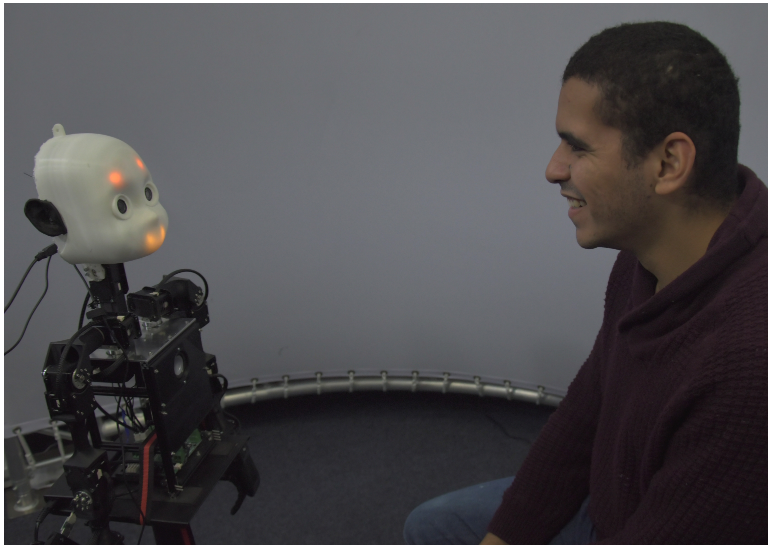


Experiment Setup

- Experiment conducted in two User Groups:
  - User Group – I:
    - Weights initialised randomly for each participant.
  - User Group – II:
    - Weights initialised once for the whole group
- 10 interaction rounds with each participant each of which required the participant to enact 5 emotions viz. Anger, Happiness, Neutral, Sadness and Surprise.
- Robot action rewarded by an affirmative (Happiness) or a negative (Anger) expression.

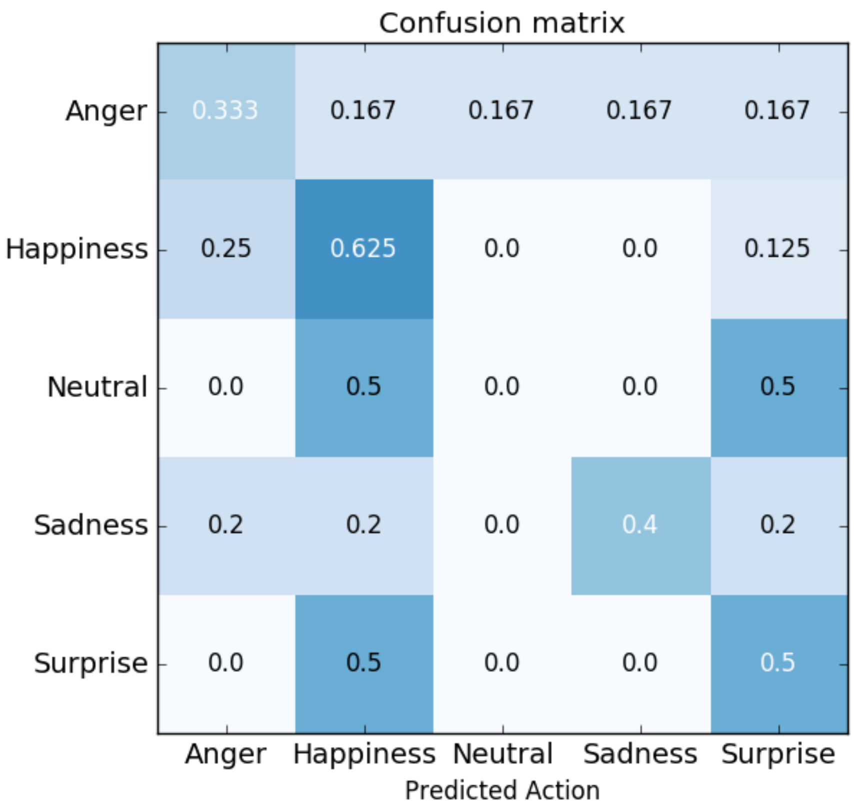
Experiment Setup

Camera Booth for participants

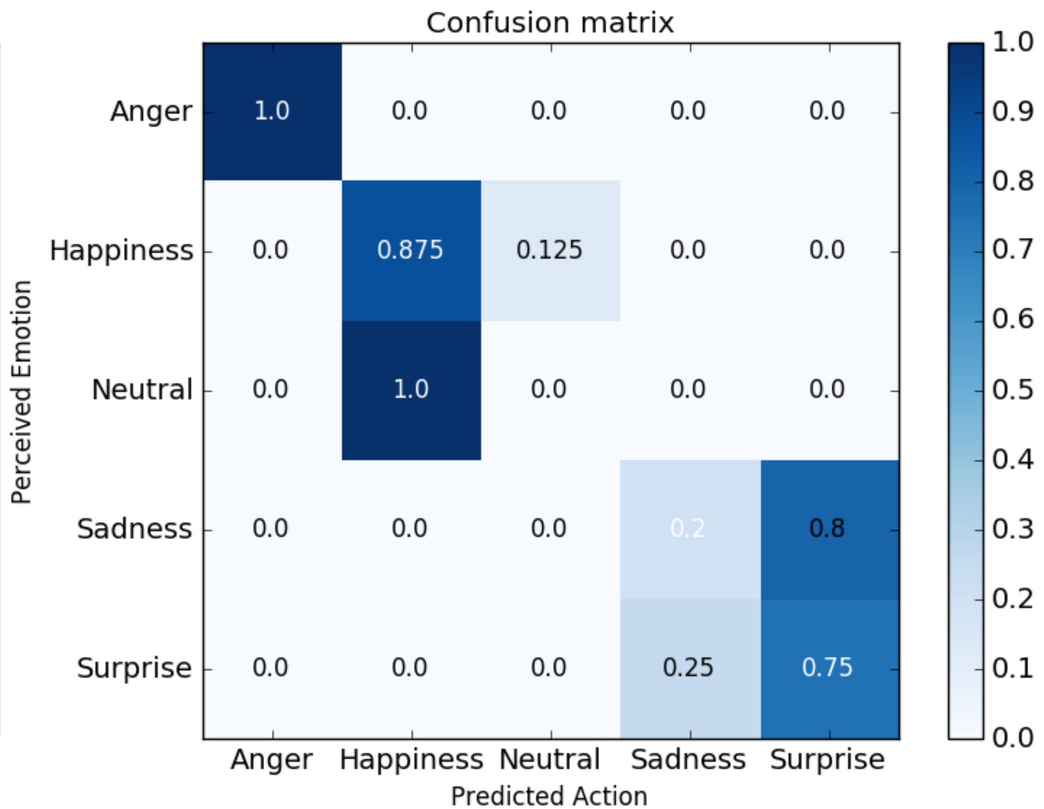


Interaction with the Robot

Experiment Results – User Group - I

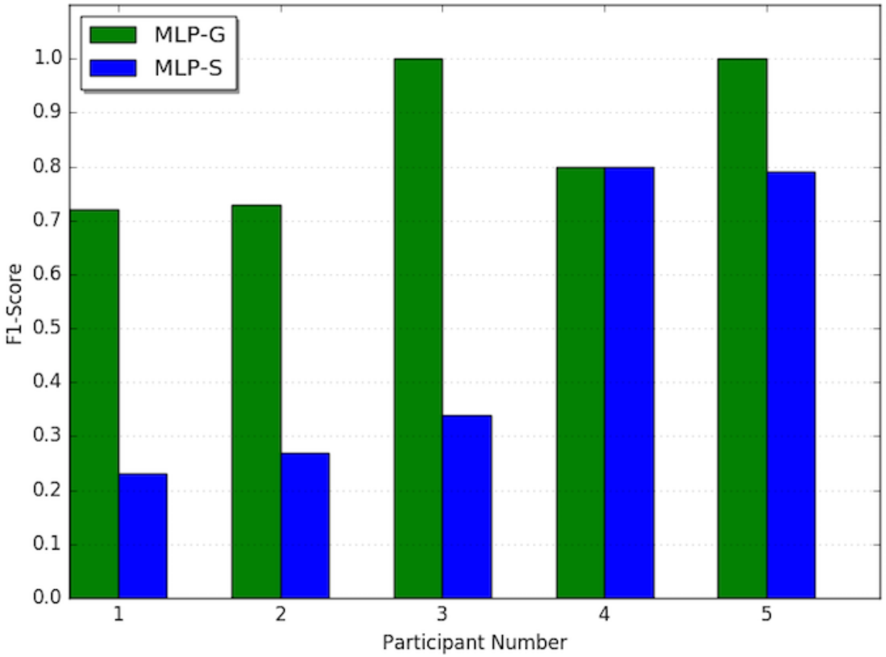


Normalised Confusion Matrix for **MLP-S** for the last epoch scores for 5 participants

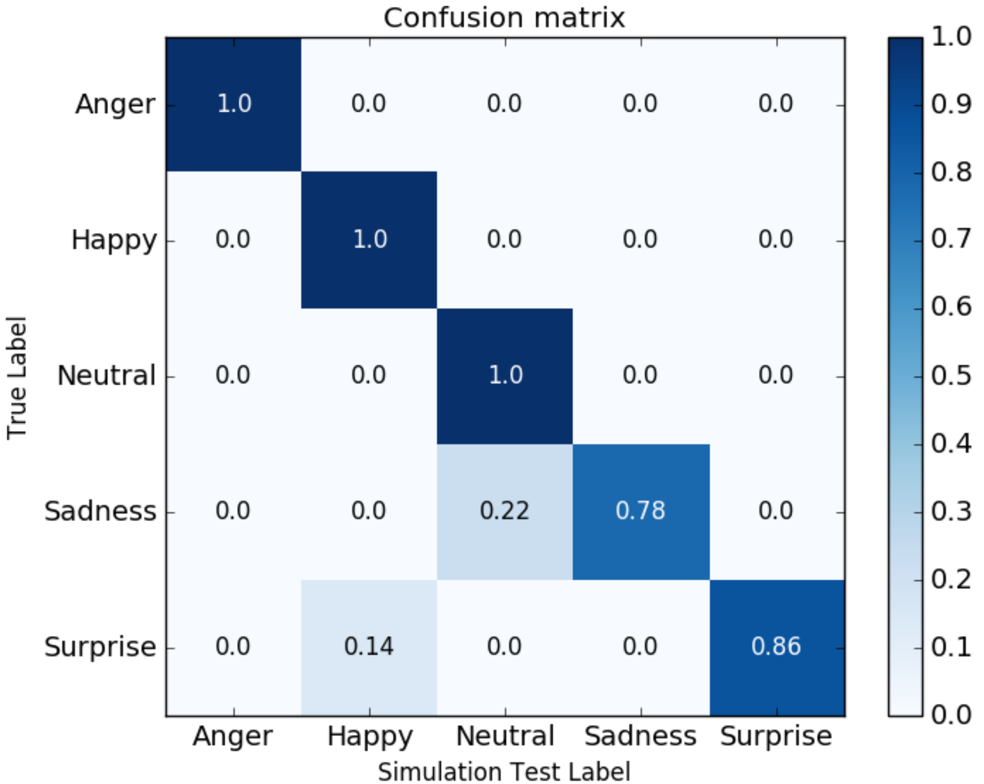


Normalised Confusion Matrix for **MLP-G** for the last epoch scores for 5 participants

Experiment Results – User Group - II

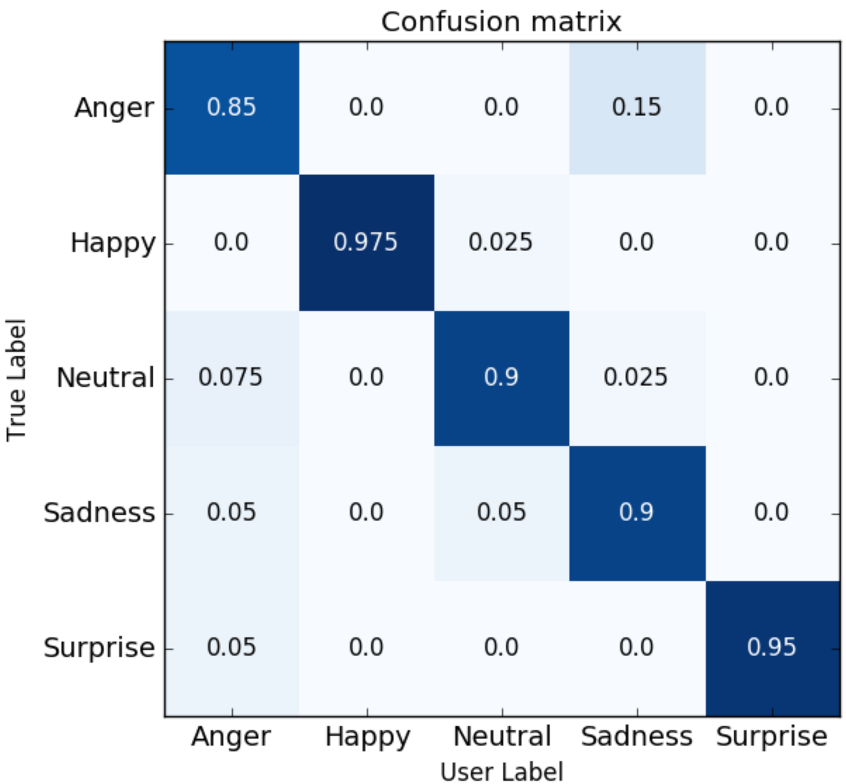


F1-Score for **continued training** with 5 participants

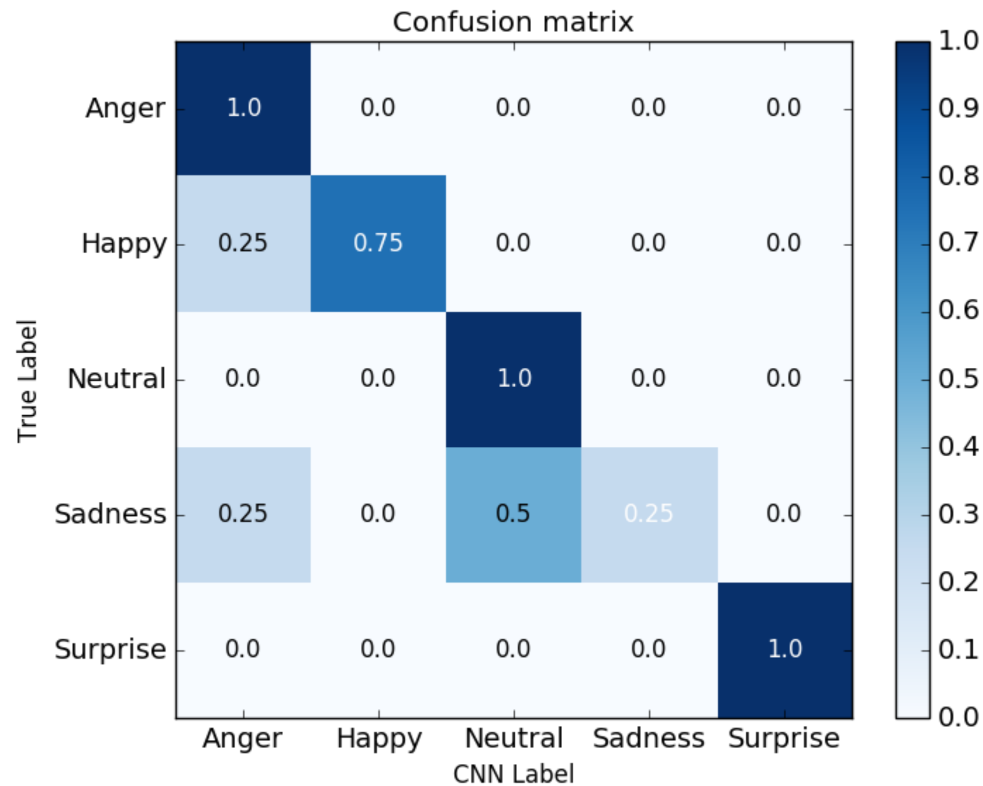


Final Confusion Matrix for the **Extended Simulation**

CNN compared to Human Annotators



Normalised Confusion Matrix (**Annotators**) for 20 interactions
Cohen's Weighted Kappa (κ) = 0.86
Kendall Tau distance (K) = 0.834

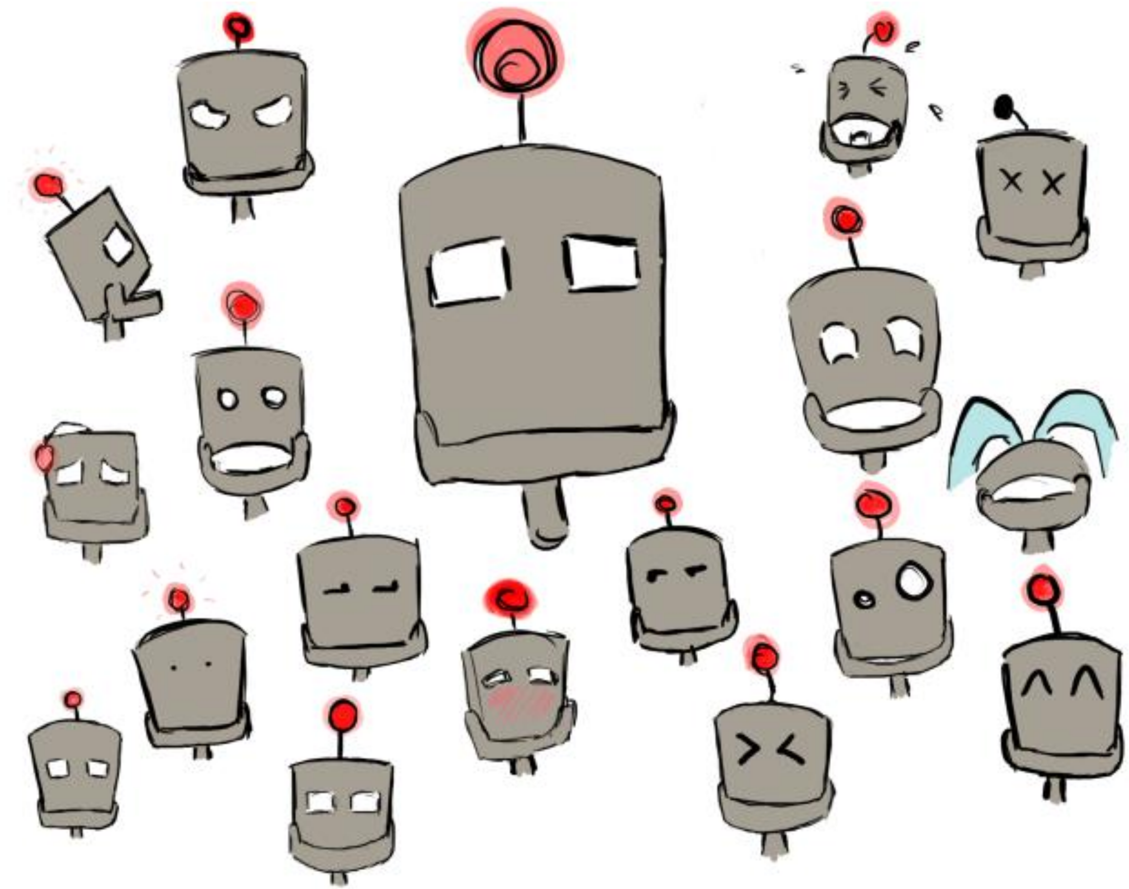


Normalised Confusion Matrix (**CNN Labels**) for 20 interactions
Cohen's Weighted Kappa (κ)= 0.81
Kendall Tau distance (K) = 0.829

Conclusion

- Two branches of the model take inspiration from the generic and specific perception of emotions in humans.
- **MLP-G** acts on the generic, “first impression” model of the network whereas the **MLP-S** learns to adapt to a particular individual.
- Continued training improves performance. MLP-G provides initial guidance but eventually MLP-S outperforms MLP-G adapting to each individual.
- **Future Work:**
  - Learned actions can be enhanced to cover a continuous space dealing even with multiple modalities.
  - Exploring a developmental learning approach for learning expressions.

Q & A



catmoo.co.uk

References

[1] D. Rea, J. Young, and P. Irani, “*The Roomba mood ring: an ambient-display robot*”, Proc. seventh Annual ACM/IEEE Int. Conf. Human-Robot Interact. - HRI ’12, pp. 217–218, 2012.

[2] A. Singh and J. E. Young, “*Animal-inspired human-robot interaction: A robotic tail for communicating state*”, Human-Robot Interact. (HRI), 2012 7th ACM/IEEE Int. Conf., pp. 237–238, 2012.

[3] J. C. Murray, S. Wermter, and M. Knowles, “MIRA: a learning multimodal interactive robot agent”, in *Hybrid Intelligent Systems, 2008. HIS’08. Eighth International Conference on*, 2008, pp. 947–950.

[4] Metta, Giorgio, et al. "The iCub humanoid robot: An open-systems platform for research in cognitive development." *Neural Networks* 23.8 (2010): 1125-1134.

[5] F. Hegel, F. Eyssel, and B. Wrede, “The social robot ’Flobi’: Key concepts of industrial design,” in *19th International Symposium in Robot and Human Interactive Communication*. IEEE, 2010, pp. 107–112. P.

[6] P. Ekman, “An argument for basic emotions,” Cognition & Emotion, vol. 6, no. 3-4, pp. 169–200, 1992.

[7] P. Barros and S. Wermter, “Developing crossmodal expression recognition based on a deep neural model,” Adaptive Behavior, vol. 24, no. 5, pp. 373–396, 2016.

[8] W. B. Knox and P. Stone, “TAMER: Training an agent manually via evaluative reinforcement,” in IEEE 7th International Conference on Development and Learning. IEEE, 2008, pp. 292–297.

[9] Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, “The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression,” in *Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010)*, San Francisco, USA, 2010, pp. 94–101.

Teaching Emotion Expressions to a Human Companion Robot using Deep Neural Architectures

Nikhil Churamani, Matthias Kerzel, Erik Strahl,
Pablo Barros and Stefan Wermter

Knowledge Technology Institute
Department of Informatics
Universität Hamburg

http://www.informatik.uni-hamburg.de/WTM/