

perceived personality rating as compared to the database as a whole in order to make accurate predictions.

4 CONCLUSIONS AND FUTURE WORKS

Small group interactions is most commonly found in workplace and school situations. Personal attributes of personality traits affect both individual's behaviors and his/her interaction process with other members of the group. The dynamics of the interaction process is key in dictating the final group performance outcomes. Due to the importance in objectively understanding the intricate behavior dynamics in small group interactions, there has been an increasing interest on research that build computational frameworks to automatically assess and recognize individual's personality traits. In this work, we propose an IM-aBLSTM framework that models the vocal behaviors of both the target speaker and his contextual interlocutors to improve the prediction performance on the score of ten different personality traits in the ELEA corpus.

Our method achieves a promising UAR of 87.9% over ten personality traits in high vs. low classification task, and it obtains an average Pearson correlation of 0.405 for regression task. IM-aBLSTM framework outperforms the current state-of-art personality recognition accuracy on this corpus. Furthermore, our analyses on the high interlocutor-modulated attention regions demonstrate that prosodic variations indeed vary according to each individual's personality trait. Also, the proposed IM-aBLSTM makes its improved recognition by concentrating on the interactive vocal behaviors during the conversational segments of speakers (preceding - target - follow-up) where there is more changes in the personality traits.

This work presents a preliminary personality prediction result by modeling vocal behaviors between interlocutors via embedding interaction-based attention mechanism in a BLSTM. Our future work will focus on leveraging multiple behavior modalities to advance our algorithm in modeling the relationship between target speaker and interlocutors to improve regression correlations across these ten personality traits. Additionally, we would also investigate algorithmic frameworks that jointly models individual behaviors at the group level to predict the final group performance outcome as they engage in a variety of small group interaction contexts.

REFERENCES

- [1] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2014. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473* (2014).
- [2] Cigdem Beyan, Francesca Capozzi, Cristina Becchio, and Vittorio Murino. 2018. Prediction of the Leadership Style of an Emergent Leader Using Audio and Visual Nonverbal Features. *IEEE Transactions on Multimedia* 20, 2 (2018), 441–456.
- [3] P Boersma and D Weenink. 2003. Praat-A system for doing phonetics by computer [Computer Software]. *The Netherlands: Institute of Phonetic Sciences, University of Amsterdam* (2003).
- [4] Jean Carletta, Simone Ashby, Sebastien Bourban, Mike Flynn, Mael Guillemot, Thomas Hain, Jaroslav Kadlec, Vasilis Karaiskos, Wessel Kraaij, Melissa Kronenthal, et al. 2005. The AMI meeting corpus: A pre-announcement. In *International Workshop on Machine Learning for Multimodal Interaction*. Springer, 28–39.
- [5] Oya Celiktutan and Haticce Gunes. 2017. Automatic prediction of impressions in time and across varying context: Personality, attractiveness and likeability. *IEEE Transactions on Affective Computing* 8, 1 (2017), 29–42.
- [6] Lei Chen, R Travis Rose, Ying Qiao, Irene Kimbara, Fey Parrill, Haleema Welji, Tony Xu Han, Jilin Tu, Zhongqiang Huang, Mary Harper, et al. 2005. VACE multimodal meeting corpus. In *International Workshop on Machine Learning for Multimodal Interaction*. Springer, 40–51.
- [7] Walter H Crockett. 1955. Emergent leadership in small, decision-making groups. *The Journal of Abnormal and Social Psychology* 51, 3 (1955), 378.
- [8] Michael R Cunningham. 1977. Personality and the structure of the nonverbal communication of emotion. *Journal of Personality* 45, 4 (1977), 564–584.
- [9] Sheng Fang, Catherine Achard, and Séverine Dubuisson. 2016. Personality classification and behaviour interpretation: An approach based on feature categories. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*. ACM, 225–232.
- [10] Howard S Friedman, M Robin DiMatteo, and Angelo Taranta. 1980. A study of the relationship between individual differences in nonverbal expressiveness and factors of personality and social interaction. *Journal of Research in Personality* 14, 3 (1980), 351–364.
- [11] Daniel Gatica-Perez, Oya Aran, and Dinesh Jayagopi. 2017. *Analysis of Small Groups*. Cambridge University Press, 349–367. <https://doi.org/10.1017/9781316676202.025>
- [12] James Gibson, Dogan Can, Panayiotis Georgiou, David C Atkins, and Shrikanth S Narayanan. 2017. Attention networks for modeling behaviors in addiction counseling. In *Proc. Interspeech*.
- [13] Lewis R Goldberg. 1990. An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology* 59, 6 (1990), 1216.
- [14] Alex Graves and Jürgen Schmidhuber. 2005. Framework phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks* 18, 5-6 (2005), 602–610.
- [15] J Richard Hackman and Charles G Morris. 1975. Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. In *Advances in experimental social psychology*. Vol. 8. Elsevier, 45–99.
- [16] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.
- [17] Shan-Wen Hsiao, Hung-Ching Sun, Ming-Chuan Hsieh, Ming-Hsueh Tsai, Yu Tsao, and Chi-Chun Lee. 2017. Toward Automating Oral Presentation Scoring during Principal Certification Program using Audio-video Low-level Behavior Profiles. *IEEE Transactions on Affective Computing* (2017).
- [18] Dinesh Babu Jayagopi, Hayley Hung, Chuohao Yeo, and Daniel Gatica-Perez. 2009. Modeling dominance in group conversations using nonverbal activity cues. *IEEE Transactions on Audio, Speech, and Language Processing* 17, 3 (2009), 501–513.
- [19] Heysem Kaya, Alexey A Karpov, and Albert Ali Salah. 2015. Fisher vectors with cascaded normalization for paralinguistic analysis. In *Sixteenth Annual Conference of the International Speech Communication Association*.
- [20] Diederik Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
- [21] Yun-Shao Lin and Chi-Chun Lee. 2017. Deriving Dyad-Level Interaction Representation using Interlocutors Structural and Expressive Multimodal Behavior Features. *Proc. Interspeech 2017* (2017), 2366–2370.
- [22] Joseph Edward McGrath. 1964. *Social psychology: A brief introduction*. Holt, Rinehart and Winston.
- [23] Joseph Edward McGrath. 1984. *Groups: Interaction and performance*. Vol. 14. Prentice-Hall Englewood Cliffs, NJ.
- [24] Seyedmahdad Mirsamadi, Emad Barsoum, and Cha Zhang. 2017. Automatic speech emotion recognition using recurrent neural networks with local attention. In *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*. IEEE, 2227–2231.
- [25] Shogo Okada, Oya Aran, and Daniel Gatica-Perez. 2015. Personality trait classification via co-occurrent multiparty multimodal event discovery. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ACM, 15–22.
- [26] Fabio Piansi, Massimo Zancanaro, Bruno Lepri, and Alessandro Cappelletti. 2007. A multimodal annotated corpus of consensus decision making meetings. *Language Resources and Evaluation* 41, 3-4 (2007), 409–429.
- [27] Jurgen Ruesch, Weldon Kees, Robert Goodloe Harper, Robert G Harper, Arthur N Wiens, and Joseph D Matarazzo. 1978. *Nonverbal Communication: The State of the Art*. Vol. 65. Univ of California Press.
- [28] Jorge Sánchez, Florent Perronnin, Thomas Mensink, and Jakob Verbeek. 2013. Image classification with the fisher vector: Theory and practice. *International journal of computer vision* 105, 3 (2013), 222–245.
- [29] Dairazalia Sanchez-Cortes, Oya Aran, Marianne Schmid Mast, and Daniel Gatica-Perez. 2012. A nonverbal behavior approach to identify emergent leaders in small groups. *IEEE Transactions on Multimedia* 14, 3 (2012), 816–832.
- [30] Barry Schwartz, Abraham Tesser, and Evan Powell. 1982. Dominance cues in nonverbal behavior. *Social Psychology Quarterly* (1982), 114–120.
- [31] Shikhar Sharma, Ryan Kiro, and Ruslan Salakhutdinov. 2015. Action recognition using visual attention. *arXiv preprint arXiv:1511.04119* (2015).
- [32] R Timothy Stein. 1975. Identifying emergent leaders from verbal and nonverbal communications. *Journal of Personality and Social Psychology* 32, 1 (1975), 125.
- [33] Alessandro Vinciarelli, Maja Pantic, and Hervé Bourlard. 2009. Social signal processing: Survey of an emerging domain. *Image and vision computing* 27, 12 (2009), 1743–1759.
- [34] Noreen M Webb. 1982. Student interaction and learning in small groups. *Review of Educational research* 52, 3 (1982), 421–445.