Toward Computational Models of Team Effectiveness with Natural Language Processing

Randall Spain, Michael Geden, Wookhee Min, Bradford Mott, and James Lester
Center for Educational Informatics, North Carolina State University, Raleigh, NC 27695, USA
{rdspain, mageden, wmin, bwmott, lester}@ncsu.edu

Abstract. Team communication provides a rich source of data about team processes that can impact team performance. It can provide information about team structure, team roles, connectedness, a team's cognitive state, and situational status. Analyzing team communication can thereby provide deep insight into processes underlying team collaboration and coordination. Traditional approaches for investigating team processes through dialogue analysis have historically relied upon human annotation, a process that is extraordinarily resource-intensive for the team training research community and cannot be utilized for real-time team assessment. In this paper, we discuss techniques that we are exploring to develop a team communication analysis toolkit that can perform real-time end-to-end natural language analysis on team members' spoken dialogue and generate team dialogue analytics that drive adaptive scaffolding. We discuss how team communication has traditionally been analyzed and describe the basis of our current work investigating a deep learning-based natural language processing framework that will support automated tagging of team discourse and predictions of team performance.

Keywords: Team Tutoring, Team Communication Assessment, Natural Language Processing.

1 Introduction

Investigating individuals' communication during team training and collaborative problem-solving activities can provide insight into the rich processes underlying team collaboration and coordination [1, 2]. For instance, communication data can be used to investigate a broad array of problem-solving and teamwork phenomena, including (1) *team shared understanding* (e.g., shared understanding of the task and solution goals, idea generation and refinement, connections between ideas and tasks, knowledge co-construction) (2) *team coordination* (e.g., connected talk, turn taking, information and resource sharing, participation patterns, idea sharing) and (3) *team social regulation* (e.g., management of team roles and structure, division of labor).

Despite the insight that team dialogue and speech data offer for understanding team performance, dialogue analysis for team communication has historically been extraordinarily resource-intensive for the team training research community. Human annota-

Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR-WS.org/Vol-2501/paper4.pdf

tors spend dozens of hours coding segments of team communication from small datasets. Similarly, learners' natural language communication has not been usable for informing adaptive scaffolding decisions because researchers have historically not had access to sufficiently effective natural language processing technologies. Early work examining automated assessment of team discourse explored how latent semantic analysis (LSA) could be used to build linguistic models of team communication content, sequence, and structure [3]. While LSA can create semantic representations of language, it fails to include other critical streams of communication data such as prosody, phraseology, and syntactic structure that could be complementary for team discourse research and driving real-time adaptive scaffolding.

Recent advances in deep learning-based natural language processing (NLP) show significant promise for automatically analyzing team communication data and providing capabilities beyond those associated with semantic analysis and related techniques. Deep learning-based NLP can incorporate neural language models with multiple streams of linguistic data (e.g., semantics, syntactic structure, phonology, stylistics) and multilevel discourse features (e.g., individual team members, current task, environmental factors) to produce flexible, holistic representations of team processes in real-time. However, there are many open questions regarding how we can most effectively leverage advances in deep learning-based NLP to analyze team discourse to help researchers automatically assess team communication and team performance.

To begin to address these questions, we are launching a new collaborative effort between North Carolina State University and the U.S. Army Futures Command, Combat Capabilities Development Command - Simulation and Training Technology Center to investigate the design and development of a deep learning-based NLP framework to automatically analyze team communication data, parse it into classification schemes, and provide summary statistics of critical team communication features that can be used to analyze and identify antecedents of team performance. By analyzing team discourse during training episodes, the framework will be able to assess team communication content, quality, and information exchange features, and provide insights into team processes and cognitive states that could be used to inform team assessment and feedback policies in adaptive instructional systems.

In this paper, we discuss techniques and approaches that our team is exploring to develop a team communication analysis toolkit that can perform real-time end-to-end natural language analysis on team members' spoken dialogue and generate team dialogue analytics that drive adaptive scaffolding. We begin the paper by discussing how team communication has traditionally been analyzed and highlight how early LSA-based approaches have been used to help automate this process. Then we discuss how deep learning-based approaches can provide additional linguistic analysis capabilities for analyzing team discourse. The paper concludes with a discussion of the deep learning-based NLP pipeline we are developing to support the automated tagging of team discourse and how we plan to investigate the accuracy of the tool and its ability to predict team performance using a corpus of team communication data from a joint military training exercise.

2 Research Context

Team communication plays a critical role in team performance [4]. A prevalent finding in the team literature is that communication is integral to a number of team processes and behaviors that lead to effective team performance. Models of teamwork posit that communication can enhance team performance by facilitating and improving critical team processes such as team coordination and strategy formulation [5]. For instance, communication can serve as a primary conduit through which team members share information, clarify misunderstandings, and provide guidance to other team members. In addition, communication can contribute to the development of team emergent states such as team cognition, which can foster more effective team performance [6]. Communication is also argued to directly relate to team performance because it distributes critical task related information to team members that may impact the nature of team interdependence, team responsibilities, and team task demands [7]. Analyzing team communication can thereby provide deep insight into effective team processes.

2.1 Measuring Team Communication

Team communication can be broken down into a number of elements. Three distinct aspects of communication that are often investigated in the team literature are information exchange, phraseology, and closed-loop communication [8]. Information exchange refers to passing information between members, including passing the right information to the appropriate person without being asked and providing updates on tasks or environment states, which could impact team performance. For instance, high performance teams rapidly identify current and potential problems and develop and share appropriate responses to these issues through information exchange [9]. Phraseology refers to using consistent terminology, communicating precisely, and passing complete information to team members [8, 10]. Closed-loop communication is a communication style applied in many complex task domains wherein team members confirm and cross-check information to ensure information is properly received [11].

In a recent meta-analysis of team communication and performance, Marlow, Lacerenza, Paoletti, Burke and Salas identified two general ways in which these elements of team communication have been measured by researchers [12]. The first method involves asking team members to rate the extent to which information is freely and openly shared among team members or to rate the extent to which team members share their knowledge using validated rating scales. Alternatively, trained raters can be asked to assess the quality of communication behaviors in teams [7] or to use behaviorally-anchored rating scales to assess communication behaviors that are tied to specific scenario events [13]. These rating-based measurement approaches are typically used to assess communication *quality* within teams.

The second method for examining teamwork communication within empirically based studies involves analyzing transcripts of team communication and hand-coding team communication based on a pre-established coding scheme. The frequencies at which the coded categories emerge from the data can then, in turn, be correlated with team performance measures. This *frequency*-based method has been used in several studies to examine differences in team performance. For instance, Bowers, Jentsch, Salas, and Braun examined communication patterns between high and low performing flight crews and found that higher performing crews answered uncertainty, planning, and fact statements more consistently with acknowledgments or responses than did lower performing crews [11]. They also found higher performing teams were more likely to follow communications from air traffic control with planning statements compared to lower performing crews and were more likely to follow uncertainty statements with acknowledgement statements. Achille, Schulze, and Schmidt-Nielsen found that more experienced teams used proper terminology and more acknowledgement and identification statements than less experienced teams [14].

Despite the prevalence of using frequency-based approaches to evaluate team performance, researchers from the team and collaborative learning research communities have repeatedly criticized this method because it is extraordinarily resource intensive, can be highly subjective, and offers limited insight into the dynamic and evolving nature of team processes and performance [15, 16]. Furthermore, empirical investigations of team performance suggest that more communication is not always associated with better performance, thus strictly using count or frequency data offers a limited and equivocal lens for analyzing team communication [11].

2.2 Automatically Analyzing Team Communication Content

In an attempt to move towards automatic analysis of team communication data, several researchers have explored using computational methods to identify the semantic content of team discourse. For example, Foltz and colleagues used LSA to automatically categorize the content of team discourse and predict team performance in a number of different task domains [17]. LSA is a statistical computational method that decomposes documents to a vector representation of their semantic meaning by applying singular value decomposition on a matrix of word frequencies by documents. The LSA vector representations can then be used to find the similarity of two documents by taking the cosine distance of their respective semantic vectors. Early investigations found LSA to be generally successful at automatically tagging discourse segments [16] and that the outputs of LSA could be used to examine team communication content [17], detecting patterns of communication and identifying locations of communication breakdowns [18], and analyzing team cognition [19]. Moreover, they found that by applying LSA-based algorithms, they could analyze and tag an hour of team transcripts in under a minute [16], thus highlighting a critical advantage of using statistical-based text analytic approaches compared to manual coding and tagging of team transcripts.

The LSA-based approaches used in previous work provide a promising initial foray into automated team communication assessment, however, the approach suffers from several limitations. LSA is unable to account for a number of linguistic features which detrimentally impacts the quality of its semantic representations, such as polysomy, word ordering, and syntactic structure. Another limitation of LSA-based approaches is that the cosine method does not easily incorporate other linguistic features (phonetics, phonology, and stylistics) or hierarchical representations of the discourse (e.g., who is speaking, what task is being performed, environmental factors), and the cosine similarity metric is confounded by the length of the documents being compared [20]. Additionally, LSA is more sensitive to corpora training [21] and performs less well than modern neural language models such as fastText, ELMo, and BERT [22] on a wide range of NLP tasks.

Recent advances in deep learning-based NLP shows significant promise for automatically analyzing team communication data. Deep learning-based NLP techniques learn multiple levels of higher-level features from lower-level data through deep neural networks. A key advantage of deep learning is its feature extraction capabilities, which reduces the need for feature engineering by human experts that is often expensive in terms of time and effort. Automatically assessing overall team performance involves integrating evaluations of each team member's performance into a holistic representation of the team. Traditionally, this has involved a simple average of each team member's performance [23], however, this approach has been criticized for assuming an individual's optimal performance is the same as the team's optimal performance [24]. Deep learning can be used to more flexibly model the inter-relations of different team members for assessment of overall team performance.

3 Team Communication Analysis Pipeline

In our current work, we are developing a generalized team communication analysis pipeline using a deep learning-based NLP framework that can support the analysis of team communication data and predict team performance. The pipeline takes raw speech communication input from team members, analyzes and converts this data into sets of language features, and generates predictions of team performance and team process states. The NLP pipeline contains several key components that perform the automated speech recognition and dialogue analysis required for real-time analysis of team communication and prediction of team performance. We describe each of these components in more depth below (Figure 1).

The first key component in the NLP pipeline is automatic speech recognition (ASR), which converts team spoken communication into text for the pipeline's linguistic processing. One of the primary challenges of ASR is stationary and non-stationary environmental noise, which can corrupt the speech signal and negatively impact the transcribed text. In many complex task environments such as air traffic control, military operations, and first responder situations, communication between team members is often constrained by task conditions and personnel are trained to use routinized verbal interactions and a common vernacular to make communication more effective and efficiency. Since verbal interactions are somewhat routinized, the difficulties associated with generalized ASR are diminished [25].

Prior work on ASR is grounded in hidden Markov models and Gaussian mixture models designed to capture the temporal dynamics of speech and predict textual representations by determining fitness between the hidden states and the acoustic inputs [26]. More recently, researchers have investigated deep neural networks, the underlying machine learning technique of deep learning, to support ASR [27]. Common tools that support deep neural network-based ASR include Google's Cloud Speech-to-Text, IBM Watson, and Microsoft Bing Speech as well as Kaldi, an open-source speech recognition toolkit [28]. These tools provide real-time speech recognition capabilities that will automatically translate spoken language into text that can be further analyzed for syntactic and semantic features. Recent analysis shows that these ASR engines offer considerable accuracy [25, 29] and can be fairly robust to environmental noise [30].

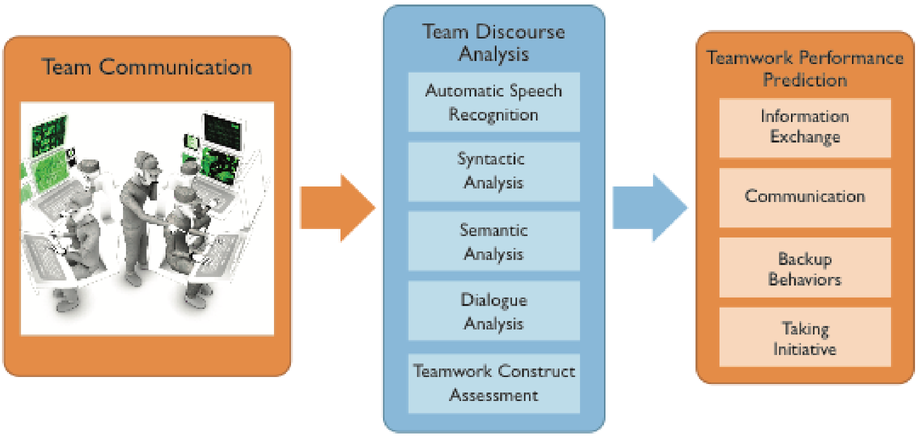

Fig. 1. NLP pipeline for team dialogue modeling.

Next, the textual translation of team members' spoken words are passed through a series of syntactic analyses, such as utterance segmentation (breaking into sentences), part-of-speech tagging (identifying the part-of-speech of each word such as noun, verb, and adjective), text lemmatization (finding the unconjugated form of each word), and dependency parsing (identifying the parent-child relationships between words by building a parse tree of an utterance). This line of syntactic analysis generates a multifaceted, structured representation that can be used to understand the natural language communicated through team conversations.

Following syntax analysis, the NLP pipeline performs a series of semantic analyses to determine the meaning of the spoken language utterances. Semantic analyses will include word sense disambiguation (identifying meanings of words), named-entity recognition (identifying phrases representing concepts such as places, names, and organizations), co-reference resolution (linking each pronoun with its associated word referring to a sequence of sentences), semantic role labeling (identifying the abstract role that arguments of a predicate can take in an event, such as agent, theme, and location), and sentiment analysis (recognizing the speaker's affective state). The semantic analyses are then followed by dialogue act classification, which will be used to recognize and classify dialogue acts or common themes inherent in the spoken team communication data.

3.1 Deep Learning Framework

The NLP pipeline described above will be supported with a deep-learning NLP framework that performs a series of team discourse analysis tasks. To analyze natural language team dialogue, we will use long short-term memory networks (LSTMs), a variant of recurrent neural networks [31], to guide the semantic role labeling, sentiment analysis, dialogue act classification, and individual performance prediction based on team members' utterances. Recurrent neural networks are specifically designed for modeling time-series data and are well suited for analyzing and learning patterns within communication data [32].

As the initial effort for team communication dialogue analysis, we plan to induce a three-task LSTM-based dialogue model that predicts the sentiment and the dialogue act (e.g., response, agreement) for each utterance as well as team-level communication performance. The model will take a series of words that appear in an utterance to predict the sentiment and the dialogue act(s) of the utterance. The sentiment will be predicted using both distributed representations of words [e.g., 33] and acoustic features (e.g., prosodic features including pitch contour and loudness extracted from the speech data) [e.g., 34]. For dialogue act classification, we will adopt a targeted subset of the 42 dialogue acts presented in Stolcke et al. including statement, opinion, question, answer, and summarize, that occur in the dialogue found in our team communication datasets [35]. To create the most effective dialogue analysis system, we will identify the dialogue acts that play central roles in conversation during team-based missions. For instance, Bowers et al. identified eight communication categories that were prominent in analyzing aviation cockpit team communication [11].

To predict team-level communication performance, all of the individual LSTM models will be aggregated into one single architecture. The time-series predictions for the individual members' sentiment, dialogue act, and performance will be used as input to make sequential predictions of the team-level performance in a hierarchical architecture. The output layer of the team-level performance classification model will predict team-performance labels, which will be presented to researchers as a summative evaluation of the team performance informed by individual team members' models. Both the individual and team-level predictive models are end-to-end trainable with a labeled speech dataset.

3.2 Target Data Set

Working with our partners at the U.S. Army Combat Capabilities Development Command - Simulation and Training Technology Center, we will investigate the predictive accuracy of the pipeline using team communication and performance data from the Squad Overmatch project [36]. The Squad Overmatch Project began in 2013 with the goal of improving decision-making under stress by integrating realistic combat exercises through a scenario-based training approach. Seventy-one total squad members participated in the final evaluation event which included six squads completing virtual and live training events. Teamwork behaviors were assessed according to information exchange, communication, supporting behaviors, and taking initiative. Team communication was assessed according to information completeness, phraseology, and closed-feedback loop practices. The team communication dataset includes audio and transcribed recordings of team communication from the scenario-based events and expert ratings of teamwork and team performance. This rich dataset will allow us to analyze team communication dynamics (e.g., dialogue acts and information exchange sequences among team members) and predict team performance at both the individual and team level.

One of the goals of the team communication assessment framework is to build generalizable team communication discourse tagging models which can assess teamwork skills, such as communication, cooperation, and coordination, across new teams and tasks. Furthermore, by using measures of team performance, such as ratings provided by experts and objective team performance scores derived from training or mission rehearsal events, the framework aims to learn what features of language are associated with different kinds of team performance.

4 Conclusion

Teamwork is a complex, dynamic, and multidimensional phenomenon. One of the most challenging aspects of conducting team-based research is developing valid, reliable, and practical computational models of teamwork skills. Such models must capture the dynamic and interdependent sequencing and timing of team members' actions in order to assess underlying team processes. Recent advances in NLP have created the opportunity to build team communication dialogue models that can perform real-time end-to-end natural language analysis on team members' spoken dialogue and generate team dialogue analytics that drive adaptive scaffolding. A significant goal of this effort is to support the analysis of team states in synthetic-based collective training events in order to develop more effective adaptive instructional systems for collective training. Automating team communication assessment offers immense opportunity for identifying bottlenecks and breakdowns in team communication and offering instructive coaching and feedback at the individual and team level.

5 References

1. Smith-Jentsch, K. A., Johnston, J. H., Payne, S. C.. Measuring team-related expertise in complex environments. Making decisions under stress: Implications for individual and team training, **1**, pp. 61-87 (1998).
2. Rousseau, V., Aubé, C., Savoie, A.Teamwork behaviors: A review and an integration of frameworks. Small Group Research, **37**(5), pp. 540-570 (2006).
3. Gorman, J. C., Foltz, P. W., Kiekel, P. A., Martin, M. J., Cooke, N. J. Evaluation of Latent Semantic Analysis-based measures of team communications content. In Proceedings of the Human Factors and Ergonomics Society annual meeting,Vol. 47, No. 3, pp. 424-428. Sage CA: Los Angeles, CA: SAGE Publications (2003).

4. Sottilare, R. A., Burke, C. S., Salas, E., Sinatra, A. M., Johnston, J. H., Gilbert, S. B. Towards a design process for adaptive instruction of teams: A meta-analysis. International Journal of Artificial Intelligence in Education, **28**, pp. 225-264. (2018).
5. Marks, M. A., Mathieu, J. E., Zaccaro, S. J.: A temporally based framework and taxonomy of team processes. Academy of Management Review, **26**(3), pp. 356-376 (2001).
6. Stout, R. J., Cannon-Bowers, J. A., Salas, E. The role of shared mental models in developing team situational awareness: Implications for training. In E. Salas (ed) Situational Awareness, pp. 287-318. Routledge (2017).
7. Smith-Jentsch, K. A., Zeisig, R. L., Acton, B., McPherson, J. A. Team dimensional training: A strategy for guided team self-correction. In J. A. Cannon-Bowers & E. Salas (eds.) Making decisions under stress: Implications for individual and team training, pp. 271-297. Washington, DC, US: American Psychological Association. (1998).
8. Wilson, K. A., Salas, E., Priest, H. A., Andrews, D. Errors in the heat of battle: Taking a closer look at shared cognition breakdowns through teamwork. Human Factors, 49(2), pp. 243-256. (2007).
9. Prince, C., Salas, E.: Training and research for teamwork in the military aircrew. In E. L. Wiener, B. G. Kanki, & R. L. Helmreich (Eds.), Cockpit Resource Management, pp. 337-366. San Diego, CA, US: Academic Press. (1993).
10. Klein, G., Feltovich, P. J., Bradshaw, J. M., Woods, D. D. Common ground and coordination in joint activity. Organizational Simulation, **53**, pp. 139-184. (2005).
11. Bowers, C. A., Jentsch, F., Salas, E., Braun, C. Analyzing communication sequences for team training needs assessment. Human Factors, **40**(4), pp. 672-679. (1998).
12. Marlow, S., Lacerenza, C., Paoletti, J., Burke, C.., Salas, E. Does team communication represent a one-size-fits-all approach? A meta-analysis of team communication and performance. Organizational Behavior and Human Decision Processes, **144**, pp. 145-170. (2018).
13. Dwyer, D. J., Fowlkes, J. E., Oser, R. L., Salas, E., Lane, N. E. Team performance measurement in distributed environments: The TARGETS methodology. In Brannick, M. T., Salas, E., & Prince, C. (Eds.). Series in applied psychology. Team performance assessment and measurement: Theory, methods, and applications, pp. 149-166. Mahwah, NJ. (1997).
14. Achille, L. B., Schulze, K. G., Schmidt-Nielsen, A. An analysis of communication and the use of military terms in Navy team training. Military Psychology, **7**(2), pp. 95. (1995).
15. Csanadi, A., Eagan, B., Kollar, I., Shaffer, D. W., Fischer, F. When coding-and-counting is not enough: Using epistemic network analysis (ENA) to analyze verbal data in CSCL research. International Journal of Computer-Supported Collaborative Learning, **13**(4), pp. 419-438. (2018).
16. Foltz, P. W., Martin, M. J., Abdelali, A., Rosenstein, M., Oberbreckling, R. Automated team discourse modeling: Test of performance and generalization. In Proceedings of the Annual Meeting of the Cognitive Science Society, **28**, pp. 28-33. (2006).
17. Foltz, P. W., Martin, M. J. Automated communication analysis of teams. In E. Salas, G. Goodwin, C. Burke. Team Effectiveness in Complex Organizations: Cross-Disciplinary Perspectives and Approaches, pp. 411 – 432. Taylor and Francis Group. (2008).
18. Kiekel, P. A., Gorman, J. C., Cooke, N. J. Measuring speech flow of co-located and distributed command and control teams during a communication channel glitch. In Proceedings of the human factors and ergonomics society annual meeting, **48**(3), pp. 683-687. (2004).
19. Gorman, J. C., Martin, M. J., Dunbar, T. A., Stevens, R. H., Galloway, T. Analysis of semantic content and its relation to team neurophysiology during submarine crew training. *In International Conference on Augmented Cognition*, pp. 143-152. Springer, Berlin, Heidelberg. (2013).

20. Hu, X., Cai, Z., Franceschetti, D., Penumatsa, P., Graesser, A., Louwerse, M., McNamara, D., Tutoring Research Group. LSA: First dimension and dimensional weighting. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 25, no. 25. (2003).
21. Altszyler, E., Sigman, M., Ribeiro, S., Slezak, D. F. Comparative study of LSA vs Word2vec embeddings in small corpora: A case study in dreams database. arXiv preprint arXiv:1610.01520. (2016).
22. Vu, X. S., Vu, T., Tran, S. N., Jiang, L. ETNLP: A Toolkit for Extraction, Evaluation and Visualization of Pre-trained Word Embeddings. arXiv preprint arXiv:1903.04433. (2019).
23. Langan-Fox, J., Code, S., Langfield-Smith, K. Team mental models: Techniques, methods, and analytic approaches. Human Factors, **42**(2), pp. 242-271. (2000).
24. Garbis, C. Artman, H. Team communication and coordination as distributed cognition. In 9th Conference of Cognitive Ergonomics, pp. 151-156. (1998).
25. Foltz, P. Automating the assessment of team collaboration through communication analysis. In R. Sottilare, A. Graesser, X. Hu, and A.M. Sinatra (Eds.). Design Recommendations for Intelligent Tutoring Systems: Volume 6 - Team Tutoring. pp 197-185, Orlando, FL: U.S. Army Research Laboratory. (2018).
26. Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A. R., Jaitly, N., ... Sainath, T. Deep neural networks for acoustic modeling in speech recognition. IEEE Signal processing magazine, **29**. (2012).
27. Graves, A., Mohamed, A. R., Hinton, G. Speech recognition with deep recurrent neural networks. In IEEE international conference on acoustics, speech and signal processing (ICASSP), pp. 6645-6649. IEEE. (2013).
28. Cheng, G., Povey, D., Huang, L., Xu, J., Khudanpur, S., Yan, Y. Output-Gate Projected Gated Recurrent Unit for Speech Recognition. In Proceedings of Interspeech 2018, pp. 1793-1797. (2018).
29. Ravanelli, M., Parcollet, T., Bengio, Y. The PyTorch-Kaldi Speech Recognition Toolkit. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 6465-6469. IEEE. (2019).
30. Santos, R. M., Matos, L. N., Macedo, H. T., Montalvão, J. Speech Recognition in Noisy Environments with Convolutional Neural Networks. In 2015 Brazilian Conference on Intelligent Systems (BRACIS), pp. 175-179. IEEE. (2015).
31. Hochreiter, S., Schmidhuber, J. Long short-term memory. Neural Computation, **9**(8),pp. 1735-1780. (1997).
32. Serban, I. Sordoni, A., Lowe, R., Charlin, L, Pineau, J., Courville, A., Bengio, Y., A hierarchical latent variable encoder-decoder model for generating dialogues. In Thirty-First AAAI Conference on Artificial Intelligence. (2017).
33. Pennington, J., Socher, R., Manning, C. Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp. 1532-1543. (2014).
34. Eyben, F., Wöllmer, M., Schuller, B.: OpenSMILE: the munich versatile and fast open-source audio feature extractor. In Proceedings of the 18th ACM international conference on Multimedia, pp. 1459-1462. ACM. (2010).
35. Stolcke, A., Ries, K., Coccaro, N., Shriberg, E., Bates, R., Jurafsky, D., Taylor, P., Martin, R., Van Ess-Dykema, C., Meteer, M.. (Dialogue act modeling for automatic tagging and recognition of conversational speech. Computational linguistics, **26**(3), pp. 339-373. (2000).
36. Johnston, J. H. Team performance and assessment in GIFT: Research recommendations based on lessons learned from the Squad Overmatch Research Program. In Proceedings of the Sixth Annual GIFT Users Symposium (vol. 6, pp. 175-187). US Army Research Laboratory, Orlando: FL. (2018).