International Conference on Artificial Intelligence and Networking (ICAIN-2024)

24th & 25th September 2024

Using Multimodal AI for Cognitive Behavior Analysis: Case Study of Deception Detection

Authors

Gargi Joshi[1], Jugal Shroff[1], Isha Pendharkar[1], Aditya Desai[1], Priyanka Jain[3], N.K.Jain[3], Rahee Walambe[1, 2,*], Ketan Kotecha[1, 2*]

Symbiosis Institute of Technology, Symbiosis International Deemed University, Pune, India[1]

Symbiosis Centre for Applied Artificial Intelligence, Symbiosis International Deemed University, Pune, India[2]

Centre for Development of Advanced Computing (C-DAC), Delhi, India[3]

1. INDEX:
2. ABSTRACT
3. INTRODUCTION
4. LITERATURE REVIEW
5. RESEARCH GAPS
6. PROPOSED METHODOLOGY
7. RESULTS AND DISCUSSIONS
8. COMPARATIVE ANALYSIS
9. CONCLUSION AND FUTURE WORK
10. REFERENCES
11. FUNDING ACKNOWLEDGEMENT

2. ABSTRACT:

Deception detection has gained widespread importance due to its applicability in critical domains like national security, judiciary, interrogation, and courtroom trials. Distinguishing between deceit and honest/ truthful behavior is a vital decision. Deception is a crucial act of humans with many complex, diversified physiological and cognitive aspects. The existing methods of deceit detection with a single modality approach need to provide a comprehensive analysis of the task. Detecting deception by integrating and combining diverse heterogeneous modalities builds a comprehensive and complete picture of the underlying task with a holistic understanding of cognitive and emotional states. In the experimental study carried out on the "Bag of Lies "multimodal deception detection dataset, the individual models achieved an accuracy rate of 60.9% for audio, 56.73% for EEG, 59.7% for video, and 64.63% for the gaze modality, respectively. In contrast, the multimodal fusion or integration of all four modalities improved task performance and accuracy by up to 80.39% compared to their unimodal counterparts. Surpassing the existing reported results on the dataset. The insights gained from the study will aid in building improved, advanced deception detection algorithms and systems in the future.

3. INTRODUCTION:

The results observed are significantly better than the previously reported results on the corpus for the individual, as well as a fusion of modalities. Accuracy obtained using Gaussian Naive Bayes and Bernoulli Naive Bayes has been a new approach towards the EEG modality.

Multimodal DNN-based fusion techniques improve existing state-of-the-art models and deepen the understanding of more complex concepts through improved feature extraction and classification techniques. Our fusion of multimodalities with video has enhanced accuracy using a Deep Neural Network with varying frames per video compared to the previously reported results on the corpus.

The methodology adopted for fusing all four modalities with a deep learning-based multimodal score level late fusion approach provides improved performance for the deception detection task with an improved accuracy of 80.39%, surpassing the previously reported results on the bag of lies dataset.

4. LITERATURE REVIEW :

Deception Study	Author	Details
Deception detection through facial and linguistic behavior in conversational dialogue	(Soldner et al., 2019) (Krishnamurthy et al., 2023)	facial and linguistic cues in conversational dialogue are explored
Multimodal deception is carried out in videos with rich information in acoustic, temporal language	(Rill-Garcia et al., 2019) (Karimi, 2018) (Wu et al., 2018) (Dinges et al., 2023)	Acoustic verbal visual and vocal features of deception are explored
A study of deception detection with machine learning techniques	(Constancio et al., 2023) (and Belavadi et al., 2020)	verbal and non-verbal cues from human subject stimuli such as facial expressions, audio, and video text
Deception detection through facial expressions and pulse rate during job interviews	(Tsuchiya et al., 2023) (Mathur & Matarić, 2020)	Facial expressions and landmarks with pulse rate are explored
speech utterances and gaze	(X. Guo et al., 2023) (Fathima Bareeda et al., 2021) (Gallardo-Antolín & Montero, 2021)	Verbal and vocal features along with visual features such as gaze are explored
verbal and non-verbal cues such as gait and gestures with a combination of contact and non-contact modalities	(Randhavane et al., 2019)	Combination of verbal and non verbal features with contact and non contact modalities are used
convolutional self-attention for attention-based representation learning for deceit detection	(King, 2019) (Bic, 2023)	Attention based deception detection is carried
Analysis evaluation and future research directions for multimodal deception detection are discussed	(Ulizia et al., 2024).	Survey paper on aanalysis evaluation and future research directions for multimodal deception detection is presented

5. RESEARCH GAPS :

1. Lack in exploration of neuro cognitive modalities for analyzing their potential in Multimodal deception detection.
2. Despite of previous studies significant improvements are expected in feature extraction and feature selection techniques to further improve the task of automated deception detection with advanced deep learning-based architectures and fusion methods.

6. PROPOSED METHODOLOGY :

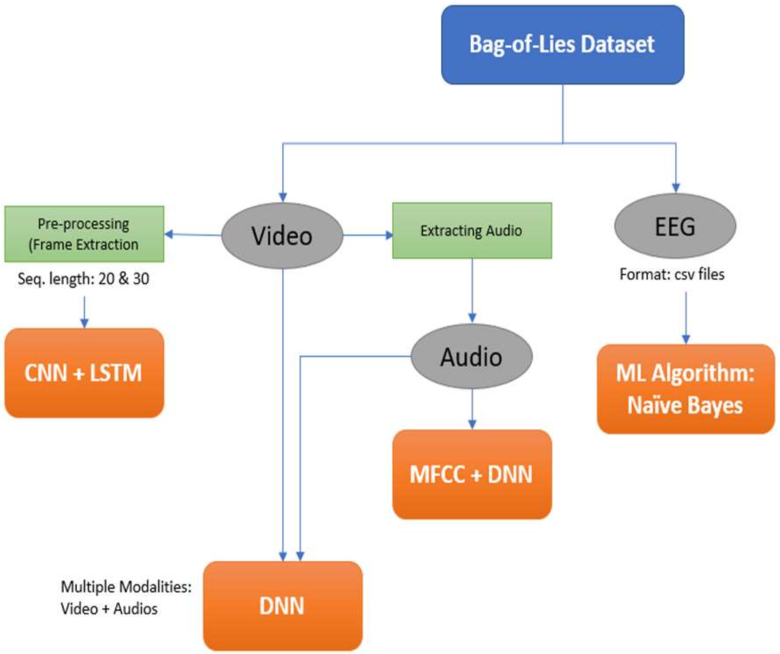


Fig 1. Proposed model architecture for multimodal deception detection

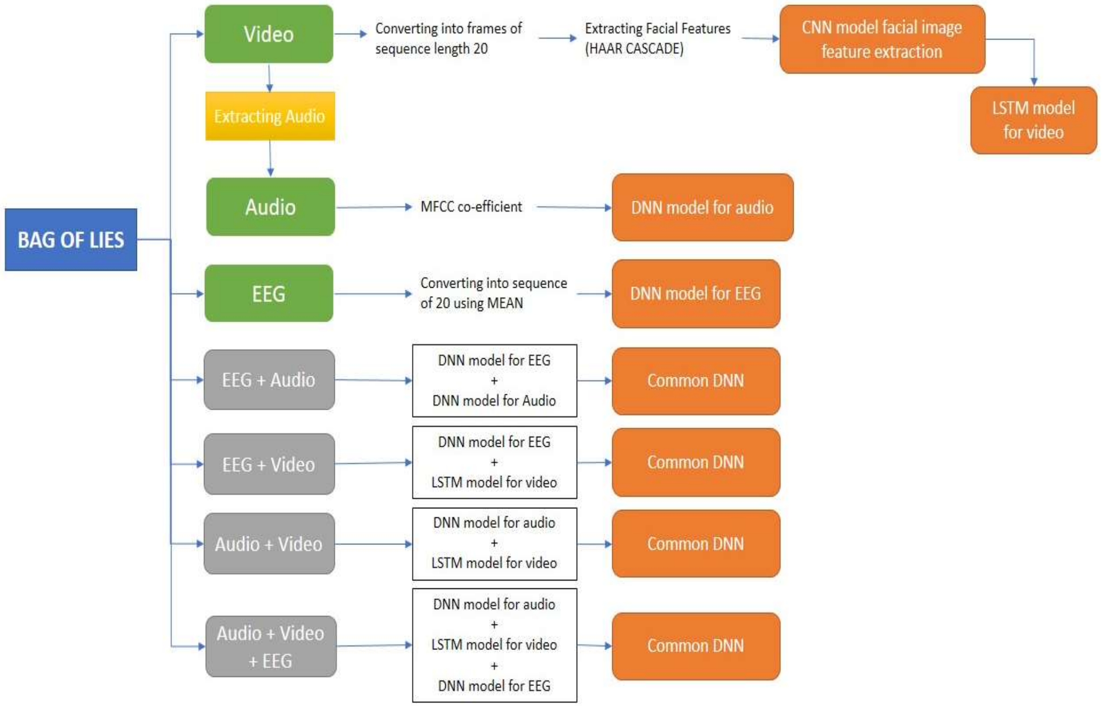


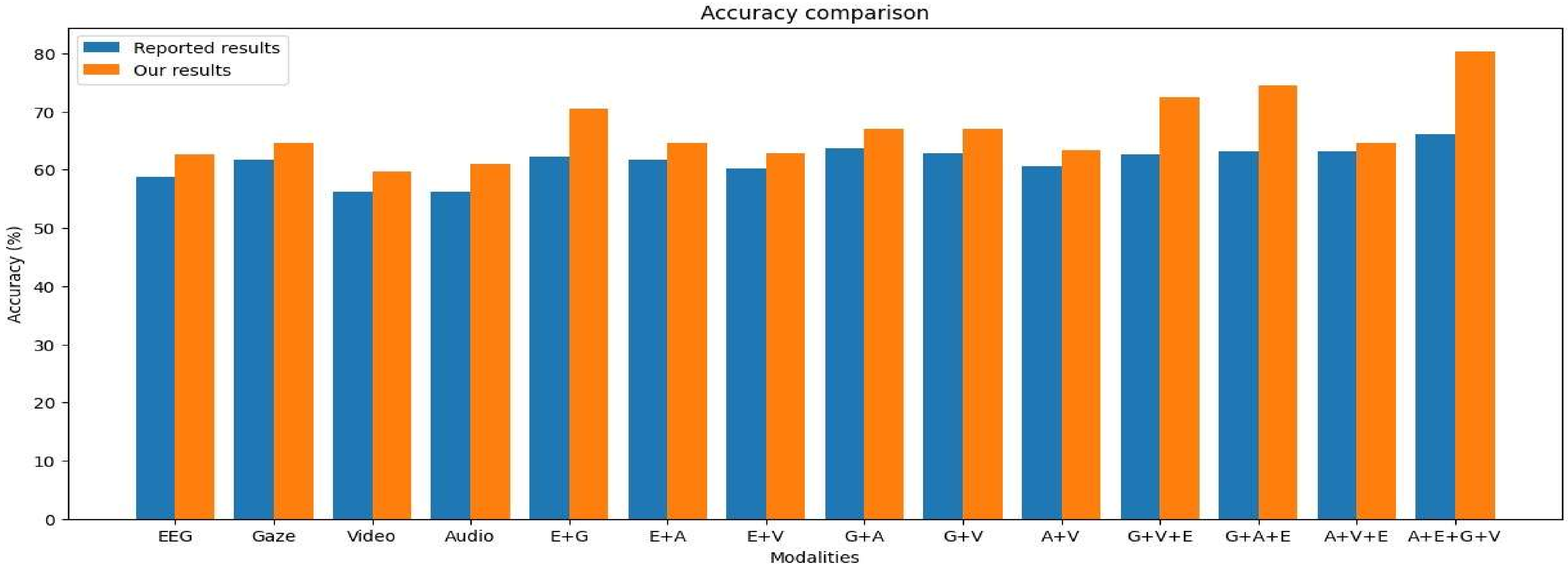
Fig 2. Pipeline for data processing and model development for the Bag of Lies dataset.

Set A: Data which contains all four modalities, basically a subset of set B

Set B: Data that contains at least three modalities (may or may not contain EEG).

Modality	Accuracy
Audio	60.9%
EEG (ML)	GaussianNB: 56.73% BernoulliNB: 53.79% (Set A)
EEG (LSTM)	62.7% (Set A)
Video (Facial + CNN + LSTM)	59.7%
Gaze (LSTM)	64.63%
Audio + EEG	64.7% (Set A)
Audio + Video	63.41%
Audio + Gaze	67.07%
EEG + Video	62.74% (Set A)
EEG + Gaze	70.5% (Set A)
Video + Gaze	67.07%
Audio + EEG + Video	64.70 (Set A)
Audio + Gaze + Video	69.51%
Audio + EEG + Gaze	74.50 (Set A)
Video + EEG + Gaze	72.54% (Set A)
Audio + EEG + Gaze + Video	80.39 (Set A)

8. COMPARATIVE ANALYSIS OF OUR RESULTS WITH BAG OF LIES DATASET:



9. CONCLUSION & FUTURE WORK :

In the privacy and security aspect, deception detection is a crucial task that can be solved by integrating and fusing multiple modalities with advanced model architectures and wiser use of different features, assisting the task with improved performance. In the future, these insights will lead us toward building more advanced deception detection systems and algorithms involving multiple modalities to build a comprehensive and all-inclusive deception detection system. The gaze and EEG modalities are complementary for the task. The overall fusion of all the modalities audio, EEG, gaze, and video achieves the highest accuracy of 80.39%, showing that multimodal fusion provides improved performance and better results than their counterparts. In the future, increasing the number of modalities and sample size with advanced processing architectures will further improve task performance, opening diverse avenues for multimodal deception detection. The outcomes of the study will aid in building robust and accurate deception detection models with behavioral insights. They will lay a foundation for advanced deception detection models ensuring data privacy, bias, and cultural context with cognitive behavioral aspects influenced by multiple verbal, nonverbal, and neurophysiological modalities.

10. REFERENCES :

1. Abdul, Z. K., & Al-Talabani, A. K. (2022). Mel Frequency Cepstral Coefficient and its Applications: A Review. *IEEE Access*, *10*(November), 122136–122158. https://doi.org/10.1109/ACCESS.2022.3223444

2. Baltrusaitis, T., Ahuja, C., & Morency, L. P. (2019). Multimodal Machine Learning: A Survey and Taxonomy. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/TPAMI.2018.2798607

3. Belavadi, V., Zhou, Y., Bakdash, J. Z., Kantarcioglu, M., Krawczyk, D. C., Nguyen, L., Rakic, J., & Thuriasingham, B. (2020). MultiModal Deception Detection: Accuracy, Applicability and Generalizability*. *Proceedings - 2020 2nd IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications, TPS-ISA 2020*, pp. 99–106. https://doi.org/10.1109/TPS-ISA50397.2020.00023

4. Bhatt, P., Sethi, A., Tasgaonkar, V., Shroff, J., Pendharkar, I., Desai, A., Sinha, P., Deshpande, A., Joshi, G., Rahate, A., Jain, P., Walambe, R., Kotecha, K., & Jain, N. K. (2023). Machine learning for cognitive behavioral analysis: datasets, methods, paradigms, and research directions. *Brain Informatics*, *10*(1). https://doi.org/10.1186/s40708-023-00196-6

5. Bic, B. (2023). *Automatic Deceit Detection Through Multimodal Analysis of High-Stake Court-Trials*. *XX*(X), pp. 1–16. https://doi.org/10.1109/TAFFC.2023.3322331

6. Camara, M. K., Postal, A., Maul, T. H., & Paetzold, G. (2022). *Can lies be faked? Comparing low-stakes and high-stakes deception video datasets from a Machine Learning perspective*. http://arxiv.org/abs/2211.13035

7. Constancio, A. S., Tsunoda, D. F., de Fátima Nunes Silva, H., da Silveira, J. M., & Carvalho, D. R. (2023). Deception detection with machine learning: A systematic review and statistical analysis. In *PLoS ONE* (Vol. 18, Issue 2 February). https://doi.org/10.1371/journal.pone.0281323

8. Dinges, L., Fiedler, M.-A., Al-Hamadi, A., Hempel, T., Abdelrahman, A., Weimann, J., & Bershadskyy, D. (2023). *Automated Deception Detection from Videos: Using End-to-End Learning Based High-Level Features and Classification Approaches*. http://arxiv.org/abs/2307.06625

9. Fathima Bareeda, E. P., Shajee Mohan, B. S., & Ahammed Muneer, K. V. (2021). Lie Detection using Speech Processing Techniques. *Journal of Physics: Conference Series*, *1921*(1). https://doi.org/10.1088/1742-6596/1921/1/012028

10. Gallardo-Antolín, A., & Montero, J. M. (2021). Detecting deception from gaze and speech using a multimodal attention LSTM-based framework. *Applied Sciences (Switzerland)*, *11*(14). https://doi.org/10.3390/app11146393