Bag-of-Lies: A Multimodal Dataset for Deception Detection

Viresh Gupta*
IIIT-Delhi, India
viresh16118@iiitd.ac.in

Mohit Agarwal*
IIIT-Delhi, India
mohit15060@iiitd.ac.in

Manik Arora
IIIT-Delhi, India
manik15053@iiitd.ac.in

Tanmoy Chakraborty
IIIT-Delhi, India
tanmoy@iiitd.ac.in

Richa Singh
IIIT-Delhi, India
rsingh@iiitd.ac.in

Mayank Vatsa
IIIT-Delhi, India
mayank@iiitd.ac.in

Abstract

Deception detection is a pervasive issue in security. It has been widely studied using traditional modalities, such as video, audio and transcripts; however, there has been a lack of investigation in using modalities such as EEG and Gaze data due to the scarcity of a publicly available dataset. In this paper, a new multimodal dataset is presented, which provides data for deception detection by the aid of various modalities, such as video, audio, EEG and gaze data. The dataset explores the cognitive aspect of deception and combines it with vision. The presented dataset is collected in a realistic scenario and has 35 unique subjects providing 325 annotated data points with an even distribution of truth (163) and lie (162). The benefits provided by incorporating multiple modalities for fusion on the proposed dataset is also investigated. It is our assertion that the availability of this dataset will facilitate the development of better deception detection algorithms which are more relevant to real world scenarios.

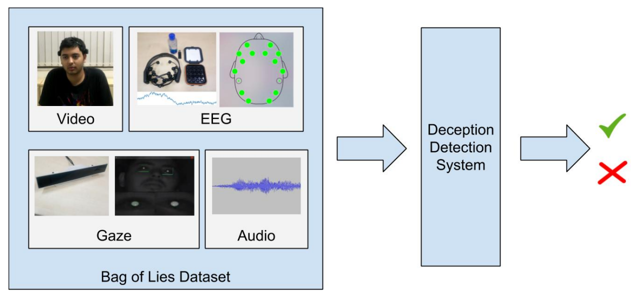

Figure 1: Proposed Bag-of-Lies Dataset, contains Video, Audio, Gaze and EEG data for subjects that are describing a stimuli image

1. Introduction

An act of deception is performed when a person tries to convince others about a false fact by providing inaccurate evidence. This can be done using either lies, misrepresentation of facts or omissions [18]. Deception is widespread in the society. It is prevalent in many forms, broadly seen as high stake environments and casual deception. The challenge is to develop methods that detect deceitful behaviour.

High stake deception occurs when the speakers are invested into making the statements, and their statements will have a significant impact directly – e.g., courtroom trials, where a deceiving statement can lead to a guilty defendant being acquitted without any charges. Casual deception can be seen in online reviews [23] to influence buyer's decisions, social media posts [31] etc. to amass people for/against a cause and thus indirectly affects the society as a whole.

Deception detection is highly useful in criminal investigations, where very often the criminal is trying to deceive the law enforcers to avoid facing punishments. Also, an easy means of deception detection can prevent the spread of rumours via social media or other networking sites. The ability of humans to detect deception without any special aids is limited – around 54% as reported in [5]. Thus there is a need for systems that aid in deception detection. In the past, several models have been proposed for deception detection. Examples include physiological methods such as the famous Polygraph test [30] or the more recent functional magnetic resonance imaging (fMRI) based tests [7]. These methods suffer from following two limitations – (i) they require sophisticated equipment setup, (ii) they are overt in nature and require a trained operator to use these methods. Therefore, their applicability to real life is quite low. The widely used polygraph test has not been scientifically proven to detect deception at individual level [3]. Also, these methods especially cannot be applied on cases

*Equal contribution by student authors

Table 1: Existing datasets for deception detection.

Data Set	Subjects	Modalities (number)	Total	Collection Strategy
CSC [13]	32	Only Audio (1)	-	Hypothetical Scenario
ReLiDDB [21]	40	Only Audio (1)	-	Hypothetical Scenario
Open Domain [27]	512	Only Text (1)	7168	Crowdsourcing
EEG-P300 [32]	11	Only EEG (1)	88	Hypothetical Scenario
Real Life Trials [25]	56	Video, Audio, Text (3)	121	Realistic Scenario
Multi Modal [28]	30	Video, Audio, Thermal, Physiological (4)	150	Hypothetical Scenario
Bag-of-Lies (proposed)	**35**	**Video, Audio, EEG, Gaze (4)**	**325**	**Realistic Scenario**

such as a deceitful YouTube video.

Researchers have proposed behavioural methods [17], such as using involuntary facial expressions [36]. These micro-expressions are hard to detect for the untrained eye and thus not useful for the masses without any special training. Thus there is a need for automated deception detection method, and the surge in computational power has paved way for construction of automated systems using data-driven approaches. Data-driven approaches also have an added advantage of being unobtrusive and can leverage as much information as is available. For example, they can work with just video, and also incorporate text annotations when available (such as the subtitles in a Youtube video), making them an ideal candidate for deception detection systems.

Other data-driven methods proposed for deception detection use modalities such as Video, Audio, Text, Electroencephalogram (EEG), Gaze [13, 32, 24, 21, 20, 27, 26], but none of them consider all modalities together due to the lack of a comprehensive dataset. An attempt was made in [28] to construct a multimodal dataset for casual deception using physiological and thermal measurements as well. However, in the experiment the participants were instructed about lying/saying a truth depending upon the scenario they were in rather than a scenario where participants would do so out of their will. For high-stakes deception, [34] took courtroom trial dataset from [25] and used vision techniques for deception detection.

1.1. Existing Deception Detection Datasets

Existing work on deception detection has widely used the datasets as described in Table 1. Many of the datasets focus only on one modality, and have unnatural stimuli, making the deception a forced one, and often lack in volume. For example, [21, 13, 32, 27] focused only on one modality. [28] aimed at constructing a multimodal dataset in a casual setting. However, the dataset was collected by eliciting deceptive and truthful statements in pre-defined scenarios on the basis of fixed roles assigned to the participants (a participant does not deceive of their own will). [25] built a multimodal dataset in high stakes scenario. However, it may not be helpful for detection of casual deception. However, to the best of our knowledge, there is no dataset for casual deception detection using multiple modalities and objective judgment which motivated us to develop a dataset of such kind.

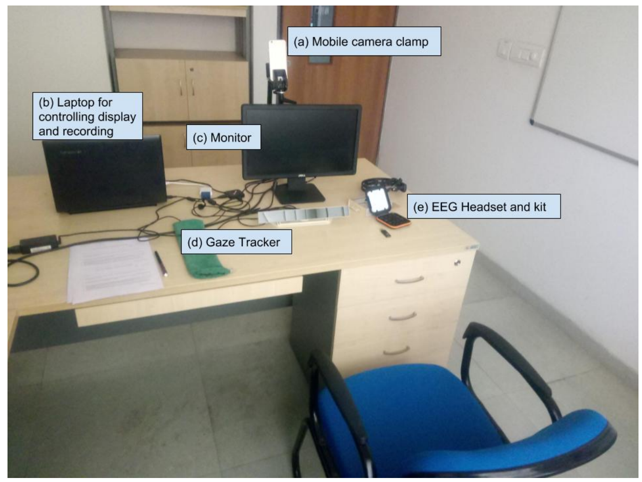

Figure 2: The apparatus used for data collection.

1.2. Contributions

This paper presents a benchmark dataset, termed as **Bag-of-Lies**. The proposed dataset consists of multiple modalities such as video, audio, EEG and Eye Gaze from 35 unique subjects collected using a carefully designed experiment. It has a total of 325 annotated recordings consisting of 162 lies and 163 truths. First, a novel methodology for collection of casual deception data is proposed, which even though is objective (free of hypothetical scenarios), allows the participants to lie freely and naturally. Next, data is recorded with the designed experiment using university students who volunteered for the study. Lastly, a comprehensive analysis over the dataset is performed to demonstrate its utility in the task of deception detection. *To the best of our knowledge, this is the first multi-modal dataset that provides an objective deception scenario for casual deceptions.*

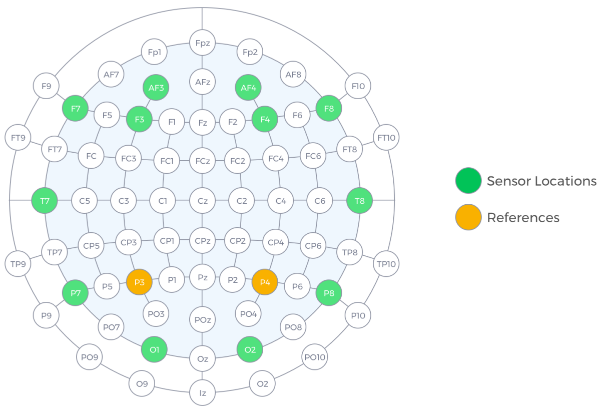

Figure 3: Electrode sensor reference used for Emotiv EPOC+.

Figure 4: Sample Images used in experiment. The images are descriptive and hence ideal for the experiment. Images taken are Creative Commons licensed

2. Proposed Bag-of-Lies Dataset

Most of the deception datasets so far are based on subjective interviews where the participants were either told beforehand if they have to lie or not, or the participants were given a hypothetical scenario where they had to express truthful and deceptive opinion. However, there exists no dataset for casual deception detection that captures multiple modalities and provides a real objective goal simultaneously. Existing studies suggest that EEG is a potentially unexplored area for deception detection which showed promising results in a pilot study [32]. Also there has been past work which suggests that eye blink [11], pupil dilation [10] and eye movements [24] are related to deception detection. Inspired from the past work, this work proposes *Bag-of-Lies*, a novel dataset which is objective in nature, and yet allows natural casual deception scenarios, i.e., the participants were free to choose whether they wanted to be honest or deceive. It combines multiple modalities such as video, audio, EEG and Eye Gaze as shown in Figure 1. Table 2 presents dataset statistics. The database will be made publicly available to the research community[1].

Table 2: Statistics of the Bag-of-Lies Dataset.

Modality	**Video, Audio, and Gaze**	**EEG**
Records	325	201
Subjects	35	22
Min, Max	3, 11	6, 10
Truth - Lie	163 - 162	108 - 93

2.1. Materials

A normal camera and microphone in a smart-phone was used for recording the video and audio of participants. This ensures that the dataset matches realistic scenarios where high definition recordings might not be available, such as a YouTube video blog / CCTV footage, etc. A 14-Channel Emotiv EPOC+EEG headset (wirelessly connected) was used for the collection of EEG data and Gazepoint GP3 Eye Tracker for collecting gaze data. In addition, 21 distinct, content-heavy and descriptive images were collected for use in the experiment. The smart-phone was mounted on a table top clamp camera stand and the volunteers were seated in front of a monitor which was used to calibrate the eye tracker and display visual stimulus. The experimental setup is shown in Figure 2.

The electrode placement system followed for calibration of EEG device is shown in Figure 3 [1]. The EEG data collected has 13 channels (channel AF3 being unavailable due to driver issues, available channels - F3, FC5, F7, T7, P7, O1, O2, P8, T8, F8, AF4, FC6, F4), sampled at 2048 Hz internally and output filtered to 128 Hz. Since it has been found that the frontal lobe is responsible for cognitive activities such as lying [35, 21], the channels collected with the headset are expected to help with the task of deception detection.

The Gaze data is compliant with the GazePoint Open Gaze API [2] and provides with the following 26 columns - CNT, TIME, TIMETICK, FPOGX, FPOGY, FPOGS, FPOGD, FPOGID, FPOGV, BPOGX, BPOGY, BPOGV, CX, CY, CS, USER, LPCX, LPCY, LPD, LPS, LPV, RPCX, RPCY, RPD, RPS, and RPV. These channels encode information such as the best position of gaze, left pupil size, and right pupil size, which can be further used to calculate features such as fixations, pupil size variations and saccades.

[1]Bag-of-Lies will be made available at http://iab-rubric.org/resources.html

2.2. Participants

The subjects of this study were 35 university students, all comfortable with English language and from different backgrounds. There were 10 female and 25 male participants, each of which were shown 6-10 images (varying depending on the subjects) from the selected set of images. Figure 4 shows sample images shown to the participants. Their responses were recorded using the above setup, thus providing us with 325 recordings. The recordings are variable in length, ranging from 3.5 seconds to 42 seconds. For some subjects, the thick hair were an obstruction for recording EEG using our device, and thus for those volunteers only video, audio and Gaze was recorded. Participants were requested to avoid significant head movements during the recording phase owing to the sensitive calibration of EEG and Gaze instruments.

2.3. Procedure

The experiment is designed as follows:

- A participant is seated in front of a monitor with an Eye Gaze sensor mounted at the bottom. They are fitted with the EEG headset. Next, they are instructed to follow the eye gaze calibration procedure.
- Each participant is shown 6-10 images from the selected image set, one at a time.
- The participant is then asked to describe the image that they are viewing on the screen.
- The participant is free to describe the image honestly or deceptively.

The participant is asked to provide a description of the image. They are free to choose between giving an honest description or lie about the image shown. Thus the experiment is objective, i.e the description does not involve imagining a hypothetical scenario. Also the experiment does not force participants into providing a deceptive description. Therefore, the choice of deceiving on a picture is with the participant and completely natural. The truth for the descriptions given by participant is annotated manually by the authors according to whether the participant's description actually matches the image shown or not.

3. Experiments

For establishing some initial analysis over the curated dataset, experiments using different modalities are performed and the effect of combining several modalities using late fusion is observed. More formally, the problem is - given a set of modalities $(m_{1i}, m_{2i}, ...m_{ni})$ for i^{th} sample, classify the sample i amongst the two classes – *truth* or *lie*.

3.1. Protocol

The dataset is divided into two sets – A and B. Set A consists of all the users for whom all four - EEG, Gaze, audio and video are available (22 unique users). Set B is a superset of Set A, and all three - Gaze, audio and video are available in the dataset (35 unique users).

For the purpose of evaluation, Set A is further divided into two-folds (11 users in each fold), and Set B into three-folds (12, 12 and 11 users, respectively). The results reported are average of cross validation over folds in respective sets. The papers [27, 25, 32, 13] with a similar problem statement have also used accuracy (or error) as a suitable metric for this task.

3.2. Method

Different standard data-driven techniques are applied on the proposed Bag-of-Lies dataset for analyzing the performance of different modalities.

3.2.1 Video

For the classification using video, 20 frames are picked from each of the recorded videos by selecting a single representative frame from $duration/20$ sized video chunks. Local Binary Pattern (LBP [22]) features are extracted for these frames and are concatenated together into a single feature vector. The order of the concatenation is the same as they appear in the video. This combined feature vector is then used further for classification using Support Vector Machine (SVM) [6], Random Forest [14] and Multilayer Perceptron (MLP) [12, 33]).

3.2.2 Audio

Audio from all the videos are extracted and processed further to calculate various frequency-based properties such as zero crossing rate, spectral centroid, spectral bandwidth, spectral rolloff, chroma frequencies and mel frequency cepstral coefficients (MFCC) [19, 9]. These are combined into a 26 dimensional feature vector which is then used for two-class audio classification using Random Forest and K-Nearest Neighbour (KNN) [4] classifiers.

3.2.3 EEG

For experiments involving EEG, features are constructed by two means: (1) the raw data is divided into sub-samples of a fixed window size of 32. The sub-samples are created using sliding window over the time series data with an overlap of 4 between two consecutive windows. These are fed into a CNN [15] architecture known to generalize over several tasks in the presence of limited training data (architecture described by [16] as 1-D convolutions along time

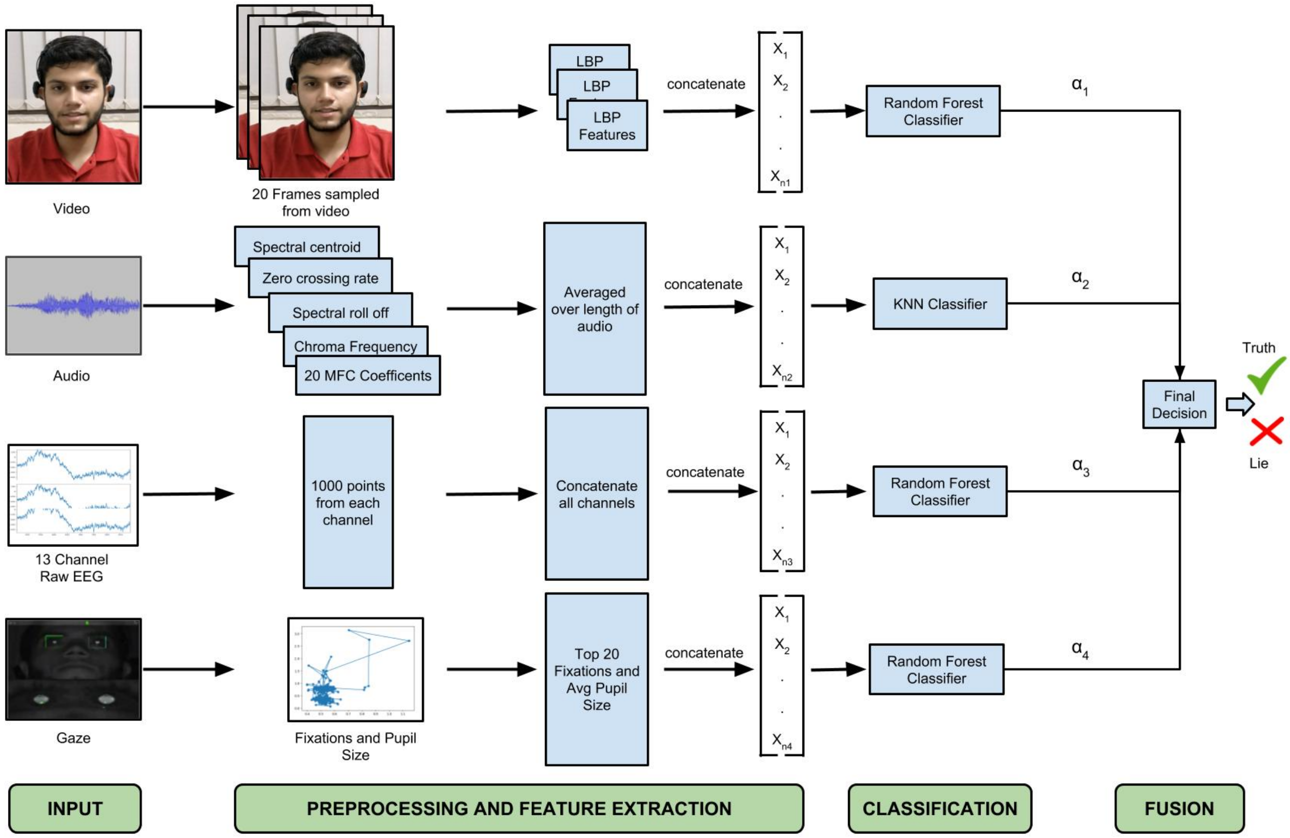

Figure 5: Architecture with the best performing baselines and score level fusion used for analysing the proposed dataset. Fusion using the strategy as described in Section 3.2.5.

axis followed by depth-wise 2-D convolutions followed by a separable 2-D convolution connected to dense layers that give output probabilities using softmax). (2) For the other classifiers, the raw data is processed to keep 1000 time points. Larger data-points are truncated while smaller ones are padded with zeros. This 1000 dimensional feature vector is then used with Random Forest for classification.

3.2.4 Gaze

For experiments with Gaze Data, fixations, eye blinks and pupil size during the recording are calculated as features. Fixations are events when the participant is focusing on one portion of the screen for a long time. These are calculated using a slightly modified version of the PyGaze analysis library [8]. The duration of a fixation, the location x, y of the fixation are used as features. A 64 dimensional feature vector is constructed using top 20 fixations (ranked by their duration) and the number of eye blinks, average pupil size and standard deviation of pupil size along with the total number of fixations (thus amounting to $20\times3+4 = 64$ dimensions).

3.2.5 Combining Modalities

We perform a late fusion [29] of the decisions from best performing classifiers on all modalities in order to determine the final prediction of a sample recording. The effect of combining all four modalities in all possible permutations is measured and summarised in Table 3. For combining the decisions, score level late fusion is used:

$$C_i = \arg\max_k \left(\sum_{j=1}^{n} \alpha_j P_{ijk}\right)$$

where k represents classes, i represents a sample, n represents modality ($n = 4$ for Bag-of-Lies dataset), and P_{ijk} is the prediction score for i^{th} sample belonging to class k as given by modality j (weighted per modality by α_j). The values of hyper-parameter α_j were chosen by searching between values from 0 to 5 with a step size of 0.2. An overview of this process is shown in Figure 5. Score level late fusion was chosen since it is easier to weigh individual modalities and thus effectively inspect the contribution of each modality towards deception detection. Better models can be developed in future to handle multiple modalities

Table 3: Results for Deception Detection on Bag-of-Lies Dataset. Set A and Set B are as defined in section 3.1.

Modality	**Method**	**Average Accuracy**	
		Set A	**Set B**
Only EEG	Random Forest	58.71	-
	EEG Net	54.25	-
	MLP	53.79	-
Only Gaze	Random Forest	61.70	57.11
	MLP	57.71	53.51
Only Video	LBP + SVM	55.21	53.25
	LBP + Random Forest	56.20	55.26
	LBP + MLP	54.22	49.90
Only Audio	Random Forest	53.24	54.89
	KNN	53.22	56.22
EEG + Gaze	Score level fusion of best performing algorithms on various modalities	62.22	-
EEG + Audio		61.69	-
EEG + Video		60.20	-
Gaze + Audio		63.69	59.42
Gaze + Video		62.19	62.71
Audio + Video		60.68	58.24
Gaze + Video + EEG		62.70	-
Gaze + Audio + EEG		63.21	-
Audio + Video + EEG		63.18	-
Gaze + Video + Audio		64.69	60.09
All four		66.17	-

together.

4. Analysis

Table 3 presents the average classification accuracy (%) over folds for both the sets A (22 users including EEG) and B (35 users without EEG).

- Using only individual modalities, it can be observed that utilizing gaze data gives the best accuracy of 61.70% and 57.11% for Sets A and B respectively amongst all modalities. This result depicts that gaze is an important modality and can provide us with essential insights to deception detection which other existing datasets lack as shown in Table 1.

- On performing score level fusion with two modalities at a time, it is observed that the results improve by a significant margin (e.g., 3% with Audio + EEG as compared to only EEG in Set A and 5.6% with Gaze + Video vs. only Gaze in Set B) as compared to using the individual modalities. This indicates that pairing up modalities is beneficial for the task of deception detection.

- Similarly when using three modalities simultaneously, the results improve further and are better than using a modality individually or pairs of modalities. Furthermore using all four modalities, the results are better than using any subset of the modalities. The relevance of gaze and EEG together in a dataset is further validated during the experiments where giving higher weights to these modalities performs better than giving all modalities the same weight. Thus multiple modalities are highly beneficial for the task and by using advanced architectures and features, better classifiers can be built to aid the deception detection process which was previously not possible due to the paucity and unavailability of such a comprehensive dataset.

5. Conclusion

In this paper, a new multimodal dataset for casual deception detection having gaze, EEG, audio and video information is presented. The dataset presents deception in a natural setting, with minimal external intervention and multiple modalities opening avenues for interesting work in the problem of deception detection which will help lay foundations for creating systems to tackle the problem in real-time. The dataset will be made publicly available to the research community. It is our assertion that this will promote research on this important topic.

A preliminary analysis using existing frequently used feature extractors and classifiers showed that gaze features have the ability to provide deeper understanding of the problem. Gaze data is unique to Bag-of-Lies dataset, and thus the dataset presents researchers with new challenges to incorporate this potential along with the existing techniques. Also the dataset has EEG data, and presents interesting research avenues for exploring the cognitive aspects of deception and this knowledge can be efficiently utilized to train a better and advanced model. Note that deception detection is an interesting challenge both for computer vision and privacy and security communities. The fact that it can immediately be aided by multiple modalities, as shown by a simple late fusion in section 3.2.5, reinforces the usefulness of the proposed dataset. It can be further supported by having a thorough human benchmark on the proposed dataset and comparing the performance of this dataset with other datasets having similar modalities. There is a huge scope of improvement by using better features for all the modalities, having a more complex network and deploying more effective multimodal fusion techniques. Thus, we believe that presence of the proposed Bag-of-Lies dataset will significantly facilitate research on deception detection and thereby assist in building more robust and practical deception detection systems.

References

[1] Electrode positions for Emotive EPOC+ headset. https://emotiv.gitbook.io/epoc-user-manual/epoc+_headset_details. 3

[2] GazePoint's open gaze API. https://www.gazept.com/dl/Gazepoint_API_v2.0.pdf. 3

[3] Deceiving the law. *Nature Neuroscience*, 11:1231 EP –, Nov 2008. Editorial. 1

[4] N. S. Altman. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46(3):175–185, 1992. 4

[5] Jr. Charles F. Bond and Bella M. DePaulo. Accuracy of deception judgments. *Personality and Social Psychology Review*, 10(3):214–234, 2006. PMID: 16859438. 1

[6] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, Sep 1995. 4

[7] Qian Cui, Eric J. Vanman, Dongtao Wei, Wenjing Yang, Lei Jia, and Qinglin Zhang. Detection of deception based on fmri activation patterns underlying the production of a deceptive response and receiving feedback about the success of the deception after a mock murder crime. *Social Cognitive and Affective Neuroscience*, 9(10):1472–1480, Oct 2014. 1

[8] Edwin S. Dalmaijer, Sebastiaan Mathôt, and Stefan Van der Stigchel. Pygaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods*, 46(4):913–921, Dec 2014. 5

[9] S. Davis and P. Mermelstein. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(4):357–366, August 1980. 4

[10] Daphne P. Dionisio, Eric Granholm, William A. Hillix, and W F Perrine. Differentiation of deception using pupillary responses as an index of cognitive processing. *Psychophysiology*, 38 2:205–11, 2001. 3

[11] Kyosuke Fukuda. Eye blinks: new indices for the detection of deception. *International Journal of Psychophysiology*, 40(3):239 – 245, 2001. Psychophysiology in. 3

[12] G. E. Hinton. Connectionist learning procedures. *Artif. Intell.*, 40(1-3):185–234, Sept. 1989. 4

[13] Julia Hirschberg, Stefan Benus, Jason M. Brenier, Frank Enos, Sarah Friedman, Sarah Gilman, Cynthia Girand, Martin Graciarena, Andreas Kathol, Laura Michaelis, Bryan L. Pellom, Elizabeth Shriberg, and Andreas Stolcke. Distinguishing deceptive from non-deceptive speech. In *INTERSPEECH*, 2005. 2, 4

[14] Tin Kam Ho. Random decision forests. In *Proceedings of the Third International Conference on Document Analysis and Recognition (Volume 1) - Volume 1*, ICDAR '95, pages 278–, Washington, DC, USA, 1995. IEEE Computer Society. 4

[15] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012. 4

[16] Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, and Brent J Lance. EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of Neural Engineering*, 15(5):056013, jul 2018. 4

[17] Bella M DePaulo, James J Lindsay, Brian E Malone, Laura Muhlenbruck, Kelly Charlton, and Harris Cooper. Cues to deception. *Psychological Bulletin*, 129:74–118, 02 2003. 2

[18] James Edwin Mahon. A definition of deceiving. *International Journal of Applied Philosophy*, 21(2):181–194, 2007. 1

[19] Paul Mermelstein. Distance Measures for Speech Recognition–Psychological and Instrumental. In *Joint Workshop on Pattern Recognition and Artificial Intelligence*, 1976. 4

[20] Nicholas Michael, Mark Dilsizian, Dimitris Metaxas, and Judee K Burgoon. Motion Profiles for Deception Detection Using Visual Cues. In *Proceedings of the 11th European conference on Computer vision: Part VI*, pages 462–475. 2010. 2

[21] H. Nasri, W. Ouarda, and A. M. Alimi. Relidss: Novel lie detection system from speech signal. In *IEEE/ACS 13th International Conference of Computer Systems and Applications*, pages 1–8, Nov 2016. 2, 3

[22] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, July 2002. 4

[23] Myle Ott, Claire Cardie, and Jeff Hancock. Estimating the prevalence of deception in online review communities. In *Proceedings of the 21st International Conference on World Wide Web*, pages 201–210, 2012. 1

[24] Jinie Pak and Lina Zhou. Eye movements as deception indicators in online video chatting. In *AMCIS*, 01 2011. 2, 3

[25] Verónica Pérez-Rosas, Mohamed Abouelenien, Rada Mihalcea, and Mihai Burzo. Deception Detection using Real-life Trial Data. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, pages 59–66. ACM Press, 2015. 2, 4

[26] Verónica Pérez-Rosas and Rada Mihalcea. Cross-cultural deception detection. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 440–445, 2014. 2

[27] Verónica Pérez-Rosas and Rada Mihalcea. Experiments in Open Domain Deception Detection. (September):1120–1125, 2015. 2, 4

[28] Veronica Perez-Rosas, Rada Mihalcea, Alexis Narvaez, and Mihai Burzo. A multimodal dataset for deception detection. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation*. European Language Resources Association (ELRA), 2014. 2

[29] Arun Ross and Anil Jain. Information fusion in biometrics. *Pattern Recogn. Lett.*, 24(13):2115–2125, Sept. 2003. 5

[30] Leonard Saxe. Science and the cqt polygraph. *Integrative Physiological and Behavioral Science*, 26(3):223–231, Jul 1991. 1

[31] Michail Tsikerdekis and Sherali Zeadally. Online deception in social media. *Commun. ACM*, 57:72–80, 2014. 1

[32] Arjon Turnip, M Faizal Amri, Hanif Fakrurroja, Artha Ivonita Simbolon, M. Agung Suhendra, and Dwi Esti Kusumandari. Deception detection of eeg-p300 component classified by svm method. In *Proceedings of the 6th International Conference on Software and Computer Applications*, pages 299–303, 2017. 2, 3, 4

[33] C. Van Der Malsburg. Frank rosenblatt: Principles of neurodynamics: Perceptrons and the theory of brain mechanisms. In Günther Palm and Ad Aertsen, editors, *Brain Theory*, pages 245–248, Berlin, Heidelberg, 1986. Springer Berlin Heidelberg. 4

[34] Zhe Wu, Bharat Singh, Larry S. Davis, and V. S. Subrahmanian. Deception Detection in Videos. *AAAI 18*, dec 2017. 2

[35] Lijun Yin, Martin Reuter, and Bernd Weber. Let the man choose what to do: Neural correlates of spontaneous lying and truth-telling. *Brain and Cognition*, 102:13 – 25, 2016. 3

[36] Z. Zhang, V. Singh, T. E. Slowe, S. Tulyakov, and V. Govindaraju. Real-time automatic deceit detection from involuntary facial expressions. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–6, June 2007. 2