



## A Deep Learning Powered System to Lie Detection While Online Study

Le Quang Thao<sup>1,2\*</sup>, Duong Duc Cuong<sup>2</sup>, Nguyen Nhan Nhi<sup>3</sup>, Nguyen Duc Tam<sup>4</sup>

<sup>1</sup> Faculty of Physics, VNU University of Science, Hanoi 100000, Vietnam

<sup>2</sup> University of Science, Vietnam National University, Hanoi 100000, Vietnam

<sup>3</sup> VNU HUS High School for the Gifted Students, Hanoi 100000, Vietnam

<sup>4</sup> Aachen University of Applied Sciences, Aachen 52066, North Rhine-Westphalia, Germany

Corresponding Author Email: [thao1q@hus.edu.vn](mailto:thao1q@hus.edu.vn)

<https://doi.org/10.18280/ts.390314>

### ABSTRACT

**Received:** 9 April 2022

**Accepted:** 11 May 2022

**Keywords:**

*lie detection, contrastive learning, MobileNetV2, self-supervised learning*

Many education facilities have recently switched to online learning due to the COVID-19 pandemic. The nature of online learning makes it easier for dishonest behaviors, such as cheating or lying during lessons. We propose a new artificial intelligence - powered solution to help educators solve this rising problem for a fairer learning environment. We created a visual representation contrastive learning method with the MobileNetV2 network as the backbone to improve predictability from an unlabeled dataset which can be deployed on low power consumption devices. The experiment shows an accuracy of up to 59%, better than several previous research, proving the usability of this approach.

## 1. INTRODUCTION

Deception has come to be recognized as an instinctive human trait because of its usefulness for highly intelligent, highly social species such as humans. Deceptions can be found anywhere in daily human life, from small humorous lies without malignancy, to ill-willed lies that bring misery to others. DePaulo et al. [1] show that most lies are self-centered, mainly used to gain personal benefits or avoid punishments. Although their deception is essentially not encouraged and actively practiced, or even in some cases, severely punished; research has shown that children's ability to deliver lies continuously improves as they grow up, even when they are just beginning to learn to speak [2, 3]. Furthermore, a large survey showed that the age group of 13-17 had the highest lying rate, up to 74% [4, 5]. Teenagers lie on average 1 to 5 times a day, while the grown-ups lie on average twice a day throughout their adult life. While we cannot deny the role of lying, such as when it is used to avoid offending others or during special circumstances like wars [6], it can pose serious risks and have severe consequences to one's relationships [7]. Liars themselves usually experience a sense of guilt, apprehension, loss of dignity, and trustworthiness [8]. ten Brinke et al. found a link between deceptive behaviors and health, in which short-term adverse effects that occur in liars include increased blood pressure and cortisol, vasoconstriction, and drain of emotion-regulating brain regions [9]. Patients and Stern lying about their true symptoms is usually a major obstruction to the diagnostic process, or they may purposefully mangle in hope of escaping legal conviction [10].

In criminal interrogation, the police have to use polygraphs that test blood pressure, pulse, respiration and skin conductivity, requiring skin-contact devices. Naturally those kinds of measures are impractical for daily life uses. The interlocutors are contingent on observing the subjects' behaviors, facial expressions, and speech to judge the dishonesty. However, these operations require considerable

experience, and sometimes the decision could be affected by prejudice [11, 12]. Deception is the use of wrong words and phrases that cause misunderstanding and distrust in relationships and can even seriously affect relevant aspects of society. Therefore, it is necessary to change and improve the methods of detecting liars. But most efforts so far have focused on simulating real-life setting scams. This is not only impractical, but also biased when participants are instructed to perform the act of lying [13, 14]. Other methods that focus on analyzing brain activity with neuro-imaging pictures have shown that differences in regions of the brain when these are associated with sureness and when with lying, meaning that the brain is being uncertain and having to imagine, can differentiate subject's behaviors. Similarly, measuring electrical activity or hemoglobin signals also helps define physiological features for lying [15-18]. While the most accurate method for lie detection is still using professional polygraph machines, measuring detailed biological cues like heartbeat or blood pressure, with the accuracy of 81 to 91% [19]. It is very invasive and requires professionals to operate such intricate machines. As a result, this is not suitable for modern society and in some areas going on in the current context, such as during online classes, exposing liars seems to be unfeasible as teachers could not recognize any behavioral and facial cues just looking through the blurred screens. Lacking awareness would make teachers easily overlook a cheating student; therefore, the appearance of a supporting device is urgent.

## 2. RELATED WORKS

There has been a lot of research showing that there are more objective indicators of lying behavior. Recently, the identification of small behavioral cues such as eye movements and speech is gaining attention as an application in telling lies. With the help of technical innovations, the process becomes

portable in autonomous systems and less invasive during processing. In 2003 Dionisio et al. [20] pointed out that there is a correlation between the increase in the volunteers' pupil size with their deception behaviors. Webb et al. [21] in 2010 showed oculomotor cues (blinking, saccades, pupillary dilation, etc.) can be detected when a person is lying. Facial expressions also play a critical role in the identification of deception. Ekman [22] defined micro-expressions as relatively short involuntary expressions, which can be indicative of deceptive behavior. Moreover, these expressions were analyzed using smoothness and asymmetry measurements to further relate them to an act of deceit [23]. Tian et al. [24] considered features such as face orientation and facial expression intensity. However, all of the research above was conducted without a constraint on the inference time, which means they cannot be applied in real time yet, and none has a software system to perform lie detection on facial expressions automatically, on the fly. Despite that, these previous works provide a valuable scientific ground for us to conduct this research. About automation facial features recognitions, Owayjan et al. [25] extracted geometric-based features from facial expressions, but the goal of that research is just to classify human emotion based on facial expression. Pfister and Pietikäinen [26] developed a micro-expression dataset to identify expressions that are clues for deception and achieved an accuracy of 55~70% depending on the inference time (the higher the accuracy, the lower the fps), but their system was still hard to deploy on lower end devices.

While there are many other approaches that undertake the lie detection problem, none of them, to the extent of our knowledge, are built with the goal to have a broad range of supported devices, especially the ones in the lower end. As more and more educational institutions are trying offering online learning parallel to the traditional on-site learning method as a permanent option, not every educator has excessive computational power to run complex supporting tools that they need for their job. Unlike onsite learning, educators are much more limited in ways of preventing cheating or deceptions in classes, which is already really common even in offline classes. Although there are several works related to this problem, none has tried to solve this more and more pressing problem. Because of that, we propose this work which describes a progressive research of an intelligent system designed to detect liars from online class learning platforms that can run in real-time even on less computationally powerful devices. Our chosen dataset is based on the one performed from consenting volunteers on a real

traditional liars' game data show, since one main drawback of existing algorithms is that they don't possess a large enough labeled dataset. In this paper, we explored an alternative scenario: self-supervised learning (SSL) [27], where the input data itself provides supervisor signals. The data is pre-processed into a set of user images from the video frames. These images will be encoded into latent representations using a MobileNetV2 [28] model, that is pre-trained by a self-supervised method on a dataset of human faces. The model will create universal representations of the human face and then use the fully connected layer (FC) to classify the extracted action units as part of dishonest or honest behavior. Testing results show that the model can perform slightly better than normal human perception level (by having a higher recognition rate), depending on input conditions such as input image quality or individual facial features.

### 3. MATERIALS & METHODS

#### 3.1 Schematic design

Our lie detection system performs two tasks as shown in Figure 1. In the first one, we use simple framework for contrastive learning of visual representations (SimCLR) [29] architecture to pre-train the MobilenetV2 backbone on the CelebA [30] dataset. In this approach, the MobilenetV2 backbone will learn how to extract enhanced features of faces in a sample that is close together while far from different samples. This pre-trained MobilenetV2 backbone's knowledge will be transferred to the next stage classification task. Finally, video data from online classes is preprocessed with the OpenCV tool to crop facial areas. We obtain a face dataset with two labels telling the truth and lying. This dataset will be trained by the MobilenetV2 backbone pre-trained with the FC layer attached to perform the classification function.

#### 3.2 Dataset preparation

In this research, we use two separate datasets for two different steps: first we used the free and publicly available CelebA dataset to pre-train the model for advanced facial features from photos using SimCLR, then we created our own facial feature dataset to train the model to be able to detect lies based on the speaker's facial expression using transfer learning on the previously pretrained model.

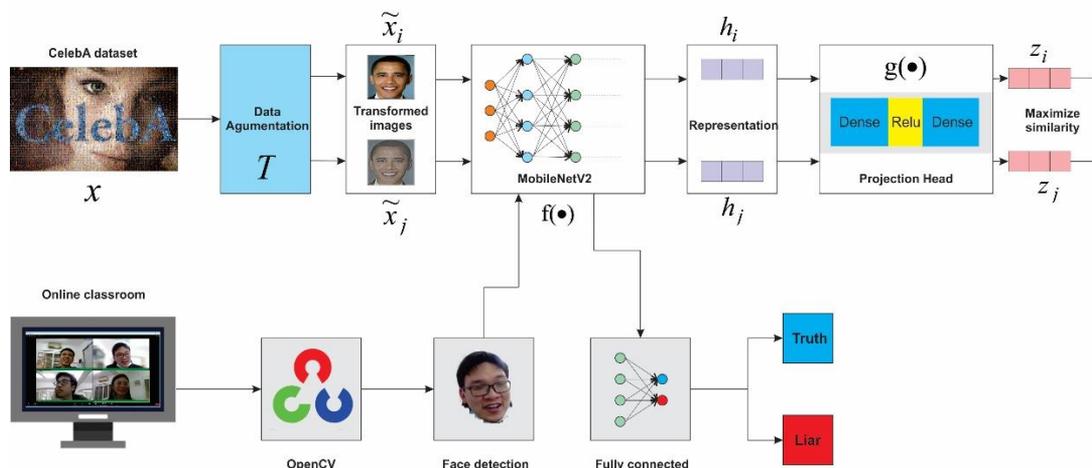


Figure 1. Liar detection system model

### 3.2.1 CelebA dataset

The CelebA [30] dataset is a public and free dataset consisting of over 200000 facial images of over 10000 celebrities, with different input conditions (variants) such as scaling, rotating, focal length, noise, or different backgrounds. With its variety and publicity, CelebA is suitable for training many different types of models, such as facial recognition, facial detection, localizing facial features, or generating new faces from existing features in the dataset. In this research, we utilize the CelebA dataset using the SimCLR method so that the model can learn complex facial features in the input photos.

### 3.2.2 Creating a new dataset for transferred learning

In an effort to detect dishonesty during online courses, we assembled records from digital devices like laptops and softwares, such as Google Meet or Zoom meeting, that serve for online conferences. The volunteers were disposed of in a famous card game ‘I Doubt It’ [31] with the setup as seen in Figure 2.



(a) Setup for data acquisition



(b) Raw datasets from video

**Figure 2.** Dataset preparation

There would be four players joining in the Zoom Meeting by computer and the others sat behind them and marked the cards, laid down by the players, as ‘True’ or ‘Lie’. The sequences of the notes in the handouts would then be collated to indicate the condition of being truthful or deceitful. This crucial step allows us to classify and label the data preparing for the training.

‘I Doubt It’ is a multiple-player card game with the ultimate objective of removing all the cards without being caught lying, using a standard 52 cards deck evenly dealt to 4 people, which means that each player gets 13 cards. The game starts with a chosen player, the leader, putting down one card face down and at the same time uttering its rank or that of another card. The next player, clockwise, will either choose to pass his turn, or to follow suit and put down a card while stating the same rank, or to challenge the previous player by saying ‘True or Lie’ and revealing the latest card. With the third option, there are two possibilities; If the previous player lied about the rank of his card, he has to take all the cards put down in that round

as his. Otherwise, the challenger takes the cards. After a card is revealed, the game continues with a new round with a new leader. The game ends when the time is over and the winner is the one with the fewest cards. A total of 16 volunteers were involved in this experiment, including 8 males and 11 females between the ages of 18 to 21. All participants can use their personal recordings for research purposes. The data generated will be further augmented to increase the dataset’s variety, thus also increase its representativity. This training set is augmented using conventional image transformations where it is randomly sheared, shifted, flipped, and zoomed as, during actual online classes, there are possibilities of various angles, brightness levels, or blurs due to dreadful cameras. Finally, we have information from 123 records of 16 different faces. After data augmentation, we got a total of 10064 frames comprising 5429 lies (45.00%) and 4635 true statements (55.00%) dataset, which was then split for training and testing. The dataset distribution is shown in Table 1.

**Table 1.** Distribution of dataset

	Train	Test	Total
Lie	4089	1340	5429
True	3301	1334	4635
Total	7390	2674	10064

### 3.3 Self-supervised learning

The main idea of SSL [27] is to create free supervisory labels from visual data and use free supervision to obtain generalizable and transferable representations. One of the simplest forms of a pretext task is to reconstruct the input image using a generative model. The latent representation in the generative model is thought to capture the high-level structures and semantic manifolds of the input distribution. Instead of relying on annotations, self-supervised learning algorithms generate labels from data by exposing relationships between parts of the data, a step believed to be critical to achieving human-level intelligence. The question “how are the images different from each other?” motivated contrastive learning, one important task of SSL. The main idea of contrastive learning is to group an image and its slightly different variations into a latent space, while still maximizing its distance to the other groups. SimCLR [29] is a recent and simple implementation of contrastive learning on visual representation.

### 3.4 MobileNet architecture

The MobileNet [28] deep learning architecture is designed to be able to run on as little computational power as possible. Before MobileNet, it was practically unfeasible to train complex deep learning models on lower power consumption devices, especially consumer graded ones. By using depthwise separable convolutions, the architecture separates the traditional convolution into 2 steps: depthwise convolution and pointwise convolution. Only in the pointwise convolution step does the number of channels increase (which means the output feature maps have more channels than their inputs). By separating the standard convolution into two parts then only allowing the numbers of channels to increase in the second one, MobileNet can greatly reduce the parameter size and lower the system requirement.

### 3.5 SimCLR based MobilenetV2 backbone

We chose the SimCLR architecture for the self-supervised pre-training of the mobileNetV2 backbone shown as in Figure 3.

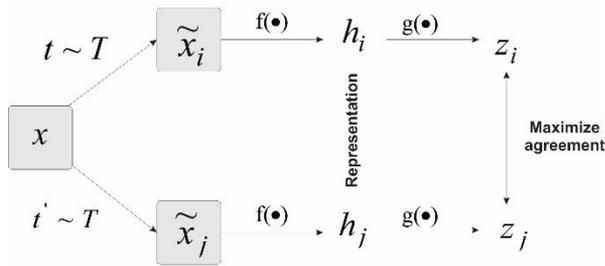


Figure 3. Framework for contrastive learning architecture

The main operating structure is that we need to use the over 200,000 unlabeled facial photos from the CelebA [30] dataset to train our lie detection image recognition model. Assume we have  $x$  for every image’s two differently augmented versions ( $\tilde{x}_i$  and  $\tilde{x}_j$ ) at each iteration, all two of them are encoded into an 1-D feature vector. The encoder network is consist of two separated components:  $f(\cdot)$  is the base encoder network and  $g(\cdot)$  is the projection head. The base network is in most cases a deep CNN, since we use MobilenetV2 for this is the most popular and lightweight architecture commonly used for mobile devices. We would use this base network to extract a representation vector ( $h_i$ ) from the augmented input data:  $f(\tilde{x}_i) = h_i$ . The representation  $h$  will then be mapped into a special space by the projection head  $g(\cdot)$ , after that we would apply contrastive loss there. Finally when the training with contrastive learning is finished, the projection head  $g(\cdot)$  can be dispose of since our pre-trained feature extractor only involves  $f(\cdot)$ . The reason for this is that the projection head  $g(\cdot)$ ’s representations when tested, showed worse performance than those of the base network  $f(\cdot)$  so it would not be suitable when fine-tuning the network for another task.

To maximize the similarity between  $z_i$  and  $z_j$  (the two augmented versions) in the Figure 3, the loss can be formally written as:

$$l_{i,j} = -\log \frac{\exp(\frac{sim(z_i, z_j)}{\tau})}{\sum_{k=1}^{2N} 1_{[k \neq j]} \exp(\frac{sim(z_i, z_k)}{\tau})} \quad (1)$$

where, the hyperparameter temperature  $\tau$  representing how peaked the distribution is. With this hyperparameter, we can balance the influence of both positive labeled input (similar patches) and negative ones (dissimilar patches) on the model’s output, since many similarity metrics are bounded. the temperature parameter allows us to balance the influence of many dissimilar image patches versus one similar patch. The function  $sim$ , a similarity metric used in SimCLR is defined as below:

$$sim(z_i, z_j) = \frac{z_i^T \cdot z_j}{\|z_i\| \cdot \|z_j\|} \quad (2)$$

### 3.6 Training

In the pre-training process, we improve the backbone model, trained with 100 epochs in CelebA dataset rich in diverse

annotations, including 10,177 identities, 202,599 face images, 5 landmark locations, and 40 binary attributes annotations per image. In our research, we selected 80% of those to train and 20% to test. The MobileNetV2 backbone will learn universal representations of the human face to extract the key features of similar samples while separating the other distinctive ones. This pre-trained model would then be transferred to learning in the second classification step.

In the transfer learning process, we use OpenCV to crop the face area in the files that have been labeled “Lie” and “True” as mentioned before, these images would then be used in the pre-trained model MobileNetV2 backbone for the extracted feature that has been enhanced by SSL and then classified by FC layer. In this stage was trained with 100 epochs, the batch-size is equal 20, and the learning rate having been optimized with the Adam [32] algorithm is 0.001. Both tasks were done in the Google Colab environment with Python and the Pytorch library with the usage of NVIDIA Tesla P100 PCIe 3.0.

## 4. RESULT

### 4.1 Improved SimCLR-mobileNetV2

In our experiments, we used a MobileNetV2 encoder since it is the most commonly used lightweight CNN architecture. More complex architectures will have better performance but also have a higher system requirement, which is not suitable for our goal. While being lightweight, MobileNetV2 is still capable of training our model with the large, unlabeled CelebA dataset, enabling us to train on much larger uncurated datasets with billions of images. By respectively altering the transformed augmentation and hyperparameter temperature we evaluate the outcomes of the model. All the results used a fixed batch size with 256 and Adam optimizer. After the best model checkpoints are saved, we track the accuracy of the top-1 and top-5 with metric validations. Results are shown in Table 2 and the training process as shown in Figure 4.

Table 2. Top-1/Top-5 accuracy results on ImageNet with SimCLR-mobileNetV2

Transform	Temp	Opt	acc-1	acc-5
AutoAugment	0.1	Adam	87.49%	94.30%
AutoAugment	0.01	Adam	87.01%	94.89%
RandAugment	0.1	Adam	91.09%	96.35%
<b>RandAugment</b>	<b>0.01</b>	<b>Adam</b>	<b>91.24%</b>	<b>96.87%</b>

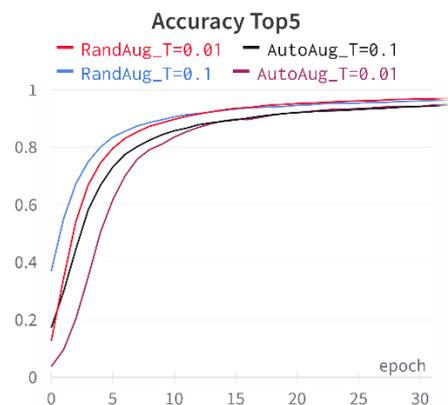


Figure 4. Top-5 accuracy when changes in augmentation and hyperparameter

AutoAugment automatically searches for improved data augmentation policies while RandAugment practices data augmentation with a reduced search space. The results show that AutoAugment outperforms all other data augmentation strategies on our model.

We follow that embedding with  $\tau = 0.01$  is distributed better and evenly, although the embedding with  $\tau = 0.1$  is more reasonable and locally clustered and globally separated. Smaller temperature benefits training more than higher ones, but extremely low temperatures are harder to train due to numerical instability. Indeed, as shown in Figure 4, small temperatures tend to slowly convergence of accuracy but it gives better results.

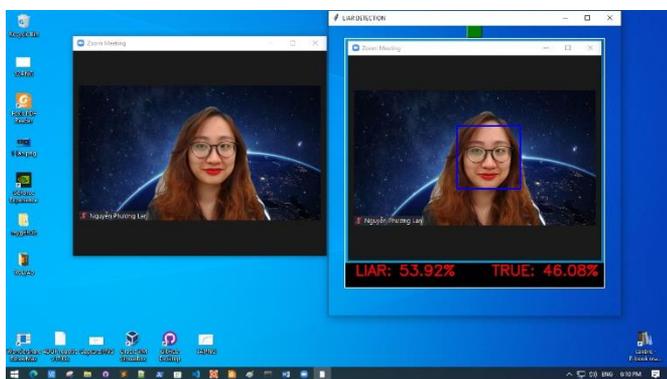
### 4.2 Comparison of Self-supervised learning to baseline approach

Choosing the optimal parameters after the previous training step, we use the pre-trained MobileNetV2 backbone model to extract the features for the pre-prepared data, then utilize the neural network-based classifier of the FC layer to detect liars. For comparing the performance of SSL, we also conducted deep learning in the pre-trained MobileNetV2 model on ImageNet dataset. The outcomes of both above methods are shown in Table 3. There are many metrics to evaluate the performance of a predictive model, but we choose the accuracy and F1-score here since they are relatively simple yet effective to represent the prediction performance. As the overall comparisons of the best performers are presented in Table 3, the origin model for lie detection yields reasonably modest results with an accuracy of approximately 57.89%. On the other hand, the SSL based model produces more accurate estimates, achieving the best accuracy so far in this dataset of 59.15% and F1-score is 54.38%.

**Table 3.** Comparison with baseline models

Model	Accuracy	F1 score
MobileNetV2	57.89%	49.21%
SimCLR-mobileNetV2	59.15%	54.38%

Although there is a notable improvement, this accuracy in our opinion is still quite low, which shows that lie detection is a challenging problem, even for human perception level.



**Figure 5.** Test results when assessing a liar

### 4.3 Real-time, real-life lie detection demonstration

This section evaluates the highest accuracy achieved by the best-performing models from the previous parts for

individual-based lie detection. We conducted the core lie task in this experiment by inviting volunteers to play the ‘Lying Card Game’ in front of the computers, where players put a card and say out loud. We used the pre-trained model to detect the lies from the images acquired by the camera. Our statement is considered a lie if more than 30% of the frames are predicted as “lie” by the proposed program shown as in Figure 5.

## 5. DISCUSSION

Our experiments show that the multimodal system can identify liars with an accuracy above 59%, which is significantly above the chance level; however, some observations that are made based on these data should not be interpreted as scientific conclusions but point to future work. The motivation behind this research stems from the need for accuracy to classify the right suspects as liars without misclassifying genuine people. Although it is better than previous studies, it is still quite inaccurate in our opinion and we believe that this result can be improved further in the future with those above mentioned plans. The lack of more successful studies on this pressing problem also proves that lie detection is one of the more complex problems, but we believe that by following these improvement plans, we can still produce a better result.

## 6. CONCLUSIONS

In this research, we have demonstrated a new usage for SSL, an approach for machine learning that is designed to be more independent to human feedback, while still able to deliver relatively good predictions, on a pre-existing backbone neural network architecture. One of its applications is to pre-train computer vision models on a large set of unlabeled data and to transfer the learned knowledge into downstream tasks, for example here we applied it to the lie detecting problem during online lectures. Most of the existing SSL implementations use large networks as encoders, which are not always applicable, especially for smaller scales. One of the goals of our research here is to tackle that problem, so that SSL-powered models can be deployed on more low power-consumption devices. We believe that our results in this paper can be even further improved and have a plan to revise it in the future. Some works to be done are:

- Acquiring a larger dataset as better data variety normally plays an important role in better generalization.
- Try using a more improved version of backbone network, such as MobileNetV3 instead of MobileNetV2. However, inference time must be taken into consideration, since newer backbone network versions may have better prediction results but also increase required computational power.
- Increase model’s complexity after upgrading hardware components, to find the diminishing return point between model’s accuracy and system requirements.

## REFERENCES

[1] DePaulo, B.M., Ansfield, M.E., Kirkendol, S.E., Boden, J.M. (2004). Serious lies. *Basic and Applied Social Psychology*, 26(2-3): 147-167.

- <https://doi.org/10.1080/01973533.2004.9646402>
- [2] Talwar, V., Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child Development*, 79(4): 866-881. <https://doi.org/10.1111/j.1467-8624.2008.01164.x>
- [3] Xu, F., Bao, X., Fu, G., Talwar, V., Lee, K. (2010). Lying and truth-telling in children: From concept to action. *Child Development*, 81(2): 581-596. <https://doi.org/10.1111/j.1467-8624.2009.01417.x>
- [4] Hancock, J.T., Thom-Santelli, J., Ritchie, T. (2004). Deception and design: The impact of communication technology on lying behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 129-134. <http://dx.doi.org/10.1145/985692.985709>
- [5] Wyudhijit Bhattacharjee. (2022). Why We Lie: The Science Behind Our Deceptive Ways, <https://www.nationalgeographic.com>, accessed on 12 Jan. 2022.
- [6] Ekman P. (2001). *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*. New York, NY: WW Norton & Company.
- [7] Peterson, C. (1996). Deception in Intimate Relationships. *International Journal of Psychology*, 31(6): 279-288. <https://doi.org/10.1080/002075996401034>
- [8] Bok, S. (1999). *Lying: Moral Choice in Public and Private Life*, NY: Vintage Books.
- [9] ten Brinke, L., Lee, J.J., Carney, D.R. (2015). The physiology of (dis) honesty: Does it impact health? *Current Opinion in Psychology*, 6: 177-182. <http://dx.doi.org/10.1016/j.copsyc.2015.08.004>
- [10] Palmieri, J.J., Stern, T.A. (2009). Lies in the doctor-patient relationship. *Primary care companion to the Journal of Clinical Psychiatry*, 11(4): 163-168. <https://doi.org/10.4088/PCC.09r00780>
- [11] Voeller, J.G. (Ed.). (2014). *Social and Behavioral Research for Homeland Security*. John Wiley & Sons.
- [12] Gannon, T.A., Beech, A.R., Ward, T. (2009). Risk assessment and the polygraph. *The Use of the Polygraph in Assessing, Treating and Supervising Sex Offenders: A Practitioner's Guide*, 129-154. <https://doi.org/10.1002/9780470743232.ch8>
- [13] Tsiamyrtzis, P., Dowdall, J., Shastri, D., Pavlidis, I.T., Frank, M.G., Ekman, P. (2007). Imaging facial physiology for the detection of deceit. *International Journal of Computer Vision*, 71(2): 197-214. <https://doi.org/10.1007/s11263-006-6106-y>
- [14] Dcosta, M., Shastri, D., Vilalta, R., Burgoon, J.K., Pavlidis, I. (2015). Perinatal indicators of deceptive behavior. In *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 1: 1-8. <https://doi.org/10.1109/FG.2015.7163080>
- [15] Bhutta, M.R., Hong, K.S., Naseer, N., Khan, M.J. (2015). Spontaneous lie detection using functional near-infrared spectroscopy in an interactive game. In *2015 10th Asian Control Conference (ASCC)*, pp. 1-5. <https://doi.org/10.1109/ASCC.2015.7244666>
- [16] Bhutta, M.R., Hong, M.J., Kim, Y.H., Hong, K.S. (2015). Single-trial lie detection using a combined fNIRS-polygraph system. *Frontiers in Psychology*, 6: 709. <https://doi.org/10.3389/fpsyg.2015.00709>
- [17] Li, F., Zhu, H., Xu, J., Gao, Q., Guo, H., Wu, S., He, S. (2018). Lie detection using fNIRS monitoring of inhibition-related brain regions discriminates infrequent but not frequent liars. *Frontiers in Human Neuroscience*, 12: 71. <https://doi.org/10.3389/fnhum.2018.00071>
- [18] Lai, Y.F., Chen, M.Y., Chiang, H.S. (2018). Constructing the lie detection system with fuzzy reasoning approach. *Granular Computing*, 3(2): 169-176. <https://doi.org/10.1007/s41066-017-0064-3>
- [19] Gaggioli, A. (2018). Beyond the truth machine: emerging technologies for lie detection. *Cyberpsychology, Behavior, and Social Networking*, 21(2): 144-144. <https://doi.org/10.1089/cyber.2018.29102.csi>
- [20] Dionisio, D.P., Granholm, E., Hillix, W.A., Perrine, W.F. (2001). Differentiation of deception using pupillary responses as an index of cognitive processing. *Psychophysiology*, 38(2): 205-211. <https://doi.org/10.1111/1469-8986.3820205>
- [21] Webb, A.K., Honts, C.R., Kircher, J.C., Bernhardt, P., Cook, A.E. (2009). Effectiveness of pupil diameter in a probable-lie comparison question test for deception. *Legal and Criminological Psychology*, 14(2): 279-292. <https://doi.org/10.1348/135532508X398602>
- [22] Ekman, P. (2009). *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*. W W Norton & Co.
- [23] Ekman, P. (2003). Darwin, deception, and facial expression. *Annals of the New York Academy of Sciences*, 1000(1): 205-221. <https://doi.org/10.1196/annals.1280.010>
- [24] Tian, Y.L., Kanade, T., Cohn, J.F. (2005). Facial expression analysis. In *Handbook of Face Recognition*, 247-275.
- [25] Owayjan, M., Kashour, A., Al Haddad, N., Fadel, M., Al Souki, G. (2012). The design and development of a lie detection system using facial micro-expressions. In *2012 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA)*, pp. 33-38. <https://doi.org/10.1109/ICTEA.2012.6462897>
- [26] Pfister, T., Pietikäinen, M. (2012). Electronic imaging & signal processing automatic identification of facial clues to lies. *SPIE Newsroom*, January.
- [27] Khosla, P., Teterwak, P., Wang, C., Sarna, A., Tian, Y., Isola, P., Krishnan, D. (2020). Supervised contrastive learning. *Advances in Neural Information Processing Systems*, 33: 18661-18673. <https://doi.org/10.48550/arXiv.2004.11362>
- [28] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4510-4520. <https://doi.org/10.48550/arXiv.1801.04381>
- [29] Chen, T., Kornblith, S., Norouzi, M., Hinton, G. (2020). A simple framework for contrastive learning of visual representations. In *International Conference on Machine Learning*, pp. 1597-1607. <https://doi.org/10.48550/arXiv.2002.05709>
- [30] Large-scale CelebFaces Attributes (CelebA) Dataset. (2022). <http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html>, accessed on 12 Jan. 2022.
- [31] I doubt it, how to play. (2022). <https://bicyclecards.com>, accessed on 12 Jan. 2022.
- [32] Kingma, D.P., Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. <https://doi.org/10.48550/arXiv.1412.6980>