

An Artificial Intelligence Lie Detection Model using ResNet-18 Network

Saif Mohammed Saif Al Salti,
Alazhar Saif Musallam Al Sharji, Sherimon P.C
Arab Open University
Muscat
Sultanate of Oman

Vinu Sherimon
University of Technology and Applied Sciences
Muscat
Sultanate of Oman

Remya Revi K.
Saintgits College of Engineering
Kottayam, Kerala
India

Abstract—An Artificial Intelligence (AI) Lie Detection system is developed using the deep learning technique which detects the expression of ‘fear’ from face of a liar. The proposed lie detection model uses the transfer learning approach in the deep learning technique to classify the facial expressions into ‘fear’ or ‘neutral’. A pre-trained lightweight deep Convolutional Neural Network architecture called ResNet-18 is used in this work. Publicly available benchmarked Facial Expression Recognition-2013 (FER-2013) dataset is used for the experimental analysis. The model performed well with a classification accuracy of 82.85%.

Keywords—Lie detection; artificial intelligence; deep learning; transfer learning; micro-expression

I. INTRODUCTION

Lie detection is significant process in criminal investigations, security verifications in airports and in many other situations. Psychophysiological or Polygraph is one of the widely used lie detection method [1]. Other methods like Voice Stress Analysis (VSA), functional Magnetic Resonance Imaging (fMRI), Analysis of non-verbal behavior, etc. are yet to be validated. Many have questioned process of the polygraph lie testing method due to the lack of its scientific backing. Also, people are discovering several ways to cheat the polygraph test by consuming antiperspirant, sedatives, etc.

Analysis of facial micro-expressions is another significant lie detection method, where the investigators can detect lies by observing emotional leakage from the facial expressions of liars [2][3]. When an analyst interrogates the people and observe the body and facial expression, it takes a lot of time for the observation and conclusion, and people may only be able to analyze one or two observations at a time. So, in this research, we develop an intelligent system that utilizes human facial micro-expressions to detect lies. Studies shows that the

emotional leakage in the form of ‘fear’ can be observed from human faces when they are lying [4]. This expression of ‘fear’ arises in the face due to the tensions inside a person while lying. Hence in this work, we develop an Artificial Intelligence (AI) Lie Detection model that identifies the expression of ‘fear’ from face of a liar. The proposed lie detection model uses the transfer learning approach in the deep learning technique to classify the facial expressions into ‘fear’ or ‘neutral’. Here, we use a lightweight pre-trained Convolutional Neural Network (CNN) called ResNet-18 architecture to develop the model. The experiments are conducted using Facial Expression Recognition-2013 (FER-2013) dataset [5].

The upcoming sections discuss about the related works, proposed method, experimental result analysis, and conclusion.

II. RELATED WORKS

The following are the brief discussion of works related to research on the lie detection using AI and machine learning techniques.

Owayjan et al. developed a LabVIEW computer vision software based on a mathematical algorithm to analyze the facial micro-expressions to detect lies [6]. Singh et al. developed an image processing method to detect lies based on eye blink rates. They used the Haar cascade algorithm to locate eye region from the face [7]. Soumya Barathi designed a lie detection system by analyzing facial micro-expression, body language, and speech. She used Principal Component Analysis, Speeded Up Robust Features, and k-means algorithm in the lie detection system [8].

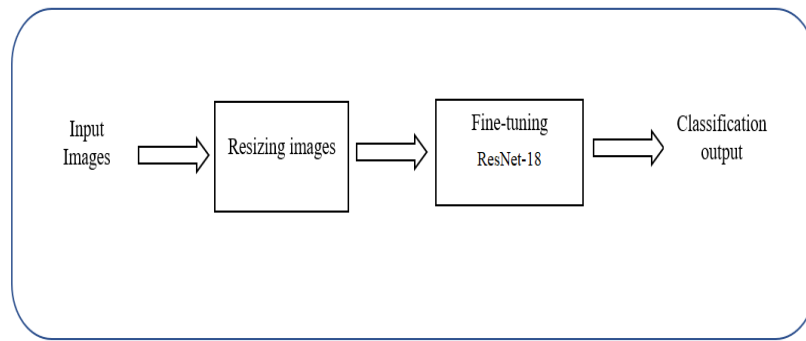


Fig. 1. Outline of the proposed method

Shen et al. analyzed the facial expression of fear to detect lies from videos. They used Random Forest, k-nearest neighbors, and bagging classification algorithms [4]. Diaz et al. designed a machine learning model to detect lies from facial expressions. They extracted handcrafted features from the facial images as well as extracted deep features from the facial images. [9]. Monaro et al. developed lie detection model by analyzing facial micro-expressions. Features are extracted using Open Face, and classifiers used are Support Vector Machine and Long Short-Term Memory network. In addition, they designed a 3-dimensional CNN architecture to extract features directly from the images and to classify the facial expressions [10].

III. PROPOSED METHOD

The proposed lie detection system is a two-class classification problem, to classify the facial expressions into 'fear' or 'neutral'. The proposed method uses a deep neural network architecture for the classification, and outline of the method is displayed in Fig. 1.

The deep learning techniques which use CNNs give remarkable results in the computer vision tasks [11-13]. A CNN is designed to mimic the human visual system and basically used in image recognition tasks [14]. To obtain accurate and reliable classification results, a CNN should be trained on huge quantity of labelled image data. But, in most real-life circumstances, there is a shortage in the availability of the number of labelled data, and hence, training the network from scratch is a difficult procedure. So, to overcome this problem, we can utilize a pre-trained CNN using the transfer learning approach [15-18]. The lie detection is also a small sample size classification problem, and hence, we use the transfer learning approach to develop the classification model. In this work, we utilize a lightweight pre-trained CNN architecture called ResNet-18, which is already trained on a large number of images from the ImageNet dataset. The specifications of ResNet-18 are: 11.7 million learnable parameters, depth is 44, size is 18 MB, input size is 224 x 224 x 3 [19]. The layers of this pre-trained network learned the various basic features as well as discriminative features [20]. Hence, in this research work, we fine-tune the ResNet-18 to classify the facial expressions into 'fear' or 'neutral' in the lie detection problem. To fine-tune the network using the transfer learning, as a first step we remove the Fully Connected (FC) layer of the ResNet-18, and is replaced with a new FC layer

with two classes for the lie detection problem. Then, we train the reorganized network using the dataset for the two-class classification.

IV. RESULT AND DISCUSSION

A Graphic Processing Unit (GPU) based system having NVIDIA GTX 1060 card, 6.0 GB RAM is used to perform the experiments. The experiments and analysis are done on MATLAB platform with help of Deep Learning Toolbox. The next subsections briefs about the dataset used, evaluation metrics, training options, and experiments results and analysis.

A. Dataset Used

The 'fear' and 'neutral' facial expression images from the Facial Expression Recognition-2013 (FER-2013) dataset are used for the experimental analysis. A total number of 10,000 images are used, i.e., 5,000 images each in 'fear' and 'neutral' classes. The images are in png format with a size of 48 x 48.

B. Evaluation Policies

The performance evaluation metrics used for assessing the model are "Accuracy, Precision, Recall, and F1-score" are, and are calculated based on (1), (2), (3) and (4), respectively. They are calculated from the confusion matrix of the model using "True Negative (TN), True Positive (TP), False Positive (FP), and False Negative (FN)".

$$Accuracy(\%) = 100 \times \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (1)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (2)$$

$$Recall = \frac{TP}{(TP+FN)} \quad (3)$$

$$F1 - score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \quad (4)$$

C. Training Options

The options used to train the network is provided in Table 1. From the dataset, 60% are used for training the model, 20% are used for validating the model, and 20% are used for testing the model. The size of the images is resized to 224 x 224 x 3 resolution based on the input size of the ResNet-18 network.

TABLE I. TRAINING OPTIONS

Learning rate	0.0001
No. of epochs	50
Mini-batch size	64
Optimizer	adam

TABLE II. PERFORMANCE EVALUATION

Accuracy (%)	Precision	Recall	F1-score
82.85	0.781	0.863	0.819

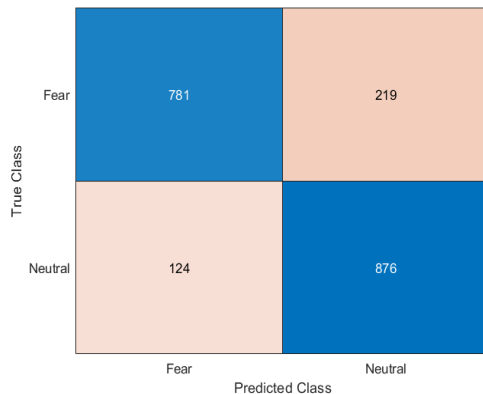


Fig. 2 Confusion matrix of the model

D. Experimental Results and Analysis

The ResNet-18 network is fine-tuned on the dataset for classifying into the two categories using the above-mentioned training options. The confusion matrix obtained while testing the model is shown in Fig. 2. The values of the performance metrics are evaluated from the confusion matrix, and is given in Table 2. The analysis shows that, Accuracy, Precision, Recall, and F1-score achieved by the model are 82.85%, 0.781, 0.863, and 0.819, respectively. This experimental analysis shows that the proposed deep learning lie detection method achieved a decent performance.

V. CONCLUSION

In this work, a deep learning model is developed to detect lies by analyzing human facial expression. We utilized the emotional leakage in the form of 'fear' to detect lies. An intelligent lie detection model using ResNet-18 deep neural network is developed. The model is evaluated using the (FER-2013) dataset. The model exhibited a good performance with a detection accuracy of 82.85%.

ACKNOWLEDGMENT

The research leading to these results has received funding from the Research Council (TRC) of the Sultanate of Oman under the Block Funding Program "BFP/URG/ICT/21/152".

REFERENCES

- [1] J. Winkler-Galicki, J. Bartkowiak-Wieczorek, D. Synowiec, R. Dąbrowska, and E. Mądry, "Polygraph analyses: technical and practical background," *J. Med. Sci.*, vol. 91, no. 1, pp. 39–46, 2022, doi: 10.20883/medical.e590.
- [2] R. Tamuli, S. Samantara, S. Sarkar, and S. Adhikari, "Lie Detection Using Facial Micro-Expressions," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 4, no. 11, pp. 25–30, 2020.
- [3] S. Jia, S. Wang, C. Hu, P. J. Webster, and X. Li, "Detection of Genuine and Posed Facial Expressions of Emotion: Databases and Methods," *Front. Psychol.*, vol. 11, pp. 1–12, 2021, doi: 10.3389/fpsyg.2020.580287.
- [4] X. Shen, G. Fan, C. Niu, and Z. Chen, "Catching a Liar Through Facial Expression of Fear," *Front. Psychol.*, vol. 12, pp. 1–10, 2021, doi: 10.3389/fpsyg.2021.675097.
- [5] M. SAMBARE, "FER-2013 Learn facial expressions from an image." <https://www.kaggle.com/datasets/msambare/fer2013> (accessed Jul. 02, 2022).
- [6] M. Owayjan, A. Kashour, N. Al Haddad, M. Fadel, and G. Al Souki, "The design and development of a lie detection system using facial micro-expressions," in *2nd International Conference on Advances in Computational Tools for Engineering Applications*, ACTEA 2012, 2012, pp. 33–38. doi: 10.1109/ICTEA.2012.6462897.
- [7] B. Singh, P. Rajiv, and M. Chandra, "Lie detection using image processing," in *Proceedings of the 2nd International Conference on Advanced Computing and Communication Systems (ICACCS -2015)*, 2015, pp. 1–5. doi: 10.1109/ICACCS.2015.7324092.
- [8] Soumya Barathi. C, "Lie Detection based on Facial Micro Expression, Body Language and Speech Analysis," *Int. J. Eng. Res. Technol.*, vol. 5, no. 02, pp. 337–343, 2016, doi: 10.17577/ijertv5is020336.
- [9] N. Rodríguez-Díaz, D. Aspandi, F. M. Sukno, and X. Binefa, "Machine learning-based lie detector applied to a novel annotated game dataset," *Futur. Internet*, vol. 14, no. 1, pp. 1–18, 2022, doi: 10.3390/fi14010002.
- [10] M. Monaro, S. Maldera, C. Scarpazza, G. Sartori, and N. Navarin, "Detecting deception through facial expressions in a dataset of videotaped interviews: A comparison between human judges and machine learning models," *Comput. Human Behav.*, vol. 127, pp. 1–10, 2022, doi: 10.1016/j.chb.2021.107063.
- [11] N. Sharma, V. Jain, and A. Mishra, "An Analysis of Convolutional Neural Networks for Image Classification," in *Procedia Computer Science*, 2018, vol. 132, pp. 377–384. doi: 10.1016/j.procs.2018.05.198.
- [12] F. Sultana, A. Sufian, and P. Dutta, "Advancements in image classification using convolutional neural network," in *2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, 2018, pp. 122–129. doi: 10.1109/ICRCICN.2018.8718718.
- [13] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, and Y. Miao, "Review of image classification algorithms based on convolutional neural networks," *Remote Sens.*, vol. 13, no. 22, pp. 1–51, 2021, doi: 10.3390/rs13224712.
- [14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [15] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: An astounding baseline for recognition," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 512–519. doi: 10.1109/CVPRW.2014.131.
- [16] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," in *Advances in Neural Information Processing Systems*, 2014, pp. 3320–3328.
- [17] N. Becherer, J. Pecarina, S. Nykl, and K. Hopkinson, "Improving optimization of convolutional neural networks through parameter fine-tuning," *Neural Comput. Appl.*, vol. 31, no. 8, pp. 3469–3479, 2019, doi: 10.1007/s00521-017-3285-0.
- [18] R. N. D'souza, P. Y. Huang, and F. C. Yeh, "Structural Analysis and Optimization of Convolutional Neural Networks with a Small

- Sample Size,” Sci. Rep., vol. 10, no. 1, pp. 1–13, 2020, doi: 10.1038/s41598-020-57866-2.
- [19] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.
- [20] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, “Learning and transferring mid-level image representations using convolutional neural networks,” in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2014, pp. 1717–1724. doi: 10.1109/CVPR.2014.222.