




The Journal is Available at DOI: 10.5281/zenodo.11212637.
International Journal of Computer Aided Design and Applications, 2024
Journal Original Website: <http://ijcada.ajournals.org/>



Cognitive Perception of People Based on Facial Recognition and Subjective Estimation

Fatima Isiaka 

Department of Computer Science, Nasarawa State University, Keffi, Nigeria

Zainab Adamu

Department of Computer Science, Ahmadu Bello University, Zaria, Nigeria.

History

Received: 12nd, April 2024

Revised: 23th, April 2024

Accepted: 26th, April 2024

Published: 3th, May 2024

Keywords:

Face recognition system,

Human cognition,

Knowledge acquisition,

Real-time cognitive perception,

Skin conductance,

Cognitive response

Abstract:

In face recognition system implementation, human factors play a considerable role in the design and interpretation of remote sensing of images for pattern classification. Image sensing and cognition are the critical topics in image interpretation that brings together current state of art procedures which are widely recognised and tackled, especially in facial recognition image sensing with complex input data. It addresses most themes around expertise including the methods for knowledge acquisition and modelling of the expertise behind the modules and systems. This paper uses a facial recognition system at its production stage to test the model's reliability and accuracy based on users' subjective perception and real-time cognitive perception of the subjects used for the test procedure. Eighty people were interviewed while we took their skin conductance response (SCR) and they interacted in real-time with the window output of the facial recognition system, this is to obtain a subjective view of their cognitive response to the interface based on the best match of a particular face to its foreground image. The result shows that most participants considered the image fit to be above average match to the original greyscale image and should not be part of feature classification for a facial recognition system.

NOTE: The article and abstract is published under the International Journal of Computer Aided Design and Applications.

1 INTRODUCTION

In recent times, the cognitive perception of people has moved towards the use of more intangible stimuli in the need to optimise experimental control. In cognitive perception, most cognitive psychology experiments evaluate the best model in human cognition using contemporarily small and well-controlled stimuli set of image sensory in facial recognition. This particular method is in contrast to some current research fields in system performance testing and user perception in computer vision. One of the focuses is on using a large-scale database of natural images to test the accuracy of a

match based on the subjective user's perception of the resultant search. The natural images also provide a controlling tool for characterising the statistical scenario of the captured image for better evaluation of psychological theories behind the accuracy and match and also for bringing insight into cognitive mechanisms to real-life applications. This looked at how the challenges of using natural images of the facial recognition system as stimuli in experiments can be addressed by using increased sample sizes and basic representations from computer visions to help in system usability and also to develop novel experimental methods. We first approach this by giving a brief literature review on a database of facial systems to explore some questions about human cognition on six different levels, high match, weak (low) match, average match, above-average match, and below-average match, and above-high match. This set of information obtained from the cognitive perception of the users is set as a basic attribute in test matches of the input matrix and used to classify the best fit for an image.

2 LITERATURE REVIEW

Research on Cognitive Perceptual Internet (CPI) [17, 5, 13, 4, 11] introduces a possible architecture and did some discussions on the directions for the state realisation of the method, which has enabled the convergence of both wired and wireless communications in image sensing networks and mobile sensing with machine learning techniques [25, 23, 24, 16, 26, 1, 21]. The concept behind this stems from the fact that technologies like smartphones and wearable sensors are very now becoming outstanding due to pervasive diffusion and the increasing amount of embedded sensors. The data collected by these forms of technology can provide unprecedented insights into the real world that be inferred through cognitive perception and the evolvement of the digital sixth sense [22, 9, 3, 10]. The paper describes how online interaction can behave like a sensing brain, which allows for the concept of the internet of senses in a network-based cognitive perception and action capabilities that are built upon the mobile crowd-sensing phenomenon [12, 2, 14, 6]. The new concept provided by the CPI is envisioned as a form of an efficient geo-referenced library of knowledge where the physical world is depicted as a constant reality. Such ideas are acquired and augmented through sensors and multiuser distributed learning processes.

The possibility of accommodating practical sensors in addition to normal reactive ones such as antennas, cameras, and inertial measurement nuts has explored the massive availability of mobile terminal perception of users as well as its range of new applications. The proposal of the Internet of Senses becomes a reality in the application of CPI. Other papers describe survey theories and ideas about image processing where some illustrative examples are taken from the human information processing laboratory (HIPL) at N.Y.U [18, 15, 19, 8]. The procedures started at the initial stage where they developed the concepts of multiple stable states and path dependence in a basic visual-motor task i.e. the vergence of the eyes and this shows how these can be encompassed in potential theory.

The second part considered two examples of human information extraction from complex dynamic, visual displays, the extraction of shape and motion of 3D wire objects from 2D images, and also the extraction of meaning from displays of deaf signers' communication in American Sign Languages (ASL)[20, 7]. These methods were used for the adaptation of the cognitive perception of the user for process optimisation. The method adopted in this paper uses the correlates of physiological measures and user perception to obtain an aggregate of users' cognitive perception through image sensing. This is also supported by the subject perception of the users.

3 METHOD

Figure 1 shows a current window interface of a facial recognition system that has undergone visual and automated test performance for the accuracy and reliability of the system. This interface is used as a visual stimulus to obtain user cognitive perception of the realistic match of the foreground image to the background image based on boundary assessment in a subjective view that correlates to their physiological response (SCR) in real time. Eighty student participants were selected for the study, where an SCR is placed on their hand facing an eye tracker (Figure 2) as they interact with the visual stimuli. The task is to search between the foreground and background image detected as the best fit for a facial recognition system and to identify the best fit between the two visions based on the criteria in Table 1. This is compared with the resulting output produced by the inference engine.

Performance measures in facial recognition systems are sometimes not enough to rely on for realistic accuracy, hence users' cognitive perception is set as an additional parameter to test and understand the accuracy from a cognitive perceptive internet point of view. The rationale for using SCR is also to determine its correlates to subject views in real-time thereby authenticating the subject cognition of the participants with regards to stress, neutral, complex, and average mood-to-subject interaction (Figure 3). Due to the complexity of personal and behaviour data, physical systems though very accurate can also be prone to error just as deep learning methods. Subjective views are also very

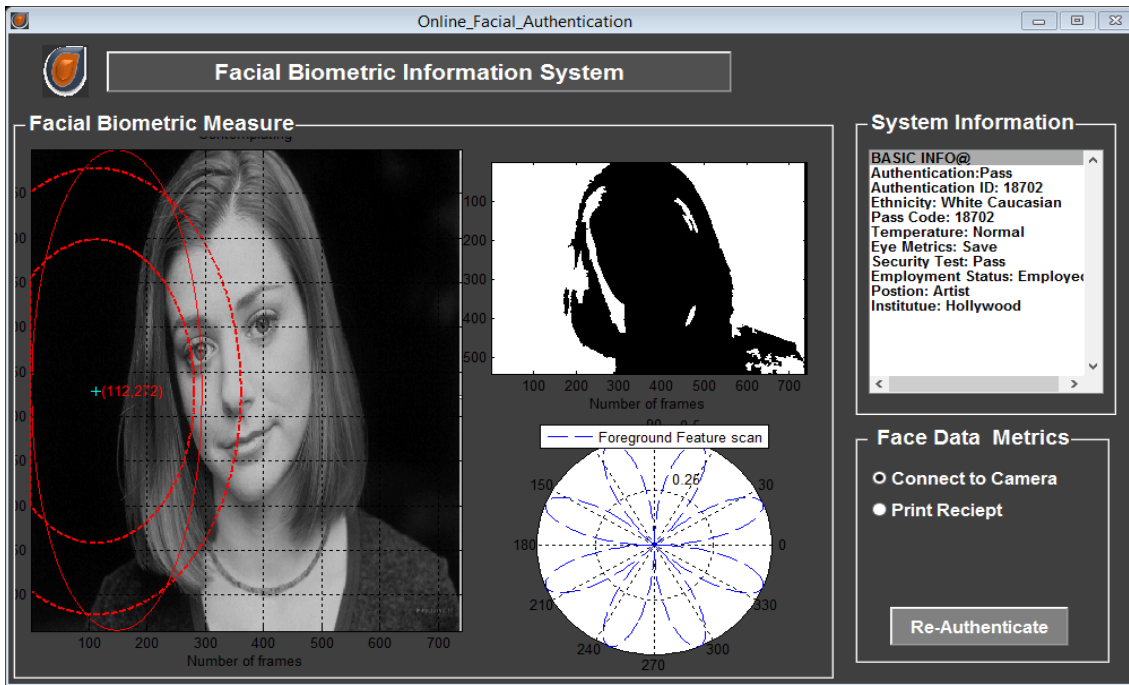


Figure 1: Interface of the image recognition system both in foreground and background feature classification

Table 1: Estimated cognitive response of participants to visual stimuli.

Cognitive perception	Vision type	Assigned value	Gender
High	Foreground	2.0	Female
Average	Background	3.5	Female
Low	Foreground	3.0	Male
Below Average	Foreground	4.3	Female
Above Average	Background	5.4	Male
Above High	background	3.3	Female

important in a realistic environment. The measures from their subjective views are discussed in the section below, and the rate of acceptance and judgment is discussed here.

4 RESULT

The data generated from their interaction and subjective views are simulated to 800 and the performance score is based on a scale of 105 in performance rate. The simulated data was divided into four segments containing both genders, below middle and average age. Figure 4 shows the performance rate of female students in response to their perception of the facial recognition system. The acceptance rate was scored above average in terms of best fit of background image to foreground image, others voted for below average and high facial and feature match to the input image as compared to low and average match. Some participants from that category also perceive the image as a low match and nothing could be concluded from that aspect.

Figure 5 shows the performance rate of Male participants in regards to their view of the facial recognition system. The acceptance rate for the best fit or match to the foreground and background image lies between above-average, high, and below-average certain groups of people in the group, while others are seen to have a very low perception of the image match. The variations in acceptance are another topic to be discussed to determine if this margin in acceptance is due to poor eye color variation in the lens aspect. The low acceptance rate signifies the measure of the likelihood that the biometric system is correctly accepted and mismatched the input image. This is also similar to the low recognition rate as a measure of the likelihood that the system has incorrectly rejected and mismatched the input image. Both high and above high acceptance and recognition rates are false in the level of accuracy as similar to a

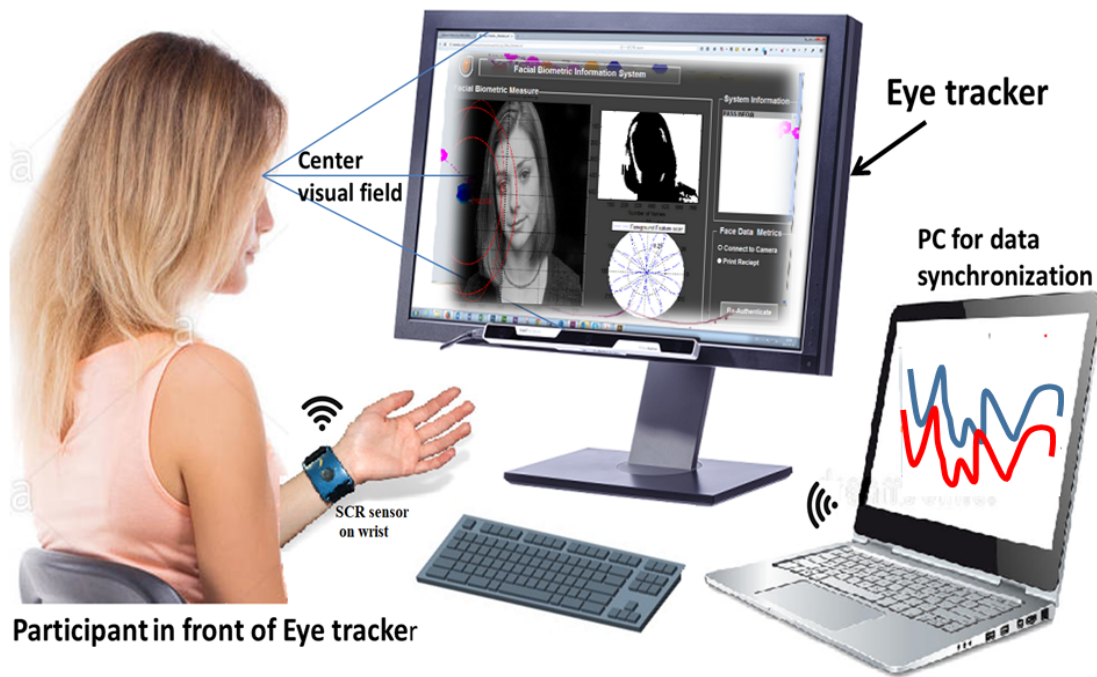


Figure 2: Participant sitting in front of Eye tracker containing visual stimuli of facial recognition system .

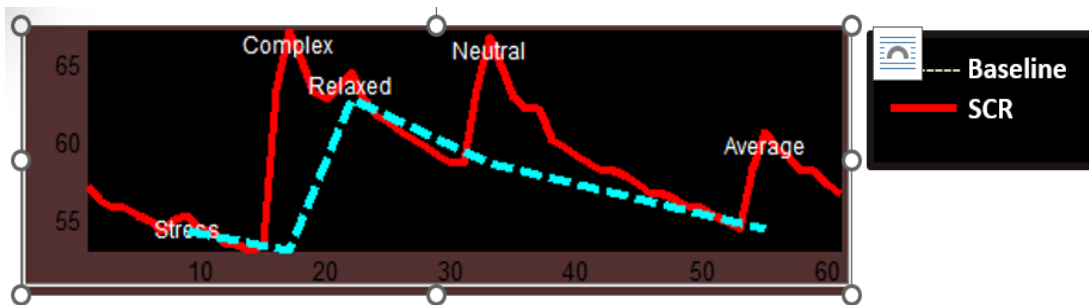


Figure 3: Physiological reading of a participant in response to the subject view of the visual interface. .

perfect match.

Most inference engines in a biometric system can run in two different modes as regards identification verification. The process of identification is trying to find out a person's identity by experimenting with the biometric pattern computed from the input image. The identification process shows that the system is trained with the same patterns both in the foreground and background for several input images. A template is calculated on the trained dataset and this pattern produces a match for every template. This process is based on automatic and system testing. Users' perception can give not only an authentic view but also a cognitive perceptive internet point of view from direct interaction with a person in a subjective manner. Figure 6 shows the performance rate of older participants concerning the accuracy in the match to the input image to its background form, above average, below average match and high match are the two most rated perceptions from this group and can be related to its accuracy from the dynamic control system.

Figure 7 indicates the performance rate of younger participants with above average, below average matches and high matches are the two most rated perceptions from this group as with the previous participants, there is an unequal variation among this set of groups. The performance in the subjective view is more accurate as compared to the control systems' predictions of an exact match. A person can more accurately identify a perfect match physically than a controlled digital procedure but since a person's predictive potential can be mimicked, a machine can also be trained to learn on its own and become more accurate than a subjective viewpoint. In this case, deep learning can also be adapted for a perfect match with minimal error.

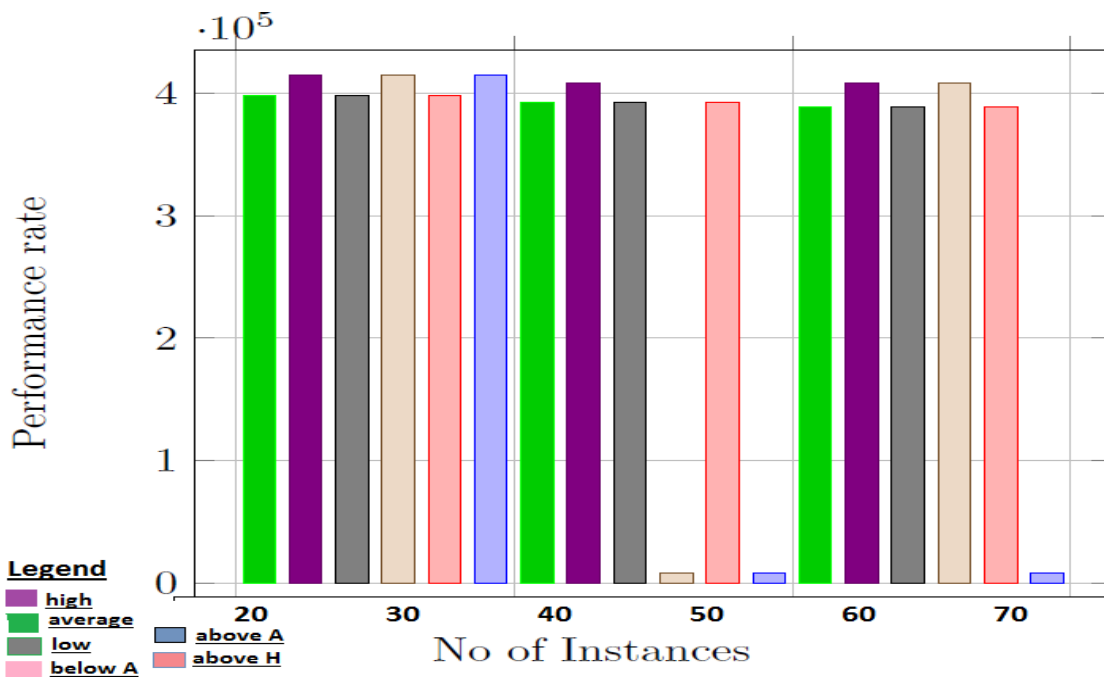


Figure 4: Performance rate of Female participants in response to facial recognition stimuli. .

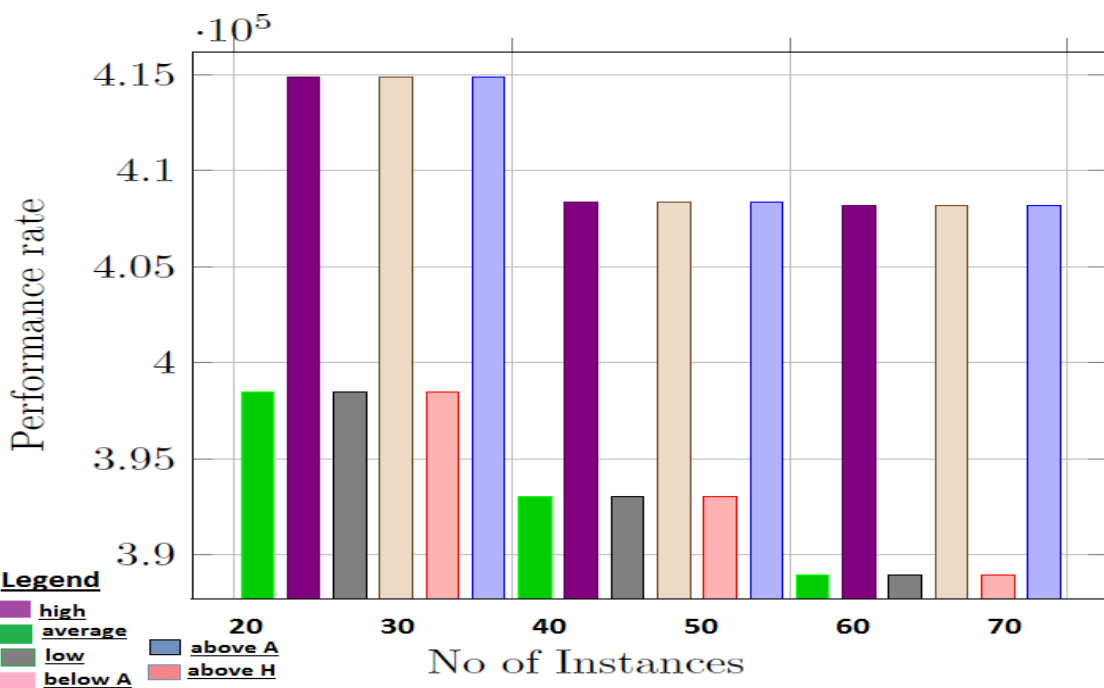


Figure 5: Performance rate of participants in respect to their cognitive perception of the input image.

5 CONCLUSION

This paper seeks to investigate a narrative comparison of the cognitive perception of people to image sensing in real-time subjective estimation. An experimental study was carried out with participants between older to younger participants, both female and male, who were asked to rate their judgment about a facial recognition system as a stimulus to determine the likelihood of the best match to an input image between the foreground and background version. Their physiological response was also taken while they interacted with the visual stimuli. Most participants rated the stimuli

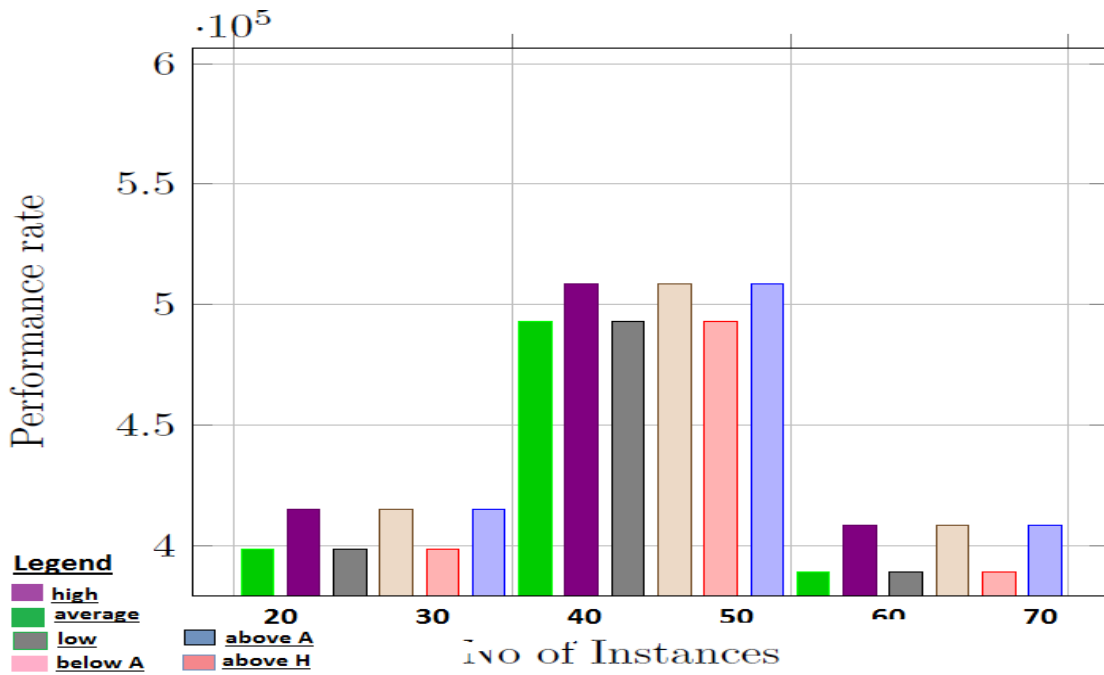


Figure 6: Performance rate of Middle-aged participants on their cognitive perception of the mapped input image.

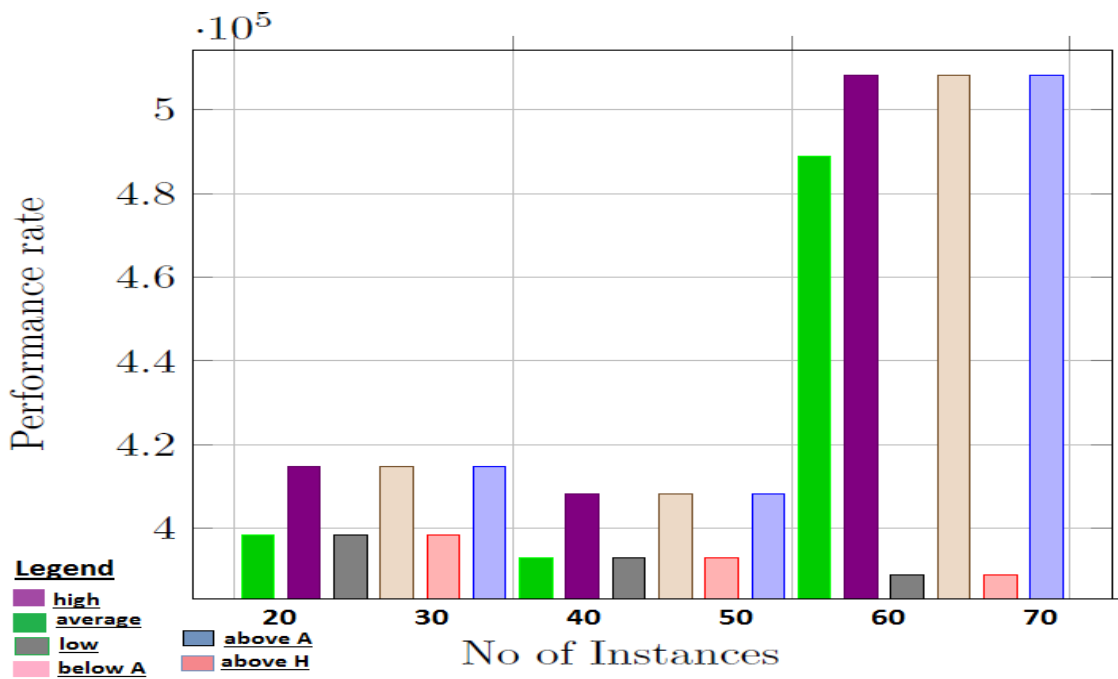


Figure 7: Performance rate of younger participants concerning the input image.

as above average and a high match between the two segments. This accuracy is also close to the predictive performance of the inference engine which was discussed in the pilot study of the previous paper. The future perspective would be to compare these results to deep learning patterns with participants' subjective views as part of the model's attributes, this would be in the form of an analytical model with its controls for parameter settings.

ACKNOWLEDGMENTS

The authors would wish to thank Nasarawa State University, Keffi and Ahamdu Bello University, Zaria, for their support and contributions to this paper.

References

- [1] Abusubaih, M. A. and Khamayseh, S. (2021). Performance of machine learning-based techniques for spectrum sensing in mobile cognitive radio networks. *IEEE Access*, 10:1410–1418.
- [2] Agarwal, N., Chauhan, S., Kar, A. K., and Goyal, S. (2017). Role of human behaviour attributes in mobile crowd sensing: a systematic literature review. *Digital Policy, Regulation and Governance*, 19(2):168–185.
- [3] Aloulou, W., Ayadi, F., Ramadani, V., and Dana, L.-P. (2024). Dreaming digital or chasing new real pathways? unveiling the determinants shaping saudi youth’s digital entrepreneurial intention. *International Journal of Entrepreneurial Behavior and Research*, 30(2/3):709–734.
- [4] Bilde, P. E., Kliim-Due, M., Rasmussen, B., Petersen, L. Z., Petersen, T. H., and Nielsen, J. B. (2011). Individualized, home-based interactive training of cerebral palsy children delivered through the internet. *BMC neurology*, 11:1–9.
- [5] Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., and Wilmer, J. B. (2012). Is the web as good as the lab? comparable performance from web and lab in cognitive/perceptual experiments. *Psychonomic bulletin and review*, 19:847–857.
- [6] Guo, B., Wang, Z., Yu, Z., Wang, Y., Yen, N. Y., Huang, R., and Zhou, X. (2015). Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm. *ACM computing surveys (CSUR)*, 48(1):1–31.
- [7] Jones, M., Lawler, M. J., Hintz, E., Bench, N., Mangrubang, F., and Trullender, M. (2014). Head mounted displays and deaf children: Facilitating sign language in challenging learning environments. In *Proceedings of the 2014 conference on Interaction design and children*, pages 317–320.
- [8] Kriegeskorte, N. (2015). Deep neural networks: a new framework for modeling biological vision and brain information processing. *Annual review of vision science*, 1:417–446.
- [9] Lai, T., Zeng, X., Xu, B., Xie, C., Liu, Y., Wang, Z., Lu, H., and Fu, S. (2024). The application of artificial intelligence technology in education influences chinese adolescent’s emotional perception. *Current Psychology*, 43(6):5309–5317.
- [10] Mele, G., Capaldo, G., Secundo, G., and Corvello, V. (2024). Revisiting the idea of knowledge-based dynamic capabilities for digital transformation. *Journal of Knowledge Management*, 28(2):532–563.
- [11] Nguyen, V. T., Sonkusare, S., Stadler, J., Hu, X., Breakspear, M., and Guo, C. C. (2017). Distinct cerebellar contributions to cognitive-perceptual dynamics during natural viewing. *Cerebral Cortex*, 27(12):5652–5662.
- [12] Pasolini, G., Guerra, A., Guidi, F., Decarli, N., and Dardari, D. (2020). Crowd-based cognitive perception of the physical world: Towards the internet of senses. *Sensors*, 20(9):2437.
- [13] Peng, M., Jing, W., Cai, M., and Zhou, Z. (2017). Perceptual difference between internet words and real-world words: temporal perception, distance perception, and perceptual scope. *Acta Psychologica Sinica*, 49(7):866.
- [14] Restuccia, F., Ghosh, N., Bhattacharjee, S., Das, S. K., and Melodia, T. (2017). Quality of information in mobile crowdsensing: Survey and research challenges. *ACM Transactions on Sensor Networks (TOSN)*, 13(4):1–43.
- [15] Saul, L. K., Weiss, Y., and Bottou, L. (2005). *Advances in neural information processing systems 17: proceedings of the 2004 conference*, volume 17. MIT Press.
- [16] Shao, S., Fan, M., Yu, C., Li, Y., Xu, X., and Wang, H. (2022). Machine learning-assisted sensing techniques for integrated communications and sensing in wlns: Current status and future directions. *Progress In Electromagnetics Research*, 175.

- [17] Sheth, A. (2016). Internet of things to smart iot through semantic, cognitive, and perceptual computing. *IEEE Intelligent Systems*, 31(2):108–112.
- [18] Sperling, G., Pavel, M., Cohen, Y., Landy, M., and Schwartz, B. (1983). Image processing in perception and cognition. In *Physical and Biological Processing of Images: Proceedings of an International Symposium Organised by the Rank Prize Funds, London, England, 27–29 September, 1982*, pages 359–378. Springer.
- [19] Sperling, G. and PSYCHOLOGY, N. Y. U. N. D. O. (1994). *Visual Motion Perception and Visual Information Processing*. AFOSR/NL.
- [20] Stokoe Jr, W. C. (2005). Sign language structure: An outline of the visual communication systems of the american deaf. *Journal of deaf studies and deaf education*, 10(1):3–37.
- [21] Sun, Y., Peng, M., Zhou, Y., Huang, Y., and Mao, S. (2019). Application of machine learning in wireless networks: Key techniques and open issues. *IEEE Communications Surveys and Tutorials*, 21(4):3072–3108.
- [22] Wang, B., Zhou, H., Li, X., Yang, G., Zheng, P., Song, C., Yuan, Y., Wuest, T., Yang, H., and Wang, L. (2024). Human digital twin in the context of industry 5.0. *Robotics and Computer-Integrated Manufacturing*, 85:102626.
- [23] Wang, J., Gao, Q., Ma, X., Zhao, Y., and Fang, Y. (2020). Learning to sense: Deep learning for wireless sensing with less training efforts. *IEEE Wireless Communications*, 27(3):156–162.
- [24] Wu, H., Han, Z., Wolter, K., Zhao, Y., Ko, H., et al. (2019). Deep learning driven wireless communications and mobile computing.
- [25] Zhang, C., Patras, P., and Haddadi, H. (2019). Deep learning in mobile and wireless networking: A survey. *IEEE Communications surveys and tutorials*, 21(3):2224–2287.
- [26] Zhang, R., Jing, X., Wu, S., Jiang, C., Mu, J., and Yu, F. R. (2020). Device-free wireless sensing for human detection: The deep learning perspective. *IEEE Internet of Things Journal*, 8(4):2517–2539.