Applied Data Science and Analysis
Vol.2025, **pp**. 3–6
DOI: https://doi.org/10.58496/ADSA/2025/002; ISSN: 3005-317X
https://mesopotamian.press/journals/index.php/ADSA

Editorial Article

Limitations of Deep Learning vs. Human Intelligence: Training Data, Interpretability, Bias, and Ethics

Yahya Layth Khaleel [1,*], Mustafa Abdulfattah Habeeb [1], Thierry Oscar Codjo EDOH [2,3]

[1] *College of Computer Science and Mathematics, Tikrit University, Iraq.*

[2] *RFW-Universität Bonn, Germany.*

[3] *Ecole Supérieure Multinationale des Télécommunications, Dakar, Senagal*

Dear Editor:

Deep Learning (DL) has brought a paradigm shift in innumerable fields and allowed machines to learn and decide to a very high extent. Its advantage is that it can analyze large sets of data, identify rather complex patterns, and learn from experience. DL models are widely used to perform complicated tasks [1], [2]. The state of the art in DL includes neural networks with two or more tiers and it has made progressive improvements in fields like image and voice identification, writing comprehension, language generation, and even decision-making self-governing systems [3]. These have led to growing concern and anticipation in the efficiency and rate of change that AI can bring about in industries and people [4].

Nonetheless, it is important to recognize that despite all the magnificent accomplishments of deep learning, the systems built using it are hardly replicas of the human brain [5]. Although DL is very effective in many different fields, its current potential is still rather limited if compared to the human brain [6]. These are nonetheless important to note because they allow us to keep in check the expectations we have about the potential of DL. Although DL can accomplish amazing things, there are some tasks, for instance, creative thinking, that the DL will never be capable it [7].

The basic concept that underpins all DL methods is that of using data and this data has to be categorized to train the models [8]. A major challenge commonly associated with DL models has to do with the need for large volumes of labelled data to feed the networks [9]. Experts, for instance, can acquire knowledge for applying a certain procedure or solving a particular problem from just one or few samples, whereas; deep models or learners require a vast sample to be able to perform basic tasks optimally [10]. Choosing only the data over using it in the training process comes with key difficulties, especially for applications in fields with a small number of labeled data or where such data is expensive to come by [11]. Besides, these models are usually dependent on the quality and amount of data fed into the model and can reflect the errors and prejudices of the feed data [12].

One of the challenges of DL is that it is often described as a ‘black box’ and it is not always easy to know how it arrives at its decisions [13]. This sets a preliminary constraint in comprehending how its techniques work [14]. Several authors have dealt with this problem using Explainable Artificial Intelligence (XAI) [15]. However, since XAI methods give results in the form of graphs and tables, these are often relatively intricate and give rise to another problem [16]. Additionally, with the abundance of these methods, another challenge arises, in particular, further work is in identifying which method is better to apply for the interpretation and which one is more clear or related to human thinking to explain the black box and make it more comprehensible [17].

Also, the DL models are perturbable, where a slight change in the input data results in an opposite output or different [18]. This is a key strength of people over models in that such models are very sensitive to context changes in contrast to human cognition which hardly changes even with minor differences in the input it receives [19]. This decision-making approach is a significant concern because it may pose a threat, especially in fields that require safety and certainty such as self-driving vehicles and data protection from hackers [20].

Furthermore, the DL models perform exceptionally in one context but fail in generalizing to other contexts of similar problems. While human intelligence can easily transfer information and knowledge to various contexts; it was found that DL models can generalize the info only within the given datasets [21]. Such a form of deficiency does not allow these models to change as influenced by varying circumstances and situations in the real world while at the same time lacking the common-sense aspect, which is a key foundational requirement of good models [22].

So, the aforementioned limitations must be considered for a more conservative approach to using DL technologies. While these models have undoubtedly enhanced a wide range of abilities and contributed to self-sufficiency, they are not superior

*Corresponding author. Email: yahya@tu.edu.iq

to the human mind [23]. The problem arises when practitioners use DL without fully understanding its limitations, leading to adverse effects, especially in critical fields like healthcare, autonomous systems, and decision-making processes [24]. For instance, in healthcare, DL models may misdiagnose rare conditions [25]; in autonomous systems, they have failed to recognize objects, causing accidents [26]; and in decision-making, such as credit scoring, DL can inherit biases from training data, leading to unfair outcomes [27].

In DL, one of the major issues that leads to bias is the training data [28]; it is possible to arise and the results manifest differently from the existing prejudice if the data set is imbalanced or erroneous. Bias is also a problem that is not easily apparent because the decision-making process of models is concealed within what is referred to as the 'black box' [29]. On the other hand, the bias that emanates from the human being can be attributed to the experiences and the society in which he lives, and even though such bias is unconscious, it is something that can be understood and changed by the individual. The effect of bias in DL can be far-reaching due to the nature of these systems, that is, they are scalable while a natural bias will affect few, but when it does it can be influential (see Figure 1).

Thus, it can be stated that DL is a valuable addition to the arsenal of AI techniques; however, one must remember certain constraints remain [30]. To date, human intelligence which can learn from small data sets, justify the involved decisions, and handle new situations has not been surpassed. In essence, as we proceed with constructing the use of AI technologies in our society, we must have this knowledge in mind to allow for the use of these gadgets with the weaknesses that are characteristic of them with the strengths of allowing for the use of these tools in supplement of human decisions.

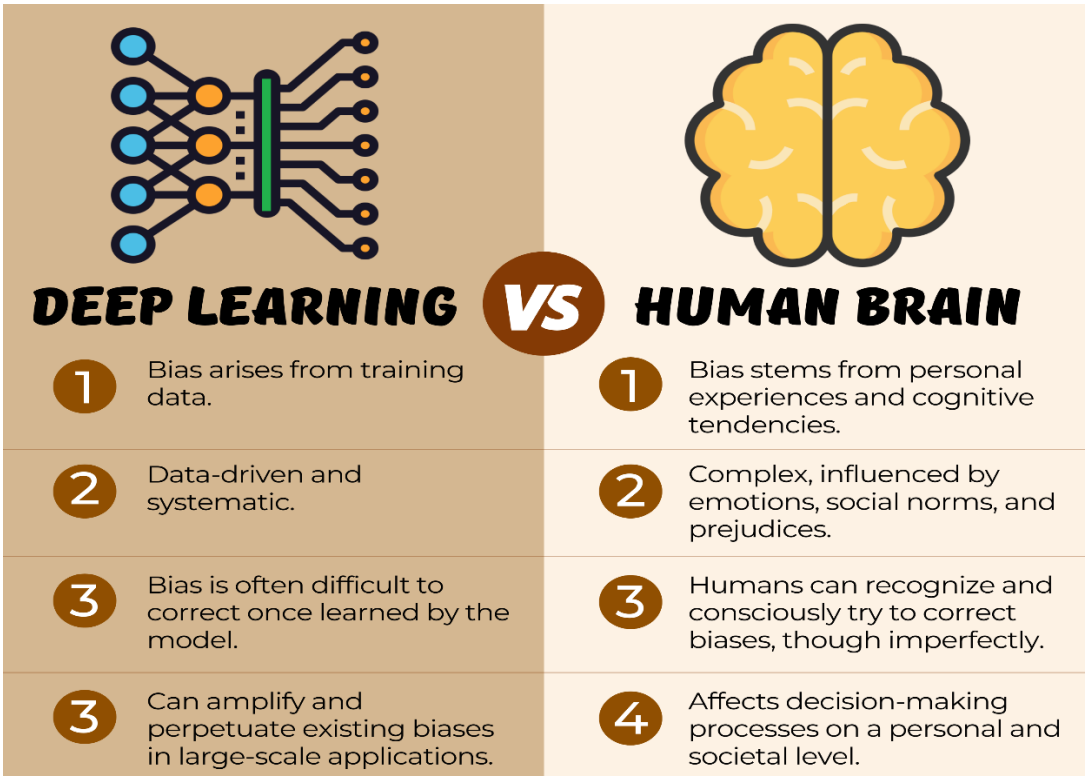

Fig. 1. DL VS. Human Brain Based on Current Limitations Factors.

Conflicts of Interest

The author's paper explicitly states that no funding was received from any institution or sponsor.

Funding

None.

Acknowledgment

None

References

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[2] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agric.*, vol. 147, pp. 70–90, 2018, doi: https://doi.org/10.1016/j.compag.2018.02.016.

[3] M. Alam, M. D. Samad, L. Vidyaratne, A. Glandon, and K. M. Iftekharuddin, "Survey on Deep Neural Networks in Speech and Vision Systems," *Neurocomputing*, vol. 417, pp. 302–321, Dec. 2020, doi: 10.1016/j.neucom.2020.07.053.

[4] S. Minaee, A. Abdolrashidi, H. Su, M. Bennamoun, and D. Zhang, "Biometrics recognition using deep learning: a survey," *Artif. Intell. Rev.*, vol. 56, no. 8, pp. 8647–8695, 2023, doi: 10.1007/s10462-022-10237-x.

[5] Y. Matsuo *et al.*, "Deep learning, reinforcement learning, and world models," *Neural Networks*, vol. 152, pp. 267–275, 2022, doi: 10.1016/j.neunet.2022.03.037.

[6] O. I. Obaid, "From Machine Learning to Artificial General Intelligence: A Roadmap and Implications," *Mesopotamian J. Big Data*, vol. 2023, no. SE-Articles, pp. 81–91, Aug. 2023, doi: 10.58496/mjbd/2023/012.

[7] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electron. Mark.*, vol. 31, no. 3, pp. 685–695, 2021, doi: 10.1007/s12525-021-00475-2.

[8] A. M. Rahmani *et al.*, "Machine Learning (ML) in Medicine: Review, Applications, and Challenges," *Mathematics*, vol. 9, no. 22, 2021, doi: 10.3390/math9222970.

[9] H. Song, M. Kim, D. Park, Y. Shin, and J. G. Lee, "Learning From Noisy Labels With Deep Neural Networks: A Survey," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 34, no. 11, pp. 8135–8153, 2023, doi: 10.1109/TNNLS.2022.3152527.

[10] G. Algan and I. Ulusoy, "Image classification with deep learning in the presence of noisy labels: A survey," *Knowledge-Based Syst.*, vol. 215, p. 106771, 2021, doi: 10.1016/j.knosys.2021.106771.

[11] F. K. H. Mihna, M. A. Habeeb, Y. L. Khaleel, Y. H. Ali, and L. A. E. Al-Saeedi, "Using Information Technology for Comprehensive Analysis and Prediction in Forensic Evidence," *Mesopotamian J. CyberSecurity*, vol. 4, no. 1, 2024, doi: 10.58496/MJCS/2024/002.

[12] T. Alqahtani *et al.*, "The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research," *Res. Soc. Adm. Pharm.*, vol. 19, no. 8, pp. 1236–1242, 2023, doi: 10.1016/j.sapharm.2023.05.016.

[13] V. Buhrmester, D. Münch, and M. Arens, "Analysis of Explainers of Black Box Deep Neural Networks for Computer Vision: A Survey," *Mach. Learn. Knowl. Extr.*, vol. 3, no. 4, pp. 966–989, 2021, doi: 10.3390/make3040048.

[14] C. B. Azodi, J. Tang, and S. H. Shiu, "Opening the Black Box: Interpretable Machine Learning for Geneticists," *Trends Genet.*, vol. 36, no. 6, pp. 442–455, 2020, doi: 10.1016/j.tig.2020.03.005.

[15] A. S. Albahri *et al.*, "A systematic review of trustworthy artificial intelligence applications in natural disasters," *Comput. Electr. Eng.*, vol. 118, 2024, doi: 10.1016/j.compeleceng.2024.109409.

[16] B. Brożek, M. Furman, M. Jakubiec, and B. Kucharzyk, "The black box problem revisited. Real and imaginary challenges for automated legal decision making," *Artif. Intell. Law*, vol. 32, no. 2, pp. 427–440, 2024, doi: 10.1007/s10506-023-09356-9.

[17] T. Dhar, N. Dey, S. Borra, and R. S. Sherratt, "Challenges of Deep Learning in Medical Image Analysis—Improving Explainability and Trust," *IEEE Trans. Technol. Soc.*, vol. 4, no. 1, pp. 68–75, 2023, doi: 10.1109/tts.2023.3234203.

[18] Y. L. Khaleel, M. A. Habeeb, A. S. Albahri, T. Al-Quraishi, O. S. Albahri, and A. H. Alamoodi, "Network and cybersecurity applications of defense in adversarial attacks: A state-of-the-art using machine learning and deep learning methods," vol. 33, no. 1, 2024, doi: doi:10.1515/jisys-2024-0153.

[19] C. Wang *et al.*, "How indoor environmental quality affects occupants' cognitive functions: A systematic review," *Build. Environ.*, vol. 193, p. 107647, 2021, doi: 10.1016/j.buildenv.2021.107647.

[20] A. Qayyum, M. Usama, J. Qadir, and A. Al-Fuqaha, "Securing Connected Autonomous Vehicles: Challenges Posed by Adversarial Machine Learning and the Way Forward," *IEEE Commun. Surv. Tutorials*, vol. 22, no. 2, pp. 998–1026, 2020, doi: 10.1109/COMST.2020.2975048.

[21] M. Pichler and F. Hartig, "Machine learning and deep learning—A review for ecologists," *Methods Ecol. Evol.*, vol. 14, no. 4, pp. 994–1016, 2023, doi: 10.1111/2041-210X.14061.

[22] K. Zhou, Z. Liu, Y. Qiao, T. Xiang, and C. C. Loy, "Domain Generalization: A Survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 4, pp. 4396–4415, 2023, doi: 10.1109/TPAMI.2022.3195549.

[23] T. J. Sejnowski, "The unreasonable effectiveness of deep learning in artificial intelligence," *Proc. Natl. Acad. Sci.*, vol. 117, no. 48, pp. 30033–30038, 2020, doi: 10.1073/pnas.1907373117.

[24] G. S. Nadella, S. Satish, K. Meduri, and S. S. Meduri, "A Systematic Literature Review of Advancements, Challenges and Future Directions of AI And ML in Healthcare," *Int. J. Mach. Learn. Sustain. Dev.*, vol. 5, no. 3, pp. 115–130, 2023.

[25] N. Sharma and N. Jindal, "Emerging artificial intelligence applications: metaverse, IoT, cybersecurity, healthcare - an overview,"*Multimed. Tools Appl.*, vol. 83, no.19, pp. 57317–57345, 2024, doi: 10.1007/s11042-023-17890-6.

[26] A. S. Mueller, J. B. Cicchino, and D. S. Zuby, "What humanlike errors do autonomous vehicles need to avoid to maximize safety?," *J. Safety Res.*, vol. 75, pp. 310–318, 2020, doi: https://doi.org/10.1016/j.jsr.2020.10.005.

[27] R. R. Waller and R. L. Waller, "Assembled Bias: Beyond Transparent Algorithmic Bias," *Minds Mach.*, vol. 32, no. 3, pp. 533–562, 2022, doi: 10.1007/s11023-022-09605-x.

[28] A. S. Albahri, Y. L. Khaleel, and M. A. Habeeb, "The Considerations of Trustworthy AI Components in Generative AI; A Letter to Editor," *Appl. Data Sci. Anal.*, vol. 2023, pp. 108–109, 2023, doi: 10.58496/adsa/2023/009.

[29] M. Shah and N. Sureja, "A Comprehensive Review of Bias in Deep Learning Models: Methods, Impacts, and Future Directions," *Arch. Comput. Methods Eng.*, 2024, doi: 10.1007/s11831-024-10134-2.

[30] M. A. Habeeb, Y. L. Khaleel, and A. S. Albahri, "Toward Smart Bicycle Safety: Leveraging Machine Learning Models and Optimal Lighting Solutions," in *Proceedings of the Third International Conference on Innovations in Computing Research (ICR'24)*, K. Daimi and A. Al Sadoon, Eds., Cham: Springer Nature Switzerland, 2024, pp. 120–131.